

EXAMINING THE ROLE OF MEDIATION IN THE ASSOCIATIONS OF INDIVIDUAL CHARACTERISTICS WITH PROSTATE SPECIFIC ANTIGEN (PSA) AND PROSTATE CANCER RISK

Sean Harrison

*A dissertation submitted to the University of Bristol in accordance with the requirements for award of
the degree of Doctor of Philosophy in the Faculty of Health Sciences*

Bristol Medical School

Submitted September 2017

Word count: 63,891

Abstract

The aim of this thesis is to identify plausible individual characteristics that have an association both with prostate cancer and prostate-specific antigen (PSA), and to precisely estimate these associations for one characteristic, including assessment of any potential mediation between the characteristic and PSA through prostate cancer.

I used both an assessment of previous studies and expert opinion to choose the characteristic of interest: body mass index (BMI). I then used a combination of aggregate data (AD) and individual participant data (IPD) meta-analyses to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA, and Mendelian Randomisation (MR) to assess whether any of the associations are causal. The AD and IPD meta-analyses showed no evidence that BMI has a linear association with prostate cancer risk, limited evidence that BMI has a positive linear association with advanced prostate cancer, and strong evidence that BMI has a negative association with PSA. It is likely that the association between BMI and PSA was non-linear, with PSA decreasing more as BMI increases. It is possible that the observed linear association between BMI and advanced prostate cancer was due to bias from testing for prostate cancer using PSA. The MR showed no evidence of any causal relationships, but was underpowered to detect effects of the magnitude seen in the meta-analyses.

Overall, I conclude that there is no strong, consistent evidence that BMI is associated with prostate cancer risk, there is some evidence BMI has a positive association with advanced prostate cancer risk, and strong evidence that BMI has a strong, negative, non-linear association with PSA. There was no evidence of mediation of the association between BMI and PSA through prostate cancer. Future studies examining risk factors for prostate cancer should consider whether the associations could be biased by an association with PSA.

Dedication and Acknowledgements

I would like to offer my profound thanks to my supervisors – Kate Tilling, Hayley Jones and Emma Turner – for everything they have done towards this thesis. Their expertise, insight and encouragement made all this possible. I cannot overstate the depth of my gratitude to all of them.

Additionally, I would like to thank Sharen O’Keefe, who helped me innumerable times, and all the researchers I have worked with in Bristol Medical School, who have made it such a great place to work.

Finally, I would like to thank my parents for their continued support.

I dedicate this thesis to Amy, who makes each day worthwhile, for everything she has done. And to Lily as well, of course.

Author's Declaration

This thesis is the work of the author. I would like to thank Rosie Lennon, who assisted with the data extraction and risk of bias assessment in the aggregate data meta-analyses. I would also like to thank Julian Higgins, Hayley Jones, Richard Martin and Sarah Lewis for their guidance and work developing the albatross plot and writing the journal paper. I would also like to thank Richard Martin, Kate Tilling, Emma Turner, Athene Lane, Andrew Simpkin, Michael Davis, Jenny Donovan, Freddie Hamdy, David Neal and Sarah Lewis for their assistance investigating the clinical utility of the age-BMI-adjusted PSA model and writing the journal paper. I am a Wellcome Trust Funded PhD student with grant code 102432/Z/13/Z. The funding source had no role in the design, conduct of the study, collection, management, analysis and interpretation or preparation, review, or approval of the thesis.

I acknowledge and am grateful to all the studies that allowed me to use their data in this thesis. Specifically, I would like to thank Professor Monique Roobol and the ERSPC-Rotterdam study, Professor Ruud Bosch and the Krimpen study, and the PCPT, PLCO, PRACTICAL and ProtecT studies. The opinions in this thesis are my own and do not necessarily reflect the views of any of the study researchers or funders. For PCPT: SWOG is a member of the National Clinical Trials Network supported by the National Cancer Institute (NCI). This manuscript was prepared using a limited access data set obtained from SWOG and does not necessarily reflect the opinions or view of SWOG or the NCI. For PLCO: I thank the NCI for access to NCI's data collected by the PLCO screening trial. The statements contained herein are solely mine and do not represent or imply concurrence of endorsement by NCI. For ProtecT: ProtecT was supported by Cancer Research UK project grant grants C11043/A4286, C18281/A8145, C18281/A11326 and C18281/A15064. The ProtecT study is supported by the UK National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme, HTA 96/20/99; ISRCTN20141297. I would like to acknowledge the support of the National Cancer Research Institute (NCRI) formed by the Department of Health, the Medical Research Council (MRC) and Cancer Research UK. The NCRI provided funding through ProMPT (Prostate Mechanisms of Progression and Treatment) and this support is also gratefully acknowledged.

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

Contents

Title Page	i
Abstract	iii
Dedication and Acknowledgements.....	v
Author's Declaration.....	vii
Contents	ix
List of Figures.....	xix
List of Tables	xxiv
CHAPTER 1. INTRODUCTION.....	27
 1.1. Aim of this Thesis	27
 1.2. Background for Prostate Cancer and Prostate-Specific Antigen.....	27
1.2.1. The Prostate.....	27
1.2.2. Prostate Cancer	28
Gleason score.....	28
Prostate cancer staging.....	29
1.2.3. Prostate Specific Antigen	30
PSA screening for prostate cancer	30
1.2.4. Improving PSA as a Test for Prostate Cancer.....	32
 1.3. Description of Thesis Chapters	34
Chapter 2: Evidence synthesis methods	34
Chapter 3: Creating and using a Sankey diagram to narrow the focus of a systematic review of prostate cancer and prostate-specific antigen.....	34
Chapter 4: Systematic review and meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA	34
Chapter 5: Individual participant data meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA.....	35
Chapter 6: Combining aggregate data and individual participant data to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA.....	35

Chapter 7: Using Mendelian Randomisation to examine the causal effect of BMI on prostate cancer risk and PSA	35
Chapter 8: Discussion.....	35
1.4. Objectives	36
CHAPTER 2. EVIDENCE SYNTHESIS METHODS.....	37
2.1. Aims	37
2.2. Systematic Reviews	38
2.3. Methods for Evidence Synthesis.....	40
2.3.1. Evidence synthesis methods that do not require an effect estimate	40
Narrative synthesis	40
Vote Counting and Sign Tests	40
Harvest Plot.....	41
Combining P values.....	41
2.3.2. Meta-analysis.....	42
Aggregate data meta-analysis.....	42
Individual Patient Data Meta-analysis	42
Fixed-effect and random-effects meta-analyses	42
Forest plots	44
Heterogeneity and inconsistency	44
Advantages, issues and limitations	45
2.3.3. Albatross Plot	46
Generation of effect contours	47
Strengths and limitations	49
2.4. Summary.....	49
CHAPTER 3. CREATING AND USING A SANKEY DIAGRAM TO NARROW THE FOCUS OF A SYSTEMATIC REVIEW OF PROSTATE CANCER AND PROSTATE-SPECIFIC ANTIGEN.....	51
3.1. Aim	51
3.2. Background	52
3.2.1. Sankey diagrams	52
3.2.2. Sankey diagrams in epidemiology	53
3.2.3. Causal diagrams	53

3.3. Generation of the Sankey diagram.....	55
3.4. Expert Opinion	58
3.5. Results	60
3.5.1. Sankey diagram.....	60
3.5.2. Expert opinion.....	63
3.6. Variable Selection	64
The associations between BMI, prostate cancer and PSA	66
3.7. Discussion	67
3.7.1. Strengths.....	67
3.7.2. Limitations	67
3.8. Summary.....	68

CHAPTER 4. SYSTEMATIC REVIEW AND AGGREGATE DATA META-ANALYSIS OF ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA69

4.1. Aims	69
4.2. Methods.....	70
4.2.1. Data Sources	70
4.2.2. Eligibility criteria	70
4.2.3. Data extraction	70
4.2.4. Risk of bias assessment.....	72
4.2.5. MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies.....	73
4.3. Methods to Derive Standard Effect Estimates	74
4.3.1. Statistical Analysis of BMI-Prostate Cancer Association	76
4.3.1.A. Weighted regression for correlated outcomes	76
4.3.1.B. Mean BMI not presented for each level of BMI	77
4.3.1.C. Mean differences presented.....	77
4.3.1.D. Matched studies using levels of BMI, and ORs not reported for levels.....	77
4.3.1.E. Other information presented.....	78
4.3.2. Statistical Analysis of BMI-PSA Association	79
4.3.2.A. Weighted regression.....	79
4.3.2.B. Percentage change in PSA for a 5 kg/m ² increase in BMI.....	79
4.3.3. Transformations of PSA	80

4.3.3.A. Mean and SD of PSA presented	81
4.3.3.B. Median PSA presented	81
4.3.3.C. Geometric mean PSA presented.....	81
4.3.3.D. Odds ratios presented	82
4.3.3.E. Ratio of geometric means presented.....	82
4.3.3.F. Other information presented.....	83
4.3.4. Combining data.....	84
Meta-analysis.....	84
Meta-regression.....	84
Funnel plots	84
Albatross plots	85
4.4. Results	86
4.4.1. BMI and prostate cancer	88
4.4.2. BMI and Advanced Prostate Cancer	99
4.4.3. BMI and PSA.....	105
4.5. Discussion	111
4.5.1. Comparison with previous research: BMI-prostate cancer	112
4.5.2. Comparison with previous research: BMI-advanced prostate cancer.....	114
4.5.3. Comparison with previous research: BMI-PSA	115
4.5.4. Strengths and limitations.....	115
4.5.5. Methodology	117
4.6. Conclusion.....	119
CHAPTER 5. INDIVIDUAL PARTICIPANT DATA META-ANALYSIS OF ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA	121
5.1. Aims	121
5.2. Background on handling missing data	122
5.2.1. Missing data.....	122
5.2.2. Multiple Imputation.....	123
5.2.3. Multiple imputation in IPD meta-analysis	123
5.2.4. Imputation with non-linear associations	125
5.3. Background on Meta-analysis in IPD	125
5.4. Methods.....	127

5.4.1. Studies	127
European Randomised Study of Screening for Prostate Cancer: Rotterdam	127
Krimpen.....	128
Prostate Cancer Prevention Trial	128
Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial.....	128
Prostate Testing for Cancer and Treatment.....	129
5.4.2. Risk of bias assessment.....	129
5.4.3. Data cleaning	130
5.4.4. Exploratory analysis.....	130
Complete case analysis	130
Checking for heterogeneity	131
Checking for interactions	131
5.4.5. Imputation	132
5.4.6. Linear Analysis	133
Analysis 1: BMI and prostate cancer.....	134
Analysis 2: BMI and advanced prostate cancer	134
Analysis 3: BMI and PSA.....	134
Analysis 4: Prostate cancer and PSA	134
5.4.7. Non-linear Analysis	134
5.5. Results	136
5.5.1. Exploratory analyses.....	136
Complete case analysis 1: BMI and prostate cancer.....	137
Complete case analysis 2: BMI and advanced prostate cancer	138
Complete case analysis 3: BMI and PSA.....	139
Checking for heterogeneity 1: BMI and PSA	141
Checking for heterogeneity 2: BMI and prostate cancer risk	142
Checking for heterogeneity 3: PSA and prostate cancer risk.....	143
Checking for interactions	145
5.5.2. Imputation	146
5.5.3. Linear analyses.....	149
Linear analysis 1: BMI and prostate cancer	149
Linear analysis 2: BMI and advanced prostate cancer	151
Linear analysis 3: BMI and PSA	153
Linear analysis 4: prostate cancer and PSA.....	156
5.5.4. Non-linear analyses	158
Non-linear analysis 1: BMI and prostate cancer	158
Non-linear analysis 2: BMI and advanced prostate cancer.....	160

Non-linear analysis 3: BMI and PSA	162
5.6. Discussion	164
5.6.1. BMI and prostate cancer	164
5.6.2. BMI and advanced prostate cancer	164
5.6.3. BMI and PSA.....	164
5.6.4. Prostate cancer and PSA.....	165
5.6.5. Methodology	165
5.6.6. Strengths and limitations.....	166
5.7. Conclusion.....	168
CHAPTER 6. COMBINING AGGREGATE DATA AND INDIVIDUAL PARTICIPANT DATA TO ESTIMATE THE ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA.....	169
6.1. Aims	169
6.2. Methods.....	170
6.2.1. Studies	170
6.2.2. Linear associations.....	170
6.2.3. Non-linear associations.....	171
6.3. Results: Linear analysis.....	173
6.3.1. BMI and prostate cancer	173
6.3.2. BMI and advanced prostate cancer	175
6.3.3. BMI and PSA.....	176
6.3.4. Summary diagram.....	177
6.4. Results: Non-linear analysis	178
6.4.1. BMI and prostate cancer	178
6.4.2. BMI and advanced prostate cancer	182
6.4.3. BMI and PSA.....	186
6.5. Discussion	190
6.5.1. Heterogeneity	191
6.5.2. Strengths and Limitations	192
6.5.3. Implications.....	193
6.6. Conclusion.....	194

**CHAPTER 7. USING MENDELIAN RANDOMISATION TO EXAMINE THE CAUSAL EFFECT OF BODY MASS INDEX ON
PROSTATE CANCER RISK, ADVANCED PROSTATE CANCER RISK AND PROSTATE-SPECIFIC ANTIGEN195**

7.1. Aims	195
7.2. Background	196
7.2.1. Problems with observational epidemiology	196
7.2.2. Instrumental variable analysis	197
7.2.3. Mendelian Randomisation.....	198
Bidirectional MR	200
7.2.4. Limitations of MR.....	201
7.3. Methods.....	203
7.3.1. Selection of instruments.....	203
7.3.2. Setting.....	203
7.3.3. Exposures and outcomes	204
Genetic risk scores	204
Genetic risk score exclusions	205
Outcomes.....	206
7.3.4. Statistical analysis	206
Main analyses	206
Sensitivity analysis 1: Assessment of confounding	207
Sensitivity analysis 2: Analysis of all SNPs.....	207
Sensitivity analysis 3: Non-linearity	207
7.4. Results	208
7.4.1. Clinical characteristics and their associations with genetic risk scores	208
7.4.2. Main analyses	210
7.4.3. Sensitivity analysis 1: Assessment of confounding	213
7.4.4. Sensitivity analysis 2: Analysis of all SNPs.....	213
7.4.5. Sensitivity analysis 3: Non-linearity	214
7.5. Discussion	215
7.5.1. Summary of Findings	215
7.5.2. Strengths and Limitations	216
7.5.3. Comparison with previous studies.....	217
7.5.4. Implications.....	218
7.6. Conclusion.....	218

CHAPTER 8. DISCUSSION.....	219
8.1. Introduction	219
8.2. Thesis Summary	220
8.3. Summary of Clinical Results and Comparison with Previous Research	224
BMI and prostate cancer.....	224
BMI and advanced prostate cancer	224
BMI and PSA.....	225
8.4. Summary of Methodological Findings	230
Albatross plot.....	230
Probabilistic matching algorithm	231
Sankey diagram and expert opinion	231
Analysis methods	231
8.5. Strengths and Limitations.....	233
Strengths.....	233
Limitations	233
8.6. Implications of this Thesis	236
8.7. Future Directions.....	237
Is an age-BMI-adjusted PSA model clinically useful?	237
Calculating the indirect effect with a binary mediator using aggregate data.....	238
8.8. Overall Conclusion	240
REFERENCES	241
APPENDICES.....	265
Appendix 1: Abbreviations.....	265
Appendix 2: Creation and Validation of a Probabilistic Matching Algorithm for Deduplication	269
Appendix 2.1 Rationale for deduplication.....	269
Appendix 2.2 Deduplication algorithms	270
Appendix 2.3 Version 1	271
Appendix 2.4 Version 2	276
Appendix 2.5 Algorithm Validation	281

Appendix 2.6 Discussion	287
Appendix 3: Question Sheet Given to ProtecT Lead Nurses	289
Appendix 4: Information Sheet for ProtecT Lead Nurses.....	291
Appendix 5: Risk of Bias Questions	293
Appendix 6: MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies	294
Appendix 7: Calculating the SD of log-PSA using the ratio of geometric means	296
Appendix 8: Table showing the effect alleles, GWAS effect size and allele frequencies of SNPs included in the BMI, PSA, log-PSA and prostate cancer genetic risk scores	300
Appendix 9: The albatross plot: A novel graphical tool for presenting results of diversely reported studies in a systematic review	305
Appendix 10: Investigating the prostate specific antigen, body mass index and age relationship: is an age-BMI-adjusted PSA model clinically useful?.....	314

List of Figures

Figure 1.1 A visualisation of Gleason's Pattern for classifying prostate cancers.....	29
Figure 1.2 Diagram showing the associations between age, prostate cancer status, prostate cancer diagnosis and PSA	33
Figure 3.1 Sankey diagram showing how energy is transferred in a light bulb	52
Figure 3.2 Causal diagram showing the associations between variables of interest (individual characteristics), prostate cancer and PSA. Each circle (node) represents a variable, and each arrow indicates the assumed direction of causality between the variables.....	54
Figure 3.3 Sankey diagram showing the number of studies associating prostate cancer and PSA (left hand side) with selected individual characteristics (right hand side).....	62
Figure 3.4 Causal diagram showing all associations between variables of interest, PSA and prostate cancer.....	65
Figure 4.1 Graph of a normal distribution (mean = 0, SD = 1) and its corresponding log-normal distribution (created by exponentiating the normal distribution)	80
Figure 4.2 PRISMA flow diagram showing the number of studies in each stage of the systematic review	87
Figure 4.3 Forest plot for the association between BMI and prostate cancer	96
Figure 4.4 Funnel plot for the association between BMI and prostate cancer	97
Figure 4.5 Albatross plot for the association between BMI and prostate cancer	98
Figure 4.6 Forest plot for the association between BMI and advanced prostate cancer.....	102
Figure 4.7 Funnel plot for the association between BMI and advanced prostate cancer.....	103
Figure 4.8 Albatross plot for the association between BMI and advanced prostate cancer.....	104
Figure 4.9 Forest plot for the association between BMI and PSA	108
Figure 4.10 Funnel plot for the association between BMI and PSA	109
Figure 4.11 Albatross plot for the association between BMI and log-PSA	110
Figure 5.1 Forest plot of BMI and prostate cancer (complete case analysis), showing the OR for prostate cancer for a 5 kg/m ² increase in BMI	137

Figure 5.2 Forest plot of BMI and advanced prostate cancer (complete case analysis), showing the OR for advanced prostate cancer for a 5 kg/m ² increase in BMI	138
Figure 5.3 Forest plot of BMI and PSA (direct effect, complete case analysis), showing the percentage change in PSA for a 5 kg/m ² increase in BMI	139
Figure 5.4 Forest plot of BMI and PSA (total effect, complete case analysis), showing the percentage change in PSA for a 5 kg/m ² increase in BMI	140
Figure 5.5 Line graph of BMI against log-PSA, using restricted cubic splines of BMI. Limited to BMIs between 20 kg/m ² and 40 kg/m ² . Knots at BMI quantiles (24.8 kg/m ² , 27 kg/m ² and 30 kg/m ²)	141
Figure 5.6 Line graph of BMI against prostate cancer risk, using restricted cubic splines of BMI. Limited to BMIs between 20 kg/m ² and 40 kg/m ² . Knots at BMI quantiles (24.8 kg/m ² , 27 kg/m ² and 30 kg/m ²)	142
Figure 5.7 Line graph of log-PSA against prostate cancer risk, limited to the middle 98% of PSA values for each study (to remove outliers), using restricted cubic splines of log-PSA. ERSPC = ERSPC-Rotterdam. ERSPC-Rotterdam, PLCO and ProtecT were also limited to PSA values above 3.0 ng/ml, 4.0 ng/ml and 3.0 ng/ml respectively as prostate cancer status was not observed lower than this. Knots at rough overall log PSA quantiles (-0.5, 0 and 0.5).....	144
Figure 5.8 Line graph of prostate cancer risk against log-PSA, limited to the middle 98% of PSA values for each study to remove outliers, using restricted cubic splines of log-PSA. ERSPC = ERSPC-Rotterdam	148
Figure 5.9 Forest plot of BMI and prostate cancer risk, showing the OR for prostate cancer for a 5 kg/m ² increase in BMI	150
Figure 5.10 Forest plot of BMI and advanced prostate cancer risk, showing the OR for advanced prostate cancer for a 5 kg/m ² increase in BMI	152
Figure 5.11 Forest plot of BMI and log PSA (direct effect), showing the percentage change in PSA for a 5 kg/m ² increase in BMI	154
Figure 5.12 Forest plot of BMI and log PSA (total effect), showing the percentage change in PSA for a 5 kg/m ² increase in BMI	155
Figure 5.13 Forest plot of the association between prostate cancer and PSA (multiplicative), showing the percentage change in PSA from prostate cancer	157
Figure 5.14 Predicted prostate cancer risk for values of BMI between 20 kg/m ² and 40 kg/m ² , for a 60-year-old man from the ERSPC-Rotterdam study with no family history of prostate cancer.....	159

Figure 5.15 Predicted advanced prostate cancer risk for values of BMI between 20 kg/m ² and 40 kg/m ² , for a 60-year-old man from the ERSPC-Rotterdam study with no family history of prostate cancer	161
Figure 5.16 Predicted log-PSA for values of BMI between 20 kg/m ² and 40 kg/m ² , for a 60-year-old man from the ERSPC-Rotterdam study without prostate cancer	163
Figure 6.1 Diagram showing the assumed directions of effect for BMI, PSA and prostate cancer	170
Figure 6.2 Forest plot of OR of prostate cancer for a 5 kg/m ² increase in BMI.....	174
Figure 6.3 Forest plot of OR of advanced prostate cancer for a 5 kg/m ² increase in BMI	175
Figure 6.4 Forest plot of the percentage change in PSA for a 5 kg/m ² increase in BMI.....	176
Figure 6.5 Summary diagram showing the results from all meta-analyses assuming linear associations between BMI, prostate cancer and PSA. OR for prostate cancer and percentage change in PSA for a 5 kg/m ² increase in BMI	177
Figure 6.6 Forest plot of the OR for prostate cancer for overweight versus normal weight BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants ..	180
Figure 6.7 Forest plot of the OR for prostate cancer for obese versus normal BMI categories, n1 = number of normal weight participants, n3 = number of obese participants	181
Figure 6.8 Forest plot of the OR for advanced prostate cancer for overweight versus normal weight BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants ..	184
Figure 6.9 Forest plot of the OR for advanced prostate cancer for obese versus normal weight BMI categories, n1 = number of normal weight participants, n3 = number of obese participants.....	185
Figure 6.10 Forest plot of the percentage change in PSA between overweight and normal BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants ..	188
Figure 6.11 Forest plot of the percentage change in PSA between obese and normal BMI categories, n1 = number of normal weight participants, n3 = number of obese participants ..	189
Figure 7.1 The causal associations between age, BMI and prostate cancer, showing confounding by age.....	196
Figure 7.2 Causal diagram of instrumental variable analysis, where the instrumental variable is (1) associated with the exposure, (2) not associated with the outcome except through the exposure, (3) not associated with confounders that are associated with the exposure and outcome	198

Figure 7.3 Causal diagram of an MR using BMI and PSA as an example, with the same assumptions as in IV analysis.....	200
Figure 7.4 Bidirectional MR, where genetic variants associated with BMI and PSA are regressed against PSA and BMI respectively, to discover in which direction(s) an association may flow	201
Figure 7.5 Summary results for the main analysis: effect estimates are for a unit increase in genetic risk score of the exposure variable. Results for PSA-BMI and PSA-prostate cancer associations are from the PSA genetic risk score, not the log-PSA genetic risk score	212

List of Tables

Table 3.1 Weighted scores of seven and above from deduplication of all 10,444 duplicate references	56
Table 3.2 Total number of references that mentioned each of the 12,868 keywords in conjunction with prostate cancer or PSA.....	60
Table 3.3 The number of references mentioning keywords referring to individual characteristics in conjunction with prostate cancer or PSA	61
Table 3.4 Results of Expert Opinion.....	63
Table 3.5 Variables that were well-studied, and/or thought to have strong associations with prostate cancer and PSA on expert opinion	64
Table 4.1 The different methods of calculating standardised effect estimates and the number of studies on which the methods were applied.....	75
Table 4.2 Data extracted from studies examining the association between BMI and prostate cancer	90
Table 4.3 Risk of bias of studies examining the association between BMI and prostate cancer (see Section 2.3.2 and Appendix 5)	94
Table 4.4 Data extracted from studies examining the association between BMI and advanced prostate cancer.....	100
Table 4.5 Risk of bias of studies examining the association between BMI and advanced prostate cancer (see Section 2.3.2 and Appendix 5).....	101
Table 4.6 Data extracted from studies examining the association between BMI and PSA	106
Table 4.7 Risk of bias of studies examining the association between BMI and PSA (see Section 2.3.2 and Appendix 5)	107
Table 4.8 Results for Kurahashi et al. (100)	118
Table 5.1 Imputation models.....	133
Table 5.2 Summary of included studies for original data	136
Table 5.3 Summary of missing information	146
Table 5.4 Summary of included studies for imputed data, averaged over 100 imputed datasets	147
Table 5.5 Results for the linear analysis of the association between BMI and prostate cancer	149

Table 5.6 Results for the linear analysis of the association between BMI and prostate cancer	151
Table 5.7 Results for the linear analysis of the association between BMI and PSA (direct and total effects)	153
Table 5.8 Results for the linear analysis of the association between prostate cancer and PSA	156
Table 5.9 Results for the cubic analysis of the association between BMI and prostate cancer	158
Table 5.10 Results for the cubic analysis of the association between BMI and advanced prostate cancer	160
Table 5.11 Results for the cubic analysis of the association between BMI and log-PSA.....	162
Table 6.1 Summary BMI and number of men with and without prostate cancer for normal, overweight and obese BMI categories.....	179
Table 6.2 Results for the categorical analysis of the association between BMI and prostate cancer	179
Table 6.3 Summary BMI and number of men with and without advanced prostate cancer for normal, overweight and obese BMI categories	183
Table 6.4 Results for the categorical analysis of the association between BMI and advanced prostate cancer.....	183
Table 6.5 Summary BMI and log-PSA values for normal, overweight and obese BMI categories	187
Table 6.6 Results for the categorical analysis of the association between BMI and PSA.....	187
Table 7.1 Table of clinical characteristics for each study.....	209
Table 7.2 Results for the MR of outcomes (columns) against genetic risk scores (rows), for all men and men without prostate cancer	211
Table 7.3 Results for the MR of confounders against genetic risk scores for men without prostate cancer only.....	213
Table 7.4 Table showing the results of LR tests for each genetic risk score (row) against each outcome (column).....	214
Table 8.1 Summary of overall results	223
Table 8.2 Results from Chapter 4, systematic review and AD meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA.....	226
Table 8.3 Clinical results from Chapter 5, IPD meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA.....	227

Table 8.4 Results from Chapter 6, combining AD and IPD to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA	228
Table 8.5 Results from Chapter 7, using MR to examine the causal relationships between BMI, prostate cancer risk, advanced prostate cancer risk and PSA.....	229

CHAPTER 1. INTRODUCTION

1.1. Aim of this Thesis

The aim of this thesis is to identify plausible individual characteristics that have an association both with prostate cancer and prostate-specific antigen (PSA), choose one modifiable characteristic that has been well-studied in relation to both prostate cancer and PSA, and then precisely estimate the associations between this characteristic, prostate cancer, advanced prostate cancer and PSA, accounting for any mediation between the characteristic and PSA through prostate cancer. This research could be helpful in making PSA testing for prostate cancer more sensitive and specific, resulting in fewer prostate biopsies in men without prostate cancer, and more biopsies in men with cancer. This could be either through characteristic-specific PSA thresholds for determining whether a prostate biopsy is required, for example age-adjusted PSA categories, or through adjustments to PSA values to account for differences in the characteristic, for example by adjusting PSA to be lower in older men.

1.2. Background for Prostate Cancer and Prostate-Specific Antigen

The aim of this chapter is to provide background information on the prostate, prostate cancer, PSA, and the role of PSA in prostate cancer diagnosis. I will also describe how this research could help to make PSA testing better at detecting prostate cancer. Finally, I will describe the structure and content of this thesis.

1.2.1. The Prostate

About the size of a walnut, the prostate normally weighs around 11 grams (range 7 to 16 grams, and is located beneath the bladder, surrounding the urethra. The function of the prostate is to produce a thick, slightly alkaline fluid which nourishes and aids spermatozoa, accounting for approximately 30% of semen.

Various pathologies affect the prostate. Benign prostatic hyperplasia (BPH) is a common complaint; as a man ages, the prostate tends to increase in size, obstructing the urethra and inhibiting urination, although this is rarely painful. Prostatitis, inflammation of the prostate, can occur in men of any age, although is most common in men aged between 30 and 50 years. In contrast to BPH, prostatitis is characterised by pain, especially when urinating or ejaculating, as well as the difficulties in urination that accompanies an increase in the size of the prostate.

1.2.2. Prostate Cancer

Prostate cancer is the most serious common pathology of the prostate. The International Agency for Research on Cancer (IARC, part of World Health Organisation (WHO)), estimates prostate cancer is the second most common cancer in men globally, with an estimated 1.1 million diagnoses in 2012 (15% of cancer diagnoses in men) (1). In the United Kingdom (UK), prostate cancer is the most commonly diagnosed cancer in men, with an estimated 46,690 diagnoses in 2014 (2). In general, most prostate cancers are slow growing, but can metastasise to the bones, lungs and brain. Worldwide, there were an estimated 307,000 deaths from prostate cancer in 2012 (1), and in the UK, 11,287 men died from prostate cancer in 2014 (3).

Prostate cancer incidence varies dramatically depending on the population. By region, there was almost a 25-fold difference in the age-standardised rate of prostate cancer between Australia/New Zealand (111.6 cases per 100,000) and South-Central Asia (4.5/100,000) in 2012 (1). Ethnicity and environment are both associated with prostate cancer incidence: in the United States of America (USA), the Surveillance, Epidemiology and End Results program (SEER) estimated the age-adjusted incidence of black men (214.5/100,000) was almost three times that of Asian/Pacific Island men (74.0/100,000) (4). Prostate cancer risk is also strongly associated with increasing age (4), thus the age distributions of different regions will also have effects on the incidence. Any study of prostate cancer must therefore stratify by age and ethnicity when making comparisons between populations, as if the distributions of age and ethnicity are different between men with and without prostate cancer there could be confounding in the measured association between the characteristic and prostate cancer.

Gleason score

As prostate cancer advances, several changes happen to the cancerous cells. The first change is that the cells become less specialised in function, and this allows histopathologists to grade a sample of a prostate tumour from a biopsy. When a tumour first develops, the cells are generally similar in appearance to normal prostate cells, except their replication is uncontrolled. As the cells acquire more mutations they become less specialised and more uniform in appearance and function.

A Gleason score is given to each prostate tumour analysed by histopathologists. The score is comprised of two numbers equalling a total, the first number representing the Gleason's Pattern that comprises most of the sample, and the second number representing the next most prevalent Pattern. For example, a biopsy may be given a Gleason score of 3+4=7, which means most of the sample is comprised of cells that have "distinctly infiltrative margins", and a lesser number of cells would have "irregular masses of neoplastic glands". Gleason scores are typically used to help in deciding whether

treatment is necessary, and whether the tumour has progressed if repeat biopsies are performed. **Figure 1.1** shows a visualisation of Gleason's Pattern for classifying prostate cancers (5). Prostate cancers are usually considered as high-grade if the Gleason score is 8 or higher.

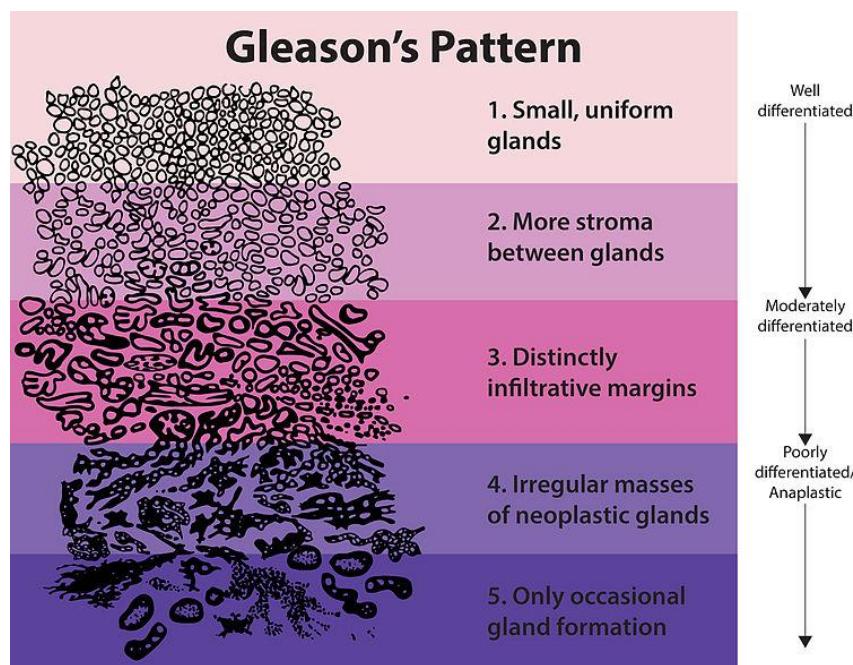


Figure 1.1 A visualisation of Gleason's Pattern for classifying prostate cancers

Prostate cancer staging

Prostate tumours themselves are staged using the TNM (Tumour, Node, Metastases) classification of malignant tumours. As the tumour progresses, so does its T (Tumour) score. At T1, the tumour is clinically unapparent. At T2, the tumour is confined to the prostate gland. At T3, the tumour has breached the capsule of the prostate, and at T4, the tumour has spread to adjacent structures, for example the bladder or surrounding muscle.

Once the requisite set of mutations have occurred to allow tumour cells to survive outside of the prostate, cells will initially be deposited and grow in lymph nodes (N1), and then may grow in other tissues (M1) such as bones, the liver, the lungs and the brain. Once tumour cells begin to grow outside the prostate and nodes, the cancer is metastatic and standard treatment options are no longer curative. Prostate cancer is often defined as *advanced* if the T-score is above 2 (sometimes called *locally advanced*), or nodes or metastases are present (N1 or M1), although this definition may vary between studies. Prostate cancer is *localised* if the T-score is less than 2, and nodes and metastases are not present.

Some tumours never progress beyond T1 or T2, while others are aggressive and metastasize before treatment can begin. Presently, there is no accurate way to predict when or whether a tumour will

metastasise (6). This makes it difficult to treat prostate cancer: it is impossible to know if the cancer will prove fatal and so require treatment, or never cause a problem to the man and so any treatment would be overtreatment. Overtreatment is a large public health problem, as it is commonly quoted that “more men die with prostate cancer than of it” (7).

1.2.3. Prostate Specific Antigen

PSA is a glycoprotein enzyme produced (almost) exclusively by the prostate. The role of PSA is to liquefy semen through proteolysis (protein destruction), which helps the semen move more freely.

PSA is not secreted directly into the blood when the prostate is undamaged, so young, healthy men have very low concentrations of PSA in the blood. However, as men age the prostate cellular walls become more damaged, and more PSA is inadvertently released into the blood (8,9). In addition, BPH, prostatitis and prostate cancer all cause damage to the prostate, and so increase PSA concentrations in the blood. As prostate cancer advances, PSA levels generally rise from relatively low levels of less than 3.0 or 4.0 ng/ml (8,10–12), up to the potentially extremely high levels of more than 1,000 ng/ml (13).

PSA has been proposed as a marker for tracking progression of prostate cancer for this reason (14), but a recent study showed that in men from US and European populations, there was no difference in the annual PSA percentage change over time between men without prostate cancer, and men with PSA-detected prostate cancer (15). PSA has also been proposed as a marker of prostate cancer recurrence after treatment (16). Here, PSA is a useful marker as if a man has their prostate removed, their blood PSA concentration should be undetectable as no other tissues are generally capable of producing PSA. If their PSA concentration starts to rise, however, then this indicates the recurrence of cancer (or the presence of metastases) (17). PSA is also used as a test for prostate cancer, both in men who have symptoms of prostate cancer, and in screening men who have no symptoms (18).

PSA screening for prostate cancer

PSA has also been proposed as a marker for presence of prostate cancer (19). PSA screening for prostate cancer is explicitly testing for cancer in the absence of symptoms of cancer. Current guidelines on the use of PSA screening vary by country, but in general screening for prostate cancer using PSA tests is not encouraged and the decision to have a PSA test should be shared between the patient and clinician (20). In practice, because the symptoms of prostate cancer are similar to those of BPH and prostatitis, and both diseases are common, PSA tests are performed frequently even in the absence of a formal screening programme; in the UK, men without prostate cancer and aged between 45 and 69 years in 2002 were found to have a 39.2% (95% CI 39.0 to 39.4%) chance of having

a PSA test over a 10-year period (21). Current guidance in the UK is to offer a man a PSA test if he requests it, as well as if he presents with symptoms (18).

As prostatitis (22), diet (23), ethnicity (24), and various drugs (25,26) have all been shown to be associated with PSA, high levels of PSA do not necessarily indicate prostate cancer. The American Cancer Society state that for a PSA test with a 4.0 ng/ml threshold, the sensitivity and specificity for detecting any prostate cancer are 0.21 and 0.91 respectively (27). This means that for men in general, 9% of men *without* prostate cancer have a PSA above 4.0 ng/ml, and 21% of men *with* prostate cancer had a PSA below 4.0 ng/ml. These results were based on the Prostate Cancer Prevention Trial (PCPT) (28), which offered prostate biopsies to all men in the study regardless of PSA levels. The proportion of men without prostate cancer who have a PSA above the threshold for biopsy will depend on the population being tested, as PSA is associated with many variables and the underlying prevalence of prostate cancer may be different.

In addition to causing many men to receive unnecessary biopsies, allowing men to request a PSA test can cause inequality in who has a PSA test. This affects the chance of a man receiving a biopsy, and therefore potential diagnosis and treatment of prostate cancer. In the UK, the chance of receiving a PSA test between 2002 and 2011 was associated with a man's deprivation (as well as age, region and any previous tests): among the least deprived, the chance of receiving a PSA test was 46.3%, whereas among the most deprived, the chance was only 31.9% (21). If PSA could be improved as a test for prostate cancer and adopted as a screening test (rather than relying on men asking for a test), more men may be offered the test and inequality reduced.

However, increasing the number of screening tests may in turn increase overtreatment of low-risk disease (29). Several studies have assessed the benefits of PSA as a screening test for prostate cancer where a positive test result (e.g. ≥ 4 ng/ml) leads to a prostate biopsy (29–31). Even so, the overall evidence is that there is no clear advantage to using PSA tests to screen all men for prostate cancer, owing in part to the poor sensitivity and specificity of PSA as a test for prostate cancer, and also to overtreatment. Even in the absence of a formal screening programme in the UK or evidence of an advantage to using PSA screening, the proportion of men currently being offered a PSA test is high, and would therefore be worthwhile to improve PSA testing for prostate cancer. This would reduce the number of healthy men undergoing an unnecessary prostate biopsy, and increase the number of men with prostate cancer receiving a biopsy.

1.2.4. Improving PSA as a Test for Prostate Cancer

One method of improving PSA testing is to adjust PSA test results for the effect of other variables on PSA, making PSA more specific to prostate cancer. Previous studies consistently show that age is associated with a higher PSA, and this has been accounted for in PSA testing using age-bands of biopsy thresholds; as a man ages, the PSA threshold for biopsy also increases (18). PSA has also been associated with: BMI (9); diet (23); ethnicity (24); genetics (32); certain drugs (including opiates (25), diuretics (26) and statins (33)); and disease (including prostatitis (22), BPH (34) and diabetes (35)).

Removing the effect of each of these variables on PSA would make PSA a better test for prostate cancer. However, many of the variables associated with PSA have also been associated with prostate cancer: age (36), BMI (37), diet (38), ethnicity (39), prostatitis (40), BPH (41) and diabetes (42). Since prostate cancer increases PSA levels, some of the association between the variables and PSA may go through prostate cancer, making it difficult to account for just the direct effect of the variable on PSA. For example, older men have a greater risk of prostate cancer, so also have on average a higher level of PSA. To improve PSA as a test for prostate cancer, only the effect of age on PSA must be accounted for, not the effect on PSA from having a greater prevalence of prostate cancer in older men.

The problem becomes more complex as the chance of receiving a prostate biopsy is often dependent on PSA levels. This means that some (or all) of the estimated association between a variable, such as age, and prostate cancer may be due to its association with PSA, and how this association changes the chance of receiving a prostate biopsy. This is because prostate cancer status is only observed in men diagnosed with prostate cancer, so in a study with no screening for prostate cancer (formal or otherwise), men will only be diagnosed if they present with symptoms, are offered a biopsy, and that biopsy is positive. In a study where the population is screened for prostate cancer using a PSA test, men will be diagnosed if their PSA is above the threshold, and their biopsy is positive. For example, age may also be associated with a higher level of PSA even in the absence of prostate cancer, so older men are more likely to have a PSA above the biopsy threshold, and so more likely to be offered a biopsy than younger men. Older men are thus more likely to be diagnosed with prostate cancer, all else being equal. This creates an apparent positive association between age and prostate cancer due to differences in diagnosis, even if no true effect of age on prostate cancer status existed. Therefore, the observed positive association between age and prostate cancer could be from the effect of age on PSA, and equally, the observed association between age and PSA could be from the effect of age on prostate cancer risk. This is summarised in **Figure 1.2**.

The aim of this thesis is to precisely estimate the association between a modifiable variable, prostate cancer and PSA, so that the effect of the variable on PSA may be accounted for in future research and PSA testing.

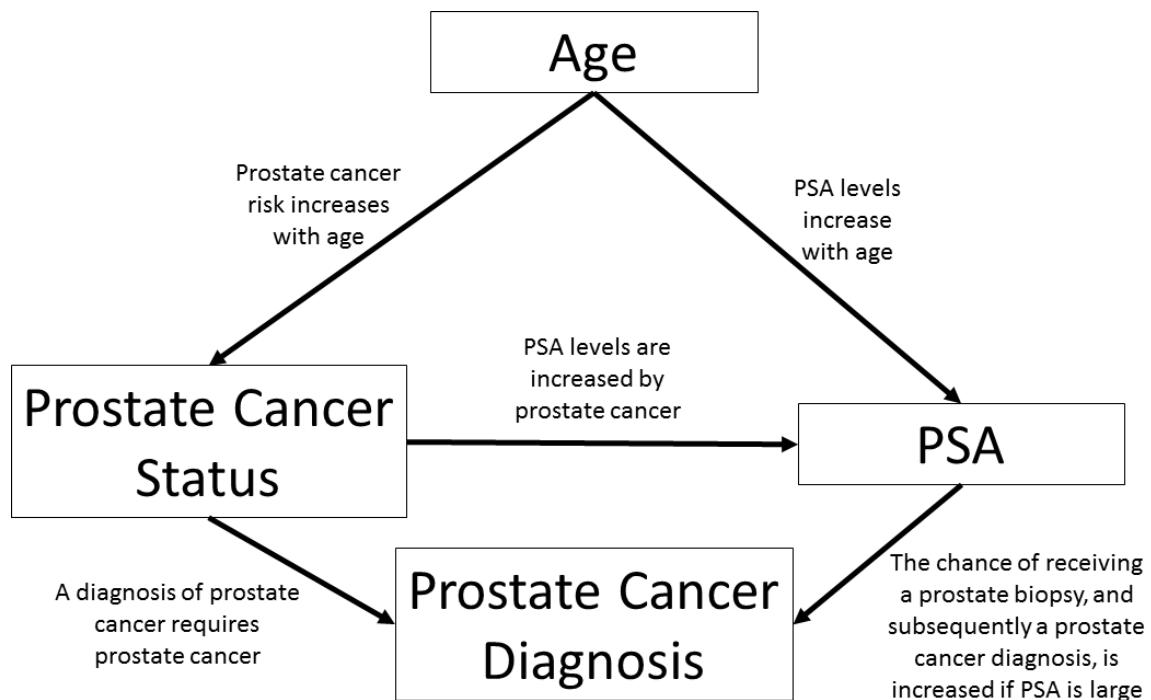


Figure 1.2 Diagram showing the associations between age, prostate cancer status, prostate cancer diagnosis and PSA

1.3. Description of Thesis Chapters

Chapter 2: Evidence synthesis methods

Aim: The aim of this chapter is to detail the methods of synthesising data from different studies to inform an overall conclusion that are relevant to this thesis. Evidence synthesis methods can be applied to any collection of data, but to reduce the potential for bias, a systematic review can be conducted to find all relevant studies pertaining to a research question. I discuss how systematic reviews are conducted and why they are important for evidence, then describe existing methods of evidence synthesis relevant to this thesis. Finally, I discuss a novel method of evidence synthesis I developed for when data could not be combined in meta-analysis, the albatross plot.

Output: A review of evidence synthesis methods relevant to this thesis, and a description of the albatross plot.

Chapter 3: Creating and using a Sankey diagram to narrow the focus of a systematic review of prostate cancer and prostate-specific antigen

Aim: The aim of this chapter is to find individual characteristics that have well-studied associations with both prostate cancer and PSA, use expert opinion to guide us toward a characteristic known to have strong associations with both prostate cancer and PSA, so we can then identify a single well-studied, modifiable, characteristic to study further. This chapter comprises a scoping study, a description of the practical issues that forced novel solutions (a probabilistic matching algorithm), a Sankey diagram, and expert opinion that was sought to complement the results of the diagram.

Output: A selection of individual characteristics that are well-studied in relation to their associations with both prostate cancer and PSA, with one characteristic (BMI) selected for further examination. Also, a description of the probabilistic matching algorithm I developed to find duplicate references that may contain typographical or other errors.

Chapter 4: Systematic review and meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA

Aim: The aim of this chapter is to systematically review the literature which examines the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA. Data were extracted from all relevant papers identified in **Section 3.3** and aggregate data (AD) meta-analyses performed.

Output: Systematic reviews and meta-analyses of the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA.

Chapter 5: Individual participant data meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA

Aim: The aim of this chapter is to conduct an individual participant data (IPD) meta-analysis to quantify the associations between: (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, (3) BMI and PSA, and (4) prostate cancer and PSA. Using individual participant data (IPD) allows us to investigate these associations more thoroughly than with AD alone as all covariates (that were measured) and missing data can be accounted for, allowing us to control for potential confounding and bias using the same set of covariates in all studies. In addition, non-linearity of the associations and possible interactions can be assessed.

Output: IPD meta-analyses of the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, (3) BMI and PSA, and (4) prostate cancer and PSA.

Chapter 6: Combining aggregate data and individual participant data to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA

Aim: The aim of this chapter is to combine the AD and IPD to calculate estimates of all associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA. I estimated both linear and non-linear associations.

Output: Meta-analyses using both AD and IPD of the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA.

Chapter 7: Using Mendelian Randomisation to examine the causal effect of BMI on prostate cancer risk and PSA

Aim: The aim of this chapter is to use Mendelian randomisation (MR) to determine whether BMI has a causal effect on prostate cancer risk, advanced prostate cancer risk or PSA, and to examine whether prostate cancer or PSA could have a causal effect on BMI. In the previous three chapters, we found little evidence that BMI was associated with prostate cancer, but strong evidence that BMI was associated with a reduced PSA, likely non-linearly. However, as all evidence was observational, causality cannot be established from these studies alone.

Output: Assessment of whether the associations between BMI and prostate cancer, BMI and advanced prostate cancer and BMI and PSA are causal.

Chapter 8: Discussion

Aim: In this chapter, I present overviews of all results and methodologies, compare the results with previous research, and consider the strengths and limitations of this thesis as a whole. Finally, I discuss

the implications of this research and potential future directions this work could take, and end with an overall conclusion.

Output: A discussion and conclusion of this thesis.

1.4. Objectives

The objectives of this thesis are to:

Chapter 1	Provide background information on prostate cancer and PSA, as well as using PSA as a screening test for prostate cancer.
Chapter 2	Describe evidence synthesis methodologies relevant to this thesis.
Chapter 3	Identify individual characteristics that have a plausible association with both prostate cancer and PSA, and have a large body of evidence examining this association, then select a characteristic to examine further.
Chapter 4	Perform a systematic review and aggregate data meta-analysis of studies examining the associations between the chosen characteristic, prostate cancer, advanced prostate cancer and PSA.
Chapter 5	Identify and collect data from large, well conducted prostate cancer studies, then perform an individual participant data meta-analysis of the associations between the characteristic, prostate cancer, advanced prostate cancer and PSA.
Chapter 6	Combine the results of the aggregate and individual participant data to estimate the associations between the characteristic, prostate cancer, advanced prostate cancer and PSA as precisely as possible.
Chapter 7	Perform a Mendelian randomisation analysis to assess evidence for causality between the characteristic, prostate cancer, advanced prostate cancer and PSA.
Chapter 8	Summarise the results, strengths and limitations from the thesis, and indicate what direction future work may take.

CHAPTER 2. EVIDENCE SYNTHESIS METHODS

2.1. Aims

The aim of this chapter is to detail the methods of synthesising data from different studies to inform an overall conclusion that are relevant to this thesis. Evidence synthesis methods can be applied to any collection of data, but to reduce the potential for bias, a systematic review can be conducted to find all relevant studies pertaining to a research question. I discuss how systematic reviews are conducted and why they are important for evidence, then describe existing methods of evidence synthesis relevant to this thesis. Finally, I discuss a novel method of evidence synthesis I developed for when data could not be combined in meta-analysis, the albatross plot. Julian Higgins, Hayley Jones, Sarah Lewis and Richard Martin all contributed towards the development of the albatross plot.

2.2. Systematic Reviews

Reviewing previous evidence before starting new research is essential (43). It is not sufficient to assume that any individual study has the correct answer to a research question, as random chance, problems with study design (e.g. bias, confounding), and the representativeness of the study population can mean the results from studies do not represent the truth. Although conducting a large well-conducted study with a diverse population could answer a research question, a systematic review of previously published studies would likely cost less, take less time and give a more generalisable result (44). If high-quality evidence for a research question already exists, choosing to not use published data could potentially waste resources and could be considered unethical. A systematic review uses a defined search strategy to attempt to find all relevant literature to answer a clearly specified research question. The search strategy should be published with the review, allowing others to replicate the search (45).

The process of a systematic review is commonly specified in a protocol, which may or may not be published before the review itself, and generally follows an accepted methodology. First, the research question is defined in terms of the population, intervention (or exposure), control (or exposure) and outcome of interest (PICO). The inclusion criteria for studies is set using PICO and any other information that may be relevant. Second, existing literature is searched using a search criterion designed to find as many relevant studies as possible, while also minimising the number of irrelevant studies. This has been made much easier through the implementation of reference databases, which record a high proportion of all published literature in different fields. Data sources that are not formally published with peer-review, such as technical reports, conference abstracts and other non-peer-reviewed literature are often searched along with peer-reviewed literature. These sources are called grey literature, and their inclusion may help to reduce publication bias, where negative results are not published as frequently as positive results, or prejudice for or against particular hypotheses by journal editors (46).

Studies found with the literature search are assessed against the inclusion criteria. Initially, titles and abstracts are assessed, followed by full-texts. This reduces the burden of collecting full-texts for all studies, which can be a time-consuming and costly process depending on which journals an institution has access to. Studies fulfilling the inclusion criteria are assessed for their methodological quality in terms of the likely biases the study may have, for example randomised trials can be assessed using the Cochrane risk of bias tool (47), and data are extracted. Data may then be combined or summarised in some way; how the data is combined depends on the data itself, and is discussed in the following sections. Once combined, a conclusion is made in the context of the research question. Systematic

reviews often include more than one author, and much of the work is duplicated to avoid missing any useful papers and to ensure accuracy when extracting data.

2.3. Methods for Evidence Synthesis

In the following section, I describe existing methods of evidence synthesis that do not require an effect size to be reported in each study, and different meta-analysis methods, which do require a reported effect size. Methods that do not require an effect size are potentially more inclusive than those that do, as they require less information, but they also are unable to estimate a combined effect size and are thus only used for determining whether there is evidence for or against a hypothesis. During this thesis, I developed a novel method of evidence synthesis (the albatross plot) that does not require an effect size, but nonetheless allows rough estimation of the typical effect magnitude for an association between two variables, and this is described last in this section.

2.3.1. Evidence synthesis methods that do not require an effect estimate

Narrative synthesis

Narrative synthesis consists of summarising the methods and results of each study to give an overall impression of the study's conclusion. Once all studies have been summarised, an overall summary may be given, answering the research question. In general, narrative syntheses can be difficult to interpret and there is a risk that conscious or unconscious bias may affect the way in which the results are presented (48). As such, narrative syntheses are generally reserved for situations where the analyses of included studies are presented in such a way that estimation of a comparable effect estimate for every study is impossible, and other statistical information are unavailable, for example P values or the total number of participants. Narrative synthesis may also be used to complement other forms of evidence synthesis (48). Alternatively, a narrative synthesis may be performed if the studies are considered too different to be combinable in any meaningful way, for example in analysis type or study design. Narrative synthesis is the most inclusive method of evidence synthesis, as it requires only that the studies answer the same underlying research question.

Vote Counting and Sign Tests

Vote counting consists of counting the number of studies reporting a positive, negative or null association using a pre-defined P value threshold (such as $P = 0.05$). A summary of how many studies had positive, negative and null associations is then presented, and conclusions drawn as to whether there is evidence for or against the association in the research question. Needing only P values, this method of evidence synthesis is inclusive and simple to perform.

Sign tests use a similar principle, but instead of counting studies using a P value threshold, the direction of effect is counted as being positive, negative or mixed regardless of either the size or

significance of the effect. This makes sign test more inclusive even than vote counting, but dismisses much of the information from each study.

Both vote counting and sign tests provide no indication of the magnitude of any associations, and information is lost through either the dichotomisation of P values into “significant” and “not significant” or the disregard of the P value entirely, and as such the methods have been widely criticised (49–51). For vote counting it is also the case that the P value of a study is dependent upon the number of participants in the study, so larger studies could skew the results in favour of an association just by being large. However, sign tests are only marginally less inclusive than narrative reviews, requiring only the effect direction of an association, and vote counting only requires a P value in addition to the effect direction, making both methods amenable to evidence synthesis of limited or poorly reported data.

Harvest Plot

Harvest plots are graphical representations of vote counting, with the extension that each study is assigned a measure of confidence in the results, so that studies the researcher determines to be more reliable (such as randomised controlled trials, RCTs) are given more weight (52). Since the plot is informed by vote counting, this method of evidence synthesis has the same disadvantages of arbitrary dichotomisation of P values and no indication of an actual effect size. However, by weighting the studies according to confidence in the results, a more complete picture is drawn than relying on vote counting alone. The Harvest plot requires more information than vote counting to assess the reliability of included studies, but is still a very inclusive method of evidence synthesis.

Combining P values

Combining P values using the methods of Fisher (53) or Stouffer (54) allows for a more statistical approach to evidence synthesis than with previous methods. Fisher’s and Stouffer’s methods pool all P values and produces an overall P value that can be used to test a null hypothesis (that there is no effect or association). Fisher’s method is shown below:

$$X_{2k}^2 = -2 \sum_{i=1}^k \ln(p_i) \quad (1)$$

where p_i is the *one-sided* P value for the i^{th} study, and X^2 is the test statistic, which has a chi-squared distribution with $2k$ degrees of freedom, where k is the number of studies. When the P values of studies tend to be small, X^2 will be large, and thus the combined P value will be low, indicating there is evidence against the null hypothesis that there is no association. Stouffer’s method is similar but uses a normal distribution instead of a chi-squared distribution.

The advantage of these methods is that all information from the P value of each study is used, unlike vote counting which dichotomises the P value into significant or non-significant with an arbitrary threshold. The disadvantage is that once again no indication is provided of the magnitude of effect, and any heterogeneity across the studies can be disguised in the combined P value. The formula also requires one-sided P values, which may not be reported, although a reported two-sided P value can be transformed to a one-sided P value if the effect direction is known. Combining P values is potentially slightly less inclusive than vote counting, as specific P values are required rather than P less or more than a threshold (e.g. P < 0.05), but these methods are still very inclusive.

2.3.2. Meta-analysis

Aggregate data meta-analysis

An aggregate data (AD) meta-analysis combines reported effect estimates from published studies to estimate an overall effect estimate (49). A weight is given to each study, often according to the inverse variance of the study; as variance is affected by the number of participants in a study, studies with the largest number of participants tend to have the largest weight. I describe the most basic and widely-used approaches to AD meta-analysis; alternative weighting schemes are available, and other methods can handle particular data type more appropriately, e.g. by using binomial likelihoods for count data.

Individual Patient Data Meta-analysis

Individual patient data (IPD) meta-analysis differs from AD meta-analysis in that instead of using published effect estimates to estimate the overall effect estimate, individual level data is used. One key advantage of this method is that the precise analysis of interest for the meta-analysis can be performed in all studies, for example ensuring that all studies have adjusted for the same set of potential confounders. The summary effect estimate for IPD meta-analysis can be calculated in either one or two-stages. Two-stage IPD meta-analysis is equivalent to AD meta-analysis, where effect estimates are estimated in each study individually and then combined, whereas in one-stage IPD meta-analysis the summary effect estimate is calculated in one analysis, accounting for clustering within studies. In general, combined estimates from one- and two-stage IPD meta-analyses are very similar (55). The following section is applicable for two-stage IPD meta-analysis, while one-stage IPD meta-analysis uses slightly different methods.

Fixed-effect and random-effects meta-analyses

Both AD and IPD meta-analyses can be performed assuming fixed-effect or random-effects models. In a fixed-effect meta-analysis, all studies are assumed to have estimated the same underlying effect,

such that any differences in the effect estimates for each study are from random (sampling) error (49). There are various options for weighting schemes, but I will describe the most common, which is the *inverse variance* weight scheme. A weight of $\frac{1}{v_i}$ is given to each study, where v_i is the variance of the effect estimate of the i^{th} study. The effect estimate should be transformed to an approximately normal measure prior to taking the weighted average, for example odds ratios (ORs) are transformed to log-ORs. The variance of the transformed effect estimate is used to weight each study, and the weighted summary effect estimate is estimated, which if necessary can then be back-transformed for interpretability.

Fixed-effect meta-analyses are performed when one can be confident that the underlying effect is constant between studies, which in practice often means that populations being combined are similar or the effect of an intervention is unlikely to be affected by any differences in population.

A random-effects meta-analysis allows for study-specific effects, and assumes that the true effect estimated by each study varies and is assumed to be randomly and normally distributed on the transformed scale (e.g. log-ORs) between studies, with a variance of τ^2 . The between-study variance is estimated (usually using the “method of moments” or DerSimonian and Laird estimate), and used to modify the weights of studies when calculating the summary effect estimate and variance. The weight of each study is $\frac{1}{v_i + \tau^2}$ where v_i is the variance of the i^{th} study and τ^2 is the estimate of the between-study variance. So long as the between-study variance is larger than 0, smaller studies will be given more weight in random-effects meta-analysis than in a fixed-effect meta-analysis. If the between-study variance is 0, then the random-effects model is equivalent to the fixed-effect model.

In both fixed-effect and random-effects inverse-variance meta-analysis, the summary effect estimate and variance are estimated in the same way. The effect estimates (θ_i) from each study can be (transformed) ORs, relative risks (RRs), standardised mean differences (SMDs) or other types of estimates:

$$\hat{\theta}_{Summary} = \frac{\sum w_i \theta_i}{\sum w_i} \quad (2)$$

$$Variance\ of\ \hat{\theta}_{Summary} = \frac{1}{\sum w_i} \quad (3)$$

where θ_i is effect estimate and w_i is the weight from the i^{th} study on an approximately normal scale.

Forest plots

The results of meta-analyses are often displayed in a forest plot (56). The forest plot shows the effect estimate, confidence interval (CI, usually 95%) and weight of all included studies, as well as the summary result and its SE. Studies which are not consistent with the summary estimate can be seen, and possibly described to give an indication as to why they might be inconsistent.

Heterogeneity and inconsistency

There can be heterogeneity between studies when the studies are different in some way, for example in terms of their population, measurements, interventions or follow-up times. These differences may cause the effect estimates to be different across studies, beyond that explained by random sampling variation, as the effect estimate may change depending on these differences. Random-effects meta-analysis allows for heterogeneity across studies.

Inconsistency between the study effect estimates can be measured, which gives an indication how much of the observed variation between effect estimates is due to natural error from chance, and how much is from heterogeneity between studies. The I^2 statistic (57) is used for this purpose, which estimates the percentage of the total variation across studies that is due to heterogeneity rather than chance. While the amount of heterogeneity itself can be difficult to interpret as it depends on the scale of the (transformed) effect estimates, I^2 lies on a 0-100% scale. A high I^2 statistic, while technically indicating ‘inconsistency’, is widely taken as a proxy for heterogeneity.

Meta-regression can be used to assess whether study-level covariates are associated with the effect estimate in each study, and thus how much heterogeneity can be explained by study-level differences (58,59). For example, the effect estimates may be stronger in studies with older participants, and when accounted for the residual heterogeneity may be lower across all studies. The Q-test for heterogeneity can also be conducted to assess any heterogeneity between subgroups in a meta-analysis (49).

Funnel plots are scatter plots of the (transformed) effect estimates from studies against the precision of the estimates (usually the SE of the estimate), where asymmetry of the plot (small study effects) can be assessed (60–62). Asymmetry in the plot can be for a number of reasons: publication bias, poor quality studies (including not adjusting for confounders), true heterogeneity, artefacts and chance (62). Interpreting asymmetry in funnel plots is dependent on the context, which may make certain reasons for asymmetry more likely, and symmetry in the funnel plot only implies no evidence of publication bias etc., not that these biases do not exist.

The risk of bias can also be assessed in included studies. In general, the aim of a risk of bias assessment is to quantify the risk of bias in specific domains, such as (for aetiological studies) confounding, selective reporting, exposure and outcome measurement, selection of participants and missing data. Often an overall risk of bias for each study is given, and the risk of bias may be assessed by more than one reviewer to ensure consistency, as the risk of bias may not always be clear. Different tools are available to help assess risk of bias, for example the Cochrane risk of bias tool (47) can be used with RCTs, the ROBINS-I tool (63) can be used with non-randomised studies, and the Critical Appraisal Skills Programme (CASP) case-control and cohort questionnaires can be used with observational studies (64,65).

Advantages, issues and limitations

The main advantage of meta-analysis over the methods of evidence synthesis that do not use the effect estimate is the ability to calculate summary effect estimates with CIs, pooling data from all included studies into a single estimate. Forest plots also provide a graphical representation of the data, and allow for assessment of heterogeneity and bias.

One complication of meta-analysis is the amount of information required. For AD meta-analysis, while the effect estimate and SE for an association (or information to calculate them) are generally well-reported, this might not be true of all studies in a systematic review. As such, studies that could have otherwise provided meaningful information are necessarily excluded, for example studies presenting a P value but no effect estimate. For IPD meta-analysis, it can be difficult to gain access to the IPD from all studies, further limiting which studies can be included, although IPD and AD studies can be combined in a single meta-analysis.

Additionally, AD studies may have adjusted for different confounders in the analysis and it is challenging to include this information in a meta-analysis. Meta-analysis was developed for use with RCTs, which are less likely to be subject to confounding than observational studies. In meta-analyses of observational studies, it is common to include the fully-adjusted results from studies (i.e. results that had the highest number of confounders included as covariates in the analysis method), and sensitivity analyses using less adjusted results can be performed. Generally, studies will have adjusted for different confounders, and in some situations, this could lead to studies giving different results, even if the underlying effect size was the same. In a random-effects meta-analysis this would increase τ^2 and give a larger summary variance. In IPD studies, different confounders may have been measured, limiting which variables can be controlled for in all studies, although missing data can be imputed with IPD (see **Section 5.2**).

2.3.3. Albatross Plot

One of the largest problems with AD meta-analysis is that sufficient information may not be available for all studies to combine in meta-analysis. Previous studies have shown that in RCTs (66), prognostic studies (67) and observational studies (68) outcomes can be poorly reported, and for RCTs and observational studies the likelihood of the effect estimate being reported in a usable way for a meta-analysis is decreased if the association is not statistically significant (66,68). Therefore, meta-analyses may be biased away from the null simply because studies without statistical significance in the association of interest had to be excluded. In this section, I describe a new method, the albatross plot, which I created in part to help to reduce this bias (48). The albatross plot paper is presented in **Appendix 9** in full. I am the first author of this paper, on which much of this section is based. Julian Higgins, Hayley Jones, Sarah Lewis and Richard Martin all made valuable contributions towards the albatross plot, both in development and the paper itself.

The albatross plot is a graphical representation of studies included in a systematic review. One of the key advantages of the plot is that it requires only the number of participants in a study, the P value and effect direction of an association. Because the effect estimate is not required, the albatross plot is more inclusive than a meta-analysis, and can be used either in place of a meta-analysis to estimate an effect magnitude (when meta-analysis is not possible), or in combination with a meta-analysis to assess whether studies excluded from the meta-analysis due to insufficient information were consistent with those that were included. When the effect estimate is not reported, the evidence synthesis methods that do not require an effect size (**Section 2.3.1**) could also be used, but none of these methods provide any indication of the typical effect magnitude, since they do not consider the size of the study that produced each P value, nor can they assess inconsistency between studies with and without sufficient information for meta-analysis.

The basic albatross plot is a scatter plot of study sample sizes against P values, separated into positive and negative associations according to the observed direction of effect. The albatross plot allows the P values to be interpreted in the context of the study sample size, as small studies appear towards the bottom of the plot and large studies towards the top. Small P values from strong negative results appear at the left of the plot and small P values from strong positive results appear at the right of the plot, with studies with null results towards the middle. Both the sample size axis and the P value axis are plotted on the log scale for improved visual interpretation. Many examples of albatross plots are provided in the paper, **Appendix 9**.

Two types of enhancement to the basic albatross allow approximate examination of effect sizes and their heterogeneity. First, we superimpose contours on the plot to reflect different hypothetical effect sizes that would have given rise to particular P values. These contours will be specific to the type of data (and statistical methods) used to calculate the P values, and are to be interpreted very approximately. The contours typically resemble large flying birds, giving rise to the name “albatross plot”. Second, different subgroups of studies can be drawn using different colours or symbols to facilitate identification of subgroup effects.

Generation of effect contours

Albatross plots acquired their name from the effect contours placed upon them, where we aim to plot N as a function of P for a given effect size. These contours are specific to the type of effect estimate used to generate the P values in each study, for example ORs from logistic regression or beta coefficients from linear regression. Approximate effect size contour lines are based on the general assumption that the P values were derived from Wald tests. A Wald test involves division of the effect size estimate (b) by its standard error (SE) to calculate a Z statistic:

$$Z_P = \frac{b}{SE} \quad (4)$$

This statistic is compared to a standard normal distribution to obtain the P value. Conversely, a Z-statistic can be obtained from a (reported) P value using the same distribution, so we write $Z_p \equiv \Phi^{-1}(P)$, where Φ^{-1} denotes the inverse of the standard normal distribution function. For each study, we can calculate Z_p from the given P-value.

In general, the SE is proportional to the inverse of the square root of the total number of participants (N) in the study, so that we can write:

$$SE = \frac{\phi}{\sqrt{N}} \quad (5)$$

The quantity ϕ may be a fixed number, or it may involve the effect size itself, and it may additionally involve other quantities that need to be specified to define an effect size contour uniquely. Whether ϕ is a fixed value or whether other quantities need to be specified depends on the type of effect estimate used to calculate the P value. For example, an SMD with equal group sizes has a fixed value of ϕ , whereas an OR, the value of ϕ is dependent on the baseline risk of the outcome and the ratio of participants between groups (exposed versus unexposed).

Rearranging (4) and (5), we can express the sample size in the form:

$$N = \frac{\phi^2}{b^2} Z_P^2 \quad (6)$$

To obtain a contour corresponding to a hypothetical effect size b , we need to determine the quantity ϕ appropriate for the choice of effect size, and then plot N as a function of ϕ , b and P . Practically, the value of ϕ is fixed (either naturally, or by choice of values of, for example, the baseline risk of the outcome), as is b since the contours are for specific effect sizes.

For some effect sizes, including the SMD for a study with equal group sizes, this effect size is all that needs to be specified to determine ϕ . As the SE of the SMD with equal group sizes is (48):

$$\text{SE} = \sqrt{\frac{8 + \text{SMD}^2}{2N}} \quad (7)$$

We can say:

$$\phi = \sqrt{\frac{8 + \text{SMD}^2}{2}} \quad (8)$$

The effect contours for the SMD are therefore defined as:

$$N = \frac{8 + \text{SMD}^2}{2\text{SMD}^2} Z_P^2 \quad (9)$$

In other cases, further variables must be specified, such as the ratio of group sizes or the baseline risk. Values for the additional variables might be chosen using the most common values in the included studies; for instance, for trials the ratio of group sizes is often 1, so there are equal number of participants in the intervention and control arms of the trial. See the full albatross plot paper for more details (48), **Appendix 9**.

In other cases, adjustment of the sample size is sometimes advisable if the further variables (on which ϕ depends) differ markedly between studies. For example, in an albatross plot of the association between a continuous exposure and binary outcome, the ratio of group sizes may be very different between case-control and cohort studies, as in the former the ratio is generally specified (e.g. 1 control for every case), but in the latter the ratio is dependent on the incidence of the outcome. To make the studies more comparable in an albatross plot, the sample size of each study can be adjusted to the effective sample size, the sample size the study would have required to give the same P value and

effect estimate if the ratio of group sizes (or baseline risk etc.) were different. Adjusting to the effective sample size is an effective way of making individual studies in the albatross plot more comparable.

Strengths and limitations

Albatross plots provide a clear summary of the collected evidence in a review. It is visually clear if the studies generally convey a positive or a negative effect size, as in this case the studies will cluster on one side of the plot. If there is no association, points will fall evenly around the null in the centre of the plot. If the studies have generally a similar effect size, the points will fall around an effect size contour; if there is heterogeneity of effect size, then the points will be scattered across contours. Studies with smaller sample sizes will have larger variances than equivalent studies with more participants, resulting in both more variation around the true value and points that are more clustered around the null than larger studies (since for the same effect size, the P value will be larger and thus closer to the null). However, many small studies can still point towards a single effect size contour if the underlying effects are homogeneous.

Furthermore, outlying studies can be identified with ease, and a brief narrative synthesis (**Section 2.3.1**) conducted alongside the albatross plot might propose possible explanations for the findings in these studies. In addition, albatross plots can also be used as a sensitivity analysis with a meta-analysis. Sometimes the information required for meta-analysis is not available from studies (the effect estimate or SE may not be given), meaning they are excluded from the meta-analysis. An albatross plot requires less information, and thus could include both the studies that had information for meta-analysis and the studies that were excluded, to determine if there were any meaningful differences between studies with and without enough information for meta-analysis.

2.4. Summary

In this section, I have described different methods of evidence synthesis without using the effect size, meta-analysis, and albatross plots, the novel method I developed for graphically representing studies included in a systematic review. Albatross plots are particularly useful for assessing whether studies excluded from a meta-analysis due to insufficient information are consistent with studies that were included.

CHAPTER 3. CREATING AND USING A SANKEY DIAGRAM TO NARROW THE FOCUS OF A SYSTEMATIC REVIEW OF PROSTATE CANCER AND PROSTATE-SPECIFIC ANTIGEN

3.1. Aim

The aim of this chapter is to find individual characteristics that have well-studied associations with both prostate cancer and PSA, use expert opinion to guide us toward a characteristic known to have strong associations with both prostate cancer and PSA, so we can then identify a single well-studied, modifiable, characteristic to study further. This chapter comprises a scoping study, a description of the practical issues that forced novel solutions (a probabilistic matching algorithm), a Sankey diagram, and expert opinion that was sought to complement the results of the diagram.

3.2. Background

3.2.1. Sankey diagrams

Sankey diagrams are a graphical representation of the flow of information (69). The flow of information is between defined points, and these points can be anything real or conceptual so long as information (such as energy, money, resources etc.) flows between them. The thickness of the lines corresponds to how much information flows between the points. Sankey diagrams are often used when representing the flow of raw materials in a population or region (70–72), including energy generation and distribution; the original Sankey diagram was of the thermal efficiency of steam-engines in 1898 (69).

A simple theoretical Sankey diagram I designed is shown in **Figure 3.1**, displaying how energy is transferred in a relatively inefficient light bulb. Of each 100 Joules (J) put into the light, 30 J is converted to light energy, and 70 J is converted to heat energy. The thickness of the lines shows clearly that more energy is lost through heat than is used to create light.



Figure 3.1 Sankey diagram showing how energy is transferred in a light bulb

The construction of Sankey diagrams requires knowledge of the relative amounts of information as it passes through a system. With a lightbulb, the proportion of energy is represented in the thickness of the lines, with actual units of energy written directly onto the lines. Where the energy goes is also displayed. In other studies, Sankey diagrams show the origins and end uses of raw materials (70,73). In medicine, Sankey diagrams could conceptually show the movement of people through treatments, going from initial diagnosis through first-line to palliative treatments, showing how many people are cured, die or continue treatment as time progresses. In 2017, Wang et al. showed the sequence of care transitions in the last 6 months of life for Medicare beneficiaries in the US in a Sankey diagram (74).

3.2.2. Sankey diagrams in epidemiology

Epidemiological and medical literature has not often used Sankey plots. I searched Embase and Medline for “Sankey” in all fields (sankey.mp) in July 2017, and 50 papers were returned. Most of these studies included Sankey diagrams¹, and were almost universally related to the flow of resources such as rare earth elements; the study by Wang et al. (74) was in the minority of studies related to epidemiology.

However, we considered that Sankey diagrams could be a useful tool in deciding which variables to consider for inclusion in this study. Our aim was to find the variables with the largest amount of information relating them to prostate cancer and PSA. As we couldn’t measure the information contained within each paper directly (without extracting data from tens of thousands of studies), we assumed that the variables with the greatest number of papers examining their associations with prostate cancer and PSA contained the most information.

3.2.3. Causal diagrams

An ideal characteristic to investigate further would have a strong, well-studied association with both prostate cancer and PSA. This can be expressed in a causal diagram (75), also called directed acyclic graphs (DAGs), which show the causal effects between variables. In a causal diagram, variables that are causally associated are joined by an arrow in the direction of causality, the arrows indicate a *direct* effect of one variable on another (76). Variables can also have *indirect* effects on others, where the effect is transmitted through one or more other variables rather than directly. The *total* effect of one variable on another is the summation of both the *direct* effect, and any *indirect* effects (75). A *confounder* is a variable that affects both the exposure and outcome in an association, while a *mediator* is any variable that transmits an *indirect* effect between other variables.

The characteristics we are searching for using the Sankey diagram would satisfy the causal diagram in **Figure 3.2**, where the characteristics have a *direct* effect on PSA, and an *indirect* effect on PSA through prostate cancer. In this diagram, prostate cancer is a *mediator* of the effect of the variables on PSA, rather than a *confounder* of the variable-PSA association (76). Controlling for prostate cancer when estimating the effect of the variables on PSA would estimate the *direct* effect, while not controlling for prostate cancer would estimate the *total* effect (75). In this diagram, all variables that affect prostate cancer and PSA are *confounders* of the association between prostate cancer and PSA. We did

¹ Two of the most interesting papers did not include a Sankey diagram, but were written by Dr Sankey: “The Way to a Man’s Heart Is through His Stomach: What about Horses?” (305) and “Do horses have a concept of person?” (306)

not consider that PSA could affect prostate cancer or any other variables: the biological function of PSA is quite narrow, to cleave proteins so sperm can be liberated, and as far as we know there have been no reports of damage caused by having PSA levels even 10,000 times the normal values. Additionally, we did not consider that prostate cancer could affect PSA through any other variables, i.e. it only has a *direct* effect on PSA.

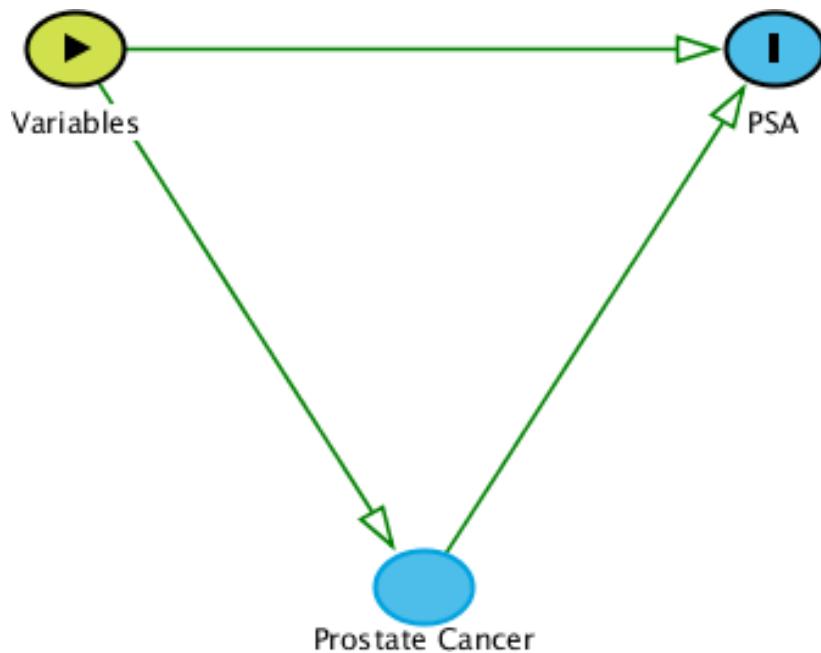


Figure 3.2 Causal diagram showing the associations between variables of interest (individual characteristics), prostate cancer and PSA. Each circle (node) represents a variable, and each arrow indicates the assumed direction of causality between the variables

3.3. Generation of the Sankey diagram

We designed the Sankey diagram to have PSA and prostate cancer nodes on the left, both connected to nodes on the right of the diagram indicating individual characteristics, or “variables”. The thickness of each line connecting the nodes was determined by the number of references that appeared to study the variable of interest and either prostate cancer or PSA. The initial steps in creating the diagram were:

1. Find all papers looking at prostate cancer and/or PSA
2. Remove duplicate papers
3. Compile a list of keywords associated with each paper
4. Search the titles, abstracts and keywords of all papers for each keyword in the compiled list
5. Search the titles, abstracts and keywords of all papers for prostate cancer and PSA to determine whether each paper looked at prostate cancer, PSA, or both
6. Count the number of papers associating each keyword with prostate cancer and PSA

In February 2015, Medline, Embase, BIOSIS and CINAHL were searched for all papers (called references from now on) that included the text words “PSA”, “prostate specific antigen”, “prostate cancer”, “prostatic neoplasm” or “prostate neoplasm” in combination with one of the following study types: “case control”, “trial”, “cohort”, “systematic review” or “meta-analysis”. Study types were specified to reduce the number of editorials, comments, case reports, cell-lines studies and other study types that would not be informative in a systematic review. Web of Science and Ovid were used to search the databases and download the references; in total, 44,068 references were found. All references were initially compiled in Endnote to allow a single text file to be exported, which contained the title, authors, year of publication, journal, page numbers, keywords (if available) and abstract for each reference.

Some references contained carriage returns in their abstracts, and we needed to remove these so each reference could be contained in a single row of data within a text file. This was accomplished by including an extra piece of information (the text string: “End of reference”) at the end of each reference, removing all carriage returns, and replacing the extra string with a carriage return. Once completed, the text file was imported into Stata (version 13) for further processing.

As the references were found using multiple databases, duplicate references existed. This was an issue, studies that were present in more than one database would also cause bias in the Sankey diagram. To save deduplicating 44,068 reference manually, I developed a probabilistic matching algorithm to find and remove duplicates, which is fully described in **Appendix 2**. Briefly, the algorithm matches chunks of text from the titles of references, and should there be enough of a match, the algorithm creates a weighted score for how identical the references are. The score is between one

and ten: a score of ten between two references is a perfect match, whereas a score of one means only the titles were similar (but not identical). A threshold of seven was applied for automatically removing duplicate references, allowing for typographical errors or other differences between references.

In total, 10,444 duplicates were identified amount the 44,068 references and removed using the probabilistic matching algorithm, leaving a total of 33,624 references for inclusion. **Table 3.1** shows the weighted score from the deduplication algorithm for duplicate references.

Table 3.1 Weighted scores of seven and above from deduplication of all 10,444 duplicate references

Weighted score	Total duplicates (%)
7	201 (1.9)
8	256 (2.5)
9	536 (5.1)
10	9,451 (90.5)
Total	10,444

To determine the list of variables that could be associated with prostate cancer and PSA, I took a hypothesis-free approach. As opposed to pre-specifying the variables of interest, all keywords associated with papers looking at prostate cancer or PSA were compiled. These keywords were assumed to represent all the variables that have been associated with either prostate cancer or PSA.

As keywords were only available from Medline, I searched Medline for “Prostate cancer” and “PSA” to compile a list of references. The keywords from each of these references were extracted, made lower case and spaces removed, then compiled into a single list. Duplicate keywords were removed, and some common keywords were grouped if their meanings were similar:

- Sex hormones: androgen, estrogen (included in oestrogen), testosterone, sex hormone, estradiol
- BMI: body mass index, BMI, obese, obesity, weight, anthropo*
- Diet: diet, eating, food, vegetable, fruit
- Insulin-like growth factor: insulin, growth, IGF
- Cholesterol: cholesterol, LDL, density lipid
- Smoking: smoking, smoke, tobacco

The titles and abstracts of all deduplicated 33,624 references were searched for each keyword, or group of keywords. The titles and abstracts were also searched for prostate cancer and PSA (and all synonyms) to determine if the reference examined prostate cancer, PSA, or both.

The keywords that were associated with the most references were manually inspected to find variables that represented individual characteristics, such as BMI and sex hormones. Keywords that were generic or inapplicable, such as “male”, “case-control study” or “clinical” were removed, as were treatments of prostate cancer, such as LHRH antagonists and radiotherapy.

The Sankey diagram was created for the most commonly studies individual characteristics, with the thickness of the lines proportional to the number of references associating the characteristics with prostate cancer and PSA. The thickest lines thus represent the characteristics associated with the most papers, and these characteristics were considered to have the largest quantity of information. The Sankey diagram was created using a precursor of TeMMPo, an online tool for identifying the quantity of literature suggesting specific mechanisms between and exposure and outcome (77). Some individual characteristics, such as ethnicity and family history of prostate cancer, could not be included in the Sankey diagram as the characteristics included too many potential relevant terms or were expected to not be reported in the titles or abstracts of a study. For example, keywords related to ethnicity could have included the actual ethnicity of the participants, such as Swedish or Korean.

3.4. Expert Opinion

One issue with the Sankey diagram is that it weights each characteristic on the number of studies that potentially measure the association between the characteristic and prostate cancer or PSA, and not the size of the association. The aim of the Sankey diagram is to find characteristics which are well-studied, but the characteristics ideally must also have strong associations with prostate cancer and PSA. A characteristic could be well-studied, but have a very small influence on PSA or prostate cancer. To avoid this issue, expert opinion was sought to determine which of the most-studied characteristics from the Sankey diagram were thought to have the largest effects on PSA and prostate cancer. Lead nurses working on the ProtecT study, a large randomised trial of different treatments for prostate cancer (31), were approached and asked to provide their opinions on the strengths of these associations.

Characteristics that ranked highly in the Sankey diagram were presented for expert opinion to discover which were thought of as having the largest effects of prostate cancer and PSA. The variables for selection were a combination of six high-ranking Sankey diagram variables (age, BMI, benign prostatic hypertrophy (BPH), diet, insulin-like growth factors (IGFs) and sex hormones) and six other variables we thought to be potential confounders of associations with prostate cancer and PSA (ethnicity, family history of prostate cancer, finasteride, kallikreins, tumour suppressors, and other cancers). Ethnicity was presented here as it was impossible to include it in the Sankey diagram (**Section 3.3**). The remaining other variables were presented either because there was a reasonable number of studies looking at their associations with prostate cancer and PSA, or because we felt these variables could have strong effects on both prostate cancer and PSA. Although we were ideally searching for a modifiable variable to study, we included non-modifiable variables, such as BPH and other cancers, to gauge expert opinion on their possible impact as confounders.

During a lead nurse meeting, I gave a short presentation on the aims of this thesis and the concept of mediation, but I did not mention prostate cancer or PSA in relation to any variables. I did not use prostate cancer, PSA or any of the variables on the question sheet in the examples of mediation. Rather, I used an example using food shopping, weight change and eating ice-cream. This was because I did not want to bias the responses toward any particular variables by mentioning them in context. The nurses had the chance to ask questions after the presentation. Each nurse was then asked to read through an information sheet and complete a table ranking the five variables they considered most associated with both prostate cancer and PSA. An “other” category was provided in case any variables the nurses thought were important were missed off the list. I asked the nurses to also complete a short

background section to provide context on their expertise in the field. The question and information sheets given to the ProtecT nurses are available in **Appendix 3** and **Appendix 4**.

For each completed form, the variable ranked highest received five points, the second four points etc. Any variable left blank received no points. The total number of points a variable received across responses was counted. In total, eleven nurses completed forms, allowing for a maximum of 55 points for each variable.

3.5. Results

3.5.1. Sankey diagram

A list of 12,868 unique keywords was generated from the 94,886 references looking at prostate cancer or PSA found using Medline. For the 33,624 references found using multiple databases (and limited by study type), 11,180 references mentioned PSA (33.3%) and 25,254 references mentioned prostate cancer (75.1%). The total number references that mentioned each of the 12,868 keywords in conjunction with PSA or prostate cancer are shown in **Table 3.2**.

Table 3.2 Total number of references that mentioned each of the 12,868 keywords in conjunction with prostate cancer or PSA

No. References	Keyword & PCa (%)	Keyword & PSA (%)
0	7,394 (57.5)	8,377 (65.1)
1-9	3,205 (24.9)	2,870 (22.3)
10-99	1,474 (11.5)	1,097 (8.5)
100-999	590 (4.6)	403 (3.1)
≥1000	205 (1.6)	121 (0.9)

PCa = prostate cancer

There were 524 keywords mentioned with PSA in at least 100 references, and 795 keywords mentioned with prostate cancer in at least 100 references. The number of references that mention high-scoring individual characteristics with prostate cancer and PSA are shown in **Table 3.3**.

The Sankey diagram is shown in **Figure 3.3**. As the number of studies examining any of the variables considered with prostate cancer was far larger than with PSA (25,254 versus 11,180), the number of the papers studying PSA were up-weighted by a factor of $25,254/11,180 = 2.26$. This made the total thickness of lines connecting PSA to variables comparable to those connecting prostate cancer to variables.

From the Sankey diagram (and **Table 3.3**), we assumed the variables most likely to have had their association with prostate cancer and PSA measured in a reasonably large number of studies were:

- Age
- Benign prostatic hyperplasia (BPH)
- BMI
- Diet
- IGF
- Sex hormones

Table 3.3 The number of references mentioning keywords referring to individual characteristics in conjunction with prostate cancer or PSA

Keyword	Total references: Keyword + PCa	Total references: Keyword + PSA
Aged	2246	1271
Aging	1737	932
Alkaline Phosphatase	135	101
Anxiety	191	94
Aspirin	109	52
Betacarotene	137	19
BMI*	1846	1069
BPH*	1040	866
Candida	637	385
Cholesterol*	401	211
Choline	117	77
Diabetes	531	217
Diarrhea	134	66
Diet*	1282	311
Estramustine	145	83
Fatty Acids	140	24
Finasteride	231	159
FSH	267	176
Glutathione	130	26
Gonadotropin	121	51
HIV	200	106
Hypertension	191	91
IGF*	1385	547
Interleukin	104	42
Kallikrein	73	65
LHRH	310	143
Lutein	333	150
Marital Status	101	41
Prostatitis	138	119
Sex Hormones*	4098	1972
Smoking*	1294	626

*Combined keywords

Grey rows indicate the characteristics assessed by expert opinion

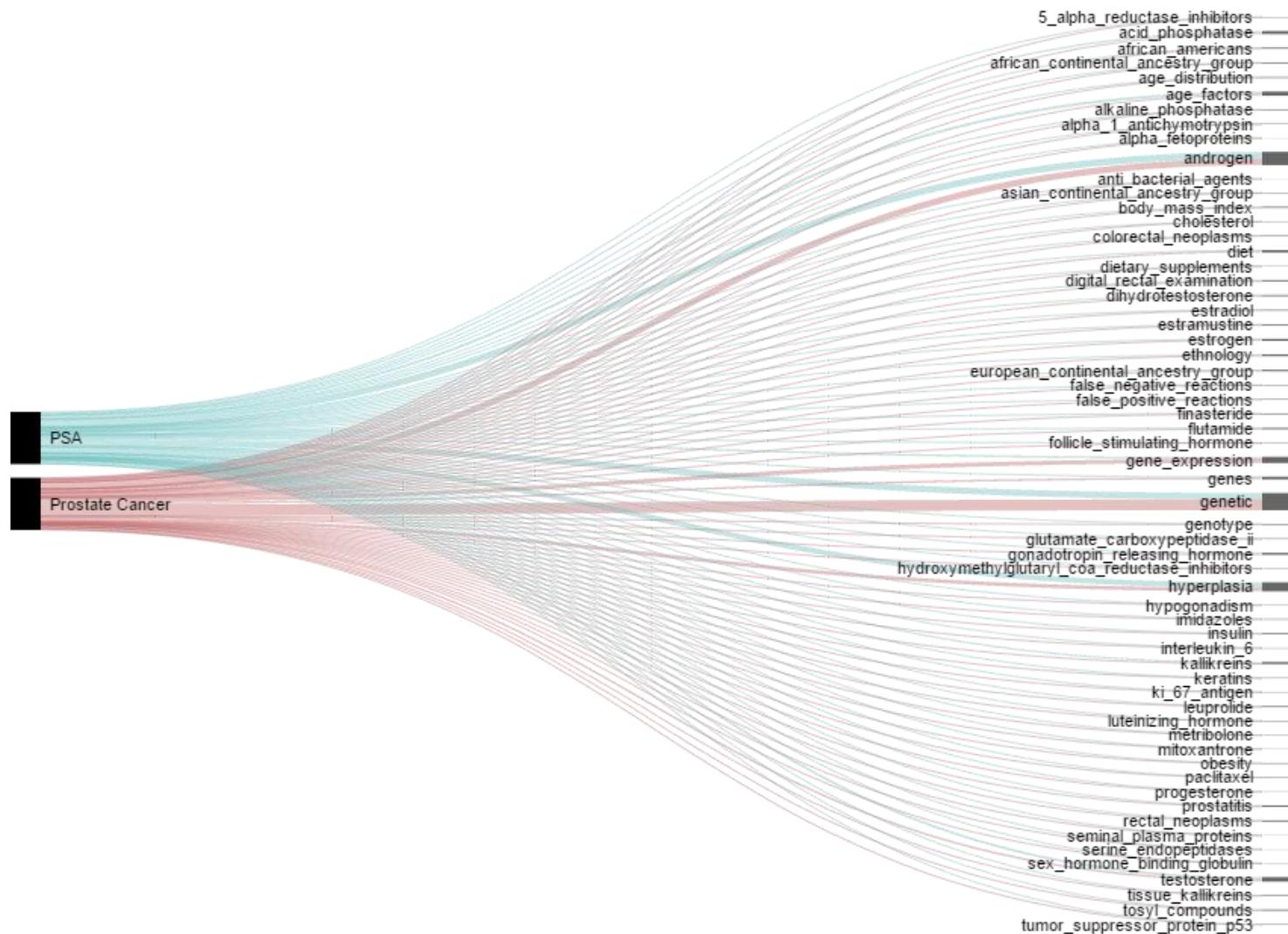


Figure 3.3 Sankey diagram showing the number of studies associating prostate cancer and PSA (left hand side) with selected individual characteristics (right hand side)

3.5.2. Expert opinion

In total, eleven lead nurses completed the question sheet, rating the top five variables they thought would have the largest associations with prostate cancer and PSA. All nurses were involved in urology or uro-oncology, with an average of over 14 years of prostate cancer expertise and 27 years of healthcare or research expertise. Six nurses had additional qualifications other than nursing qualifications (RGN/other). As researchers involved in both a large research trial on prostate cancer and nurses working with men with prostate cancer, ProtecT lead nurses represented excellent expert opinion on prostate cancer and PSA.

The results of the consultation with ProtecT lead nurses showed that age, sex-hormones (termed androgens on the question sheet), family history of prostate cancer and finasteride use were considered to have the largest associations with both prostate cancer and PSA (**Table 3.4**). There were no additional variables given in the “Other” category: the only entry was to provide more information when specifying “ethnicity”, detailing black African ethnicity.

Table 3.4 Results of Expert Opinion

Variable	Expert Score
Age	32
Androgens	29
Family history	23
Finasteride	20
BPH	13
Ethnicity	12
IGF	12
Diet	10
Tumour Suppressors	8
Kallikreins	6
BMI	3
Other cancers	1

3.6. Variable Selection

Following the results of both the Sankey diagram and the expert opinion, we considered the variables that had both many studies associating them with prostate cancer and PSA, and potentially large effects, **Table 3.5**.

Table 3.5 Variables that were well-studied, and/or thought to have strong associations with prostate cancer and PSA on expert opinion

Variable	Expert Score	Total references: Keyword + PSA	Total references: Keyword + PCa
Age	32	1,271 (aged)	2,246 (aged)
Sex hormones	29	1,972	4,098
BPH	13	866	1,040
Ethnicity	12	NA	NA
BMI	3	1,069	1,846

We aimed to carry out systematic reviews and meta-analyses of the associations between the factor chosen and both PSA and prostate cancer. Thus, the ranking of the variable in the Sankey diagram was our prime criterion in choosing a variable to study further. The expert opinion was to ensure that we did not choose a variable which was of no interest to experts, or which was widely thought to have little or no effect. Using these two criteria, we identified BMI as the variable to study further in this thesis.

Finasteride use, although highly ranked by expert opinion, was not studied alongside PSA and prostate cancer as commonly as other variables. IGF was moderately ranked by experts, but was likely overestimated by the keyword search used in the Sankey diagram. Family history was ranked much higher by the experts. However, we assumed that there is no mechanism by which a family history of prostate cancer should be able to affect PSA, excepting when a genetic predisposition towards higher PSA results in more diagnoses of prostate cancer, as PSA testing is used to detect prostate cancer. Given we decided to not consider genetics as variables for the systematic review, family history was thus also not considered further. Diet was ranked moderately by expert opinion, but the broadness of the topic does not allow for systematic review and meta-analysis of “diet” as a whole. Ethnicity was moderately ranked by expert opinion, and although “ethnicity” was not a keyword, many prostate cancer studies report ethnicity and is thus more prevalent than can be shown using keywords.

I created a causal diagram to show how the variables of interest related to each other, **Figure 3.4**, by using the following principles, which were derived using biological plausibility:

1. Age and ethnicity cannot be affected by any other variables
2. Sex hormones, BMI and BPH are all affected by age and ethnicity
3. BMI may affect both BPH and sex hormones
4. All variables could affect both prostate cancer and PSA
5. Prostate cancer affects PSA, but PSA does not affect prostate cancer
6. Unmeasured confounding may exist, but is not included in the graph

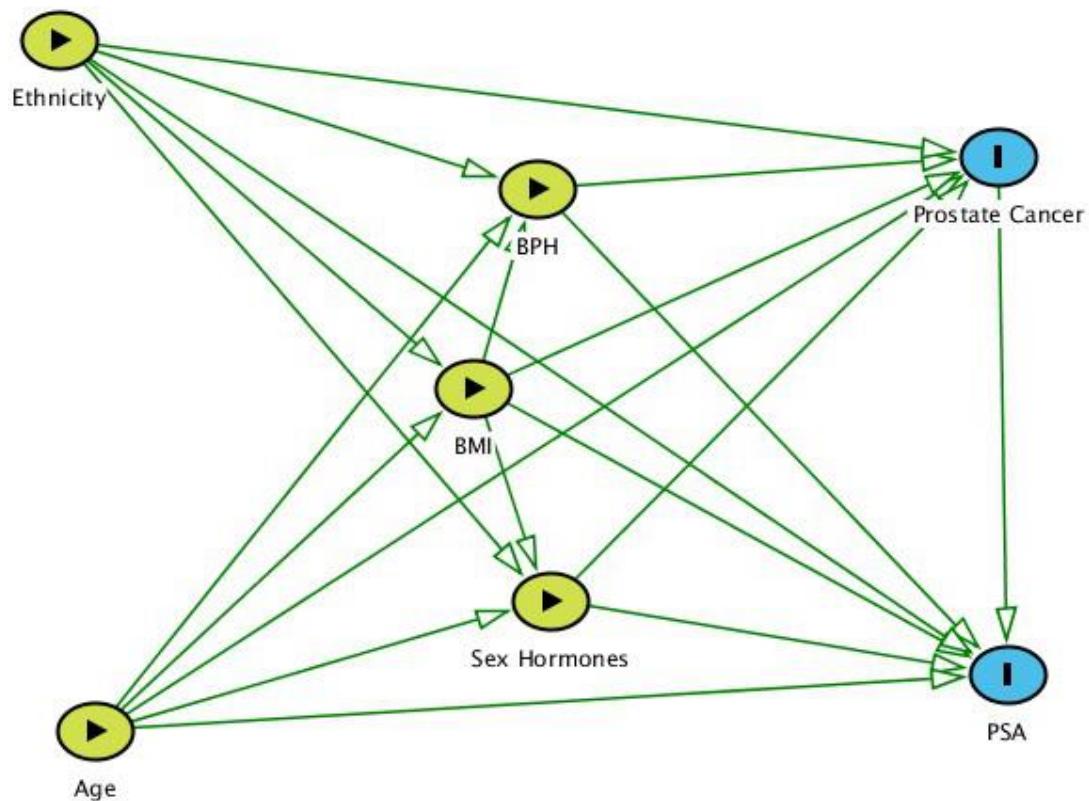


Figure 3.4 Causal diagram showing all associations between variables of interest, PSA and prostate cancer

BMI is the only variable of interest identified in our systematic approach that is easily modifiable through lifestyle changes. While BPH and sex hormones are possibly modifiable through lifestyle changes, drugs and other treatments are a more common alternative. Age and ethnicity are not modifiable, but represent confounders in the associations between the other characteristics, prostate cancer and PSA. As such, we chose BMI as the variable to carry forward to later chapters.

The associations between BMI, prostate cancer and PSA

Many studies have observed a negative association between BMI and PSA, and one possible explanation is haemodilution; as BMI increases, blood volume increases without a commensurate increase in the PSA protein, reducing the concentration of PSA (78–80). There is some disagreement in the literature as to whether BMI affects prostate cancer risk in total, or increases the risk of high-grade prostate cancer and lowers the risk of low-grade prostate cancer (81). It is possible that as BMI increases, an increase in IGFs and oestrogens and a decrease in levels of sex hormone binding globulin changes the risk of both low- and high-grade cancers (82).

However, testing for prostate cancer using BMI may also induce an association between BMI and prostate cancer risk. If haemodilution occurs with increasing BMI, then men with higher BMIs will on average be diagnosed with fewer prostate cancers when testing with PSA, simply because their PSA is, on average, lower. This would occur even if there were no true association between BMI and prostate cancer risk, and is similar to the association between age, prostate cancer and PSA (**Figure 1.2**), although an increase in BMI has been associated with a decrease in PSA rather than an increase. Additionally, if men with higher BMIs are biopsied later because their PSA is on average lower, they may be diagnosed with prostate cancer later than if they had a smaller BMI, leading to more instances of high-grade or advanced prostate cancer among men with high BMIs and fewer instances of low-grade or localised prostate cancer.

BMI may also affect the likelihood of finding prostate cancer on biopsy, again affecting the observed association between BMI and prostate cancer risk. As the number of samples taken at biopsy increases, the number of cancers diagnosed also increases (83), indicating that some cancers will be missed if a limited number of biopsy cores is taken. There is an association between the size of the prostate and BMI (84), as the volume of the prostate increases with BMI. This means there may be differential classification of prostate cancer in larger men, resulting in bias in the estimation of the association between BMI and prostate cancer.

BMI is a measure of adiposity, which can also be measured using weight, weight gain, waist or hip circumference and waist-to-hip ratio. However, BMI appears to be the most commonly used metric for adiposity when considering cancer; an umbrella review of the systematic reviews and meta-analyses associating adiposity and cancer at major anatomical sites, 131 out of 194 analyses (68%) from previous studies were between BMI and cancer, with the remaining 32% split between other adiposity measures (85). In this thesis, only BMI will be considered, although future studies could also consider other measures of adiposity.

3.7. Discussion

In this chapter, I used a systematic approach to select an individual characteristic (BMI) that was both well-studied and could have strong associations with both prostate cancer and PSA. A probabilistic matching algorithm to account for slight differences between duplicate references was also developed.

3.7.1. Strengths

The Sankey diagram, combined with expert opinion, proved a helpful tool when deciding which variable to consider further in this thesis. If either the Sankey diagram or expert opinion were used in isolation, there would be a risk of choosing variables that were well-studied but with weak (or no) association with the outcomes (Sankey), or choosing variables that have been examined in few studies or received a large amount of media interest (expert opinion).

The probabilistic matching algorithm I developed was useful in reducing the risk of bias in the Sankey diagram. While standard perfect-match deduplication software (such as in Endnote or Ovid) works well in general, the deduplication is limited and doesn't account for typographical errors in references, which we found to be quite common. There may also be other limitations in the standard software, such as having to manually remove duplicates (Endnote), or restrictions to less than 6,000 references (Ovid). The probabilistic matching algorithm I developed avoids these limitations, and allowed many duplicate studies to be found and removed prior to creation of the Sankey diagram.

3.7.2. Limitations

One of the largest limitations of our approach is that the Sankey diagram does not measure the amount of evidence directly, rather the number of studies that have measured both an exposure and outcome. Therefore, it is not known from the Sankey diagram whether the studies relating two variables all have few participants, all have very small associations, or whether they reported the association between the variables at all. Another concern is that when searching titles and abstracts for keywords, short keywords will appear inside other words: for example, "state" occurs in prostate, and would thus be over-represented. Similarly, "growth" could refer either to the growth of an individual or a prostate, but also within the variable "insulin-like growth factor" (IGF). Equally, insulin occurs as a variable by itself, and also within IGF. As the combined IGF group included two terms that could refer to different variables, the estimation of the number of references including IGF is likely an overestimate. Additionally, some variables cannot be reliably included in a Sankey diagram, for example "family history of prostate cancer" and "ethnicity", as these variables may be well-studied,

but not included in the list of keywords or abstracts of studies. The risk of these problems was mitigated by also asking for expert opinion, and including characteristics we thought may be well-studied, but which we could not measure directly in the Sankey diagram (e.g. ethnicity).

Interpretation of the Sankey diagram requires some degree of subjectivity, balancing the number of studies that included the keyword against the possibility of false positives, and there may be some subjectivity in the responses of the expert opinion. In addition, the Sankey diagram and expert opinion assessed two slightly different questions: the Sankey diagram assessed the number of studies mentioning two variables, whereas the expert opinion assessed the perceived strength of the association between variables. However, the overall degree of subjectivity is likely lower by using the approach we took than relying on either the Sankey diagram or expert opinion alone. While the order of scoping and expert opinion could be reversed, with the opinion informing which variables to search for, this may cause variables to be missed because the experts may not have all the available information. However, as the aim was not to find *all* the relevant variables, but to find those with enough information to be worthwhile studying and choose one of them, we felt the limitations of our approach would not overly affect the outcome of this section.

The probabilistic matching algorithm had a high sensitivity and specificity when assessed in a subset of studies (**Appendix 2**), though some non-duplicates were incorrectly tagged as duplicates (false positives). These studies tended to have very similar titles, and were published in the same journals at roughly the same time and as such were relatively rare. In general, the number of false positives should be very low using the algorithm, and the weighted score can be used to identify both perfect duplicate (accounting for punctuation, capitalisation etc.) and highly similar references.

3.8. Summary

After creation of a Sankey diagram and assessment of the expert opinion, we chose BMI as the variable to study further, as it was modifiable, easily-measured and has likely associations with both prostate cancer and PSA.

CHAPTER 4. SYSTEMATIC REVIEW AND AGGREGATE DATA META-ANALYSIS OF ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA

4.1. Aims

The aim of this chapter is to systematically review the literature which examines the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA.

Data were extracted from all relevant papers identified in **Section 3.3** and aggregate data (AD) meta-analyses performed. The results of this chapter and the results from the individual participant data (IPD) meta-analyses in **Chapter 5** are combined in **Chapter 6** to calculate the most precise estimates of the associations between BMI, prostate cancer, advanced prostate cancer and PSA.

4.2. Methods

4.2.1. Data Sources

The list of 33,624 de-duplicated papers compiled in **Section 3.3** examining prostate cancer and/or PSA was used as the primary data source for this chapter. In addition, three previous meta-analyses of the association between BMI and prostate cancer (86–88) were searched for any papers not identified in the database searches.

4.2.2. Eligibility criteria

From the complete list of papers, I restricted studies to those that investigated the association between BMI and either prostate cancer or PSA (or both). I did this by searching for “bmi”, “bodymassindex”, “anthropo*” and “weight” in de-spaced and lower-case version of the abstracts and titles of all studies. The asterisk in “anthropo*” refers to any words starting with “anthropo” followed by any number of other letters, for instance “anthropometric” or “anthropometry”.

We included original articles that were published in peer reviewed journals (including any supplements and meeting abstracts). Reviews, books, commentaries and letters were excluded. We included studies that measured an association between BMI and prostate cancer, BMI and advanced prostate cancer, or BMI and PSA. If the abstract did not specifically mention BMI but mentioned height or weight, we acquired the full text to determine if BMI was calculable from data included in the publication.

We excluded studies examining pre-malignant disease if there was no mention of prostate cancer or PSA (e.g. grade prostatic intraepithelial neoplasia). Prostate cancer studies were included whether the control subjects had a history of benign prostatic hypertrophy (BPH) or not. Human RCTs, case-control, cohort, non-randomised experimental studies and systematic reviews were included. All animal and cell-line studies were excluded. Studies where BMI was measured after diagnosis of prostate cancer were excluded, as this increases the likelihood of reverse causality, i.e. that cancer affected BMI, rather than the reverse. Studies that we considered a critical risk of bias (see **Section 4.2.4**) were also excluded prior to analysis, these consisted of studies that screened for prostate cancer using PSA, and studies that did not account for age in either the study design or analysis.

4.2.3. Data extraction

I screened the abstracts of all papers for inclusion and retrieved full texts for all studies that met the inclusion criteria, or looked like they might with further information. For instance, full texts were

sought if no abstract was available or if the abstract did not include sufficient information to decide on inclusion. If after every effort was made to locate a copy of the full text of a study (including requesting the full text from the British library), no copy could be found, the study was excluded.

I screened all full texts for inclusion, and one of three independent reviewers (KT, ET, HJ) reviewed the first 60 full texts as well to check for consistency. We resolved any inconsistency with discussion to clarify screening criteria. A random subset of the remaining studies (30 full texts) were also reviewed by the independent reviewers to check for drift from inclusion/exclusion criteria.

For studies meeting the inclusion criteria after full text review, an independent extractor (RL) and I extracted all relevant data, with disagreements resolved by discussion. The first ten extractions were performed by all 5 reviewers (SH, RL, KT, ET, HJ) to check for consistency.

We extracted population and study data for all studies, including: study name, mid-year of recruitment, study design, population (location, ethnicity, whether selection into study was based on PSA, sample size), BMI ascertainment type (measured or asked), outcome, statistical measure (including details of any adjustments for potential confounders) and the time between the measurement of BMI and outcome. We categorised studies as “before” if BMI was measured on average at least two years before diagnosis (prospective studies), and “same time” if BMI was measured on average less than two years before diagnosis. In general, “before” studies were cohort studies and “same time” studies were case-control studies. By restricting the “before” criterion to two years or more between BMI measurement and diagnosis, we reduced the risk of reverse confounding in this subgroup of studies, where prostate cancer affected BMI rather than BMI affecting prostate cancer risk.

We extracted data that were (or could be transformed to) an odds ratio (OR) for the association between BMI and prostate cancer risk and advanced prostate cancer risk, and a beta-coefficient for the association between BMI and log-PSA. We also extracted SEs of the effect estimates. For example, the mean and standard deviation (SD) of BMI for men with and without diagnosed prostate cancer could be used to estimate the standardised mean difference (SMD) in BMI from prostate cancer, and ORs could be estimated from SMDs. If both adjusted and unadjusted results were given, we recorded both results, but the most-adjusted model was used in the meta-analysis.

If the data were not sufficient to calculate an effect estimate and SE for the association, or when the data did not permit transformation to a beta coefficient or an OR, we extracted a P value, number of participants and effect direction for use in an albatross plot. If multiple analyses were presented, we extracted the results from the analysis which best answered the question “does having a larger BMI

change a man's risk of prostate cancer or PSA level". For instance, we extracted P values from fully-adjusted analyses in preference to unadjusted analyses, and P values from analyses where BMI was treated continuously in preference to analyses where BMI was categorised. If only analyses with BMI as a categorical variable were presented, we extracted the P value from the comparison of the obese ($BMI \geq 30 \text{ kg/m}^2$) and normal weight ($BMI < 25 \text{ kg/m}^2$) categories, but only if a P for trend was not presented.

When several papers reported on the same study, we extracted data from all papers but the paper with the smallest SEs were used in the meta-analysis. These tended to be papers with the most information (e.g. were analysed when more participant data had been collected) or different analytic method (e.g. continuous effect estimate versus levels of BMI).

In addition to effect estimates for the association between BMI and prostate cancer, we extracted effect estimates for the association between BMI and advanced prostate cancer. We determined the effect estimate to be for advanced prostate cancer if the individual studies labelled the effect as such, or if the effect was for locally advanced, extra-prostatic, nodular or metastatic prostate cancer. High-grade prostate cancer was not considered equivalent to advanced prostate cancer.

4.2.4. Risk of bias assessment

As both case-control and cohort studies were included in the search, we used an assessment tool I created for a previous meta-analysis (89) to assess risk of bias in each study. Both RL and I assessed risk of bias independently using this tool, with disagreements resolved by discussion. This tool uses the categories of assessment from a draft of the ROBINS-I tool (63), and questions from the CASP case-control and cohort questionnaires (64,65) to aid in assessing risk of bias. ROBINS-I could not be used directly as it is explicitly used for studies of interventions (Risk Of Bias In Non-randomised Studies - of Interventions).

We assessed risk of bias in six categories: confounding, selection of participants, missing data, outcome measurement, exposure measurement and results' reporting. Each category contained questions designed to help assess the risk of bias, and these varied depending on the study type (prostate cancer or PSA) and design (e.g. cohort, RCT, case-control), see **Appendix 5** for a list of questions. We assigned an overall and category-specific risk of biases; either low, moderate, high, critical or unclear.

We determined whether the category-specific risk of bias was low, moderate or high risk of bias, which was determined subjectively by both data extractors based on the questions in the risk assessment

tool. We gave a category-specific risk of unclear if there was insufficient information to assign a risk. We based the overall risk of bias on a subjective combination of the category-specific risk of biases, looking at the maximum risk of bias that could have been introduced into the study by each category. In most cases, the risk of bias from confounding was considered with the highest weight when deciding the study's overall risk of bias, as we thought the potential for bias was greatest for studies that did not adjust for confounding factors, such as ethnicity.

In addition, we determined that a study had a critical risk of bias if age was not accounted for in either the design or analysis of the study and, for BMI-prostate cancer case control studies, if there was more than a 5-year difference in the mean or median ages of cases and controls. We considered such studies to be at critical risk of bias because age is strongly associated with BMI (90), PSA (9) and prostate cancer risk (9).

We also determined that studies had a critical risk of bias if the design of the study was such that entry into the study was conditional upon PSA levels, both for the association between BMI and PSA (conditioning on the outcome) and the association between BMI and prostate cancer (conditioning on a collider). For the association between BMI and prostate cancer, we excluded studies, for example, if the association between BMI and prostate cancer was presented only for men with a PSA above a threshold, or studies where the population was drawn from men who were biopsied because of having high PSA. This is because both BMI and prostate cancer are thought to affect PSA, so conditioning on PSA will introduce collider bias (91), where two variables are associated simply because something they both affect has been conditioned upon.

Studies with a critical risk of bias were excluded prior to analysis and were not considered further.

4.2.5. MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies

The MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies (92) were used to ensure reporting of this systematic review was complete. A checklist detailing the guidelines is available in **Appendix 6**.

4.3. Methods to Derive Standard Effect Estimates

In this section, I describe the statistical methods I used to derive standardised effect estimates for three meta-analyses, between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA. The association between prostate cancer and PSA was not examined with AD, as we felt this was better examined using IPD in **Chapter 5**.

Initially, I describe the methods used to transform the results of the BMI-prostate cancer studies to the standardised effect estimate, the log-OR for prostate cancer for a 5 kg/m^2 increase in BMI. I then describe the methods used to transform the results of BMI-PSA studies to the same scale, the change in log-PSA for a 5 kg/m^2 increase in BMI (a beta coefficient). Continuous linear estimates of the associations between BMI, prostate cancer, advanced prostate cancer and log-PSA were estimated because this was the only way to combine data from studies that variously measured linear effect estimates with studies that reported categorical effect estimates and those that reported summarised statistics for BMI in men with and without prostate cancer. All meta-analyses required the standardised effect estimate (log-OR or beta coefficient) as well as the SE of the estimate for each study. **Table 4.1** shows all the methods used to transform results from studies to standardised effect estimates, and the number of studies to which each method was applied.

We used log-PSA as the outcome for the BMI-PSA analysis as we assumed the association between BMI and PSA was multiplicative, where an increase in BMI causes a percentage change in PSA (as opposed to an absolute change). We considered a multiplicative model a better fit to the theory of haemodilution (79,80,93), which suggests that PSA decreases because it is diluted when an increase in BMI causes a greater blood volume. In effect, the amount of PSA as a substance doesn't change but becomes more dilute or concentrated as BMI changes, hence the difference in PSA attributable to a change in BMI is multiplicative. This implies that a man with a PSA of 10 ng/ml is going to have a larger change in PSA due to a change in BMI than a man with a PSA of 0.1 ng/ml.

In most BMI-PSA studies, the reported outcome was PSA, and so a transformation of PSA results to log-PSA results was required; these methods are described in the third part of this section. In the final part of this section, I describe the methods of meta-analysis used in this chapter.

Table 4.1 The different methods of calculating standardised effect estimates and the number of studies on which the methods were applied

Section	Method	Number of Relevant studies
4.3.1 Statistical Analysis of BMI-Prostate Cancer Association		
4.3.1.A	Weighted regression for correlated outcomes	35
4.3.1.B	Mean BMI not presented for each level of BMI	33
4.3.1.C	Mean differences presented	20
4.3.1.D	Matched studies using levels of BMI, and ORs not reported for levels	8
4.3.1.E	Other information was presented	29
4.3.2 Statistical Analysis of BMI-PSA Association		
4.3.2.A	Weighted regression	8
4.3.2.B	Percentage change in PSA for a 5 kg/m ² increase in BMI	10
4.3.3 Transformations of PSA		
4.3.3.A	Mean and SD of PSA presented	6
4.3.3.B	Median PSA presented	1
4.3.3.C	Geometric mean PSA presented	1
4.3.3.D	Odds ratios presented	1
4.3.3.E	Ratio of geometric means presented	2
4.3.3.F	Other information presented	3

4.3.1. Statistical Analysis of BMI-Prostate Cancer Association

4.3.1.A. Weighted regression for correlated outcomes

Many studies ($n=35$) (37,94–127) that examined the association between BMI and prostate cancer risk reported the ORs for prostate cancer (often adjusted for potential confounders) at different levels of BMI, with reference to a baseline level. To combine the ORs from different levels into a single linear effect estimate of the OR for prostate cancer for a 5 kg/m^2 increase in BMI, I used generalised least squares for trend estimation (GLST) as proposed by Greenland and Longnecker (128) (the GLST command in Stata (129)). This method estimates the variance-covariance matrix of the log-ORs, and allows adjusted ORs to be manipulated (with some assumptions) to account for the correlation between ORs. The correlation appears because the same reference group is used in all ORs. The GLST method also requires that the dose (BMI) in the reference group is 0, and therefore all values of BMI were reduced by the BMI in the reference group.

The **GLST** command estimates beta coefficients using the following methodology for both case-control and cohort data in Stata (128,129). First, the expected number of men with and without prostate cancer are estimated for each category of BMI, given the adjusted OR and total number of participants for each category (or total follow-up time for cohort studies), using a fitting algorithm (to solve a non-closed equation). This is necessary because unadjusted and adjusted ORs may be different due to differences in the confounders between men with and without prostate cancer, and so the effective number of men in each group may differ. The variance-covariance matrix (described above) is estimated from the fitted values. These estimates are used to estimate the OR per unit change in BMI. I re-expressed this OR as for a 5 kg/m^2 increase in BMI, calculated by multiplying the log-OR and SE per unit increase in BMI by five, then exponentiating.

For each study, the GLST command requires the number of men with prostate cancer and total number of men (or total follow-up time), as well as the mean BMI, in each category of BMI, in addition to the OR and SE for each non-reference category of BMI. The number of participants and OR are both easily found, and the SE of log-OR can be calculated from the 95% confidence interval (CI) for the OR if not given directly:

$$\text{SE} = \frac{\ln(\text{CI}_u) - \ln(\text{CI}_l)}{3.92} \quad (10)$$

where CI_u and CI_l are the upper and lower CIs, respectively.

This methodology has some assumptions that are usually viable under most circumstances:

1. the crude OR parameters are approximately equal to the adjusted OR parameters, due to the collapsibility of ORs (130)
2. the correlation matrices are approximately equal between crude and adjusted ORs
3. the variances of the crude ORs can be estimated using the standard methods

The third assumption is almost always viable, except in matched studies; however, there is evidence that matching does not influence the variances excessively (131). The first two assumptions will be viable when the adjustment variables are only weakly related to the exposure and outcome. While age is correlated both with BMI and prostate cancer risk, it is unlikely to be a strong enough association to invalidate the assumptions. Additionally, as Greenland and Longnecker (128) note:

"some set of externally specified constraints is necessary in order to allow estimation to proceed when the covariate-specific data are unreported, and assumptions 1-3 are far more reasonable than assuming that the L's [ORs] are uncorrelated."

4.3.1.B. Mean BMI not presented for each level of BMI

The GLST calculation (128) requires the mean BMI in each category of BMI. Almost all studies (33 of 35, 94%) presented the range of BMI values (e.g. 25–29.9 kg/m²), but not the mean or median value. However, the mean BMI can be estimated using the method presented by Chêne and Thompson (132). This method assumes a normal distribution for BMI; although BMI is generally positively skewed, we considered the degree of skewness insufficient to materially bias the results. The approach additionally requires the number of participants and upper limits of each category of BMI.

4.3.1.C. Mean differences presented

Some studies (n=20) (133–152) presented the mean BMI for all cases and controls separately. For these studies, the standardised mean difference (SMD) was calculated using standard methods (49). SMDs are equivalent to log-ORs per SD increase in exposure (153). Therefore, these were divided by the pooled SD of BMI and multiplied by five to give the log-OR for a 5 kg/m² increase in BMI.

4.3.1.D. Matched studies using levels of BMI, and ORs not reported for levels

Some studies matched on key variables (e.g. age) to reduce confounding, but did not adjust for these variables in the analysis and presented the number of cases and controls by level of BMI. Additionally, some studies reported the number of men with and without prostate cancer in different categories of BMI without reporting the ORs for prostate cancer versus a reference category. In both types of study (n=8) (154–161), the unadjusted OR for prostate cancer could be estimated from the number of men with and without prostate cancer, but this may have been biased from confounding if the age of men in the different BMI categories was very different. Therefore, rather than using the potentially biased unadjusted ORs for prostate cancer, I estimated the overall mean and SD of BMI for both cases and

controls in these studies using the method of Chêne and Thompson (132), then estimated the SMD in BMI for prostate cancer as in **Section 4.3.1.C** above.

4.3.1.E. Other information presented

Some studies (n=17) (162–178) presented the OR per unit increase in BMI; these studies required no transformation or calculation other than multiplying the log-OR by five.

Studies that did not present enough data to estimate an OR with SE (n=12) (179–190) had a P value extracted.

4.3.2. Statistical Analysis of BMI-PSA Association

4.3.2.A. Weighted regression

Most studies ($n=9$) (35,80,82,191–196) looking at the association between BMI and PSA gave the mean PSA values for levels of BMI. Firstly, all PSA results were transformed to log-PSA results (see **Section 4.3.3**). Once transformed, the log-PSA values in each level of BMI needed to be combined to give a continuous effect estimate. I thus estimated the change in log-PSA for a 5 kg/m^2 increase in BMI, which was exponentiated (see below) to give the percentage change in PSA for a 5 kg/m^2 increase in BMI.

Variance weighted least squares (VWLS) linear regression is a simplified version of GLST, where there is no correlation between effect measures in the different categories that needs to be accounted for. There is no correlation because the mean PSA is presented for each level of BMI independently, rather than each non-reference level being compared to a reference level. VWLS regression requires an estimation of the variance of the mean log-PSA in each category of BMI, so for studies presenting the number of participants and the SD of log-PSA for each level of BMI:

$$\nu_i = \frac{\sigma_i^2}{n_i} \quad (11)$$

where for each category i of BMI, ν_i is the variance of the mean PSA, σ_i^2 is the variance of PSA and n_i is the number of participants.

4.3.2.B. Percentage change in PSA for a 5 kg/m^2 increase in BMI

A consequence of using log-PSA as opposed to observed PSA is that the beta coefficient of the linear regression represents the change in log-PSA per unit change in BMI. As this is not an intuitive scale, the result was exponentiated to give the ratio change in PSA per unit change in BMI.

The coefficient from VWLS regression (and its SE) can be converted to a percentage change in PSA for a 5 kg/m^2 increase in BMI:

$$\% \text{ change in PSA} = 100(e^{5\beta} - 1) \quad (12)$$

This conversion was applied to final effect estimates from the meta-analysis of BMI and PSA to give the results in an interpretable way.

4.3.3. Transformations of PSA

A variety of statistical methods were required to transform summary PSA results to log-PSA values.

We assume a log-normal distribution for PSA (197), with mean μ_n and standard deviation σ_n . A log-normal distribution is an exponentiated normal distribution, so becomes normally distributed when the distribution is logged.

Normal distributions have a defined relationship with log-normal distributions; the mean and SD from one distribution can be used to calculate the mean and SD of the other distribution. **Figure 4.1** shows a graph of a normal distribution with a mean of 0 and an SD of 1, and the corresponding log-normal distribution created by exponentiating the normal distribution.

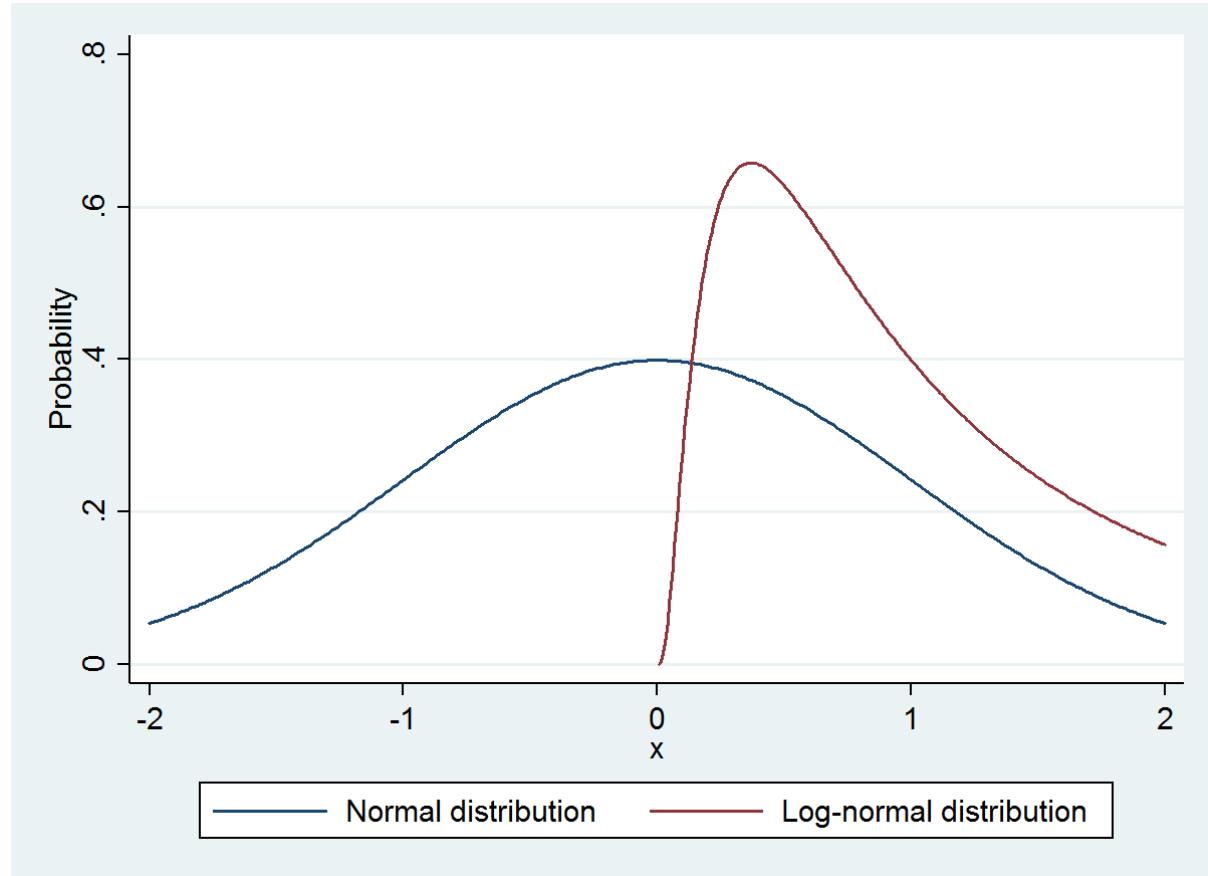


Figure 4.1 Graph of a normal distribution (mean = 0, SD = 1) and its corresponding log-normal distribution (created by exponentiating the normal distribution)

Most BMI-PSA studies looked at the association between BMI and PSA, not log-PSA, requiring the conversion of the mean and SD of PSA to the mean and SD of log-PSA. A similar transformation was also required when the median PSA and SD were presented, and when the geometric mean and SD were presented. Most studies required some transformation for the data to be combinable, but there

were also some studies without sufficient information for inclusion in a meta-analysis; these studies had P values extracted for use in an albatross plot instead.

4.3.3.A. Mean and SD of PSA presented

The mean and SD of a normal distribution in terms of the mean and SD of a log-normal distribution are:

$$\mu_n = \ln\left(\frac{\mu_o^2}{\sqrt{\sigma_o^2 + \mu_o^2}}\right) \quad (13)$$

$$\sigma_n = \sqrt{\ln\left(\frac{\sigma_o^2}{\mu_o^2} + 1\right)} \quad (14)$$

where values with the subscript o are the log-normally distributed observed mean and SD of PSA, while values with the subscript n are the normally distributed mean and SD of log-PSA. The equations presented here are presented in Higgins et al. (2008) (198). To convert data to a common scale (log-PSA), studies presenting the mean and SD of the observed PSA ($n=6$) (79,80,82,191,192,199) were transformed using these equations prior to further analysis.

4.3.3.B. Median PSA presented

One paper (196) reported the median PSA with an interquartile range. Assuming a log-normal distribution, the median PSA is equal to the exponentiated mean of log-PSA. The inter-quartile range (IQR) was used to estimate the SD of observed PSA, as the first and third quartiles for PSA are the exponentiated quartiles for log-PSA:

$$\sigma_o = \frac{\ln(IQR_u) - \ln(IQR_l)}{1.35} \quad (15)$$

where IQR_u and IQR_l are the upper and lower bounds of the interquartile range respectively (200). These equations were used on the single paper reporting the median PSA with the IQR.

4.3.3.C. Geometric mean PSA presented

One paper (194) reported the geometric mean and (arithmetic) SD of PSA for each level of BMI, rather than the arithmetic mean and SD. The geometric mean of a distribution is by definition equal to the exponentiated mean of the logged distribution (198), so the required mean log-PSA is simply calculated as the logarithm of the reported geometric mean.

This paper gave a 95% CI around the geometric mean (which was estimated from linear regression with age as a covariate to give an age-adjusted mean), which was converted to the SD of log-PSA using the following equation from the Cochrane handbook (200):

$$\sigma_o = \sqrt{N} \times \frac{\ln(\text{CI}_u) - \ln(\text{CI}_l)}{3.92} \quad (16)$$

where CI_u is the upper 95% confidence bound and CI_l is the lower 95% confidence bound. This standard method was used with the single paper reporting the geometric mean and (arithmetic) SD of PSA for each level of BMI.

4.3.3.D. Odds ratios presented

One study (201) presented the OR for each level of BMI having a PSA above a certain amount compared to a reference BMI level. This study also presented the PSA for different levels of BMI, but did not adjust for age, confounding the PSA values with the age-BMI association and giving those results a critical risk of bias.

Given the ORs only give the ratio of odds of having a PSA higher than a threshold for each level of BMI compared to the baseline level, there was insufficient information to estimate the mean PSA value (with an SD) for each level of BMI. As such there is no conceivable way of converting this study's results to a change in log-PSA given a 5 kg/m² increase in BMI.

Therefore, for this study, the P value for the largest category of BMI was taken for inclusion in the albatross plot, as the study does tell us something about the association between BMI and PSA, albeit with insufficient data for inclusion in the meta-analysis.

4.3.3.E. Ratio of geometric means presented

Two studies (35,195) presented the ratio (or percentage change) of geometric means of PSA with 95% CIs for two levels of BMI with respect to a baseline level of BMI (overweight and obese versus normal-weight). These effects are equivalent to the exponentiated difference in log-PSA between the men who were normal weight and those who were overweight or obese. The ratios of geometric means were estimated in both papers using linear regression. Although the geometric means of log-PSA for each level of BMI were also presented, the SDs of log-PSA (which are required for use of the method presented in **Section 4.3.3.C**) were not. I devised an approach to estimate these SDs using the ratios and their SEs, which made two assumptions:

1. Log-PSA was normally distributed within each level of BMI
2. The variance of log-PSA in the second level of BMI was equal to the mean of the variances of log-PSA for the first and third BMI levels

These assumptions were necessary due to mean ratios providing effect estimates and SEs for all but one level of BMI: the baseline BMI level does not have an effect estimate or SE since it is included in all other estimates. The full derivation of my approach to estimating the log-PSA SDs is available in

Appendix 7. Once the SDs were estimated, VWLS could be used to compute the change in log-PSA for a 5 kg/m² increase in BMI.

4.3.3.F. Other information presented

Two studies (193,202) presented the log-PSA change per unit increase in BMI; no transformations were required for these data. One study (203) presented information on the effect of BMI changes on the change in PSA per year. I was unable to estimate the beta coefficient and SE from this information and could therefore not include this study in the meta-analysis, so the P value was taken and used in the albatross plot.

4.3.4. Combining data

Meta-analysis

Random-effects and fixed-effect meta-analyses (see **Section 2.3.2**) were both used to determine the overall/average estimate of the association between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA. The results were given as the OR for prostate cancer or advanced prostate cancer, or percentage change in PSA for a 5 kg/m^2 increase in BMI. Forest plots were produced for all analyses. Forest plots for the association between BMI and PSA were presented as the percentage change in PSA, where, for example, an effect of 0.95 represents a 5% decrease in PSA. The I^2 statistic (57) was used as a measure of inconsistency between studies, and was reported with the P value for Cochran's Q statistic, a test for heterogeneity (204). If individual papers presented more than one type of data, results where BMI was treated as a linear continuous variable were prioritised, and the data type with the smallest SE was taken if there were no continuous results.

Because both case-control and cohort studies were allowed by the inclusion criteria for studies measuring the association between BMI and prostate cancer (and advanced prostate cancer), both ORs and hazard ratios (HRs) were collected. Rate ratios were assumed to be roughly equivalent to HRs (205). ORs and HRs are not equivalent, so to assess the validity of the assumption that ORs and HRs could be combined, I performed subgroup analyses for prostate cancer and advanced prostate cancer to determine if case-control studies (which tended to report ORs) and cohort studies (which tended to report HRs) gave different effect estimates.

Meta-regression

Meta-regression (58) (**Section 2.3.2**) was used to determine if the effect estimates of individual studies included in the meta-analysis varied due to study-level factors. For the associations between BMI and prostate cancer and advanced prostate cancer, we considered ethnicity (non-Caucasian versus Caucasian); mid-year of recruitment; the mean BMI in the study; and the risk of bias of the study (high versus medium). For the association between BMI and PSA, we considered ethnicity (split into Caucasian and non-Caucasian); mid-year of recruitment; the mean age at diagnosis; time between BMI measurement and diagnosis; and the risk of bias of the study (medium or high). The risk of bias was not low in any study, as all studies were observational and thus potentially subject to unmeasured confounding.

Funnel plots

A funnel plot (61) was drawn to assess for small study effects (asymmetry of the funnel plot) (60) (**Section 2.3.2**).

Albatross plots

An albatross plot (48) (see **Chapter 2.3.3**) was created using the P values from all studies in the meta-analysis, along with all studies that did not have a critical risk of bias, but which could not be included in the meta-analysis. The studies without sufficient data for meta-analysis were marked differently from those in the meta-analysis, and the plot was examined to determine if inclusion of the remaining studies would have altered interpretation of the evidence. The albatross plot was adjusted for unequal numbers of men with and without prostate cancer, where the effective sample size for each study was estimated for a given case-control ratio of one.

4.4. Results

In total, 3,308 papers were found that had keywords for BMI and prostate cancer or PSA and had an eligible study design. After title and abstract screening, 519 papers remained. During full-text screening: 58 studies were removed for not having enough information to calculate either an effect estimate of BMI or a P value; 217 papers were removed for not reporting an association between BMI and prostate cancer, advanced prostate cancer or PSA; 11 studies were removed for being exact duplicate studies; 10 abstracts or posters were removed as a full text was available; 33 review studies were removed; and one study could not be found, even after contacting the British Library (206). Five additional relevant studies were found in previous meta-analyses, which left 194 papers. Data were extracted from all remaining papers and assessed for risk of bias. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram is presented in **Figure 4.2**.

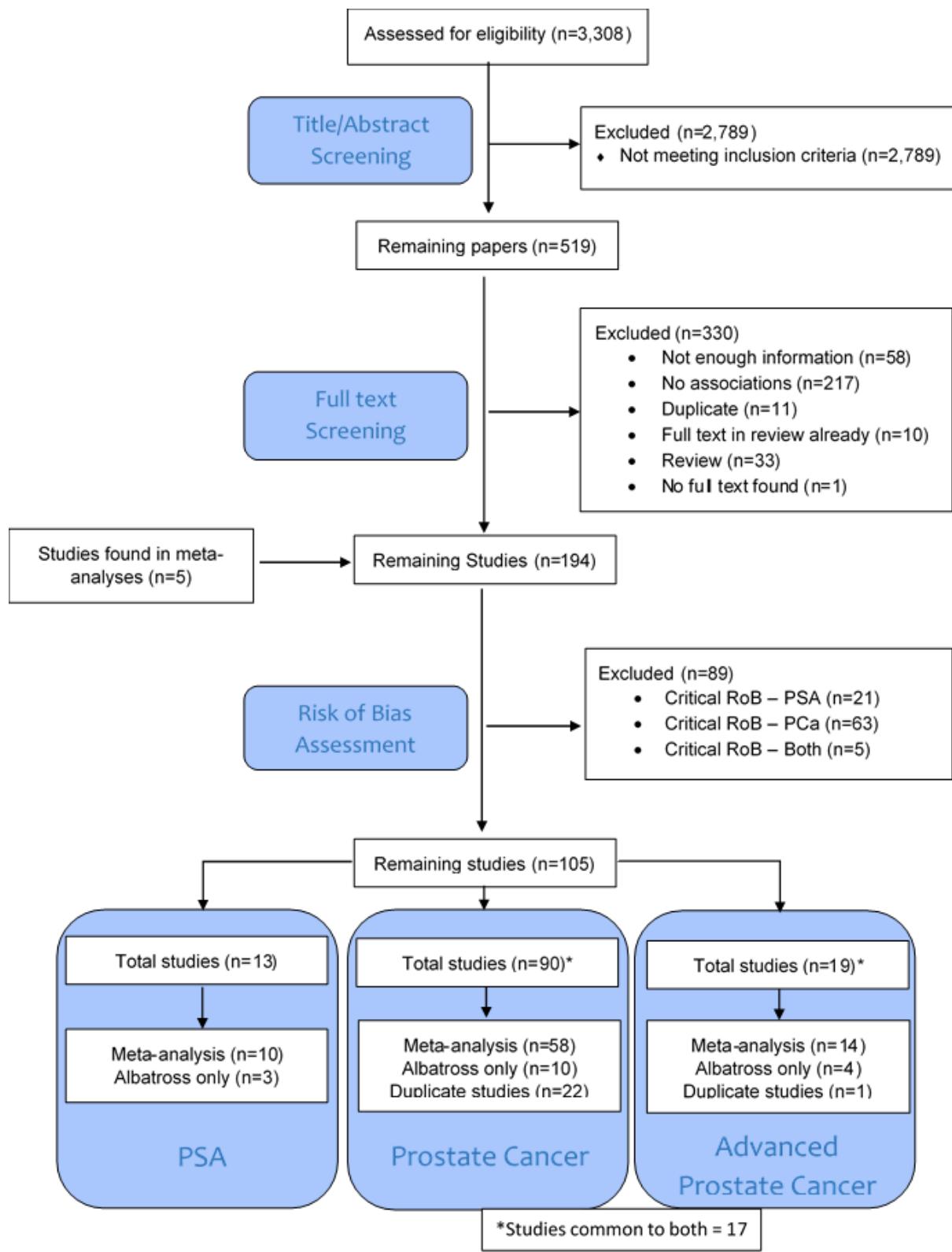


Figure 4.2 PRISMA flow diagram showing the number of studies in each stage of the systematic review

4.4.1. BMI and prostate cancer

Data were extracted from 160 studies examining the association between BMI and prostate cancer or advanced prostate cancer. After assessing risk of bias, 68 studies were excluded because of a critical risk of bias, leaving 92 studies.

Reasons for critical risk of bias were: bias due to confounding in 33 studies (49%); bias due to selection of participants in 31 studies (46%); bias due to measurement in exposure in two studies (3%); and bias due to selective or insufficient reporting in two studies (3%).

Of the 92 studies without a critical risk of bias, 90 examined the association between BMI and prostate cancer, and 19 examined the association between BMI and advanced prostate cancer (17 studies examined both associations). In total, 22 papers used information from the same studies and were considered duplicates in the BMI and prostate cancer analysis, and 1 study was considered a duplicate in the BMI and advanced prostate cancer analysis, leaving 68 and 18 studies respectively for analysis.

In total, 58 studies were included in the meta-analysis of BMI and any prostate cancer; 10 studies (15%) were not included due to insufficient data, but were included in the albatross plot. All studies examining the association between BMI and prostate cancer are detailed in **Table 4.2**, with the results of the risk of bias assessment in **Table 4.3**.

The random-effects meta-analysis (**Figure 4.3**) estimated the average OR for prostate cancer for a 5 kg/m² increase in BMI to be 1.00 (95% CI 0.97 to 1.02, P = 0.71). There was a strong evidence for inconsistency in effect estimates across studies ($I^2 = 67.3\%$, P < 0.001). The fixed-effect analysis showed essentially the same result (OR = 1.01, 95% CI 1.00 to 1.01, P = 0.21). Subgroup analyses showed no evidence of a difference between studies where BMI was recorded prior to diagnosis (before, n = 39) or studies where BMI was recorded at diagnosis (same time, n = 19) (P for heterogeneity 0.11). In total, 9,252,407 men were included in the meta-analysis; of these, 162,470 men had prostate cancer (1.8%).

One study was markedly different from all others: Chia et al. (2012) (157) estimated an OR of 0.14 (95% CI 0.052 to 0.38), compared with an estimated OR of between 0.58 and 1.41 for all other studies. This study can be seen to the far left of the forest plot. The I^2 value fell to 65.1% when this study was removed, which does not change the overall impression that there was strong evidence of inconsistency in effect estimates across studies. This study presented the number of men with and without prostate cancer in categories of BMI, and used frequency matching with 5-year age groups but did not present ORs for prostate cancer. I thus used the method presented in **Section 4.3.1.D** to

estimate the SMD for BMI between men with and without prostate cancer, as calculating the unadjusted ORs for different categories of BMI could have been confounded by age. Even so, men without prostate cancer were still 3.1 years younger than men with prostate cancer ($P < 0.01$), which may have biased the result. Additionally, there were a relatively high proportion of men with advanced prostate cancer (59 out of 240 men with prostate cancer, 25%); this could also possibly explain the discrepancy, assuming the case-mix of other studies was generally less advanced, and advanced prostate cancers could cause weight loss (reverse causation).

The funnel plot (**Figure 4.4**) did not show convincing evidence of any small study effects. One study had to be removed from the funnel plot as the SE was so large that it obscured the view of all other studies (163). This study presented the results of a logistic regression of prostate cancer with BMI as a covariate, but did not detail the units of BMI. We assumed the regression coefficient for BMI was for a 1 kg/m^2 increase in BMI, but it is possible it was for a different amount of BMI. As the effect estimate was consistent with our results ($\text{OR} = 1.04$, for an assumed unit increase in BMI, 95% CI 0.5 to 2.85), inclusion of this study would be unlikely to change the interpretation of no convincing evidence of systematic bias on the funnel plot.

The albatross plot (using effective sample sizes, **Figure 4.5**) showed that the eleven studies without sufficient information for meta-analysis were spread evenly across both positive and negative effect sizes, consistent with the null result seen in the meta-analysis.

Meta-regression did not show evidence of any variation in results due to ethnicity (ratio of ORs for white versus non-white = 1.01, 95% CI 0.93 to 1.09, $P = 0.85$), overall risk of bias of the study (ratio of ORs for high versus medium = 0.99, 95% CI 0.94 to 1.03, $P = 0.52$), years between BMI measurement and diagnosis (ratio of ORs for a five-year increase = 1.00, 95% CI 0.99 to 1.01, $P = 0.53$), or mean age at diagnosis (ratio of ORs for a five-year age increase = 1.00, 95% CI 0.99 to 1.00, $P = 0.81$). There appeared to be a negative effect of mid-year of recruitment on the effect estimate (ratio of ORs for a five-year increase = 0.98, 95% CI 0.97 to 0.99, $P = 0.003$): I^2 reduced to 35.8% after this was accounted for in meta-regression.

Table 4.2 Data extracted from studies examining the association between BMI and prostate cancer

Author	Year*	Study Name	Study Location	Ethnicity*	Mid-year*	Variables Adjusted ⁺	Effect estimate*	N total (cases)	P value*	Effect type*
Meta-analysis: BMI measured at least two years before prostate cancer diagnosis (before)										
Severson (127)	1988		Oahu, Hawaii	Japanese (100%)	1966	1	1.31 (0.92 to 1.87)	7,994 (174)	0.13	Categorical
Mills (126)	1989	7th Day	USA	White (100%)	1974	1	1.13 (0.84 to 1.52)	35,161 (161)	0.63	Categorical
Andersson (110)	1997	CIOWE	Sweden		1972	1	1.08 (0.99 to 1.17)	137,179 (2,364)	0.10	Categorical
Cerhan (111)	1997	Iowa 65+	USA		1981	1	1.32 (0.9 to 1.92)	1,049 (70)	0.10	Categorical
Heikkila (144)	1999		Finland		1969	1,5	0.84 (0.6 to 1.16)	466 (166)	0.45	SMD
Nilsen (108)	1999	NHSSN	Norway		1983	1	1.02 (0.9 to 1.16)	22,248 (642)	0.77	Categorical
Schuurman (176)	2000	NLCS	Netherlands		1986	1,4,6	1.00 (0.83 to 1.21)	2,240 (675)	10	Continuous
Lee (125)	2001	HAHS	USA		1988	1,4,9,10	1.05 (0.82 to 1.34)	8,930 (439)	0.71	Categorical
Engeland (106)	2003		Norway		1966	1,16	1.07 (1.05 to 1.09)	951,459 (33,314)	0.001	Categorical
Giles (123)	2003		Australia		1995	1	1.04 (0.94 to 1.14)	2,885 (1,409)	0.47	SMD
Jonssonni (122)	2003		Sweden		1961	1,2,4,17	0.98 (0.66 to 1.46)	662 (331)	0.94	Categorical
Kuriyama (105)	2005		Japan		1984	1	1.12 (0.65 to 1.94)	12,263 (45)	0.99	Categorical
Liu (101)	2005		USA	White (91%)	2000	1,2,8,12	0.96 (0.76 to 1.23)	902 (434)	0.73	Categorical
Oh (104)	2005	KNHIC	Korea		1992	1,4,5,9,10,11	1.41 (1.15 to 1.74)	781,283 (387)	0.0012	Categorical
Porter (118)	2005	SEER	USA	Caucasian (95%)	1994	1,2,4,8,9,12,13	0.84 (0.70 to 1.00)	1,456 (753)	0.04	Categorical
Baillargeon (99)	2006	SABOR	USA	Caucasian (86%)	2001	1,2,18	0.78 (0.56 to 1.09)	229 (104)	0.22	Categorical
Haheim (170)	2006	OS	Norway		1972	1	1.10 (0.96 to 1.27)	1,5933 (507)	0.18	Continuous
Kurahashi (100)	2006	JPHC	Japan		1990	1,4,5,9,19	1.20 (0.95 to 1.50)	49,850 (311)	0.13	Categorical
Lukanova (117)	2006	NSHDC	Sweden		1985	1,9,20	0.92 (0.78 to 1.09)	33,424 (461)	0.31	Categorical
Lundqvist (169)	2007	Older cohort Younger cohort	Scandinavia		1975	1,2,4,9,11,8,7	1.00 (0.72 to 1.39) 1.16 (0.71 to 1.89)	874 (437) 430 (215)	0.91 0.58	Continuous
Machova (98)	2007		Czech Republic		1995	1,9,15,21	0.96 (0.77 to 1.21)	17,335 (338)	0.76	Categorical
Rodriguez (37)	2007	CPS-II	USA	White (98%)	1992	1,2,4,7,8, 9,11,12,13	0.98 (0.94 to 1.01)	69,991 (5,252)	0.14	Categorical
Wright (116)	2007	NIH-AARP	USA		1995	1,2,4,7,8,9	0.97 (0.95 to 0.99)	287,760 (9,986)	0.0008	Categorical
Pischon (178)	2008	EPIC	Europe		1996	1,5,8,9,10,11, 15	0.96 (0.9 to 1.02)	129,502 (2,446)	0.20	Continuous
Farhat (148)	2009	MrOS	USA		2001		0.90 (0.76 to 1.06)	4,597 (255)	0.21	SMD
Hernandez (114)	2009	Multiethnic cohort	USA	Multiethnic	1994	1,2,4,8,9,19,24	1.00 (0.95 to 1.05)	83,879 (5,554)	0.62	Categorical

Author	Year*	Study Name	Study Location	Ethnicity*	Mid-year*	Variables Adjusted†	Effect estimate*	N total (cases)	P value*	Effect type*
Wallström (115)	2009	MDCS	Sweden		1994	1,6,7,9,10,11, 12,15,25,26	0.96 (0.85 to 1.08)	10,548 (817)	0.58	Categorical
Burton (168)	2010	GAC	Scotland		1950	6,9,15	1.00 (0.72 to 1.39)	9,549 (211)	0.89	Continuous
Stocks (96)	2010	SCWC	Sweden		1981	1,9,27	1.04 (1.00 to 1.08)	336,159 (10,002)	0.40	Categorical
Mori (113)	2011		Japan		2007	1,5	0.58 (0.33 to 1.03)	394 (117)	0.051	Categorical
Bassett (166)	2012	MCCS	Australia		1992	2,8	1.06 (0.97 to 1.16)	16,514 (1,374)	0.22	Continuous
Häggström (112)	2012	Me-Can	Norway, Sweden, Austria		2011	1,9	0.99 (0.95 to 1.04)	296,539 (6,673)	0.43	Categorical
Shafique (95)	2012	Midspan study	Scotland		1973	1,6,9,28	1.02 (0.88 to 1.18)	12,924 (650)	0.79	Categorical
Kopp (164)	2013	DCH	Denmark		1995	1	0.77 (0.63 to 0.95)	668 (334)	0.01	SMD
Rao (165)	2013		USA	Caucasian (82%)	2003	1,2,7,9,29	0.95 (0.93 to 0.98)	54,4197 (34,275)	0.0001	Continuous
Bhaskaran (173)	2014	CPRD	UK		1986	1,6,7,9,10,20	0.98 (0.96 to 1.01)	5,264,901 (24,901)	0.0042	Continuous
Bhavsar (152)	2014	CLUE II	Washington County, USA	White (98%)	1989	1,2	0.88 (0.68 to 1.13)	536 (268)	0.25	SMD
Lai (135)	2014	HPFS	USA	White (94%)	1986	1	0.97 (0.87 to 1.07)	2,628 (1,314)	0.50	SMD
Yu (133)	2014	PROtEuS	Canada	White (87%)	2007	1	0.89 (0.83 to 0.96)	3,866 (1,904)	0.01	SMD
Meta-analysis: BMI measured less than two years before diagnosis of prostate cancer (same time)										
Whittemore (146)	1995		Hawaii, LA, SF, Vancouver	Black (100%)	1990	1,2	0.91 (0.80 to 1.04)	1,071 (531)	0.16	SMD
				Chinese-American (100%)			1.14 (0.89 to 1.47)	555 (283)	0.29	
				Japanese-American (100%)			1.12 (0.91 to 1.39)	655 (326)	0.29	
				White (100%)			1.02 (0.89 to 1.17)	1,019 (515)	0.72	
Andersson (177)	1996	AMORIS	Sweden		1990	1	1.00 (0.76 to 1.32)	477 (249)	0.83	Continuous
Lagiou (145)	1998		Greece		1993	1	0.75 (0.41 to 1.36)	91 (43)	0.34	SMD
Hsieh (155)	1999		Athens, Greece		1995	1	0.92 (0.40 to 2.10)	572 (320)	0.84	SMD
Villeneuve (107)	1999	NECSS	Canada		1995	1,2,4,5,6,9, 10,12	1.05 (1.02 to 1.09)	2,810 (1413)	0.37	Categorical
Hsing (143)	2001		China		1994	1	0.90 (0.65 to 1.25)	434 (128)	0.55	SMD
Sharpe (124)	2001		Canada		1982	1,2,4,10	1.22 (0.95 to 1.59)	869 (396)	0.12	Categorical

Author	Year*	Study Name	Study Location	Ethnicity*	Mid-year*	Variables Adjusted†	Effect estimate*	N total (cases)	P value*	Effect type*
Cui (171)	2004	NHANES III	USA	White (38%); Black (28%); Mexican-American (30%)	1997	1,2,8,9,10,12	0.73 (0.56 to 0.96)	8,815 (95)	0.023	Continuous
Friedenreich (121)	2004		Canada		1999	1,3,4,5,8,10, 11,13	1.04 (0.91 to 1.19)	2,049 (987)	0.57	Categorical
Bidoli (154)	2005		Italy		1996	1	1.03 (0.91 to 1.16)	2,737 (1,290)	0.62	SMD
Bradbury (103)	2005		UK		1996	1,5,7,9	0.96 (0.84 to 1.11)	3,200 (730)	0.04	Categorical
Gallus (97)	2007		Italy		1996	1,4,8,5,11	1.12 (0.84 to 1.51)	648 (219)	0.39	Categorical
Ahn (140)	2008	PLCO	USA	Caucasian (89%)	1997	1	0.93 (0.82 to 1.06)	1,530 (749)	0.19	SMD
Chamie (174)	2008		USA		2002	1,2,9,18,22,2 3	1.10 (0.95 to 1.28)	13,144 (363)	0.21	Continuous
Chia (157)	2012		Singapore		0	1,2	0.14 (0.05 to 0.38)	396 (212)	0.00011	SMD
Nemesure (167)	2012	PCBP	Barbados	Black (100%)	2007	1	1.00 (0.90 to 1.11)	1,904 (963)	1	Continuous
Edwards (134)	2013	NMHS	USA	European (86%)	0		0.94 (0.85 to 1.04)	1,777 (871)	0.26	SMD
Möller (162)	2013	CAPS	Sweden		2001	1,5,7,8,11	0.93 (0.79 to 1.09)	1,601 (935)	0.38	Continuous
Salem (163)	2013		Iran		2010	1,30,31	1.22 (0.02 to 94.37)	511 (194)	0.35	Continuous
Albatross only: BMI measured at least two years before prostate cancer diagnosis (before)										
Le Marchand (188)	1994		Hawaii		1977	1,2,6	Positive effect	8,881 (198)	0.20	RR Trend
Veierød (187)	1997		Norway		1980	1	Positive effect	25,778 (70)	0.02	IRR Trend
Habel (186)	2000	KPMCP	California, USA	White (79%), Black (12%)	1968	1,2	Positive effect	29,117 (28,285)	0.90	RR BMI>27.9 vs BMI<22.7
Putnam (185)	2000		Iowa, USA		1987	1,4,12	Positive effect	1,555 (81)	0.08	RR trend
Cox (181)	2006		New Zealand	European descent	1997	1	Positive effect	495 (199)	0.08	RR BMI q5 vs BMI q1
Littman (190)	2007	VITAL	USA	White (93%)	2001	1,2,4,13	Positive effect	34,868 (808)	0.13	HR trend
Attner (179)	2012	Skaine study	Sweden		2006	1	Positive effect	30,199 (3545)	0.08	RR obesity (B ^a)
Harding (189)	2015	ANZDCC	Australia/NZ		1993	1	Positive effect	40,734 (2,866)	0.63	HR BMI SD increase
Albatross only: BMI measured less than two years before diagnosis of prostate cancer (same time)										
Hsing (184)	2000		China		1994	1,8,12,19	Positive effect	709 (238)	1	OR trend
Pan (183)	2004	NECSS	Canada		1995	1,5,8,9,11,12	Positive effect	4,348 (1,801)	0.026	RR trend

Author	Year*	Study Name	Study Location	Ethnicity*	Mid-year*	Variables Adjusted†	Effect estimate*	N total (cases)	P value*	Effect type*
*Year = publication year, Ethnicity = Ethnicity (%) if specified, Mid-year = the mid-year of recruitment for each study (publication date – 1 year if not reported) Effect estimate = OR for prostate cancer for a 5 kg/m ² increase in BMI (meta-analysis) or effect direction (albatross), P value = P value for effect estimate, Effect type = original presentation of data (meta-analysis) or to which effect estimate the P value refers (albatross)										
+Variables adjusted for: 1 = Age, 2 = Ethnicity, 3 = DRE, 4 = Family History, 5 = Area, 6 = Income, 7 = Diabetes, 8 = Education, 9 = Smoking, 10 = Alcohol, 11 = Physical Activity, 12 = Diet, 13 = PSA test, 14 = BMI at younger age, 15 = Height, 16 = Cohort, 17 = Twins, 18 = PSA, 19 = Marital Status, 20 = Year, 21 = Hypertension, 22 = Finasteride, 23 = Agent Orange Exposure, 24 = Birthplace, 25 = Co-habitation Status, 26 = Birth Country, 27 = Blood Pressure, 28 = Cholesterol, 29 = Angiotensin receptor blocker, 30 = Calcium, 31 = Sex hormones										
Study Acronyms: AMORIS = The Swedish Apolipoprotein MOrtality RISK study, ANZDCC = Australia and New Zealand Diabetes and Cancer Collaboration, CAPS = Cancer of the Prostate in Sweden, CIOWE = The Construction Industry's Organization for Working Environment (Safety and Health), CLUE II = Campaign Against Cancer and Heart Disease, CPRD = Clinical Practice Research Datalink, CPS-II = Cancer Prevention Study II, DCH = Danish Diet, Cancer and Health, EPIC = European Prospective Investigation into Cancer and Nutrition, GAC = Glasgow Alumni Cohort, HAHS = Harvard Alumni Health Study, HPFS = Health Professionals Follow-Up Study, Iowa 65+ = Iowa 65+ rural health study, JPHC = Japan Public Health Center-based Prospective Study, KNHIC = Korea National Health Insurance Company, KPMCP = Kaiser Permanente medical care program, MCCS = Melbourne Collaborative Cohort Study, MDCS = Malmo Diet and Cancer, Me-Can = Metabolic syndrome and Cancer, MrOS = Osteopathic Fracture in Men study, NECSS = Canadian National Enhanced Cancer Surveillance System, NHANES = National Health and Nutrition Examination Survey, NHSSN = National Health Screening Service in Norway, NIH-AARP = National Institute of Health American Association of Retired Persons diet and health study, NLCS = The Netherlands cohort study, NMHS = Nashville Men's Health Study, NSHDC = North Sweden Health and Disease Cohort, OS = Oslo Study, PCBP = Prostate Cancer in a Black Population, PLCO = Prostate, lung, colorectal and ovarian cancer screening trial, PRTOTeuS = Prospective, randomized, multicentre, open label, phase II / III study to assess efficacy and safety of ranibizumab 0.5 mg intravitreal injections plus panretinal photoocoagulation (PRP) versus PRP in monotherapy in the treatment of subjects with high risk proliferative diabetic retinopathy, SABOR = San Antonio Center for Biomarkers of Risk of prostate carcinoma, SCWC = Swedish Construction Workers Cohort, SEER = Surveillance, Epidemiology and End Results Program, VITAL = Vitamins and lifestyle cohort										

Table 4.3 Risk of bias of studies examining the association between BMI and prostate cancer (see Section 2.3.2 and Appendix 5)

Author	Overall RoB	Confounding	Participants	Missing Data	Measurement of Outcome	Measurement of Exposure	Selective Reporting
Meta-analysis: BMI measured at least two years before prostate cancer diagnosis (before)							
Severson (127)	High	Medium	Low	Medium	Medium	Medium	Low
Mills (126)	High	Medium	Medium	Low	Medium	Medium	Low
Andersson (110)	Medium	Medium	Low	Medium	Medium	Low	Low
Cerhan (111)	High	Medium	Low	Medium	Medium	Medium	Low
Heikkila (144)	Medium	Medium	Low	Low	Medium	Low	Low
Nilsen (108)	Medium	Medium	Low	Low	Medium	Low	Low
Schuurman (176)	High	Medium	Low	Medium	Medium	Medium	Low
Lee (125)	High	Medium	Low	Medium	Medium	Medium	Low
Engeland (106)	Medium	Medium	Low	Medium	Medium	Low	Low
Giles (123)	Medium	Medium	Low	Low	Medium	Medium	Low
Jonssoni (122)	Medium	Medium	Low	Low	Medium	Medium	Low
Kuriyama (105)	Medium	Medium	Low	Medium	Medium	Medium	Low
Liu (101)	Medium	Medium	Medium	Low	Medium	Medium	Low
Oh (104)	Medium	Medium	Low	Medium	Low	Medium	Low
Porter (118)	Medium	Medium	Medium	Medium	Medium	Medium	Low
Baillargeon (99)	Medium	Medium	Low	Low	Medium	Low	Low
Haheim (170)	Medium	Medium	Low	Low	Low	Low	Low
Kurahashi (100)	Medium	Medium	Low	Medium	Medium	Medium	Low
Lukanova (117)	High	High	Low	Medium	Medium	Medium	Low
Lundqvist (169)	Medium	Low	Low	Medium	Low	Medium	Low
Machova (98)	Medium	Medium	Low	Low	Medium	Low	Low
Rodriguez (37)	High	Medium	Low	Medium	Medium	Medium	Low
Wright (116)	Medium	Medium	Low	Medium	Medium	Medium	Low
Pischon (178)	Medium	Medium	Low	Medium	Medium	Low	Low
Farhat (148)	High	High	Low	Medium	Medium	Medium	Low
Hernandez (114)	Medium	Medium	Low	Medium	Medium	Medium	Low
Wallström (115)	High	High	Low	Medium	Medium	Low	Low
Burton (168)	High	High	Low	Low	Low	Low	Low
Stocks (96)	Medium	Medium	Low	Low	Low	Low	Low
Mori (113)	High	Medium	Medium	Low	Medium	Medium	Low
Bassett (166)	High	High	Low	Medium	Medium	Low	Low
Häggström (112)	Medium	Medium	Low	Medium	Medium	Medium	Low
Shafique (95)	Medium	Medium	Low	Low	Low	Low	Low
Kopp (164)	Medium	Medium	Low	Low	Low	Low	Low
Rao (165)	High	Medium	High	Low	Unclear	Unclear	Low

Author	Overall RoB	Confounding	Participants	Missing Data	Measurement of Outcome	Measurement of Exposure	Selective Reporting
Bhaskaran (173)	Medium	Medium	Low	Medium	Medium	Medium	Low
Bhavsar (152)	Medium	Medium	Low	Low	Medium	Low	Low
Lai (135)	Medium	Medium	Low	Low	Medium	Low	Low
Yu (133)	High	High	Low	Medium	High	Medium	Low
Meta-analysis: BMI measured less than two years before diagnosis of prostate cancer (same time)							
Whittemore (146)	High	High	Medium	Medium	Medium	Medium	Low
Andersson (177)	Medium	Medium	Medium	Medium	Medium	Low	Low
Lagiou (145)	High	High	Low	Low	Medium	Low	Low
Hsieh (155)	High	High	Medium	Medium	Medium	Medium	Low
Villeneuve (107)	Medium	Medium	Low	Medium	Medium	Medium	Low
Hsing (143)	High	High	High	High	Medium	Low	Low
Sharpe (124)	Medium	Medium	Low	Medium	Medium	Medium	Low
Cui (171)	Medium	Medium	Low	Low	Medium	Medium	Low
Friedenreich (121)	Medium	Medium	Low	Low	Medium	Low	Low
Bidoli (154)	High	High	Medium	Low	Medium	Medium	Low
Bradbury (103)	Medium	Medium	Low	Low	Medium	Medium	Low
Gallus (97)	High	Medium	Medium	Low	Medium	Medium	Low
Ahn (140)	High	High	Medium	Low	Low	Low	Low
Chamie (174)	High	Medium	Medium	Medium	Medium	Medium	Low
Chia (157)	High	High	Medium	Medium	Low	Medium	Low
Nemesure (167)	High	High	Medium	Low	Low	Low	Low
Edwards (134)	High	High	Medium	Low	Low	Low	Low
Möller (162)	High	Medium	Low	Medium	Medium	Medium	Low
Salem (163)	High	High	High	Low	Low	Unclear	Low
Albatross only: BMI measured at least two years before prostate cancer diagnosis (before)							
Le Marchand (188)	High	Medium	Low	Medium	Medium	Medium	Low
Veierød (187)	High	Medium	Low	Medium	Medium	Medium	Low
Habel (186)	Medium	Medium	Low	Medium	Medium	Low	Low
Putnam (185)	High	Medium	Low	Medium	Medium	Medium	Low
Cox (181)	High	Medium	Medium	Medium	Medium	Medium	Low
Littman (190)	Medium	Medium	Low	Low	Medium	Medium	Low
Attner (179)	Medium	Medium	Low	Medium	Medium	Medium	Low
Harding (189)	Medium	Medium	Low	Medium	Medium	Low	Low
Albatross only: BMI measured less than two years before diagnosis of prostate cancer (same time)							
Hsing (184)	High	Medium	Low	Medium	Medium	Medium	Low
Pan (183)	Medium	Medium	Low	Medium	Medium	Medium	Low

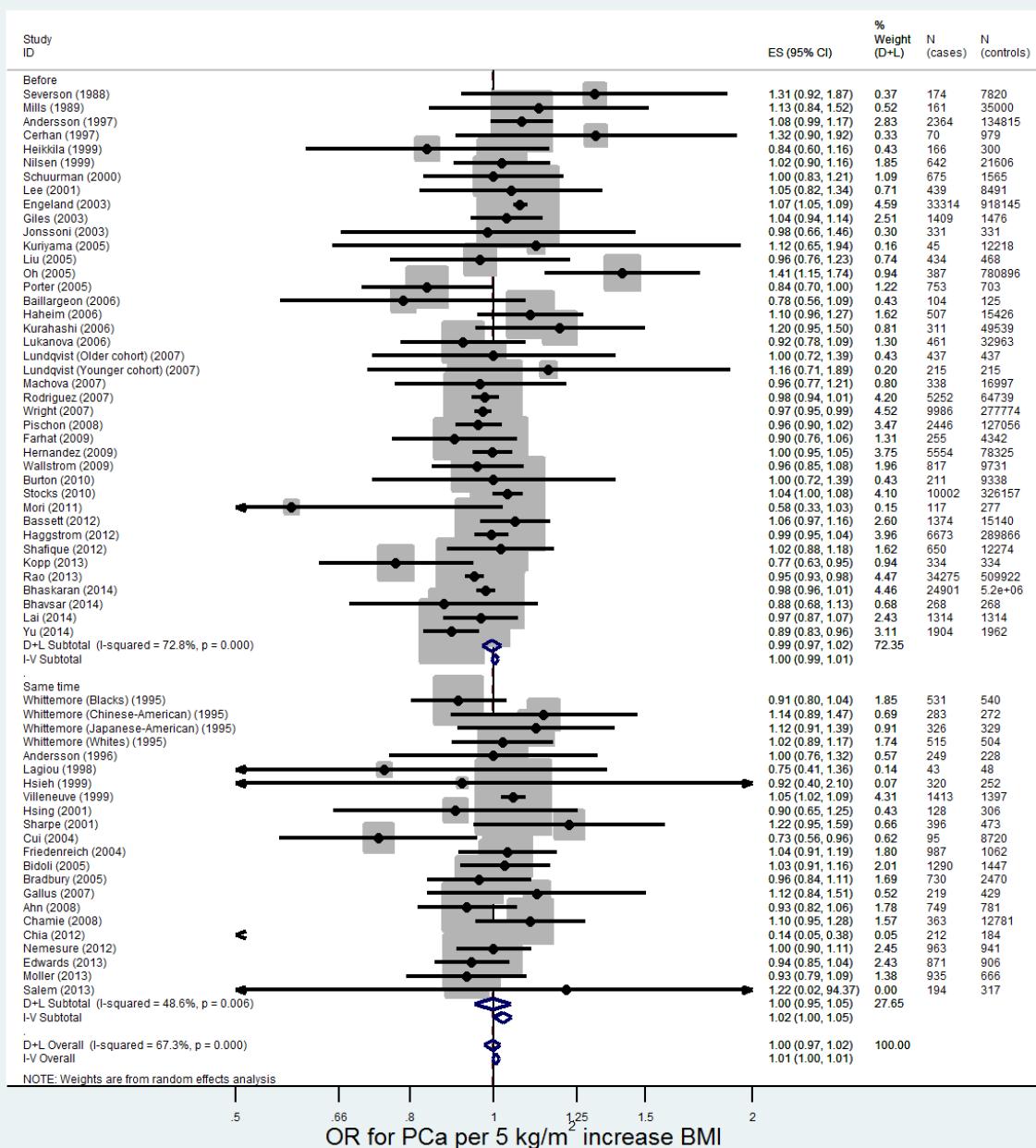


Figure 4.3 Forest plot for the association between BMI and prostate cancer

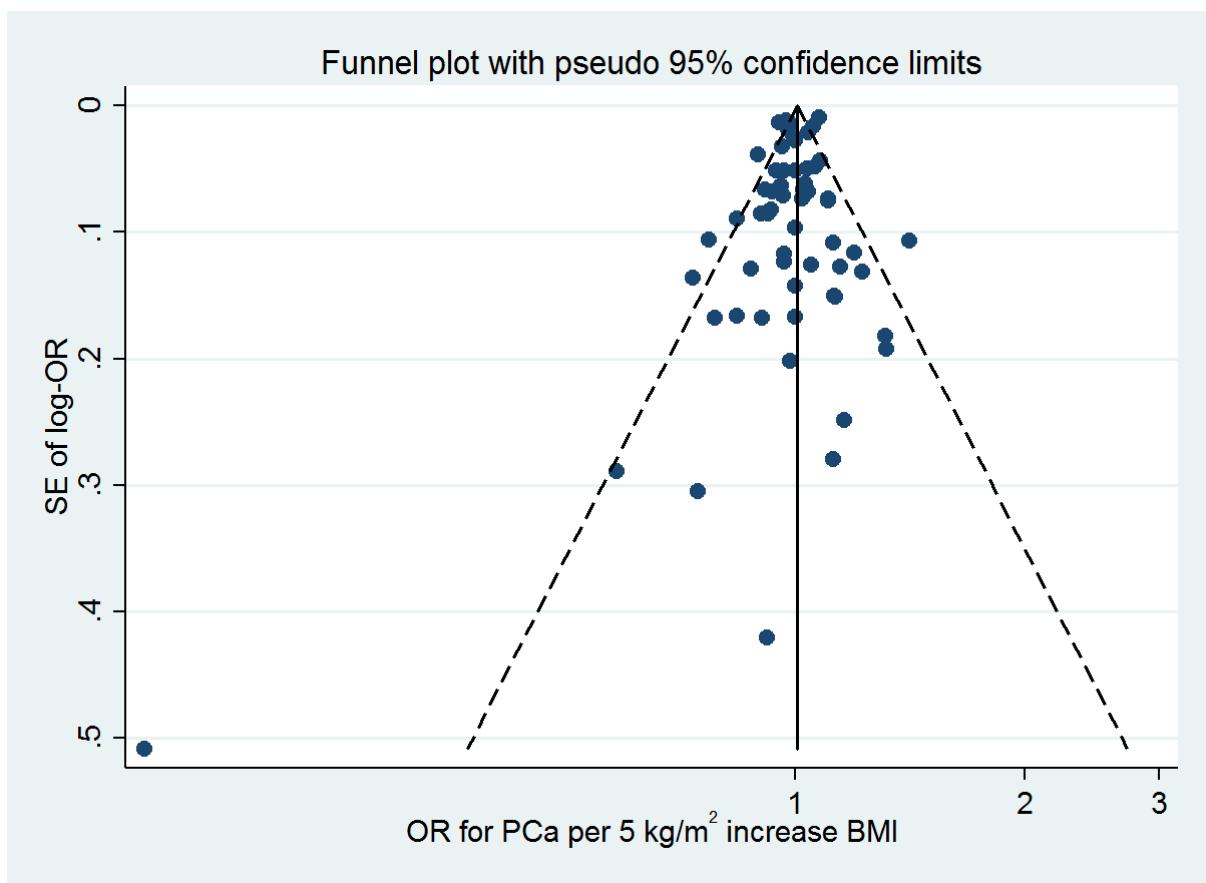


Figure 4.4 Funnel plot for the association between BMI and prostate cancer

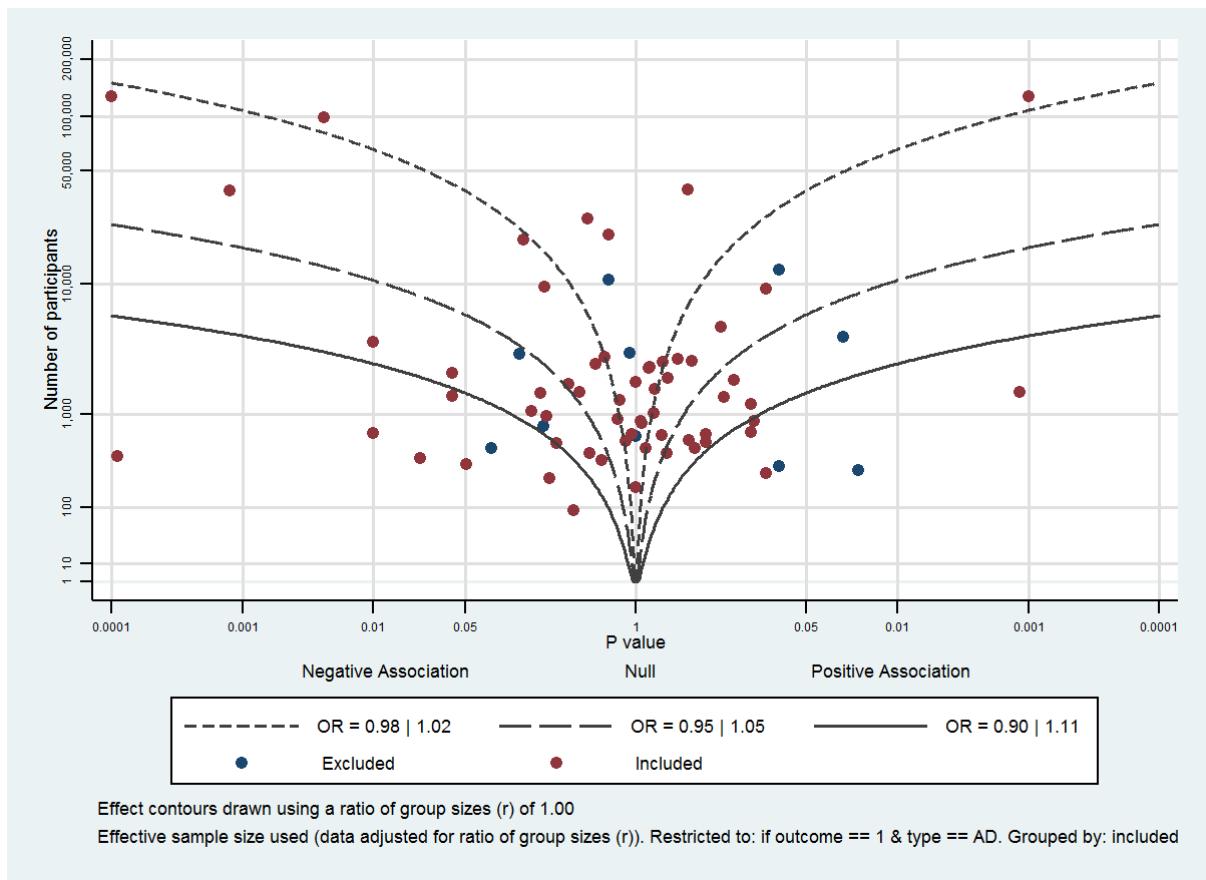


Figure 4.5 Albatross plot for the association between BMI and prostate cancer

4.4.2. BMI and Advanced Prostate Cancer

Of the 18 studies examining the association between BMI and advanced prostate cancer, 14 studies had enough information to be included in the meta-analysis. The remaining 4 studies (22%) were included in the albatross plot. The studies examining the association between BMI and advanced prostate cancer are detailed in **Table 4.4**, with the results of the risk of bias assessment in **Table 4.5**.

The random-effects meta-analysis (**Figure 4.6**) estimated the average OR for a 5 kg/m² increase in BMI to be 1.05 (95% CI 0.99 to 1.10, P = 0.09). There was a moderate evidence for inconsistency in effect estimates across studies ($I^2 = 35.1\%$, P = 0.0032). The fixed-effect analysis showed essentially the same result (OR = 1.05, 95% CI 1.01 to 1.09, P = 0.023). This result did not vary between studies where BMI was recorded prior to diagnosis (before, n=12) or studies where BMI was recorded at diagnosis (same time, n=2), although as only two studies measured BMI at the time advanced prostate cancer was diagnosed there was limited power to detect a difference (P for heterogeneity = 0.28). In total, 1,053,109 men were included in the meta-analysis; of these, 8,357 men (0.79%) had advanced prostate cancer.

The funnel plot (**Figure 4.7**) did not show convincing evidence of any small study effects. The albatross plot (using effective sample sizes, **Figure 4.8**) showed that the four studies without sufficient information for meta-analysis all showed a positive association between BMI and advanced prostate cancer risk. Two small studies Cerhan (1997) (111) and Putnam (2000) (185) showed inconsistently strong effects (111,185) (cases = 17, controls = 814 and cases = 18, controls = 1,474, respectively), the P values for which were taken from the P value for trend across categories of BMI. In the Cerhan study, there were four categories of BMI, with between 3 and 5 men with advanced prostate cancer in each, and a P for trend of 0.1. Equally, in the Putnam study, there were three categories of BMI, with between 2 and 11 men with advanced prostate cancer in each, and a P for trend of 0.02. Because these studies were very small (the next smallest study had at least five times as many men with advanced prostate cancer), they did not change our overall interpretation of a possibly slightly positive association between BMI and advanced prostate cancer.

Meta-regression did not show evidence of any variation in results due to ethnicity (ratio of ORs for whites versus non-whites = 0.86, 95% CI 0.53 to 1.38, P = 0.42), overall risk of bias of the study (ratio of ORs for high versus medium = 0.95, 95% CI 0.74 to 1.22, P = 0.62), mid-year of recruitment (ratio of ORs for a five-year increase = 0.97, 95% CI -0.88 to 1.07, P = 0.44), years between BMI measurement and diagnosis (ratio of ORs for a five-year increase = 0.99, 95% CI 0.95 to 1.03, P = 0.51), or mean age at diagnosis (ratio of ORs for a five-year increase = 1.02, 95% CI 0.98 to 1.05, P = 0.31).

Table 4.4 Data extracted from studies examining the association between BMI and advanced prostate cancer

Table 4.5 Risk of bias of studies examining the association between BMI and advanced prostate cancer (see Section 2.3.2 and Appendix 5)

Author	Overall RoB	Confounding	Participants	Missing Data	Measurement of Outcome	Measurement of Exposure	Selective Reporting
Meta-analysis: BMI measured at least two years before prostate cancer diagnosis (before)							
Giovannucci (109)	High	Medium	Low	Medium	Medium	Medium	Low
Schuurman (176)	High	Medium	Low	Medium	Medium	Medium	Low
Robinson (158)	Medium	Medium	Low	Medium	Low	Medium	Low
Kurahashi (100)	Medium	Medium	Low	Medium	Medium	Medium	Low
Rodriguez (37)	High	Medium	Low	Medium	Medium	Medium	Low
Wright (116)	Medium	Medium	Low	Medium	Medium	Medium	Low
Pischon (178)	Medium	Medium	Low	Medium	Medium	Low	Low
Hernandez (114)	Medium	Medium	Low	Medium	Medium	Medium	Low
Wallström (115)	High	High	Low	Medium	Medium	Low	Low
Stocks (96)	Medium	Medium	Low	Low	Low	Low	Low
Discacciati (175)	Medium	Medium	Low	Medium	Medium	Medium	Low
Bassett (166)	High	High	Low	Medium	Medium	Low	Low
Meta-analysis: BMI measured less than two years before diagnosis of prostate cancer (same time)							
Chamie (174)	High	Medium	Medium	Medium	Medium	Medium	Low
Möller (162)	High	Medium	Low	Medium	Medium	Medium	Low
Albatross only: BMI measured at least two years before prostate cancer diagnosis (before)							
Cerhan (111)	High	Medium	Low	Medium	Medium	Medium	Low
Putnam (185)	High	Medium	Low	Medium	Medium	Medium	Low
Littman (190)	Medium	Medium	Low	Low	Medium	Medium	Low
Albatross only: BMI measured less than two years before diagnosis of prostate cancer (same time)							
Hsing (184)	High	Medium	Low	Medium	Medium	Medium	Low

101

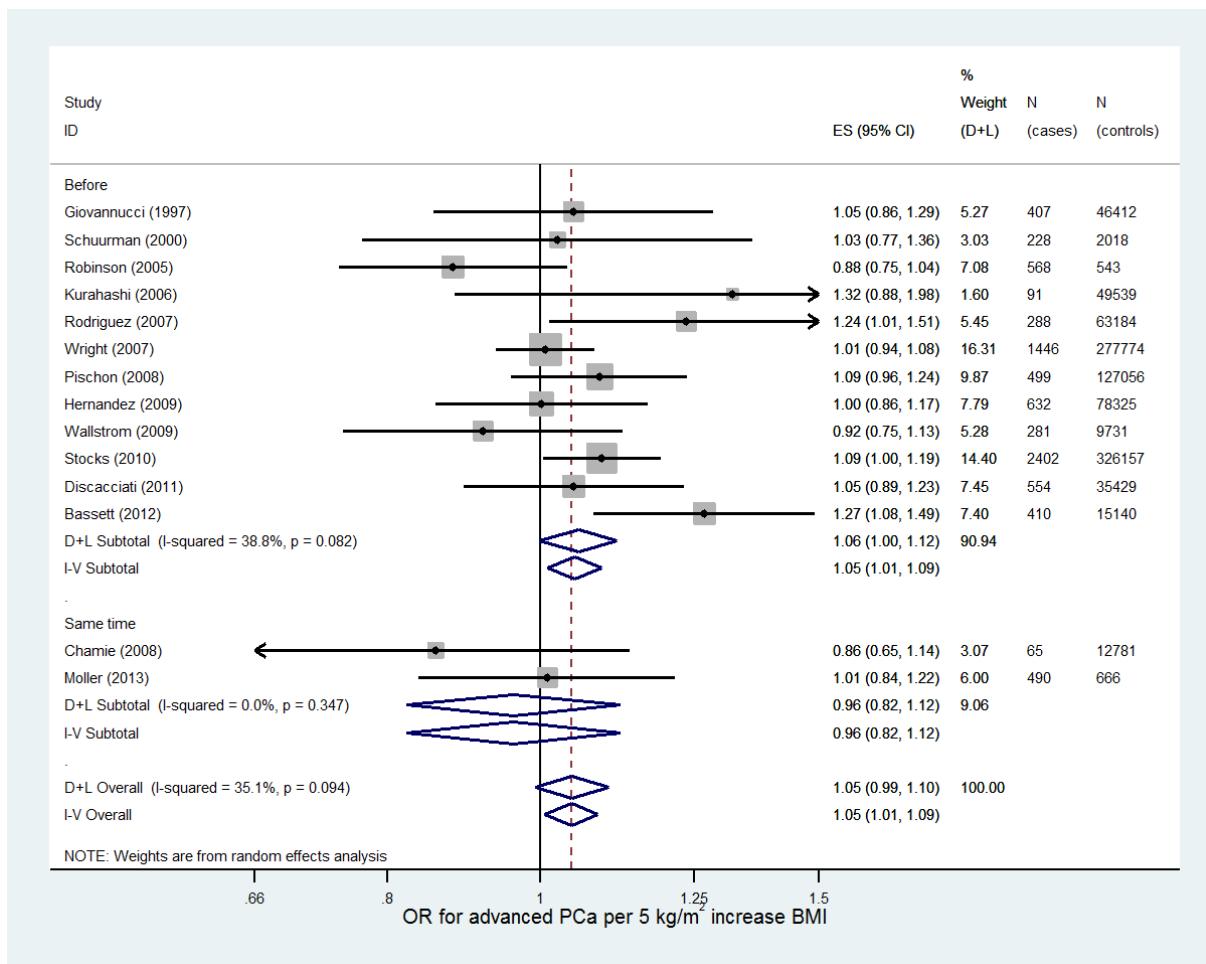


Figure 4.6 Forest plot for the association between BMI and advanced prostate cancer

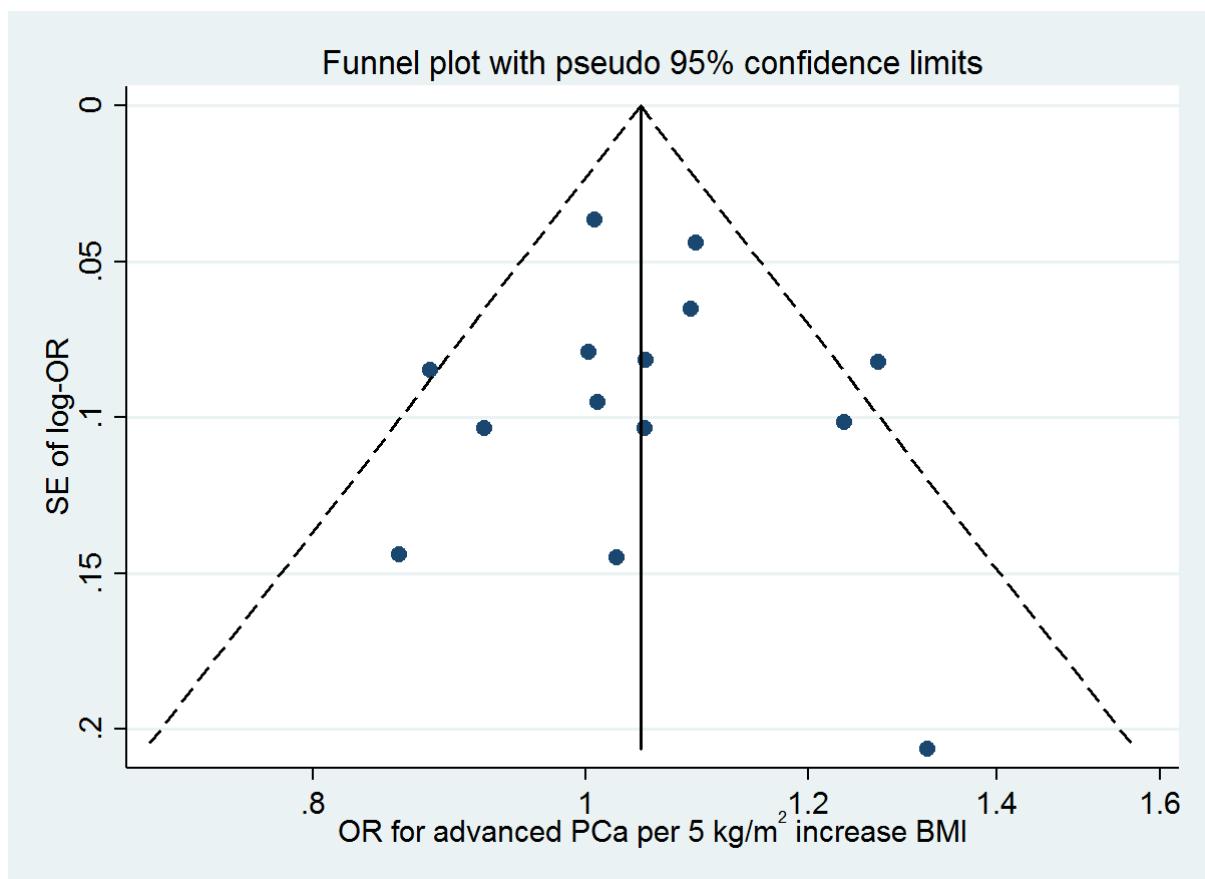


Figure 4.7 Funnel plot for the association between BMI and advanced prostate cancer

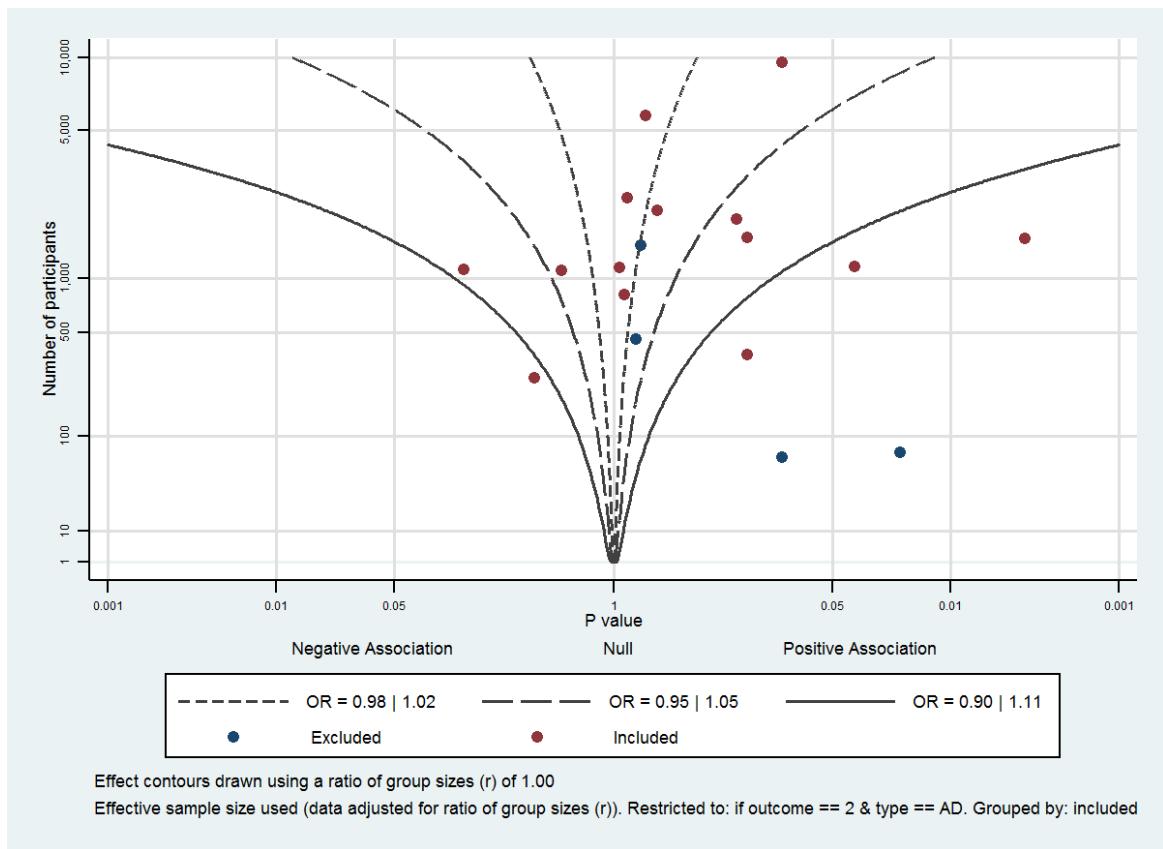


Figure 4.8 Albatross plot for the association between BMI and advanced prostate cancer

4.4.3. BMI and PSA

Data were extracted from 39 studies examining the association between BMI and PSA. After assessing risk of bias, 26 studies were excluded because of a critical risk of bias, leaving 13 studies.

Reasons for critical risk of bias were: bias due to confounding (e.g. not adjusting for age in study design or analysis) in 12 studies (46%); bias due to selection of participants (e.g. selecting on PSA levels) in 14 studies (54%).

In total, 10 studies were included in the meta-analysis of BMI and PSA; three studies were not included due to insufficient data, but were included in the albatross plot. All included studies are detailed in **Table 4.6**, with the results of the risk of bias assessment in **Table 4.7**.

The random-effects meta-analysis (**Figure 4.9**) estimated the average percentage change in PSA for a 5 kg/m² increase in BMI to be -5.16% (95% CI -6.85% to -3.44%, P < 0.001). There was strong evidence for inconsistency in effect estimates across studies ($I^2 = 63.7\%$, P < 0.001). The fixed-effect analysis showed essentially the same result with narrower CIs (percentage change in PSA = -5.16%, 95% CI -6.01% to -4.30%, P < 0.001). In total, 63,648 men were included in the meta-analysis.

The funnel plot (**Figure 4.10**) showed some evidence of small-study effects, as three of four smallest studies (35,192,195) showed a larger association than the average across studies. To be symmetric, the funnel plot would include small studies with a null association. The albatross plot (**Figure 4.11**) showed that of the three studies without sufficient information for meta-analysis, two were consistent with the meta-analysis effect size (199,207), and one was inconsistent as PSA increased as BMI increased (201). This was the only study showing a positive effect, and was also the only study where the main ethnicity of men was Black; all other studies were conducted in Caucasian or Asian populations. This inconsistent study is not enough to change the interpretation of the evidence overall, but does caution that results may not be applicable to all ethnicities.

Meta-regression did not show evidence of any variation in results due to ethnicity (difference in beta coefficients for white versus non-white = 0.025, 95% CI -0.21 to 0.26, P = 0.79), mid-year of recruitment (difference in beta coefficients for a five-year increase = -0.051, 95% CI -0.13 to 0.024, P = 0.13), mean BMI of the study (difference in beta coefficients for a 5 kg/m² increase = -0.097, 95% CI -0.046 to 0.27, P = 0.50) or overall risk of bias of the study (difference in beta coefficients high versus medium = 0.11, 95% CI -0.061 to 0.27, P = 0.15).

Table 4.6 Data extracted from studies examining the association between BMI and PSA

Author	Year	Study	Study Name	Ethnicity	Mid-	Variables	Participants	Effect estimate	P value	Effect type
Meta-analysis										
Baillargeon (82)	2005	USA	SABOR	White (87%)	2002	1,2	2,770	-1.51% (-4.52% to 1.60%)	0.34	Categories
Freedland (191)	2006	USA		White (61%), Black (29%)	1996	1,2,5	1,414	-1.36% (-6.70% to 4.27%)	0.63	Categories
Bañez (80)	2007	USA	Duke	White (84%), Black (15%)	1995	1,2,5	1,974	-7.39% (-10.9% to -3.72%)	0.00011	Categories
			Johns Hopkins	White (91%), Black (6%)			10,287	-2.35% (-4.16% to -0.51%)	0.012	
			SEARCH	White (52%), Black (41%)			1,373	-6.4% (-9.42% to -3.28%)	<0.0001	
Price (192)	2008	USA		White (37%), Black (59%)	2006	1,2,3,4,8,9	535	-9.34% (-15.7% to -2.55%)	0.0078	Categories
Culp (193)	2009	USA	NHANES	White (54%)	2004	1,2	3,152	-5.35% (-8.00% to -2.62%)	0.00015	Categories
Muller (35)	2009	Germany	ESTHER	German	2001	1	777	-8.37% (-16.94% to 1.09%)	0.081	Categories
Park (194)	2009	Korea		Asian	2006	1	38,410	-7.17% (-8.59% to -5.72%)	<0.0001	Categories
Wright (195)	2011	USA			2003	1,6,7	770	-10.4% (-18.1% to -2.07%)	0.016	Categories
Chamie (202)	2013	USA		White (55%), Hispanic (26%)	2006	1,2,5	573	-4.88% (-13.8% to 4.92%)	0.32	Continuous
Bhindi (196)	2014	Canada	Genitourinary BioBank	European (73%), Asian (12%)	2012		1,613	-4.07% (-7.52% to -0.49%)	0.026	Categories
Albatross only										
Wallner (207)	2011	USA	OCS	Caucasian	1990		545	Negative effect	0.18	PSA intercept
Ikuerowo (201)	2012	Nigeria		African	2011	1,2	1,954	Positive effect	0.07	Obese ≥ 4.0 ng/ml
Yang (199)	2013	Korea		Asian	2010	1	20,509	Negative effect	0.082	Obese ≥ 4.0 ng/ml
<p>Year = publication year, Ethnicity (%) if specified, Mid-year = the mid-year of recruitment for each study (publication data – 1 year if not reported), Effect estimate = Percentage change in PSA for a 5 kg/m² increase in BMI (meta-analysis) or effect direction (albatross), P value = P value for effect estimate, Effect type = original presentation of data (meta-analysis) or to which effect estimate the P value refers (albatross)</p> <p>Study acronyms: ESTHER = Epidemiological study on chances of prevention, early detection, NHANES = National Health and Nutrition Examination Survey, and treatment optimization of chronic diseases in the elderly, OCS = Olmsted County Study, SABOR = San Antonio Center for Biomarkers of Risk of prostate carcinoma, SEARCH = Shared equal access regional cancer hospital</p> <p>Variables adjusted: 1 = Age, 2 = Ethnicity, 3 = DRE, 4 = Family History, 5 = Prostate Cancer Characteristics, 6 = Statins & aspirin, 7 = Diabetes, 8 = Education, 9 = Vasectomy</p>										

Table 4.7 Risk of bias of studies examining the association between BMI and PSA (see Section 2.3.2 and Appendix 5)

Author	Overall RoB	Confounding	Selection of Participants	Missing Data	Measurement of Outcome	Measurement of Exposure	Selective Reporting
Meta-analysis							
Baillargeon (82)	Medium	Medium	Low	Low	Medium	Low	Low
Freedland (191)	High	Medium	High	Low	Low	Medium	Low
Bañez (80)	High	Medium	High	Low	Low	Medium	Low
Price (192)	Medium	Medium	Low	Low	Low	Unclear	Low
Culp (193)	High	High	Low	Low	Low	Unclear	Low
Muller (35)	High	High	Low	Low	Low	Medium	Low
Park (194)	High	High	Medium	Low	Low	Low	Low
Wright (195)	High	Medium	Medium	Medium	Low	Medium	Low
Chamie (202)	Medium	Medium	Medium	Low	Low	Medium	Low
Bhindi (196)	High	High	Low	Low	Low	Low	Low
Albatross only							
Wallner (207)	Medium	Medium	Low	Low	Low	Low	Low
Ikuerowo (201)	High	High	Low	Low	Low	Low	Low
Yang (199)	Medium	Medium	Low	Low	Low	Low	Low

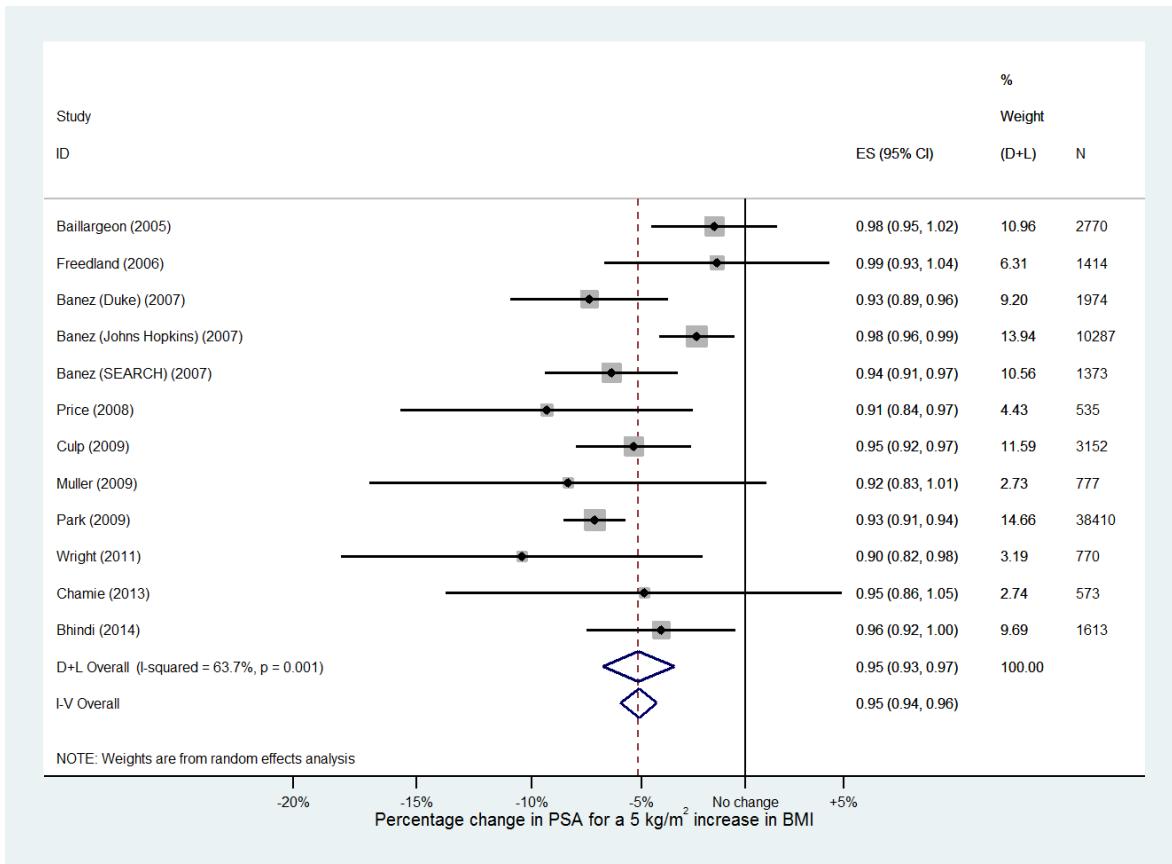


Figure 4.9 Forest plot for the association between BMI and PSA

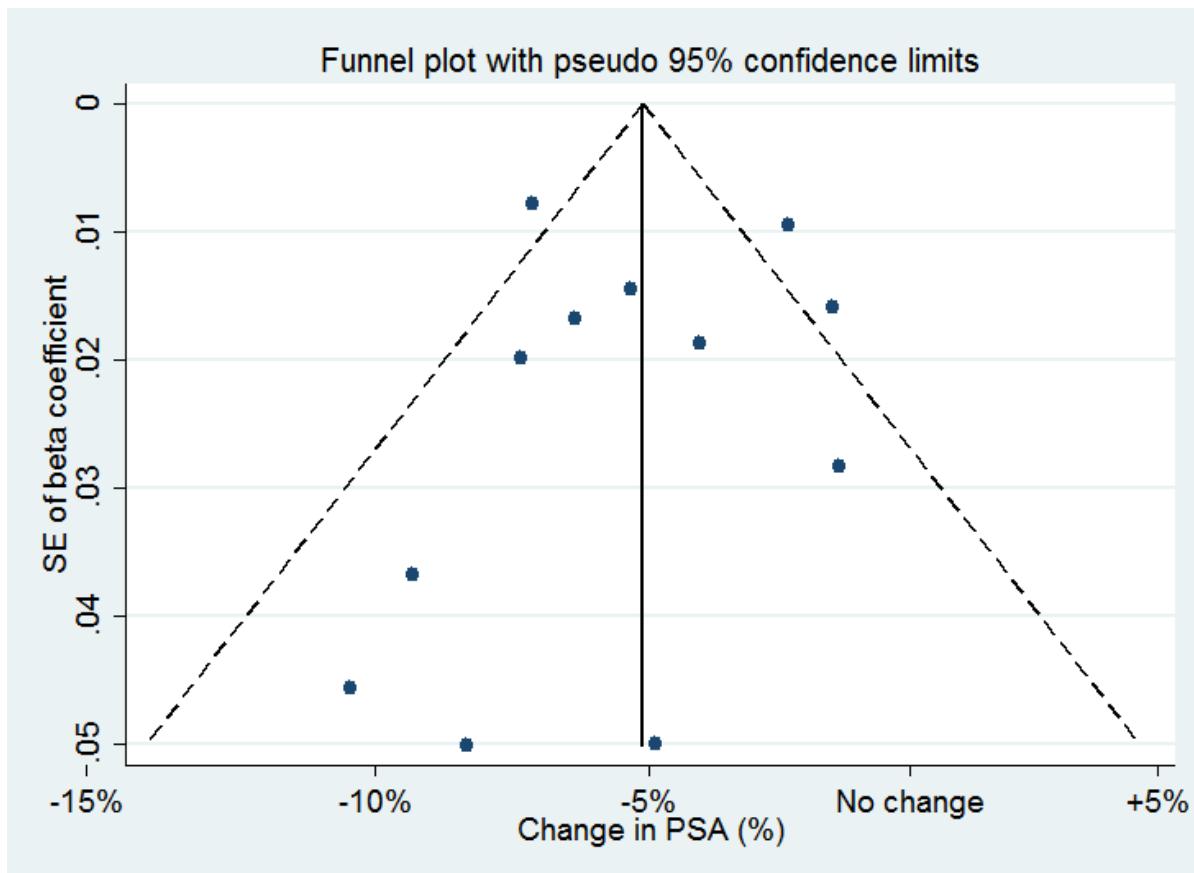


Figure 4.10 Funnel plot for the association between BMI and PSA

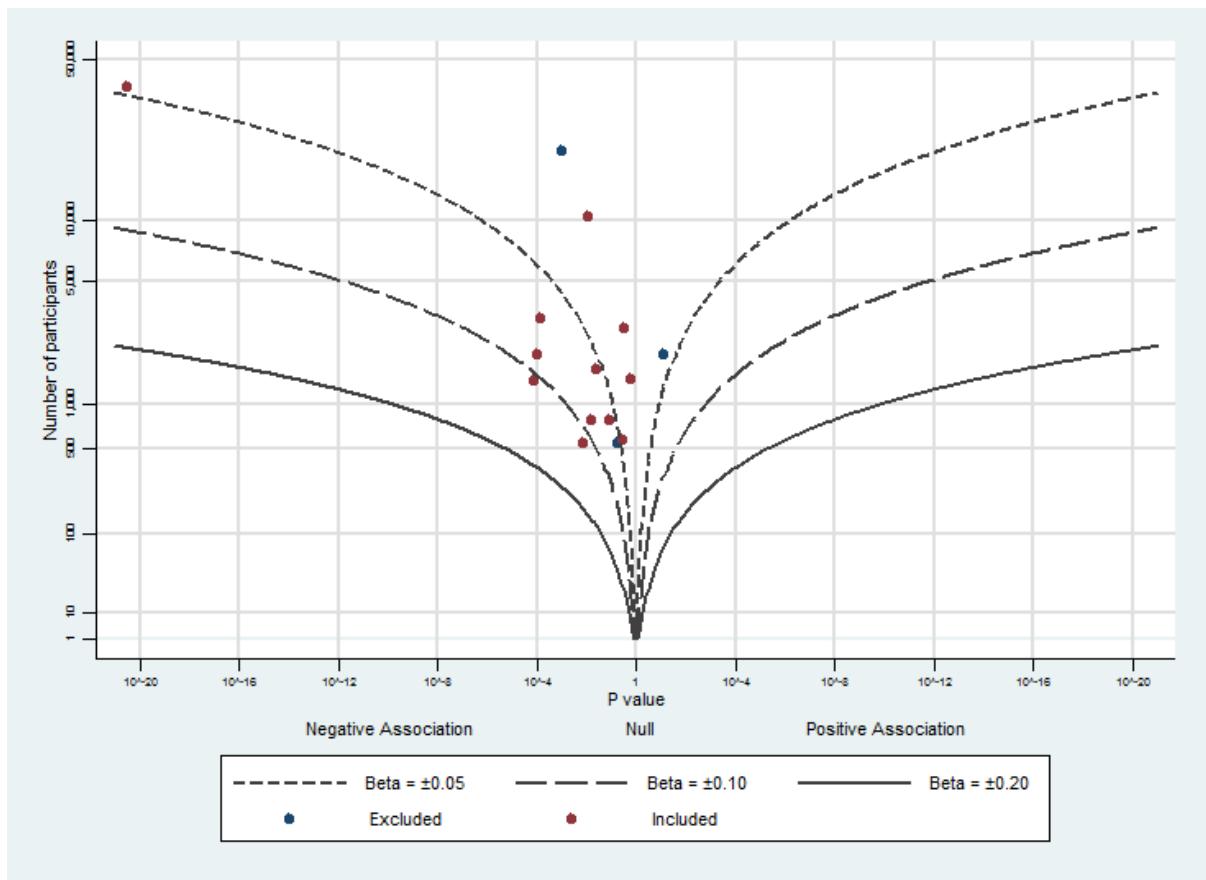


Figure 4.11 Albatross plot for the association between BMI and log-PSA

4.5. Discussion

From analysis of AD, there was no evidence to suggest there is an association between BMI and prostate cancer risk. This was true both for studies where BMI was measured at diagnosis of prostate cancer, and studies where BMI was measured before diagnosis. Equally, the number of years between measurement of BMI and diagnosis of prostate cancer did not appear to alter the association between BMI and prostate cancer risk. However, the mid-year of study recruitment was associated with the effect estimate, showing that for every 5-year increase the OR for prostate cancer for a 5 kg/m² increase in BMI was reduced (ratio of ORs = 0.98, 95% CI 0.97 to 0.99, P = 0.003). This could be a result of increased testing for prostate cancer using PSA: there was strong evidence to suggest that as BMI increases, PSA decreases, which means men with a larger BMI are less likely to be offered a biopsy following PSA testing (as PSA levels would fall below the threshold triggering further investigation) and are therefore less likely to receive a diagnosis of prostate cancer. The apparent reduction in OR for prostate cancer due to the mid-year of recruitment could also possibly be explained by changing demographics of participants over time, but meta-regression showed no evidence of an association of effect estimates with ethnicity or mean age at diagnosis, making it more likely that the association between BMI and prostate cancer diagnosis has changed over time. We therefore conclude that there is no evidence for a linear association between BMI and prostate cancer. It is, however, possible that PSA testing is masking a positive association between BMI and prostate cancer; it is impossible to ascertain the extent of the bias from the AD.

Additionally, there was some evidence to suggest that, on average across studies, as BMI increases, the risk of advanced prostate cancer increases. This association may be differently affected by testing for prostate cancer: the bias may be negative for the same reasons as for prostate cancer, but may be positive as the men with cancer who were not biopsied when testing with PSA due to a lower PSA level may present later with more advanced cancers. Additionally, there may be collider bias from conditioning on prostate cancer, since any unmeasured confounders associated with both prostate cancer and advanced prostate cancer would induce an association between BMI and advanced prostate cancer. However, since the overall association between BMI and prostate cancer appears to be negligible, this is not a primary concern.

4.5.1. Comparison with previous research: BMI-prostate cancer

Meta-analysis	Effect Estimate (95% CI)	Participants (cases)	Restrictions
This analysis	1.00 (0.97 to 1.02)	9,252,407 (162,470)	None
MacInnis (2006) (88)	1.05 (1.01 to 1.08)	2,818,767 (55,521)	None
Renehan (2008) (208)	1.03 (1.00 to 1.07)	3,029,338 (70,421)	Prospective only
Discacciati (2012) (87)	0.94 (0.91 to 0.97)	1,033,009 (19,130)	Prospective, localized PCa only
Hu (2014) (86)	1.15 (0.98 to 1.34)	21,737 (5,989)	Screening only
Zhang (2014) (209)	1.00 (0.95 to 1.06)	2,342,066 (73,851)	Prospective only, RR for obese versus normal weight
Tsilidis (2016) (210)	1.00 (0.97 to 1.03)	3,798,746 (88,632)	None

Six previous meta-analyses have looked at the association between BMI and prostate cancer. A meta-analysis by MacInnis et al. (2006) (88) showed a small positive association between BMI and prostate cancer (OR for a 5 kg/m² increase in BMI = 1.05, 95% CI 1.01 to 1.08), including 56 studies and 2,818,767 participants, with 55,521 men with prostate cancer (2.0%). The slight difference between our and the MacInnis meta-analyses is likely due to the inclusion of newer studies: when we restricted our analysis to studies that were published before 2006, the average random-effects OR for a 5 kg/m² increase in BMI was estimated to be 1.04 (95% CI 1.01 to 1.07), consistent with the MacInnis estimate. In comparison, when we restricted our analysis to studies published in or after 2006, the estimated average random-effects OR for a 5 kg/m² increase in BMI was 0.98 (95% CI 0.96 to 1.00). This difference in estimates by publication year may reflect changing PSA testing practices, which could induce a negative bias into the estimate as BMI lowers PSA levels. Alternatively, the populations studied may have changed over time: the meta-regression showed that the mid-year of recruitment may decrease the effect estimate, but it does not account for differences in populations over time. The MacInnis meta-analysis also showed no difference between cohort studies (BMI measure before outcome) and case-control studies (BMI measured at the same time as outcome).

A meta-analysis by Renehan et al. (2008) (208) of 27 prospective studies of BMI and incidence of prostate cancer included 3,029,338 participants with 70,421 men with prostate cancer (2.3%). The RR for prostate cancer for a 5 kg/m² increase in BMI was reported to be 1.03 (95% CI 1.00 to 1.07), which is consistent with the MacInnis meta-analysis, and written shortly afterwards. It is therefore likely that the slightly higher effect estimate in the Renehan meta-analysis is due to the same differences as those between our and the MacInnis meta-analyses.

A meta-analysis by Discacciati et al. (2012) (87) showed a negative association between BMI and localised prostate cancer risk (OR for a 5 kg/m² increase in BMI = 0.94, 95% CI 0.91 to 0.97), from 12 studies with 1,033,009 participants, with 19,130 men with prostate cancer (1.9%). The included studies were limited to prospective studies and included many of the studies included as “before” studies in this meta-analysis. The focus on localised prostate cancers likely caused the difference in

effect sizes seen between this and our meta-analyses. While there were more localised than advanced cancers in all included studies in the Discacciati et al. meta-analysis, the OR for advanced cancers for a 5 kg/m² increase in BMI in the same meta-analysis was 1.09 (95% CI 1.02 to 1.16), so across all cancers there may have been no association between BMI and prostate cancer. As we did not extract information on localised prostate cancers for our meta-analysis, it is difficult to make further comparisons.

A meta-analysis by Hu et al. (2014) (86) indicated that BMI is positively associated with prostate cancer (OR for a 5 kg/m² increase in BMI = 1.15, 95% CI 0.98 to 1.34), and involved nine studies with 21,737 participants with 5,989 men with prostate cancer (27.6%). All included studies detected prostate cancer through PSA screening programmes; these studies were excluded from our meta-analysis given the previously known association of BMI and PSA and the potential for collider bias from only knowing the prostate cancer status of men with PSA levels above a threshold. In theory, as BMI is associated with a reduction in PSA levels, the effect of screening should be to make BMI appear protective for prostate cancer, so it is unlikely that the focus on screening led to the positive association between BMI and prostate cancer. Another explanation may be due to chance, as the Hu meta-analysis involved a substantially smaller number of studies and participants than in our analysis, leading to an imprecise effect estimate: note that the 95% CI for the Hu OR included the OR for our meta-analysis. Additionally, studies included in the Hu meta-analysis adjusted for prostate volume, which has been associated with both prostate cancer (211) and BMI (203) and so may have introduced collider bias if changes in prostate volume were caused by BMI and prostate cancer. The populations investigated may also have contributed to the differences, with five of the nine (56%) studies in the Hu analysis conducted in Asian populations (212–216), compared to primarily Caucasian populations in our meta-analysis (9% of studies had mostly Asian populations). Although ethnicity did not appear to be associated with effect size in our meta-regression, there may have been insufficient power to detect a difference, or there may be only be an association with ethnicity in screening studies.

A meta-analysis by Zhang et al. (2014) involved 14 prospective studies, with 2,342,066 participants and 73,851 men with prostate cancer (3.2%). The study reported an RR for prostate cancer for obese versus normal weight men of 1.00 (95% CI 0.95 to 1.06), consistent with our finding that the OR for prostate cancer did not change with increasing BMI.

A random-effects dose-response meta-analysis was conducted by Markozannes et al. (2016) (210) using data from the World Cancer Research Fund (WCRF) as part of the continuous update project (38). Markozannes included 39 studies with 3,798,746 participants and 88,632 men with prostate cancer (2.3%) for the association between BMI and prostate cancer (excluding studies on mortality),

including many of the same studies we included in our meta-analysis. The RR for prostate cancer for a 5 kg/m² increase in BMI was 1.00 (95% CI 0.97 to 1.03), entirely consistent with our results.

Finally, an umbrella review of systematic reviews and meta-analysis by Kyrgiou et al. (2017) (85) concluded that there was no strong evidence for an association between BMI and prostate cancer risk, but did not present an effect estimate.

4.5.2. Comparison with previous research: BMI-advanced prostate cancer

Meta-analysis	Effect Estimate (95% CI)	Participants (cases)	Restrictions
This analysis	1.05 (0.99 to 1.10)	1,053,116 (8,361)	None
Discacciati (2012) (87)	1.09 (1.02 to 1.16)	1,080,790 (7,067)	Prospective
Hu (2014) (86)	1.37 (1.19 to 1.57)	15,254 (1,486)	Screening, high-grade PCa only
Markozannes (2016) (210)	1.08 (1.04 to 1.12)	1,676,220 (11,204)	Advanced, high-grade and fatal PCa

There were three previous meta-analysis that estimated the association between BMI and advanced prostate cancer. The Discacciati meta-analysis (87) showed a positive association between BMI and advanced prostate cancer (OR for a 5 kg/m² increase in BMI = 1.09, 95% CI 1.02 to 1.16), and involved 12 prospective studies with 1,080,790 participants with 7,067 men with prostate cancer (0.65%). In our meta-analysis, the average random-effects OR for advanced prostate cancer for a 5 kg/m² increase in BMI in prospective studies was estimated to be 1.05 (95% CI 0.99 to 1.10), and we included approximately the same number of men in total, although a greater proportion of men with prostate cancer: 1,053,116 participants with 8,361 men with advanced prostate cancer (0.79%). We could not include some of the studies in the Discacciati meta-analysis as the studies lacked data to compute an effect estimate for a 5 kg/m² increase in BMI, but these were included in the albatross plot (111,185,190). Notably, the Cerhan (111) and Putnam (185) studies showed very strong positive associations, although their weights were very low in the Discacciati analysis (0.36% each). The MacInnis study (2003) (172) was not included in our meta-analysis because it used data from the Melbourne Collaborative Cohort Study (MCCS), the same as Bassett (2012) (166), which we used instead. The later Bassett study had a smaller OR for advanced prostate cancer than the MacInnis study (1.27 versus 1.51). In addition, we included four studies not included in the Discacciati meta-analysis (114,158,162,174): all these studies had null or negative effect estimates, drawing our estimate closer to the null. Finally, although our study and Discacciati both used the GLST method (**Section 4.3.1.A**) for estimating linear associations, we used a slightly different method to estimate the mean BMI values in each BMI category (**Section 4.3.1.B**), which drew some study estimates closer to the null (see **Section 4.5.5**).

The Hu meta-analysis (86) also showed a positive association between BMI and high-grade prostate cancer (OR for a 5 kg/m² increase in BMI = 1.37, 95% CI 1.19 to 1.57), and involved 8 screening studies

with 15,254 participants with 1,486 men with high-grade prostate cancer (9.7%). High-grade prostate cancer (usually defined as a Gleason score of 8-10) is not equivalent to advanced prostate cancer (usually defined as a high TNM score), and thus this meta-analysis measures a different (but related) outcome to our meta-analysis. It is notable, however, that the OR in the Hu meta-analysis showed such a large positive effect. This effect may be due to the different outcome, but could also be explained by differences in inclusion criteria (screening studies only in the Hu analysis) and therefore differences in the sample populations.

Markozannes also conducted a meta-analysis of the BMI and advanced, high-grade or fatal prostate cancer using WCRF data, which included 23 studies with 1,676,220 participants and 11,204 men with advanced/high-grade/fatal prostate cancer (0.67%) (210). The RR for advanced/high-grade/fatal prostate cancer for a 5 kg/m² increase in BMI was 1.08 (95% CI 1.04 to 1.12), consistent with the Discacciati meta-analysis. The effect estimate may be increased in the WCRF estimate by the inclusion of high-grade and/or fatal prostate cancers, use of RRs rather than ORs, or from methodological differences in calculation of the mean BMI values in each BMI category (as in the Discacciati meta-analysis, see **Section 4.5.5**).

Finally, an umbrella review of systematic reviews and meta-analysis by Kyrgiou et al. (2017) (85) concluded that there was weak evidence for a positive association between increasing BMI and advanced prostate cancer risk.

4.5.3. Comparison with previous research: BMI-PSA

We could only find one previous review of the association between BMI and PSA, which did not include a meta-analysis or estimated effect size (217). Their conclusion was that many studies reported an inverse association between BMI and PSA, in agreement with our findings.

4.5.4. Strengths and limitations

The largest strength of our analysis was the large number of studies and participants included from many different populations at different time points. All pooled effect estimates were precise; the estimated null association of BMI with prostate cancer is statistically very precise, and this appears not to be a lack of evidence for an association but evidence for a lack of association. By including both studies where BMI was measured before and at the same time prostate cancer, we could compare different study types: that there appeared to be little difference between these two study types for all outcomes could be evidence the findings are robust. There was no evidence of small-study effects for in the associations between BMI, prostate cancer and advanced prostate cancer.

However, there are limitations. Many of the studies included in the meta-analysis compared men with a diagnosis of prostate cancer versus men without a diagnosis of prostate cancer. As PSA testing is a common method for detecting prostate cancer, and because there appears to be a negative association between BMI and PSA, men with high BMIs may have had prostate cancer but not received a diagnosis. This would cause a negative association between BMI and prostate cancer risk, which may have been observed in the meta-regression, where mid-year of study was associated with a decrease in the OR for prostate cancer, possibly indicating an increased uptake of PSA testing over time. Bias from testing for prostate cancer with PSA was limited as much as possible from excluding studies that exclusively screened for prostate cancer, but as PSA testing is used in general practice the bias could not be entirely removed. The proportion of prostate cancers detected by testing with PSA likely varied in each study, potentially accounting for some of the residual heterogeneity in studies examining the association between BMI and prostate cancer and advanced prostate cancer. This implies that there is a possible positive association between BMI and prostate cancer risk, which is unobserved because of the unknown amount of bias from testing for prostate cancer with PSA.

Because the studies may not have used the same definition of advanced prostate cancer, and because advanced prostate cancers could be locally advanced prostate cancer, nodes or metastatic cancer, these studies may be relatively heterogeneous. This may have attenuated any association between BMI and advanced prostate cancer.

Heterogeneity was also relatively large across the studies examining the associations between BMI and PSA. This could be partly due to many of studies having a very large number of participants, reducing within-study variation and so leading to high I^2 values (since I^2 is the ratio of variance due to heterogeneity to total variance). There was also some evidence of small-study effects in the funnel plot. However, it is also possible the association between BMI and PSA varies by population, though our meta-regressions did not find any explanatory factors. Additionally, we estimated the *total* effect of BMI on PSA, not accounting for the *indirect* effect through prostate cancer, and so there may be variation from different levels of prostate cancer in each study. Although the association between BMI and prostate cancer appears negligible, the *total* effect of BMI on PSA is likely approximately equal to the *direct* effect.

There was a moderate risk of bias for all studies, as all studies could have been biased by unobserved confounding. We attempted to limit effects of bias due to confounding by identifying key confounders and only including studies without a critical risk of bias. There was also no evidence that the studies with a medium risk of bias had a systematically different effect estimates than those with a high risk of bias. Bias may also have been introduced by the assumptions we made: we assumed BMI had a

multiplicative association with PSA, and thus calculated log-PSA from reported PSA levels; we also assumed that ORs and HRs would be approximately equal. However, the assumption of the multiplicative association has a theoretical justification (in haemodilution), and there was no difference in the subgroup analysis of ORs versus HRs for prostate cancer or advanced prostate cancer.

4.5.5. Methodology

While standard methods were used for most of this meta-analysis, I developed and used albatross plots (**Section 2.3**) to assess the extent to which studies without sufficient information for inclusion in meta-analysis would have affected the results had they been included. The albatross plot for the association of BMI and prostate cancer confirmed that the studies without sufficient data were consistent with those studies with sufficient information, but in both the associations between BMI and PSA and BMI and advanced prostate cancer, there were studies which were inconsistent and required explanations or limitations to the generalisability of the main findings. Overall, both methods allowed the inclusion of more studies, improving precision and identifying outliers, enhancing the interpretability of the results.

Additionally, I used a different method of calculating the mean BMI in subgroups of BMI than Discacciati (87) in their meta-analysis. Discacciati computed the mean BMI in each category of BMI by taking the midpoint between the lower and upper bounds; for open-ended BMI categories (e.g. BMI < 25 kg/m²), they assumed the unknown bound was the same magnitude as the neighbouring categories. We assumed, however, that BMI had an underlying normal distribution, and estimated the mean BMI level in category using the method developed by Chêne and Thompson (132). This likely led to more accurate estimates of the effect of BMI on prostate cancer risk.

For instance, our calculations for the OR for advanced prostate cancer for a 5 kg/m² increase in BMI for Kurahashi (100) gave a smaller OR than MacInnis (1.32 versus 1.54), which a result of how the BMI means are estimated (**Table 4.8**). With the Discacciati method, BMI is assumed to be uniformly distributed, and the open-ended categories are assumed to have the same number of men as the neighbouring categories; neither of these assumptions are likely in practice. Our assumption that BMI is normally distributed is likely to better fit observed BMI values, and this leads to the BMI means for the open-ended categories being more extreme than the Discacciati method. As the BMI means are now further apart across the whole study, the estimated OR for a 5 kg/m² is necessarily smaller.

Table 4.8 Results for Kurahashi et al. (100)

BMI subgroup	BMI mean (our method)	BMI mean (Discacciati method)	HR (95% CI)
<21.9	20.1	21.2	1
22-23.4	22.6	22.7	0.9 (0.49 to 1.67)
23.5-24.9	24.1	24.2	1.33 (0.75 to 2.37)
≥ 25	26.8	25.7	1.38 (0.8 to 2.39)
5 kg/m ² increase	OR = 1.32 (0.88 to 1.98)	OR = 1.54 (0.86 to 2.76)	NA

This method affected all studies where the OR, RR or HR for prostate cancer was presented for different categories of BMI, and so had a small but meaningful effect on the summary estimate for both prostate cancer and advanced prostate cancer. As BMI is a reasonably normally distributed variable, we feel this methodology is more appropriate than assuming a uniform distribution, and improves the accuracy of the individual study estimates.

4.6. Conclusion

In conclusion, the AD evidence suggest that prostate cancer risk is not meaningfully associated with BMI (random-effects OR for a 5 kg/m² increase in BMI = 1.00 95% CI 0.97 to 1.02, P = 0.71), although testing for prostate cancer with PSA may induce a negative bias. The evidence also suggests there is a slight positive non-significant association between BMI and advanced prostate cancer (OR for a 5 kg/m² increase in BMI = 1.05, 95% CI 0.99 to 1.10, P = 0.09), and there is a negative association between BMI and PSA (PSA change for a 5 kg/m² increase in BMI = -5.16%, 95% CI -6.85% to -3.44%, P < 0.001).

This evidence suggests it is unlikely that the association seen between BMI and PSA is mediated through the BMI-prostate cancer and prostate cancer-PSA associations, since there was no evidence for an association between BMI and prostate cancer. However, individual patient data analyses (**Chapters 5 and 6**) may allow further investigation of non-linear associations, better adjustment for confounding, and investigation of missing data.

CHAPTER 5. INDIVIDUAL PARTICIPANT DATA META-ANALYSIS OF ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA

5.1. Aims

The aim of this chapter is to conduct an individual participant data (IPD) meta-analysis to quantify the associations between: (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, (3) BMI and PSA, and (4) prostate cancer and PSA. Using individual participant data (IPD) allows us to investigate these associations more thoroughly than with AD alone as all covariates (that were measured) and missing data can be accounted for, allowing us to control for potential confounding and bias using the same set of covariates in all studies. In addition, non-linearity of the associations and possible interactions can be assessed.

5.2. Background on handling missing data

In **Chapter 4**, the AD studies likely included some men who were tested for prostate cancer using PSA, and so received prostate biopsies conditional on having a high PSA level. Studies which exclusively screened for prostate cancer using PSA were excluded, but as PSA testing is a common method of detecting prostate cancer, bias from PSA testing could remain in the measured association between BMI and prostate cancer risk (**Section 4.5.4**). However, if the screening protocol of a study (or biopsy status of a man) is known and the IPD is available, it is possible to account for the missing prostate cancer data using imputation if the factors influencing selection for biopsy are known and measured. This reduces the risk of bias from screening, but also allows more data to be included to estimate the association between BMI, prostate cancer and advanced prostate cancer.

In this section, I will describe missing data, how imputation deals with missing data, and how imputation can be used with IPD.

5.2.1. Missing data

Data can be missing for many reasons, but missingness mechanisms fall into three distinct categories: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR) (218). Data are MCAR if missingness is not dependent on any observed or unobserved variable. In practice, data are rarely MCAR, although one example would be if the electronic scales short-circuited when trying to measure weight to calculate BMI. Data are MAR if missingness is dependent only on observed variables. For example, cancer status would be MAR if a PSA threshold was used to determine who is offered a prostate biopsy for cancer, and all men had PSA measured. Data are MNAR if they are not MCAR or MAR. For example, prostate cancer status would be MNAR if men of higher socio-economic status (SES) were both more likely to seek screening for prostate cancer with PSA and less likely to actually have prostate cancer, but data on SES were not available. If data are MNAR, there is no viable universal method of accounting for the missing data that does not involve extra assumptions, and sensitivity analyses are recommended (219). There is no way to distinguish between MAR and MNAR given only the observed data.

One way to analyse data where some data is missing is to perform a complete case analysis, where any observations with missing data are excluded (220). If data are MCAR, a complete case analysis will be unbiased (218), although the precision will be reduced. If data are MAR, then a complete case analysis may be biased (221). Therefore, we excluded screening studies in our AD meta-analysis examining the association between BMI and prostate cancer (**Chapter 4**). However, so long as the data

are MAR, then imputation can be used to remove the bias. Additionally, imputation can improve the power of an analysis as fewer participants need to be excluded from the study due to missing data.

5.2.2. Multiple Imputation

Imputation is replacement of missing data using information from the participants with data.

Multiple imputation (MI) generates many imputed datasets, where the missing values are repeatedly estimated with appropriate error (222,223). Each dataset is then analysed separately, and combined to give a single result, for example using Rubin's rules (224). Rubin's rules simply average the coefficients from the analysis, and compute the variance of the combined estimate as shown:

$$\bar{\beta} = \frac{\sum_{i=1}^M \beta_i}{M}$$

$$Var_{within} = \frac{\sum_{i=1}^M Var_i}{M}$$

$$Var_{between} = \frac{\sum_{i=1}^M (\beta_i - \bar{\beta})^2}{M - 1}$$

$$Var_{total} = Var_{within} + Var_{between} + \frac{Var_{between}}{M}$$

where for each imputed dataset $i = 1, \dots, M$, β_i is the effect estimate for the exposure on the outcome, $\bar{\beta}$ is the combined effect estimate across imputed datasets, Var_i is the variance of the effect estimate, and Var_{total} is the total variance of the effect estimate across all imputations.

MI with Multiple Imputation by Chained Equations (MICE) is commonly used (225–229). Briefly, missing values are filled in M times using a series of equations relating variables to one another to generate M complete datasets, which are analysed separately and then combined as above to give an overall estimate (230). Variables can be continuous or non-continuous (227). There is no universal criteria for how many datasets need to be created (standard texts on multiple imputation suggest 3 or 5 datasets is adequate), although White (2011) suggests that the number of datasets should be more than 100 times the fraction of missing information (225).

5.2.3. Multiple imputation in IPD meta-analysis

When using multiple datasets from different studies, MI can either be performed separately in each study using a within-study model, or across studies using a stratified model (231). Assuming an analysis with two continuous exposures (one complete, the second incomplete) and one complete continuous

outcome, the within-study imputation method imputes missing values of the second exposure variable using the following model:

$$x_{2is} = \alpha_{0s} + \alpha_{1s}x_{1is} + \alpha_{2s}y_{is} + \epsilon_{is}$$

$$\epsilon_{is} \sim N(0, \sigma_s^2)$$

where for each participant i and study s , x_1 and x_2 are both exposure variables (complete and non-complete respectively), y is the outcome, and ϵ_{is} is the residual error variance, which can differ between studies. The α_{1s} and α_{2s} parameters are estimated separately in each study, as is the residual error variance. No information is shared between studies in the imputation: as such, there is no way to impute a variable that is completely missing in a dataset.

In the (homoscedastic) stratified imputation method, the following model is used to impute missing data:

$$x_{2is} = \alpha_{0s} + \alpha_1x_{1is} + \alpha_2y_{is} + \epsilon_{is}$$

$$\epsilon_{is} \sim N(0, \sigma^2)$$

In this model, α_1 and α_2 are assumed to be identical across all studies and the error distribution is assumed to be homogeneous across studies, but with fixed study-level intercepts. Because information is shared between studies, variables that are completely missing in one dataset can still be imputed, assuming there are no systematic, unmeasured differences between the study where x_2 was completely missing and those where it was measured. The study-level intercepts cannot be estimated directly for completely missing variables and will be fixed at a value (e.g. 0), although in many analyses this will only affect the intercept term, not the parameter of interest (231). The model is homoscedastic because the variance of the error distribution is assumed to be homogeneous across studies; a heteroscedastic model could be used, where the variance of the error distribution is study-specific, although this would again make it difficult to impute variables that are completely missing in a study, as it is unclear what value the variance should take (231).

Once imputed, both Rubin's rules and meta-analysis must be performed to give an overall effect estimate. Rubin's rules combine imputations, whereas meta-analyses combine studies, and either can be applied first. When Rubin's rules are applied first, the result from each study is estimated across all imputations (giving as many results as studies), then all studies are combined in meta-analysis. When Rubin's rules are applied second, each imputed dataset is meta-analysed separately (giving as many results as imputed datasets), then the results are combined. Burgess (2013) recommends using

Rubin's rules followed by meta-analysis (231), which has the advantage of allowing forest plots to be produced showing the results from each study.

5.2.4. Imputation with non-linear associations

All variables included in the analysis must also be included in the imputation model, to prevent bias (225). More generally, the analysis method must be congenial with the imputation method, i.e. all the assumptions made by the analysis model must also be made by the imputation model(s), and vice versa. If the association between two variables is non-linear in the analysis model, then the imputation model(s) will have to include the non-linear terms. For instance, to estimate a non-linear association between BMI and log-PSA, BMI, BMI^2 and BMI^3 may all be variables in the regression model. These terms must then be included in the imputation model for any missing PSA values. However, if BMI is also incomplete, then BMI, BMI^2 and BMI^3 will all need imputing (from models using the remaining data). As the non-linear terms are entirely derived from the linear term, it is problematic to include them in the imputation – e.g. to include BMI in the model to impute BMI^2 .

One approach to this is to passively impute the non-linear terms, so that only the linear term is imputed and the non-linear terms are calculated from this at each imputation. However, only including the linear term in the imputation model will discount the non-linear association in the imputation; it will still be possible to analyse the data using non-linear terms, but the imputation will be less precise (225). A different approach is to treat each non-linear term as “just another variable” (225), and impute them all separately. This means the values in a single imputation may not be mathematically related as desired, e.g. a quadratic term may not be equal to the linear term squared. However, this approach has been shown to perform well with minimal bias in the analysis model (232), and will be used in this chapter.

Another approach is to not assume any particular association (e.g. linear, quadratic, cubic) but instead to use in both analysis and imputation model(s) categories for the exposure. For BMI, the normal weight ($BMI < 25 \text{ kg/m}^2$), overweight ($25 \leq BMI < 30 \text{ kg/m}^2$) and obese ($BMI \geq 30 \text{ kg/m}^2$) categories could be used. This obviates the need to treat non-linear terms as “just another variable”, but categorisation generally loses information, i.e. men with BMIs of 20 kg/m^2 and 24 kg/m^2 would be treated identically.

5.3. Background on Meta-analysis in IPD

IPD meta-analysis can be performed in one or two stages, although generally, one- and two-stage IPD meta-analyses give very similar results (233,234). In two-stage IPD meta-analysis, the first stage is to

analyse each study separately to estimate coefficients and standard errors for the association of interest, and the second stage is to bring together these results using standard meta-analysis methods (**Section 2.3.2**). In contrast, a one-stage meta-analysis combines all IPD in a single meta-analysis based on a regression model (e.g. linear, logistic) stratified by trial (235). In order to incorporate random-effects to allow for heterogeneity between studies, hierarchical or mixed-effect regression models are used (236,237). Homoscedastic stratified imputation is congenial with one-stage meta-analysis, and also can be used with two-stage fixed-effect and random-effects meta-analyses (238). In this chapter, I will use both one-stage and two-stage meta-analyses to check the results are consistent.

5.4. Methods

In this section, I describe the studies for which we requested IPD, how the data were cleaned, the specific imputation methods used in this analysis and how we assessed whether the assumptions used in the imputation were valid, how the potential interaction between age and BMI was assessed, and how the linear and non-linear effects were estimated.

5.4.1. Studies

Five prospective studies looking at prostate cancer were approached to provide IPD for this analysis (28,30,31,239,240). These studies were chosen because they were large studies of prostate cancer with known screening protocols (or the biopsy status of each man was known) and many variables measured. All studies except the Krimpen study would have been excluded from our AD meta-analysis (since they screened for prostate cancer using PSA), and no papers using data from the Krimpen study were found in the AD meta-analysis. Therefore, there is no overlap in studies between this chapter and **Chapter 4**. For all studies, we used PSA and BMI data measured at baseline, and information on whether a man was diagnosed with prostate cancer at any time-point during the study; we did not record how long until the time of diagnosis, and as such did not estimate hazard ratios (HRs).

For every IPD study, we requested data measured at baseline on BMI and PSA, as well as age, family history of prostate cancer and ethnicity. We also requested data on prostate cancer status (including TNM [tumour, node, metastases] and Gleason scores). The five studies are described briefly below.

European Randomised Study of Screening for Prostate Cancer: Rotterdam

The European Randomised Study of Screening for Prostate Cancer (ERSPC) (240) was a multicentre, randomised trial assessing PSA testing in eight European countries. Eligible men aged 50–74 years were identified from population registries and randomly assigned to screening or no intervention (control); only data from the screening arm were used in this chapter as the non-screening arm did not have baseline PSA levels. The primary outcome was prostate cancer mortality. For this analysis, only the screened arm of the Dutch cohort (Rotterdam) was included due to difficulties accessing the full dataset. From October 1991 to December 2000, men were invited to enter the study, and men randomised to the screening arm underwent three tests: digital rectal exam (DRE); transrectal ultrasonography (TRUS); and a PSA test. Up to November 1997, men were indicated for biopsy if their PSA level was above 4.0 ng/ml or either of the other two tests were positive. After November 1997, all men with a PSA above 3.0 ng/ml were indicated for prostate biopsy. Based on these criteria, we assumed 17% of men in ERSPC-Rotterdam were biopsied. Men were invited for a second biopsy after four years, but only the PSA values, age, family history of prostate cancer and prostate cancer

diagnoses from the initial round of screening were used in this chapter. BMI was not measured in this cohort, and had to be imputed for our analyses, as did prostate cancer status for all men not biopsied. Ethnicity was not measured in this cohort, and we assumed all participants were white – in 1993, only about 9% of Dutch people were not Caucasian (241).

Krimpen

In the Krimpen study (239), men aged 50–75 years were recruited from all general practices in Krimpen aan den IJssel, the Netherlands, between August 1995 and January 1998 in a longitudinal study to determine the prevalence of BPH. Men were not screened for prostate cancer at baseline, but may have had PSA tests as part of their general care. Prostate cancer status was recorded for those men who received biopsies (26% of all men), and those with prostate cancer at the beginning of the study were excluded. The last prostate cancer diagnosis was made in 2004. Data for PSA, age, family history of prostate cancer and BMI were recorded at baseline and three follow-up rounds; in this chapter, only the baseline measures were used. Ethnicity was not recorded, although the study investigators informed us that more than 90% of men in the Krimpen study were Caucasian; as with ERSPC-Rotterdam, we thus assumed that all men in Krimpen were of white ethnicity.

Prostate Cancer Prevention Trial

In the Prostate Cancer Prevention Trial (PCPT) (28), from January 1994 through May 1997, men with a PSA level under 3.0 ng/ml underwent randomisation to either finasteride (a 5 α reductase inhibitor) or placebo in the USA. The men underwent annual DREs and measurement of PSA. At the end of 7 years, all the men in whom prostate cancer had not been diagnosed were offered an end-of-study biopsy; in total, 61% of men received a prostate biopsy over the course of the study. Data collection, including BMI, PSA, age, ethnicity and family history of prostate cancer, and prostate-cancer assessments continued until June 2004. In this chapter, PSA, age, family history of prostate cancer and BMI from baseline were used, and prostate cancer diagnosis on follow-up.

Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial

In the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial (30), men aged 55–74 years were enrolled at 10 screening centres in the USA between November 1993 and July 2001. Men were randomly assigned to the intervention – organised screening of annual PSA testing for 6 years and annual DRE for 4 years – or usual care, which sometimes included opportunistic screening. The screening test was positive if the man had a PSA above 4.0 ng/ml or if the DRE was positive, although the decision to offer a biopsy was made by local health-care providers. Screening was completed in October 2006. All incident prostate cancers through 13 years of follow-up or through to the end of 2009 were ascertained. Given the PSA threshold and number of men with prostate cancer, we

assumed 24% of men received a prostate biopsy. Only participants in the screening arm of PLCO had a recorded PSA test (taken at screening), so only the screening arm was included here. In this chapter, PSA, age, family history of prostate cancer and BMI from baseline were used, and prostate cancer diagnosis on follow-up.

Prostate Testing for Cancer and Treatment

In the Prostate Testing for Cancer and Treatment (ProtecT) study (31), PSA screening was conducted between 1999 and 2009. Men with a PSA above 3.0 ng/ml and below 19.9 ng/ml were invited to biopsy, while men with a PSA above 20 ng/ml were referred to usual care. Given the PSA threshold, we assumed 15% of men received prostate biopsies. Participants with localised prostate cancer were offered randomisation to one of three treatments: prostatectomy, radical radiotherapy or active monitoring. BMI was only available for a subset of the participants included in an accompanying study, thus BMI had to be imputed for the remaining men. In this chapter, PSA, age, family history of prostate cancer and BMI from baseline, and only prostate cancers found after initial biopsy were used.

5.4.2. Risk of bias assessment

The risk of bias assessment used in the AD meta-analysis (**Appendix 5**) was used to assess the risk of bias for the IPD studies. Overall, none of the studies except PCPT were assessed to be at high risk of bias due to: confounding (since ethnicity and age could be controlled); selection of participants; missing data (we assumed prostate cancer status to be MAR given PSA levels, and other variables with missing data to be MCAR); measurement of outcome (as the screening study protocols were defined, and biopsy status was known in the Krimpen study); or measurement of exposure. Bias due to selective reporting was not applicable here. Therefore, overall all studies except for PCPT were classed as a medium risk of bias, since in terms of the associations between BMI, prostate cancer, advanced prostate cancer and PSA these were all observational studies.

PCPT represents an important source of information for our study: because men were offered a prostate biopsy independent of their PSA, PCPT has less potential for bias in the estimate of the association between BMI and PSA for men with an initial PSA below 3.0 ng/ml. However, PCPT can be expected to estimate each of the overall associations between BMI, prostate cancer and PSA with bias, because only men with a PSA below 3.0 ng/ml were included. In measuring the association between BMI and prostate cancer, conditioning on PSA will lead to collider bias. In measuring the association between BMI and PSA, and prostate cancer and PSA, conditioning on the outcome will attenuate the association. Therefore, PCPT was only used for the imputation of prostate cancer status (and other missing variables) in the four other studies, and not included in the analysis.

5.4.3. Data cleaning

Data from all studies were imported into Stata for analysis. Age and BMI at baseline were calculated from available data if not provided. PSA was logarithmically transformed for analysis due to its log-normal distribution and our assumption that due to haemodilution, any change in BMI would give a percentage change in PSA, not an absolute change (**Section 4.3.3**). For 33 men where PSA was recorded as 0.00 ng/ml, and this was recoded as 0.02 ng/ml, a standard amount used when PSA is undetectable (242). Although this could have been a missing PSA value, it could also be for men who naturally do not have a detectable PSA level from genetic mutations (32), or from misinterpretation of a PSA level of <0.02 ng/ml. Advanced prostate cancer was defined as having locally advanced disease (a T-stage 3 or above, $T \geq 3$), nodal involvement ($N=1$) or metastases ($M=1$) (243) (TNM described in **Section 1.2.2**). Non-white participants were excluded from the dataset ($N = 5,838$), as there were relatively few participants that were in this group, making imputation in these groups unfeasible.

5.4.4. Exploratory analysis

In advance of imputation, I conducted a complete case analysis assuming men without a biopsy did not have prostate cancer, which approximates the AD studies where PSA testing may have occurred and the comparison groups were men without a diagnosis of prostate cancer. This was to assess the potential effect of screening on all results, which could have affected the AD study results, and to check the imputed results are plausible given the complete case results and the expected bias from screening for prostate cancer using PSA.

I also explored heterogeneity between studies and checked for interactions between age and BMI for all associations, excluding men without a biopsy from these analyses to reduce the risk of bias from screening. These analyses were to check whether the assumptions of study-specific intercepts but shared associations between the key variables of interest would be valid for the imputation and analysis models. Additionally, we wanted to check whether an interaction term between age and BMI was necessary in the imputation and analysis models.

Complete case analysis

For the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer and (3) BMI and PSA, I conducted complete case analyses without imputation of prostate cancer status. The complete case analyses assumed prostate cancer status was known for all men regardless of biopsy status, i.e. we assume that all men with no diagnosis of prostate cancer and who did not undergo a biopsy do not have prostate cancer. This is the same assumption as in the AD studies where

men without prostate cancer were defined as not having a prostate cancer diagnosis, rather than a negative prostate biopsy. Both two-stage IPD random-effects and fixed-effect meta-analysis results were presented in forest plots.

Checking for heterogeneity

I compared the observed associations between (1) BMI and prostate cancer risk, (2) BMI and log-PSA, and (3) log-PSA and prostate cancer risk for men with complete information on these variables in the IPD studies. Men recorded or assumed not to have had a biopsy were excluded, as were any men missing values for the exposure or outcome variables. For the association between BMI and prostate cancer, for each study I regressed prostate cancer status against BMI using logistic regression and cubic splines with knots at quartiles of BMI across all participants (24.8 kg/m^2 , 27 kg/m^2 and 30 kg/m^2), with age and family history of prostate cancer as covariables. I plotted fitted values of prostate cancer risk against BMI for each study (at age 60 years and no family history of prostate cancer) for BMI values between 20 kg/m^2 to 40 kg/m^2 . For the association between BMI and log-PSA, I regressed log-PSA against BMI using cubic splines, with age and prostate cancer status as covariables. I plotted fitted values of log-PSA against BMI for each study (at age 60 years and no prostate cancer) for BMI values between 20 kg/m^2 to 40 kg/m^2 . As ERSPC-Rotterdam did not measure BMI, it could not be included in these two graphs.

To check the assumption of study-specific intercepts but shared effect estimates was valid for the model imputing prostate cancer status, I regressed prostate cancer status against log-PSA using cubic splines with knots at rough log-PSA quartiles (-0.5, 0 and 0.5), with age and BMI as covariables. I plotted fitted values of prostate cancer risk against log-PSA for each study (at age 60 years and a BMI of 25 kg/m^2) for log-PSA values between -1.6 and 2.4 (PSA values between 0.2 and 11.0 ng/ml). I limited the results for ERSPC-Rotterdam, PLCO and ProtecT to above the study-specific PSA thresholds for biopsy.

For all graphs, we compared the lines for each study to determine whether the assumption of study-specific intercepts but shared effect estimates was valid. We considered the assumptions valid if there were no substantial differences between the slopes of studies. We did not conduct formal tests of heterogeneity using pooled data.

Checking for interactions

I assessed whether an interaction between age and BMI should be included in the imputation and analysis models for the associations between BMI and prostate cancer, and BMI and PSA. To do this, I regressed prostate cancer status against age, BMI, family history, a dummy variable for study and an

interaction term for age and BMI. A second regression replaced prostate cancer with advanced prostate cancer as the outcome. For PSA, I regressed log-PSA against age, BMI, prostate cancer status, a dummy variable for study and an interaction term for age and BMI. In all regressions, I excluded men recorded or assumed not to have had a biopsy, or missing values for the exposure or outcome variables. For all outcomes, we assessed the effect estimate for the interaction term to determine whether an interaction term would be necessary in the imputation and analysis.

5.4.5. Imputation

To remove the potential for bias from screening for all associations, I imputed prostate cancer status for all men assumed not to have received a biopsy. In addition, some men were missing data for BMI (including all the men in ERSPC-Rotterdam), age, family history of prostate cancer or PSA, all of which were also imputed. We assumed prostate cancer status was MAR, conditional only on PSA levels, while the other variables were assumed MCAR. As MI uses all variables in the model to predict all missing values, the imputation was unbiased even if any missingness is related to any of the variables in the model: age, BMI, log-PSA, family history of prostate cancer, prostate cancer status.

For each study where biopsy status was not recorded, we determined whether participants likely received biopsies. The Krimpen study recorded the biopsy status for all participants, so no assumptions were necessary. For ERSPC-Rotterdam, we assumed all men with PSA above 4.0 received a biopsy, and all men with a PSA above 3.0 ng/ml if the biopsy occurred in November 1997 or later, as well as men with a diagnosis of prostate cancer. For PLCO, we assumed participants received a biopsy if any recorded PSA measurement was above 4.0 ng/ml (the study threshold) or if they received a diagnosis of prostate cancer. For ProtecT, we assumed participants received a biopsy if their PSA was above 3.0 ng/ml (the study threshold) or if they received a diagnosis of prostate cancer. For PCPT, we assumed all participants received a biopsy; all participants without prostate cancer were invited to receive a biopsy at the end of the study regardless of PSA level. Although not all participants will have opted for a biopsy, we assumed that any missingness was MCAR, or at least not associated with any of the variables in our analysis. Therefore, assuming all participants had a biopsy should not bias any of the associations between BMI, prostate cancer and PSA, and therefore not bias the imputation model. For all men, if a biopsy was assumed to have been received and there was no diagnosis of prostate cancer, we assumed the man did not have prostate cancer. Prostate cancer status was treated as missing and therefore imputed if the man was not assumed to have had a prostate biopsy.

MICE was used to impute prostate cancer status, family history of prostate cancer, log-PSA, BMI and age where missing. The MI impute package in Stata was used for the imputation (244). The

homoscedastic stratified method of imputation was used, allowing for study-specific intercepts but shared effect estimates. Within-study imputation was not used, since this would prevent the unbiased estimation of prostate cancer risk in PCPT from informing the missing data from the screening studies. **Table 5.1** shows the full imputation models for all variables. In total, each imputed dataset was the product of 1000 iterations, and 100 new datasets were created for each model. Rubin's rules were applied before meta-analysis. Advanced prostate cancer status was imputed in a separate model in the same way as above, but replacing prostate cancer with advanced prostate cancer.

Table 5.1 Imputation models

Variable being imputed	Regression model	Variables included in regression
Age	Linear regression	BMI, log-PSA, family history, prostate cancer
BMI	Linear regression	Age, log-PSA, family history, prostate cancer
Log-PSA	Linear regression	Age, BMI, family history, prostate cancer
Family history	Logistic regression	Age, BMI, log-PSA, prostate cancer
Prostate cancer	Logistic regression	Age, BMI, log-PSA, family history
Advanced prostate cancer	Logistic regression	Age, BMI, log-PSA, family history

5.4.6. Linear Analysis

In all the analyses in this section, the associations between BMI and all outcomes, and prostate cancer and PSA, were assumed to be linear, giving results that could be compared and combined with the AD meta-analysis. The main analyses were:

1. The association between BMI and prostate cancer risk
2. The association between BMI and advanced prostate cancer risk
3. The *direct* and *total* associations between BMI and log-PSA
4. The association between prostate cancer and log-PSA

All analyses were conducted using the imputed datasets, using a two-stage IPD fixed-effect meta-analysis model, after applying Rubin's rules. To check for consistency, one-stage IPD fixed-effect and two-stage random-effects meta-analyses were also conducted. I conducted all analyses in Stata using the **mi estimate** command for analysing imputed datasets, and for the one-stage fixed-effect meta-analysis I used either the **regress** (continuous) or **logit** (binary) commands, treating study as a factor variable to allow for study-specific intercepts. For the two-stage meta-analyses, I used the **metan** command after estimating the effects in each study.

The PCPT effect estimate was compared with the other studies to observe the effect of restricting the analysis by PSA, but was not included in any meta-analyses due to bias from restricting PSA to less than 3.0 ng/ml. Forest plots were produced for each analysis showing the results for the two-stage IPD random-effects and fixed-effect meta-analyses, and all results including BMI were displayed for a

5 kg/m^2 increase for clarity. Results from one-stage IPD fixed-effect meta-analyses were reported in tables. Details of the regression models for all analyses are summarised below.

Analysis 1: BMI and prostate cancer

Prostate cancer was regressed against BMI, with age and family history as covariates using logistic regression for each study. The coefficient and SE for BMI were multiplied by 5 for each study, then exponentiated to give the change in prostate cancer odds for a 5 kg/m^2 increase in BMI. PSA was not included in the prostate cancer analysis, as we assumed that both BMI and prostate cancer affect PSA, meaning including PSA as a covariate would cause collider bias (see **Figure 3.4**).

Analysis 2: BMI and advanced prostate cancer

Advanced prostate cancer was analysed identically to prostate cancer, using logistic regression and the imputed advanced prostate cancer dataset.

Analysis 3: BMI and PSA

For the *direct* multiplicative effect estimate for BMI and PSA, log-PSA was regressed against BMI, with age, family history and prostate cancer as covariates using linear regression for each study. The coefficient and SE for BMI were then multiplied by 5 for each study, then exponentiated to give the percentage change in PSA for a 5 kg/m^2 increase in BMI. These analyses were repeated without prostate cancer as a covariate to estimate the *total* multiplicative effect estimate for BMI and PSA.

Analysis 4: Prostate cancer and PSA

Log-PSA was regressed against prostate cancer, with age, BMI and family history as covariates using linear regression for each study. The coefficients and CI limits from each study were exponentiated to give the percentage change in PSA from prostate cancer. We assumed a multiplicative model for the practical reason that log-PSA was the outcome in the regression with BMI, and while prostate cancer could cause an additive increase in PSA, we thought a multiplicative increase at least as likely.

5.4.7. Non-linear Analysis

In additional analyses, we allowed BMI to have a cubic relationship with each of prostate cancer and log-PSA. Splines were not used in the imputation or analysis models as 1) linear splines are biologically implausible, 2) cubic splines are relatively uninterpretable, and 3) a cubic model appeared to fit the associations seen in the complete case analysis, and would be more parsimonious than a spline model. Because the imputation and analysis models should be congenial, new imputed datasets were created including the non-linear terms. The imputation equations were the same as for the linear model, but with the addition of BMI^2 and BMI^3 , which were both treated as “just another variable” (**Section 5.2.4**).

All the linear analyses involving BMI were repeated using BMI^2 and BMI^3 as additional variables in the regressions, however only one-stage fixed-effect models were conducted, rather than two-stage or random-effects models.

As all three BMI variables (BMI , BMI^2 and BMI^3) have coefficients in each analysis, forest plots could not capture the association between BMI and each outcome and would be uninterpretable, as each variable depends on the others. As such, instead of forest plots, line graphs for each outcome were produced, displaying the expected log-PSA, prostate cancer risk and advanced prostate cancer risk from the summary effect estimate, for BMI values of 20 and 40 kg/m^2 for a man of 60 years and no family history of prostate cancer. For the association between BMI and log-PSA, we plotted this curve for a man without prostate cancer: given the analysis model, the line graph would be the same shape at a different baseline level for a man with prostate cancer, since we assumed no interaction between BMI and prostate cancer. These graphs were produced using the **nlcom** (non-linear combination of estimators) command in Stata using the results from one-stage IPD fixed-effect meta-analysis, which accounts for the covariance between estimates from a regression model.

In addition to the imputation and analysis where BMI was treated as a cubic variable, we performed additional analyses in which BMI was imputed and analysed in categories (normal weight, overweight and obese). The only difference in the imputation was that the categorical BMI was estimated using ordinal logistic regression, rather than linear regression. The results of these analyses are reported in **Chapter 6** when combined with AD, as we felt non-linear regression with cubic terms was sufficient to analyse non-linear associations.

5.5. Results

5.5.1. Exploratory analyses

Discussion of the complete case analysis results, and their comparison with imputed results, are in the discussion of this chapter (**Section 5.6**).

The summary statistics of men in the original (non-imputed) IPD studies are shown in **Table 5.2**. Across all studies, the average age was 62.0 years (SD: 5.7 years), the average BMI was 27.6 kg/m² (SD: 4.1 kg/m²) and the average log-PSA was 0.15 (SD: 0.87), for which the equivalent PSA is 1.16 ng/ml.

Table 5.2 Summary of included studies for original data

	ERSPC-Rotterdam	Krimpen	PLCO	ProtecT	Total	PCPT
Participants	19,970	1,661	33,025	41,412	96,068	17,815
PCa (%)	1,013 (5.1)	58 (3.5)	3,890 (11.8)	2,292 (5.5)	7,253 (7.5)	2,254 (12.7)
Advanced PCa* (%)	227 (1.1)	7 (0.4)	488 (1.5)	221 (0.5)	943 (1.0)	49 (0.3)
Age (mean, [SD])	63.6 (5.6)	61.5 (6.6)	62.7 (5.3)	60.0 (5.5)	61.7 (5.7)	63.3 (5.7)
BMI (mean, [SD])	NA	26.0 (2.9)	27.6 (4.2)	27.5 (4.0)	27.5 (4.1)	27.7 (4.1)
Log-PSA (mean, [SD])	0.33 (0.93)	0.27 (0.89)	0.16 (0.88)	0.10 (0.91)	0.17 (0.91)	0.03 (0.63)
Family history PCa (%)	1,427 (7.1)	133 (9.2)	2,482 (7.7)	2,076 (5.7)	6,118 (6.8)	2,768 (15.5)

*Advanced prostate cancer defined as a T-score 3 or above ($T \geq 3$), nodal involvement ($N=1$) or metastases ($M=1$)

PCa = prostate cancer, ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

Complete case analysis 1: BMI and prostate cancer

The average OR for prostate cancer for a 5 kg/m^2 increase in BMI using two-stage IPD random-effects meta-analysis was estimated to be 0.94 (95% CI 0.91 to 0.98), **Figure 5.1**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0\%$, $P = 0.72$).

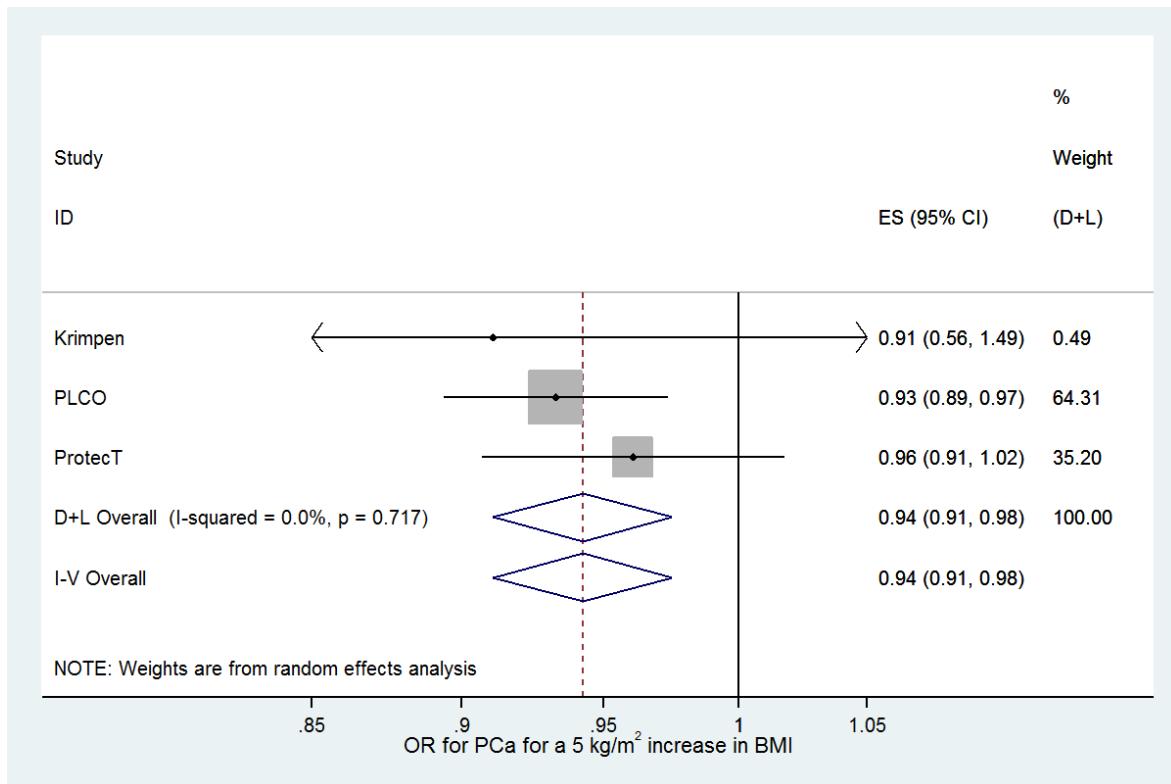


Figure 5.1 Forest plot of BMI and prostate cancer (complete case analysis), showing the OR for prostate cancer for a 5 kg/m^2 increase in BMI

Complete case analysis 2: BMI and advanced prostate cancer

The average OR for advanced prostate cancer for a 5 kg/m^2 increase in BMI using two-stage IPD random-effects meta-analysis was estimated to be 0.98 (95% CI 0.89 to 1.08), **Figure 5.2**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0\%$, $P = 0.94$).

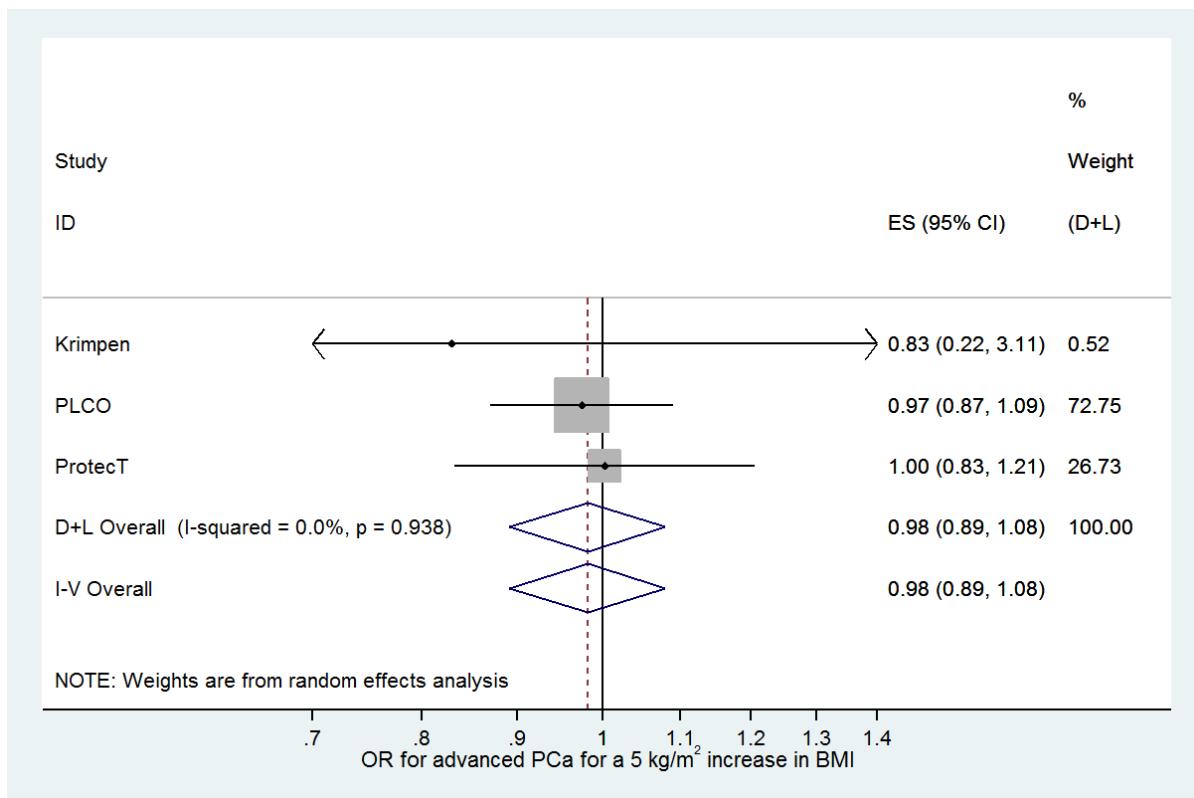


Figure 5.2 Forest plot of BMI and advanced prostate cancer (complete case analysis), showing the OR for advanced prostate cancer for a 5 kg/m^2 increase in BMI

Complete case analysis 3: BMI and PSA

The average *direct* effect estimate for PSA for a 5 kg/m² increase in BMI using two-stage IPD random-effects meta-analysis was estimated to be -6.15% (95% CI -6.82% to -5.48%), **Figure 5.3**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0\%$, $P = 0.41$).

The average *total* effect estimate for PSA for a 5 kg/m² increase in BMI using two-stage IPD random-effects meta-analysis was estimated to be -6.66% (95% CI -7.59% to -5.73%), **Figure 5.4**. There was little evidence of inconsistency in the effect estimates across studies ($I^2 = 24.1\%$, $P = 0.27$).

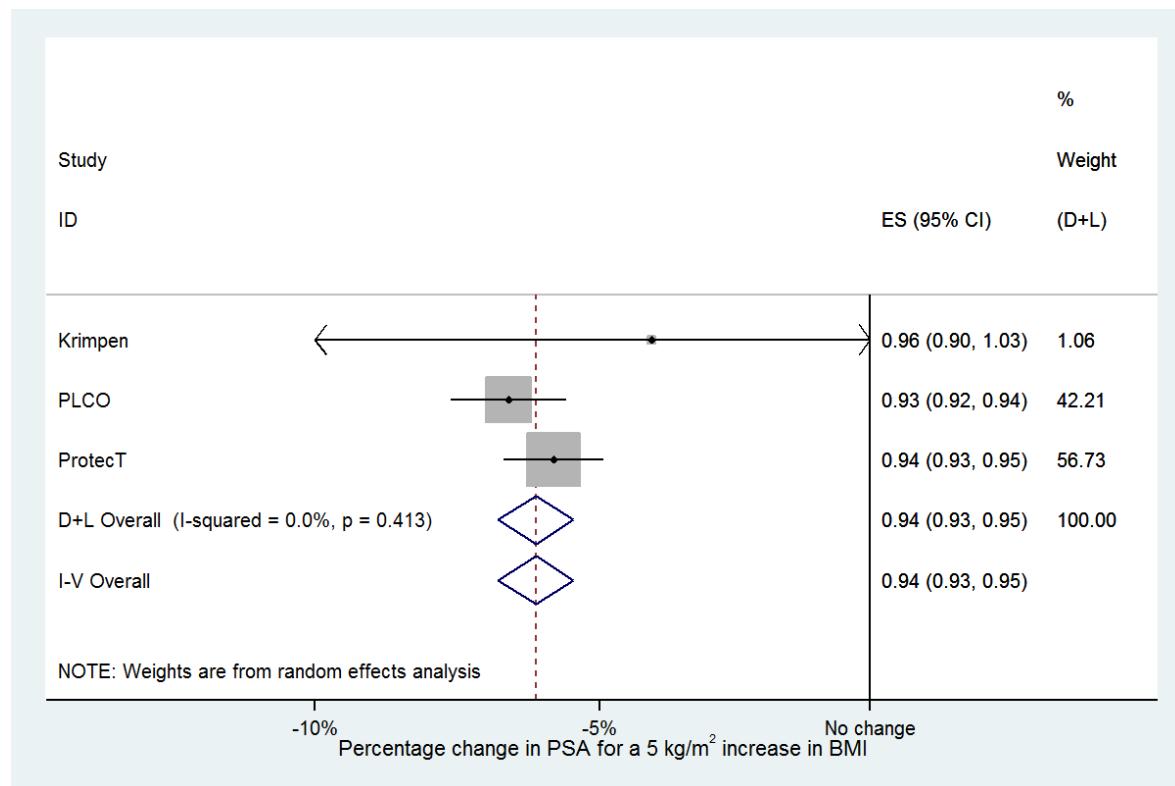


Figure 5.3 Forest plot of BMI and PSA (direct effect, complete case analysis), showing the percentage change in PSA for a 5 kg/m² increase in BMI

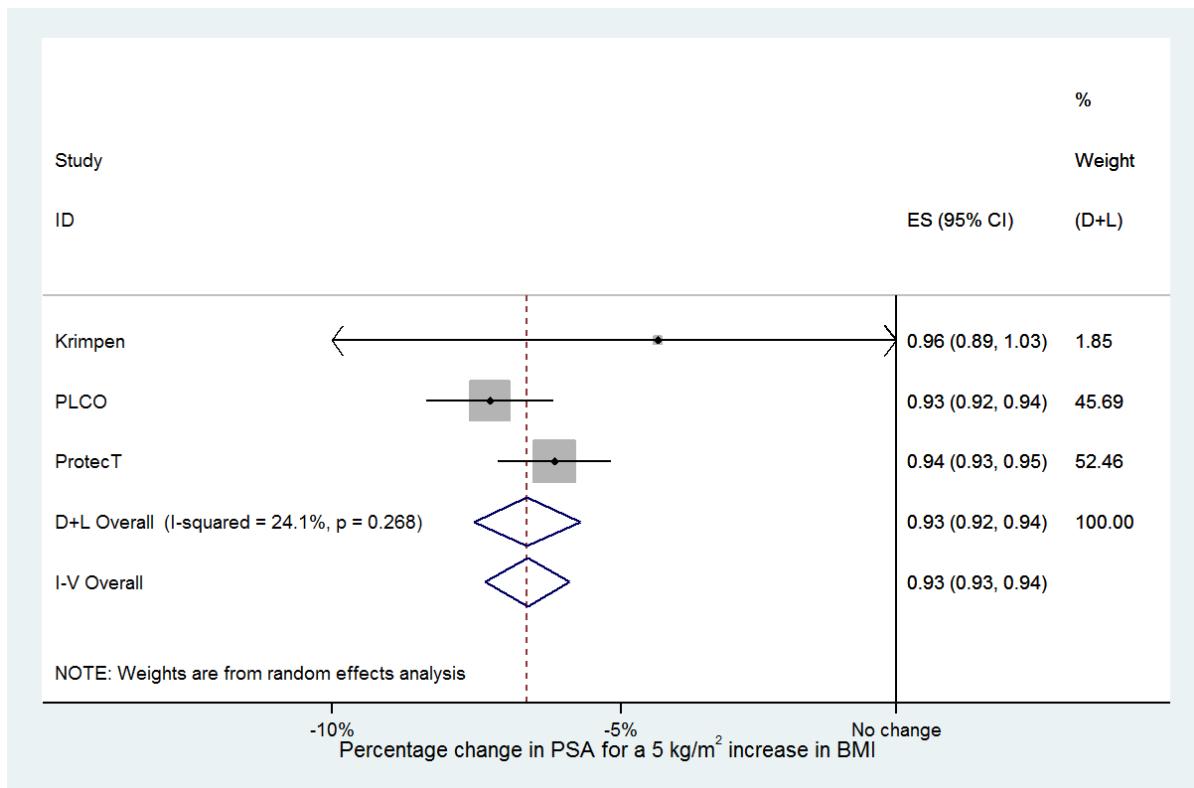


Figure 5.4 Forest plot of BMI and PSA (total effect, complete case analysis), showing the percentage change in PSA for a 5 kg/m² increase in BMI

Checking for heterogeneity 1: BMI and PSA

The line graph of log-PSA against BMI, using restricted cubic splines of BMI, is shown in **Figure 5.5**. The slope of the curve is seen to be similar across PCPT, PLCO and ProtecT: as BMI increases, log-PSA first increases then decreases, albeit with different baseline levels of log-PSA in each study. The lines remain relatively parallel until around 35 kg/m^2 , where PLCO starts to converge with PCPT. However, as most men have a BMI below this amount, we considered the assumption of study-specific intercepts but fixed-effect estimates to be reasonable for the association between BMI and log-PSA for the majority of men. The Krimpen study estimated a more variable association, which is likely explained by the low number of participants overall, potentially allowing small numbers of men with undiagnosed prostate cancer to have a large influence on log-PSA. The variability also helps to justify the use of cubic terms rather than splines in the non-linear analysis, since the association may be overfitted with splines.

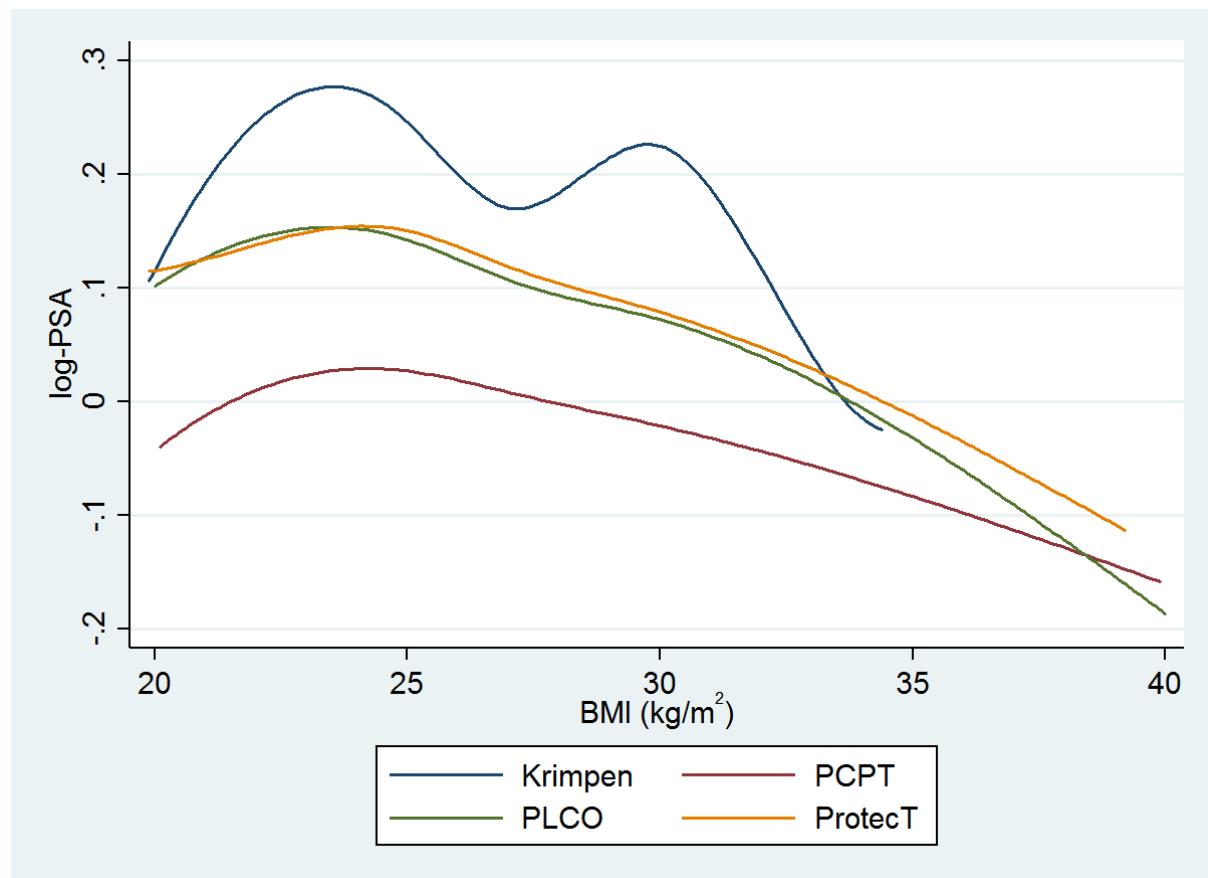


Figure 5.5 Line graph of BMI against log-PSA, using restricted cubic splines of BMI. Limited to BMIs between 20 kg/m^2 and 40 kg/m^2 . Knots at BMI quantiles (24.8 kg/m^2 , 27 kg/m^2 and 30 kg/m^2)

Checking for heterogeneity 2: BMI and prostate cancer risk

The probability of diagnosed prostate cancer for each study across the same BMI range for each study is shown in **Figure 5.6**, using cubic splines for BMI. In all studies, the distribution of prostate cancers across BMI appear similar, apart from fluctuation caused by the low number of participants and men with prostate cancer in the Krimpen study, and the decrease in risk in the PLCO study at lower BMI values, also likely caused by relatively low numbers of men with prostate cancer. We therefore assumed that study-specific intercepts but fixed-effect was valid for the association between BMI and prostate cancer.

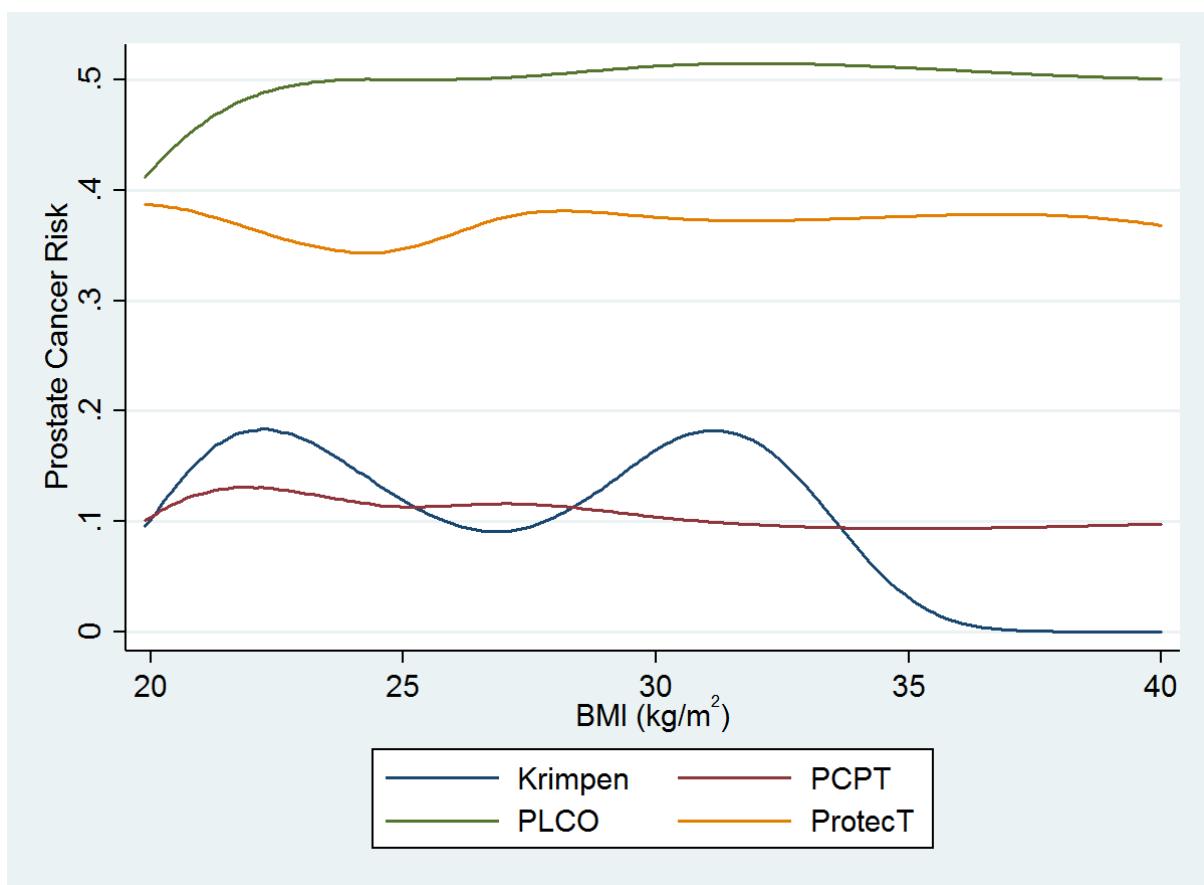


Figure 5.6 Line graph of BMI against prostate cancer risk, using restricted cubic splines of BMI. Limited to BMIs between 20 kg/m² and 40 kg/m². Knots at BMI quantiles (24.8 kg/m², 27 kg/m² and 30 kg/m²)

Checking for heterogeneity 3: PSA and prostate cancer risk

The probability of prostate cancer for each study for different log-PSA levels is shown in **Figure 5.7**, using cubic splines for log-PSA. ERSPC-Rotterdam and PLCO and ProtecT were limited to PSA values above 3.0 ng/ml, 4.0 ng/ml and 3.0 ng/ml respectively. The association between log-PSA and prostate cancer risk seemed similar for PCPT, which had data up to 3.0 ng/ml, and ProtecT, which had data above 3.0 ng/ml, with the two lines almost becoming a single continuous line. Participants in PLCO had a higher baseline rate of prostate cancer, and prostate cancer risk did not decrease as much as in other studies with PSA, although this may still have decreased had PLCO biopsied men with lower PSA levels. Participants in ERSPC-Rotterdam followed broadly the same distribution as ProtecT. The distribution for Krimpen followed a similar distribution as the others, although with an increase in risk at low PSA levels where there would be relatively few participants (only 6 men with prostate cancer had a log-PSA value less than 0.5 in Krimpen). Although there was some heterogeneity in the associations between log-PSA and prostate cancer risk, we assumed that study-specific intercepts but fixed-effect was likely valid.

In addition, non-white participants were removed and the USA, UK and Netherlands are all western countries with broadly similar incidences of prostate cancer (age standardised rates per 100,000 in 2012: 97.2 in North America (1), 111.1 in the UK (39), and 124.5 in the Netherlands (39)). Thus, since none of the analyses revealed any large discrepancies between the associations of the variables of interest, and ethnicity, age and family history of prostate cancer were all able to be controlled, we were satisfied imputing with fixed-effect but study-specific intercepts was a valid assumption.

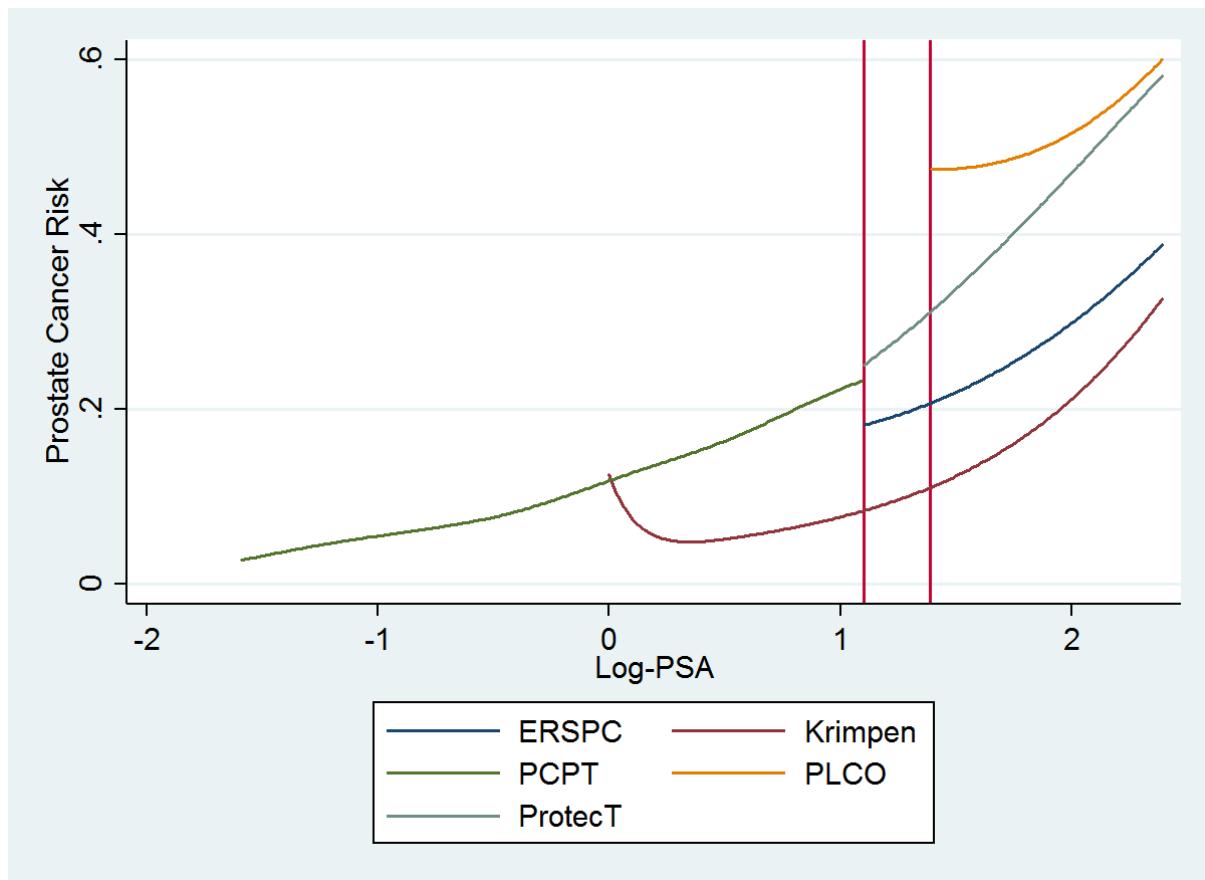


Figure 5.7 Line graph of log-PSA against prostate cancer risk, limited to the middle 98% of PSA values for each study (to remove outliers), using restricted cubic splines of log-PSA. ERSPC = ERSPC-Rotterdam. ERSPC-Rotterdam, PLCO and ProtecT were also limited to PSA values above 3.0 ng/ml, 4.0 ng/ml and 3.0 ng/ml respectively as prostate cancer status was not observed lower than this. Knots at rough overall log PSA quantiles (-0.5, 0 and 0.5)

Checking for interactions

In the regression with prostate cancer as the outcome, the OR for the interaction term was 1.000 (95% CI 0.999 to 1.002, P = 0.59), indicating no interaction between age and BMI for prostate cancer risk. When advanced prostate cancer was the outcome, the OR for the interaction term was 1.003 (95% CI 0.999 to 1.007, P = 0.10), indicating no interaction between age and BMI for advanced prostate cancer risk. When log-PSA was the outcome, the beta coefficient for the interaction term was 0.0002 (95% CI -0.0002 to 0.0005, P = 0.30), indicating no interaction between age and BMI for log-PSA.

We therefore did not include an interaction term between age and BMI in the imputation or analysis models for any outcome.

5.5.2. Imputation

The proportion of missing information for all variables is shown in **Table 5.3**. As data in PCPT was only used to impute missing values in the other studies, and was not included in any analyses, it is not included in the totals. There were 13,050 men with full information on all variables (14%), 59,004 men missing information for one variable (61%), 23,887 men missing information for two variables (25%), and 137 men missing information for three or four variables (0.1%). The percentage of men missing prostate cancer status is very high, but we believed that by using PCPT we could impute prostate cancer status without biasing the analyses. The other variables had relatively low levels of missingness, excepting BMI in ERSPC, although we also felt this could be imputed without bias.

Table 5.3 Summary of missing information

	ERSPC-Rotterdam	Krimpen	PLCO	ProtecT	Total for analysis	PCPT
Missing data						
PCa (N [%])	16,552 (82.9)	1,235 (74.4)	25,168 (76.2)	35,273 (85.2)	78,228 (81.4)	0 (0.0)
Age (N [%])	3 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.0)	0 (0.0)
BMI (N [%])	19,970 (100)	60 (3.6)	385 (1.2)	52 (0.1)	20,467 (21.3)	180 (1.0)
Log-PSA (N [%])	0 (0.0)	2 (0.1)	2,620 (7.9)	0 (0.0)	2,622 (2.7)	0 (0.0)
Family history PCa (N [%])	0 (0.0)	208 (12.5)	797 (2.4)	4,850 (11.7)	5,855 (6.1)	0 (0.0)
Missing data for:						
0 variables	0 (0.0)	359 (21.6)	7,253 (22.0)	5,438 (13.1)	13,050 (13.6)	17,635 (99.0)
1 variable	3,417 (17.1)	1,113 (67.0)	22,696 (68.7)	31,778 (76.7)	59,004 (61.4)	180 (1.0)
2 variables	16,551 (82.9)	175 (10.5)	2,960 (9.0)	4,191 (10.1)	23,887 (24.9)	0 (0.0)
3 variables	2 (0.0)	14 (0.8)	110 (0.3)	5 (0.0)	131 (0.1)	0 (0.0)
4 variables	0 (0.0)	0 (0.0)	6 (0.0)	0 (0.0)	6 (0.0)	0 (0.0)
PCa = prostate cancer, N = number of participants, ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study						

The summary statistics for each study, averaged across 100 imputed datasets, are shown in **Table 5.4**.

This table is separated into men who we assumed received a biopsy (first 5 rows), did not receive a biopsy (next 5 rows), and all participants (other rows). PCPT was not included in the totals column. We assumed 17,840 men received a biopsy (19%), and of these, 7,253 had prostate cancer (41%) and 943 men had advanced prostate cancer (5%). By contrast, we assumed 78,228 men did not receive a biopsy (81%), and when imputed, on average 19,095 of these men had prostate cancer (24%) and 1,043 men had advanced prostate cancer (1%). These results seem plausible when considering men without a prostate cancer diagnosis generally had lower PSA values.

Once imputed, the mean age, BMI and log-PSA did not change materially from the complete case analysis. Across all participants (those who we assumed received a biopsy and not), 26,348 men were estimated to have prostate cancer (34%), and 1,986 men advanced prostate cancer (2%).

Table 5.4 Summary of included studies for imputed data, averaged over 100 imputed datasets

	ERSPC-Rotterdam	Krimpen	PLCO	ProtecT	Total	PCPT
Men who were biopsied (prostate cancer status not imputed)						
Participants (% of total)	3,418 (17.1)	426 (25.6)	7,857 (23.8)	6,139 (14.8)	17,840 (18.6)	17,815 (100)
PCa (%)	1,013 (29.6)	58 (13.6)	3,890 (49.5)	2,292 (37.3)	7,253 (40.7)	2,254 (12.7)
Advanced PCa (%)*	227 (6.6)	7 (1.6)	488 (6.2)	221 (3.6)	943 (5.3)	49 (0.3)
No PCa (%)	2,405 (70.4)	368 (86.4)	3967 (50.5)	3847 (62.7)	10,587 (59.3)	15,561 (87.3)
Men who were not biopsied (prostate cancer status imputed)						
Participants (% of total)	16,552 (82.9)	1,235 (74.4)	25,168 (76.2)	35,273 (85.2)	78,228 (81.4)	0
PCa (%)	2,587 (15.6)	106 (8.6)	9,304 (37.0)	7,098 (20.1)	19,095 (24.4)	0
Advanced PCa (%)*	238 (1.4)	7 (0.6)	563 (2.2)	235 (0.7)	1,043 (1.3)	0
No PCa (%)	13,965 (84.4)	1,129 (91.4)	15,864 (63.0)	28,175 (79.9)	59,133 (75.6)	0
All participants (men who were and were not biopsied combined)						
Participants	19,970	1,661	33,025	41,412	96,068	17,815
PCa (%)	3,600 (18.0)	164 (9.9)	13,194 (40.0)	9,390 (22.7)	26,348 (33.7)	2,254 (12.7)
Advanced PCa (%)*	465 (2.3)	14 (0.9)	1,051 (3.2)	456 (1.1)	1,986 (2.1)	49 (0.3)
No PCa (%)	16,370 (82.0)	1,497 (90.1)	19,831 (60.0)	32,022 (77.3)	69,720 (89.1)	15,561 (87.3)
Age (mean, [SD])	63.6 (5.6)	61.5 (6.6)	62.7 (5.3)	60.0 (5.5)	61.7 (5.7)	63.3 (5.7)
BMI (mean, [SD])	25.9 (4.1)	26.0 (3.0)	27.6 (4.2)	27.5 (4.0)	27.2 (4.1)	27.7 (4.1)
Log-PSA (mean, [SD])	0.33 (0.93)	0.27 (0.89)	0.16 (0.88)	0.10 (0.91)	0.17 (0.91)	0.03 (0.63)
Family history PCa (%)	1,427 (7.1)	151 (9.1)	2,543 (7.7)	2,348 (5.7)	6,470 (6.7)	2,768 (15.5)

*Advanced prostate cancer defined as a T-score 3 or above ($T \geq 3$), nodal involvement ($N=1$) or metastases ($M=1$)

PCa = prostate cancer, N = number of participants, ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

The probability of prostate cancer, imputed for each study, for different log-PSA levels is shown in **Figure 5.8**. The distributions of prostate cancer appear plausible in the screening studies, becoming higher as log-PSA increases.

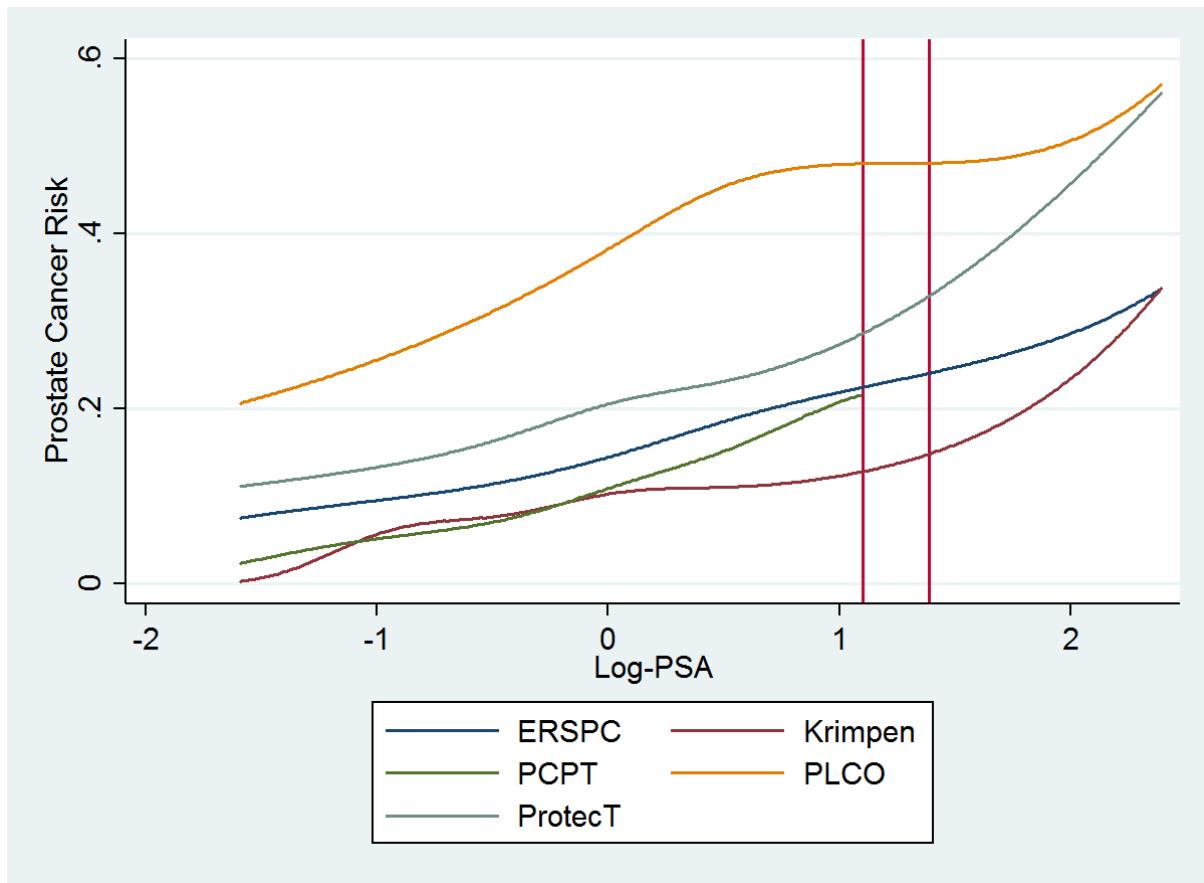


Figure 5.8 Line graph of prostate cancer risk against log-PSA, limited to the middle 98% of PSA values for each study to remove outliers, using restricted cubic splines of log-PSA. ERSPC = ERSPC-Rotterdam

5.5.3. Linear analyses

Linear analysis 1: BMI and prostate cancer

The overall OR for prostate cancer for a 5 kg/m² increase in BMI using two-stage IPD fixed-effect meta-analysis was estimated to be 0.98 (95% CI 0.95 to 1.01, P = 0.16), **Table 5.5** and **Figure 5.9**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0$, P = 0.99). There was no material difference in the results of the different meta-analysis methods.

Table 5.5 Results for the linear analysis of the association between BMI and prostate cancer

Study	OR for prostate cancer for a 5 kg/m ² increase in BMI (95% CI)	P value
ERSPC-Rotterdam	0.97 (95% CI 0.90 to 1.04)	0.40
Krimpen	0.99 (95% CI 0.69 to 1.42)	0.95
PCPT	0.91 (95% CI 0.86 to 0.96)	0.00070
PLCO	0.98 (95% CI 0.94 to 1.02)	0.39
ProtecT	0.98 (95% CI 0.93 to 1.03)	0.42
One-stage FE meta-analysis*	0.98 (95% CI 0.95 to 1.02)	0.32
Two-stage FE meta-analysis*	0.98 (95% CI 0.95 to 1.01)	0.16
Two-stage RE meta-analysis*	0.98 (95% CI 0.95 to 1.01)	0.16

*All meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial,

PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

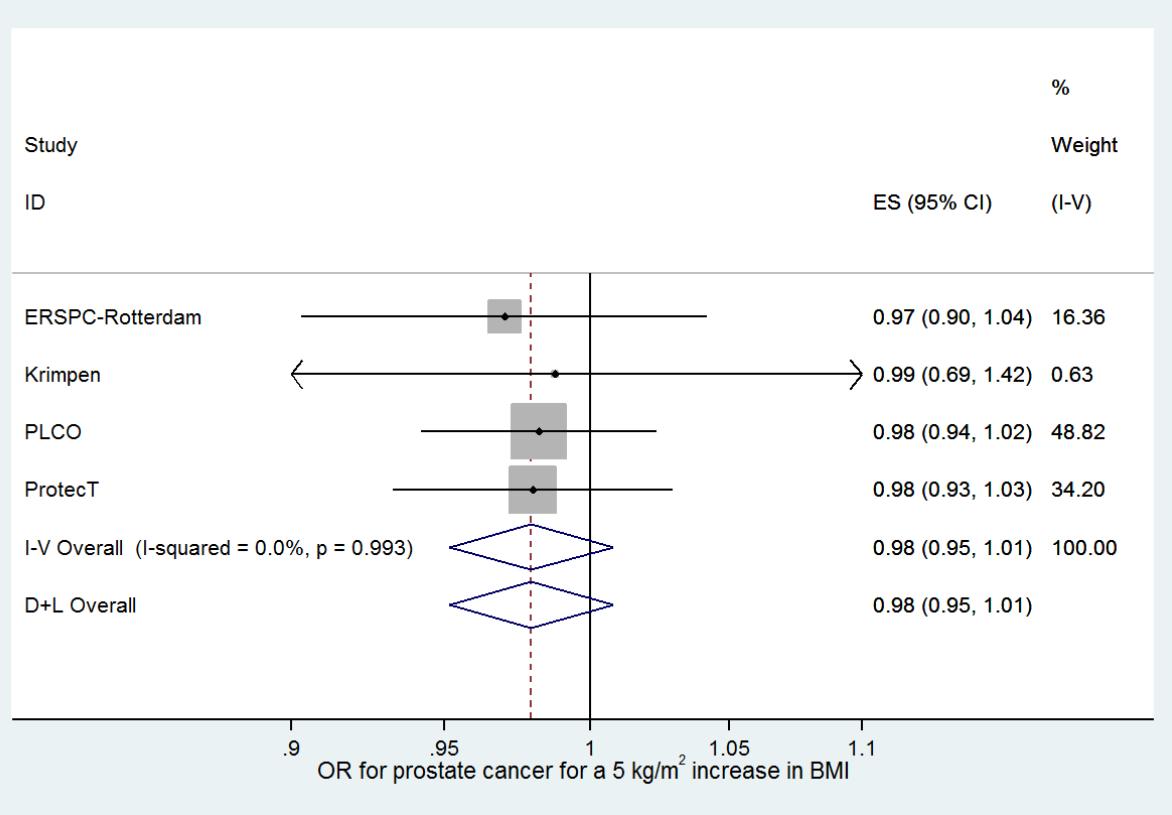


Figure 5.9 Forest plot of BMI and prostate cancer risk, showing the OR for prostate cancer for a 5 kg/m² increase in BMI

Linear analysis 2: BMI and advanced prostate cancer

The overall OR for advanced prostate cancer for a 5 kg/m² increase in BMI using two-stage IPD fixed-effect meta-analysis was 1.00 (95% CI 0.93 to 1.08, P = 0.98), **Table 5.6** and **Figure 5.10**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0$, P = 1.00). There was no material difference in the results of the different meta-analysis methods.

Table 5.6 Results for the linear analysis of the association between BMI and prostate cancer

Study	OR for advanced prostate cancer for a 5 kg/m ² increase in BMI (95% CI)	P value
ERSPC-Rotterdam	0.98 (95% CI 0.83 to 1.16)	0.81
Krimpen	0.97 (95% CI 0.32 to 2.94)	0.96
PCPT	0.80 (95% CI 0.55 to 1.16)	0.24
PLCO	1.00 (95% CI 0.91 to 1.11)	0.94
ProtecT	1.00 (95% CI 0.87 to 1.17)	0.95
One-stage FE meta-analysis*	1.00 (95% CI 0.92 to 1.09)	0.94
Two-stage FE meta-analysis*	1.00 (95% CI 0.93 to 1.08)	0.98
Two-stage RE meta-analysis*	1.00 (95% CI 0.93 to 1.08)	0.98

*All meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

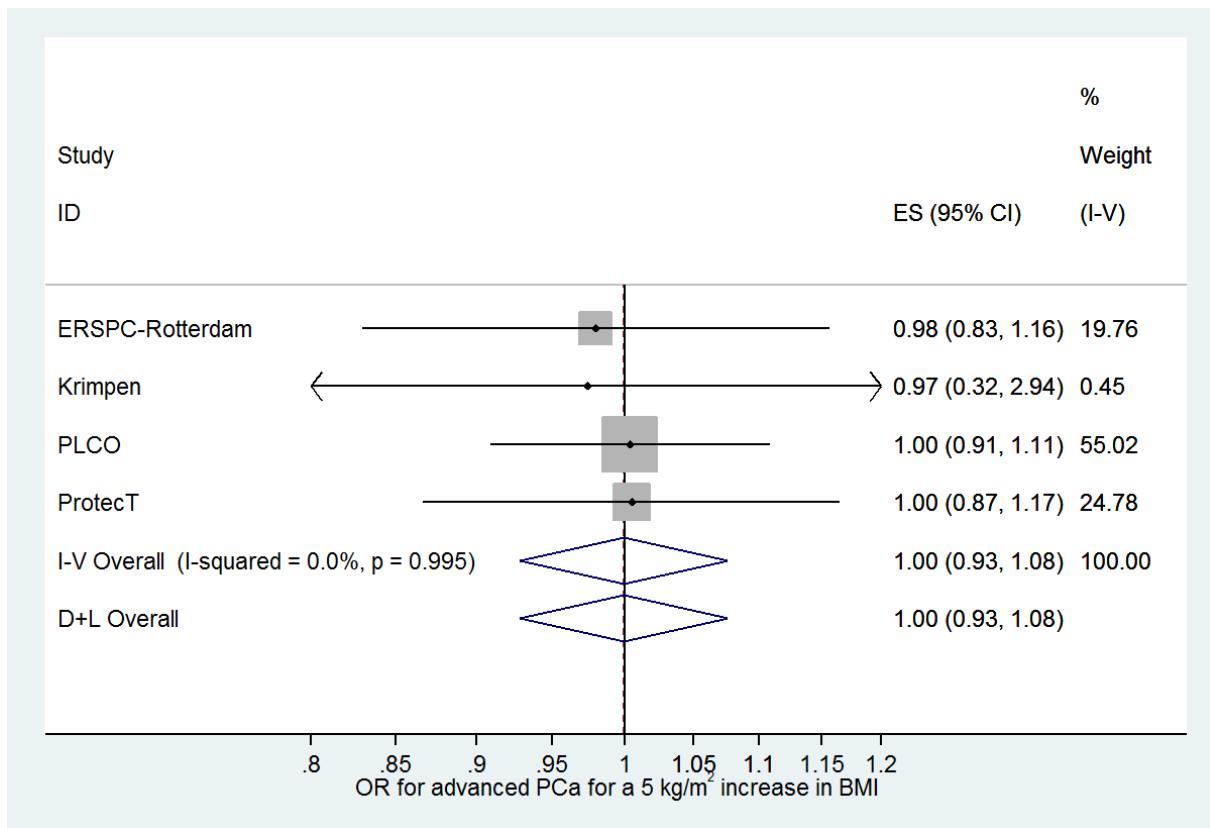


Figure 5.10 Forest plot of BMI and advanced prostate cancer risk, showing the OR for advanced prostate cancer for a 5 kg/m² increase in BMI

Linear analysis 3: BMI and PSA

The overall *direct* effect estimate for a 5 kg/m² increase in BMI on PSA using two-stage IPD fixed-effect meta-analysis was estimated to be a percentage change of -6.51% (95% CI -7.21% to -5.81%, P < 0.001), **Table 5.7** and **Figure 5.11**. There was no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0$, P = 0.51). There was no material difference in the results of the different meta-analysis methods.

The overall *total* effect estimate for a 5 kg/m² increase in BMI on PSA using two-stage IPD fixed-effect meta-analysis was estimated to be a percentage change of -6.64% (95% CI -7.33% to -5.95%, P < 0.001), **Table 5.7** and **Figure 5.12**. There was similarly no evidence of inconsistency in the effect estimates across studies ($I^2 = 0.0$, P = 0.50), and no material difference in the results of the different meta-analysis methods.

Table 5.7 Results for the linear analysis of the association between BMI and PSA (direct and total effects)

Study	Percentage change in PSA for a 5 kg/m ² increase in BMI (95% CI)	P value
Direct Effect		
ERSPC-Rotterdam	-6.78% (-8.89% to -4.61%)	<0.001
Krimpen	-4.27% (-10.89% to 2.84%)	0.23
PCPT	-4.02% (-5.08% to -2.94%)	<0.001
PLCO	-7.07% (-8.16% to -5.96%)	<0.001
ProtecT	-6.04% (-7.04% to -5.02%)	<0.001
One-stage FE meta-analysis*	-6.44% (-7.19% to -5.67%)	<0.001
Two-stage FE meta-analysis*	-6.51% (-7.21% to -5.81%)	<0.001
Two-stage RE meta-analysis*	-6.51% (-7.21% to -5.81%)	<0.001
Total Effect		
ERSPC-Rotterdam	-6.95% (-9.06% to -4.79%)	<0.001
Krimpen	-4.32% (-10.94% to 2.8%)	0.23
PCPT	-4.31% (-5.38% to -3.22%)	<0.001
PLCO	-7.21% (-8.30% to -6.10%)	<0.001
ProtecT	-6.18% (-7.17% to -5.18%)	<0.001
One-stage FE meta-analysis*	-6.57% (-7.32% to -5.82%)	<0.001
Two-stage FE meta-analysis*	-6.64% (-7.33% to -5.95%)	<0.001
Two-stage RE meta-analysis*	-6.64% (-7.33% to -5.95%)	<0.001

*All meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

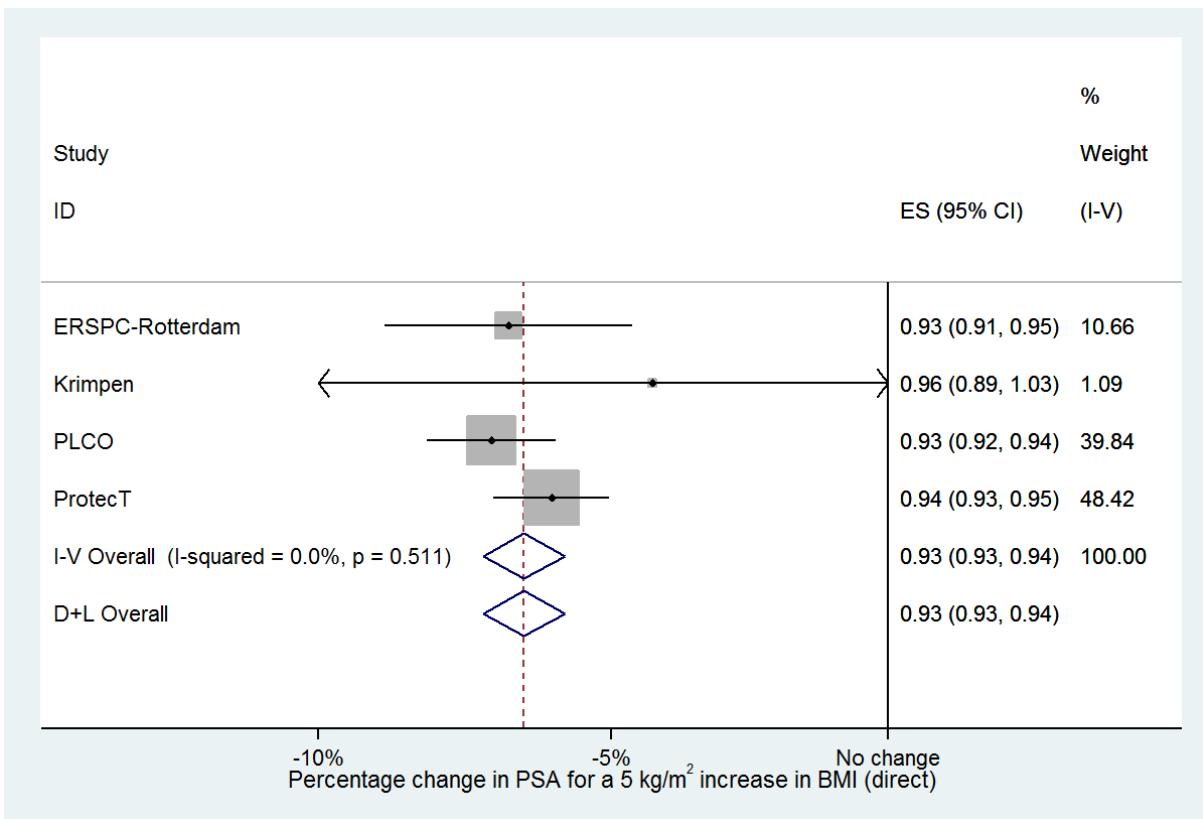


Figure 5.11 Forest plot of BMI and log PSA (direct effect), showing the percentage change in PSA for a 5 kg/m² increase in BMI

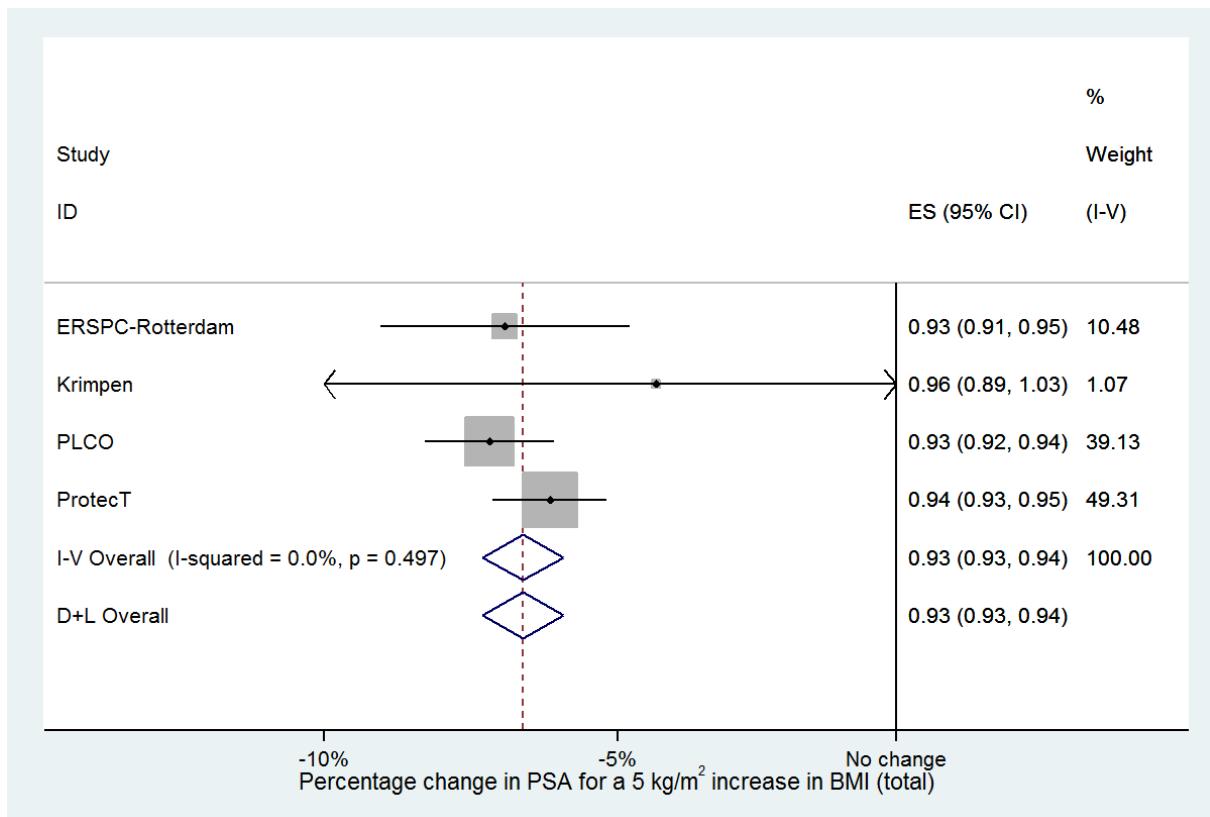


Figure 5.12 Forest plot of BMI and log PSA (total effect), showing the percentage change in PSA for a 5 kg/m² increase in BMI

Linear analysis 4: prostate cancer and PSA

The overall percentage change in PSA from prostate cancer using two-stage IPD fixed-effect meta-analysis was estimated to be 43.9% (95% CI 40.6% to 47.2%, $P < 0.001$), **Table 5.8** and **Figure 5.13**. There was strong evidence of inconsistency in effect estimates across studies ($I^2 = 87.1\%$, $P < 0.001$), likely due to differences in prostate cancer severity between the study populations and the precision of the effect estimates. However, the estimated change in PSA from prostate cancer was still relatively similar between studies (ranging from a 34% to a 66% increased risk), so although inconsistency was large, we considered that imputation assuming study-specific intercepts but fixed-effect estimates and analysis was still reasonable, especially as there was no substantial differences in the pooled estimates from the different meta-analysis methods.

Table 5.8 Results for the linear analysis of the association between prostate cancer and PSA

Study	Percentage change in PSA from prostate cancer (95% CI)	P value
ERSPC-Rotterdam	50.70% (42.67% to 59.18%)	<0.001
Krimpen	66.19% (36.24% to 102.73%)	<0.001
PCPT	34.08% (30.45% to 37.81%)	<0.001
PLCO	36.29% (31.99% to 40.72%)	<0.001
ProtecT	53.41% (46.86% to 60.25%)	<0.001
One-stage FE meta-analysis*	45.51% (40.95% to 50.23%)	<0.001
Two-stage FE meta-analysis*	43.87% (40.57% to 47.24%)	<0.001
Two-stage RE meta-analysis*	48.34% (37.20% to 60.39%)	<0.001

*All meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

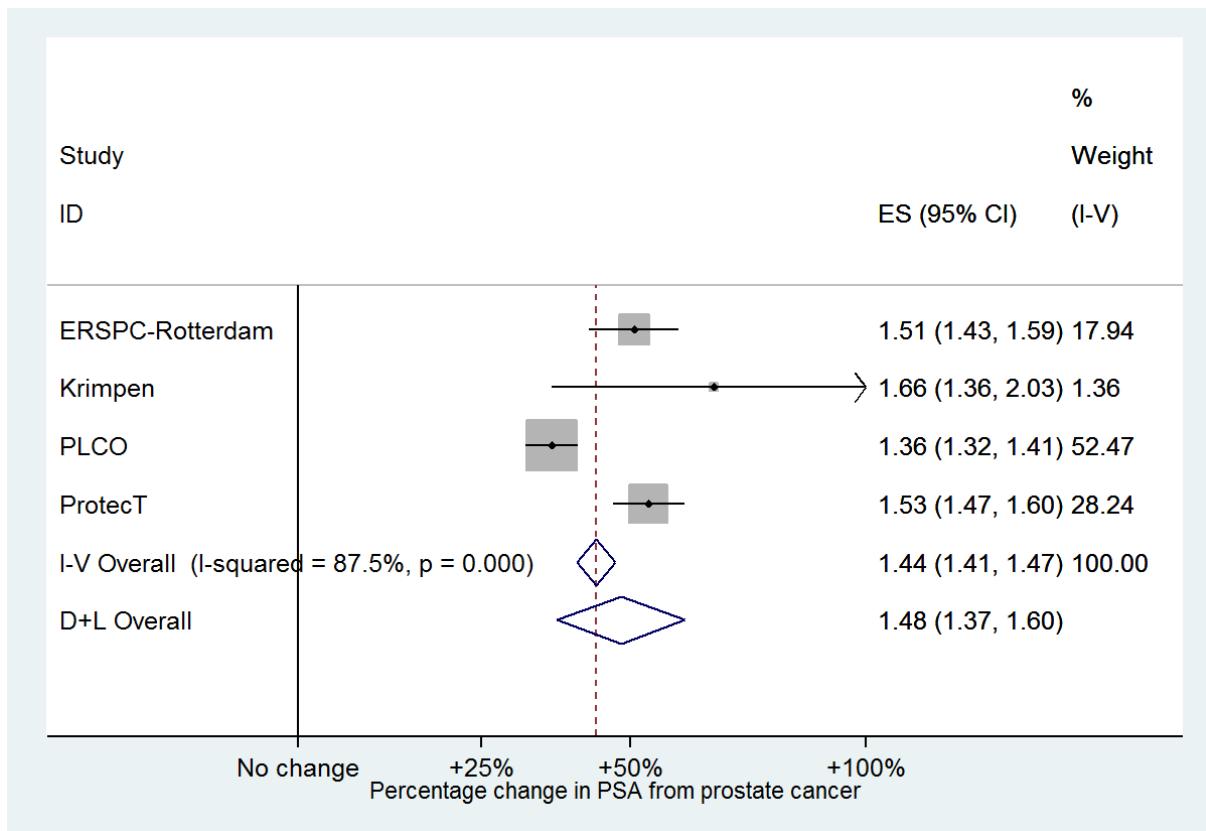


Figure 5.13 Forest plot of the association between prostate cancer and PSA (multiplicative), showing the percentage change in PSA from prostate cancer

5.5.4. Non-linear analyses

Non-linear analysis 1: BMI and prostate cancer

The individual log-OR estimates for prostate cancer for BMI, BMI^2 and BMI^3 are shown in **Table 5.9**.

Figure 5.14 shows the predicted prostate cancer risk and 95% CI for BMI values between 20 kg/m² and 40 kg/m², for a 60-year-old man in the ERSPC-Rotterdam study with no family history of prostate cancer. A fixed-effect model was used to estimate the predicted prostate cancer risk, thus other studies would have the same distribution, but higher or lower depending on the study-specific intercept. There is a suggestion of a non-linear association between BMI and prostate cancer, but the CIs are wide and the results are consistent with no association between BMI and prostate cancer. In addition, all the BMI terms have wide CIs that cross the null, indicating there is no evidence of an association between BMI and prostate cancer.

To provide some numerical context, the predicted prostate cancer risk for a man with a BMI of 20 kg/m² was 0.16 (95% CI 0.14 to 0.18), of 25 kg/m² was 0.17 (95% CI 0.16 to 0.18), of 30 kg/m² was 0.17 (95% CI 0.16 to 0.18) and of 35 kg/m² of 0.16 (95% CI 0.14 to 0.18).

Table 5.9 Results for the cubic analysis of the association between BMI and prostate cancer

Study	BMI	BMI^2	BMI^3
	Log-OR (95% CI)	Log-OR (95% CI)	Log-OR (95% CI)
ERSPC-Rotterdam	0.81 (-0.99 to 2.61)	-0.023 (-0.077 to 0.03)	0.00019 (-0.00032 to 0.00071)
Krimpen	0.45 (-16.06 to 16.96)	-0.012 (-0.589 to 0.564)	0.00009 (-0.00659 to 0.00676)
PCPT	0.72 (-1.17 to 2.61)	-0.026 (-0.084 to 0.032)	0.00027 (-0.00032 to 0.00085)
PLCO	1.23 (-0.03 to 2.5)	-0.036 (-0.074 to 0.003)	0.00031 (-0.00007 to 0.0007)
ProtecT	0.68 (-0.74 to 2.1)	-0.019 (-0.062 to 0.025)	0.00015 (-0.00029 to 0.00058)
One-stage FE meta-analysis*	0.89 (-0.12 to 1.9)	-0.025 (-0.056 to 0.006)	0.00021 (-0.00009 to 0.00051)

*Meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

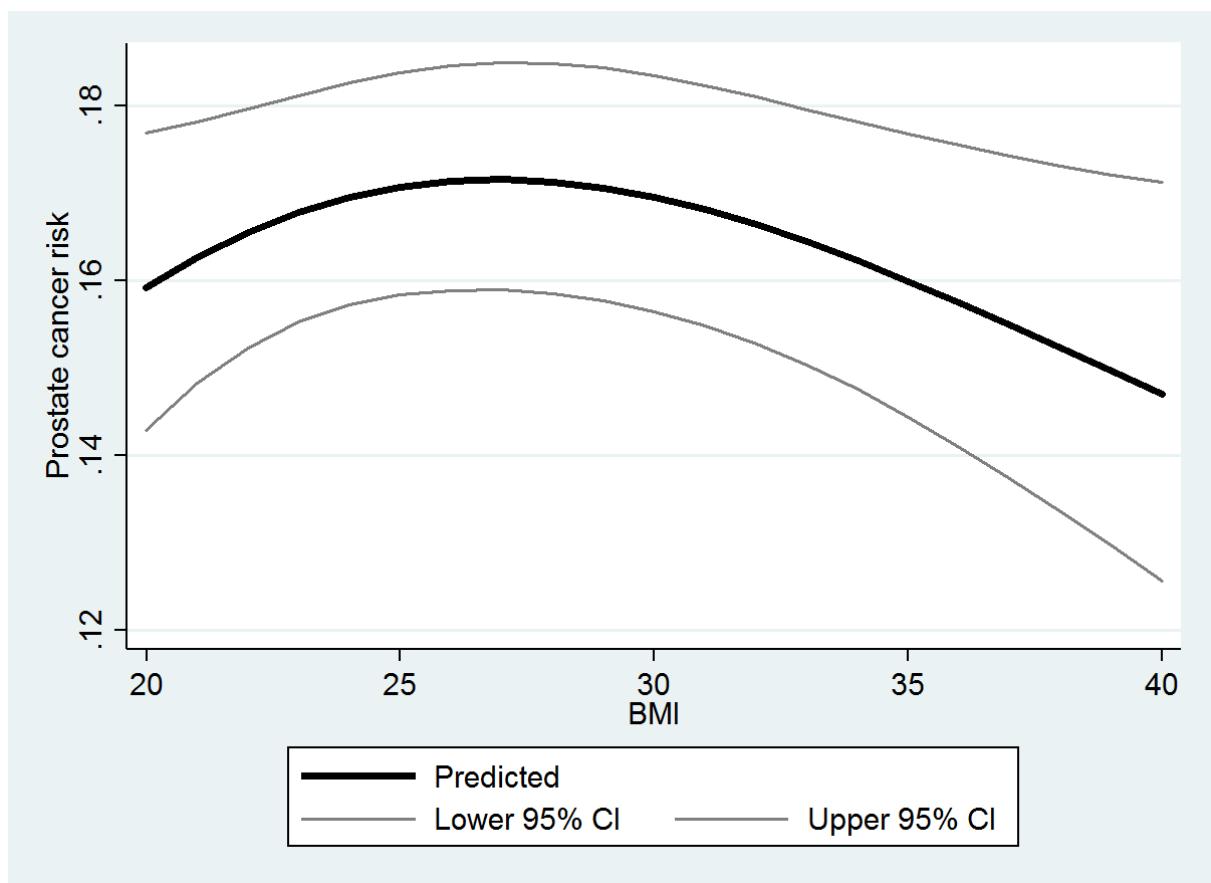


Figure 5.14 Predicted prostate cancer risk for values of BMI between $20 \text{ kg}/\text{m}^2$ and $40 \text{ kg}/\text{m}^2$, for a 60-year-old man from the ERSPC-Rotterdam study with no family history of prostate cancer

Non-linear analysis 2: BMI and advanced prostate cancer

The individual log-OR estimates for advanced prostate cancer for BMI, BMI^2 and BMI^3 are shown in

Table 5.10. Figure 5.15 shows the predicted advanced prostate cancer risk and 95% CI for BMI values between 20 kg/m^2 and 40 kg/m^2 , for a 60-year-old man in the ERSPC-Rotterdam study with no family history of prostate cancer. Similar to prostate cancer, there is the suggestion of a non-linear association between BMI and advanced prostate cancer, but the CIs are wide and the results are consistent with no association between BMI and advanced prostate cancer. Equally, all the BMI terms have wide CIs that cross the null, so as with prostate cancer, there is no evidence of an association between BMI and advanced prostate cancer.

To provide some numerical context, the predicted probability of advanced prostate cancer for a man with a BMI of 20 kg/m^2 was 0.019 (95% CI 0.013 to 0.025), of 25 kg/m^2 was 0.020 (95% CI 0.017 to 0.024), of 30 kg/m^2 was 0.021 (95% CI 0.017 to 0.025) and of 35 kg/m^2 of 0.019 (95% CI 0.015 to 0.025).

Table 5.10 Results for the cubic analysis of the association between BMI and advanced prostate cancer

Study	BMI	BMI^2	BMI^3
	Log-OR (95% CI)	Log-OR (95% CI)	Log-OR (95% CI)
ERSPC-Rotterdam	0.84 (-4.16 to 5.85)	-0.023 (-0.17 to 0.13)	0.00016 (-0.0013 to 0.0016)
Krimpen	-1.99 (-40.74 to 36.76)	0.054 (-1.27 to 1.38)	-0.00046 (-0.015 to 0.014)
PCPT	-8.35 (-20.02 to 3.32)	0.235 (-0.15 to 0.62)	-0.00213 (-0.0062 to 0.0020)
PLCO	0.90 (-3.57 to 5.37)	-0.017 (-0.16 to 0.13)	0.00004 (-0.0014 to 0.0015)
ProtecT	-0.14 (-4.67 to 4.40)	0.006 (-0.14 to 0.15)	-0.00009 (-0.0015 to 0.0013)
One-stage FE meta-analysis*	0.77 (-2.17 to 3.70)	-0.017 (-0.11 to 0.073)	0.00009 (-0.00081 to 0.00099)

*Meta-analysis results exclude PCPT

ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

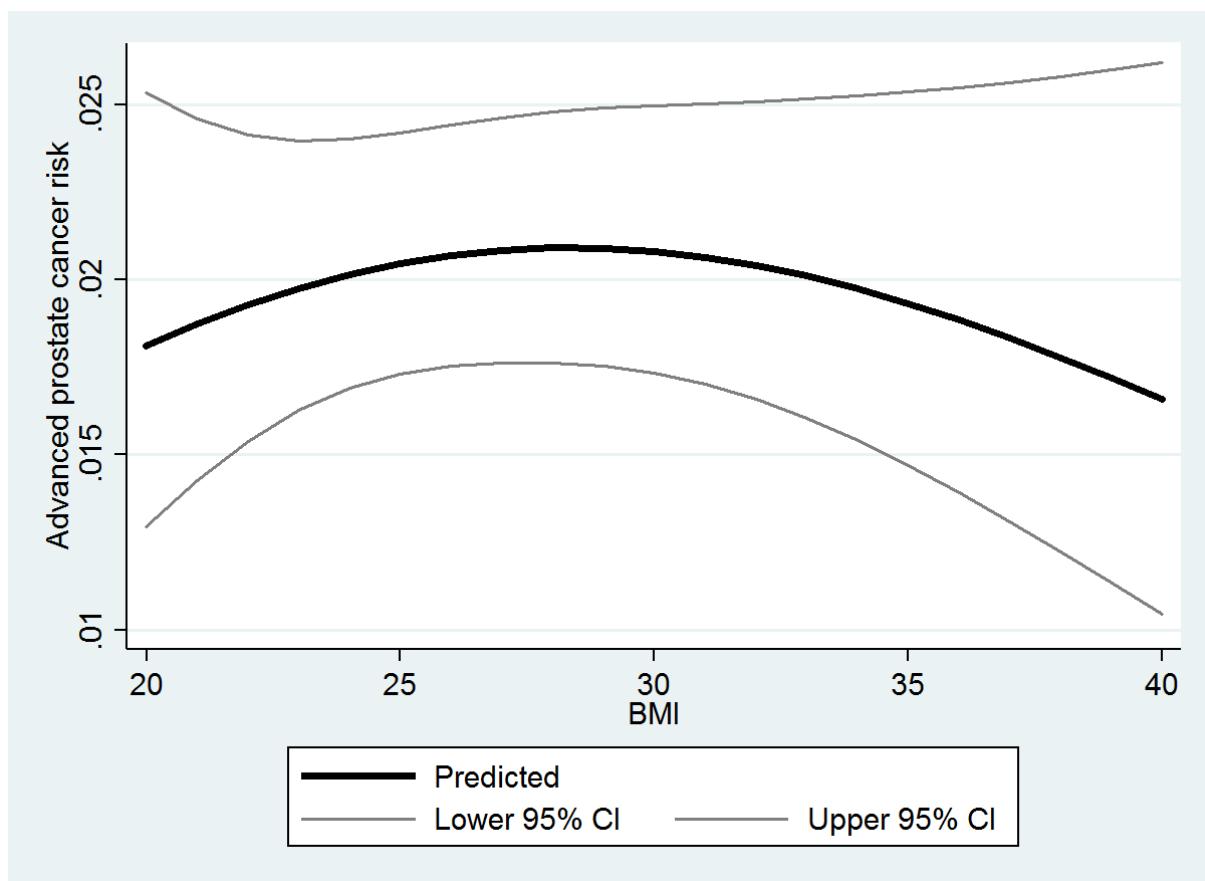


Figure 5.15 Predicted advanced prostate cancer risk for values of BMI between $20 \text{ kg}/\text{m}^2$ and $40 \text{ kg}/\text{m}^2$, for a 60-year-old man from the ERSPC-Rotterdam study with no family history of prostate cancer

Non-linear analysis 3: BMI and PSA

The effect estimates of BMI, BMI^2 and BMI^3 for log-PSA are shown in **Table 5.11**. **Figure 5.16** shows the predicted log-PSA and 95% CI for BMI values between 20 kg/m^2 and 40 kg/m^2 , for a 60-year-old man without prostate cancer in the ERSPC-Rotterdam study. A clear non-linear association is shown, with log-PSA increasing with BMI until a BMI of around 23 kg/m^2 , then decreasing. The CIs for all BMI term coefficients also do not cross the null, implying that a non-linear model fits the data well.

To provide some numerical context, the predicted PSA for a man without prostate cancer with a BMI of 20 kg/m^2 was 1.16 ng/ml (95% CI 1.13 to 1.19 ng/ml), of 25 kg/m^2 was 1.17 ng/ml (95% CI 1.15 to 1.19 ng/ml), of 30 kg/m^2 was 1.10 ng/ml (95% CI 1.08 to 1.12 ng/ml) and of 35 kg/m^2 of 1.00 ng/ml (95% CI 0.98 to 1.02 ng/ml).

Table 5.11 Results for the cubic analysis of the association between BMI and log-PSA

Study	BMI	BMI^2	BMI^3
	Beta ⁺ (95% CI)	Beta ⁺ (95% CI)	Beta ⁺ (95% CI)
ERSPC-Rotterdam	0.66 (0.04 to 1.28)	-0.020 (-0.039 to -0.002)	0.00018 (0 to 0.00036)
Krimpen	1.31 (-2.06 to 4.68)	-0.043 (-0.161 to 0.075)	0.00044 (-0.00094 to 0.0018)
PCPT	0.56 (0.16 to 0.96)	-0.017 (-0.029 to -0.004)	0.00014 (0.00002 to 0.00027)
PLCO	0.66 (0.29 to 1.02)	-0.02 (-0.032 to -0.009)	0.00018 (0.00006 to 0.00029)
ProtecT	0.51 (0.26 to 0.76)	-0.016 (-0.023 to -0.009)	0.00014 (0.00007 to 0.00021)
One-stage FE meta-analysis*	0.59 (0.39 to 0.80)	-0.019 (-0.025 to -0.012)	0.00016 (0.0001 to 0.00022)

*Meta-analysis results exclude PCPT
Beta is the coefficient for each BMI term in the regression of log-PSA
ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study

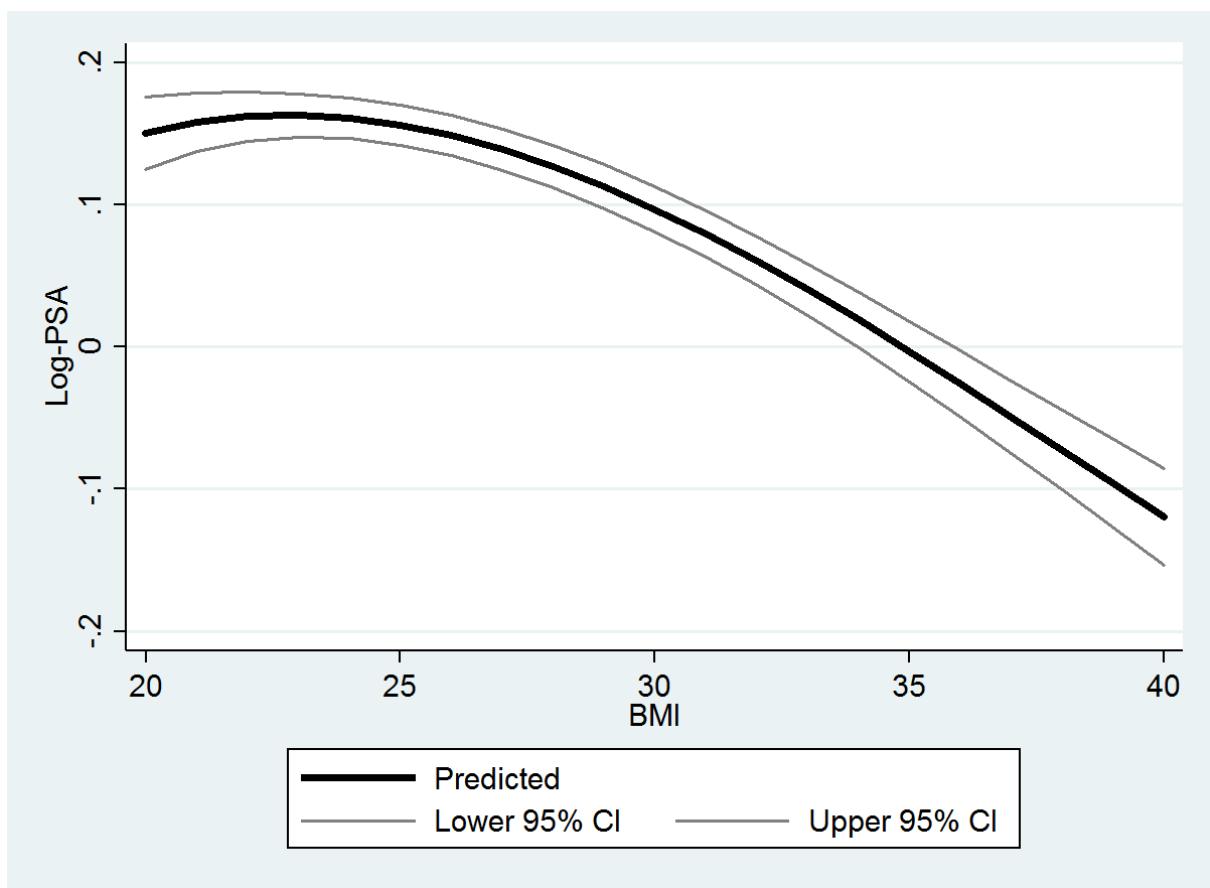


Figure 5.16 Predicted log-PSA for values of BMI between 20 kg/m² and 40 kg/m², for a 60-year-old man from the ERSPC-Rotterdam study without prostate cancer

5.6. Discussion

In this analysis, I used IPD to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA, estimating both linear and non-linear associations.

5.6.1. BMI and prostate cancer

Using the imputed datasets, there was little evidence of a linear association between BMI and prostate cancer risk (OR for a 5 kg/m² increase in BMI = 0.98, 95% CI 0.95 to 1.01), consistent with the AD meta-analysis (OR = 1.00, 95% CI 0.97 to 1.02). There was a suggestion of a non-linear association, with the risk of prostate cancer increasing up to around 27 kg/m², then decreasing, but the probability of prostate cancer was very similar (0.16 to 0.17) between BMIs of 20 to 35 kg/m², and all BMI terms in the cubic regression had wide CIs that crossed the null. In the complete case analysis, there was an apparent negative association between BMI and prostate cancer (OR for a 5 kg/m² increase in BMI 0.94, 95% CI 0.91 to 0.98), which was more pronounced in PLCO than ProtecT. This is expected as screening induces a negative association between BMI and prostate cancer, and PLCO had a higher screening threshold than ProtecT. PCPT showed a negative association between BMI and prostate cancer, likely because it restricted participation to men with a PSA of less than 3.0 ng/ml.

In summary, we found no strong evidence that BMI was associated with prostate cancer.

5.6.2. BMI and advanced prostate cancer

Using the imputed datasets, there was no evidence of a linear association between BMI and advanced prostate cancer (OR for a 5 kg/m² increase in BMI = 1.00, 95% CI 0.93 to 1.08), unlike the AD meta-analysis which showed a small non-significant positive effect (OR = 1.05, 95% CI 0.99 to 1.10). There was again a suggestion of a non-linear association, with the risk of advanced prostate cancer increasing up to around 28 kg/m², then decreasing, but the probability of advanced prostate cancer was very similar (0.019 to 0.021) between BMIs of 20 to 35 kg/m², and again all BMI terms in the cubic regression had wide CIs that crossed the null. The complete case analysis gave a very similar result to the imputed analysis (OR for a 5 kg/m² increase in BMI = 0.98, 95% CI 0.89 to 1.08).

In summary, we found no strong evidence that BMI was associated with advanced prostate cancer.

5.6.3. BMI and PSA

Using the imputed datasets, there was evidence of a negative linear association between BMI and PSA (percentage change in PSA (*direct* effect estimate) for a 5 kg/m² increase in BMI = -6.51%, 95% CI -

7.21% to -5.81%), greater than, but still consistent with, the AD meta-analysis result (percentage change = -5.16%, 95% CI -6.85% to -3.44%). There was evidence of a non-linear association, with log-PSA increasing with BMI up to around 23 kg/m^2 , then decreasing, and the predicted log-PSA values for BMIs between 20 and 35 kg/m^2 reflect this. The graph of the association using imputed data (**Figure 5.16**) was very similar to the graph using only men with prostate biopsies and cubic splines of BMI (**Figure 5.5**). There were no material differences between the *direct* and *total* effect estimates, which is consistent with no (or very little) association between BMI and prostate cancer. In the complete case analysis, the *total* effect estimate was consistent with the imputed results, implying that imputing prostate cancer status did not change the *total* effect estimate of BMI on PSA. However, the *direct* effect estimate was underestimated, likely because there was an apparent negative association between BMI and prostate cancer from the bias induced by screening.

Overall, we found strong evidence that BMI was associated with a decrease in PSA, and the association is likely non-linear, possibly with a slight increase in PSA at lower BMI levels.

5.6.4. Prostate cancer and PSA

Using the imputed datasets, prostate cancer was associated with an increase in PSA (percentage increase in PSA from prostate cancer = 43.9%, 95% CI 40.6% to 47.2%). There was a large amount of inconsistency in the effect estimates between these studies. One potential reason for this is differences in the severity of prostate cancer between studies.

5.6.5. Methodology

Overall, the differences between the complete case analyses and the imputed analyses highlight the importance of imputing missing data when missingness would bias the results. We believe that using a stratified imputation model was the best choice for our data, given the systematically missing data which made within-study imputation a poor choice. However, imputation may potentially also induce bias in some associations. For example, although we tested whether the assumption of fixed-effect estimates but study-specific intercepts was valid in all analyses, and we believe the different studies had similar populations, because the imputation model borrowed information from all studies, any bias in the PCPT study results may have induced bias in the screening studies. This would be undetectable as we cannot compare the imputed values with the true values. Equally, we imputed BMI for all the ERSPC-Rotterdam participants, and have no way to assess whether the imputed values are consistent with the true values (although the populations in the IPD are quite similar). It is, however, unlikely that the imputation would impart more bias in all the associations we estimated

than it removed by reducing the effect of screening, although the method may overestimate the precision of the estimates.

The predicted risk of prostate cancer in the imputed data seems viable. In ERSPC-Rotterdam, 30% of biopsied men had prostate cancer, compared to an estimated 16% of men without a biopsy; in PLCO, 50% of biopsied men had prostate cancer, compared to an estimated 37% of men without a biopsy; and in ProtecT, 37% of biopsied men had prostate cancer, compared to an estimated 20% of men without a biopsy. The decreased risk of prostate cancer in men with lower PSA levels (and men who were not diagnosed with prostate cancer in routine care) is expected, and fits with results from other studies (245,246). In ERSPC-Rotterdam, BMI was completely imputed, with a mean value of 25.9 kg/m², very close to the mean in Krimpen (26.0 kg/m²) but slightly lower than in other studies; this is encouraging, since both the Krimpen and ERSPC-Rotterdam populations were from the Netherlands. For all men, prostate cancer risk was quite variable between studies; 10% in Krimpen versus 40% in PLCO. One reason for this difference is that men were followed up for longer in PLCO and had more rounds of biopsies (ProtecT only had one round and we only used information from the first round in ERSPC-Rotterdam), and so men had a larger amount of time to be diagnosed with prostate cancer. However, there may also be differences in the study populations, although unless some of the differences are effect modifiers of any of the measured associations, then different baseline rates of prostate cancer would not affect the results of this chapter.

The use of IPD allowed us to conduct the analyses in multiple ways, allowing us to determine whether different assumptions affected the overall results. We used both one- and two-stage meta-analysis models in the linear analyses, and both fixed-effect and random-effects models in the two-stage IPD meta-analysis models, all of which gave broadly the same results. We also tested for an interaction between age and BMI, which we determined not to be an issue for our analyses.

5.6.6. Strengths and limitations

The strengths of this study were the large amount of IPD, allowing for imputation of missing data, notably prostate cancer status in the large proportion of men who were probably not biopsied, control over which variables were included in each analysis and the capability to conduct non-linear analyses. The included studies had similar populations, which potentially increased the precision of the summary estimates as inconsistency was generally very low. However, this also limits the findings to white men from developed countries.

The most obvious limitation was that prostate cancer status was not known for a large proportion of men, due to only subgroups of men being offered biopsies. While imputation is a useful method in

determining likely men with prostate cancer, it cannot replace the certain knowledge of which men had cancer. In addition, the imputation relies on the missing data being either MAR (conditional only on variables included in the imputation model) or MCAR, and on the imputation models being specified correctly, although the results will be robust to all variables being MAR given all other variables. As the imputed data cannot be compared with the true values, we can never be sure the imputed data is completely valid, although the results seem plausible. PCPT provided an unbiased estimate (with respect to PSA) of prostate cancer status, which may have improved the imputations in other studies. It is also true that even with a biopsy, cancers may be missed. Larger men tend to have larger prostates (84), possibly increasing the percentage of missed cancers in those with large BMIs, compounding the problem potentially caused by increasing BMIs decreasing PSA and thus biopsy rates for larger men. This problem could not be solved with imputation without a study where prostate cancer was diagnosed perfectly.

We assumed a multiplicative association between prostate cancer and PSA, rather than an additive association. This was for the practical reason that log-PSA was used as an outcome (and covariate in the imputation), but also because it is impossible to categorically state whether prostate cancer causes a percentage change in PSA or an absolute change. We could also not include variables not measured by the IPD studies, such as exercise, alcohol intake, smoking, socioeconomic status and other variables that may be related both to BMI and prostate cancer or PSA: therefore, the results from this chapter are likely subject to residual confounding. However, we controlled for the variables we thought most likely to confound the results, limiting the effect of any bias on the results.

For the associations between BMI, prostate cancer and advanced prostate cancer, we did not consider follow-up time, calculating ORs rather than hazard ratios (HRs). In **Section 4.4**, there was no difference between the results of studies where BMI was measured less or more than 2 years before the average time to prostate cancer diagnosis, but there may still be differences between the ORs and HRs in prospective studies. However, a previous study estimating the association between BMI and prostate cancer in PLCO gave an HR for a 5 kg/m^2 increase in BMI of 0.97 (95% CI 0.94 to 1.00) (247), consistent with our imputed result (OR = 0.98, 95% CI 0.94 to 1.02), despite the previous study including the non-screened arm of PLCO, not imputing prostate cancer status and adjusting for more confounders.

5.7. Conclusion

Using IPD, we have determined that there no likely no association between BMI and prostate cancer or advanced prostate cancer (in either linear or non-linear models of BMI), but there the is strong evidence for a non-linear negative association between BMI and log-PSA. We also determined that prostate cancer is associated with an increase in PSA, although the percentage change varied between studies.

CHAPTER 6. COMBINING AGGREGATE DATA AND INDIVIDUAL PARTICIPANT DATA TO ESTIMATE THE ASSOCIATIONS BETWEEN BMI, PROSTATE CANCER, ADVANCED PROSTATE CANCER AND PSA

6.1. Aims

The aim of this chapter is to combine the AD and IPD to calculate estimates of all associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA. I estimated both linear and non-linear associations.

6.2. Methods

6.2.1. Studies

In this chapter, I combine the AD results from the systematic review conducted in **Chapter 4** with the IPD results from the studies in **Chapter 5**.

AD studies examining the association between BMI and PSA rarely controlled for prostate cancer status, and often could not fully control for prostate cancer as it was unobserved unless a biopsy was performed, and the chance of biopsy may be determined by PSA levels. Studies that did not control for prostate cancer status measured the *total* effect of BMI on PSA, rather than the *direct* effect. Therefore, I only estimate the *total* effect estimate of BMI and PSA in this chapter (**Figure 6.1**).

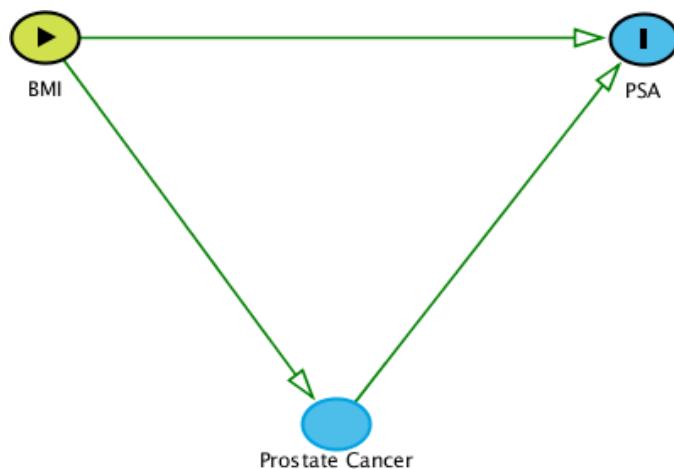


Figure 6.1 Diagram showing the assumed directions of effect for BMI, PSA and prostate cancer

6.2.2. Linear associations

I combined the linear effect estimates from the AD and the IPD for the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA, using random-effects and fixed-effect meta-analyses. Although fixed-effect models are more congenial with the imputation model used for the IPD analysis, the AD study populations were much more heterogeneous than the IPD studies, warranting the use of random-effects meta-analysis. The methodology for calculation of each study-specific effect estimate remained the same as in the previous two chapters, detailed in **Section 4.3** and **Section 5.4**. Results were presented in forest plots, expressed as the OR for prostate cancer or advanced prostate cancer, and percentage change in PSA, for a 5 kg/m^2 increase in BMI. The forest plots were subgrouped by study analysis, AD or IPD, and a Q-test for heterogeneity was conducted to assess any difference in effect estimates between the AD and IPD (49). A summary diagram of the associations between BMI, prostate cancer and PSA was created. We did not conduct

a systematic review and meta-analysis of the association between prostate cancer and PSA, so this association was only estimated from the IPD.

6.2.3. Non-linear associations

There was strong evidence from the IPD meta-analysis (**Section 5.5.4**) that the association between BMI and PSA was non-linear, and some evidence the associations between BMI, prostate cancer, and advanced prostate cancer were possibly non-linear. We therefore investigated these non-linear associations here.

The only AD studies that could be included in a non-linear meta-analysis were those presenting results by categories of BMI; ORs for prostate cancer or advanced prostate cancer, or observed means and SDs of PSA or log-PSA. The studies still had to account for age in the study design or analysis, and maximally adjusted results were used. Studies presenting linear effect estimates and mean differences in BMI between men with and without prostate cancer could not be used to estimate non-linear effect estimates, and were excluded from this analysis. The following BMI categories were used: normal weight ($BMI < 25 \text{ kg/m}^2$), overweight ($25 \text{ kg/m}^2 \leq BMI < 30 \text{ kg/m}^2$), and obese ($BMI \geq 30 \text{ kg/m}^2$). The ORs for other categories of BMI were not used (such as morbidly obese, $BMI \geq 35 \text{ kg/m}^2$), though if possible I combined the mean and SD of PSA for different categories with neighbouring categories. For example, I combined obese ($30 \text{ kg/m}^2 \leq BMI < 35 \text{ kg/m}^2$) and morbidly obese categories ($BMI \geq 35 \text{ kg/m}^2$) to give the mean and SD of PSA for men with a $BMI \geq 30 \text{ kg/m}^2$ (248). I used the reported ORs for overweight and obese men versus normal weight men for the prostate cancer and advanced prostate cancer meta-analyses, and estimated the mean difference (MD) in log-PSA between overweight or obese and normal weight men for the PSA meta-analysis. The estimated average MD in log-PSA from the meta-analysis was exponentiated to give the percentage change in PSA between BMI groups.

The mean BMI values, median mid-year of recruitment, and the average estimated linear effect estimates from random-effects meta-analysis were compared between AD studies that could and could not be included in the non-linear meta-analyses, and a Q-test for heterogeneity (49) was conducted between the groups to determine if there might be a risk of bias resulting from which studies could be included.

In the IPD studies, categorical effect estimates (ORs and MDs) were estimated using the methods in **Section 5.2.4**, using datasets imputed with BMI as a categorical variable (normal weight, overweight or obese), with effect estimates calculated across datasets using dummy variables for categorised BMI.

The AD and IPD effect estimates were combined using both random-effects and fixed-effect meta-analysis, and forest plots were produced for the differences between overweight and normal weight men and between obese and normal weight men. The weighted mean BMI across all studies was calculated in each category of BMI (248) to assess whether the difference in mean BMI between normal weight and overweight groups was comparable to that between overweight and obese groups.

6.3. Results: Linear analysis

6.3.1. BMI and prostate cancer

All 58 AD studies (**Table 4.2**) and 4 IPD studies (**Table 5.4**) were included in this analysis. In these studies, there were 9,302,337 participants and 188,244 men had prostate cancer (2.0%).

In random-effects meta-analysis combining the AD and IPD studies, the average OR for prostate cancer for a 5 kg/m² increase in BMI across studies was estimated to be 0.99 (95% CI 0.97 to 1.01, P = 0.57),

Figure 6.2. In the fixed-effect model, the overall OR was 1.00 (95% CI 1.00 to 1.01, P = 0.42). There was strong evidence of inconsistency across the entire set of studies ($I^2 = 65.8\%$, $P < 0.001$), driven by the AD studies. However, there was no strong evidence of heterogeneity between the pooled effect estimates from AD and IPD ($P = 0.09$).

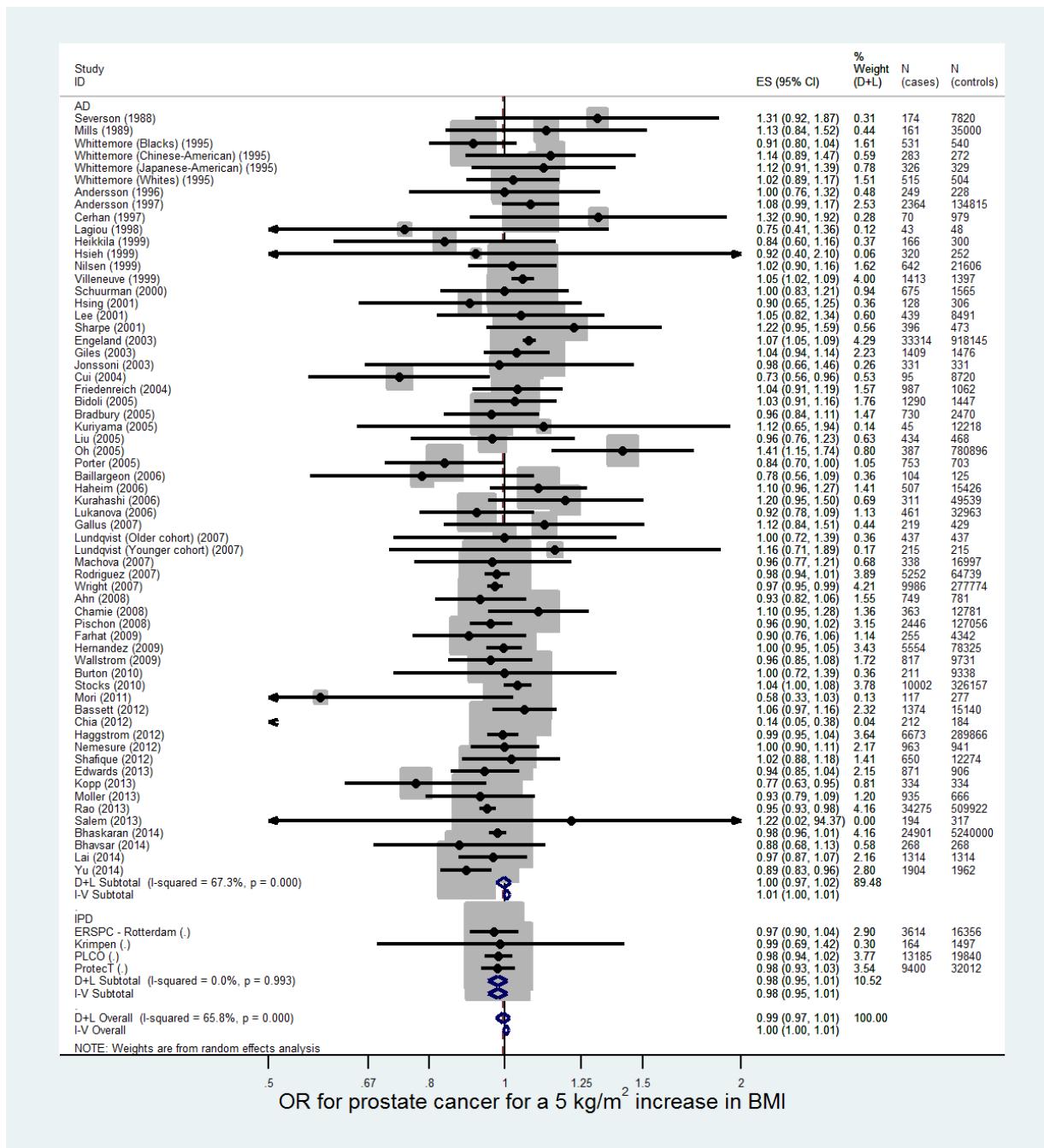


Figure 6.2 Forest plot of OR of prostate cancer for a 5 kg/m² increase in BMI

6.3.2. BMI and advanced prostate cancer

All 14 AD studies (**Table 4.4**) and 4 IPD studies (**Table 5.4**) were included in this analysis. In these studies, there were 1,149,184 participants and 10,354 men had advanced prostate cancer (0.90%).

In random-effects meta-analysis combining the AD and IPD studies, the average OR for advanced prostate cancer for a 5 kg/m² increase in BMI across studies was estimated to be 1.04 (95% CI 0.99 to 1.08, P = 0.09), **Figure 6.3**. In the fixed-effect model, the overall OR was 1.04 (95% CI 1.00 to 1.07, P = 0.04). There was little evidence of inconsistency across the entire set of studies ($I^2 = 20.0\%$, P = 0.22), and no evidence of heterogeneity between the pooled effect estimates from AD and IPD (P = 0.28).

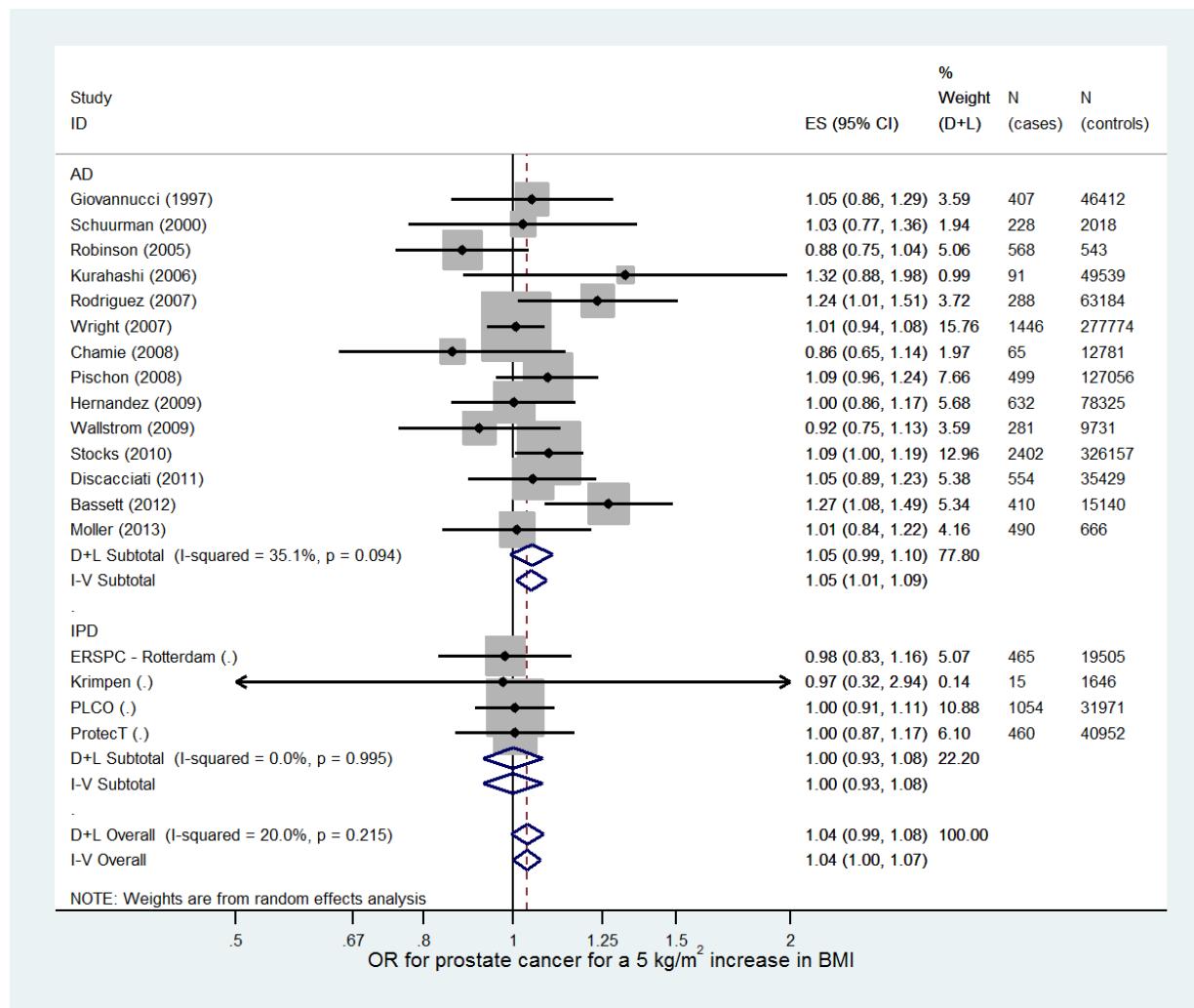


Figure 6.3 Forest plot of OR of advanced prostate cancer for a 5 kg/m² increase in BMI

6.3.3. BMI and PSA

All 10 AD studies (**Table 4.6**) and 4 IPD studies (**Table 5.4**) were included in this analysis. In these studies, there were 159,716 participants, with 63,648 men in the AD studies (39.9%) and 96,068 in the IPD studies (60.1%).

In random-effects meta-analysis combining the AD and IPD studies, the average percentage change in PSA for a 5 kg/m² increase in BMI was estimated to be -5.62% (95% CI -6.73% to -4.51%, P < 0.001), **Figure 6.4**. There was strong evidence of inconsistency across the entire set of studies ($I^2 = 62.6\%$, P < 0.001), and evidence of heterogeneity between the pooled effect estimates from AD and IPD (P = 0.007), where the IPD studies gave a combined estimate further from the null than the AD studies. Fixed-effect results were similar, although the IPD studies were given more weight, drawing the effect estimate slightly farther from the null: the overall percentage change in PSA for a 5 kg/m² increase in BMI was -6.03% (95% CI -6.56% to -5.49%, P < 0.001).

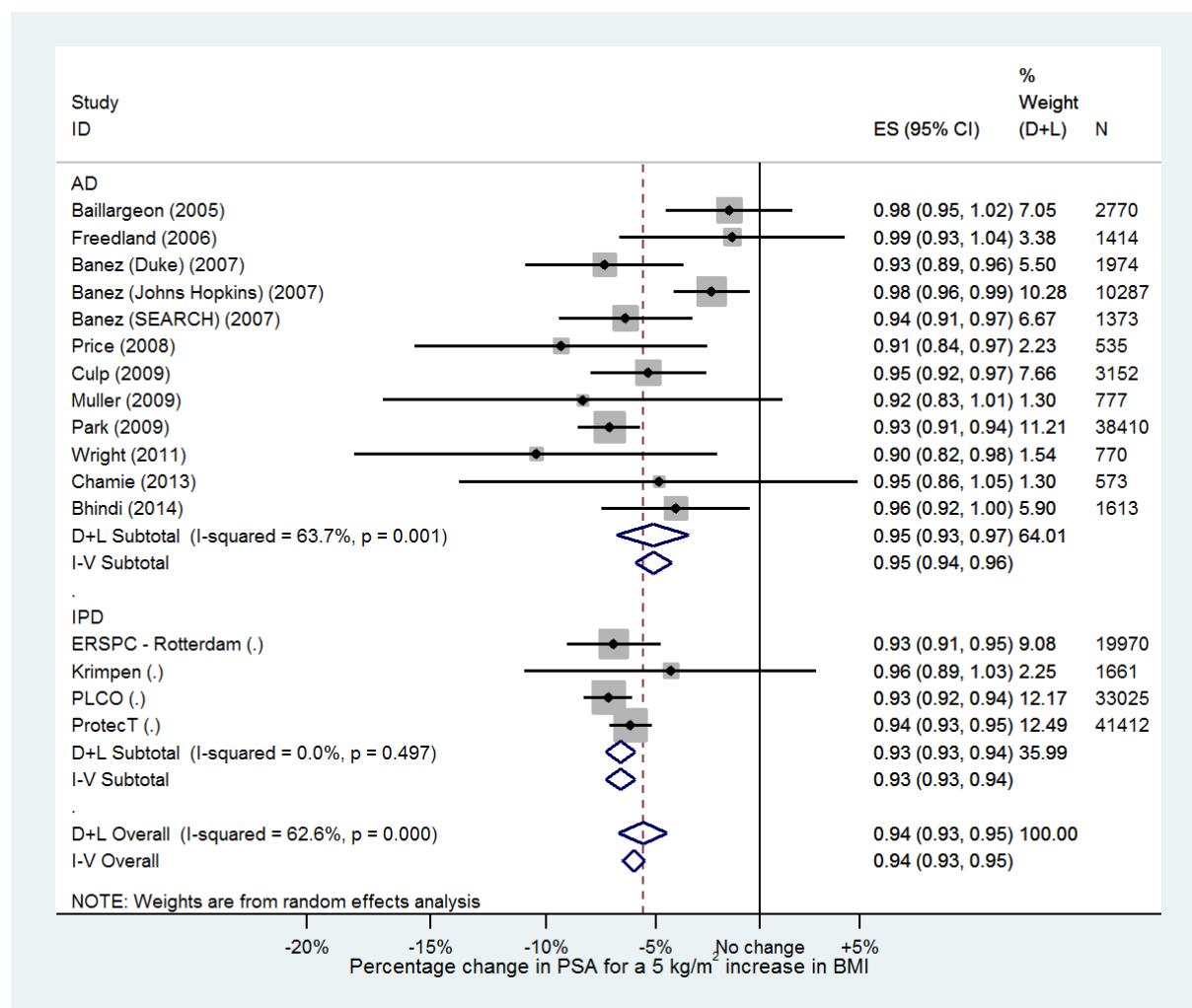


Figure 6.4 Forest plot of the percentage change in PSA for a 5 kg/m² increase in BMI

6.3.4. Summary diagram

The summary diagram, which shows the linear estimates from the meta-analyses of AD and IPD studies for the associations between BMI and prostate cancer, and BMI and PSA, and the association between prostate cancer and PSA using just the IPD studies, is shown in **Figure 6.5**. In the linear model, there would only be a very small (if any) *indirect* effect of BMI on PSA through prostate cancer, since there is only a very small, non-significant negative association between BMI and prostate cancer. Thus, the *total* and the *direct* linear effects of BMI on PSA would be almost identical.

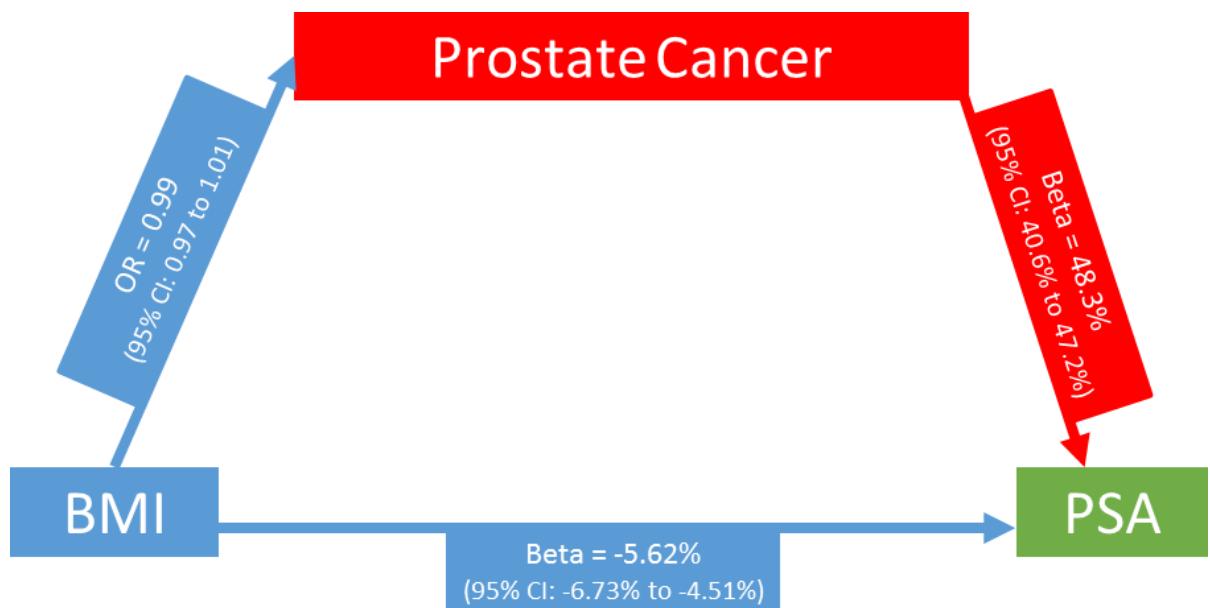


Figure 6.5 Summary diagram showing the results from all meta-analyses assuming linear associations between BMI, prostate cancer and PSA. OR for prostate cancer and percentage change in PSA for a 5 kg/m² increase in BMI

6.4. Results: Non-linear analysis

In all analyses, AD studies with categorical data were combined with the four IPD studies used in the IPD meta-analysis.

6.4.1. BMI and prostate cancer

The AD used for this analysis consists of nine of 58 studies (37,95,98,99,101,114,116,118,122) from the meta-analysis of BMI and prostate cancer that presented ORs for overweight and obese men versus normal weight men. Only six of the AD studies presented ORs for overweight men, whereas all nine had ORs for obese men versus normal weight men. The AD studies included in the non-linear analysis were all prospective, and when compared with prospective studies not included in the non-linear analysis, had a slightly higher mean BMI (26.6 kg/m^2 versus 25.3 kg/m^2), a later mid-year of recruitment (median of 1994 versus 1986) and a stronger pooled linear effect estimate from random-effects meta-analysis (OR for a 5 kg/m^2 increase in BMI = 0.97, 95% CI 0.96 to 0.99 in the included studies, versus OR = 1.00, 95% CI 0.97 to 1.04 in the excluded studies, P for heterogeneity < 0.001). The difference in effect estimates is likely due to the difference in mid-year of recruitments, as this was shown to be associated with effect size with meta-regression in **Section 4.4.1**. Combining across the included AD and the IPD, there were 571,206 participants and 49,764 men had prostate cancer (8.7%). The weighted mean BMI across all studies was 22.6 kg/m^2 for the normal BMI category, 27.3 kg/m^2 for the overweight category, and 32.2 kg/m^2 for the obese category.

Table 6.1 shows the mean BMI, total number of men and number of men with prostate cancer in each category of BMI, **Table 6.2** shows the ORs for prostate cancer for each study for overweight and obese versus normal weight men, with forest plots presented in **Figure 6.6** and **Figure 6.7**. For the random-effects meta-analysis, the average OR for prostate cancer between overweight and normal weight men was estimated to be 1.01 (95% CI 0.98 to 1.05, P = 0.37), with no evidence of inconsistency across the entire set of studies (I^2 = 0.0%, P = 0.98), and the average OR for prostate cancer between obese and normal weight men was estimated to be 0.96 (95% CI 0.92 to 0.99, P = 0.02), with no evidence of inconsistency across the entire set of studies (I^2 = 0.0%, P = 0.98). Fixed-effect models gave the same results. There was no evidence for heterogeneity between the effect estimates for the AD and IPD (P = 0.92 and P = 0.90 respectively).

Given these results, there is some evidence for a non-linear association between BMI and prostate cancer risk, with no reduction in risk in overweight versus normal men but lower risk in obese men.

Table 6.1 Summary BMI and number of men with and without prostate cancer for normal, overweight and obese BMI categories.

Author/Study	Normal weight (<25 kg/m ²)			Overweight (25-29.9 kg/m ²)			Obese (≥30 kg/m ²)		
	BMI (kg/m ²)	Cases	Controls	BMI (kg/m ²)	Cases	Controls	BMI (kg/m ²)	Cases	Controls
Aggregate Data									
Jonssoni (122)	23.1	190	185	27.1	125	136	30.8	13	8
Liu (101)	22.6	106	119				31.7	106	123
Porter (118)	22.2	195	173				31.0	178	175
Baillargeon (99)	22.7	21	22	27.6	50	47	33.0	33	56
Machova (98)	23.6	42	2,159	27.6	210	10,273	31.6	86	4,565
Rodriguez (37)	21.6	1,935	23,167				32.1	556	7,809
Wright (116)	21.2	3,076	NK	27.5	5,054	NK	32.8	1,532	NK
Hernandez (114)	23.0	2,366	32,779	27.2	2,462	33,438	31.3	666	11,245
Shafique (95)	22.7	279	5,499	27.1	320	5,729	31.4	51	1,046
Individual Participant Data									
ERSPC-Rotterdam	22.1	1,488	6,625	27.3	1,665	7,525	32.1	461	2,206
Krimpen	23.2	62	551	27.0	87	802	31.9	16	144
PLCO	23.1	3,384	5,098	27.2	6,733	9,906	33.3	3,067	4,837
ProtecT	23.1	2,473	8,413	27.3	4,854	16,249	33.1	2,073	7,350
<i>ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study *NK = not known</i>									

Table 6.2 Results for the categorical analysis of the association between BMI and prostate cancer

Author/Study	OR for prostate cancer	
	Overweight vs normal (95% CI)	Obese vs normal (95% CI)
Aggregate Data		
Jonssoni (122)	0.90 (0.61 to 1.32)	1.60 (0.61 to 4.18)
Liu (101)		0.96 (0.61 to 1.51)
Porter (118)		0.77 (0.56 to 1.06)
Baillargeon (99)	1.45 (0.73 to 2.89)	0.72 (0.36 to 1.45)
Machova (98)	1.05 (0.76 to 1.46)	0.97 (0.66 to 1.42)
Rodriguez (37)		0.94 (0.85 to 1.04)
Wright (116)	1.00 (0.96 to 1.05)	0.97 (0.91 to 1.03)
Hernandez (114)	1.04 (0.97 to 1.11)	0.94 (0.85 to 1.04)
Shafique (95)	1.02 (0.86 to 1.20)	1.03 (0.76 to 1.39)
Individual Participant Data		
ERSPC-Rotterdam	1.00 (0.88 to 1.13)	0.95 (0.80 to 1.12)
Krimpen	0.98 (0.62 to 1.54)	1.05 (0.50 to 2.19)
PLCO	1.03 (0.94 to 1.12)	0.96 (0.86 to 1.06)
ProtecT	1.02 (0.93 to 1.12)	0.96 (0.87 to 1.07)
<i>ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study</i>		

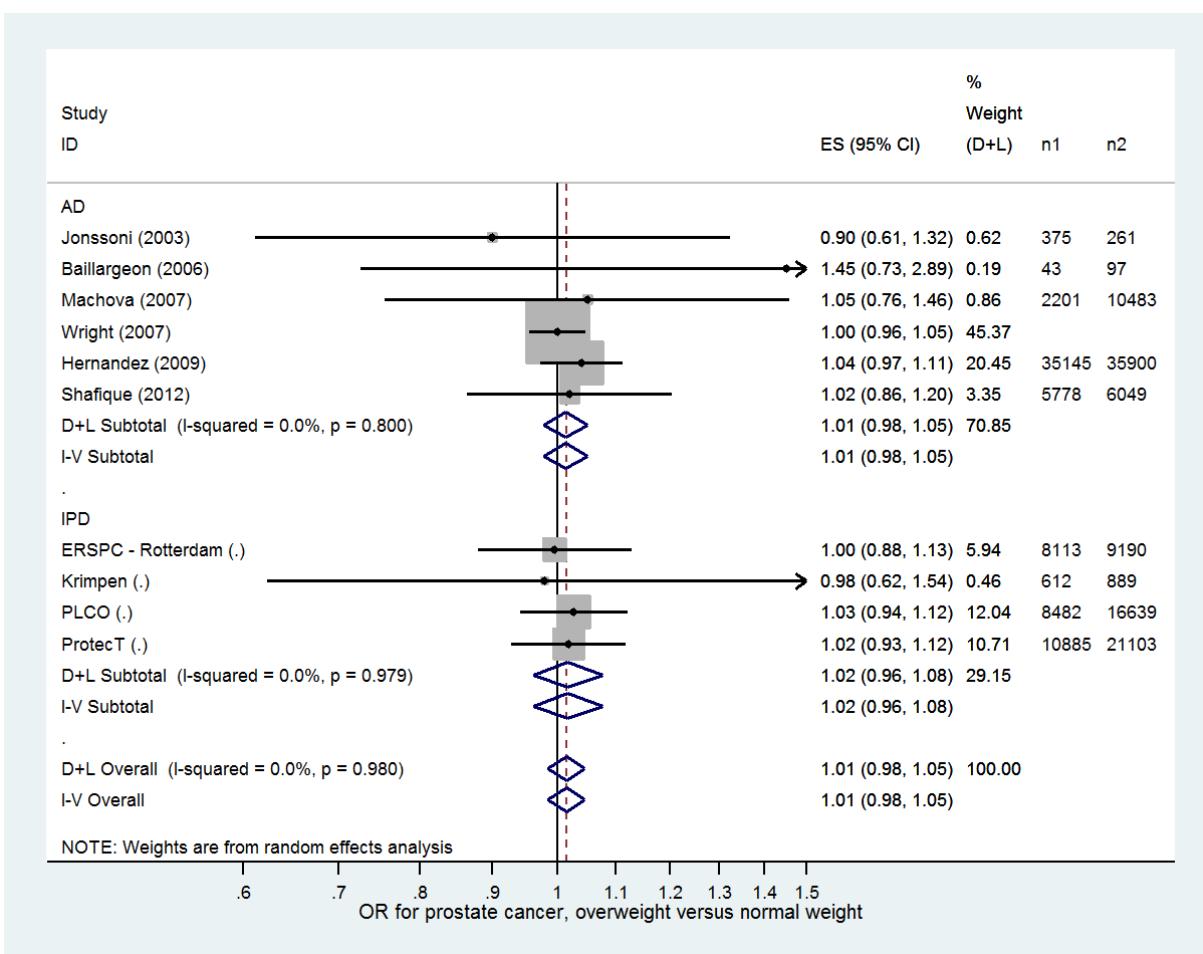


Figure 6.6 Forest plot of the OR for prostate cancer for overweight versus normal weight BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants

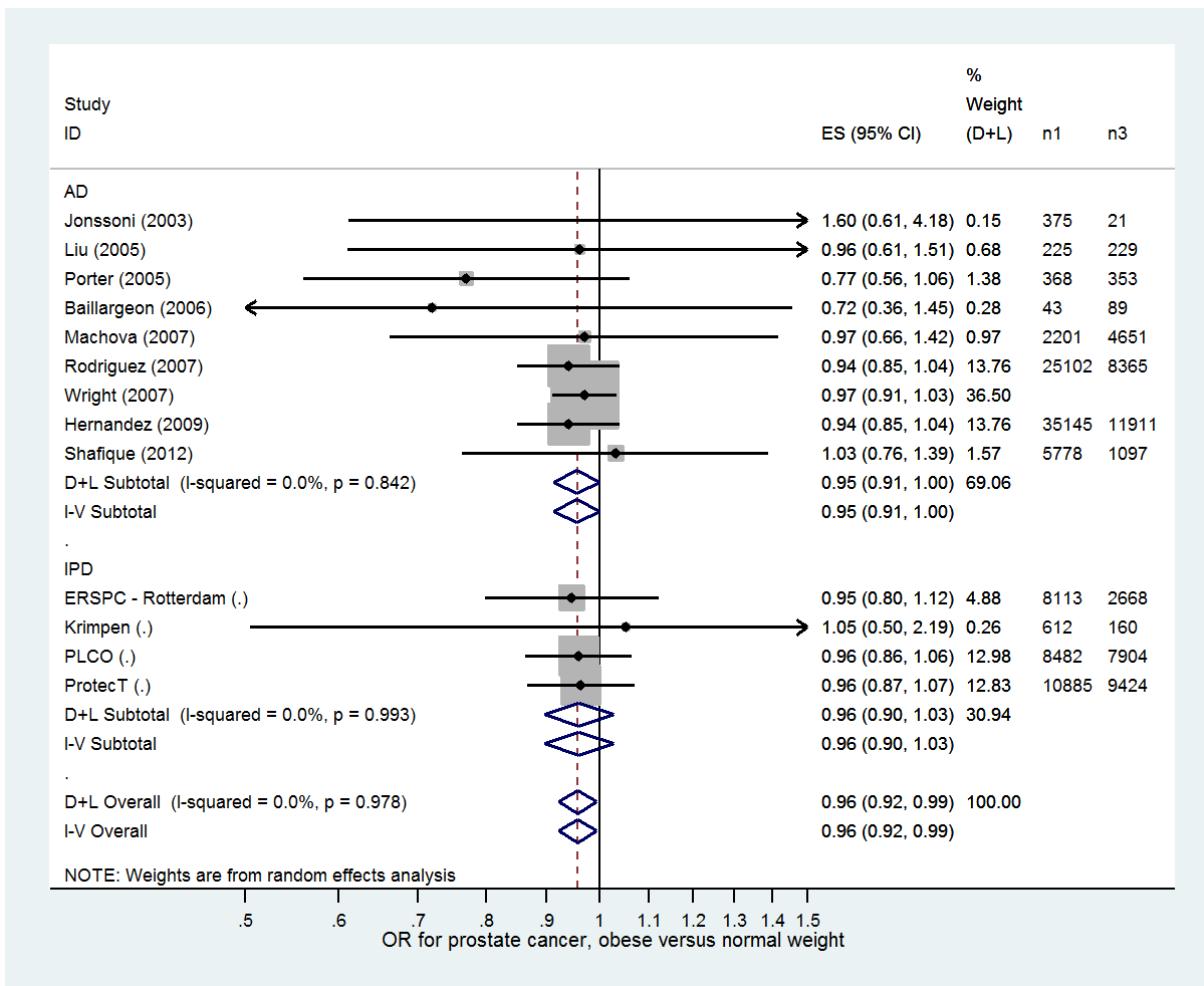


Figure 6.7 Forest plot of the OR for prostate cancer for obese versus normal BMI categories, n1 = number of normal weight participants, n3 = number of obese participants

6.4.2. BMI and advanced prostate cancer

The AD used for this comparison consists of three of 14 studies (37,114,116) from the meta-analysis of BMI and advanced prostate cancer that presented ORs for overweight and obese versus normal weight men. These studies were all prospective studies, and when compared with prospective studies not included in the non-linear analysis, had a slightly higher mean BMI (26.3 kg/m^2 versus 25.7 kg/m^2), and the same pooled linear effect estimate with random-effects meta-analysis (OR for a 5 kg/m^2 increase in BMI = 1.06, 95% CI 1.00 to 1.12, versus OR = 1.06, 95% CI 0.99 to 1.14, P for heterogeneity = 0.33). Only two of the AD studies had presented ORs for advanced prostate cancer for overweight versus normal weight men, whereas all three presented ORs for obese versus normal weight men. Across both the AD and IPD, there were 517,717 participants and 4,359 men had advanced prostate cancer (0.84%). The weighted mean BMI across all studies was 22.9 kg/m^2 for the normal BMI category, 27.3 kg/m^2 for the overweight category, and 32.2 kg/m^2 for the obese category.

Table 6.3 shows the mean BMI, total number of men and number of men with advanced prostate cancer in each category of BMI, **Table 6.4** shows the ORs for advanced prostate cancer for each study for overweight and obese versus normal weight men, with forest plots presented in **Figure 6.8** and **Figure 6.9**. For the random-effects meta-analysis, the average OR for advanced prostate cancer between overweight and normal weight men was estimated to be 1.04 (95% CI 0.95 to 1.13, P = 0.39), with no evidence of inconsistency in the entire set of studies ($I^2 = 0.0\%$, P = 0.98), and between obese and normal weight men was estimated to be 1.08 (95% CI 0.97 to 1.20, P = 0.29), with little evidence of inconsistency in the entire set of studies ($I^2 = 16.2\%$, P = 0.31). Fixed-effect models gave very similar results. There was no evidence for heterogeneity between the effect estimates for the AD and IPD (P = 0.91 and P = 0.16 respectively). The AD results had an estimated average OR of 1.14 (95% CI 1.00 to 1.30, P = 0.05) for obese versus normal weight men, while the IPD results had an estimated average OR of 0.97 (95% CI 0.81 to 1.16, P = 0.73), indicating a possible discrepancy. However, this difference was due to the large effect seen in a single study (37), which also meant the three AD studies were relatively inconsistent with each other ($I^2 = 53.7\%$, P = 0.12).

Given the results overall, and inconsistency in the AD results, there is very little evidence for a non-linear association between BMI and advanced prostate cancer risk.

Table 6.3 Summary BMI and number of men with and without advanced prostate cancer for normal, overweight and obese BMI categories

Author/Study	Normal weight (<25 kg/m ²)			Overweight (25-29.9 kg/m ²)			Obese (≥30 kg/m ²)		
	BMI (kg/m ²)	Cases	Controls	BMI (kg/m ²)	Cases	Controls	BMI (kg/m ²)	Cases	Controls
Aggregate Data									
Rodriguez (37)	22.7	92	23,167				31.5	46	7,809
Wright (116)	22.3	424	NK	27.5	726	NK	32.0	256	NK
Hernandez (114)	23.0	267	32,779	27.2	281	33,438	31.3	77	11,245
Individual Participant Data									
ERSPC-Rotterdam	22.1	196	7,913	27.3	211	8,980	32.2	58	2,611
Krimpen	23.2	6	606	27.0	6	883	31.9	3	157
PLCO	23.1	261	8,221	27.2	552	16,086	33.3	241	7,663
ProtecT	23.1	124	10,761	27.3	238	20,865	33.1	97	9,326
<i>ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study</i>									
<i>*NK = not known</i>									

Table 6.4 Results for the categorical analysis of the association between BMI and advanced prostate cancer

Author/Study	OR for advanced prostate cancer	
	Overweight vs normal (95% CI)	Obese vs normal (95% CI)
Aggregate Data		
Rodriguez (37)		1.54 (1.06 to 2.23)
Wright (116)	1.03 (0.91 to 1.16)	1.14 (0.97 to 1.33)
Hernandez (114)	1.07 (0.88 to 1.3)	0.93 (0.69 to 1.25)
Individual Participant Data		
ERSPC-Rotterdam	0.98 (0.71 to 1.35)	0.93 (0.59 to 1.47)
Krimpen	0.74 (0.16 to 3.33)	1.98 (0.34 to 11.6)
PLCO	1.09 (0.88 to 1.34)	1.00 (0.78 to 1.27)
ProtecT	0.99 (0.75 to 1.29)	0.91 (0.65 to 1.28)
<i>ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study</i>		

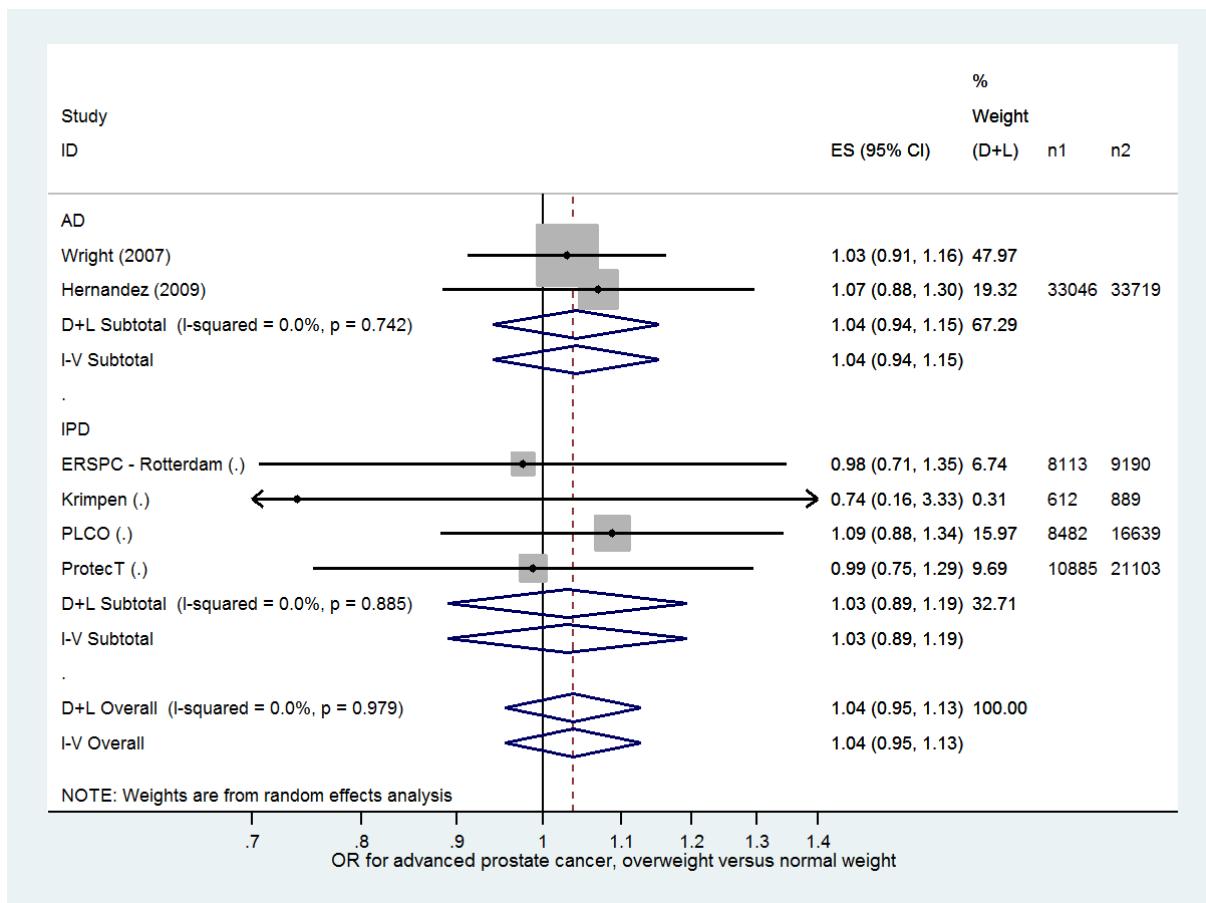


Figure 6.8 Forest plot of the OR for advanced prostate cancer for overweight versus normal weight BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants

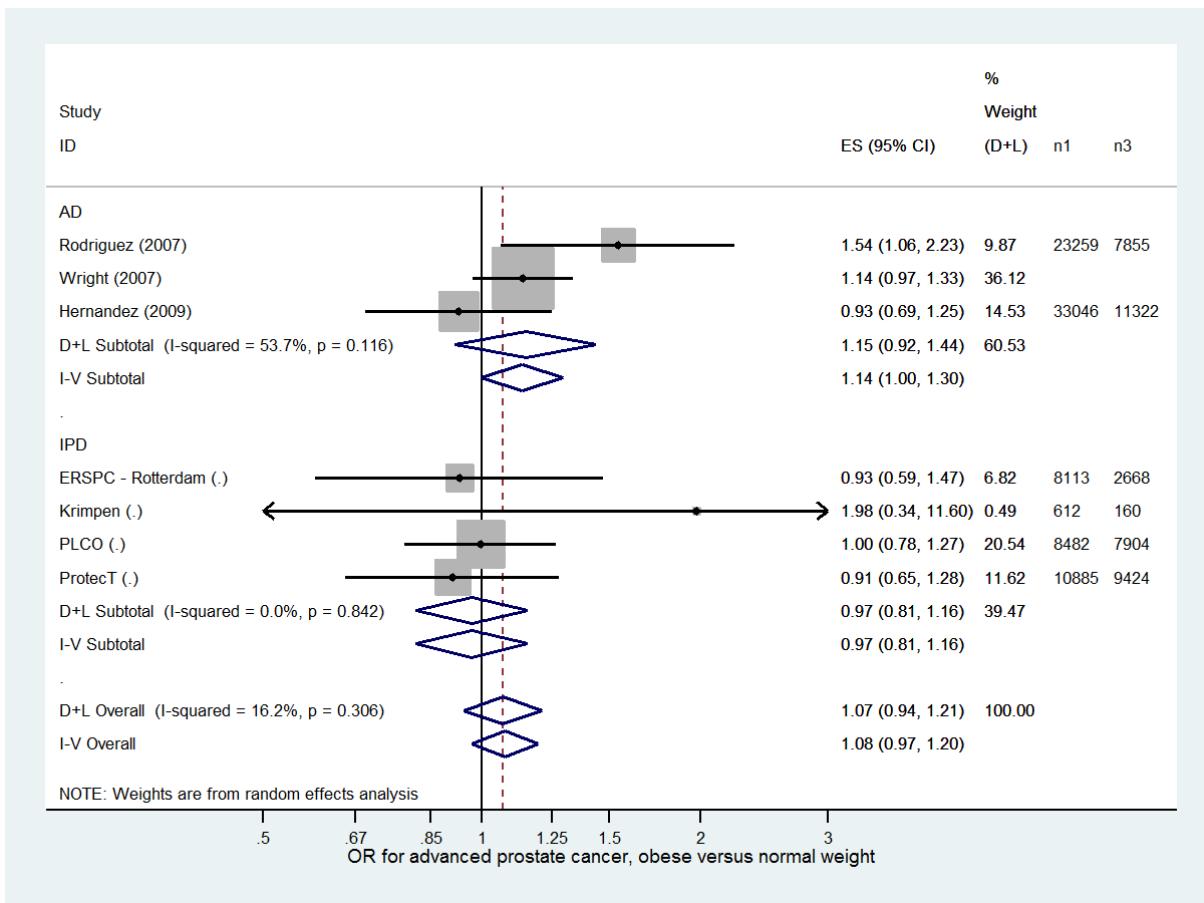


Figure 6.9 Forest plot of the OR for advanced prostate cancer for obese versus normal weight BMI categories, n1 = number of normal weight participants, n3 = number of obese participants

6.4.3. BMI and PSA

The AD used for this comparison consists of eight of ten studies (35,80,82,191,192,194–196) from the meta-analysis of BMI and log-PSA. The mean BMI could not be compared between studies included in the non-linear analysis and the two studies not included, as these two studies did not report the mean BMI. However, the estimated average linear effect estimate between BMI and the percentage change in PSA was very similar between included and excluded studies with random-effects meta-analysis (percentage change in PSA for a 5 kg/m² increase in BMI (included) = -5.18%, 95% CI -7.14% to -3.18%, versus percentage change (excluded) = -5.31%, 95% CI -7.86% to -2.69%, P for heterogeneity = 0.88). Overall, there were 155,991 participants included in this analysis, 59,923 from the AD studies (38.4%) and 96,068 from the IPD studies (62.6%). The weighted mean BMI across all studies was 22.7 kg/m² for the normal BMI category, 27.2 kg/m² for the overweight category, and 32.9 kg/m² for the obese category.

Table 6.5 displays the average log-PSA in each BMI subgroup for all included studies, and **Table 6.6** displays the percentage MD in PSA for all comparisons, with forest plots presented in **Figures 6.10** and **Figure 6.11**. For the random-effects meta-analysis, the average percentage change in PSA between overweight and normal weight men was estimated to be -3.93% (95% CI -5.73% to -2.10%, P < 0.001), with strong evidence of inconsistency in the entire set of studies ($I^2 = 57.0\%$, P = 0.004), and the average percentage change in PSA between obese and normal weight men was estimated to be -11.1% (95% CI -13.4% to -8.72%, P < 0.001), with strong evidence of inconsistency in the entire set of studies ($I^2 = 56.9\%$, P = 0.004). There was no evidence of heterogeneity in the results for the AD and IPD (P = 0.78 and P = 0.16 respectively).

The difference in log-PSA between the obese and normal groups (-0.116) was more than twice the difference between the overweight and normal weight groups (-0.040). Therefore, although the difference in mean BMI between obese and overweight groups (5.7 kg/m²) is larger than between overweight and normal groups (4.5 kg/m²), it appears that PSA decreases more as BMI increases, indicating a non-linear association in both the AD and IPD.

Table 6.5 Summary BMI and log-PSA values for normal, overweight and obese BMI categories

Author/Study	Normal weight (<25 kg/m ²)		Overweight (25-29.9 kg/m ²)		Obese (≥30 kg/m ²)				
	Mean BMI (kg/m ²)	Mean Log-PSA (SD)	N	Mean BMI (kg/m ²)	Mean Log-PSA (SD)	N	Mean BMI (kg/m ²)	Mean Log-PSA (SD)	N
Aggregate Data									
Baillargeon (82)	21.2	-0.41 (0.92)	519	27.8	-0.62 (1.07)	1,318	34.6	-0.49 (0.84)	933
Freedland (191)	21.8	1.57 (0.89)	397	27.7	1.54 (0.9)	684	33.5	1.56 (0.90)	333
Bañez (80) (Duke)	22.6	1.68 (0.66)	452	27.6	1.70 (0.63)	972	32.7	1.54 (0.65)	550
Bañez (80) (Johns Hopkins)	22.7	1.60 (0.59)	2,982	27.4	1.60 (0.59)	5,661	32.0	1.55 (0.60)	1,644
Bañez (80) (SEARCH)	21.8	1.90 (0.56)	357	27.7	1.81 (0.57)	611	33.6	1.76 (0.58)	405
Price (192)	21.7	-0.01 (0.75)	144	27.7	-0.03 (0.72)	247	33.7	-0.23 (0.78)	144
Muller (35)	23.0	-0.12 (0.82)	178	27.5	-0.16 (0.86)	408	31.9	-0.27 (0.90)	191
Park (194)	22.7	-0.14 (0.76)	22,255	26.9	-0.19 (0.76)	15,160	31.0	-0.31 (0.66)	995
Wright (195)	22.6	0.17 (0.82)	219	27.4	0.12 (0.88)	363	32.2	-0.06 (0.95)	188
Bhindi (196)	22.7	1.70 (0.46)	475	27.4	1.71 (0.48)	809	31.9	1.61 (0.51)	329
Individual Participant Data									
ERSPC-Rotterdam	22.1	0.41 (0.95)	8,113	27.3	0.31 (0.92)	9,190	32.1	0.22 (0.91)	2668
Krimpen	23.2	0.32 (0.88)	612	27.0	0.25 (0.90)	889	31.9	0.19 (0.92)	160
PLCO	23.1	0.23 (0.88)	8,482	27.2	0.18 (0.86)	16,639	33.3	0.05 (0.90)	7904
ProtecT	23.1	0.15 (0.89)	10,885	27.3	0.12 (0.91)	21,103	33.1	0.02 (0.92)	9424
ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study									

Table 6.6 Results for the categorical analysis of the association between BMI and PSA

Author/Study	Percentage PSA Mean Difference	
	Overweight vs normal (95% CI)	Obese vs normal (95% CI)
Aggregate Data		
Baillargeon (82)	-19.05 (-27.09 to -10.13)	-7.79 (-15.98 to 1.20)
Freedland (191)	-3.45 (-13.61 to 7.91)	-1.59 (-13.66 to 12.16)
Bañez (80) (Duke)	2.15 (-4.90 to 9.72)	-12.66 (-19.48 to -5.27)
Bañez (80) (Johns Hopkins)	-0.41 (-2.98 to 2.23)	-4.63 (-7.99 to -1.15)
Bañez (80) (SEARCH)	-8.98 (-15.47 to -2.00)	-13.28 (-20.00 to -5.98)
Price (192)	-2.03 (-15.70 to 13.86)	-19.82 (-32.75 to -4.40)
Muller (35)	-4.49 (-17.74 to 10.89)	-14.61 (-28.39 to 1.83)
Park (194)	-4.98 (-6.46 to -3.48)	-16.43 (-20.34 to -12.32)
Wright (195)	-4.24 (-17.07 to 10.58)	-20.34 (-32.92 to -5.39)
Bhindi (196)	0.55 (-4.73 to 6.11)	-8.73 (-14.68 to -2.36)
Individual Participant Data		
ERSPC-Rotterdam	-7.10 (-10.73 to -3.33)	-11.61 (-16.46 to -6.48)
Krimpen	-4.89 (-13.08 to 4.08)	-6.57 (-19.74 to 8.75)
PLCO	-3.51 (-5.79 to -1.18)	-12.66 (-15.08 to -10.17)
ProtecT	-3.53 (-5.51 to -1.51)	-10.99 (-13.15 to -8.78)
ERSPC = European Randomised Study of Screening for Prostate Cancer, PCPT = Prostate Cancer Prevention Trial, PLCO = Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial, ProtecT = Prostate Testing for Cancer and Treatment study		

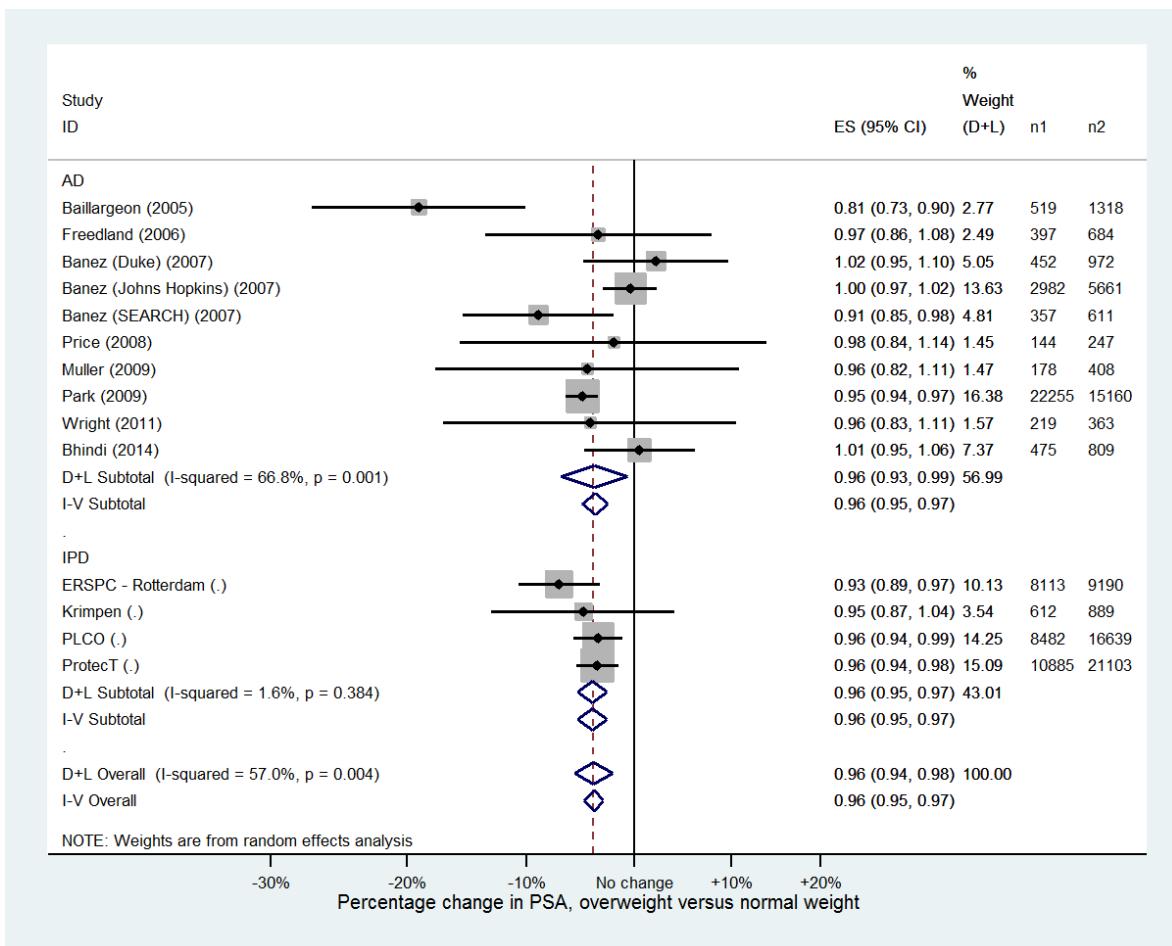


Figure 6.10 Forest plot of the percentage change in PSA between overweight and normal BMI categories, n1 = number of normal weight participants, n2 = number of overweight participants

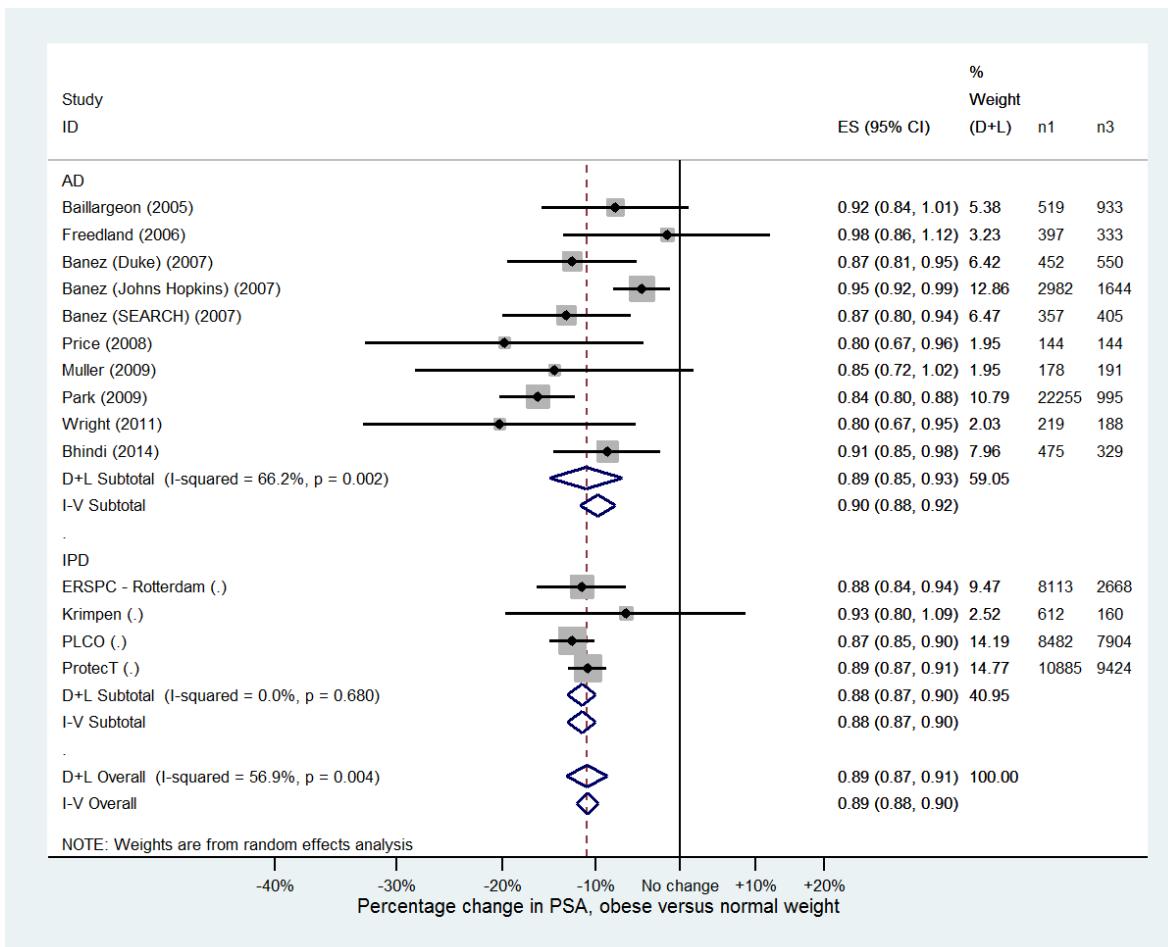


Figure 6.11 Forest plot of the percentage change in PSA between obese and normal BMI categories, n1 = number of normal weight participants, n3 = number of obese participants

6.5. Discussion

This chapter combines both the AD from **Chapter 4** and the IPD from **Chapter 5** to estimate the associations between (1) BMI and prostate cancer, (2) BMI and advanced prostate cancer, and (3) BMI and PSA.

In the linear models using random-effects meta-analysis, estimates of the average effect across all studies for a 5 kg/m² increase in BMI were:

- OR for prostate cancer: 0.99 (95% CI 0.97 to 1.01)
- OR for advanced prostate cancer: 1.04 (95% CI 0.99 to 1.08)
- Percentage change in PSA: -5.62% (95% CI -6.73% to -4.51%)

In the categorical models using random-effects meta-analysis, estimates of the average effect across all studies were:

- OR for prostate cancer:
 - Overweight versus normal weight: 1.01 (95% CI 0.98 to 1.05)
 - Obese versus normal weight: 0.96 (95% CI 0.92 to 0.99)
- OR for advanced prostate cancer:
 - Overweight versus normal weight: 1.04 (95% CI 0.95 to 1.13)
 - Obese versus normal weight: 1.08 (95% CI 0.97 to 1.20)
- Percentage change in PSA:
 - Overweight versus normal weight: -3.93% (95% CI -5.73% to -2.10%)
 - Obese versus normal weight: -11.1% (95% CI -13.4% to -8.72%)

There was no evidence for an association between BMI and prostate cancer in the linear analysis, but some evidence that there was a negative association between obesity and prostate cancer in the categorical analysis, consistent with the non-linear association between BMI and prostate cancer in the more detailed non-linear IPD analysis (**Section 5.5.4**). It is possible that obese men have a reduced risk of prostate cancer, but equally there may be a reduced risk of being diagnosed with prostate cancer due to the negative bias from testing for prostate cancer with PSA. This bias was likely present in the AD studies, and while the imputation of prostate cancer status in the IPD studies reduced the bias, it may not have removed it entirely. In addition, obese men with prostate cancer may have a larger risk of missed diagnoses due to having larger prostates (203), which are associated with less diagnosed prostate cancer (211,249), inducing a negative association with diagnosed prostate cancer rather than true prostate cancer.

There was some evidence of a positive association between BMI and advanced prostate cancer in the linear analysis, but in the categorical analysis the IPD showed very little association compared to the AD. Therefore, the apparent association in the AD may be due to testing for prostate cancer with PSA, or limited number of studies, including the single study with a strong effect (37). If BMI decreases PSA, then testing with PSA will induce a negative association between BMI and prostate cancer, and a potential positive association between BMI and advanced prostate cancer. This is because men with prostate cancer and larger BMIs may have a smaller PSA than comparably smaller men, and thus are not being offered a biopsy until later in their disease progression. Larger prostates increasing the risk of a false negative on biopsy may also contribute to this problem.

There was strong evidence of a negative association between BMI and PSA, which is likely non-linear, decreasing more quickly between overweight and obese than normal weight and overweight, in keeping with the non-linear analysis in **Section 5.5.4**. A recent study assessing the percentage change in PSA for the same BMI categories as this analysis in 15,326 Swedish men in the STHLM-2 study gave very similar results (250); a percentage change in PSA of -3.71% (95% CI -7.18% to -0.24%) between overweight and normal weight men, and -11.7% (95% CI -17.1% to -6.17%) between obese and normal weight men.

6.5.1. Heterogeneity

There were large amounts of inconsistency between AD studies in the linear analyses of BMI and prostate cancer, and advanced prostate cancer. In **Section 4.4.1**, we showed that the between-study heterogeneity decreased substantially if mid-year of recruitment was used in meta-regression. This likely indicates that a proportion of the inconsistency in results may be because of increased PSA testing on prostate cancer diagnosis, or possibly changes in the populations in studies over time.

We hypothesise that the heterogeneity in studies examining the association between BMI and PSA may be due to two main reasons other than population differences. The first reason is that the distribution of BMI in the population will affect the linear effect estimate, since the association between BMI and PSA appears non-linear. Populations with low BMIs may even have positive effect estimates, as the IPD showed that the association between BMI and PSA increased until a point in the normal weight category, while populations with high BMIs would show highly negative effect estimates (**Figure 5.16**). The second reason is that various transformations were required for the AD studies (**Section 4.3.3**) to estimate the percentage change in PSA per unit increase in BMI, and if any of the necessary assumptions were violated bias could be introduced. Specifically, in the Baillargeon (82) study mean PSA values decreased consistently with BMI, from 1.01 ng/ml in normal weight men

to 0.95 ng/ml in overweight men, to 0.91 in obese men. However, because the SD of PSA was calculated to be higher in the overweight group, the mean log-PSA values for each category were -0.41, -0.62 and -0.48 respectively, meaning that in the categorical analysis, the Baillargeon study would have a positive effect estimate for the difference in PSA between obese and overweight men. However, the association of BMI and log-PSA may be more meaningful than the association between BMI and PSA given the haemodilution theory of how BMI and PSA are associated, and most studies did not seem to be as affected by the calculation of log-PSA values as the Baillargeon study.

6.5.2. Strengths and Limitations

Overall, the strengths and limitations from **Section 4.5.4** and **Section 5.6.6** remain, since the included studies are drawn from these chapters. The largest strength of this analysis is the number of participants: in the linear analyses, 189,874 men with prostate cancer and 9,303,353 men in total contributed to the association between BMI and prostate cancer; 10,420 men with advanced prostate cancer and 1,149,184 men in total to the association between BMI and advanced prostate cancer; and 159,716 men contributed to the association between BMI and PSA. In the categorical analyses, 50,830 men with prostate cancer and 571,206 men in total contributed to the association between BMI and prostate cancer; 4,425 men with advanced prostate cancer and 517,717 men in total to the association between BMI and advanced prostate cancer; and 155,991 men to the association between BMI and PSA. This means the results are very precise, and generally involve populations from many countries, improving generalisability.

The limitations remain the same as the previous two chapters. Testing for prostate cancer using PSA undoubtedly has induced some bias in the association between BMI and prostate cancer, and while this was accounted for in the IPD, it could not be in the AD, although we suspect that the change in effect estimates over time in the AD studies might reflect changes in PSA testing. Other population level differences in studies in the AD were not accounted for in this analysis, and the populations in the IPD were relatively homogeneous, coming all from developed countries with similar diets and being restricted to white men only.

An additional limitation is that the non-linear analysis was restricted to comparisons of categories of BMI, as only the IPD studies could be analysed using non-linear regression models such as the cubic models fitted in **Section 5.5.4**. AD Studies presenting prostate cancer risk or PSA results for more categories of BMI could in theory have been used to construct a non-linear effect estimate as in **Section 5.4.7**, including BMI, BMI^2 and BMI^3 in the GLST and VWLS estimations of the associations between BMI and prostate cancer or advanced prostate cancer, and BMI and PSA respectively.

However, such models would require fitting four parameters per study and, as such, at least five BMI categories would be required for these models to not be saturated (six categories for studies presenting ORs). This approach was therefore not attempted. In future studies, IPD from all AD studies could be sought so more studies could be included in a non-linear analysis, however, we lacked the time to attempt this.

Finally, we were unable to assess mediation of the association between BMI and PSA through prostate cancer using the AD. The AD studies estimating the association between BMI and PSA did not often account for prostate cancer status, and therefore the results are for the *total* effect estimate of BMI on PSA. In principle, we could have estimated the *indirect* effect of BMI on PSA through prostate cancer by estimating the number of men with prostate cancer in each study (possibly accounting for men who did not receive a prostate biopsy), and using the estimates from the AD and IPD studies for the associations between BMI and prostate cancer, and prostate cancer and PSA. The *direct* effect could then be estimated by subtracting the *indirect* effect from the measured *total* effect (251). However, there are currently no exact methods for assessing mediation using AD with a continuous exposure and binary mediator or outcome, and as BMI had very little association with prostate cancer in either the linear or non-linear analyses, it is very unlikely that being able to account for mediation in the AD studies would have impacted on the results.

6.5.3. Implications

The main implication of this analysis is that PSA levels are lower in obese and overweight men than normal weight men and the association between BMI and PSA is non-linear. Therefore, it could be beneficial to account for BMI, assuming a non-linear association, when considering whether a man may be indicated for prostate biopsy after a PSA test. This could be either by using a non-linear function of the continuous measure of BMI, or by using categories of BMI, although the former would be more precise and better accommodate men with more extreme BMI levels.

As there was little robust evidence of an association between BMI and prostate cancer, and as the slight positive association between BMI and advanced prostate cancer was only present in the AD (which may be biased), there is no evidence to state that increasing or decreasing BMI would alter prostate cancer risk. However, there is possible evidence that testing for prostate cancer with PSA biases any association between BMI and prostate cancer, and therefore any study seeking to estimate the association between any risk factor and prostate cancer must consider if the risk factor is associated with PSA, and either impute prostate cancer status in men without biopsies, or offer biopsies to all men regardless of PSA levels.

6.6. Conclusion

In conclusion, there was some evidence to suggest that obesity is associated with reduced prostate cancer risk, but this is likely biased by how prostate cancer was detected. There was also some evidence that as BMI increased the risk of advanced prostate cancer increased, but this association is also likely biased by how prostate cancer was detected, and the AD and IPD showed inconsistent associations. There was strong evidence for an association between BMI and PSA, which is likely non-linear where PSA decreases more with BMI as BMI increases. It may be beneficial to consider BMI or obesity with the results of a PSA test when determining whether to offer a prostate biopsy, especially in men with very large BMI values where the greatest differences in PSA are likely seen.

Any associations using observational data cannot establish causality. In **Chapter 7** I use Mendelian Randomisation to assess whether any association between BMI and prostate cancer or PSA is likely causal.

CHAPTER 7. USING MENDELIAN RANDOMISATION TO EXAMINE THE CAUSAL EFFECT OF BODY MASS INDEX ON PROSTATE CANCER RISK, ADVANCED PROSTATE CANCER RISK AND PROSTATE-SPECIFIC ANTIGEN

7.1. Aims

The aim of this chapter is to use Mendelian randomisation (MR) to determine whether BMI has a causal effect on prostate cancer risk, advanced prostate cancer risk or PSA, and to examine whether prostate cancer or PSA could have a causal effect on BMI. In the previous three chapters, we found little evidence that BMI was associated with prostate cancer, but strong evidence that BMI was associated with a reduced PSA, likely non-linearly. However, as all evidence was observational, causality cannot be established from these studies alone.

7.2. Background

7.2.1. Problems with observational epidemiology

Observational epidemiology are studies in which participants are observed, with measurements taken of both exposures and outcomes. The aim is generally to find and quantify causal relationships between an exposure and outcome: in health research, the outcome is often a disease (e.g. prostate cancer) or related to a disease (e.g. survival), and the exposure is a variable thought to cause the outcome. Observational studies offer no intervention to the participants, except the collection of information.

One problem with observation epidemiology is *confounding* (**Section 3.2.3**). For example, age is a likely confounder of the association between BMI and prostate cancer, as increased age potentially causes both BMI (252) and higher risk of prostate cancer (9). Therefore, in general the higher the BMI, the higher the age, and thus the higher the risk of having prostate cancer independent of any effect of BMI. **Figure 7.1** shows these associations. While some confounders can be measured and accounted for, unmeasured confounders will almost always exist and affect the results in a possibly unpredictable way.

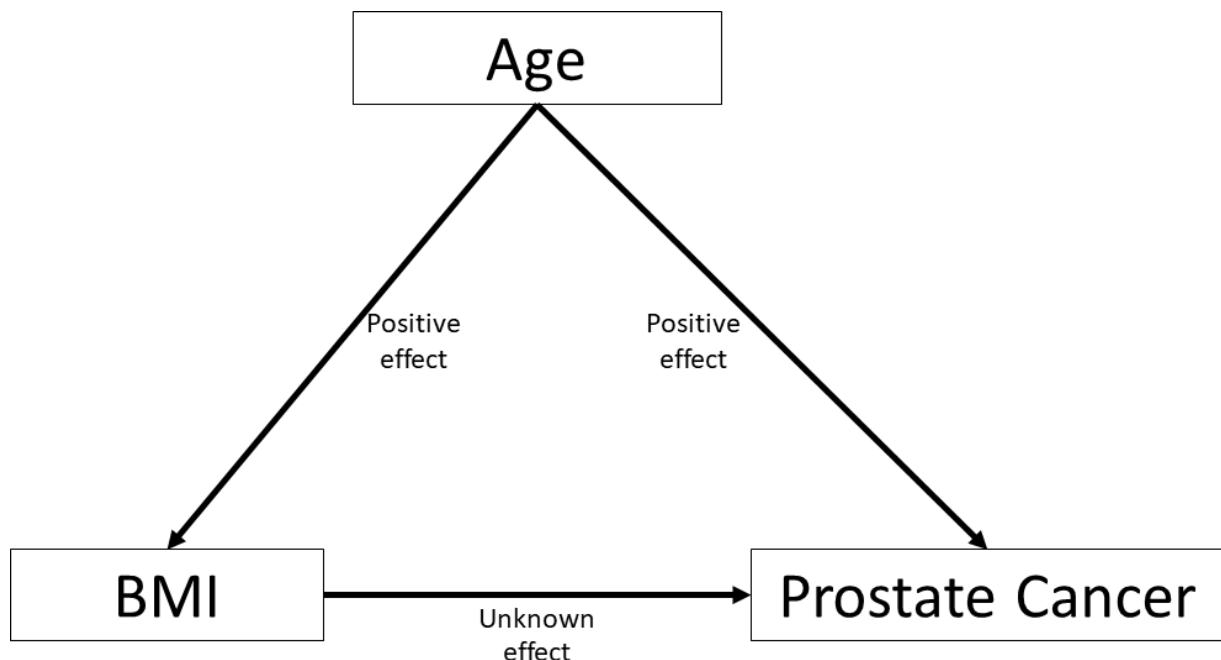


Figure 7.1 The causal associations between age, BMI and prostate cancer, showing confounding by age

A separate problem is *reverse causation*, when the presumed outcome actually causes the presumed exposure, rather than the reverse. In cross-sectional observational studies, it is very difficult to establish causality as it is often impossible to know whether the outcome changed because of the

exposure, or vice versa. For some variables reverse causation could not exist, for example if the exposure is unmodifiable, like age. However, for other variables the supposed outcome could cause the exposure, for example if prostate cancer causes a change in a man's BMI, either biologically or because being tested for prostate cancer may also make a man consciously choose to lose weight.

Often biological plausibility is necessary to state whether an association between variables could be causal; in 1965, Bradford Hill proposed nine criteria to provide evidence of a causal relationship, including the strength, consistency, specificity, temporality, biological gradient, plausibility and coherence of an association, and also including evidence from experiments and analogy (253). For observational studies, temporality could aid in determining causality. Cohort studies measure exposures before measuring the outcome, limiting the possibility of reverse causation (assuming the outcome can be perfectly detected, so that at the beginning of the study no participants had to outcome of interest). However, because unmeasured confounding may still exist, no non-experimental study can prove causation absolutely (254).

The problems of confounding and reverse causality are not found with (properly conducted) RCTs, where participants are randomised into two or more groups to receive different interventions. A direct comparison of interventions is made between randomised groups; with enough participants, randomisation should ensure that differences in both measured and unmeasured confounders between the groups are due only to chance (255). Since the researchers initiate the intervention, there can be no reverse causation. Bias can be introduced if the RCTs are conducted poorly, but otherwise causal attributions can be made from the results of an RCT.

7.2.2. Instrumental variable analysis

RCTs are, however, not always possible. Variables cannot always be randomised for practical (e.g. asking someone to change their home town) or ethical reasons (e.g. asking someone to take up smoking). In these situations, it can be useful to conduct an instrumental variable (IV) analysis (256) to attempt to limit the effects of confounding and reverse causation in estimating the causal effect of an exposure on an outcome.

IV analysis requires an additional variable, called an *instrumental variable* (IV), which has the following assumptions (shown in **Figure 7.2**) (257):

- 1) The IV must be associated with the exposure
- 2) The IV must ONLY be associated with the outcome through the exposure
- 3) The IV must be independent of unobserved confounders that are associated with the exposure and outcome, after conditioning on observed confounders

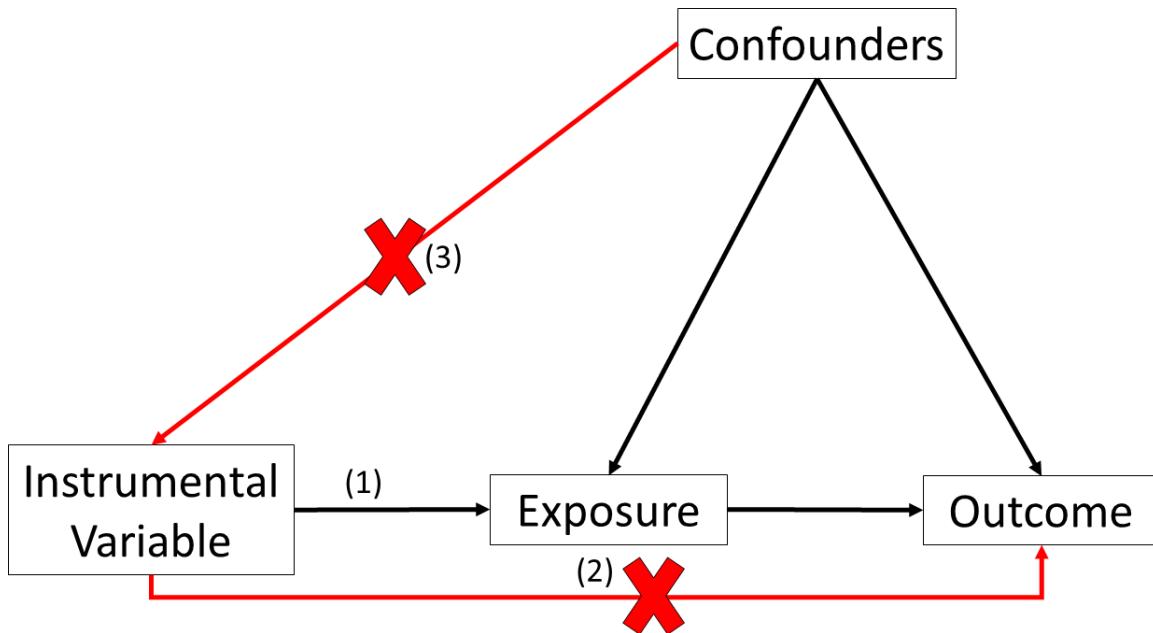


Figure 7.2 Causal diagram of instrumental variable analysis, where the instrumental variable is (1) associated with the exposure, (2) not associated with the outcome except through the exposure, (3) not associated with confounders that are associated with the exposure and outcome

Assuming all three assumptions are satisfied, regression of the outcome against the IV is sufficient to show causality. In practice, however, both the second and third assumptions are untestable, since there is almost an infinite number of variables that could transmit the effect of the IV on the outcome or confound the association (258). Therefore, the IV must be carefully chosen to satisfy all three assumptions as much as possible, and biological plausibility is often necessary in finding IVs in health research.

The effect of the exposure on the outcome can be estimated using a variety of different methods, including two-stage least-squares (2SLS) regression, limited-information maximum likelihood (LIML) and generalised method of moments (GMM) (259)(260). All methods give a causal estimate of the effect of the exposure on the outcome which is unaffected by confounding or reverse causation, so long as the three assumptions hold true.

7.2.3. Mendelian Randomisation

Mendelian Randomisation (MR) provides a biologically plausible method of IV analysis in health research (261). The IVs in MR are genetic variants, natural variations in the genetic code that give rise to observable differences in the exposure variable. Single nucleotide polymorphisms (SNPs), point changes in the genetic code where a single nucleotide is changed for a different nucleotide, are often used as the genetic variant, although other changes to genes could be used (e.g. repeats, additions etc.).

The following properties of genetic variants help to satisfy the assumptions necessary for IV analysis (257,261):

- 1) Genetic variants are often chosen because they have previously been shown in genome-wide association studies (GWAS) to be associated with the exposure
- 2) Genetic variants tend to affect individual genes; if the gene is known to be directly associated with the exposure, then there is less risk of a separate path from the variant to the outcome, although pleiotropy may exist, which is where one genetic position influences more than one phenotype
- 3) Potential confounding is limited with genetic variants, as at conception one variant from each parent was randomly given to the child

MR can be thought of as a natural RCT, where instead of an intervention, individuals are randomised to a certain set of genetic variants at birth, independent of their environment. People with variants that favour, for example, higher BMIs will have a larger BMI on average than those people with variants in favour of a lower BMI. An MR analysis is analogous to an intention-to-treat RCT analysis, where participants are analysed by their randomisation group, regardless of whether they complied with the treatment protocol (261).

The following diagram (**Figure 7.3**) represents an MR using BMI and PSA as an example exposure and outcome respectively. In this hypothetical example, we consider that there may be confounders that cause changes in both BMI and PSA. As such, we decide to use MR, and instead of using BMI as the exposure variable, we use a genetic variant. This genetic variant is believed to (1) be associated with BMI, (2) not be associated with PSA (except through BMI), and (3) is not associated with any of the confounders that cause BMI or PSA.

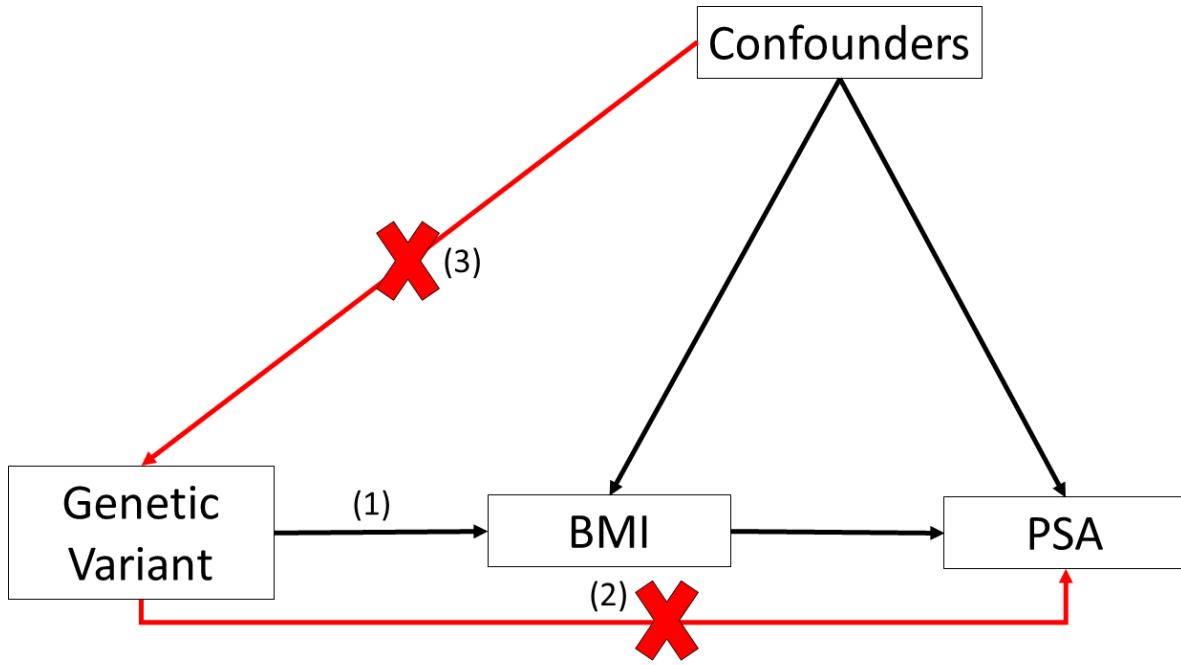


Figure 7.3 Causal diagram of an MR using BMI and PSA as an example, with the same assumptions as in IV analysis

Bidirectional MR

Reverse causation can be tested using MR. In a bidirectional MR, the putative outcome is regressed on a genetic variant for the exposure (to examine whether exposure causes outcome), and the exposure is regressed on a genetic variant for the putative outcome (to examine whether outcome causes exposure) (261). Depending on the associations seen, it is possible to discover whether the exposure causes the outcome or vice versa, or whether there are causal relationships in both directions.

The following diagram (**Figure 7.4**) shows a bidirectional MR focusing on BMI and PSA. First, PSA is regressed on a genetic variant (A) known to be associated with BMI (red line). Second, BMI is regressed on a genetic variant (B) known to be associated with PSA (blue line). We can then look at the relative associations: if there is a strong association between variant A and PSA, then BMI likely causes a change in PSA, and likewise for variant B and BMI. Both associations could be strong, indicating that BMI causes a change in PSA and PSA causes a change in BMI.

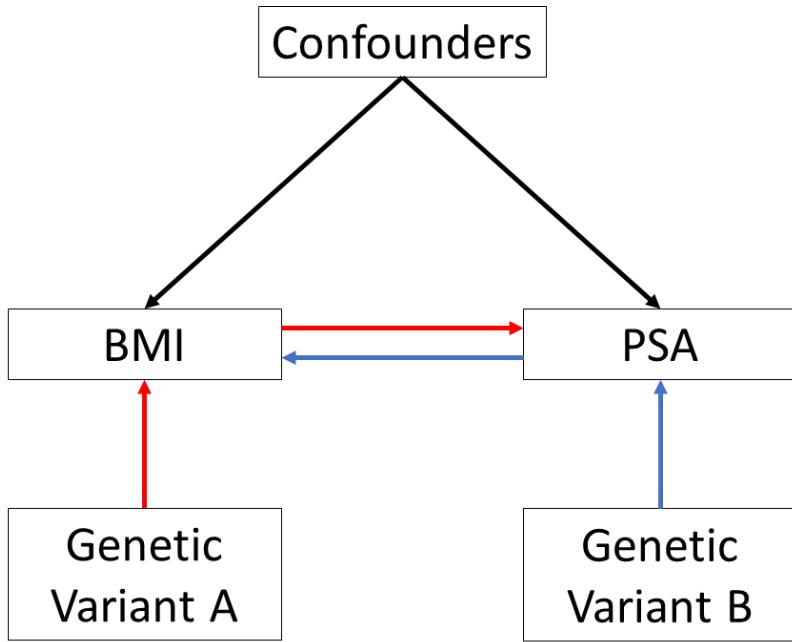


Figure 7.4 Bidirectional MR, where genetic variants associated with BMI and PSA are regressed against PSA and BMI respectively, to discover in which direction(s) an association may flow

7.2.4. Limitations of MR

Statistical power for an MR analysis depends on both the strength of association between the variant and the exposure, and the strength of the association between the exposure and outcome. However, individual genetic variants are rarely strongly associated with a trait. Because the power of an MR analysis is dependent on the strength of the association between the variant and the exposure, MR generally requires more participants to find the same association between exposure and outcome than an analogous observational study (262). Another problem related to the strength of the association between genetic variant and exposure is that if the IV does not explain much of the variability in the exposure (a weak instrument), this may cause bias in small studies (263).

One way to increase statistical power and reduce weak instrument bias, and therefore reduce the required number of participants, is to combine genetic variants into a genetic risk score. Practically, this is often a summation of each risk alleles, where every variant that increases the exposure (or risk of exposure) adds to the total risk score. Each variant can be weighted depending on the strength of association it has with the exposure, so variants with stronger associations have proportionately more weight in the score. By combining variants, the strength of the association with the exposure increases, resulting in larger statistical power (261,264). Combining variants, rather than regressing each variant individually, also reduces the chances of finding associations by chance through multiple

testing. Additionally, summary data from GWAS studies can be used in MR, increasing the potential power of an analysis (265).

Another problem with MR is pleiotropy, which is where one gene may cause many phenotypes (266). This can invalidate the second assumption of IV analysis, as there may be a path from the gene to the outcome that does not go through the exposure. To counter this, genetic variants are often investigated to ensure that it is biologically plausible the variant is both associated with the exposure and *only* associated with the exposure. Pleiotropy can be detected and corrected for using MR-Egger (267), which can be used as a sensitivity analysis to assess the robustness of the results of an MR analysis. Genetic risk scores can also reduce the possibility of pleiotropy. As more variants are added into the risk score, the chances that all variants are pleiotropic and cause the outcome through paths that have the same direction of effect decreases. However, some pleiotropic variants may cause problems if their effects are large.

A further problem with MR is that populations have different levels of genetic variants (population stratification) (268). This can be problematic, as a variant could be associated with a trait not because of any direct effect, but because of a confounding variable, the population the samples were taken from. This invalidates the third assumption of IV analysis, but can usually be accounted for by including principal components in any regression models in the MR (268). Principal components are variables that reflect the genetic structure of the population, and as such will hopefully remove population level confounders.

A final complication of MR is that genetic variants tend to be associated with the exposure across a person's life, and as such indicate long-term levels of exposure (269). Across a population, genetic variants that increase BMI will be associated with higher BMI across all members of the population, regardless of the age. As such, MR cannot determine the effects of short-term changes in exposure; this is beneficial if the exposure is subject to measurement error or changes rapidly, but potentially disadvantageous when considering interventions to alter levels of an exposure, as the short-term effect may not be associated with the long-term effect.

7.3. Methods

The main aim of this analysis was to determine if BMI has a causal effect on prostate cancer risk, advanced prostate cancer risk or PSA, by conducting MR with a genetic risk score for BMI. Causal effect estimates were not calculated as we were only concerned with establishing if BMI causes changes in prostate cancer risk, advanced prostate cancer risk or PSA. Bidirectional MR was used to determine whether any changes in BMI are caused by prostate cancer or PSA.

7.3.1. Selection of instruments

I searched the NHGRI-EBI GWAS Catalog (National Human Genome Research Institute-European Bioinformatics Institute genome-wide association catalogue) for SNPs associated with BMI, PSA or prostate cancer in previous GWAS studies. The Catalog is a quality controlled, manually curated, literature-derived collection of all published genome-wide association studies assaying at least 100,000 SNPs and all SNP-trait associations with P values of less than 10^{-5} (270,271). The chances any individual SNP has an association with the trait by chance are reduced by specifying the P value must be this low.

I generated a list of SNPs that were associated with BMI (up to **March 2016**), with information obtained about the effect allele, and the effect direction and magnitude.

7.3.2. Setting

Participants in this analysis were men of European ancestry from 55 independent studies of prostate cancer included in the prostate cancer association group to investigate cancer associated alterations in the genome consortium (PRACTICAL) (272,273).

Most participants included in PRACTICAL had their blood analysed by one of two genotyping arrays, iCOGS (Illumina Custom Infinium genotyping array) or OncoArray, although some may have been analysed using both arrays (11.4%). The iCOGS array was designed for the Collaborative Oncological Gene-environment Study (COGS) to test genetic variants related to three hormone related cancers, including prostate cancer, and consists of 211,155 SNPs (272)(274). The OncoArray array was designed to evaluate genetic variants for association with the risk of breast, ovarian, prostate, colorectal and lung cancer, and consists of around 600,000 SNPs (275). The two arrays measured different SNPs, so most participants had missing data for some SNPs because some SNPs were only measured in one array. The missing allele dosages of SNPs measured in only one array were replaced with the mean dose across men who were measured by the other array (mean substitution (276)) i.e. any dosages

missing from iCOGS were replaced with the mean dosages from OncoArray and vice versa. This allowed construction of a genetic risk score for all men, and would only bias the analyses if there were systematic differences between the participants measured using the OncoArray and iCogs chips.

I requested all relevant SNPs found in the Catalog from PRACTICAL: of 522 SNPs associated with BMI, PRACTICAL had data for 285; of 18 SNPs associated with PSA, PRACTICAL had data for 10; of 200 SNPs associated with prostate cancer, PRACTICAL had data for 171. The 10 SNPs associated with PSA in PRACTICAL were comprised of 6 SNPs associated with observed PSA, and 4 SNPs associated with log-PSA: the difference was the effect estimate, where a unit change in PSA for SNPs associated with PSA, and a percentage change in PSA for SNPs associated with log-PSA.

From the PRACTICAL data, I excluded participants if their BMI was less than 10 kg/m² (n=1) or more than 100 kg/m² (n=1), since BMI values this extreme are likely errors.

7.3.3. Exposures and outcomes

Genetic risk scores

I constructed genetic risk scores (277) for BMI, prostate cancer, PSA and log-PSA using SNPs previously reported in GWAS to be associated with each variable respectively. To construct the risk scores, the SNPs required reported effect estimates on the same scale and effect alleles from these GWAS, meaning some SNPs were not included in the genetic risk scores. The scales for the effect estimates were, for an increase of 1 effect allele on the outcome: BMI in kg/m²; OR for prostate cancer; PSA in ng/ml; or PSA in percentage change, which was converted to an increase in log-PSA. SNPs reported as a gene-gene interaction, or for related outcomes (e.g. BMI change over time) were excluded from the genetic risk scores. For BMI, the GWAS did not report a combinable effect estimate (e.g. no effect reported or a Z-score reported) or reported an unrelated measure (e.g. obesity-related traits) for 101 SNPs, an effect allele for 11 SNPs, and a unit of BMI for 27 SNPs; in total, 162 SNPs were included in the BMI genetic risk score. For prostate cancer, 71 SNPs did not report an effect size or effect allele, so 100 SNPs were included in the genetic risk score for prostate cancer. For PSA, the GWAS did not report an effect size or effect allele for 3 SNPs, so 3 SNPs were included in the PSA genetic risk score. All 4 SNPs associated with log-PSA were included in the log-PSA genetic risk score. It is worth noting the SNPs associated with PSA (rather than log-PSA) came from only one study of a Korean population with a relatively low number of participants (n=1,575) (278), and thus may not be expected to perform well in Caucasian populations. A full list of SNPs included in the genetic risk score, how many men had information on dosage, effect allele, minor allele frequency, GWAS effect allele, GWAS effect size is available in **Appendix 8**.

Allele dosages were coded as 0, 1 and 2 to represent the number of effect alleles each man possessed, then used to construct the genetic risk scores. The dosages for each variable were summed in each man, with weight equal to the effect estimate of the SNP on the variable in the GWAS. The genetic risk scores are thus a weighted sum of all risk alleles found to be associated with each variable, where the higher the score the more risk alleles a man has.

Every effect estimate for BMI from the original GWAS were kg/m² changes in BMI, therefore assuming all SNPs were independent and every additional effect allele of each SNP caused the same change in BMI, a 1-unit increase in the genetic risk score is roughly equal to a 1-unit increase in BMI. Equally, a unit increase in the risk scores for PSA and log-PSA are equal to a unit increase in PSA and log-PSA respectively. A unit increase in the genetic risk score for prostate cancer is on the log-scale, and would thus be equal to an OR of 2.72.

I regressed the measured values of BMI, PSA, log-PSA and prostate cancer against their respective genetic risk scores, dummy variables for study and assay type, and all principal components to determine how well the genetic risk scores were associated with their phenotypes. This was repeated for BMI, PSA and log-PSA in men in the ProtecT study (31) without prostate cancer (and without the dummy variables for study and assay type) to compute the R² value, the proportion of variation in the phenotypes explained by the genetic risk score. I selected the ProtecT study for this because men without prostate cancer were well-defined, either with a PSA less than 3.0 ng/ml or a raised PSA but no evidence of prostate cancer on biopsy, and by using only a single study, the R² value was not inflated for differences in study or assay type. I calculated the predicted power of the associations between the genetic risk score for BMI and prostate cancer, advanced prostate cancer and PSA, based on the number of participants with data in PRACTICAL, and the R² value for the association between the genetic risk score for BMI and observed BMI in men without prostate cancer in ProtecT (262).

As the two genotyping arrays measured different SNPs, both the genetic risk scores and population characteristics for men measured with iCOGS may have been different to those measured with OncoArray. I included a dummy variable indicating by which array the men were genotyped in all analyses to account for this.

Genetic risk score exclusions

I excluded participants from the SEARCH study (cases = 2,934, controls = 1,454), CAPS study (cases = 1,153, controls = 664), PLCO study (cases = 678, controls = 980) and the Cambridge sub-study of the EPIC study (cases = 27, controls = 194) when creating the BMI risk score, because these studies discovered some of the genetic variants associated with BMI (279). Because most SNPs associated

with prostate cancer and PSA were discovered in PRACTICAL studies, we could not exclude these studies in the same way.

Outcomes

The outcomes for the main analysis were prostate cancer status, advanced prostate cancer status and log PSA levels. In the bidirectional MR, BMI was the outcome variable. Advanced prostate cancer was defined as cancer with a T-stage of 3 or above ($T \geq 3$), an N-stage of 1 ($N=1$) and any evidence of metastases ($M=1$), as in IPD meta-analysis (**Section 5.4.3**). Advanced prostate cancer was only used as an outcome as no reported SNPs were associated with specifically advanced prostate cancer, so a genetic score for advanced prostate cancer was not created.

7.3.4. Statistical analysis

Main analyses

In the first analysis, I regressed log-PSA, prostate cancer status and advanced prostate cancer status separately against the genetic score for BMI, with dummy variables for study and assay type, and principal components as covariates. In the second analysis, the bidirectional MR, I regressed BMI against the genetic scores for prostate cancer risk, PSA and log-PSA separately, with principal components, study and assay type as covariates. In the third analysis, I regressed log-PSA against the genetic score for prostate cancer risk, and in the fourth analysis I regressed prostate cancer status and advanced prostate cancer status against the genetic scores for PSA and log-PSA.

These analyses are summarised below:

1. Prostate cancer status, advanced prostate cancer and PSA regressed separately against the genetic risk score for BMI (examines whether BMI causes prostate cancer, advanced prostate cancer, or PSA)
2. BMI regressed against the genetic risk scores for prostate cancer risk, log-PSA and PSA separately (examines whether prostate cancer or PSA cause BMI)
3. Log-PSA regressed against the genetic risk score for prostate cancer risk (examines whether prostate cancer causes PSA)
4. Prostate cancer status and advanced prostate cancer status regressed separately against the genetic risk scores for log-PSA and PSA (examines whether PSA causes prostate cancer)

All analyses were fixed-effect, assuming no heterogeneity between studies. I assessed whether this assumption was valid by calculating the I^2 metric (57) with inverse-variance fixed-effect meta-analysis.

We assumed that no strong evidence of inconsistency ($I^2 < 50\%$) was sufficient for fixed-effect analyses to be valid.

Sensitivity analysis 1: Assessment of confounding

I regressed measured confounders (age, smoking status, alcohol intake) against all four genetic risk scores in men without prostate cancer to check the assumption the IVs were not associated with confounders. Principal components, chip type and study were all included as covariates in these regressions, and a P values of less than 0.0042 was considered evidence of association to account for multiple tests (4 genetic risk scores with 3 outcomes = 12 tests with Bonferroni correction: $0.05/12 = 0.0042$).

Sensitivity analysis 2: Analysis of all SNPs

I regressed BMI, prostate cancer status and log-PSA against all SNPs associated with BMI, prostate cancer risk, PSA and log-PSA separately, with principal components, study and assay type as covariates. This was to determine if any individual SNPs were associated with any outcome other than the outcome of the GWAS, as individual SNPs may be better determinants of the exposure than others. Missing values of SNPs were left missing, i.e. the mean values of iCOGS were not used if the SNP was not included in OncoArray and vice versa. SNPs were determined to be of interest if the P value for an association they had with any outcome was less than 0.000035, which accounted for multiple comparisons (280) (479 SNPs with 3 outcomes = 1,437 tests with Bonferroni correction: $0.05/1,437 = 0.000035$).

Sensitivity analysis 3: Non-linearity

I looked for evidence of non-linearity in associations between genetic risk scores and all outcomes, as there was some evidence from previous chapters of a possible non-linear association between BMI and prostate cancer, and a likely non-linear association between BMI and log-PSA. I conducted two analyses: in the first, I calculated the square and cube of all risk scores, then used a likelihood ratio (LR) test to compare the likelihood of the linear model in the main analysis with the likelihood of a cubic model. I recorded the P value of the LR test to determine if the non-linear model fit much better than the linear model. In the second analysis, I categorised each genetic risk score into quintiles. For each risk score and outcome, I used an LR test to compare the likelihood of a model using the quintiles as continuous variables and quintiles as categorical variables. Again, I used the P value of the LR test to determine if the non-linear categorical model fit much better than the linear categorical model.

7.4. Results

7.4.1. Clinical characteristics and their associations with genetic risk scores

The clinical characteristics of men in each of the 55 studies contributing to this study are presented in

Table 7.1. There were 97,224 participants, 41,124 men had prostate cancer (42%) and 7,010 men had advanced prostate cancer (7.2%).

The genetic risk score for BMI showed reasonable agreement with observed values of BMI: a unit increase in the BMI genetic risk score gave a 0.54 kg/m^2 increase in BMI (95% CI 0.43 to 0.65 kg/m^2) in all participants, and 0.63 kg/m^2 (95% CI 0.41 to 2.84 kg/m^2 , $r^2 = 1.52\%$) in men without prostate cancer in ProtecT. A unit increase in the prostate cancer genetic risk score gave a 0.30 increase in the log-OR for prostate cancer (95% CI 0.29 to 0.31) in all participants, and 0.35 (95% CI 0.28 to 0.42, pseudo $r^2 = 1.88\%$) in men in ProtecT (both men with and without prostate cancer). A unit increase in log-PSA genetic risk score gave a 0.082 increase in log-PSA (95% CI -0.076 to 0.24) in all participants, and 0.44 increase (95% CI -0.075 to 0.95, $r^2 = 14.8\%$) in men without prostate cancer in ProtecT. A unit increase in PSA genetic risk score gave an 18.6 ng/ml increase in PSA (95% CI -100.6 to 137.5 ng/ml) in all participants, and a 0.48 ng/ml increase (95% CI -0.06 to 1.03 ng/ml, $r^2 = 16.1\%$) in men without prostate cancer in ProtecT.

The power was 15% to detect the percentage change in PSA for a 5 kg/m^2 increase in BMI (-5.62%) seen when combining the AD and IPD studies (**Section 6.3.3**), with 48,233 participants with a recorded PSA, alpha = 0.05 and $r^2 = 1.52\%$ for the association between BMI and the genetic risk score for BMI. For a power over 90%, a 5 kg/m^2 increase BMI would have to cause a -18% change PSA. For a power over 90%, the OR for prostate cancer for a SD increase in BMI (4.1 kg/m^2 in PRACTICAL) would have to be 0.84 or 1.19, with 97,224 participants with 41,124 men with prostate cancer (42%), alpha = 0.05 and $r^2 = 1.52\%$. For a power over 90%, the OR for advanced prostate cancer for a SD increase in BMI would have to be 0.72 or 1.29, with 97,224 participants with 9,606 men with prostate cancer (9.9%), alpha = 0.05 and $r^2 = 1.52\%$.

Table 7.1 Table of clinical characteristics for each study

Study	No PCa (N)	PCa (N)	Advanced PCa (N)	Age* (years)	BMI* (kg/m ²)	PSA ⁺ (ng/ml)	Family History (%)
AHS	1,179	471	0	59.8	27.5		10.8
ATBC	1,913	1,279	163	63.6	26.4		4.0
Aarhus	545	1,076	279	64.3		11	10.2
CAPS	664	1,153	248	66.7	26.3	13	14.4
COH	259	257	50	59.0	28.8	5	25.3
COSM	1,122	2,289	247	67.8	25.6	9.4	14.6
CPCS1	3,014	1,382	103	60.6		12.7	8.2
CPCS2	1,227	710	58	59.5		8.7	14.7
Canary PASS	0	362		62.1			
CeRePP	645	922	206	65.7		10	23.5
EPIC	1,079	722	22	61.7		8.4	
ERSPC	65	71	0	71.2			
ESTHER	320	329	91	65.1	27.6	6.9	3.5
FHCRC	1,116	1,168	0	59.9	27.1	6.3	16.8
Gene-PARE	0	230		66.2		6.3	
HPFS	968	1,101	23	62.8		6.6	100.0
Hamburg-Zagreb	149	146	44	61.7	27.3	12	0.0
IMPACT	860	49	3	53.7	27.7	4.1	13.5
IPO-Porto	246	554	239	54.8		7.4	49.6
KUL	0	2	2	60.9		7	
KULEUVEN	103	164	92	66.9	27.2	9	30.0
LAAPC	280	440	0	66.1			
MALAYSIA	0	1	1	78.4		4373	0.0
MAYO	488	767	332	65.3		7.8	23.3
MCCS	1,485	2,365	143	58.8	26.9	5.4	35.8
MCC_Spain	395	518	36	66.9	27.6	7.6	12.5
MDACC_AS	0	483		64.7		4	23.0
MEC	902	909	0	67.4	26.2		10.5
MOFFITT	285	783	20	64.2	29.1	5	19.6
PCMUS	229	343	144	68.0	27.7	11.5	2.1
PHS	251	614	30	65.9		7	15.2
PLCO	980	678	0	66.1	27.7	5.9	8.8
PPF-UNIS	185	246	49	67.9	26.5	5.5	25.3
PRAGGA	99	129	19	64.6	28.4	10	12.0
PROCAP	236	659	0	65.1		8.3	
PROFILE	20	13	1	54.7	26.5	3.9	84.8
PROGReSS	322	673	67	67.0		9.8	13.2
PROMPT	1	1	0	67.8			
Poland	675	923	45	66.3		12	4.8
ProMPT	13	992	354	64.8		8	64.9
ProtecT	2,848	1,545	59	61.1	26.9	3.1	6.4
QLD	1,271	2,257	10	62.0	28.1	6.4	24.6
RAPPER	0	1,778		70.1			
SEARCH	1,454	2,934	311	60.2	27.3	9.3	17.5
SFPCS	205	279	0	64.8	28.0		15.1
SPAG	170	40	10	65.2		11	80.0
STHLM-1	2,224	2,006	147	66.7			17.1
STHLM-2	1,480	3,016	0	65.5		7	
SWOG-PCPT	1,044	1,063	12	67.0	27.5	2.1	22.2
SWOG-SELECT	2,014	1,451	11	65.4	28.5	4.3	20.5
TORONTO	392	607	0	63.6	27.7	5.9	22.7
UKGPCS	5,100	10,098	2,637	60.6	27.5	5.4	26.1
ULM	357	1,005	358	63.1	27.2	8.4	33.5
UTAH	245	440	0	64.0			33.1
WUGS	0	1,607	344	61.0	29.9	5	31.0
All studies	41,124	56,100	7,010	63.1	27.3	6.2	19.3

PCa = prostate cancer, *Mean, [†]Median

7.4.2. Main analyses

None of the analyses showed strong evidence of inconsistency of the effect estimates across all studies ($I^2 > 50\%$), so I used a fixed-effect model in all analyses, with study- and assay-specific intercepts. Results for the main analyses are shown in **Table 7.2**, and summarised in **Figure 7.5**.

For the first analysis, examining the causal effects of BMI, for a unit increase in the genetic risk score for BMI, the overall OR for prostate cancer was estimated to be 0.95 (95% CI 0.92 to 0.99, $P = 0.014$), the OR for advanced prostate cancer was estimated to be 0.98 (95% CI 0.92 to 1.04, $P = 0.54$), and the percentage change in PSA was estimated to be -0.69% (95% CI -3.69% to 2.40%, $P = 0.66$). When men with prostate cancer were excluded, the percentage change in PSA was estimated to be -0.072% (95% CI -4.87% to 4.96%, $P = 0.98$).

For the second analysis, examining the causal effects of prostate cancer and PSA for the bidirectional MR, for a unit increase in the genetic risk score for: prostate cancer, the overall change in BMI was estimated to be 0.016 kg/m² (95% CI -0.011 to 0.043, $P = 0.25$); log-PSA, the overall change in BMI was estimated to be 0.16 kg/m² (95% CI -0.37 to 0.70, $P = 0.56$); and PSA, the overall change in BMI was estimated to be -0.054 kg/m² (95% CI -0.37 to 0.27, $P = 0.74$).

In the third analysis, examining the causal effect of prostate cancer on PSA, for a unit increase in the genetic risk score for prostate cancer, the overall change in PSA was estimated to be 7.09% (95% CI 6.23% to 7.96%, $P < 0.001$).

In the fourth analysis, examining the causal effects of PSA on prostate cancer and advanced prostate cancer, for a unit increase in the genetic risk score for log-PSA, the overall OR for prostate cancer was estimated to be 1.57 (95% CI 1.28 to 1.91, $P < 0.001$) and the overall OR for advanced prostate cancer was estimated to be 1.30 (95% CI 0.95 to 1.78, $P < 0.001$). For a unit increase in the genetic risk score for PSA, the overall OR for prostate cancer was estimated to be 1.79 (95% CI 1.59 to 2.02, $P < 0.001$) and the overall OR for advanced prostate cancer was estimated to be 0.83 (95% CI 0.69 to 1.00, $P = 0.046$).

Table 7.2 Results for the MR of outcomes (columns) against genetic risk scores (rows), for all men and men without prostate cancer

Risk score	SNPs	Prostate cancer			Advanced prostate cancer			PSA			BMI		
		OR (95% CI)	P value	N	OR (95% CI)	P value	N	% Change (95% CI)	P value	N	Beta (95% CI)	P value	N
Regression with all men													
BMI	162	0.95 (0.92 to 0.99)	0.014	84,677	0.98 (0.92 to 1.04)	0.543	77,847	-0.69 (-3.69 to 2.40)	0.658	44,751	0.54 (0.43 to 0.65)	<0.001	34,806
PCa	100	1.35 (1.34 to 1.37)	<0.001	92,761	1.13 (1.11 to 1.14)	<0.001	84,273	7.09 (6.23 to 7.96)	<0.001	47,636	0.016 (-0.011 to 0.043)	0.252	42,532
Log-PSA	4	1.57 (1.28 to 1.91)	<0.001	92,761	1.30 (0.95 to 1.78)	0.1	84,273	8.60 (-7.36 to 27.32)	0.309	47,636	0.16 (-0.37 to 0.70)	0.555	42,532
PSA	3	1.79 (1.59 to 2.02)	<0.001	92,761	0.83 (0.69 to 1.00)	0.046	84,273	41.87 (28.95 to 56.09)	<0.001	47,636	-0.054 (-0.37 to 0.27)	0.741	42,532
Regression with men without prostate cancer													
BMI	162							-0.072 (-4.87 to 4.96)	0.98	6,831	0.63 (0.49 to 0.76)	<0.001	19,728
Log-PSA	4							43.17 (11.79 to 83.36)	0.004	6,831	0.14 (-0.62 to 0.90)	0.718	19,728
PCa = prostate cancer													
All regressions adjusted for principal components, array type and study													

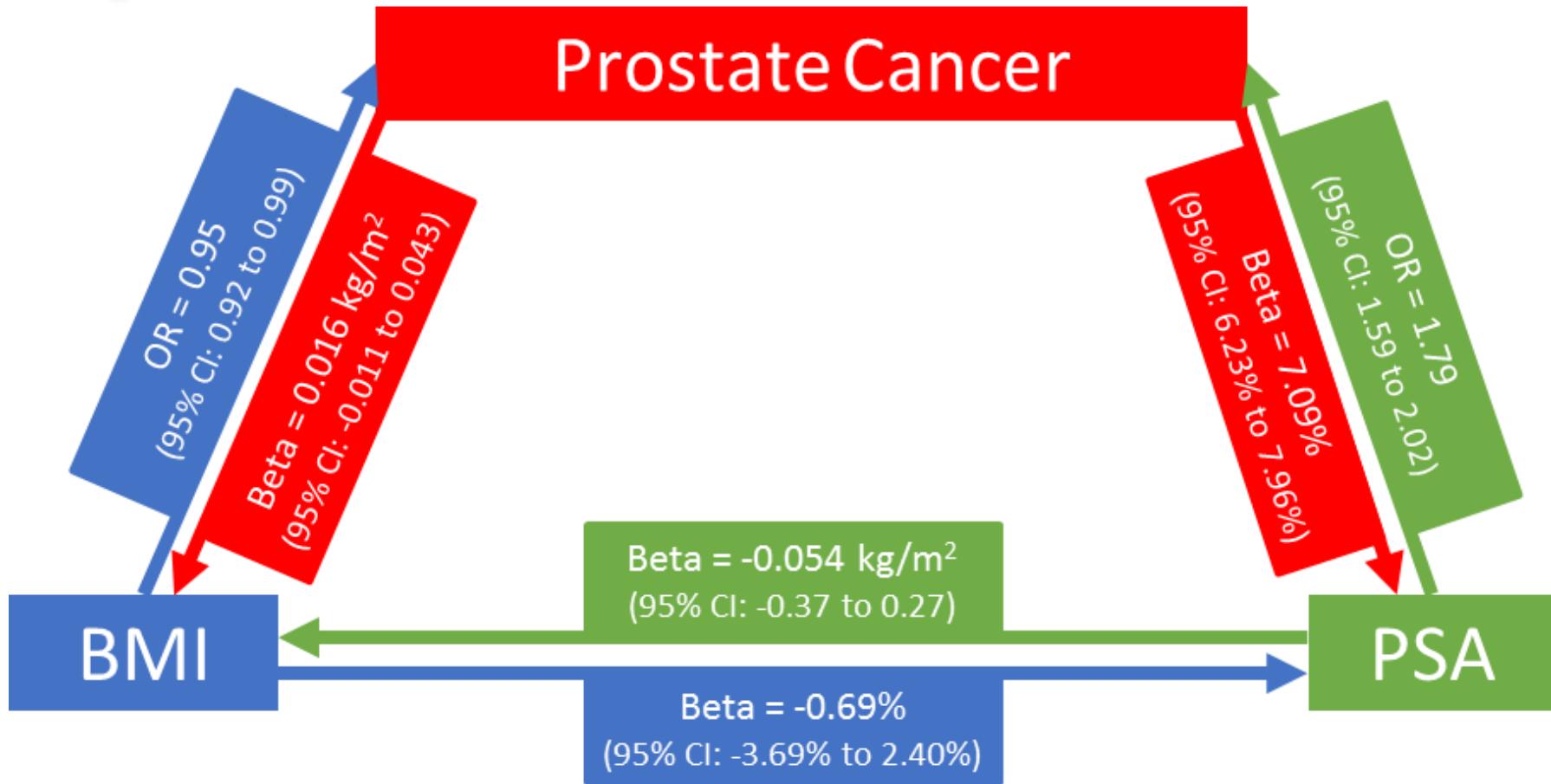


Figure 7.5 Summary results for the main analysis: effect estimates are for a unit increase in genetic risk score of the exposure variable. Results for PSA-BMI and PSA-prostate cancer associations are from the PSA genetic risk score, not the log-PSA genetic risk score

7.4.3. Sensitivity analysis 1: Assessment of confounding

Table 7.3 shows the P values for associations between the genetic risk scores and age, smoking and alcohol. There was a small association between the genetic risk score for prostate cancer and age, which may be explicable by collider bias with PSA, as many of the PRACTICAL studies will diagnosed men with prostate cancer with PSA testing. None of the other genetic risk scores showed a large association with any of the confounders (accounting for multiple tests), increasing the likelihood that the third assumption in **Section 7.2.2**, that the IV is not associated with confounders, although as ever there may still be unobserved confounding.

Table 7.3 Results for the MR of confounders against genetic risk scores for men without prostate cancer only

Risk score	Age		Smoking		Alcohol	
	Effect (95% CI)*	P value	Effect (95% CI)*	P value	Effect (95% CI)*	P value
BMI	0.09 (-0.15 to 0.33)	0.48	0.02 (-0.02 to 0.05)	0.30	-0.03 (-0.15 to 0.09)	0.65
PCa	-0.09 (-0.15 to -0.02)	0.01	0.00 (-0.01 to 0.01)	0.70	-0.01 (-0.04 to 0.02)	0.38
Log-PSA	0.58 (-0.59 to 1.76)	0.33	-0.08 (-0.25 to 0.08)	0.32	0.09 (-0.50 to 0.68)	0.76
PSA	-0.01 (-0.70 to 0.67)	0.97	0.06 (-0.04 to 0.15)	0.26	-0.31 (-0.66 to 0.05)	0.09

PCa = prostate cancer, BMI = body mass index, PSA = prostate-specific antigen
**Effects are the change in outcome for a unit increase in genetic risk score*
All regressions adjusted for principal components, array type and study, and performed only in men without prostate cancer

7.4.4. Sensitivity analysis 2: Analysis of all SNPs

In the sensitivity analysis looking at all SNPs, the SNPs associated with BMI in GWAS that were strongly associated ($P < 0.000035$) with prostate cancer were: rs7830341 (OR = 0.63, 95% CI 0.60 to 0.67, $P < 10^{-40}$, effect allele = G); rs7111341 (OR = 1.13, 95% CI 1.10 to 1.16, $P = 6.1 \times 10^{-23}$, effect allele = A); and rs205262 (OR = 1.06, 95% CI 1.03 to 1.08, $P = 0.000015$, effect allele = G). One SNP approached significance: rs2836754 (OR = 0.95, 95% CI 0.93 to 0.97, $P = 0.00029$, effect allele = G).

The only SNP previously associated with BMI that was associated with PSA was also rs7830341 (percentage change in PSA of -10.0%, 95% CI -14.3% to -5.48%, $P = 0.000025$, effect allele = G).

The SNP rs7830341 is located in a gene called prostate cancer associated transcript 1, and is thus is very likely associated with prostate cancer (and PSA) independently of BMI, breaking assumption 2 of IV analysis. rs205262 is located in the Chromosome 6 Open Reading Frame 106 gene, which has previously been associated with cholesterol (281), height (282), QRS duration and Chagas cardiomyopathy in *Tripanosoma cruzi* seropositivity (283) as well as BMI. rs7111341 is an intergenic

variant that has been associated with type-I diabetes in one other GWAS (284), and rs2836754 is an intron variant that has been associated with Crohn's disease (285).

While rs7830341 does not meet the criteria to be used as an instrument in MR, and as such there is no evidence that BMI causes a change in PSA, three SNPs not previously associated with prostate cancer were shown to be associated with changes in the risk of being diagnosed with prostate cancer.

7.4.5. Sensitivity analysis 3: Non-linearity

There was no evidence of a non-linear association between the genetic risk score for BMI and any outcome, or between the genetic risk scores for prostate cancer, log-PSA or PSA with BMI (**Table 7.4**). Therefore, there is no evidence non-linear associations fit better than linear associations for either the genetic risk score for BMI and all outcomes, or the genetic risk scores of the other variables and BMI.

Table 7.4 Table showing the results of LR tests for each genetic risk score (row) against each outcome (column)

Risk Score	LR Test	PCa		Advanced PCa		PSA		BMI	
		P value	N	P value	N	P value	N	P value	N
BMI	Cubic	0.939	84,677	0.835	77,847	0.809	44,751	0.493	34,806
	Quintiles	0.177	84,677	0.56	77,847	0.278	44,751	0.505	34,806
PCa	Cubic	<0.001	92,761	<0.001	84,273	<0.001	47,636	0.494	42,532
	Quintiles	<0.001	92,761	0.337	84,273	0.008	47,636	0.902	42,532
Log-PSA	Cubic	0.127	92,761	0.061	84,273	0.394	47,636	0.398	42,532
	Quintiles	<0.001	92,761	<0.001	84,273	<0.001	47,636	0.231	42,532
PSA	Cubic	0.471	92,761	0.648	84,273	0.984	47,636	0.866	42,532
	Quintiles	0.571	92,761	0.045	84,273	0.184	47,636	0.342	42,532

PCa = prostate cancer, BMI = body mass index, PSA = prostate-specific antigen.

The P values are from a comparison of models including a cubic and squared term versus linear models, and quintiles as categorical versus continuous variables. A low P value indicates there was evidence the more complex model (cubic model or quintiles as a categorical variable) was better than the less complex model.

7.5. Discussion

In this chapter, I examined whether differences in BMI caused a change in prostate cancer risk or PSA using MR and data from the PRACTICAL consortium.

7.5.1. Summary of Findings

Overall, there was some evidence an increase in BMI protected against prostate cancer ($OR = 0.95$, 95% CI 0.92 to 0.99), and three SNPs previously associated with BMI appear to also be associated with prostate cancer risk. However, there was no evidence for a change in PSA due to BMI, especially when men with prostate cancer were excluded from the analysis, either from the genetic risk score or from individual SNPs. There was no evidence that a change in PSA or prostate cancer caused a change in BMI. There was no evidence that the genetic risk score for BMI had a non-linear relationship with either prostate cancer or PSA.

Given we found a potential causal link between BMI and prostate cancer, but not between BMI and PSA, we could conclude that any association seen between BMI and PSA in observational studies could be explained by the effect of BMI on prostate cancer, and prostate cancer on PSA. If increasing BMI lowers prostate cancer risk, then BMI should also lower PSA even in studies that account for prostate cancer status, as many prostate cancers are undiagnosed. However, as in previous chapters, the small protective effect seen here could be due to testing for prostate cancer with PSA. We found evidence of a causal effect of log-PSA (or PSA) on prostate cancer status, which would indicate either bias due to PSA testing, or pleiotropy.

In our MR analyses, we found a small (but not statistically significant) negative effect of BMI on PSA: the percentage change in PSA for a 1-unit increase in the genetic risk score for BMI was -0.69% , 95% CI -3.69% to 2.40%. Assuming a 1-unit increase in the genetic risk score for BMI was equivalent to a unit increase in BMI, the percentage change in PSA for a 5 kg/m^2 increase in BMI would be -3.40% , 95% CI -17.1% to 12.6%; whereas if a 1-unit increase in the genetic risk score for BMI was equivalent to a 0.69 kg/m^2 increase in BMI (as seen in the association between the score and BMI in men without prostate cancer in ProtecT), a 5 kg/m^2 increase in BMI would be -5.35% , 95% CI -25.8% to 20.7%. These results are roughly consistent with the percentage change seen in the combination AD and IPD studies (**Section 6.3.3**): the percentage change in PSA for a 5 kg/m^2 increase in BMI was -5.62% , 95% CI -6.73% to -4.51%. Therefore, it is possible that with more participants, we would have seen a causal association between BMI and PSA.

7.5.2. Strengths and Limitations

The main strength of this analysis is the use of MR, which overcomes problems of confounding and reverse causation in estimating the causal effect of an exposure on an outcome.

One potential limitation is that two of the underlying assumptions of MR cannot be tested. While there was possibly an association between the prostate cancer genetic risk score and age due to collider bias, we saw no evidence of associations between the other IVs and the measured confounders, increasing our belief that assumption 3 in **Section 7.2.2** is satisfied. We included many participants from a total of 55 studies, which showed little between-study inconsistency in effect estimates, increasing both the generalisability and precision of our results. We were also able to test for non-linear associations between genetic risk scores and outcomes, although we found the linear models performed as well as the non-linear models.

The largest limitation of this analysis is that prostate cancer status was not known for many participants, only whether each man had a diagnosis of prostate cancer. Therefore, if there is a causal effect of BMI on PSA this analysis did not have the power to detect, then in any men who were tested for prostate cancer using PSA the causal effect of BMI on prostate cancer could be biased. Unfortunately, the biopsy status and biopsy protocol for included studies were not known, and so prostate cancer status could not be reliably imputed (as for the IPD in **Chapter 5**). Equally, as with previous chapters, prostate biopsies do not find all cancers, and obesity may increase the risk of missing cancer on biopsy (203).

A further limitation is that case-control studies where a secondary phenotype that is related to the risk of being a case is measured, for example PSA in this analysis, will likely be biased when compared with the whole population (286). This is because the cases and controls will likely be sampled at a different rate to the population. This is exacerbated in this analysis, as PSA testing may directly influence the risk of being diagnosed with prostate cancer. We attempted to remove some of this bias by conducting the regressions involving log-PSA or BMI as outcomes in men without diagnosed prostate cancer, although this is both statistically limited (286), and unable to account for undiagnosed prostate cancers. In addition, while most men with prostate cancer had an observed PSA value (74%), only 17% of men without cancer had a PSA value, and thus the analyses with log-PSA (or PSA) as an outcome were limited to mostly cases and a small number of controls. This may have led to the difference between estimates of the causal effect of BMI on PSA in the entire sample and only in men without prostate cancer. The reason for receiving a PSA test may be important, since men without

prostate cancer who received a PSA test had a lower BMI than those who did not (26.8 kg/m^2 versus 27.3 kg/m^2 , P value for difference < 0.001).

The genetic risk scores lacked statistical power to detect the association between BMI and PSA, as power was 15% to detect a percentage change in PSA of the same magnitude as seen in **Section 6.3.3**. In addition, we assumed that for the effect estimate for all SNPs was additive and independent, so that two copies of the effect allele had twice the effect as one copy. While unlikely to bias the results, if any of the SNP associations were not additive or independent, the genetic risk score would be less associated with the phenotype. Finally, the genetic risk scores for log-PSA and PSA involved few SNPs, and may have been associated with prostate cancer as well as PSA (independently of any effect on prostate cancer through PSA testing), making them poor instruments for PSA.

7.5.3. Comparison with previous studies

In 2015, Davies et al. (287) conducted an analysis of the effects of height and BMI on prostate cancer incidence and mortality with MR using the PRACTICAL consortium dataset. This MR only included men who were genotyped using the iCOGS array, as the study was conducted before men genotyped with the OncoArray assay were included in PRACTICAL datasets. Their results showed an OR for prostate cancer for a standard deviation (SD) increase in BMI genetic risk score (standardised to mean zero SD one) of 0.98 (95% CI 0.96 to 1.00, P = 0.07). On the same scale, our regression showed a consistent OR of 0.98 (95% CI 0.97 to 1.00, P = 0.011), while the increased number of participants likely increased the precision of the estimate (20,848 and 20,214 men with and without prostate cancer in Davies et al., and 51,609 and 40,285 men with and without prostate cancer in this analysis).

Another previous MR by Edwards et al. (134) of only 863 and 876 men with and without prostate cancer in the Nashville men's health study gave an OR of 1.07 for prostate cancer (95% CI 0.91 to 1.25, P = 0.41) for an increase in weighted BMI genetic risk score, which is consistent with our results, given the wide CI.

In 2016, Gao et al. (288) conducted an MR of the effects of adiposity-related traits on breast, ovarian, prostate, lung and colorectal cancer using the Genetic Associations and Mechanisms in Oncology Consortium, comprising 51,537 and 61,600 men with and without prostate cancer. In their analysis, the OR for prostate cancer for their genetic risk score for adult BMI was 1.01 (95% CI 0.84 to 1.21), which is consistent with our estimate. The Gao et al. genetic risk score was comprised of 77 SNPs associated with BMI in a study by Locke et al. (289), all of which we requested from PRACTICAL.

We found three SNPs associated with BMI that appear to be associated with prostate cancer risk that were not previously identified as such. However, the SNP most associated with BMI in Davies' MR (rs1558902) and also shown to have a strong association with prostate cancer risk (OR = 0.97, 95% CI 0.94, 1.00, P = 0.10), appeared to have no association with prostate cancer in our analysis (OR = 1.02, 95% CI 1.00 to 1.04, P = 0.12), despite having a strong association with BMI (change in BMI for each additional effect allele (A) = -0.30 kg/m², 95% CI -0.36 to -0.24, P = 7.6x10⁻²²). The rs1558902 SNP was in linkage disequilibrium (highly correlated) with another SNP (rs9939609) found in a 2010 study by Lewis et al. (290) of 1,550 men with prostate cancer and 1,815 age-matched men without prostate cancer in the ProtecT study. The SNP was associated in Lewis' study with an OR for prostate cancer of 0.93 (95% CI 0.85 to 1.02, P = 0.12).

Davies (287) also reported the effect of a SD increase in the genetic risk score for BMI on advanced prostate cancer risk, with an OR of 1.01 (95% CI 0.97 to 1.05, P = 0.62), consistent with our results. Edwards (134) reported an OR for advanced prostate cancer of 1.04 (95% CI 0.85 to 1.27, P = 0.68), which again is consistent with our results.

We found no MR studies looking at BMI and PSA, although Davies also reported the effect of a SD increase in the genetic risk score for BMI on PSA in men with prostate cancer in ProtecT (N = 828) as - 0.31 ng/ml (95% CI -1.32 ng/ml to 0.70 ng/ml), consistent with our results.

7.5.4. Implications

The largest implication of this analysis is that any MR with prostate cancer as an outcome should consider PSA testing as a potential biasing factor. Although we did not find evidence of an association between BMI and PSA, this may be due to the case-control nature of studies in PRACTICAL, and future MRs should look at cohort studies with measured PSA levels, especially studies where prostate cancer status is well reported, with biopsies potentially performed independently of PSA. Alternatively, imputation could be used if the biopsy status of men was known, or could be predicted based on PSA levels.

7.6. Conclusion

In conclusion, there is no evidence to support a causal effect of BMI on PSA, but some evidence of a protective effect of BMI on the risk of being diagnosed with prostate cancer. However, this may entirely due to testing for prostate cancer with PSA, rather than an association between BMI and prostate cancer risk.

CHAPTER 8. DISCUSSION

8.1. Introduction

In **Chapters 2-7**, I discussed the development of the albatross plot, the results of the Sankey diagram and expert opinion, AD and IPD meta-analyses, and MR. I also discussed the strength and limitations of each of the analyses individually, as well as comparisons with earlier research. In this chapter, I present overviews of all clinical results, compare the results with previous research, assess the methodologies used in this thesis and consider the strengths and limitations of this thesis as a whole. Finally, I discuss the implications of this research, potential future directions this work could take, and conclusions.

8.2. Thesis Summary

The aim of this thesis was to identify plausible individual characteristics that have an association both with prostate cancer and PSA, choose one modifiable characteristic that has been well-studied in relation to both prostate cancer and PSA, and then precisely estimate the associations between this characteristic, prostate cancer, advanced prostate cancer and PSA, accounting for any mediation between the characteristic and PSA through prostate cancer. Presently, PSA tests are used to test for prostate cancer, but PSA lacks sensitivity and specificity for prostate cancer, and is associated with other variables. We believed this research could be helpful in making PSA testing for prostate cancer more sensitive and specific, which would result in fewer unnecessary prostate biopsies and more necessary biopsies.

In **Chapter 1**, I introduced prostate cancer and PSA testing for prostate cancer. I discussed how using PSA to test for prostate cancer lacks sensitivity and specificity, and how improving the test to detect cancer earlier may be beneficial, but may also result in overtreatment (29). I noted how at present, the guidance for PSA testing for most countries is that the decision to have a PSA test should be shared between patient and clinician, weighing the potential benefits against the possible risks (20).

In **Chapter 2**, I introduced the evidence synthesis methods I would use to estimate the associations between an individual characteristic, prostate cancer and PSA. I also developed a method of synthesising evidence with insufficient data for meta-analysis, the albatross plot.

In **Chapter 3**, I used a combination of a Sankey diagram (the use of which appeared novel in epidemiological research), which displayed the number of studies mentioning a list of individual characteristics with prostate cancer and PSA, and expert opinion from lead nurses working on the ProtecT study to choose BMI as the characteristic to be explored in the remaining chapters. In addition to being well-studied, BMI is modifiable, unlike the other characteristics that were well-studied and thought to have strong associations with prostate cancer and PSA. To remove duplicate references that were not entirely identical (e.g. had typographical errors in the titles), I developed a probabilistic matching algorithm. This was necessary to reduce bias in the results of the Sankey diagram, as studies in some journals may be found on multiple databases while others are only found in a single database.

In **Chapter 4**, to estimate the associations between BMI and prostate cancer, advanced prostate cancer and PSA, I conducted a systematic review and meta-analysis of AD studies. I also produced albatross plots to assess whether inclusion of studies without sufficient information for meta-analysis would change our interpretation from the studies with sufficient information. I found no evidence that BMI was associated with prostate cancer risk on average across all studies, although I found that older

studies (those with earlier mid-year of recruitment) tended to estimate a slight positive association, while more recent studies tended to estimate a slight negative association. I noted that this apparent negative association in more recent studies could be due to bias resulting from PSA testing, uptake of which has increased over time. There was some evidence that higher BMI values may be associated with a higher risk of being diagnosed with advanced prostate cancer, but this could also be because of PSA testing for prostate cancer. Finally, I found strong evidence of a negative association between BMI and PSA.

In **Chapter 5** I conducted an IPD meta-analysis of four large prostate cancer studies (three of which screened for prostate cancer using PSA) to estimate all the associations between BMI and prostate cancer, advanced prostate cancer and PSA. I used multiple imputation to account for undiagnosed prostate cancers using a fifth study with relatively unbiased estimates of prostate cancer incidence at low PSA levels. My aim in doing so was to overcome the negative bias in the association between BMI and prostate cancer due to screening that was likely present in many AD studies. I assessed the potential for the associations between BMI and the other variables to be non-linear using continuous linear regression with linear, square and cubic terms for BMI. I found no evidence that BMI was associated with prostate cancer risk or advanced prostate cancer risk, but strong evidence that BMI was negatively associated with PSA, likely non-linearly.

In **Chapter 6** I combined and interpreted together the AD and IPD results, both for linear and non-linear (categorical) associations. There was some evidence that obesity was associated with a lower risk of prostate cancer, but I noted that this could be due to the negative effect of PSA testing for prostate cancer. There was no consistent evidence of a non-linear association between BMI and advanced prostate cancer risk. The continuous non-linear association between BMI and PSA seen in **Section 5.5.4** was also seen in the categorical analysis, with PSA much lower in obese men than overweight or normal weight men. The AD and IPD studies were reasonably consistent in analyses of prostate cancer and PSA, but the small number of AD studies showed a positive association between obesity and advanced prostate cancer risk, while the studies did not.

In **Chapter 7**, I conducted an MR study to assess causality in the associations between BMI and prostate cancer, advanced prostate cancer and PSA. Overall, there was no evidence that a change in BMI caused a change in prostate cancer risk, advanced prostate cancer risk or PSA, or any evidence of prostate cancer or a change in PSA changing BMI. However, the MR was underpowered to detect changes in PSA due to changes in BMI.

The overall results for the associations between BMI, prostate cancer, advanced prostate cancer and PSA are summarised in **Table 8.1**. The main clinical findings are summarised in **Section 8.3**, and the main methodological findings are discussed in **Section 8.4**.

Table 8.1 Summary of overall results

The association between BMI and prostate cancer
Linear random-effects meta-analysis, OR for prostate cancer for a 5 kg/m ² increase in BMI (AD & IPD) or per unit increase in genetic risk score for BMI (MR) estimated to be:
<ul style="list-style-type: none"> • AD: OR = 1.00 (95% CI 0.97 to 1.02, P = 0.71) • IPD: OR = 0.98 (95% CI 0.95 to 1.01, P = 0.16) • AD & IPD: OR = 0.99 (95% CI 0.97 to 1.01, P = 0.57) • MR: OR = 0.95 (95% CI 0.92 to 0.99, P = 0.014)
Non-linear random-effects meta-analysis, OR for prostate cancer between overweight and obese men versus normal weight men estimated to be:
<ul style="list-style-type: none"> • AD & IPD: OR for overweight versus normal weight = 1.01 (95% CI 0.98 to 1.05, P = 0.37) • AD & IPD: OR for obese versus normal weight = 0.96 (95% CI 0.92 to 0.99, P = 0.02)
The association between BMI and advanced prostate cancer
Linear random-effects meta-analysis, OR for advanced prostate cancer for a 5 kg/m ² increase in BMI (AD & IPD) or per unit increase in genetic risk score for BMI (MR) estimated to be:
<ul style="list-style-type: none"> • AD: OR = 1.05 (95% CI 0.99 to 1.10, P = 0.09) • IPD: OR = 1.00 (95% CI 0.93 to 1.08, P = 0.98) • AD & IPD: OR = 1.04 (95% CI 0.99 to 1.08, P = 0.09) • MR: OR = 0.98 (95% CI 0.92 to 1.04, P = 0.054)
Non-linear random-effects meta-analysis, OR for advanced prostate cancer between overweight and obese men versus normal weight men estimated to be:
<ul style="list-style-type: none"> • AD & IPD: OR for overweight versus normal weight = 1.04 (95% CI 0.95 to 1.13, P = 0.39) • AD & IPD: OR for obese versus normal weight = 1.08 (95% CI 0.97 to 1.20, P = 0.29)
The association between BMI and PSA
Linear random-effects meta-analysis, percentage change in PSA for a 5 kg/m ² increase in BMI (AD & IPD) or per unit increase in genetic risk score for BMI (MR) estimated to be:
<ul style="list-style-type: none"> • AD: percentage change = -5.16% (95% CI -6.85 to -3.44%, P < 0.001) • IPD: percentage change = -6.64% (95% CI -7.33 to -5.95%, P < 0.001) • AD & IPD: percentage change = -5.62% (95% CI -6.73 to -4.51%, P < 0.001) • MR: percentage change = -0.69% (95% CI -3.69% to 2.40%, P = 0.66)
Non-linear random-effects meta-analysis, percentage change in PSA between overweight and obese men versus normal weight men estimated to be:
<ul style="list-style-type: none"> • AD & IPD: percentage change for overweight versus normal weight = -3.93% (95% CI -5.73% to -2.10%, P < 0.001) • AD & IPD: percentage change for obese versus normal weight = -11.1% (95% CI -13.4% to -8.72%, P < 0.001)
The association between prostate cancer and PSA
Linear random-effects meta-analysis, percentage change in PSA between men with and without prostate cancer (IPD) or per unit increase in genetic risk score for prostate cancer (MR) estimated to be:
<ul style="list-style-type: none"> • IPD: percentage change = 43.9% (95% CI 40.57% to 47.24%, P < 0.001) • MR: percentage change = 7.09% (95% CI 6.23% to 7.96%, P < 0.001)

8.3. Summary of Clinical Results and Comparison with Previous Research

The clinical results from **Chapter 4** to **Chapter 7** are presented in **Table 8.2** to **Table 8.5** respectively.

BMI and prostate cancer

I found no evidence that BMI was linearly associated with prostate cancer risk in either the AD, IPD or MR. I did find limited evidence that obesity was associated with reduced prostate cancer risk in the AD and IPD, although there was evidence from AD, IPD and MR that testing for prostate cancer with PSA may induce a negative association between BMI and diagnosed prostate cancer, which may account for this negative association. In the AD, the mid-year of study recruitment was associated with a decrease in the OR for the association between BMI and prostate cancer, potentially because the number of men receiving a PSA test also increased over time (**Section 4.4.1**). In the IPD, imputing prostate cancer status in men who were not biopsied removed the negative association seen between BMI and prostate cancer seen in the complete case analysis (**Section 5.5.1**). In the MR, the genetic risk scores for log-PSA and PSA were both associated with prostate cancer risk, implying either an effect of PSA testing for prostate cancer or pleiotropy in the risk scores, or both (**Section 7.4.2**). Therefore, the negative association between obesity and prostate cancer may be entirely due to bias from PSA testing for prostate cancer, and there may be a small positive association between BMI and prostate cancer that was masked by this bias.

Compared to previous meta-analyses of the association between BMI and prostate cancer, our combined AD and IPD meta-analysis included many more studies and participants (**Section 4.5.1**), and was not restricted by study design (e.g. screening studies or prospective studies). The largest previous meta-analysis was conducted by Mazokannas in 2016 (210), including 39 studies and 3,798,746 participants (88,632 men with prostate cancer, 2.3%), whereas we included 62 studies and 9,302,337 participants (188,244 men with prostate cancer, 2.0%). The Mazokannas meta-analysis gave an OR for prostate cancer for a 5 kg/m² increase in BMI of 1.00 (95% CI 0.97 to 1.03), roughly equivalent to our OR of 0.99 (95% CI 0.97 to 1.01). We believe our results represent the best quality evidence for the linear association between BMI and prostate cancer risk to date. Our MR study also gave results that were consistent with previous MR studies (**Section 7.5.3**). To our knowledge, there have been no meta-analyses or MRs of the non-linear association between BMI and prostate cancer risk.

BMI and advanced prostate cancer

There was some evidence that higher BMI values may be associated with a higher risk of being diagnosed with advanced prostate cancer in the AD (**Section 4.4.2**), but this is also likely to be because

of testing for prostate cancer using PSA as the same effect was not seen in the IPD (**Section 5.5.3**). This could have also been a result of the likely heterogeneity in how advanced prostate cancer was defined in the AD studies.

There were a limited number of meta-analyses to compare our results to for the association between BMI and advanced prostate cancer (**Section 4.5.2**). The Mazokannes meta-analysis (210) also examined the association between BMI and advanced prostate cancer, but included high-grade and fatal prostate cancers in the outcome, and therefore included more studies and participants than our meta-analysis: 23 versus 18 studies, 1,676,220 versus 1,149,184 participants, and 11,204 (0.67%) versus 10,354 (0.90%) men with advanced prostate cancer. The RR for advanced prostate cancer for a 5 kg/m² increase in BMI was 1.08 (95% CI 1.04 to 1.12) in the Mazokannes meta-analysis, and the OR was 1.04 (95% CI 0.99 to 1.08) in our meta-analysis. This difference may have been partly caused by the inclusion of high-grade and/or fatal prostate cancers, use of RRs rather than ORs, or from methodological differences in calculation of the mean BMI values in each BMI category. Our MR analysis gave results that were consistent with previous MR studies (**Section 7.5.3**). To our knowledge, there have been no meta-analyses or MRs of the non-linear association between BMI and advanced prostate cancer risk.

BMI and PSA

There was strong evidence in both the AD and IPD (**Section 6.3.3**) that PSA decreases with increasing BMI in Caucasian and Asian men, and that this association is likely non-linear (**Section 5.5.4** and **Section 6.4.3**), with PSA decreasing more rapidly as BMI increases. We found no evidence that this is a causal association, but the MR analysis was underpowered for all analyses (**Section 7.4.2**). However, even without evidence that the association between BMI and PSA is causal, it is likely worth considering BMI when interpreting a PSA test result.

There have been relatively few previous studies looking at the association between BMI and PSA, either with meta-analysis or MR analysis (**Section 4.5.3**). Our results are consistent with the limited previous research, and the conclusion that BMI may have a negative effect on PSA is bolstered by all but one study in the AD meta-analysis showing a negative effect. To our knowledge, no meta-analyses have assessed non-linearity in the association between BMI and PSA.

Table 8.2 Results from Chapter 4, systematic review and AD meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA

BMI and prostate cancer
58 studies, representing 9,252,407 participants and 162,470 men with prostate cancer (1.8%), were included in the meta-analysis, and 68 studies were included in the albatross plot.
<ul style="list-style-type: none"> The random-effects meta-analysis estimated the average OR for a 5 kg/m² increase in BMI to be 1.00 (95% CI 0.97 to 1.02, P = 0.71) There was strong evidence of inconsistency in effect estimates across studies ($I^2 = 67.3\%$, P < 0.001), but this was reduced to 35.8% on meta-regression with mid-year of recruitment There was no evidence that the result varied by whether the BMI was recorded more than 2 years before the average diagnosis of prostate cancer No evidence of small-study effects from the funnel plot The albatross plot showed the 10 studies not included in the meta-analysis were consistent with the meta-analysis results
BMI and advanced prostate cancer
14 studies, representing 1,053,109 participants and 8,357 men with advanced prostate cancer (0.79%), were included in the meta-analysis, and 18 studies were included in the albatross plot.
<ul style="list-style-type: none"> The random-effects meta-analysis estimated the average OR for a 5 kg/m² increase in BMI to be 1.05 (95% CI 0.99 to 1.10, P = 0.09) There was moderate evidence of inconsistency in the effect estimates across studies ($I^2 = 35.1\%$, P = 0.023); no variables included in the meta-regression were strongly associated with the effect estimate There was no evidence that the result varied by whether the BMI was recorded more than 2 years before the average diagnosis of advanced prostate cancer No evidence of small-study effects from the funnel plot The albatross plot showed the four studies not included in the meta-analysis were all positively associated with advanced prostate cancer, however, the two studies with large effect sizes (111,185) were small (N = 831 and 1,492) and we therefore concluded these studies would not change the interpretation had they been included in the meta-analysis
BMI and PSA
10 studies, representing 63,648 men, were included in the meta-analysis, and 13 studies were included in the albatross plot.
<ul style="list-style-type: none"> The random-effects meta-analysis estimated the average percentage change in PSA for a 5 kg/m² increase in BMI to be -5.16% (95% CI -6.85 to -3.44%, P < 0.001) There was strong evidence of inconsistency in the effect estimates across studies ($I^2 = 63.7\%$, P < 0.001); no variables included in the meta-regression were strongly associated with the effect estimate There was some evidence of small-study effects from the funnel plot The albatross plot showed of the three studies not included in the meta-analysis, two were consistent with the meta-analysis (199,207), and one was not (201). This inconsistent study was conducted in an African population rather than the other studies, which were conducted in Asian and Caucasian populations, indicating a possible reason for the inconsistency

Table 8.3 Clinical results from Chapter 5, IPD meta-analysis of associations between BMI, prostate cancer, advanced prostate cancer and PSA

Study Participants
Four large prostate cancer studies were included: ERSPC-Rotterdam (240), Krimpen (239), PLCO (30) and ProtecT (31). Meta-analysis was used to combine linear results of each association for each study, N total = 96,068 and 26,348 men had prostate cancer on imputation (33.7%).
BMI and prostate cancer
<ul style="list-style-type: none"> The fixed-effect meta-analysis estimated the overall OR for a 5 kg/m² increase in BMI to be 0.98 (95% CI 0.95 to 1.01, P = 0.16) There was no evidence of inconsistency in effect estimates across studies ($I^2 = 0.0\%$, P = 0.99) There was little evidence of a non-linear association between BMI and prostate cancer risk, although prostate cancer risk seemed to increase with BMI until a BMI of around 27 kg/m², then decrease. However, the CIs for the association between all BMI terms and prostate cancer from non-linear regression all crossed the null The predicted prostate cancer risk for a man with a BMI of: <ul style="list-style-type: none"> 20 kg/m² = 0.16 (95% CI 0.14 to 0.18) 25 kg/m² = 0.17 (95% CI 0.16 to 0.18) 30 kg/m² = 0.17 (95% CI 0.16 to 0.18)
BMI and advanced prostate cancer
<ul style="list-style-type: none"> The fixed-effect meta-analysis estimated the overall OR for a 5 kg/m² increase in BMI to be 1.00 (95% CI 0.93 to 1.08, P = 0.98) There was no evidence of inconsistency in effect estimates across studies ($I^2 = 0.0\%$, P = 1.00) There was little evidence for a non-linear association between BMI and advanced prostate cancer risk, and the CIs for BMI from non-linear regression all crossed the null The predicted advanced prostate cancer risk for a man with a BMI of: <ul style="list-style-type: none"> 20 kg/m² = 0.019 (95% CI 0.013 to 0.025) 25 kg/m² = 0.020 (95% CI 0.017 to 0.024) 30 kg/m² = 0.021 (95% CI 0.017 to 0.025)
BMI and PSA
<ul style="list-style-type: none"> The fixed-effect meta-analysis for the <i>direct</i> effect of BMI on PSA estimated the overall percentage change in PSA for a 5 kg/m² increase in BMI to be -6.51% (95% CI -7.21% to -5.81%, P < 0.001), and the <i>total</i> effect to be -6.64% (95% CI -7.33% to -5.95%, P < 0.001) There was no evidence of inconsistency in effect estimates across studies for either the <i>direct</i> or <i>total</i> effects ($I^2 = 0.0\%$, P = 0.51 and $I^2 = 0.0\%$, P = 0.50 respectively) There was strong evidence of a non-linear association between BMI and log-PSA, as log-PSA increased with BMI until around 25 kg/m², then fell sharply as BMI increased further. The CIs for the association between all BMI terms and log-PSA from non-linear regression all crossed the null The predicted PSA values for a man with a BMI of: <ul style="list-style-type: none"> 20 kg/m² = 1.16 ng/ml (95% CI 1.13 to 1.19 ng/ml) 25 kg/m² = 1.17 ng/ml (95% CI 1.15 to 1.19 ng/ml) 30 kg/m² = 1.10 ng/ml (95% CI 1.08 to 1.12 ng/ml)
Prostate cancer and PSA
<ul style="list-style-type: none"> The fixed-effect meta-analysis estimated the overall percentage difference in PSA between men with and without prostate cancer to be 43.9% (95% CI 40.6% to 47.2%, P < 0.001) There was strong evidence of inconsistency in the effect estimates across studies ($I^2 = 87.1\%$, P < 0.001)

Table 8.4 Results from Chapter 6, combining AD and IPD to estimate the associations between BMI, prostate cancer, advanced prostate cancer and PSA

Study participants
All studies from Chapter 4 and Chapter 5 were combined in this chapter using standard meta-analysis. Non-linearity was assessed using subgroup meta-analyses.
BMI and prostate cancer
In the linear analysis, I combined 58 AD studies with 4 IPD studies, representing 9,302,337 participants with 188,244 men with prostate cancer (2.0%). In the non-linear analysis, I combined 9 AD studies with 4 IPD studies, representing 571,206 participants with 49,764 men with prostate cancer (8.7%).
<ul style="list-style-type: none"> Random-effects meta-analysis estimated that for a 5 kg/m^2 increase in BMI, the average OR for prostate cancer was a 0.99 (95% CI 0.97 to 1.01, $P = 0.57$) There was strong evidence of inconsistency in effect estimates across studies ($I^2 = 65.8\%$, $P < 0.001$) There was no evidence the AD and IPD gave different effect estimates ($P = 0.09$) Non-linear random-effects meta-analysis estimated the average OR for prostate cancer between overweight and obese men versus normal weight men to be: <ul style="list-style-type: none"> OR for overweight versus normal weight = 1.01 (95% CI 0.98 to 1.05, $P = 0.37$) OR for obese versus normal weight = 0.96 (95% CI 0.92 to 0.99, $P = 0.02$)
BMI and advanced prostate cancer
In the linear analysis, I combined 14 AD studies with 4 IPD studies, representing 1,147,184 participants with 10,354 men with advanced prostate cancer (0.90%). In the non-linear analysis, I combined 3 AD studies with 4 IPD studies, representing 517,717 men participants with 4,359 men with advanced prostate cancer (0.84%)
<ul style="list-style-type: none"> Random-effects meta-analysis estimated that for a 5 kg/m^2 increase in BMI, the average OR for advanced prostate cancer was a 1.04 (95% CI 0.99 to 1.08, $P = 0.09$) There was little evidence of inconsistency in effect estimates across studies ($I^2 = 20.0\%$, $P = 0.22$) There was no evidence the AD and IPD gave different effect estimates ($P = 0.28$) Non-linear random-effects meta-analysis estimated the average OR for advanced prostate cancer between overweight and obese men versus normal weight men to be: <ul style="list-style-type: none"> OR for overweight versus normal weight = 1.04 (95% CI 0.95 to 1.13, $P = 0.39$) OR for obese versus normal weight = 1.08 (95% CI 0.97 to 1.20, $P = 0.29$)
BMI and PSA
In the linear analysis, I combined 10 AD studies with 4 IPD studies, representing 159,716 men. In the non-linear analysis, I combined 8 AD studies with 4 IPD studies, representing 155,991 men
<ul style="list-style-type: none"> Random-effects meta-analysis estimated that for a 5 kg/m^2 increase in BMI, there was an average percentage change in PSA of -5.62% (95% CI -6.73% to -4.51%, $P < 0.001$) There was strong evidence of inconsistency in effect estimates across studies ($I^2 = 62.6\%$, $P < 0.001$) There was evidence the IPD and AD gave different estimates ($P = 0.007$) Non-linear effect random-effects meta-analysis estimated the average percentage change in PSA between overweight and obese men versus normal weight men to be: <ul style="list-style-type: none"> Change for overweight versus normal weight = -3.93% (95% CI -5.73% to -2.10%, $P < 0.001$) Change for obese versus normal weight = -11.1% (95% CI -13.4% to -8.72%, $P < 0.001$)

Table 8.5 Results from Chapter 7, using MR to examine the causal relationships between BMI, prostate cancer risk, advanced prostate cancer risk and PSA

Study participants
The PRACTICAL consortium dataset comprised 97,224 participants with 41,124 men with prostate cancer (42%) from 55 independent studies. SNPs associated with BMI, prostate cancer and PSA in GWAS were found using the NHGRI-EBI GWAS Catalog. Genetic risk scores were created using 162 SNPs for BMI, 100 SNPs for prostate cancer, 3 SNPs for PSA and 4 SNPs for log-PSA.
BMI and prostate cancer
<ul style="list-style-type: none"> A unit increase in the genetic risk score of BMI gave an OR for prostate cancer of 0.95 (95% CI 0.92 to 0.99, $P = 0.014$) A unit increase in the genetic risk score of prostate cancer caused a change in BMI of 0.016 kg/m² (95% CI -0.011 to 0.043 kg/m², $P = 0.25$) There was no evidence of non-linearity in the association between the genetic risk score for BMI and prostate cancer risk
BMI and advanced prostate cancer
<ul style="list-style-type: none"> A unit increase in the genetic risk score of BMI gave an OR for advanced prostate cancer of 0.98 (95% CI 0.92 to 1.04, $P = 0.54$) There was no evidence of non-linearity in the association between the genetic risk score for BMI and advanced prostate cancer risk
BMI and PSA
<ul style="list-style-type: none"> A unit increase in the genetic risk score of BMI caused a percentage change in PSA of -0.69% (95% CI -3.69% to 2.40%, $P = 0.66$) In men without prostate cancer only, a unit increase in the genetic risk score of BMI caused a percentage change in PSA of -0.072% (95% CI -4.87% to 4.96%, $P = 0.98$) A unit increase in the genetic risk score of log-PSA caused a change in BMI of 0.16 kg/m² (95% CI -0.37 to 0.70, $P = 0.56$) A unit increase in the genetic risk score of PSA caused a change in BMI of -0.054 kg/m² (95% CI -0.37 to 0.27, $P = 0.74$) There was no evidence of non-linearity in the association between the genetic risk score for BMI and PSA
Prostate cancer and PSA
<ul style="list-style-type: none"> A unit increase in the genetic risk score of prostate cancer caused a percentage change in PSA of 7.09% (95% CI 6.23% to 7.96%, $P < 0.001$) A unit increase in the genetic risk score of log-PSA gave an OR for prostate cancer of 1.57 (95% CI 1.28 to 1.91, $P < 0.001$) A unit increase in the genetic risk score of PSA gave an OR for prostate cancer of 1.79 (95% CI 1.59 to 2.02, $P < 0.001$)

8.4. Summary of Methodological Findings

In this thesis, I developed novel methods for presenting the results of a systematic review when there was insufficient data for a meta-analysis (the albatross plot), developed a probabilistic matching algorithm for deduplication of references for a systematic review, and used Sankey diagrams with expert opinion to assess plausible targets for systematic reviews. I used standard random-effects meta-analysis to combine AD studies, (homoscedastic) stratified imputation and fixed-effect meta-analysis to combine IPD studies, and MR to assess causality. I also assessed the possibility that all considered associations were non-linear.

Albatross plot

The albatross plot proved useful in assessing whether studies without sufficient information for inclusion in meta-analysis (excluded studies) were consistent with studies with sufficient information (included studies). The plots confirmed that for the association between BMI and prostate cancer risk, the included and excluded studies had consistent effect estimates. The plots also identified two excluded studies that had inconsistent effect estimates for the advanced prostate cancer association, which was likely due to chance from having small numbers of men with advanced prostate cancer. Finally, the plots identified one study with an inconsistent effect estimate for the BMI and PSA association, which could have been due to differences in ethnicity, an important qualification.

Since the development of the albatross plot, these plots have also been used in two published studies. One study examined the association between milk, insulin-like growth factors and prostate cancer risk (291), and one study examined the effect of conditional cash transfers and vouchers on maternity services uptake (292). In both studies, the albatross plots were used because there wasn't sufficient information to conduct meta-analyses, rather than as a sensitivity analysis assessing whether excluded studies were consistent with included studies. When used this way, an approximate effect magnitude can be estimated through visual inspection of the plot and heterogeneity can be assessed, but the estimated effect magnitude should not be reported as a precise estimate as would be estimated using meta-analysis. In addition, although albatross plots require less information than meta-analyses, a P value, total number of participants and effect direction still need to be reported for inclusion in the plot, so some studies may not even have enough information to be included in the albatross plot. However, in that case the only way those studies could be included in an evidence synthesis at all is by contacting the authors for more information, rather than using the published information.

Probabilistic matching algorithm

The probabilistic matching algorithm was both useful in reducing the risk of bias in the Sankey diagram and in saving time when screening titles and abstracts for the AD systematic review. While standard perfect-match deduplication software (such as in Endnote or Ovid) works well in general, the deduplication is limited and doesn't account for typographical errors in references (which were found to be quite common). There may also be other limitations, such as having to manually remove duplicates (Endnote), or restrictions to less than 6,000 references (Ovid). The probabilistic matching algorithm avoided these limitations, and allowed many duplicate studies to be found and removed prior to creation of the Sankey diagram and screening of abstracts in the systematic review. The algorithm has since also been used within the University of Bristol to deduplicate references when reconciling studies uploaded to the university repository (PURE) and from searches of online databases. While the algorithm had a high sensitivity and specificity when used in the assessed subset of studies, some non-duplicates were incorrectly tagged as duplicates. These studies tended to have very similar titles, and were published in the same journals at roughly the same time and as such were relatively rare. In general, the number of false positives should be very low using the algorithm, and the weighted score can be used to identify both perfect duplicates (accounting for inconsistent punctuation, capitalisation etc.) or and highly similar references.

Sankey diagram and expert opinion

The Sankey diagram, combined with expert opinion, proved a helpful tool when deciding which variables to consider in this thesis. The Sankey diagram provided us with a list of variables with many references (and so likely sufficient evidence for meta-analysis), while the expert opinion allowed us to make an informed choice about which variables would likely have strong associations or clinical relevance with the outcomes of interest. If either the Sankey diagram or expert opinion were used in isolation, there would be a risk of choosing variables that were well-studied but not associated with the outcomes, or choosing variables that have been examined in few studies or received a large amount of media interest. The Sankey diagram did not measure the amount of evidence directly, rather the number of studies that have measured both an exposure and outcome. Therefore, it was not known from the Sankey diagram whether the studies relating two variables were all very small, or whether they reported the association between the variables at all, so expert opinion was sought in addition to the Sankey diagram.

Analysis methods

The random-effects meta-analysis of the AD studies allowed study-specific effect estimates, which was appropriate given the evidence of inconsistency between studies seen in all AD analyses. Although

case-control and cohort studies of the association between BMI and prostate cancer and advanced prostate cancer were combined in the analysis, ORs and HRs represent different estimates and are not generally combinable. However, the effect estimates for case-control and cohort studies were very similar, since the combined effect estimates were close to null. Additionally, it was necessary to transform many of the effect estimates to the same scale for combination, increasing the complexity of the meta-analyses.

The (homoscedastic) stratified imputation and meta-analysis of the IPD was appropriate for all analyses involving BMI, as there was little inconsistency in the effect estimates between the IPD studies. While heteroscedastic stratified imputation or even within-study imputation may have been more appropriate when considering the relatively inconsistent association between prostate cancer and PSA between studies, this would have caused difficulties in estimating both the systematically missing prostate cancer statuses of men with PSA levels below the threshold of biopsy of the screening studies, as well as BMI in ERPSC-Rotterdam. Overall, it seems unlikely the results involving BMI would be much changed by using other imputation methods, given the high degree of consistency in the effect estimates between studies (excluding the association between prostate cancer and PSA), as well as the consistency between different meta-analysis methods for all outcomes.

The MR assessed the causal effects between BMI, prostate cancer, advanced prostate cancer and PSA with less risk of bias from confounding than the meta-analyses. However, because the potential bias from testing for prostate cancer with PSA was still present, the results are not necessarily unbiased. This is an important consideration in MR studies; although bias due to confounding may be reduced, study populations or protocols may still introduce bias.

Finally, in both the meta-analyses and MR, I assessed whether the associations considered were potentially non-linear. This was simple to assess in the IPD and MR because full data were available, however the results of the AD studies needed to be reported in specific categories of BMI to be combinable. It was fortunate that most AD studies measuring both PSA and BMI reported mean PSA or log-PSA for standard categories of BMI, as this was the association most likely to be non-linear. However, the AD results were still categorised, and thus the IPD non-linear continuous analysis was much more detailed than the non-linear categorical analysis of both the AD and IPD, allowing us to calculate the expected PSA level for any value of BMI.

8.5. Strengths and Limitations

Strengths

There are many strengths to this thesis. Using a Sankey diagram and expert opinion to systematically identify variables that might be related to both prostate cancer and PSA improved the likelihood of choosing a variable with strong associations and was well-studied. This was enhanced by the deduplication algorithm, which potentially reduced bias in the Sankey diagram. In the IPD, I imputed prostate cancer status to remove bias from PSA screening that likely affected the results of the AD meta-analysis. Albatross plots were devised and used to assess whether studies not included in the meta-analyses would have changed the overall interpretation of the evidence, had they been included, reducing the potential for reporting bias. Non-linearity was assessed in both in the IPD and some AD studies, which was important for the association between BMI and PSA, which appears to be non-linear. The causality of these associations was assessed with MR, which had less potential for bias from confounding than the meta-analyses (although in this analysis lacked power).

Overall, this thesis has combined evidence from different study designs for the associations of BMI, prostate cancer, advanced prostate cancer and PSA, while also investigating bias due to testing for prostate cancer with PSA, and overcoming this bias where possible.

Limitations

There are also many limitations to this thesis. Both the AD and IPD meta-analyses were comprised of observational studies, which may be affected by selection bias and confounding. Confounding was reduced by controlling for age in all analyses and restricting the populations to men of the same ethnicity (white) in the IPD analyses, but unobserved confounding could have remained. In addition, the inclusion of studies from many different populations may reduce the chance of consistent confounding, as distributions of confounders would be different between studies. The MR analysis was likely less subject to bias from confounding, but was underpowered to detect either a causal effect of BMI on prostate cancer risk (for a power over 90%, the OR for prostate cancer for a SD increase in BMI would have to be lower than 0.85 to higher than 1.17) or a causal effect of BMI on PSA of the same magnitude as seen in the combination of AD and IPD in **Section 6.3.3** (power = 15% for a -5.62% change in PSA for a 5 kg/m² increase in BMI).

In the AD studies and MR analysis there was likely bias in the association between BMI and prostate cancer from using PSA as a test for prostate cancer. Indirect evidence for this was shown in the meta-regression, where the ORs for the association between BMI and prostate cancer risk decreased with the mid-year of study recruitment, which may be indicative of the increased uptake of PSA testing

over time. Equally, the positive causal effects of the genetic risk scores for PSA and log-PSA on prostate cancer risk imply that either PSA testing induced bias in the MR, or that the SNPs that comprised the risk scores were also associated with prostate cancer risk (i.e. there was pleiotropy). We attempted to reduce bias from screening in the IPD by imputing prostate cancer status in unbiopsied men, although a large proportion of men were unbiopsied (81%), making it difficult to assess the validity of the imputed values. It is possible that the negative bias induced by testing for prostate cancer using PSA masked a small positive association between BMI and prostate cancer risk, although the IPD showed no association between BMI and prostate cancer even after imputation of prostate cancer status.

In addition to the association between BMI and PSA, other variables also associated with BMI may be associated with the risk of being offered a PSA test. In a recent study, we determined that the risk of receiving a PSA test over a 10-year period was associated with age, index of multiple deprivation (IMD, a measure of socio-economic status (SES)) and region of the UK (21). While age was accounted for in all studies included in our meta-analyses, SES and region were not. BMI is associated with SES both observationally and using MR with Biobank data (293), and BMI varies by region of the UK (294), implying that either of these variables could have biased the association between BMI and prostate cancer by affecting which men were offered biopsies. Other variables may be associated with both BMI and the risk of receiving a PSA test, which, depending on the causality of the relationship, may further bias the estimate of the association between BMI and prostate cancer. These biases would likely not affect screening studies, however, as there was a defined biopsy protocol, although there could be bias in which men were included in the studies.

A separate limitation is that prostate cancer may be missed on biopsy. As the number of samples taken at biopsy increases, the number of cancers diagnosed also increases (83), indicating that some cancers will be missed if a limited number of biopsy cores is taken. There is an association between the size of the prostate and BMI (84), as the volume of the prostate increases with BMI. This means there may be differential classification of prostate cancer in larger men, resulting in bias in the estimation of the association between BMI and prostate cancer. While imputation can be used to estimate whether men who never received a biopsy would have had cancer diagnosed on biopsy, it is impossible to estimate how many men had prostate cancer missed on biopsy without knowing the true prostate cancer status of men.

Another limitation is the relative homogeneity of the IPD study populations, which comprised only white men from high-income countries. Estimates of the associations of BMI, prostate cancer and PSA (especially non-linear associations) should therefore be made in studies with different and varied populations to check for generalisability, since ethnicity has been associated with both prostate cancer

risk (295) and differences in PSA (24). The AD included studies from a relatively large number of countries and ethnicities, making the results more generalisable, although the estimate of the association between BMI and PSA could not be generalised to black populations given the only study including black men had an inconsistent effect estimate. A further limitation of the linear AD meta-analysis of the association between BMI and PSA is that if the association is truly non-linear, then the BMI distribution of each population would affect the effect estimate, with populations with smaller BMIs having a smaller effect estimate compared with populations with larger BMIs. Additionally, in the AD studies advanced prostate cancer was defined by each study independently, likely increasing the heterogeneity of the ORs for the association between BMI and advanced prostate cancer. In the IPD, I defined advanced prostate cancer and the IPD analyses did not suffer from this issue.

We only considered BMI as a measure of adiposity in this study. Although BMI is measured in more studies than other measures and thus has more evidence, other measures of adiposity may find conflicting associations. However, in Markozannes' umbrella study of meta-analyses associating diet, body-size, physical activity and the risk of prostate cancer (210), weight, waist circumference and waist-to-hip ratio were all associated in the same direction with advanced prostate cancer (with no adiposity measures with contradicting evidence), while there was no evidence for an association between any adiposity measures and total prostate cancer risk. In MacInnis' meta-analysis of body size composition and prostate cancer risk (296), the results for both the waist circumference and waist-to-hip ratio meta-analyses were consistent with the BMI meta-analysis, showing slight, non-significant positive associations.

Overall, studies that have previously estimated the associations between BMI, prostate cancer and PSA may have been biased due to PSA testing to determine who receives prostate biopsies. This thesis attempted to limit the effects of bias from screening by imputing prostate cancer status in studies with a defined biopsy protocol. However, because prostate cancer is not always found on biopsy and this may be related to BMI, and because unobserved confounding may exist in all studies, bias may remain. Overall though, estimates from the diverse range of populations and follow-up times included in the meta-analyses will likely be less biased than any individual study.

8.6. Implications of this Thesis

This thesis has estimated the associations between BMI, prostate cancer, advanced prostate cancer and PSA. The association between BMI and PSA is unlikely to be linear, making linear estimates in other studies potentially misleading. Additionally, there may be bias in study estimates for the association between BMI and prostate cancer due to PSA testing for prostate cancer and missed cancers on biopsy. Therefore, future research of any association between BMI and PSA should focus on estimating non-linear associations imputing and accounting for prostate cancer status, ideally using data from studies where the chance of biopsy was not dependent on PSA, and from studies that acquired many biopsy cores (e.g. saturation biopsies). Future studies investigating risk factors for prostate cancer should assess whether the associations measured could be biased by PSA testing for prostate cancer.

The aim of the thesis was to conduct research that could help improve the sensitivity and specificity of PSA testing for prostate cancer. PSA testing for prostate cancer is used frequently around the world; in the UK, men without prostate cancer and aged between 45 and 69 years in 2002 were found to have a 39.2% (95% CI 39.0 to 39.4%) chance of having a PSA test over a 10-year period (21). Given PSA does not have a high sensitivity and specificity for prostate cancer, and that prostate cancer is the fifth leading cause of death from cancer in men (1), there is great scope for improving PSA testing by making it more specific to prostate cancer. For the association between BMI and PSA, this could be accomplished either by using BMI-specific PSA thresholds for biopsy or by calculating a BMI-adjusted PSA level, both of which would account for the negative association between BMI and PSA. Both methods would assume there is no association between BMI and prostate cancer, although if an association is found in the future this could potentially be accounted for. Ideally, both methods would use a non-linear estimate of the association between BMI and PSA, as this appears to represent the association better than a linear estimate.

8.7. Future Directions

There are two primary future directions for this work: the first is accounting for the association between BMI (and other variables such as those found **Section 3.6**: age, ethnicity, sex hormones and BPH) and PSA when testing for prostate cancer; the second is to develop methodology to calculate the *indirect* effect of an exposure on an outcome through a binary mediator, especially when using AD. Future studies could make use of more extensive IPD to precisely estimate the non-linear association between BMI and PSA, which would improve any attempts to make PSA more specific for prostate cancer by adjusting for BMI. In addition, other measures of adiposity could be considered, such as weight and waist circumference.

In addition, the methodologies developed during this thesis could be useful for future research. Sankey diagrams, combined with expert opinion, could be used to select variables to systematically review, for example selecting variables that may be associated with PSA and/or prostate cancer when attempting to improve PSA testing for prostate cancer. Albatross plots could be used in future systematic reviews, either when studies were excluded from meta-analyses due to insufficient information to assess any differences between included and excluded studies, or when meta-analyses were not possible at all to estimate an effect magnitude and assess heterogeneity. The deduplication algorithm may be useful in future systematic and rapid reviews by reducing the number of duplicate references that require screening.

Is an age-BMI-adjusted PSA model clinically useful?

During this PhD, we conducted a study using ProtecT data to determine whether using an age-BMI-adjusted PSA model would improve PSA as a screening test for prostate cancer. I conducted the analyses and was first author on the paper, which was published in *Cancer Causes and Control* at the end of 2016 (81), and is included in full in **Appendix 10**. Co-authors, including Richard Martin, Kate Tilling and Emma Turner, contributed towards the study design and analysis, and all co-authors commented on drafts of the paper.

In brief, we used multivariable linear regression to investigate the associations between log-PSA, age and BMI in 11,293 men, of which 1,836 men had prostate cancer (16.3%). From linear regression, there was a change in PSA of -5.1% for a 5 kg/m² increase in BMI (95% CI -3.4% to -6.8%), and a change in PSA of 13.6% for a 5-year increase in age (95% CI 12.0% to 15.1%). An age- and BMI-adjusted PSA value was estimated for all men using these results, which estimated the PSA the man would have had if they were 61.7 years old with a BMI of 27.2 kg/m² (the sample mean age and BMI respectively). This adjusted PSA value was then compared with the unadjusted PSA and the National Institute for Health

and Care Excellence (NICE) age-adjusted guidelines (18) for detecting prostate cancer. The age-BMI adjusted PSA model performed as well as the NICE guidelines at detecting prostate cancer, and better than the unadjusted model. However, there was little evidence that adjusting PSA for BMI improved the sensitivity or specificity for detecting prostate cancer.

At the time, we had not identified the non-linear association between BMI and PSA, and thus used a linear model of BMI and log-PSA. It is entirely possible the result would have been different if non-linear associations were used, as the sensitivity and specificity would have potentially been made worse for some men and better for others if PSA increased with BMI before it decreased. We also did not impute prostate cancer status for men with a PSA level below 3.0 ng/ml, the threshold for biopsy in ProtecT, because we did not have data from PCPT for imputation. However, we used a theoretical biopsy threshold of 4.0 ng/ml to calculate the sensitivity and specificity of the age-BMI-adjusted PSA model with relative confidence, as ProtecT offered biopsies to all men with a PSA above 3.0 ng/ml, and no PSA values below this threshold were adjusted to above 4.0 ng/ml. There may still have been bias from missing prostate cancer on biopsy more in men with larger BMI values.

A future study could be conducted that used the non-linear association between BMI and log-PSA, rather than assuming a linear association. In addition, imputing prostate cancer status for men without a biopsy, for example by randomly biopsying a small proportion of men with low PSA values to allow for robust imputation, or using a dataset where all men were biopsied independent of PSA, would allow for more accurate sensitivity to be calculated, and for prostate cancer to be controlled more effectively. The sensitivity of the age-BMI-adjusted PSA test in our study was necessarily too high, as there were men who were not biopsied who would have had prostate cancer. Ideally, a future study would involve more participants, and include men from different populations to test for the generalisability of the adjusted PSA for detecting prostate cancer. The future study could also include more variables associated with PSA in the model, adjusting potentially for ethnicity, presence of BPH, drug use (e.g. finasteride) and other variables.

Calculating the indirect effect with a binary mediator using aggregate data

In this thesis, I was unable to estimate the direct effect of BMI on PSA for studies which only measured the total effect. If the mediator (prostate cancer) were continuous, then I could use the coefficient of BMI from linear regression of prostate cancer, and multiply this with the coefficient of prostate cancer from linear regression of log-PSA to calculate the indirect effect of BMI on log-PSA through prostate cancer (297). This is not possible with a binary mediator, as logistic regression gives a coefficient of the change in log-odds of prostate cancer per unit increase in BMI, which cannot be used by itself to

calculate the indirect effect. While the *direct* and *total* effects for BMI and PSA were not too different in the IPD analyses, other variables may give different effect estimates.

There are three problems with using logistic regression coefficients to estimate the *indirect* effect. The first is that logistic regression coefficients are on a different scale to linear regression coefficients. David Kenny suggested using scaling factors to force the coefficients onto the same scale (298). This method ignores the fact that because a logistic regression coefficient changes the outcome multiplicatively, the baseline risk of the mediator affects how large the *indirect* effect would be. Valeri and VanderWeele proposed a method of calculating the *indirect* effect when the exposure was binary (the mediator and outcome can be either binary or continuous) (299) by calculating the change in probability in the mediator for two levels of the exposure. However, the baseline odds must be known for this, and the AD studies did not generally provide information about the baseline odds (only ORs). The second problem is that the Valeri and VanderWeele methods do not allow for a continuous exposure: the indirect effect of BMI on prostate cancer risk when increasing from a BMI from 20 to 21 kg/m² will be different from 29 to 30 kg/m², and so the exposure level of all participants or the distribution of the exposure must be known. The third problem is that there is no established method of using data from different studies to calculate the *indirect* effect (and thus, *direct* effect) from studies that present the *total* effect. Therefore, even if the baseline odds of prostate cancer were to be known for each AD study examining the BMI-PSA association, there is no accepted method of calculating the *indirect* effect using coefficients from other studies.

If methodological advances can be made to estimate the *indirect* effect of an exposure on an outcome when the mediator is binary, and when using AD, then the *direct* effect of BMI or other variables on PSA accounting for any effect of the variables on prostate cancer could be estimated from AD.

8.8. Overall Conclusion

In conclusion, there is no strong, consistent evidence that increased BMI is associated with prostate cancer risk. There is, however, evidence from AD, IPD and MR that PSA testing for prostate cancer may induce a negative association between BMI and diagnosed prostate cancer, and this bias may be masking a small positive association. There is some evidence that higher BMI values may be associated with a higher risk of being diagnosed with advanced prostate cancer, but this is also likely to be because of PSA testing. There was strong evidence that PSA decreases with increasing BMI in Caucasian and Asian men, and that this association is likely non-linear, with PSA decreasing more rapidly as BMI increases. We found no evidence from and MR analysis that this is a causal association, but future studies with greater power and in unselected populations are needed. PSA testing for prostate cancer could be improved through use of BMI-specific PSA thresholds, or by adjusting PSA for BMI levels.

The albatross plots proved useful in assessing whether studies without sufficient information for meta-analysis were different from those that were included, and should be included in systematic reviews where meta-analysis of some/all studies is not possible. The Sankey diagram, when combined with expert opinion, was very helpful in choosing variables to study further. The deduplication algorithm both potentially reduced bias in the Sankey diagram and saved time in the systematic review.

Overall, to improve the precision of the non-linear association between BMI and PSA, more IPD needs to be analysed, ideally in heterogeneous populations. Further studies could identify more variables associated with prostate cancer and/or PSA to develop better tests for prostate cancer, but the studies need to consider whether risk factors for prostate cancer are biased by an association with PSA.

References

1. World Health Organisation International Agency for Research on Cancer. GLOBOCAN 012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 [Internet]. 2015 [cited 2015 Dec 18]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx
2. Cancer Research UK. Prostate cancer incidence statistics [Internet]. 2017 [cited 2017 Jul 17]. Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/incidence>
3. Cancer Research UK. Prostate cancer mortality statistics [Internet]. 2017 [cited 2017 Jul 17]. Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/mortality>
4. National Cancer Institute. SEER Stat Fact Sheets: Prostate Cancer [Internet]. Nih. 2012 [cited 2015 Dec 18]. Available from: <http://seer.cancer.gov/statfacts/html/prost.html>
5. US National Institutes of Health National Cancer Institute. SEER Training Module - Prostate Cancer Morphology & Grade [Internet]. [cited 2017 Jul 17]. Available from: <https://training.seer.cancer.gov/prostate/abstract-code-stage/morphology.html>
6. Muir KR, Lophatananon A, Gnanapragasam V, Rees J. The Future of Prostate Cancer Risk Prediction. *Curr Epidemiol Rep.* 2015;2:251–6.
7. Selley S, Donovan J, Faulkner A, Coast J, Gillatt D. Diagnosis, management and screening of early localised prostate cancer. *Health technology assessment (Winchester, England)*. 1997.
8. Oesterling JE, Jacobsen SJ, Chute CG, Guess HA, Girman CJ, Panser LA, et al. Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges. *JAMA*. 1993;270:860–4.
9. Harrison S, Tilling K, Turner EL, Lane JA, Simpkin A, Davis M, et al. Investigating the prostate specific antigen – body mass index and age relationship: is an age-BMI adjusted PSA model clinically useful? 2016.
10. Oesterling JE, Jacobsen SJ, Klee GG, Petterson K, Piironen T, Abrahamsson PA, et al. Free, Complexed and Total Serum Prostate Specific Antigen: The Establishment of Appropriate Reference Ranges for their Concentrations and Ratios. *J Urol.* 1995;154(3):1090–5.
11. Morgan TO, Jacobsen SJ, McCarthy WF, Jacobson DJ, McLeod DG, Moul JW. Age-specific reference ranges for prostate-specific antigen in black men. *N Engl J Med.* 1996;335(5):304–10.
12. Choi YD, Kang DR, Nam CM, Kim YS, Cho SY, Kim SJ, et al. Age-Specific Prostate-Specific Antigen Reference Ranges in Korean Men. *Urology*. 2007;70(6):1113–6.
13. Koo KC, Park SU, Kim KH, Rha KH, Hong SJ, Yang SC, et al. Predictors of survival in prostate cancer patients with bone metastasis and extremely high prostate-specific antigen levels. *Prostate Int.* 2015;3(1):10–5.
14. Choo R, Klotz L, Danjoux C, Morton GC, DeBoer G, Szumacher E, et al. Feasibility study: watchful

- waiting for localized low to intermediate grade prostate carcinoma with selective delayed intervention based on prostate specific antigen, histological and/or clinical progression. *J Urol.* 2002;167(4):1664–9.
15. Simpkin AJ, Donovan JL, Tilling K, Athene Lane J, Martin RM, Albertsen PC, et al. Prostate-specific antigen patterns in US and European populations: comparison of six diverse cohorts. *BJU Int.* 2016;118(6):911–8.
 16. Adhyam M, Gupta AK. A Review on the Clinical Utility of PSA in Cancer Prostate. *Indian Journal of Surgical Oncology.* 2012. p. 120–9.
 17. Taylor JMG, Park Y, Ankerst DP, Proust-Lima C, Williams S, Kestin L, et al. Real-Time Individual Predictions of Prostate Cancer Recurrence Using Joint Models. *Biometrics.* 2013;69(1):206–13.
 18. NICE. NICE Clinical Knowledge Summary: Prostate Cancer - PSA Testing [Internet]. 2011 [cited 2017 Jul 17]. Available from: <https://cks.nice.org.uk/prostate-cancer#!diagnosissub:2>
 19. Tokudome S, Ando R, Koda Y. Discoveries and application of prostate-specific antigen, and some proposals to optimize prostate cancer screening. *Cancer Management and Research.* 2016. p. 45–7.
 20. Loeb S. Guideline of guidelines: prostate cancer screening. *BJU Int.* 2014;114(3):323–5.
 21. Young G, Harrison S, Turner EL, Walsh EI, Oliver SO, Ben-Shlomo Y, et al. Prostate Specific Antigen (PSA) testing of men in UK general practice: a 10-year longitudinal cohort study. *BMJ Open.*
 22. Man L-B, Li G-Z, Huang G-L, Wang J-W, Liu B-Y. [Aggressiveness and extent of prostatic inflammation relates with serum PSA levels in type IV prostatitis]. *Zhonghua Nan Ke Xue.* 2012;18(8):710–4.
 23. Burton AJ, Martin RM, Donovan JL, Lane JA, Davis M, Hamdy FC, et al. Associations of lifestyle factors and anthropometric measures with repeat PSA levels during active surveillance/monitoring. *Cancer Epidemiol Biomarkers Prev.* 2012;21(10):1877–85.
 24. Hosain GMM, Sanderson M, Du XL, Chan W, Strom SS. Racial/ethnic differences in predictors of PSA screening in a tri-ethnic population. *Cent Eur J Public Health [Internet].* 2011;19(1):30–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21526653>
 25. Safarinejad MR, Asgari SA, Farshi A, Iravani S, Khoshdel A, Shekarchi B. Opium consumption is negatively associated with serum prostate-specific antigen (PSA), free PSA, and percentage of free PSA levels. *J Addict Med [Internet].* 2013;7(1):58–65. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84877987233&partnerID=tZ0tx3y1>
 26. Marberger M, Freedland S, Andriole G, Emberton M, Pettaway C, Montorsi F, et al. Usefulness of prostate-specific antigen (PSA) rise as a marker of prostate cancer in men treated with dutasteride: lessons from the REDUCE study. *BJU Int.* 2012;109(8):1162–9.
 27. Wolf AM, Wender RC, Etzioni RB, Thompson IM, D'Amico A V, Volk RJ, et al. American Cancer Society guideline for the early detection of prostate cancer: update 2010. *CA Cancer J Clin [Internet].* 2010;60(2):70–98. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=20200110

28. Thompson IM, Goodman PJ, Tangen CM, Parnes HL, Minasian LM, Godley PA, et al. Long-term survival of participants in the prostate cancer prevention trial. *N Engl J Med* [Internet]. 2013;369(7):603–10. Available from: <http://www.nejm.org/doi/full/10.1056/NEJMoa1215932#t=article>
29. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet* [Internet]. 2014;384(9959):2027–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25108889>
30. Andriole GL, Crawford ED, Grubb 3rd RL, Buys SS, Chia D, Church TR, et al. Prostate cancer screening in the randomized Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial: mortality results after 13 years of follow-up. *J Natl Cancer Inst* [Internet]. 2012;104(2):125–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22228146%5Cnhttp://jnci.oxfordjournals.org/content/104/2/125.full.pdf>
31. Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: Study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol*. 2014;15(10):1109–18.
32. Rodriguez S, Al-Ghamdi OA, Burrows K, Guthrie PA, Lane JA, Davis M, et al. Very low PSA concentrations and deletions of the KLK3 gene. *Clin Chem* [Internet]. 2013;59(1):234–44. Available from: <http://www.clinchem.org/content/59/1/234.full.pdf>
33. Yang L, Egger M, Plattner R, Klocker H, Eder IE. Lovastatin causes diminished PSA secretion by inhibiting AR expression and function in LNCaP prostate cancer cells. *Urology* [Internet]. 2011;77(6):1508.e1–7. Available from: <http://www.sciencedirect.com/science/article/pii/S0090429511001312>
34. Orsted DD, Bojesen SE, Nielsen SF, Nordestgaard BG, Ørsted DD. Association of clinical benign prostate hyperplasia with prostate cancer incidence and mortality revisited: a nationwide cohort study of 3,009,258 men. *Eur Urol* [Internet]. 2011;60(4):691–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21705134>
35. Muller H, Raum E, Rothenbacher D, Stegmaier C, Brenner H, Mu H. Association of Diabetes and Body Mass Index with Levels of Prostate-Specific Antigen: Implications for Correction of Prostate-Specific Antigen Cutoff Values? *Cancer Epidemiol Biomarkers Prev* [Internet]. 2009;18(5):1350–6. Available from: <http://cebp.aacrjournals.org/cgi/content/abstract/18/5/1350>
36. Alibhai SMH, Krahn MD, Fleshner NE, Cohen MM, Tomlinson GA, Naglie G. The association between patient age and prostate cancer stage and grade at diagnosis. *BJU International*. 2004. p. 303–6.
37. Rodriguez C, Freedland SJ, Deka A, Jacobs EJ, McCullough ML, Patel A V, et al. Body mass index, weight change, and risk of prostate cancer in the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev*. 2007;16(1):63–9.
38. World Cancer Research Fund. World Cancer Research Fund International/American Institute for Cancer Research Continuous Update Project Report: Diet, Nutrition, Physical Activity, and Prostate Cancer [Internet]. 2014. Available from:

<http://www.wcrf.org/sites/default/files/Prostate-Cancer-2014-Report.pdf>

39. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer.* 2015;136(5):E359–86.
40. Roberts RO, Bergstrahl EJ, Bass SE, Lieber MM, Jacobsen SJ. Prostatitis as a risk factor for prostate cancer. *Epidemiology [Internet].* 2004;15(1):93–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14712152>
41. Miah S, Catto J. BPH and prostate cancer risk. *Indian J Urol [Internet].* 2014;30(2):214–8. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4066021/>
42. Lee J, Giovannucci E, Jeon JY. Diabetes and mortality in patients with prostate cancer: a meta-analysis. *Springerplus [Internet].* 2016;5(1):1548. Available from: <http://springerplus.springeropen.com/articles/10.1186/s40064-016-3233-y>
43. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet.* 2009;374(9683):86–9.
44. Murad M, Asi N, Alsawas M, Alahdab F. EBM - New Evidence Pyramid. *Evid Based Med.* 2016;21:125–6.
45. Knobloch K, Yoon U, Vogt PM. Preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement and publication bias. *J Cranio-Maxillofacial Surg.* 2011;39(2):91–2.
46. Mahood Q, Van Eerd D, Irvin E. Searching for grey literature for systematic reviews: Challenges and benefits. *Res Synth Methods.* 2014;5(3):221–34.
47. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Br Med J.* 2011;343:889–93.
48. Harrison S, Jones HE, Martin RM, Lewis SJ, Higgins JP. The albatross plot: a novel graphical tool for presenting results of diversely reported studies in a systematic review. *Res Synth Methods.* 2017;
49. Borenstein M, Hedges LV., Higgins JPT, Rothstein HR. Introduction to Meta-Analysis [Internet]. Psychotherapy research journal of the Society for Psychotherapy Research. 2009. 421 p. Available from: <http://doi.wiley.com/10.1002/9780470743386>
50. Friedman L. Why vote-count reviews don't count. *Biol Psychiatry.* 2001;(49):161–2.
51. Sterne J, Davey Smith G. Sifting the evidence-what's wrong with significance tests? *Bmj.* 2001;322(7280):226–31.
52. Ogilvie D, Fayter D, Petticrew M, Sowden A, Thomas S, Whitehead M, et al. The harvest plot: a method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol.* 2008;8:8.

53. Fisher R a., Fisher R a. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* [Internet]. 1915;10(4):507–521. Available from: <http://biomet.oxfordjournals.org/cgi/reprint/10/4/507.pdf>
54. Riley JW, Stouffer SA, Suchman EA, Devinney LC, Star SA, Williams RM. The American Soldier: Adjustment During Army Life. *American Sociological Review*. 1949. p. 557.
55. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: One-stage and two-stage approaches, and why they may differ. *Statistics in Medicine*. 2016;
56. Verhagen AP, Ferreira ML. Forest plots. *Journal of Physiotherapy*. 2014. p. 170–3.
57. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ Br Med J*. 2003;327(7414):557–60.
58. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Meta-Regression. *Introd to Meta-Analysis* [Internet]. 2009;187–203. Available from: <http://doi.wiley.com/10.1002/9780470743386.ch20>
59. Harbord RM, Higgins JPT. Meta-regression in Stata. *Stata J*. 2008;8(4):493–519.
60. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* [Internet]. 1997;315(7109):629–34. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2127453&tool=pmcentrez&rendertype=abstract>
61. Sterne JAC, Becker BJ, Egger M. The Funnel Plot. In: *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments*. 2006. p. 73–98.
62. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*. 2011;343(jul22 1):d4002–d4002.
63. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* [Internet]. 2016;i4919. Available from: <http://www.bmjjournals.org/lookup/doi/10.1136/bmj.i4919>
64. CASP. CASP Case Control Checklist [Internet]. Critical Appraisal Skills Programme (CASP). 2014 [cited 2015 Nov 24]. Available from: http://media.wix.com/ugd/dded87_63fb65dd4e0548e2bfd0a982295f839e.pdf
65. CASP. CASP cohort study checklist [Internet]. Critical Appraisal Skills Programme (CASP). 2014 [cited 2015 Nov 24]. Available from: http://media.wix.com/ugd/dded87_e37a4ab637fe46a0869f9f77dacf134.pdf
66. Chan AW, Hróbjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*. 2004;291(20):2457–65.
67. Riley RD, Abrams KR, Sutton a J, Lambert PC, Jones DR, Heney D, et al. Reporting of prognostic markers: current problems and development of guidelines for evidence-based practice in the future. *Br J Cancer*. 2003;88(8):1191–8.

68. Bekkering GE, Harris RJ, Thomas S, Mayer AMB, Beynon R, Ness AR, et al. How much of the data published in observational studies of the association between diet and prostate or bladder cancer is usable for meta-analysis? *Am J Epidemiol.* 2008;167(9):1017–26.
69. Sankey H. The Thermal Efficiency of Steam-Engines. *Proc Inst Civ Eng.* 1898;CXXXIV(IV).
70. Du X, Graedel TE. Uncovering the end uses of the rare earth elements. *Sci Total Environ.* 2013;461–462:781–4.
71. Cullen JM, Allwood JM. Mapping the global flow of aluminum: From liquid aluminum to end-use goods. *Environ Sci Technol.* 2013;47(7):3057–64.
72. Nakamura S, Kondo Y, Matsubae K, Nakajima K, Nagasaka T. UPIOM: A new tool of mfa and its application to the flow of iron and steel associated with car production. *Environ Sci Technol.* 2011;45(3):1114–20.
73. Curmi E, Richards K, Fenner R, Allwood JM, Kopec GM, Bajželj B. An integrated representation of the services provided by global water resources. *J Environ Manage.* 2013;129:456–62.
74. Wang S, Aldridge MD, Gross CP, Canavan M, Cherlin E, Bradley E. End-of-Life Care Transition Patterns of Medicare Beneficiaries. *J Am Geriatr Soc* [Internet]. 2017; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28369785>
75. Williamson EJ, Aitken Z, Lawrie J, Dharmage SC, Burgess JA, Forbes AB. Introduction to causal diagrams for confounder selection. *Respirology.* 2014. p. 303–11.
76. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* [Internet]. 1999;10(1):37–48. Available from: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=9888278&retmode=ref&cmd=prlinks%5Cnpapers2://publication/uuid/8B198F93-361C-4091-9A45-2CB165C98559>
77. World Cancer Research Fund Continuous Update Project. TeMMPO (Text Mining for Mechanism Prioritisation) [Internet]. 2017. Available from: <https://www.temmpo.org.uk/>
78. Chang IH, Ahn SH, Han JH, Kim TH, Kim YS, Myung SC. The Clinical Significance in Healthy Men of the Association Between Obesity Related Plasma Hemodilution and Tumor Marker Concentration. *J Urol* [Internet]. American Urological Association; 2009;181(2):567–73. Available from: <http://dx.doi.org/10.1016/j.juro.2008.10.030>
79. Grubb RL, Black A, Izmirlian G, Hickey TP, Pinsky PF, Mabie JE, et al. Serum prostate-specific antigen hemodilution among obese men undergoing screening in the prostate, lung, colorectal, and ovarian cancer screening trial. *Cancer Epidemiol Biomarkers Prev.* 2009;18(3):748–51.
80. Bañez LL, Hamilton RJ, Partin AW, Vollmer RT, Sun L, Rodriguez C, et al. Obesity-related plasma hemodilution and PSA concentration among men with prostate cancer. *Jama* [Internet]. 2007;298(19):2275–80. Available from: <http://jama.jamanetwork.com/article.aspx?articleid=209508%5Cnhttp://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.298.19.2275%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/18029831>
81. Harrison S, Tilling K, Turner EL, Lane JA, Simpkin A, Davis M, et al. Investigating the prostate specific antigen, body mass index and age relationship: is an age-BMI-adjusted PSA model

clinically useful? *Cancer Causes Control*. 2016;27(12):1465–74.

82. Baillargeon J, Pollock BH, Kristal AR, Bradshaw P, Hernandez J, Basler J, et al. The association of body mass index and prostate-specific antigen in a population-based study. *Cancer*. 2005;103(5):1092–5.
83. Haas GP, Delongchamps NB, Jones RF, Chandan V, Serio AM, Vickers AJ, et al. Needle biopsies on autopsy prostates: Sensitivity of cancer detection based on true prevalence. *J Natl Cancer Inst*. 2007;99(19):1484–9.
84. Freedland SJ, Wen J, Wuerstle M, Shah A, Lai D, Moalej B, et al. Obesity Is a Significant Risk Factor for Prostate Cancer at the Time of Biopsy. *Urology* [Internet]. 2008;72(5):1102–5. Available from: <http://www.sciencedirect.com/science/article/pii/S009042950800722X>
85. Kyrgiou M, Kalliala I, Markozannes G, Gunter MJ, Paraskevaidis E, Gabra H, et al. Adiposity and cancer at major anatomical sites: umbrella review of the literature. *BMJ*. 2017;j477.
86. Hu MB, Liu SH, Jiang HW, Bai P De, Ding Q. Obesity affects the biopsy-mediated detection of prostate cancer, particularly high-grade prostate cancer: A dose-response meta-analysis of 29,464 patients. *PLoS One* [Internet]. 2014;9(9):e106677. Available from: <http://dx.plos.org/10.1371/journal.pone.0106677>
87. Discacciati A, Orsini N, Wolk A. Body mass index and incidence of localized and advanced prostate cancer-a dose-response meta-analysis of prospective studies. *Ann Oncol*. 2012;23(7):1665–71.
88. MacInnis RJ, English DR. Body size and composition and prostate cancer risk: Systematic review and meta-regression analysis. *Cancer Causes Control* [Internet]. 2006;17(8):989–1003. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16933050>
89. Harrison S, Lennon R, Holly J, Higgins JPT, Gardner M, Perks C, et al. Does milk intake promote prostate cancer initiation or progression via effects on insulin-like growth factors (IGFs)? A systematic review and meta-analysis. *Cancer causes Control*. 2017;28(248):1–32.
90. Vlassopoulos a, Combet E, Lean MEJ. Changing distributions of body size and adiposity with age. *Int J Obes (Lond)* [Internet]. 2013;38(August):1–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24247373>
91. Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39(2):417–20.
92. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. *Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA* [Internet]. 2000;283(15):2008–12. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10789670
93. Li F, Shen Z, Lu Y, Yun J, Fan Y. Serum prostate-specific antigen concentration and hemodilution among Chinese middle-aged obese men: A hematocrit-based equation for plasma volume estimation is induced. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2012;21(10):1731–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22850805>

94. Shafique K, McLoone P, Qureshi K, Leung H, Hart C, Morrison DS. Coffee consumption and prostate cancer risk: further evidence for inverse relationship. *Nutr J* [Internet]. 2012;11(1):42. Available from: <http://nutritionj.biomedcentral.com/articles/10.1186/1475-2891-11-42>
95. Shafique K, McLoone P, Qureshi K, Leung H, Hart C, Morrison DS. Cholesterol and the risk of grade-specific prostate cancer incidence: evidence from two large prospective cohort studies with up to 37 years' follow up. *BMC Cancer* [Internet]. 2012;12:25. Available from: <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=22260413&retmode=ref&cmd=prlinks>
96. Stocks T, Hergens M-P, Englund A, Ye W, Stattin P. Blood pressure, body size and prostate cancer risk in the Swedish Construction Workers cohort. *Int J Cancer* [Internet]. 2010;127(7):1660–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20087861>
97. Gallus S, Foschi R, Talamini R, Altieri A, Negri E, Franceschi S, et al. Risk Factors for Prostate Cancer in Men Aged Less Than 60 Years: A Case-Control Study from Italy. *Urology* [Internet]. 2007;70(6):1121–6. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18158031>
98. Máčová L, Čížek L, Horáková D, Koutná J, Lorenc J, Janoutová G, et al. Association between obesity and cancer incidence in the population of the District Sumperk, Czech Republic. *Onkologie*. 2007;30(11):538–42.
99. Baillargeon J, Platz EA, Rose DP, Pollock BH, Ankerst DP, Haffner S, et al. Obesity, adipokines, and prostate cancer in a prospective population-based study. *Cancer Epidemiol Biomarkers Prev*. 2006;15(7):1331–5.
100. Kurahashi N, Iwasaki M, Sasazuki S, Otani T, Inoue M, Tsugane S. Association of body mass index and height with risk of prostate cancer among middle-aged Japanese men. *Br J Cancer* [Internet]. 2006;94(5):740–2. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2361195&tool=pmcentrez&rendertype=abstract>
101. LIU X, RYBICKI BA, CASEY G, WITTE JS. Relationship Between Body Size and Prostate Cancer in a Sibling Based Case-Control Study. *J Urol* [Internet]. 2005;174(6):2169–73. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0022534701689390>
102. Salinas CA, Austin MA, Ostrander EO, Stanford JL. Polymorphisms in the androgen receptor and the prostate-specific antigen genes and prostate cancer risk. *Prostate*. 2005;65(1):58–65.
103. Bradbury BD, Wilk JB, Kaye JA. Obesity and the risk of prostate cancer (United States). *Cancer Causes Control* [Internet]. 2005;16(6):637–41. Available from: <http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L41110953%5Cnhttp://dx.doi.org/10.1007/s10552-005-0383-6%5Cnhttp://sfx.library.uu.nl/sfx?sid=EMBASE&issn=09575243&id=doi:10.1007/s10552-005-0383-6&atitle=Obesity+and+the+risk+of+prostate>
104. Oh SW, Yoon YS, Shin SA. Effects of excess weight on cancer incidences depending on cancer sites and histologic findings among men: Korea National Health Insurance Corporation study. *J Clin Oncol* [Internet]. 2005;23(21):4742–54. Available from: <http://www.jco.org/cgi/doi/10.1200/JCO.2005.11.726>
105. Kuriyama S, Tsubono Y, Hozawa A, Shimazu T, Suzuki Y, Koizumi Y, et al. Obesity and risk of

cancer in Japan. *Int J Cancer*. 2005;113(1):148–57.

106. Engeland A, Tretli S, Bjørge T. Height, body mass index, and prostate cancer: a follow-up of 950000 Norwegian men. *Br J Cancer* [Internet]. 2003;89(7):1237–42. Available from: <http://www.nature.com/doifinder/10.1038/sj.bjc.6601801>
107. Villeneuve PJ, Johnson KC, Kreiger N, Mao Y, Paulse B, Dewar R, et al. Risk factors for prostate cancer: Results from the Canadian National Enhanced Cancer Surveillance System. *Cancer Causes Control*. 1999;10(5):355–67.
108. Lund Nilsen TI, Vatten LJ. Anthropometry and prostate cancer risk: A prospective study of 22,248 Norwegian men. *Cancer Causes Control* [Internet]. 1999;10(4):269–75. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10482485
109. Giovannucci E, Rimm EB, Stampfer MJ, Colditz GA, Willett WC. Height, body weight, and risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev* [Internet]. 1997;6(8):557–63. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9264267%5Cnhttp://cebp.aacrjournals.org/content/6/8/557.full.pdf>
110. Andersson S-O, Wolk A, Bergstrom R, Adami H-O, Engholm G, Englund A, et al. Body Size and Prostate Cancer: A 20-Year Follow-up Study Among 135006 Swedish Construction Workers. *JNCI J Natl Cancer Inst* [Internet]. 1997;89(5):385–9. Available from: <http://jnci.oxfordjournals.org/cgi/doi/10.1093/jnci/89.5.385>
111. Cerhan JR, Torner JC, Lynch CF, Rubenstein LM, Lemke JH, Cohen MB, et al. Association of smoking, body mass, and physical activity with risk of prostate cancer in the Iowa 65+ Rural Health Study (United States). *Cancer Causes Control* [Internet]. 1997;8(2):229–38. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9134247>
112. Häggström C, Stocks T, Ulmert D, Bjørge T, Ulmer H, Hallmans G, et al. Prospective study on metabolic factors and risk of prostate cancer. *Cancer*. 2012;118(24):6199–206.
113. Mori M, Masumori N, Fukuta F, Nagata Y, Sonoda T, Miyanaga N, et al. Weight gain and family history of prostate or breast cancers as risk factors for prostate cancer: results of a case-control study in Japan. *Asian Pac J Cancer Prev* [Internet]. 2011;12(3):743–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/21627376>
114. Hernandez BY, Park S-Y, Wilkens LR, Henderson BE, Kolonel LN. Relationship of body mass, height, and weight gain to prostate cancer risk in the multiethnic cohort. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2009;18(9):2413–21. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2742565&tool=pmcentrez&rendertype=abstract>
115. Wallström P, Bjartell a, Gullberg B, Olsson H, Wirfält E. A prospective Swedish study on body size, body composition, diabetes, and prostate cancer risk. *Br J Cancer* [Internet]. 2009;100(11):1799–805. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2695694&tool=pmcentrez&rendertype=abstract>
116. Wright ME, Chang S-C, Schatzkin A, Albanes D, Kipnis V, Mouw T, et al. Prospective study of

- adiposity and weight change in relation to prostate cancer incidence and mortality. *Cancer* [Internet]. 2007;109(4):675–84. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17211863>
117. Lukanova A, Björ O, Kaaks R, Lenner P, Lindahl B, Hallmans G, et al. Body mass index and cancer: Results from the Northern Sweden Health and Disease Cohort. *Int J Cancer* [Internet]. 2006;118(2):458–66. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16049963>
118. Porter MP, Stanford JL. Obesity and the risk of prostate cancer. *Prostate*. 2005;62(April 2004):316–21.
119. Crispo A, Talamini R, Gallus S, Negri E, Gallo A, Bosetti C, et al. Alcohol and the risk of prostate cancer and benign prostatic hyperplasia. *Urology*. 2004;64(4):717–22.
120. Dal Maso L, Zucchetto a, La Vecchia C, Montella M, Conti E, Canzonieri V, et al. Prostate cancer and body size at different ages: an Italian multicentre case-control study. *Br J Cancer* [Internet]. 2004;90(11):2176–80. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2409495&tool=pmcentrez&rendertype=abstract>
121. Friedenreich CM, McGregor SE, Courneya KS, Angyalfi SJ, Elliott FG. Case-control study of anthropometric measures and prostate cancer risk. *Int J Cancer* [Internet]. 2004;110(2):278–83. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15069694>
122. Jonsson F, Wolk A, Pedersen NL, Lichtenstein P, Terry P, Ahlbom A, et al. Obesity and hormone-dependent tumors: Cohort and co-twin control studies based on the Swedish Twin Registry. *Int J Cancer* [Internet]. 2003;106(4):594–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12845658>
123. Giles GG, Severi G, English DR, McCredie MRE, MacInnis R, Boyle P, et al. Early growth, adult body size and prostate cancer risk. *Int J Cancer*. 2003;103(2):241–5.
124. Sharpe CR, Siemiatycki J. Joint effects of smoking and body mass index on prostate cancer risk. *Epidemiology* [Internet]. 2001;12(5):546–51. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11505174>
125. Lee IM, Sesso HD, Paffenbarger RS. A prospective cohort study of physical activity and body size in relation to prostate cancer risk (United States). *Cancer Causes Control* [Internet]. 2001;12(2):187–93. Available from: <http://dx.doi.org/10.1023/A:1008952528771>
126. Mills PK, Beeson WL, Phillips RL, Fraser GE. Cohort study of diet, lifestyle, and prostate cancer in adventist men. *Cancer* [Internet]. 1989;64(3):598–604. Available from: [http://onlinelibrary.wiley.com/doi/10.1002/1097-0142\(19890801\)64:3%3C598::AID-CNCR2820640306%3E3.0.CO;2-6/abstract%5Cnhttp://onlinelibrary.wiley.com/doi/10.1002/1097-0142\(19890801\)64:3%3C598::AID-CNCR2820640306%3E3.0.CO;2-6/abstract%5Cnhttp://onlinelibrary.wiley.co](http://onlinelibrary.wiley.com/doi/10.1002/1097-0142(19890801)64:3%3C598::AID-CNCR2820640306%3E3.0.CO;2-6/abstract%5Cnhttp://onlinelibrary.wiley.com/doi/10.1002/1097-0142(19890801)64:3%3C598::AID-CNCR2820640306%3E3.0.CO;2-6/abstract%5Cnhttp://onlinelibrary.wiley.co)
127. Severson RK, Grove JS, Nomura a M, Stemmermann GN. Body mass and prostatic cancer: a prospective study. *BMJ*. 1988;297(6650):713–5.
128. Greenland S, Longnecker MP. Methods for trend estimation from summarized dose-response data, with applications to meta-analysis. *Am J Epidemiol*. 1992;135(11):1301–9.

129. Orsini N, Bellocchio R, Greenland S. Generalized least squares for trend estimation of summarized dose-response data. *Stata Journal*. 2006; p. 40–57.
130. Whittemore A. Collapsibility of multidimensional contingency tables. *J R Stat Soc Ser B (... [Internet]*. 1978;40(3):328–40. Available from: <http://www.jstor.org/stable/10.2307/2984697>
131. Thomas DC, Greenland S. The relative efficiencies of matched and independent sample designs for case-control studies. *J Chronic Dis*. 1983;36(10):685–97.
132. Chene G, Thompson SG. Methods for summarizing the risk associations of quantitative variables in epidemiologic studies in a consistent form. *Am J Epidemiol [Internet]*. 1996;144(6):610–21. Available from: <http://aje.oxfordjournals.org/content/144/6/610.abstract>
133. Yu J, Lavoue J, Parent ME. Sunlight exposure during leisure activities and risk of prostate cancer in Montreal, Canada, 2005–2009. *BMCPublic Heal*. 2014;14:756-.
134. Edwards TL, Giri A, Motley S, Duong W, Fowke JH. Pleiotropy between genetic markers of obesity and risk of prostate cancer. *Cancer Epidemiol Biomarkers Prev [Internet]*. 2013;22(September):1538–46. Available from: <http://cebp.aacrjournals.org/content/22/9/1538.short>
135. Lai GY, Giovannucci EL, Pollak MN, Peskoe SB, Stampfer MJ, Willett WC, et al. Association of C-peptide and leptin with prostate cancer incidence in the Health Professionals Follow-up Study. *Cancer Causes Control [Internet]*. 2014;25(5):625–32. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24664287>
136. Lahmann PH, Wallström P, Lissner L, Olsson H, Gullberg B. Measures of birth size in relation to risk of prostate cancer: the Malmö Diet and Cancer Study, Sweden. *J Dev Orig Health Dis [Internet]*. 2012;3(6):442–9. Available from: http://www.journals.cambridge.org/abstract_S2040174412000402
137. Fowke JH, Motley SS, Concepcion RS, Penson DF, Barocas DA. Obesity, body composition, and prostate cancer. *BMC Cancer [Internet]*. BioMed Central Ltd; 2012;12(1):23. Available from: <http://bmccancer.biomedcentral.com/articles/10.1186/1471-2407-12-23>
138. Nguyen PL, Ma J, Chavarro JE, Freedman ML, Lis R, Fedele G, et al. Fatty acid synthase polymorphisms, tumor expression, body mass index, prostate cancer risk, and survival. *J Clin Oncol*. 2010;28(25):3958–64.
139. Li H, Stampfer MJ, Mucci L, Rifai N, Qiu W, Kurth T, et al. A 25-year prospective study of plasma adiponectin and leptin concentrations and prostate cancer risk and survival. *Clin Chem [Internet]*. 2010;56(1):34–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19910504>
140. Ahn J, Peters U, Albanes D, Purdue MP, Abnet CC, Chatterjee N, et al. Serum vitamin D concentration and prostate cancer risk: A nested case-control study. *J Natl Cancer Inst*. 2008;100(11):796–804.
141. Hultdin J, Van Guelpen B, Bergh A, Hallmans G, Stattin P. Plasma folate, vitamin B12, and homocysteine and prostate cancer risk: a prospective study. *Int J Cancer [Internet]*. 2005;113(5):819–24. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15499634>

142. Stattin P, Bylund A, Biessy C, Kaaks R, Hallmans G, Adlercreutz H. Prospective study of plasma enterolactone and prostate cancer risk (Sweden). *Cancer Causes Control* [Internet]. 2004;15(10):1095–102. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15801493
143. Hsing AW, Chua S, Gao Y, Gentzschein E, Chang L, Deng J, et al. Prostate Cancer Risk and Serum Levels of Insulin and Leptin : a Population- Based Study. 2001;93(10):783–9.
144. Heikkila R, Aho K, Heliovaara M, Hakama M, Marniemi J, Reunananen A, et al. Serum testosterone and sex hormone-binding globulin concentrations and the risk of prostate carcinoma: a longitudinal study. *Cancer* [Internet]. 1999;86(2):312–5. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10421267
145. Lagiou P, Signorello LB, Trichopoulos D, Tzonou a, Trichopoulou a, Mantzoros CS. Leptin in relation to prostate cancer and benign prostatic hyperplasia. *Int J Cancer* [Internet]. 1998;76(1):25–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9533757>
146. Whittemore AS, Kolonel LN, Wu AH, John EM, Gallagher RP, Howe GR, et al. Prostate cancer in relation to diet, physical activity, and body size in blacks, whites, and Asians in the United States and Canada. *J Natl Cancer Inst* [Internet]. 1995;87(9):652–61. Available from: internal-pdf://72.80.119.144/Whittemore-1995-Prostate_cancer_in_relation_to.pdf%5Cnhttp://jnci.oxfordjournals.org/content/87/9/652.full.pdf
147. Lai GY, Helzlsouer KJ, Clipp SL, Rifai N, Platz EA. Association between C-peptide concentration and prostate cancer incidence in the CLUE II cohort study. *Cancer Prev Res* [Internet]. 2010;3(10):1334–41. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2055794
148. Farhat GN, Taioli E, Cauley JA, Zmuda JM, Orwoll E, Bauer DC, et al. The association of bone mineral density with prostate cancer risk in the osteoporotic fractures in men (MrOS) study. *Cancer Epidemiol Biomarkers Prev*. 2009;18(1):148–54.
149. Platz EA, De Marzo AM, Erlinger TP, Rifai N, Visvanathan K, Hoffman SC, et al. No association between pre-diagnostic plasma C-reactive protein concentration and subsequent prostate cancer. *Prostate*. 2004;59(4):393–400.
150. Zhu K. History of Diabetes Mellitus and Risk of Prostate Cancer in Physicians. *Am J Epidemiol* [Internet]. 2004;159(10):978–82. Available from: <http://aje.oupjournals.org/cgi/doi/10.1093/aje/kwh139>
151. Stattin P, Rinaldi S, Stenman UH, Riboli E, Hallmans G, Bergh a, et al. Plasma prolactin and prostate cancer risk: A prospective study. *Int J Cancer* [Internet]. 2001;92(3):463–5. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/11291087>
152. Bhavsar NA, Bream JH, Meeker AK, Drake CG, Peskoe SB, Dabitao D, et al. A peripheral circulating TH1 cytokine profile is inversely associated with prostate cancer risk in CLUE II. *Cancer Epidemiol Biomarkers Prev*. 2014;23(11):2561–7.
153. Feingold A. A Regression Framework for Effect Size Assessments in Longitudinal Modeling of

Group Differences. Rev Gen Psychol. 2013;17:111–21.

154. Bidoli E, Talamini R, Bosetti C, Negri E, Maruzzi D, Montella M, et al. Macronutrients, fatty acids, cholesterol and prostate cancer risk. Ann Oncol. 2005;16(1):152–7.
155. Hsieh CC, Thanos A, Mitropoulos D, Deliveliotis C, Mantzoros CS, Trichopoulos D. Risk factors for prostate cancer: a case-control study in Greece. Int J Cancer [Internet]. 1999;80(5):699–703. Available from: <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&CSC=Y&NEWS=N&PAGE=fulltext&D=med4&AN=10048970>
156. Brändstedt J, Almquist M, Manjer J, Malm J. Vitamin D, PTH, and calcium and the risk of prostate cancer: a prospective nested case-control study. Cancer Causes Control [Internet]. 2012;23(8):1377–85. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22706676>
157. Chia SE, Wong KY, Cheng C, Lau W, Tan PH. Sun exposure and the risk of prostate cancer in the Singapore Prostate Cancer Study: a case-control study. Asian Pac J Cancer Prev [Internet]. 2012;13(7):3179–85. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22994730>
158. Robinson WR, Stevens J, Gammon MD, John EM. Obesity before age 30 years and risk of advanced prostate cancer. Am J Epidemiol [Internet]. 2005;161(12):1107–14. Available from: <http://aje.oupjournals.org/cgi/doi/10.1093/aje/kwi150>
159. Stark JR, Li H, Kraft P, Kurth T, Giovannucci EL, Stampfer MJ, et al. Circulating prediagnostic interleukin-6 and C-reactive protein and prostate cancer incidence and mortality. Int J Cancer [Internet]. 2009;124(11):2683–9. Available from: <http://doi.wiley.com/10.1002/ijc.24241>
160. Bosetti C, Talamini R, Montella M, Negri E, Conti E, Franceschi S, et al. Retinol, carotenoids and the risk of prostate cancer: A case-control study from Italy. Int J Cancer. 2004;112(4):689–92.
161. Augustin LSA, Galeone C, Dal Maso L, Pelucchi C, Ramazzotti V, Jenkins DJA, et al. Glycemic index, glycemic load and risk of prostate cancer. Int J Cancer. 2004;112(3):446–50.
162. Möller E, Adami H-O, Mucci L a, Lundholm C, Bellocchio R, Johansson J-E, et al. Lifetime body size and prostate cancer risk in a population-based case-control study in Sweden. Cancer Causes Control [Internet]. 2013;24(12):2143–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24048969>
163. Salem S, Hosseini M, Allameh F, Babakoohi S, Mehrsai A, Pourmand G. Serum calcium concentration and prostate cancer risk: a multicenter study. Nutr Cancer [Internet]. 2013;65(7):961–8. Available from: <http://www.tandfonline.com/doi/abs/10.1080/01635581.2013.806936#.VuRf8fkrLIU>
164. Kopp TI, Friis S, Christensen J, Tjønneland A, Vogel U. Polymorphisms in genes related to inflammation, NSAID use, and the risk of prostate cancer among Danish men. Cancer Genet [Internet]. Elsevier Inc.; 2013;206(7–8):266–78. Available from: <http://dx.doi.org/10.1016/j.cancergen.2013.06.001>
165. Rao GA, Mann JR, Bottai M, Uemura H, Burch JB, Bennett CL, et al. Angiotensin receptor blockers and risk of prostate cancer among united states veterans. J Clin Pharmacol. 2013;53(7):773–8.
166. Bassett JK, Severi G, Baglietto L, MacInnis RJ, Hoang HN, Hopper JL, et al. Weight change and

- prostate cancer incidence and mortality. *Int J Cancer*. 2012;131(7):1711–9.
167. pinnacle, Nemesure B, Wu S-Y., Hennis A, Leske MC. Central Adiposity and Prostate Cancer in a Black Population. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2012;21(5):851–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/22402288> Cn<http://cebp.aacrjournals.org/cgi/doi/10.1158/1055-9965.EPI-12-0071>
168. Burton A, Martin R, Galobardes B, Davey Smith G, Jeffreys M. Young adulthood body mass index and risk of cancer in later adulthood: historical cohort study. *Cancer Causes Control*. 2010;21(12):2069–77.
169. Lundqvist E, Kaprio J, Verkasalo PK, Pukkala E, Koskenvuo M, Söderberg KC, et al. Co-twin control and cohort analyses of body mass index and height in relation to breast, prostate, ovarian, corpus uteri, colon and rectal cancer among Swedish and Finnish twins. *Int J Cancer* [Internet]. 2007;121(4):810–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17455257>
170. Håheim LL, Wisløff TF, Holme I, Nafstad P. Metabolic syndrome predicts prostate cancer in a cohort of middle-aged Norwegian men followed for 27 years. *Am J Epidemiol*. 2006;164(8):769–74.
171. Cui Y, Winton MI, Zhang ZF, Rainey C, Marshall J, De Kernion JB, et al. Dietary boron intake and prostate cancer risk. *Oncol Rep*. 2004;11(4):887–92.
172. MacInnis RJ, English DR, Gertig DM, Hopper JL, Giles GG. Body Size and Composition and Prostate Cancer Risk. *Cancer Epidemiol Biomarkers Prev*. 2003;12(12):1417–21.
173. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet* [Internet]. Bhaskaran et al. Open Access article distributed under the terms of CC BY; 2014;384(9945):755–65. Available from: [http://dx.doi.org/10.1016/S0140-6736\(14\)60892-8](http://dx.doi.org/10.1016/S0140-6736(14)60892-8)
174. Chamie K, DeVere White RW, Lee D, Ok J-H, Ellison LM. Agent Orange exposure, Vietnam War veterans, and the risk of prostate cancer. *Cancer* [Internet]. 2008;113(9):2464–70. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18666213>
175. Discacciati A, Orsini N, Andersson S-O, Andrén O, Johansson J-E, Wolk A. Body mass index in early and middle-late adulthood and risk of localised, advanced and fatal prostate cancer: a population-based prospective study. *Br J Cancer* [Internet]. 2011;105(7):1061–8. Available from: <http://dx.doi.org/10.1038/bjc.2011.319>
176. Schuurman AG, Goldbohm RA, Dorant E, van den Brandt PA. Anthropometry in relation to prostate cancer risk in the Netherlands Cohort Study. *Am J Epidemiol* [Internet]. 2000;151(6):541–9. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10733035
177. Andersson SO, Baron J, Bergström R, Lindgren C, Wolk a, Adami HO. Lifestyle factors and prostate cancer risk: a case-control study in Sweden. *Cancer Epidemiol Biomarkers Prev*. 1996;5(7):509–13.
178. Pischon T, Boeing H, Weikert S, Allen N, Key T, Johnsen NF, et al. Body size and risk of prostate

- cancer in the European prospective investigation into cancer and nutrition. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2008;17(11):3252–61. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=18990768
179. Attner B, Landin-Olsson M, Lithman T, Noreen D, Olsson H. Cancer among patients with diabetes, obesity and abnormal blood lipids: a population-based register study in Sweden. *Cancer Causes Control* [Internet]. 2012;23(5):769–77. Available from: [http://link.springer.com/article/10.1007%252Fs10552-012-9946-5%5Cnhttp://download.springer.com/static/pdf/443/art%25253A10.1007%25252Fs10552-012-9946-5.pdf?originUrl=http%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs10552-012-9946-5&token2=exp=1455626637~acl="](http://link.springer.com/article/10.1007%252Fs10552-012-9946-5%5Cnhttp://download.springer.com/static/pdf/443/art%25253A10.1007%25252Fs10552-012-9946-5.pdf?originUrl=http%253A%252F%252Flink.springer.com%252Farticle%252F10.1007%252Fs10552-012-9946-5&token2=exp=1455626637~acl=)
180. Sawada N, Iwasaki M, Inoue M, Sasazuki S, Yamaji T, Shimazu T, et al. Plasma testosterone and sex hormone-binding globulin concentrations and the risk of prostate cancer among Japanese men: A nested case-control study. *Cancer Sci.* 2010;101(12):2652–7.
181. Cox B, Sneyd MJ, Paul C, Skegg DCG. Risk factors for prostate cancer: A national case-control study. *Int J Cancer* [Internet]. 2006;119(7):1690–4. Available from: <http://doi.wiley.com/10.1002/ijc.22022>
182. Tavani A, Gallus S, Bertuzzi M, Dal Maso L, Zucchetto A, Negri E, et al. Diabetes mellitus and the risk of prostate cancer in Italy. *Eur Urol* [Internet]. 2005;47(3):313–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/15716192>
183. Pan SY, Johnson KC, Ugnat AM, Wen SW, Mao Y. Association of Obesity and Cancer Risk in Canada. *Am J Epidemiol.* 2004;159(3):259–68.
184. Hsing a W, Deng J, Sesterhenn I a, Mostofi FK, Stanczyk FZ, Benichou J, et al. Body size and prostate cancer: a population-based case-control study in China. *Cancer Epidemiol Biomarkers Prev.* 2000;9(12):1335–41.
185. Putnam SD, Cerhan JR, Parker AS, Bianchi GD, Wallace RB, Cantor KP, et al. Lifestyle and anthropometric risk factors for prostate cancer in a cohort of Iowa men. *Ann Epidemiol.* 2000;10(6):361–9.
186. Habel LA, Van Den Eeden SK, Friedman GD. Body size, age at shaving initiation, and prostate cancer in a large, multiracial cohort. *Prostate* [Internet]. 2000;43(2):136–43. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/10754529>
187. Veierød MB, Laake P, Thelle DS. Dietary fat intake and risk of prostate cancer: a prospective study of 25,708 Norwegian men. *Int J Cancer* [Internet]. 1997;73(5):634–8. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/9398038>
188. Le Marchand L, Kolonel LN, Wilkens LR, Myers BC, Hirohata T. Animal fat consumption and prostate cancer: a prospective study in Hawaii. *Epidemiology* [Internet]. 1994;5(3):276–82. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/8038241>
189. Harding JL, Shaw JE, Anstey KJ, Adams R, Balkau B, Brennan-Olsen SL, et al. Comparison of anthropometric measures as predictors of cancer incidence: A pooled collaborative analysis of 11 Australian cohorts. *Int J Cancer.* 2015;137(7):1699–708.

190. Littman AJ, White E, Kristal AR. Anthropometrics and prostate cancer risk. *Am J Epidemiol* [Internet]. 2007;165(11):1271–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17395597>
191. Freedland SJ, Platz EA, Presti JC, Aronson WJ, Amling CL, Kane CJ, et al. Obesity, serum prostate specific antigen and prostate size: Implications for prostate cancer detection. *J Urol*. 2006;175(2):500–4.
192. Price MM, Hamilton RJ, Robertson CN, Butts MC, Freedland SJ. Body Mass Index, Prostate-Specific Antigen, and Digital Rectal Examination Findings Among Participants in a Prostate Cancer Screening Clinic. *Urology* [Internet]. 2008;71(5):787–91. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18267334>
193. Culp S, Porter M. The effect of obesity and lower serum prostate-specific antigen levels on prostate-cancer screening results in American men. *BJU Int* [Internet]. 2009;104(10):1457–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19522868>
194. Park J-H, Cho B-L, Kwon H-T, Lee C-M, Han H-J. Effect of body mass index and waist circumference on prostate specific antigen and prostate volume in a generally healthy Korean population. *J Urol* [Internet]. American Urological Association; 2009;182(1):106-10-1. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19450837>
195. Wright JL, Lin DW, Stanford JL. The effect of demographic and clinical factors on the relationship between BMI and PSA levels. *Prostate*. 2011;71(15):1631–7.
196. Bhindi B, Margel D, Trottier G, Hamilton RJ, Kulkarni GS, Hersey KM, et al. Obesity is associated with larger prostate volume but not with worse urinary symptoms: Analysis of a large multiethnic cohort. *Urology* [Internet]. Elsevier Inc.; 2014;83(1):81–7. Available from: <http://dx.doi.org/10.1016/j.urology.2013.07.039>
197. Punglia RS, D'Amico A V., Catalona WJ, Roehl KA, Kuntz KM. Impact of age, benign prostatic hyperplasia, and cancer on prostate-specific antigen level. *Cancer*. 2006;106(7):1507–13.
198. Higgins JPT, White IR, Anzures-Cabrera J. Meta-analysis of skewed data: Combining results reported on log-transformed or raw scales. *Stat Med*. 2008;27(29):6072–92.
199. Yang WJ. The likelihood of having a serum PSA level of ≥ 2.5 or ≥ 4.0 ng ml $^{-1}$ according to obesity in a screened Korean population. *Asian J Androl* [Internet]. 2013;15(6):770–2. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84887242982&partnerID=tZOTx3y1>
200. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. In: The Cochrane Collaboration [Internet]. 2011. Available from: www.handbook.cochrane.org
201. Ikuerowo SO, Omisanjo OA, Bioku MJ, Ajala MO, Esho JO. Effect of obesity on serum prostate-specific antigen in nigerian men. *Urol Int*. 2012;89(1):52–6.
202. Chamie K, Oberfoell S, Kwan L, Labo J, Wei JT, Litwin MS. Body mass index and prostate cancer severity: Do obese men harbor more aggressive disease on prostate biopsy? *Urology* [Internet]. Elsevier Inc.; 2013;81(5):949–55. Available from: <http://dx.doi.org/10.1016/j.urology.2013.01.021>

203. Wallner LP, Morgenstern H, McGree ME, Jacobson DJ, St Sauver JL, Jacobsen SJ, et al. The effects of body mass index on changes in prostate-specific antigen levels and prostate volume over 15 years of follow-up: implications for prostate cancer detection. *Cancer Epidemiol Biomarkers Prev* [Internet]. 2011;20(3):501–8. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21242331
204. Cochran WG. The Combination of Estimates from Different Experiments. *Biometrics* [Internet]. 1954;10(1):101. Available from: <http://www.jstor.org/stable/3001666?origin=crossref>
205. Hernán MA. The hazards of hazard ratios. *Epidemiology*. 2010;21(1):13–5.
206. Baldelli R, De Marinis L, D'Amico E, Barnabei A, Pasimeni G, Mecule A, et al. Obesity and cancer. *Obes Metab*. 2008;4(1):4–11.
207. Wallner LP, Morgenstern H, McGree ME, Jacobson DJ, St. Sauver JL, Jacobsen SJ, et al. The Effects of Body Mass Index on Changes in Prostate-Specific Antigen Levels and Prostate Volume Over 15 Years of Follow-up: Implications for Prostate Cancer Detection. *Cancer Epidemiol Biomarkers Prev*. 2011;20(3):501–8.
208. Renéhan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371(November):569–78.
209. Zhang X, Zhou G, Sun B, Zhao G, Liu D, Sun J, et al. Impact of obesity upon prostate cancer-associated mortality: A meta-analysis of 17 cohort studies. *Oncol Lett*. 2015;9(3):1307–12.
210. Markozannes G, Tzoulaki I, Karli D, Evangelou E, Ntzani E, Gunter MJ, et al. Diet, body size, physical activity and risk of prostate cancer: An umbrella review of the evidence. *Eur J Cancer*. 2016;69:61–9.
211. Al-Azab R, Toi A, Lockwood G, Kulkarni GS, Fleshner N. Prostate Volume Is Strongest Predictor of Cancer Diagnosis at Transrectal Ultrasound-Guided Prostate Biopsy with Prostate-Specific Antigen Values Between 2.0 and 9.0 ng/mL. *Urology*. 2007;69(1):103–7.
212. Oh JJ, Jeong SJ, Lee BK, Jeong CW, Byun SS, Hong SK, et al. Does obesity affect the accuracy of prostate-specific antigen (PSA) for predicting prostate cancer among men undergoing prostate biopsy. *BJU Int* [Internet]. 2013;112(4):E265-71. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23432960>
213. Masuda H, Kagawa M, Kawakami S, Numao N, Matsuoka Y, Yokoyama M, et al. Body mass index influences prostate cancer risk at biopsy in Japanese men. *Int J Urol* [Internet]. 2013;20(7):701–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23186107>
214. Jeon KP, Jeong TY, Lee SY, Hwang SW, Shin JH, Kim DS. Prostate cancer in patients with metabolic syndrome is associated with low grade gleason score when diagnosed on biopsy. *Korean J Urol*. 2012;53(9):593–7.
215. Park J, Cho SY, Lee SB, Son H, Jeong H. Obesity is associated with higher risk of prostate cancer detection in a biopsy population in Korea. *BJU Int*. 2014;114(6):891–5.
216. Kobayashi T, Mitsumori K, Nishizawa K, Kawahara T, Ogura K, Ide Y. Association between body mass index and prostate cancer detection rates in Japanese urologic patients. *Urology*

- [Internet]. 2005;66(1):130–4. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=16014286
217. Skolarus TA, Wolin KY, Grubb 3rd RL. The effect of body mass index on PSA levels and the development, screening and treatment of prostate cancer. *Nat Clin Pract Urol* [Internet]. 2007;4(11):605–14. Available from: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17982437
218. Donders ART, van der Heijden GJMG, Stijnen T, Moons KGM. Review: A gentle introduction to imputation of missing values. *J Clin Epidemiol*. 2006;59(10):1087–91.
219. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. *Am J Epidemiol* [Internet]. 1995;142(12):1255–64. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/7503045>
220. Rothman KJ, Greenland S, Associate TLL. *Modern Epidemiology*, 3rd Edition. Hastings Cent Rep [Internet]. 2014;44 Suppl 2:insidebackcover. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24644503>
221. White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. *Stat Med*. 2010;29(28):2920–31.
222. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj*. 2009;338(July):1–10.
223. Lee KJ, Simpson JA. Introduction to multiple imputation for dealing with missing data. *Respirology*. 2014;19(2):162–7.
224. Rubin R. *Multiple imputation for nonresponse in surveys*. New York: Wiley; 1987.
225. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377–99.
226. Van Buuren S, Groothuis-Oudshoorn K. Multivariate Imputation by Chained Equations. *J Stat Softw* [Internet]. 2011;45(3):1–67. Available from: <http://igitur-archive.library.uu.nl/fss/2010-0608-200146/UUindex.html>
227. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16:219–42.
228. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: What is it and how does it work? *Int J Methods Psychiatr Res*. 2011;20(1):40–9.
229. Kunkel D, Kaizar EE. A comparison of existing methods for multiple imputation in individual participant data meta-analysis. *Stat Med*. 2017;
230. Carpenter JR, Kenward MG. *Multiple Imputation and its Application*. *Multiple Imputation and its Application*. 2012. 1-345 p.
231. Burgess S, White IR, Resche-Rigon M, Wood AM. Combining multiple imputation and meta-

- analysis with individual participant data. *Stat Med*. 2013;32(26):4499–514.
232. Von Hippel PT. How to impute interactions, squares, and other transformed variables. *Sociol Methodol*. 2009;39(1):265–91.
233. Debray TPA, Moons KGM, Abo-Zaid GMA, Koffijberg H, Da Riley R. Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? *PLoS One*. 2013;8(4).
234. Burke DL, Ensor J, Riley RD. Meta-analysis using individual participant data: one-stage and two-stage approaches, and why they may differ. *Stat Med*. 2016;
235. Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart LA. Statistical analysis of individual participant data meta-analyses: a comparison of methods and recommendations for practice. *PLoS One* [Internet]. 2012;7(10):e46042. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23056232>
236. Turner RM, Omar RZ, Yang M, Goldstein H, Thompson SG. A multilevel model framework for meta-analysis of clinical trials with binary outcomes. *Stat Med* [Internet]. 2000;19(24):3417–32. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1659673/>
237. Riley RD, Steyerberg EW. Meta-analysis of a binary outcome using individual participant data and aggregate data. *Res Synth Methods* [Internet]. 2010;1(1):2–19. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26056090>
238. Quartagno M, Carpenter JR. Multiple imputation for IPD meta-analysis: allowing for heterogeneity and studies with missing covariates. *Stat Med*. 2016;35(17):2938–54.
239. Blankestijn MH, Groeneveld FPMJ, Prins A, Bernsen RMD, Bohnen AM, Bosch JLHR. Strong effects of definition and nonresponse bias on prevalence rates of clinical benign prostatic hyperplasia: The Krimpen study of male urogenital tract problems and general health status. *BJU Int*. 2000;85(6):665–71.
240. Schröder FH, Hugosson J, Roobol MJ, Tammela TLJ, Zappa M, Nelen V, et al. Screening and prostate cancer mortality: Results of the European Randomised Study of Screening for Prostate Cancer (ERSPC) at 13 years of follow-up. *Lancet*. 2014;384(9959):2027–35.
241. Population and Languages of Netherlands [Internet]. 2017. Available from: <https://www.amsterdam.info/netherlands/population/>
242. Eisenberg ML, Davies BJ, Cooperberg MR, Cowan JE, Carroll PR. Prognostic Implications of an Undetectable Ultrasensitive Prostate-Specific Antigen Level after Radical Prostatectomy. *Eur Urol*. 2010;57(4):622–30.
243. Moul JW. The evolving definition of advanced prostate cancer. *Rev Urol* [Internet]. 2004;6 Suppl 8(Suppl 8):S10-7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16985915%5Cnhttp://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC1472896>
244. StataCorp. Stata multiple-imputation reference manual. College Station, TX: StataCorp LP; 2013.
245. Aus G, Damber JE, Khatami A, Lilja H, Stranne J, Hugosson J. Individualized screening interval for prostate cancer based on prostate-specific antigen level: results of a prospective,

- randomized, population-based study. Arch Intern Med [Internet]. 2005;165(16):1857–61. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16157829%5Cnhttp://archinte.jamanetwork.com/data/Journals/INTEMED/12035/loi50064.pdf>
246. Lilja H, Ulmert D, Vickers AJ. Prostate-specific antigen and prostate cancer: prediction, detection and monitoring. Nat Rev Cancer [Internet]. 2008;8(4):268–78. Available from: <http://www.nature.com/doifinder/10.1038/nrc2351>
247. Kelly SP, Graubard BI, Andreotti G, Younes N, Cleary SD, Cook MB. Prediagnostic body mass index trajectories in relation to prostate cancer incidence and mortality in the PLCO cancer screening trial. J Natl Cancer Inst. 2017;109(3).
248. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. In: The Cochrane Collaboration. 2011. p. Table 7.7.a: Formulae for combining groups.
249. Shariat SF, Roehrborn CG. Using biopsy to detect prostate cancer. Rev Urol [Internet]. 2008;10(4):262–80. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19145270%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC2615104/pdf/RIU010004_0262.pdf
250. Bonn SE, Sjolander A, Tillander A, Wiklund F, Gronberg H, Balter K. Body mass index in relation to serum prostate-specific antigen levels and prostate cancer risk. Int J Cancer [Internet]. 2016; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26914149>
251. Pearl J. An Introduction to Causal Inference. Int J Biostat. 2010;6(2).
252. Reas DL, Nygård JF, Svensson E, Sørensen T, Sandanger I. Changes in body mass index by age, gender, and socio-economic status among a cohort of Norwegian men and women (1990–2001). BMC Public Health [Internet]. 2007;7(1):269. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2222164&tool=pmcentrez&rendertype=abstract%5Cnhttp://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-7-269>
253. Bradford-Hill A. The Environment and Disease: Association or Causation? Proc R Soc Med [Internet]. 1965;58:295–300. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1898525&tool=pmcentrez&rendertype=abstract%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/14283879%5Cnhttp://www.ncbi.nlm.nih.gov/articlerender.fcgi?artid=PMC1898525>
254. Kovesdy CP, Kalantar-Zadeh K. Observational Studies Versus Randomized Controlled Trials: Avenues to Causal Inference in Nephrology. Advances in Chronic Kidney Disease. 2012. p. 11–8.
255. Kirkwood BB, Sterne J. Essential medical statistics [Internet]. Malden, MA: Blackwell Science. 2003. 1-512 p. Available from: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Essential+Medical+Statistics#0>
256. Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol [Internet]. 2000;29(4):722–9. Available from:

<http://www.ncbi.nlm.nih.gov/pubmed/10922351><http://ije.oxfordjournals.org/content/29/4/722.full.pdf%5Cn><http://ije.oxfordjournals.org/content/29/4/722.full.pdf>

257. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Smith GD. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8):1133–63.
258. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. *Epidemiology* [Internet]. 2014;25(3):427–35. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24681576>
259. Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. *Stat Methods Med Res* [Internet]. 2015;962280215597579. Available from: <http://smm.sagepub.com.proxy1.lib.uwo.ca/content/early/2015/08/14/0962280215597579.full%5Cn><http://smm.sagepub.com/cgi/doi/10.1177/0962280215597579>
260. Baum CF, Schaffer ME, Stillman S. Instrumental variables and GMM: Estimation and testing. *Stata J*. 2003;3(1):1–31.
261. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* [Internet]. 2014;23(R1):R89–98. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/25064373%5Cn><http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4170722/>
262. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. *Int J Epidemiol*. 2014;43(3):922–9.
263. Burgess S, Thompson SG. Avoiding bias from weak instruments in mendelian randomization studies. *Int J Epidemiol*. 2011;40(3):755–64.
264. Brion MJA, Shakhbazov K, Visscher PM. Calculating statistical power in Mendelian randomization studies. *Int J Epidemiol*. 2013;42(5):1497–501.
265. Burgess S, Scott RA, Timpson NJ, Smith GD, Thompson SG. Using published data in Mendelian randomization: A blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015;30(7):543–52.
266. Stearns FW. One hundred years of pleiotropy: A retrospective. *Genetics*. 2010. p. 767–73.
267. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–25.
268. Price AL, Patterson N, Plenge RM, Weinblatt ME, Shadick NA. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* [Internet]. 2006;38(8):904–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16862161>
269. Swanson SA, Tiemeier H, Ikram MA, Hernán MA. Nature as a Trialist? *Epidemiology* [Internet]. 2017;28(5):653–9. Available from: <http://insights.ovid.com/crossref?an=00001648-201709000-00004>
270. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* [Internet]. 2009;106(23):9362–7. Available from:

<http://www.ncbi.nlm.nih.gov/pubmed/19474294>

271. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 2014;42(D1).
272. Kote-Jarai Z, Easton DF, Stanford JL, Ostrander EA, Schleutker J, Ingles SA, et al. Multiple novel prostate cancer predisposition loci confirmed by an international study: The PRACTICAL consortium. *Cancer Epidemiol Biomarkers Prev.* 2008;17(8):2052–61.
273. PRACTICAL consortium. PRACTICAL Consortium [Internet]. Available from: <http://practical.cge.medschl.cam.ac.uk/>
274. Eeles RA, Olama AA AI, Benlloch S, Saunders EJ, Leongamornlert DA, Tymrakiewicz M, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. *Nat Genet* [Internet]. 2013;45(4):385–91, 391–2. Available from: <http://dx.doi.org/10.1038/ng.2560>
275. Consortium launches genotyping effort. *Cancer Discov.* 2013;3(12):1321.
276. Rubin LH, Witkiewitz K, Andre JS, Reilly S. Methods for Handling Missing Data in the Behavioral Neurosciences: Don't Throw the Baby Rat out with the Bath Water. *J Undergrad Neurosci Educ* [Internet]. 2007;5(2):A71–7. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2007001/>
277. Palmer TM, Lawlor DA, Harbord RM, Sheehan NA, Tobias JH, Timpson NJ, et al. Using multiple genetic variants as instrumental variables for modifiable risk factors. *Stat Methods Med Res* [Internet]. 2012;21(3):223–42. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3333333/>
278. Kim S, Shin C, Jee SH. Genetic variants at 1q32.1, 10q11.2 and 19q13.41 are associated with prostate-specific antigen for prostate cancer screening in two Korean population-based cohort studies. *Gene.* 2015;556(2):199–205.
279. Taylor AE, Davies NM, Ware JJ, Vanderweele T, Smith GD, Munafò MR. Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates. *Econ Hum Biol.* 2014;13(1):99–106.
280. Bland JM, Altman DG. Statistics notes: Multiple significance tests: the Bonferroni method. *BMJ* [Internet]. 1995;310(6973):170–170. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2007001/>
281. Willer CJ, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S, et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* [Internet]. 2013;45(11):1274–83. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3333333/>
282. Turro E, Greene D, Wijgaerts A, Thys C, Lentaigne C, Bariana TK, et al. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nat Genet* [Internet]. 2013;45(1):501–12. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3333333/>

er.fcgi

283. Deng X, Sabino EC, Cunha-Neto E, Ribeiro AL, Ianni B, Mady C, et al. Genome wide association study (GWAS) of chagas cardiomyopathy in trypanosoma cruzi seropositive subjects. *PLoS One.* 2013;8(11).
284. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet [Internet].* 2009;41(6):703–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/19430480%5Cnhttp://www.ncbi.nlm.nih.gov/pmc/articles/PMC2889014>
285. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, Fisher SA, et al. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet [Internet].* 2007;39(7):830–2. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2628541/>&tool=pmcentrez&rendertype=abstract
286. Lin DY, Zeng D. Proper analysis of secondary phenotype data in case-control association studies. *Genet Epidemiol.* 2009;33(3):256–65.
287. Davies NM, Gaunt TR, Lewis SJ, Holly J, Donovan JL, Hamdy FC, 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. *Cancer Causes Control.* Springer International Publishing; 2015;26(11):1603–16.
288. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol [Internet].* 2016;45(3):896–908. Available from: <https://academic.oup.com/ije/article-lookup/doi/10.1093/ije/dyw129>
289. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature [Internet].* 2015;518(7538):197–206. Available from: <http://www.nature.com/doifinder/10.1038/nature14177>
290. Lewis SJ, Murad A, Chen L, Smith GD, Donovan J, Palmer T, et al. Associations between an obesity related genetic variant (FTO rs9939609) and prostate cancer risk. *PLoS One.* 2010;5(10).
291. Harrison S, Lennon R, Holly J, Higgins JPT, Gardner M, Perks C, et al. Does milk intake promote prostate cancer initiation or progression via effects on insulin-like growth factors (IGFs)? A systematic review and meta-analysis. *Cancer Causes Control.* 2017;28(6):497–528.
292. Hunter B, Harrison S, Portela A, Bick D. The effects of cash transfers and vouchers on the use and quality of maternity care services: A systematic review. *PLoS One.* 2017;12(3).
293. Tyrrell J, Jones SE, Beaumont R, Astley CM, Lovell R, Yaghoobkar H, et al. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *BMJ [Internet].* 2016;352:i582. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC483516/>&tool=pmcentrez&rendertype=abstract
294. Baker C. Briefing Paper 3336: Obesity Statistics [Internet]. House of Commons Library. 2017 [cited 2017 Aug 5]. Available from: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2889014>

295. Kheirandish P, Chinegwundoh F. Ethnic differences in prostate cancer. *Br J Cancer*. 2011;105(4):481–5.
296. MacInnis RJ, English DR. Body size and composition and prostate cancer risk: systematic review and meta-regression analysis. *Cancer Causes Control* [Internet]. 2006;17(8):989–1003. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16933050>
297. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation Analysis. *Annu Rev Psychol* [Internet]. 2007;58(1):593–614. Available from: <http://www.annualreviews.org/doi/10.1146/annurev.psych.58.110405.085542>
298. Kenny D. Mediation with Dichotomous Outcomes [Internet]. 2013 [cited 2017 Aug 5]. Available from: davidakenny.net/doc/dichmed.pdf
299. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods* [Internet]. 2013;18(2):137–50. Available from: <http://www.scopus.com/inward/record.url?eid=2-s2.0-84879516417&partnerID=40&md5=31a61e457b695f531dacc8dc07a6f86b%5Cnhttp://psycnet.apa.org/journals/met/18/2/137.pdf>
300. OvidSP: Deduplicating search results [Internet]. 2016. Available from: <http://resourcecenter.ovid.com/site/help/documentation/ospa/en/dedupe.html>
301. Kwon Y, Lemieux M, McTavish J, Wathen N. Identifying and removing duplicate records from systematic review searches. *J Med Libr Assoc* [Internet]. 2015;103(4):184–8. Available from: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4613377&tool=pmcentrez&rendertype=abstract>
302. Rathbone J, Carter M, Hoffmann T, Glasziou P. Better duplicate detection for systematic reviewers: evaluation of Systematic Review Assistant-Deduplication Module. *Syst Rev* [Internet]. 2015;4(1):6. Available from: <http://www.systematicreviewsjournal.com/content/4/1/6>
303. Jaro MA. Probabilistic linkage of large public health data files. *Stat Med* [Internet]. 1995;14(5–7):491–8. Available from: http://onlinelibrary.wiley.com/doi/10.1002/sim.4780140510/abstract%5Cnhttp://onlinelibrary.wiley.com/store/10.1002/sim.4780140510/asset/4780140510_ftp.pdf?v=1&t=i9wxu75r&s=a0dcf152aafa1b2528453c729350ac78090f4009
304. Patrias K. Citing Medicine: The NLM Style Guide for Authors, Editors, and Publishers [Internet]. 2nd editio. Wendling D, editor. National Library of Medicine (US); 2007. Chapter 23.
305. Sankey C, Henry S, Gérecka-Bruzda A, Richard-Yris MA, Hausberger M. The way to a man's heart is through his stomach: What about horses? *PLoS One*. 2010;5(11).
306. Sankey C, Henry S, André N, Richard-Yris MA, Hausberger M. Do horses have a concept of person? *PLoS One*. 2011;6(3).

APPENDICES

Appendix 1: Abbreviations

Acronym	Description
People	
SH	Sean Harrison
ET	Emma Turner
KT	Kate Tilling
HJ	Hayley Jones
RL	Rosie Lennon
General	
AD	Aggregate Data
BMI	Body-Mass Index
BPH	Benign Prostatic Hypertrophy
CASP	Critical Appraisal Skills Programme
CI	Confidence Interval
DAG	Directed Acyclic Graph, also called a causal diagram
DRE	Digital Rectal Exam
GLST	Generalised Least Squares for Trend estimation (of summarised dose-response data)
GWAS	Genome-Wide Association Study
HR	Hazard Ratio
iCOGS	Illumina Custom Infinium genotyping array
IGF	Insulin-like Growth Factor
IPD	Individual Participant Data
IQR	Inter-Quartile Range
IV	Instrumental Variable, also called an instrument
J	Joules
LR	Likelihood Ratio
PSA	Prostate-Specific Antigen
PCa	Prostate Cancer
GLST	Generalized Least Squares for Trend estimation
MR	Mendelian Randomisation
PICO	Population, Intervention, Comparison, Outcome
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomised Controlled Trial
ROBINS-I	Risk Of Bias In Non-randomised Studies - of Interventions
RR	Relative Risk
SMD	Standardized Mean Difference
SE	Standard Error
SD	Standard Deviation
SMD	Standardised mean difference
SNP	Single Nucleotide Polymorphism
TNM	Tumour, Node, Metastasis Score
UK	United Kingdom
USA	United States of America
VWLS	Variance Weighted Least Squares

Acronym	Description
Studies and organisations	
Aarhus	Aarhus Prostate Cancer Study
AHS	Agricultural Health Study
AMORIS	The Swedish Apolipoprotein MOrtality RISk study
ANZDCC	Australia and New Zealand Diabetes and Cancer Collaboration
ATBC	Alpha-Tocopherol, Beta-Carotene (BPC3) cancer prevention study
Canary PASS	Canary Prostate Active Surveillance Study
CAPS	Cancer of the Prostate in Sweden
CeRePP	Centre de Recherche pour les Pathologies Prostatiques
CIOWE	The Construction Industry's Organization for Working Environment (Safety and Health)
CLUEII	Campaign Against Cancer and Heart Disease
COGS	Collaborative Oncological Gene-environment Study
COH	City Of Hope National Medical Center
COSM	Cohort of Swedish Men
CPCS1	Copenhagen Prostate Cancer Study 1
CPCS2	Copenhagen Prostate Cancer Study 2
CPRD	Clinical Practice Research Datalink
DCH	Danish Diet, Cancer and Health
EPIC	European Prospective Investigation into Cancer and Nutrition
ERSPC	European Randomised Study of Screening for Prostate Cancer
ESTHER	Epidemiological investigations of the chances of preventing, recognizing early and optimally treating chronic diseases in an elderly population
FHCRC	Fred Hutchinson Cancer Research Center
GAC	Glasgow Alumni Cohort
Gene-PARE	Prostate cancer radiotherapy cohort
HAHS	Harvard Alumni Health Study
HPFS	Health Professionals Follow-Up Study
IARC	International Agency for Research on Cancer
IMPACT	Identification of Men with a genetic predisposition to ProstAte Cancer: Targeted screening in men at a higher genetic risk and controls
Iowa65+	Iowa 65+ rural health study
IPO-Porto	Portuguese Oncology Institute, Porto
JPHC	Japan Public Health Center-based Prospective Study
KNHIC	Korea National Health Insurance Company
KPMCP	Kaiser Permanente medical care program
KULEUVEN	Katholieke Universiteit te Leuven
LAAPC	Los Angeles Study of Aggressive Prostate Cancer
MAYO	Mayo Clinic
MCC Spain	Multi Case Control Study-Spain
MCCS	Melbourne Collaborative Cohort Study
MDACC_AS	MD Anderson Cancer Center, Active Surveillance trial
MDCS	Malmo Diet and Cancer
MEC	Multiethnic Cohort Study
Me-Can	Metabolic syndrome and Cancer
MMCS	Melbourne Collaborative Cohort Study
MOFFITT	The Moffitt Group

Acronym	Description
MrOS	Osteopathic Fracture in Men study
NECSS	Canadian National Enhanced Cancer Surveillance System
NHANES	National Health and Nutrition Examination Survey
NHGRI-EBI	National Human Genome Research Institute-European Bioinformatics Institute
	National Health Screening Service in Norway
NICE	National Institute for Health and Care Excellence
NIH-AARP	National Institute of Health American Association of Retired Persons diet and health study
NLCS	The Netherlands cohort study
NMHS	Nashville Men's Health Study
NSHDC	North Sweden Health and Disease Cohort
OS	Oslo Study
PCBP	Prostate Cancer in a Black Population
PCMUS	Prostate Cancer study Medical University Sofia
PCPT	Prostate Cancer Prevention Trial
PHS	Physicians' Health Study
PLCO	Prostate, lung, colorectal and ovarian cancer screening trial
Poland	The Poland Group
PPF-UNIS	Prostate Project Foundation-Postgraduate Medical School, Surrey
PRACTICAL	PRostate cancer AssoCiation group To Investigate Cancer Associated aLterations in the genome
PRAGGA	PRostate cAancer Genetics in Galicia
PROCAP	PROgression in Cancer of the Prostate
PROFILE	Germline Genetic Profiling: Correlation With Targeted Prostate Cancer Screening and Treatment
PROGReSS	Prostate Cancer Group, Santiago, Spain
PROMPT	Prostate cancer: Mechanisms of progression and Treatment
ProtecT	Prostate Testing for Cancer and Treatment
PROtEuS	Prospective, randomized, multicentre, open label, phase II / III study to assess efficacy and safety of ranibizumab 0.5 mg intravitreal injections plus panretinal photocoagulation (PRP) versus PRP in monotherapy in the treatment of subjects with high risk proliferative diabetic retinopathy
QLD	Retrospective Queensland Study
RAPPER	Radiogenomics: Assessment of Polymorphisms for Predicting the Effects of Radiotherapy
SABOR	San Antonio Center for Biomarkers of Risk of prostate carcinoma
SCWC	Swedish Construction Workers Cohort
SEARCH	Study of Epidemiology and Risk factors in Cancer Heredity
SEER	Surveillance, Epidemiology and End Results Program
SFPCS	San Francisco Bay Area Prostate Cancer Study (former NC_CCPC)
SPAG	Serum Proteomic analysis for biomarkers of Aggressive prostate disease in the Guernsey population
STHLM-1	Stockholm-1 study
STHLM-2	Stockholm-2 study
SWOG-PCPT	Southwest Oncology Group - Prostate Cancer Prevention Trial
SWOG-SELECT	Southwest Oncology Group - Selenium and Vitamin E Cancer Prevention Trial
TORONTO	University of Toronto

Acronym	Description
UKGPCS	U.K. Genetic Prostate Cancer Study and The Prostate Cancer Research Foundation Study
ULM	Familial Prostate Cancer Study Ulm
UTAH	UTAH Study
VITAL	Vitamins and lifestyle cohort
WCRF	World Cancer Research Fund
WHO	World Health Organisation
WUGS	Washington University Genetics Study

Appendix 2: Creation and Validation of a Probabilistic Matching Algorithm for Deduplication

Appendix 2.1 Rationale for deduplication

The Sankey diagram uses the number of papers that appear to be associated with variables and PSA or prostate cancer, where a count of the number of references the variable appears in is used as a proxy for the total amount of information available for that variable. This ignores the number of participants in each study, but number of participants is difficult to reliably extract autonomously as the abstract would have to list the total number of participants in a standardised way. As such, the average number of participants across many studies is assumed to be relatively constant, so that 100 papers looking at age and PSA should have a similar total number of participants to 100 papers looking at BMI and PSA.

However, bias can be introduced when duplicate studies are not taken into consideration. When compiling references from multiple databases, duplicate references are inevitable, as references are not limited to single databases. It is possible that the variables considered in a study might be associated with the country of origin of the study. Thus, if that country's journals are more likely to send the paper's information to multiple databases, the variable will be over-represented unless the complete set of references is deduplicated. For example, if BMI were more commonly studied in America, and American journals were more likely to send paper data to multiple databases, then BMI could be over-represented compared to more universal variables, such as age.

For references with exactly the same information (title, authors, journal, page numbers, year of publication), deduplication is simple; both Ovid and Endnote have the capacity to remove exact duplicates, but Ovid limits this to results sets of less than 6,000 (300).

However, scoping studies often have more than 6,000 results, and references in different databases frequently are not exactly the same; authors' middle initials, Greek characters, punctuation and numbers are often dealt with differently between databases, and typographical errors are surprisingly common. Conference abstracts that become full papers were also an issue, as the same information is included in more than one paper; again, the country in which a study is conducted may influence the relative number of conference abstracts, and thus bias the results of a Sankey diagram.

In a systematic review of IGF-1 acting as an intermediate mechanism for the relationship between milk and prostate cancer, I found that of 4,945 papers from multiple databases that passed through standard deduplication in Endnote (where exact matches were removed), 831 (16.8%) were still found

to be duplicates (291). This represents a substantial risk of bias, requiring a solution that doesn't necessitate manually searching for duplicates.

In addition, deduplication of references for systematic reviews can save valuable time that would otherwise be spent doing title/abstract screening on duplicate references.

Appendix 2.2 Deduplication algorithms

A review of deduplication algorithms in 2015 (301) showed that Ovid can deduplicate records it finds from MEDLINE, Embase and CINAHL, and different reference management software packages (RefWorks, EndNote and Mendeley) can also deduplicate references. The authors stated the RefWorks produced the smallest number of false-positives (i.e. references marked as duplicates which were not); minimizing false positives is more important than maximizing sensitivity in systematic reviews, as missed references may lead to different outcomes, whereas maximizing sensitivity reduces the number of work hours.

RefWorks can use the Systematic Review Assistant-Deduplication Module (SRA-DM) (302) to deduplicate references. The module uses various techniques to identify duplicates, but does not allow for typographical errors (i.e. errors in databases, such as misspelled names or the order of words in titles).

Probabilistic matching algorithms have been used in public health to link individuals in large datasets (303) where there may be missing or corrupted data. Although no published deduplication tool has used probabilistic matching, I reasoned this would be able to account for typographical errors and other inconsistencies. Instead of matching the references in their entirety, each aspect of the reference is considered separately and a weighted score is given, representing the likelihood of two references being duplicates. Duplicates could then be removed based on some probability that the papers are identical, accounting for any superficial differences in the references between databases. This strategy also allowed conference abstracts to be found, which often have different titles from the final published paper.

To this end, I created a new deduplication algorithm that could account for small differences between references, in an effort to reduce, as much as possible, any bias in the Sankey diagram. The algorithm was developed through several iterations, and was validated on a small subset of the scoping study.

Appendix 2.3 Version 1

Appendix 2.3.1 Starting Point

In the previous systematic review (291), I devised a strategy to remove duplicates that were not exactly the same; references that had exactly the same title but different information elsewhere. These references were highlighted, checked and marked as duplicates by hand. Additionally, titles with the same first 10 characters were highlighted, checked and marked as duplicates, also by hand. Finally, references with the same journal and authors were compared, but with limited success; authors tend to have the most discrepancy between databases. Conference abstracts were considered duplicates if there was a full paper with a similar title and abstract from the same researchers.

The probabilistic matching algorithm was developed by extending the idea of matching references on their titles, allowing for slight differences, then examining more aspects of the matched reference if the titles are similar enough. The algorithm went through two main iterations, described below, and was not influenced by existing algorithms.

Appendix 2.3.2 Matching References

The initial algorithm ordered references by title, and numbered 1 to N. This was to ensure the most efficient use of resources: duplicate references were mostly likely to have a similar title, and therefore would be closer when ordered by title. As the algorithm only attempted to match a reference with others until it found a match, this reduces the number of searches so long as there are at least some duplicates.

Titles were then split into 10 individual fragments of equal length. The first title's fragments were matched against the entire title of the second reference. If 8 or more of the title fragments were matched in the second reference, then the second reference's title was split into 10 fragments and the fragments were matched to the initial title using the same process. The two references were considered a match and assigned a unique match number if the second title also matched 8 out of 10 fragments with the first title. Matching the second title to the first is important for shorter titles, where 10 equal fragments of title could result in very small and common fragments and thus a high probability of false positives.

How many title fragments required to consider the titles a “match” was a balance between the sensitivity (to detect all duplicates) and specificity (to *not* detect non-duplicates) of the algorithm. Allowing 8 or more matched fragments allows for up to 2 pieces of non-matching information (such as a typographical error, or a difference between roman numerals and numbers); allowing only one

typographical errors may not be inclusive enough, allowing 3 typographical errors allows more non-duplicates to be found. By necessitating 8 fragment matches, the minority of titles that have two discontinuous typographical errors will be found, but the algorithm wasn't so relaxed that many non-duplicates were found.

If no match was found on either of the title matches, then the initial title was compared against the next title in the list, and so on until the first title was compared against the last title. The algorithm then moved onto the second reference, attempting to match it with the third title and so on. Matches were only searched for in references numbered greater than the initial reference, since an attempt to match earlier references will have already been made.

Later references that were already matched with an earlier reference were still considered; this allowed for matching of three or more references that were considered duplicates. Each reference was allowed two matching pair numbers; one for a match with earlier references, and one for a match with later references. By keeping track of these matching numbers, groups of duplicates could be identified. An example of the matching algorithm deduplicating 10 references is shown in [Appendix](#)

Table 2.1.

Once the algorithm has completed matching all references, matched references were compared to see if the authors, year of publication, journal and page number matched. The other pieces of information were considered as whole pieces of information, not as fragments, necessitating an exact match between each variable.

Appendix Table 2.1 Example of 10 studies requiring deduplication. Duplicate references are shaded the same. A description of the first matching algorithm is described on the following page

ID	Title	Authors	Year	Journal	Page	Match	Weight
1	Short-Course Accelerated Radiotherapy in Palliative Treatment of Advanced Pelvic Malignancies: A Phase I Study	L. Caravatta, G. D. A. Padula, G. Macchia, et al.	2012	International Journal of Radiation Oncology Biology Physics	E627-E631	No	0
2	Editorial Comment on: Combined Magnetic Resonance Imaging and Magnetic Resonance Spectroscopy Imaging in the Diagnosis of Prostate Cancer: A Systematic Review and Meta-analysis	M. Seitz and C. Gratzke	2009	European Urology	591	No	0
3	Prostatic and testicular changes in spinal cord injured men	S. Gokkaya, O. Demirdal, A. Memis, et al.	2010	PM and R	S167-S168	ID 9	7
4	Transurethral resection of the prostate (TURP)	A. B. County	1998	Surgical procedures	10-20	No	0
5	The role of previous transurethral resection of the prostate (TURP) on the accuracy of nomograms to predict final pathological stage in men undergoing radical prostatectomy for prostate cancer (PCA)	D. Thuer, D. Pfister, P. Firek, et al.	2010	Journal of Urology	e726-e727	No	0
6	Comprehensive report on prostate cancer misclassification by 16 currently used low-risk and active surveillance criteria	J. R. Palisaar, J. Noldus, B. Loppenberg, et al.	2012	Bju International	E172-E181	No	0
7	Expression and Clinical Role of Growth Differentiation Factor-15 in Ovarian Carcinoma Effusions	A. J. Bock, H. T. Stavnes, T. Kempf, et al.	2010	International Journal of Gynecological Cancer	1448-1455	ID 8	6
8	Expression and Clinical Role of Growth Differentiation Factor-XV in Ovarian Carcinoma Effusions	A. J. Bock, H. T. Stavnes, T. Kempf, et al.	2010	International Journal of Gynaecological Cancer	1448-1455	ID 8, 10	6, 3
9	Prostatic and testicular changes in spinal cord injured men	S. Gokkaya, O.K. Demirdal, A. Memis, et al.	2010	PM and R	S167-8	ID 3	7
10	Expression and Clinical Role of Growth Differentiation Factor-X15 in Ovarian-Carcinoma Effusions	A. Bock, H. Stavnes, T. Kempf, et al.	2010	IJGC	1448-55	ID 8	3

Matching algorithm – version 1, example for Table 1

Initial matching

1. The title of the first reference is split into 10 equal fragments, and each segment is searched for in the title of ID 2
2. Fewer than 8 title pieces matched between IDs 1 and 2, so the fragments are searched for in the title of ID 3
3. Again, fewer than 8 title pieces were matched, so the fragments are search for in the remaining references; no matches were made
4. The title of the second reference is split into 10 equal segments, and each segment is searched for in the titles of IDs 3+
5. No matches were made, so the third reference's title is split and the titles of IDs 4+ are searched
6. 10 out of 10 fragments are matched between ID3 in ID 9
7. The title of ID 9 is split in 10 equal fragments, and these are searched for in the title of ID 3
8. All 10 fragments were matched, so IDs 3 and 9 are marked with a unique matching number
9. Once matched, ID 3 is ignored, so the matching algorithm moves on the ID 4
10. The title of ID 4 is fragmented, and the title of ID 5 is searched
11. 10 out of 10 fragments match, so the title of ID 5 is fragmented and searched for in the title of ID 4
12. Fewer than 8 fragments match, so IDs 4 and 5 are not marked as matching
13. The title of ID 4 is re-fragmented and the titles of IDs 6+ are searched
14. No matches are made with ID 4, or IDs 5 and 6
15. ID 7 matches immediately with ID 8, as the only difference is in how the number 15 is represented. The pair are marked with a unique matching number
16. ID 8 matches with ID 10; as ID 8 is already matched, ID 10 is marked with the same matching number as ID 8 to identify IDs 7, 8 and 10 as the same duplicate
17. ID 9 does not match with ID 10

Complete matching

1. With the initial matching complete, the algorithm considers the matched references
2. The first matched pair, IDs 3 and 9, are compared: the authors and page numbers are different, due to middle names and how page numbers are considered; the year and journals are the same. This represents a good match and the references should be considered duplicates, but possibly with some caution.
3. The second matched pair, IDs 7 and 8, are compared: the references are identical excepting the titles. This represents a very good match and the references should be considered duplicates automatically.
4. The third matched pair, IDs 8 and 10, are compared: only the years are identical, as middle names, journal titles and page numbers have all been dealt with different between databases. This represents a poor match, and the duplicates would have to be checked manually to ensure duplication.

Appendix 2.3.3 Weighted Probability

A weighted probability was used to give more weight to pieces of information that gave a greater degree of certainty in the duplicity of references: for example, if two titles were identical then the references were considered more likely to be the same than if the titles were 80% similar and years were identical. Considerations to the differences in how references were stored in databases were made: it was deemed unlikely that the year of a publication would be very different between databases, but it was relatively more likely for the authors and page numbers to be recorded differently (for example, by missing out middle initials or by recording pages 161-2 rather than 161-162). Therefore, the title, year of publication and journal were weighted higher than page numbers and authors, with the title given the most weight.

$$\text{Weighted score} = \text{Author} + 2 * \text{Year} + 4 * \text{Title} + \text{Journal} + 2 * \text{Page}$$

In this score, the different variables (author, journal etc.) scored 1 if they matched completely between the paired references and 0 otherwise. For example, two references that had identical authors, pages and titles but different journals and years would score 7 out of a possible 10. It is possible for two references that had similar titles (8 or 9 matching fragments) to have a weighted score of 0, as the titles could be similar enough to be considered as a duplicate pair but all variables were in some way different.

Paired references were assumed to be identical if the page numbers and journals were identical and not missing. This was because the titles must have had 8/10 fragments that were identical, so given a similar title and identical page numbers in the same journal it is very likely the references are duplicates. Thus, the weighted score of paired references with the same page numbers and journal were set to 10, a perfect match.

Appendix 2.4 Version 2

Appendix 2.4.1 Decreasing Processing Time

The number of potential matches the algorithm could find increases quickly as more references are added – each unique reference increases the number of searches by the total number of references minus 1. For example, if there were 4 references, then the first reference must be compared with references 2, 3 and 4, the second reference with references 3 and 4, and the third reference must be compared with the fourth: $3+2+1=6$. Adding a fifth reference adds one more comparison to each of those, as well as a comparison between references 4 and 5: $4+3+2+1 = 10$. Therefore, the number of searches if all references are unique is the sum of the number of references minus 1:

$$1 + 2 + 3 + \dots + (n - 1) = \frac{(n^2 - n)}{2}$$

Under the assumption that no references were duplicates, 44,068 references involved 970,972,278 searches, where each search included matching 10 fragments and the possibility the second reference would be compared to the first. Stata could not process this number of operations in a reasonable amount of time; the initial algorithm had a run time in weeks.

Therefore, version 2 of the algorithm was refined to pre-filter the searches based on title length (plus or minus 5 characters) and year of publication (plus or minus 2 years). This reduced the number of operations required to match all references, allowing for a run time of hours instead of weeks.

Appendix 2.4.2 Pre-match Modifications

It quickly became clear that pre-match modifications were necessary to the titles, page numbers, authors and journals to improve the ability of the algorithm to discern duplicates. These changes are listed below:

- All text was made lower case to ensure no differences due to differential capitalisation
- Ampersands (&) were all replaced by the word “and”
- Other common punctuation and spaces were removed, as neither provided information but were a common cause of missed duplicates
- The word “the” was removed from journals, as databases were inconsistent when including “the” in “the journal of...”
- Similarly, the word “jr” was removed from authors
- Page numbers were limited to the first page to avoid differential expression of the last page, e.g. 132-133 versus 132-3

Identifying conference abstracts was considered important as they generally provide less information than a full paper equivalent and should therefore be removed preferentially when removing duplicates. Conference abstracts were identified using a search for “supplement” in the journal name, as in “European Urology, Supplements” and where the first or last character of the page number was an “S”. As conference abstracts are usually issued as supplements, both were highly predictive; the book *Citing Medicine: The NLM Style Guide for Authors, Editors, and Publishers* states in the general rules for location (pagination) that “S” is often used for supplements (304). Some supplements append an “A” to page numbers, to indicate an abstract; therefore, should page numbers include an “S” or an “A” then they can be considered supplements.

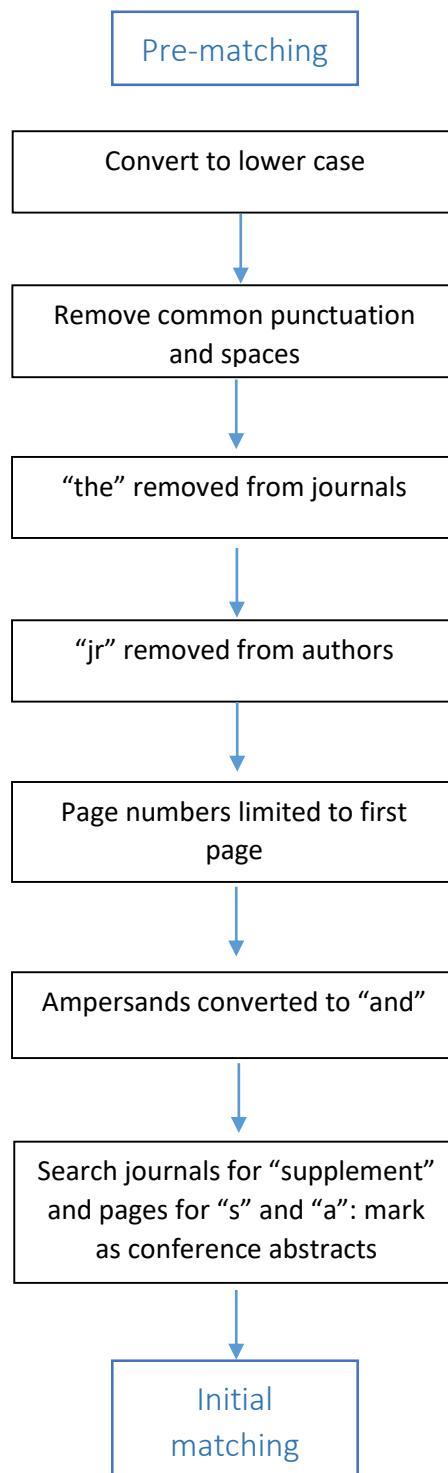
Appendix 2.4.3 Other Changes

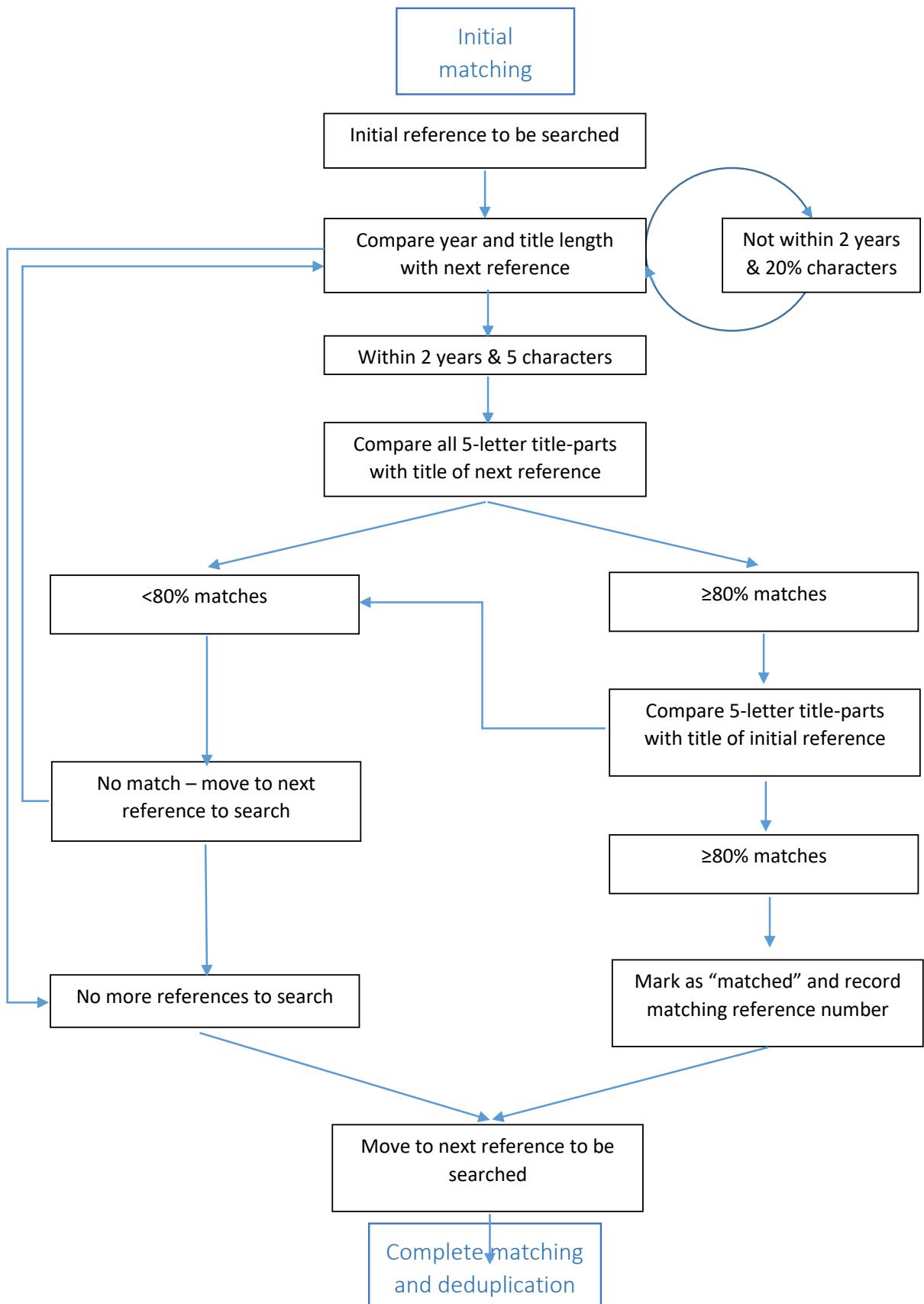
To improve the sensitivity of the algorithm, instead of splitting titles into 10 equal fragments, the title was split into fragments 5 characters in length (with 0-4 extra characters ignored at the end of the title). This means that longer titles are split into more fragments, which can be useful; typographical errors will be more prevalent as titles increase in length, therefore a minimum percentage of matching fragments would work better with both short and long titles than using a total of 10 fragments and hard threshold of 8 fragments regardless of title length. In this iteration, rather than titles having to be within 5 characters to be considered, titles must be within 20% length; this works much better for longer titles.

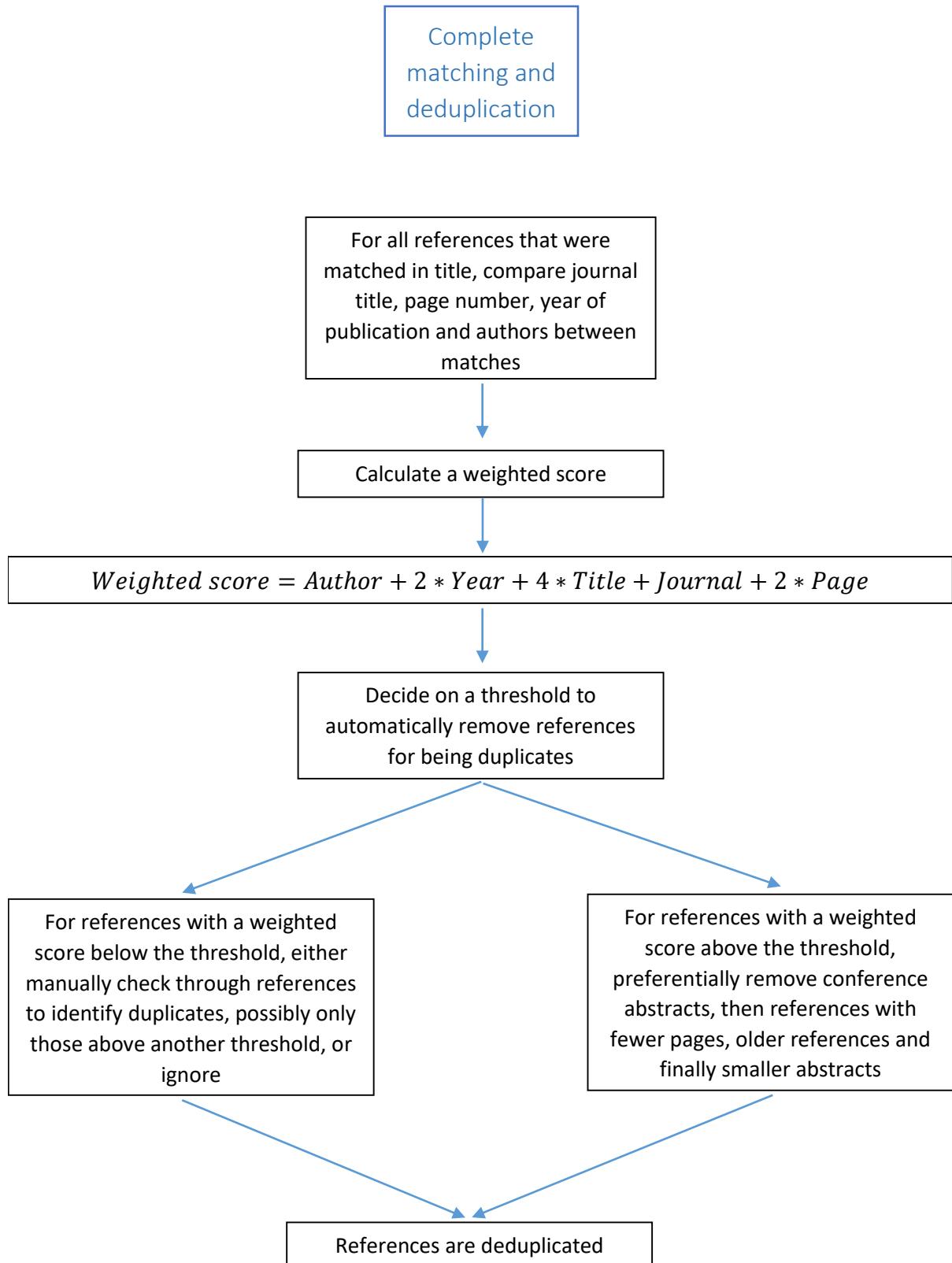
To maximize both sensitivity and specificity, it would appear to be best to automatically remove duplicates with a weighted score above a threshold, but manually search through the remained of the matches to determine if the references are in fact duplicates. This saves time in considering obvious duplicates, but makes it easier to find the remaining duplicates. A minimum threshold could also be specified to avoid manually searching through obvious non-duplicates.

The revised matching algorithm is detailed in the next pages as a flow chart, **Appendix Figure 2.1**. The key differences between the versions are in pre-match modifications, marking of conference abstracts, and using 5-character length title fragments rather than 10 title fragments.

Appendix Figure 2.1 Matching algorithm – version 2







Appendix 2.5 Algorithm Validation

Version 2 of the algorithm was validated using a subset of references from the scoping study; those references that included any of the following terms in their title or abstract (without spaces or capitalisation): *body mass index; BMI; body size; weight; height; anthropomet**; *obesity; obese*. These terms were chosen to include the subset of references which examined body-mass index and prostate cancer or PSA, a variable likely to be included in many studies and much easier to search for than age, which suffers from being a short word present in many non-age-related words, making it difficult to search for. As BMI was eventually chosen for further analysis, this also meant that in the systematic review no studies would be missed due to the algorithm.

The algorithm was applied to the 4,382 references that included the above terms to determine the number of references which could be automatically removed using different thresholds of weighting, as well as the reasons why any duplicates couldn't be found using an exact rather than probabilistic duplicate algorithm. The algorithm was applied both using pre-match modifications and not to ascertain the benefits of the modifications. Potential duplicates that were not exact matches without pre-match modifications were searched manually to determine if they were actual duplicates or false positive, and any differences between the duplicates were noted.

Appendix 2.5.1 Title-matching

Appendix Table 2.2 shows how many duplicates were found both using pre-match modifications and not, before calculating weighted scores. In summary, 2,252 references (51%) had been matched in 1,079 duplicate groups. For binary duplicates, pre-match modifications accounted increased the number of identical duplicates from 302 (30%) to 794 (79%) references.

The table shows the total number of references, duplicate groups, duplicate references and false positive references (and duplicate groups) overall and for binary duplicates (split into without and with pre-match modifications) and non-binary duplicates (just with pre-match modifications), with each split further into total, identical and non-identical matches.

Binary duplicates are matches of only 2 references, whereas non-binary duplicates can have 3 or more matched references. Non-binary duplicates are more difficult to classify, since when 3 references match, 2 may be identical matches and 1 may be non-identical, so one duplicate group provides information for both rows (identical and non-identical matches). Additionally, non-binary duplicates had an extra round of matching to identify the identical references within the duplicate group; references that didn't match exactly with any others were considered again just on the title.

Appendix Table 2.2 Duplicates found when matching titles only i.e. before calculating a weighted score

		Total references	Total duplicate groups	Total duplicate references	Total false positive references (groups)
Overall	Total non-matches	2130	-	-	-
	Total matches	2252	1077	1155	24 (13)
Binary duplicates					
Overall	Total matches	2004	1002	996	12 (6)
Without pre-match modifications	Identical matches	604	302	302	0 (0)
	Non-identical matches	1400	700	694	12 (6)
With pre-match modifications	Identical matches	1588	794	794	0 (0)
	Non-identical matches	416	208	202	12 (6)
Non-binary duplicates					
Overall	Total matches	248	75	159	12 (7)
With pre-match modifications	Identical matches	128	60	92	0 (0)
	Non-identical matches	120	65	67	12 (7)

Appendix 2.5.2 False positives (title-matching)

In total, 24 references (13 duplicate groups) were false positives, in that the titles were similar enough to be considered matches but the references were not duplicates. Reasons for false positives were as follows: 11 duplicates (61%) had titles with only a single word changed, 6 duplicates (33%) had extremely short, similar titles (range 16-30 characters) and one duplicate (6%) had a title with two additional words. All false positives had weighted scores below 7 (range 1 to 6); a weighted score of 7 or more would therefore result in a 100% specificity in this set of references.

Setting a minimum title length of 25 characters would remove all but one of the short-title false-positives, as four titles were “obesity and cancer” and one was “diet, obesity and cancer” – short titles have much more likelihood of returning false positives due to the reduced amount of information contained within. Less than 0.6% of all references had titles of 24 characters or less (after pre-match modifications), but contained 28% of the false positives (and only 0.3% of the true positives); restricting the algorithm would therefore improve specificity without impacting on sensitivity much.

False positives with one or two different words are harder to deal with, as allowing one word to change increase the sensitivity of the algorithm by a large proportion; 75 titles (6.5% of all duplicates) were only found because typographical errors were allowed. However, the weighted scores were all 4 or lower, and therefore it is highly unlikely that an automated algorithm would delete these references. It is important to note that one duplicate group had 10 references, with titles of: “*Cancer incidence among pesticide applicators exposed to [pesticide] in the agricultural health study*”; in this single group, there were 6 unique references, so 4 duplicates and 2 false-positives. Of the four duplicates, only 2 matched perfectly, as the middle initials of the other two duplicates were dealt with differently.

Appendix 2.5.3 Non-perfect duplicates

Appendix Table 2.3 shows the reasons why duplicates were not exact after pre-match modifications in each of the 5 categories: authors, year, title, journal and page numbers. Most of these differences could not have been altered with pre-match modifications, and therefore represent the increased number of duplicates that could be found using a probabilistic algorithm rather than an exact one.

Supplements, for example conference abstracts, were the most common reason for non-exact matching; supplements often had different authors and titles to the main article equivalents, and were published in different journals, possibly in different years. Even if other information was incorrect for supplements (e.g. typos in names etc.), the reference was only marked as being a supplement, with the assumption that at least some of the incorrect information was due to the authors of the supplement rather than the database of references.

While the algorithm marked many references as supplements (67 out of 2252 matched references), it was impossible to tell in many cases as no letters were added to the page numbers denoting “abstract” or “supplement”. As such, although supplement will be preferentially discarded automatically, it is necessary to also keep the newest reference with the largest number of pages, as this is likely the full paper. If necessary, the abstract with the longest abstract could be kept, as this will provide the most information on title/abstract screening.

Authors’ middle initials were different in 42 (36%) of non-exact duplicates - databases seem to record middle initials differently, some allowing multiple middle initials, some writing the initials in full - indicating a need for a name-matching algorithm to account for these differences. However, there were almost as many typographical errors (37 - 28%) - misspellings of authors’ names – as differences in middle initials, which could not be fixed.

Years of publication were generally correct, with only 4 typographical errors, i.e. different years of publication for the exactly the same article.

Titles often differentially included the language the paper was written in [in square brackets], with 19 (25%) non-exact duplicates; this could be incorporated into pre-match modifications by using a pre-defined list of common languages in square brackets and removing them from the title. Typographical errors were also relatively common, accounting for 12 (16%) of the non-exact duplicate titles.

Journal titles were the largest cause of non-exact duplicates, having differences in 145 matches (50% of the total number of non-exact duplicates). Apart from supplements, taglines (30 matches, 21%) and abbreviations (19 matches, 13%) were the most common causes of differences; taglines are additions

to the journal title, such as “an international journal”, and abbreviations are the differential use of the journal title acronym, such as JAMA for the Journal of the American Medical Association.

Page numbers were missing for many references (24 matches, 21%), and were incorrect in 8 matches (7%).

Appendix Table 2.3 Reasons why duplicates were not exact after pre-match modifications

Author		Year		Title		Journal		Page									
Supplement	48	Supplement	25	Supplement	32	Supplement	77	Supplement	76								
Middle Initial	42	Typo	4	[language]	19	Tagline	30	Missing	24								
Typo	37	Review	1	Typo	12	Abbreviation	19	Typo	8								
Editorial	4			Roman numeral	3	(country)	7	Editorial	3								
Suffix	2			Element	3	[language]	4	Review	1								
Review	1			Additional text	2	Typo	2	Roman numerals	1								
				Review	1	Editorial	2	Letters	2								
				Editorial	1	(online)	2										
				[number]	1	Review	1										
				Greek letter	1	Volume	1										
Total: 134		Total: 30		Total: 75		Total: 145		Total: 115									
Total non-exact duplicates after pre-match modifications: 269																	
<ul style="list-style-type: none"> • Supplements: conference abstracts or other articles published in addition to a main article • Middle initials: either in full or omitted entirely, e.g. A. Hart versus A. B. Hart versus A. Ben Hart • Typos: typographical errors - any differences not due to supplements, commonly misspelled words and names • Editorials: editorials of the main article • Suffixes: omitted, e.g. A. Hart III versus A. Hart • Reviews: reviews of the main article • [language]: addition of [language] to titles, e.g. [Chinese] • Roman numeral: differential use of roman numerals, e.g. IV versus 4 • Element: differential positioning of radioactivity and element in titles, e.g. Lu-177 versus 177Lu • Additional text: additional text in the title, e.g. “results from the SEARCH database” • [number]: addition of a number in brackets to a title, e.g. [1] • Greek letter: differential use of anglicised Greek letters, e.g. “beta” versus “B” • Tagline: differential use of the tagline for journals, e.g. “official journal of the epidemiology society” • Abbreviation: differential use of the abbreviated journal title • (country): addition of the country of origin for a journal in brackets, e.g. (China) • Volume: addition of the volume of the journal • Missing: page numbers that are not recorded • Letters: differential use of letters in page numbers, e.g. 63S versus 63 																	

Appendix 2.5.4 Weighted scores

Appendix Table 2.4 shows the weighted scores of the duplicates, with the number and percentage of true and false duplicates. References were manually determined to be true duplicates or not; one reference in each duplicate group was left blank to indicate the reference that would remain after deduplication. There was no meaningful difference between those marked and those not marked (mean weighted scores for unmarked = 9.55 and for duplicates = 9.56).

The references which were matched incorrectly to another reference (false positives) had an average weighted score of 3.44, with a maximum score of 6. In total, 5.3% of actual duplicates (n=62) had weighted scores below 7 and just over 85.5% of duplicates (n=988) had a weighted score of 10. If titles with fewer than 26 characters were removed, the average weighted score of false positive became 3.15 with a maximum score of 4, with actual duplicates virtually unchanged.

Therefore, assuming no unfound duplicates, the sensitivity and specificity of the algorithm for identifying and removing duplicates automatically with a weighted score of 7 was 94.7% (85.5% with a perfect score of 10) and 100% respectively.

Appendix Table 2.4 The weighted scores of references marked manually as duplicates, as well as those marked as not duplicates (false positives), both with and without titles of less than 25 characters removed. Shading indicates proposed cut-off for automatically removing duplicates. Percentages are cumulative.

Weighted score	Duplicates		False positives		Titles <25 characters removed			
					Duplicates		Duplicates	
	Total	%	Total	%	Total	%	Total	%
0	6	0.5	3	16.7	6	0.5	2	15.4
1	5	1.0	2	27.8	5	1.0	2	30.8
2	5	1.4	3	44.4	5	1.4	3	53.8
3	7	2.0	4	77.8	7	2.0	4	84.6
4	7	2.6	4	88.9	7	2.6	2	100
5	11	3.5	0	88.9	11	3.6	0	100
6	21	5.4	2	100	21	5.4	0	100
7	17	6.8	0	100	17	6.9	0	100
8	35	9.9	0	100	34	9.8	0	100
9	53	14.5	0	100	51	14.2	0	100
10	988	100	0	100	988	100	0	100
Total	1155	-	18	-	1152	-	13	-

Other weighted scores were considered, and their maximum sensitivity for 100% specificity was compared to the sensitivity of the pre-designed weighted score sensitivity of 94.7%:

1. Equal weighting for all variables: 92.8%
2. Weighting variables in order of the number of differences in **Table 3**: 94.5%

$$\text{Weighted score} = \frac{2 * \text{Author} + 5 * \text{Year} + 4 * \text{Title} + \text{Journal} + 3 * \text{Page}}{1.5}$$

3. Weighting variables inversely to above: 88.7%

$$\text{Weighted score} = \frac{4 * \text{Author} + \text{Year} + 2 * \text{Title} + 5 * \text{Journal} + 3 * \text{Page}}{1.5}$$

4. Weighting variables depending on the exact number of differences in **Table 3**: 92.8%

$$Weighted\ score = \frac{134 * Author + 30 * Year + 75 * Title + 145 * Journal + 115 * Page}{499}$$

5. Weighting variables inversely to above: 95.2%

$$Weighted\ score = 10 * \left(\frac{\frac{Author}{134} + \frac{Year}{30} + \frac{Title}{75} + \frac{Journal}{145} + \frac{Page}{115}}{\frac{1}{134} + \frac{1}{30} + \frac{1}{75} + \frac{1}{145} + \frac{1}{115}} \right)$$

6. Weighting variables similarly to above, but with whole numbers: 94.6%

$$Weighted\ score = Author + 5 * Year + 2 * Title + Journal + Page$$

Given the relative similarity between these different weightings and the complexity of the fifth weighting, the initial weighted score (score 2) was chosen as the most parsimonious score.

Appendix 2.6 Discussion

Deduplication is an important step in reducing bias in Sankey diagrams, and could save time in systematic reviews. Typographical errors and other differences between databases make exact deduplication imperfect, which can lead to bias and inefficiency. By using a probabilistic matching algorithm, the number of duplicates found can be maximized, at the cost of a tiny number of false positives.

One of the greatest advantages of using a probabilistic matching algorithm is that a limited number of typographical errors can be ignored. Another advantage is that supplements can have similar but different titles to the main article equivalents, which can't be found without using an algorithm that can handle small changes. One pertinent example is the supplement entitled "*The association between obesity-related plasma hemodilution and tumor markers concentration and the clinical significance among healthy men*", which was matched to the correct main article, entitled "*The Clinical Significance in Healthy Men of the Association Between Obesity Related Plasma Hemodilution and Tumor Marker Concentration*".

However, there is a balance between sensitivity and specificity when considering how many typographical errors to allow: two sets of two articles were missed as duplicates because the "United States" was abbreviated to "US" in the titles, resulting too many unmatched 5-letter fragments. More duplicates may have been missed as a result of similar problems; however, this problem cannot be overcome without a different approach to deduplication. Allowing two titles to diverge by more than 20% will necessarily decrease the specificity of the algorithm; in deduplication specificity needs to be as high as possible, especially if the algorithm is automated, otherwise references may be dropped without reason.

Because of this, the sensitivity of version 2 of the algorithm will be smaller than 94.7% (or indeed 85.5%) as there were at least two duplicates not found by the algorithm. No other duplicates were found when manually searching, but it is impossible to state categorically that no more duplicates existed. The only way to improve sensitivity without analysing abstracts for information using text mining would be to remove the restriction on titles to be at least 80% similar. While this could be accomplished by creating a weighted score for every reference against every other reference and using the highest score found, but this would drastically increase computing time with only a very modest increase in sensitivity.

In summary, the probabilistic matching deduplication algorithm improves upon existing deduplication algorithms by allowing typographical errors (and other changes) in the variables commonly used to

determine duplication. In the subset of references examined, perfect specificity was achieved with sensitivities above 80-90% (assuming only a small number of unfound duplicates and a small number of references with short titles).

The probabilistic matching algorithm has been used to reduce manual deduplication of references internally, with no problems encountered.

Appendix 3: Question Sheet Given to ProtecT Lead Nurses

Which variables have the strongest association with both prostate cancer and prostate-specific antigen (PSA)?

Background: We are looking for variables that have a large effect on PSA levels and that are also associated with the risk of being diagnosed with prostate cancer. The reason we are doing this is to try to develop method of making PSA testing a more accurate tool for diagnosing prostate cancer. Once we have decided which variables have the largest effect on both PSA and prostate cancer we will develop a way of removing the effect of the variable from the PSA level, while still keeping the effect the variable has on prostate cancer.

I have printed an overview of my PhD, in case you are interested in the overall project. These will be available at the end of this session.

Instructions: Below, there is a table of the top 12 variables likely associated with both prostate cancer and PSA. Please would you **rank up to five variables you consider the most associated with BOTH prostate cancer and PSA 1 to 5**, with 1 having the strongest associations and 5 the lowest. If you think less than 5 variables are strongly associated, then please rank those you think are associated. If you have other suggestions please write them in the “Other” section below, and rank the suggestion appropriately.

There are no right or wrong answers – as an expert, any information or insights you can provide would be greatly appreciated.

Variable	Rank
Age	
Androgens (testosterone, oestrogen)	
Body mass index (BMI)	
Benign prostatic hypertrophy (BPH)	
Diet	
Ethnicity	
Family history	
Finasteride (and other urinary drugs)	
Insulin-like growth factors (IGF)	
Kallikreins	
Other cancers	
Tumour suppressors	
Other (please detail)	

***Note: Score = the sum of the 10 individual scores, where each response = 6-rank (so rank 1 gives a score of 5)**

Thank you very much for your time!

Background questions

We are also interested in some basic information about you – this is anonymous, and **completely optional**.

1. Describe your specialty, e.g. urology, oncology, uro-oncology etc.

2. List your qualifications

3. How long have you worked in prostate cancer?

4. How long have you worked in healthcare or research?

Thank you again for your time!

Appendix 4: Information Sheet for ProtecT Lead Nurses

Prostate Cancer and PSA: Examining the role of mediation in the associations of individual characteristics with prostate specific antigen (PSA) and prostate cancer risk

Sean Harrison's PhD

Overview

Many papers have looked at the relationship between prostate specific antigen (PSA) levels in men and individual characteristics, such as age, body mass index (BMI), ethnicity, serum levels of various blood components, such as IGF, and components of diet. However, these exposures often also have an effect on prostate cancer risk, and have relationships with each other. This makes it difficult to identify just how strongly an exposure is related to PSA and to the risk of prostate cancer.

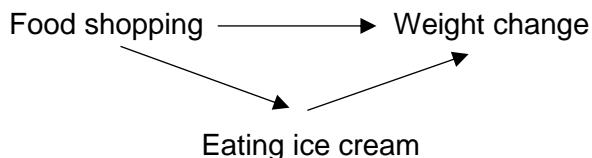
However, it is important to try to reduce the effect the exposures have on PSA, without influencing their effect on prostate cancer risk, because otherwise the PSA test will be biased. For example, a man with a high BMI may have a lower BMI because BMI decreases PSA, as having more blood decreases the PSA level. However, a high BMI also might increase prostate cancer risk due to higher levels of certain hormones. The aim is to remove the effect BMI has on PSA, without removing its effect on prostate cancer.

The first step is to identify exposures that have an effect on both prostate cancer and on PSA. We have made a Sankey Diagram, which takes information from all academic papers related to prostate cancer or PSA, and displays how many papers there are which mention selected exposures. However, just knowing that there are lots of papers looking at a particular exposure and prostate cancer and PSA doesn't mean that there is an association. For instance, there may have been 99 papers looking at an exposure, but then 1 paper later that proved there was no association. This would show up the same as an exposure that had 100 papers proving an association.

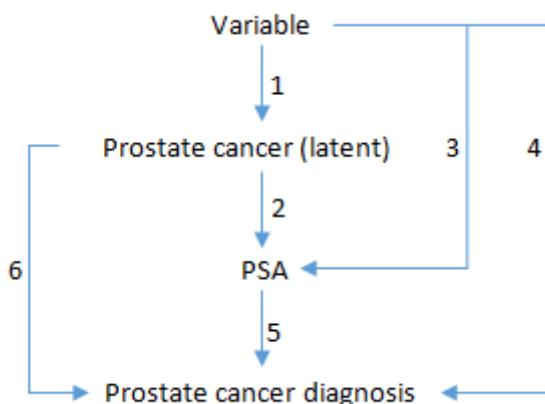
Experts can make the difference between looking at the exposure that has nothing to do with prostate cancer and PSA, and the exposure that matters most.

Mediation

Mediation is the effect a variable has on an outcome acting through another variable. In the figure below, shopping for food has an effect on weight change, and also on the mediator variable eating ice cream. The ice cream is a mediator because it also has an effect on gaining weight. Here, the effect of shopping on weight change goes through eating ice cream. If I stopped buying ice cream, and didn't replace it with something equally unhealthy, then food shopping wouldn't increase my weight by quite so much.



The primary aim of this PhD is to separate several variables' effect on PSA from their effect on PSA which acts through the effect on prostate cancer risk. The figure below shows how the variables, PSA and prostate cancer inter-relate in a directed acyclic graph (DAG).



1 = the effect of the variable on prostate cancer

2 = the effect of prostate cancer on PSA

3 = the effect of the variable on PSA – this is what I'm trying to find, so I can remove it when a man received a PSA test

4 = the effect of the variable on the chance of getting a prostate cancer diagnosis (e.g. a comorbidity that makes a PSA test more likely, like high BMI or age)

5 = the effect of screening for prostate cancer using PSA, which includes getting a prostate biopsy as PSA rarely leads to a prostate cancer diagnosis by itself

6 = the effect of having prostate cancer on getting a diagnosis of prostate cancer

Appendix 5: Risk of Bias Questions

Each section (1-6) requires a risk of bias; low, medium, high, critical or unclear. The questions in each section help in identifying risk of bias, and are designed so a “yes” answer implies some risk of bias. An overall risk of bias was given based on the maximum risk of bias each section could contribute.

1. Bias due to confounding
 - a. Is the mean age of the cases/controls NOT within 5 years, or is age NOT adjusted for in the analysis?
 - b. Are there any other confounders likely to cause bias that have not been accounted for? SES, ethnicity etc.
2. Bias in selection of participants?
 - a. Were the participants recruited in a way likely to cause bias? e.g. controls recruited in an obesity clinic
 - b. Are the participants not representative of the general population of men in any way? e.g. All men had pituitary tumours
 - c. Are there any other ways in which the selection of participants could cause bias?
3. Bias due to missing data?
 - a. Cohort: Was the follow-up time insufficient to allow for a diagnosis of prostate cancer? e.g. less than 5 years
 - b. Cohort: Were there differences in baseline measures of age and BMI between those lost to follow-up and not lost?
 - c. Cohort: Were the baseline measures of age and BMI NOT presented for those lost to follow-up?
 - d. Case-control: Were there non-responders that could have caused bias in the study?
4. Bias in measurement of outcome?
 - a. Was prostate cancer or PSA measured in a way that could have caused bias? e.g. poor follow-up of obese men
 - b. Could the participants without prostate cancer have had cancer? e.g. if controls were not assessed for cancer
5. Bias in measurement of exposure?
 - a. Was BMI measured differently between cases and controls? e.g. controls were self-reported, cases were measured
6. Bias due to selective reporting?
 - a. Was there selective reporting in the study which could have caused bias?

Appendix 6: MOOSE Guidelines for Meta-Analyses and Systematic Reviews of Observational Studies

294

Guidelines	Section reported	Notes
Introduction: Present...		
The clinical problem	4.1	
The hypothesis	4.1	
A statement of objectives that includes the study population, the condition of interest, the exposure or intervention, and the outcome(s) considered	4.1	There were no limits placed on the population of interest
Sources: Describe...		
Qualifications of searchers (eg, librarians and investigators)	NA	I conducted the search and provided all search terms
Search strategy, including time period included in the synthesis and keywords	4.2.1	
Effort to include all available studies, including contact with authors	4.2.3	No contact with authors was sought
Databases and registries searched	4.2.1	
Search software used, name and version, including special features used (eg, explosion)	NA	Online database searches were performed
Use of hand searching (eg, reference lists of obtained articles)	4.2.1	Reference lists were not searched, reference lists of previous meta-analyses were
List of citations located and those excluded, including justification	NA	All citations contained in database, which could be accessed if requested
Method of addressing articles published in languages other than English	4.2.1	
Method of handling abstracts and unpublished studies	4.2.3	
Description of any contact with authors	NA	No authors contacted
Study selection: Describe...		
Types of study designs considered	4.2.1	
Relevance or appropriateness of studies gathered for assessing the hypothesis to be tested	NA	All studies assessing BMI-PCa/PSA association relevant
Rationale for the selection and coding of data (eg, sound clinical principles or convenience)	4.2	
Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)	4.2	
Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)	4.2.4	

Guidelines	Section reported	Notes
Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results	4.2.4	
Assessment of heterogeneity	4.3.4	
Statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	4.3	
Results: Present...		
A graph summarizing individual study estimates and the overall estimate	F4.3/6/9	
A table giving descriptive information for each included study	T4.2/4/6	
Results of sensitivity testing (eg, subgroup analysis)	F4.3/5/6/8/9/11	Before versus same-time measurement of BMI/PCa (forest plot), included and excluded studies (albatross plot)
Indication of statistical uncertainty of findings	NA	All results are given with 95% CIs
Discussion: Present...		
Strengths and weaknesses	4.5.4	
Potential biases in the review process (eg, publication bias)	4.5.4	
Justification for exclusion (eg, exclusion of non-English-language citations)	4.5.4	Only exclusion was for studies with a critical risk of bias
Assessment of quality of included studies	4.5.4	
Consideration of alternative explanations for observed results	NA	Examined in later considering results from MR and IPD meta-analysis
Generalization of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)	NA	Examined in later considering results from MR and IPD meta-analysis
Guidelines for future research	NA	Examined in later considering results from MR and IPD meta-analysis
Disclosure of funding source	NA	Thesis has an acknowledgement page that details funding

Appendix 7: Calculating the SD of log-PSA using the ratio of geometric means

Two studies (35,195) presented the ratio (or percentage change) of geometric means of PSA with 95% confidence intervals (CIs) for two levels of BMI with respect to a baseline level of BMI (overweight and obese versus normal-weight). These studies effectively conducted two linear regressions:

$$\ln(\text{PSA}) = \beta_0 + \beta_1 \times \text{BMI}_{\text{over}} + \varepsilon \quad (17)$$

$$\ln(\text{PSA}) = \beta_0 + \beta_1 \times \text{BMI}_{\text{obese}} + \varepsilon \quad (18)$$

where BMI_{over} and $\text{BMI}_{\text{obese}}$ equal 1 if the man is overweight or obese respectively, and 0 otherwise. Only normal weight and overweight men are included in the first equation, and only normal weight and obese men are included in the second equation.

n_i = number of participants in group i
 n_{0i} = number of participants in groups 0 and i combined
 μ_{yi} = mean log-PSA in group i
 μ_{y0i} = mean log-PSA in groups 0 and i combined
 σ_{yi}^2 = variance of log-PSA in group i
 σ_{0i}^2 = variance of log-PSA in groups 0 and i combined
 σ_{0i}^2 = variance of BMI in groups 0 and i combined
 p_i = proportion of participants in group i

Where $i = 0$ for normal weight, 1 for overweight and 2 for obese BMI categories.

If the subscript is $0i$, that means the quantity is for two groups combined, either 01 for normal weight and overweight men, or 02 for normal weight and obese men.

β_{0i} = linear regression coefficient for group i (1 for overweight, 2 for obese)

SE_{0i} = standard error of regression coefficient for group i (1 for overweight, 2 for obese)

$\varphi(y, \mu, \sigma^2)$ = the density of log-PSA in a normal distribution, with mean μ and variance σ^2

First, the geometric mean ratios were transformed back from percentage changes in PSA to beta coefficients for the difference in log-PSA between BMI levels. This required transforming percentages into proportions (so $-5\% = 0.95$) and taking the log of the result ($\ln(0.95) = -0.051$). The 95% CIs were used to calculate the SE of each estimate. The geometric means of PSA were logged to give the arithmetic means of log-PSA. Our aim is to find the SD of log-PSA for each level of BMI (σ_{yi}).

The calculation of the SE for the coefficient for overweight or obese (i) is (255):

$$SE_{0i} = \sqrt{\frac{\sigma_{y0i}^2 - \beta_{0i}^2 \sigma_{x0i}^2}{N \sigma_{x0i}^2}} \quad (19)$$

where σ_{y0i}^2 is the variance of log-PSA in the normal-weight and either overweight or obese groups, and σ_{x0i}^2 is the variance of BMI for the same groups.

The variance of the binary BMI variable (σ_{x0i}^2) can be calculated from the number of participants in each group, allowing us to calculate the SD of log-PSA when considering both the normal and overweight or obese groups (σ_{y0i}^2) from the SE of the beta coefficients. The variance of a binary variable is (255) a function of the proportion of men with the outcome, so the proportion of overweight or obese men (p_i):

$$\sigma_{x0i}^2 = p_i(1 - p_i) \quad (20)$$

As $p_i = \frac{n_i}{n_{0i}}$ and $(1 - p_i) = \frac{n_0}{n_{0i}}$:

$$\sigma_{x0i}^2 = \frac{n_0 n_i}{n_{0i}^2} \quad (21)$$

This can be substituted into (19) to find σ_{y0i}^2 using the results of the regression of BMI groups 0 and i together:

$$SE_{0i} = \sqrt{\frac{\sigma_{y0i}^2 - \beta_{0i}^2 \frac{n_0 n_i}{n_{0i}^2}}{n_{0i} \frac{n_0 n_i}{n_{0i}^2}}} \quad (22)$$

$$SE_{0i}^2 \frac{n_0 n_i}{n_{0i}} = \sigma_{y0i}^2 - \beta_{0i}^2 \frac{n_0 n_i}{n_{0i}^2} \quad (23)$$

$$\sigma_{y0i}^2 = SE_{0i}^2 \frac{n_0 n_i}{n_{0i}} + \beta_{0i}^2 \frac{n_0 n_i}{n_{0i}^2} \quad (24)$$

$$\sigma_{y0i}^2 = \frac{n_0 n_i}{n_{0i}^2} (n_{0i} SE_{0i}^2 + \beta_{0i}^2) \quad (25)$$

The variance of log-PSA for both the normal and overweight or obese groups combined is not sufficient by itself to calculate the variance of log-PSA in each group. This is because we assume log-PSA has a normal distribution in each level of BMI, and that BMI has a potential association with log-PSA. Therefore, when both the normal and overweight or obese groups are considered, the overall variance will likely be larger than the variance in either group alone, since the mean log-PSA will be different in each group.

However, we can estimate the variances for each level of BMI. The variance of log-PSA for a single group is (255):

$$\sigma_y^2 = \frac{\sum(y_p - \mu_y)^2}{n^2} \quad (26)$$

where y_p is the value of log-PSA for each person p . Therefore, the variance of the log-PSA for two groups (normal weight and overweight or obese) can also be calculated by assuming a distribution of log-PSA is normal within each category of BMI:

$$\sigma_{y0i}^2 = \frac{\sum(y_{p0} - \mu_{y0i})^2 + \sum(y_{pi} - \mu_{y0i})^2}{n_{0i}^2} \quad (27)$$

where the mean log-PSA for both categories of BMI (0 and i) is subtracted from each person p in each BMI category and divided by the total number of participants in both groups. The mean of log-PSA for two combined groups can be calculated:

$$\mu_{y0i} = \frac{n_0\mu_0 + n_i\mu_i}{n_{0i}} \quad (28)$$

However, summation of individual values of log-PSA is impossible without the IPD. Instead, we assume that log-PSA has a normal distribution within each category of BMI, with a separate mean and unknown SD (which is the value we are attempting to estimate), and use integration to replace the summation. The probability density of the two log-PSA distributions (in BMI categories 0 and i) approximates the proportion of participants with any log-PSA value. We then multiply this proportion by log-PSA minus the mean log-PSA for both BMI groups, and the integration sums this result across the complete range of log-PSA values. There are two integration terms, one for each category of BMI, and these must be weighted by the number of participants in each group:

$$\begin{aligned} \sigma_{y0i}^2 &= \int \left(\frac{n_0}{n_{0i}} \times \varphi(y, \mu_{y0}, \sigma_{y0}^2) \times (y - \mu_{y0i})^2 \right) dx \\ &\quad + \int \left(\frac{n_i}{n_{0i}} \times \varphi(y, \mu_{yi}, \sigma_{yi}^2) \times (y - \mu_{y0i})^2 \right) dy \end{aligned} \quad (29)$$

$$\sigma_{y0i}^2 = \int \left((y - \mu_{y0i})^2 \left(\frac{n_0}{n_{0i}} \times \varphi(y, \mu_{y0}, \sigma_{y0}^2) + \frac{n_i}{n_{0i}} \times \varphi(y, \mu_{yi}, \sigma_{yi}^2) \right) \right) dy \quad (30)$$

Since the mean (μ_{yi}) log-PSA for each group is known, and the mean (μ_{y0i}) and variance (σ_{y0i}^2) of log-PSA for two groups combined are calculable, σ_{yi}^2 is the only remaining unknown in this equation. However, there are three categories of BMI and only two linear regressions. We must therefore assume something about one of the variances of log-PSA: we chose to assume that the variance of the log-PSA in the overweight category of BMI was equal to the mean of the variances of normal weight and obese categories:

$$\sigma_{y1}^2 = \frac{\sigma_{y0}^2 + \sigma_{y2}^2}{2} \quad (31)$$

We therefore have two equations and only two missing values, σ_{y1}^2 and σ_{y3}^2 :

$$\sigma_{y01}^2 = \int \left((y - \mu_{y01})^2 \left(\frac{n_0}{n_{01}} \times \varphi(y, \mu_{y0}, \sigma_{y0}^2) + \frac{n_1}{n_{01}} \times \varphi\left(y, \mu_{y1}, \frac{\sigma_{y0}^2 + \sigma_{y2}^2}{2}\right) \right) \right) dy \quad (32)$$

$$\sigma_{y02}^2 = \int \left((y - \mu_{y02})^2 \left(\frac{n_0}{n_{02}} \times \varphi(y, \mu_{y0}, \sigma_{y0}^2) + \frac{n_2}{n_{02}} \times \varphi(y, \mu_{y2}, \sigma_{y2}^2) \right) \right) dy \quad (33)$$

The normal density function does not integrate with a closed-form equation, and thus an iterative fitting algorithm must be used to find values of σ_{y0}^2 and σ_{y2}^2 that fit both equations. The algorithm cycles gradually through values of σ_{y0}^2 and σ_{y2}^2 , starting at reasonable estimates (such as $\sigma_{y01}^2 = 0.2$), until values of σ_{y0}^2 and σ_{y2}^2 are found which satisfy both equations above to a reasonable number of significant figures. The value of σ_{y1}^2 was calculated from the other two variances, and all were square-rooted to give the SD of log-PSA for each level of BMI. When the SDs of log-PSA were calculated for each level of BMI, VWLS could be used to estimate the change in log-PSA for a 5 kg/m² increase in BMI.

Appendix 8: Table showing the effect alleles, GWAS effect size and allele frequencies of SNPs included in the BMI, PSA, log-PSA and prostate cancer genetic risk scores

SNP*	EA	OA	GWAS EA	GWAS effect	EE	EO	OO	EAF	MAF
BMI SNPs									
rs10150332	G	A	C	0.13	1,924	13,192	24,467	0.215	0.215
rs1016287	A	G	T	0.02	5,776	28,018	34,810	0.288	0.288
rs10167079	G	A	G	0.024	55,348	12,542	719	0.898	0.102
rs10182181	G	A	G	0.023	15,140	34,018	19,428	0.469	0.469
rs10269783	A	G	A	0.014	10,511	32,731	25,210	0.393	0.393
rs10499694	A	G	A	0.015	16,368	34,397	17,848	0.489	0.489
rs10540	G	A	G	0.028	51,676	15,741	1,176	0.868	0.132
rs10733682	A	G	A	0.023	15,915	34,380	18,323	0.482	0.482
rs10742752	G	A	C	0.014	25,559	32,701	10,348	0.611	0.389
rs10760279	A	C	T	0.022	10,614	32,470	25,128	0.394	0.394
rs10761785	C	A	G	0.019	15,889	34,226	18,498	0.481	0.481
rs10938353	G	A	G	0.029	47,738	18,943	1,942	0.834	0.166
rs10975870	G	A	G	0.024	5,695	27,978	34,930	0.287	0.287
rs11057405	G	A	G	0.033	55,584	12,328	712	0.9	0.1
rs11084753	G	A	G	0.06	30,217	30,241	7,529	0.667	0.333
rs1109114	G	A	C	0.021	23,468	33,331	11,806	0.585	0.415
rs11126666	A	G	A	0.026	4,966	26,962	36,661	0.269	0.269
rs11150911	A	C	A	0.015	5,961	28,401	34,255	0.294	0.294
rs11165643	A	G	T	0.021	23,298	33,500	11,759	0.584	0.416
rs11168854	G	C	G	0.016	16,174	18,256	5,163	0.639	0.361
rs11170468	A	C	A	0.02	40,018	24,722	3,880	0.763	0.237
rs11583200	G	A	C	0.018	10,056	32,563	25,988	0.384	0.384
rs11607976	G	A	C	0.022	32,387	29,530	6,700	0.687	0.313
rs11611246	A	C	T	0.024	2,888	22,264	42,994	0.206	0.206
rs11672550	A	G	T	0.024	22,236	33,380	12,925	0.568	0.432
rs1167827	G	A	G	0.023	21,767	33,727	13,108	0.563	0.437
rs11688816	G	A	G	0.019	19,842	33,862	14,906	0.536	0.464
rs11727676	A	G	T	0.04	56,291	11,719	611	0.906	0.094
rs11771526	G	A	G	0.027	626	11,656	56,099	0.094	0.094
rs11787111	G	A	G	0.053	64,950	3,225	44	0.976	0.024
rs11847697	A	G	T	0.17	142	5,459	63,026	0.042	0.042
rs11866815	G	A	C	0.024	22,705	14,609	2,280	0.758	0.242
rs11997175	G	A	C	0.014	16,592	34,239	17,786	0.491	0.491
rs12286929	G	A	G	0.022	19,340	34,220	15,044	0.531	0.469
rs12401738	A	G	A	0.026	8,829	30,883	28,897	0.354	0.354
rs12429545	A	G	A	0.035	1,175	15,335	52,068	0.129	0.129
rs12446632	G	A	G	0.039	50,781	16,440	1,303	0.861	0.139
rs12450239	A	G	A	0.026	4,108	25,108	38,866	0.245	0.245
rs12454712	G	A	C	0.017	10,527	32,556	25,496	0.391	0.391
rs12546962	G	A	G	0.016	4,640	26,295	37,659	0.259	0.259
rs12566985	G	A	G	0.022	13,321	33,839	21,136	0.443	0.443
rs12680842	A	G	A	0.015	30,684	30,278	7,659	0.668	0.332
rs12894211	G	A	C	0.021	39,434	25,118	4,072	0.758	0.242
rs12900158	G	A	C	0.083	65,901	2,691	37	0.98	0.02
rs12940622	G	A	G	0.021	21,926	33,715	12,975	0.565	0.435
rs12961799	G	A	C	0.016	5,499	27,714	35,340	0.282	0.282
rs13078960	C	A	G	0.026	2,795	21,714	44,110	0.199	0.199
rs13191362	A	G	A	0.03	53,003	14,631	977	0.879	0.121
rs13201877	G	A	G	0.028	1,377	16,690	50,551	0.142	0.142

SNP*	EA	OA	GWAS EA	GWAS effect	EE	EO	OO	EAF	MAF
rs134871	G	A	C	0.013	18,034	34,083	16,480	0.511	0.489
rs1436351	A	C	T	0.016	38,624	25,726	4,247	0.751	0.249
rs1441264	A	G	A	0.018	25,034	32,675	10,918	0.603	0.397
rs1502337	A	G	T	0.014	8,535	31,508	28,579	0.354	0.354
rs150992	A	G	A	0.017	34,611	27,837	5,897	0.71	0.29
rs1514175	A	G	A	0.07	12,441	33,631	22,544	0.426	0.426
rs1546924	A	G	T	0.014	15,593	34,364	18,640	0.478	0.478
rs1557765	G	A	C	0.015	25,742	32,454	10,403	0.612	0.388
rs1558902	A	T	A	0.081	23,862	32,970	11,714	0.589	0.411
rs16907751	G	A	C	0.047	54,972	12,895	743	0.895	0.105
rs17001654	G	C	G	0.032	49,362	17,591	1,671	0.847	0.153
rs17024393	G	A	C	0.061	65	3,861	64,652	0.029	0.029
rs17094222	G	A	C	0.031	3,013	22,982	42,627	0.211	0.211
rs17203016	G	A	G	0.023	2,695	21,454	44,463	0.196	0.196
rs17513613	G	A	C	0.015	7,435	30,326	30,851	0.329	0.329
rs17522122	A	C	T	0.019	14,858	33,863	19,412	0.467	0.467
rs17724992	A	G	A	0.023	37,549	26,503	4,538	0.741	0.259
rs1808579	G	A	C	0.022	19,233	34,153	15,193	0.529	0.471
rs1928295	A	G	T	0.026	21,277	33,707	13,630	0.556	0.444
rs1979755	C	G	C	0.019	21,220	33,594	13,791	0.554	0.446
rs200810	A	G	T	0.014	27,042	31,883	9,702	0.626	0.374
rs2033529	G	A	G	0.02	5,757	28,331	34,517	0.29	0.29
rs2033732	G	A	C	0.019	38,501	25,936	4,180	0.75	0.25
rs205262	G	A	G	0.027	5,399	27,940	35,270	0.282	0.282
rs206936	G	A	G	0.06	2,836	21,850	43,933	0.201	0.201
rs2080454	C	A	C	0.017	10,518	32,173	25,917	0.388	0.388
rs2121279	A	G	T	0.029	1,325	15,760	51,067	0.135	0.135
rs2124499	G	C	G	0.014	27,075	31,809	9,677	0.627	0.373
rs2176598	A	G	T	0.024	4,255	25,628	38,744	0.249	0.249
rs2228213	G	A	G	0.019	16,460	18,009	5,132	0.643	0.357
rs2236176	A	G	T	0.02	875	10,001	28,719	0.148	0.148
rs2241423	G	A	G	0.13	41,518	23,591	3,511	0.777	0.223
rs2245368	G	A	C	0.034	2,228	19,532	46,240	0.176	0.176
rs2270204	C	A	G	0.022	4,174	25,030	39,408	0.243	0.243
rs2278491	G	A	C	0.028	53,417	14,203	998	0.882	0.118
rs2287019	G	A	C	0.15	44,101	21,705	2,733	0.802	0.198
rs2307022	A	G	A	0.016	7,656	30,512	30,458	0.334	0.334
rs2371767	C	G	C	0.018	5,680	27,366	35,541	0.282	0.282
rs2481665	A	G	T	0.019	21,803	33,742	13,075	0.564	0.436
rs253664	A	G	T	0.025	2,507	21,064	44,690	0.191	0.191
rs2579103	C	A	C	0.024	4,452	26,059	38,109	0.255	0.255
rs2612012	A	C	A	0.02	38,673	25,670	4,264	0.751	0.249
rs2770102	G	A	G	0.014	13,836	34,009	20,768	0.449	0.449
rs2815752	A	G	A	0.13	25,928	32,430	10,255	0.614	0.386
rs2836754	G	A	C	0.019	26,870	31,999	9,724	0.625	0.375
rs285575	A	G	A	0.019	15,793	34,197	18,422	0.481	0.481
rs2867125	G	A	C	0.31	46,432	19,980	2,204	0.822	0.178
rs29941	G	A	G	0.06	18,311	17,135	4,155	0.679	0.321
rs3101336	G	A	C	0.035	14,636	18,925	6,038	0.609	0.391
rs33439	A	G	T	0.023	24,637	32,857	11,023	0.599	0.401
rs355810	A	G	A	0.014	26,738	32,261	9,622	0.625	0.375
rs3783890	A	G	T	0.023	45,621	20,636	2,366	0.815	0.185
rs3810291	A	G	A	0.09	30,566	30,179	7,428	0.67	0.33
rs38313	G	A	G	0.021	17,110	34,496	16,998	0.501	0.499
rs4243830	G	A	C	0.03	827	13,473	54,233	0.11	0.11
rs4256980	G	C	G	0.019	8,712	31,207	28,676	0.354	0.354
rs4327120	A	G	T	0.04	55,004	12,876	737	0.895	0.105
rs4357530	G	A	G	0.019	7,154	30,018	31,432	0.323	0.323

SNP*	EA	OA	GWAS EA	GWAS effect	EE	EO	OO	EAF	MAF
rs4738873	A	G	A	0.02	1,910	18,410	48,299	0.162	0.162
rs4771122	G	A	G	0.09	2,305	14,648	22,628	0.243	0.243
rs4787491	G	A	G	0.016	18,770	34,161	15,646	0.523	0.477
rs4984406	A	G	T	0.019	17,426	33,895	17,281	0.501	0.499
rs4985155	A	G	A	0.021	17,119	17,701	4,473	0.661	0.339
rs4986044	G	A	C	0.016	19,638	33,925	15,041	0.534	0.466
rs543874	G	A	G	0.05	2,707	21,429	44,488	0.196	0.196
rs555267	A	C	T	0.016	7,164	29,959	31,459	0.323	0.323
rs571312	A	C	A	0.23	2,187	14,300	23,116	0.236	0.236
rs6465468	A	C	T	0.025	6,440	28,700	33,373	0.303	0.303
rs6477694	G	A	C	0.021	8,948	31,204	28,463	0.358	0.358
rs6504108	G	A	C	0.019	5,521	27,695	35,364	0.282	0.282
rs657452	A	G	A	0.022	10,319	32,708	25,489	0.389	0.389
rs6804842	G	A	G	0.02	22,964	33,453	12,199	0.578	0.422
rs6864049	G	A	G	0.016	19,565	33,859	15,111	0.532	0.468
rs6870983	G	A	C	0.018	41,737	23,391	3,475	0.779	0.221
rs6990042	C	A	G	0.019	15,400	34,188	18,995	0.474	0.474
rs7111341	A	G	T	0.021	5,431	27,835	35,313	0.282	0.282
rs713586	G	A	C	0.14	15,452	34,044	19,098	0.473	0.473
rs7138803	A	G	A	0.12	10,190	32,332	26,099	0.384	0.384
rs7141420	A	G	T	0.022	18,755	34,251	15,612	0.523	0.477
rs7143963	A	G	T	0.026	2,232	20,043	46,346	0.179	0.179
rs715	G	A	C	0.022	6,500	29,421	32,678	0.309	0.309
rs7164727	A	G	T	0.019	31,923	29,697	6,995	0.682	0.318
rs718948	G	A	C	0.016	8,524	31,330	28,757	0.353	0.353
rs7223966	G	A	G	0.017	35,231	27,799	5,589	0.716	0.284
rs7226371	G	A	G	0.026	1,733	18,008	48,865	0.157	0.157
rs7239883	G	A	G	0.023	10,660	32,438	25,517	0.392	0.392
rs7243357	A	C	T	0.025	46,119	20,180	2,316	0.819	0.181
rs7498665	G	A	G	0.15	6,309	18,869	14,163	0.4	0.4
rs751008	A	G	A	0.02	15,358	34,204	19,051	0.473	0.473
rs7550169	C	A	C	0.02	3,803	16,576	19,221	0.305	0.305
rs758747	A	G	T	0.026	2,798	15,422	21,362	0.265	0.265
rs7599312	G	A	G	0.026	36,340	27,182	5,109	0.728	0.272
rs7620457	G	A	G	0.018	2,904	21,725	43,992	0.201	0.201
rs7640424	G	A	C	0.015	33,450	28,771	6,369	0.697	0.303
rs7830341	G	A	G	0.068	63,020	5,503	105	0.958	0.042
rs7844647	A	G	T	0.016	37,402	26,496	4,728	0.738	0.262
rs7899106	G	A	G	0.044	161	6,177	62,182	0.047	0.047
rs7970953	A	G	A	0.015	5,737	28,182	34,695	0.289	0.289
rs8123881	G	A	G	0.023	1,349	16,263	51,009	0.138	0.138
rs889398	G	A	C	0.016	22,907	33,470	12,214	0.578	0.422
rs9275595	A	G	T	0.016	40,523	24,273	3,772	0.768	0.232
rs929354	A	G	T	0.023	27,918	31,549	9,155	0.637	0.363
rs9364687	C	A	G	0.02	23,046	33,243	12,291	0.578	0.422
rs943466	G	A	G	0.016	39,457	25,054	4,063	0.758	0.242
rs9507983	G	A	C	0.02	5,805	18,774	15,017	0.384	0.384
rs9540493	A	G	A	0.017	13,817	33,742	21,033	0.447	0.447
rs955423	A	C	A	0.335	24,203	33,131	11,266	0.594	0.406
rs9633835	G	A	G	0.018	27,183	32,076	9,363	0.63	0.37
rs9634489	G	A	G	0.02	17,131	34,414	17,072	0.5	0.5
rs9641123	C	G	C	0.025	11,139	32,861	24,595	0.402	0.402
rs968059	C	A	C	0.015	10,498	32,514	25,605	0.39	0.39
rs9856151	A	G	A	0.017	4,774	26,659	37,188	0.264	0.264
rs9925964	A	G	A	0.02	26,810	32,064	9,711	0.625	0.375

SNP*	EA	OA	GWAS EA	GWAS effect	EE	EO	OO	EAF	MAF
PSA SNPs									
rs2153904	C	A	C	-0.13	28,311	10,295	997	0.845	0.155
rs2659051	G	C	G	-0.2	2,680	21,287	44,657	0.194	0.194
rs4631830	G	A	C	0.12	7,796	19,269	12,536	0.44	0.44
Log-PSA SNPs									
rs11067228	A	G	A	0.0797	12,295	19,607	7,696	0.558	0.442
rs1354774	G	A	G	0.0583	9,261	31,462	27,895	0.364	0.364
rs3213764	G	A	G	0.0488	16,066	34,006	18,517	0.482	0.482
rs401681	G	A	C	0.0677	21,383	33,726	13,502	0.557	0.443
Prostate cancer SNPs									
rs10009409	A	G	T	0.077	6,348	28,912	33,347	0.303	0.303
rs10090154	A	G	T	0.5188	1,076	14,485	53,047	0.121	0.121
rs1016343	A	G	T	0.3148	3,678	23,949	40,988	0.228	0.228
rs10187424	A	G	A	0.0862	23,578	33,179	11,854	0.585	0.415
rs103294	G	A	C	0.2469	42,411	22,912	3,292	0.785	0.215
rs1041449	G	A	G	0.0583	7,801	19,573	12,212	0.444	0.444
rs10486567	G	A	G	0.1133	41,785	23,428	3,401	0.78	0.22
rs10503733	A	C	T	0.2546	4,519	26,144	37,963	0.256	0.256
rs10505483	A	G	T	0.5481	107	5,522	62,984	0.042	0.042
rs10774740	C	A	G	0.131	27,416	31,831	9,359	0.632	0.368
rs10875943	G	A	C	0.0677	5,792	28,087	34,697	0.289	0.289
rs10896449	G	A	G	0.0953	18,980	34,036	15,598	0.525	0.475
rs10934853	A	C	A	0.1133	5,648	27,927	35,007	0.286	0.286
rs11135910	A	G	A	0.1044	1,796	18,652	48,181	0.162	0.162
rs11214775	G	A	G	0.0677	35,000	27,924	5,679	0.714	0.286
rs11228565	A	G	A	0.207	3,622	23,962	41,024	0.227	0.227
rs11568818	A	G	A	0.0953	21,332	34,054	13,237	0.559	0.441
rs11650494	A	G	A	0.1398	500	10,575	57,550	0.084	0.084
rs11672691	G	A	G	0.077	39,081	25,296	4,221	0.754	0.246
rs11902236	A	G	A	0.0677	5,376	27,471	35,768	0.279	0.279
rs12051443	A	G	A	0.0583	0	0	67,338	0	0
rs1218582	G	A	G	0.0583	14,246	33,728	20,606	0.454	0.454
rs12597458	C	A	G	0.1044	13,030	19,352	7,221	0.573	0.427
rs12653946	A	G	T	0.2311	13,071	33,412	22,110	0.434	0.434
rs12682344	C	A	G	0.67	64	3,080	36,460	0.041	0.041
rs1270884	A	G	A	0.0677	16,792	34,213	17,550	0.494	0.494
rs130067	C	A	G	0.0488	3,015	22,347	43,218	0.207	0.207
rs13252298	A	G	A	0.1133	36,120	27,265	5,238	0.725	0.275
rs13254738	C	A	C	0.4637	4,291	17,315	17,994	0.327	0.327
rs13385191	G	A	G	0.1398	4,101	25,362	39,115	0.245	0.245
rs1447295	A	C	A	0.3577	1,110	14,624	52,884	0.123	0.123
rs16901979	A	C	A	0.5008	65	3,087	36,452	0.041	0.041
rs16902094	G	A	G	0.1906	1,581	17,970	49,069	0.154	0.154
rs17023900	G	A	G	0.2311	447	10,107	57,960	0.08	0.08
rs17599629	G	A	G	0.077	3,573	23,998	41,041	0.227	0.227
rs17694493	G	C	G	0.077	1,344	16,714	50,468	0.142	0.142
rs17765344	A	G	A	0.174	17,355	34,336	16,915	0.503	0.497
rs1859962	C	A	G	0.2311	17,434	34,264	16,807	0.505	0.495
rs188140481	A	T	A	1.0647	14	1,479	67,092	0.011	0.011
rs1894292	G	A	G	0.0953	18,539	34,297	15,779	0.52	0.48
rs1933488	A	G	A	0.1133	23,992	33,239	11,381	0.592	0.408
rs1983891	A	G	T	0.1398	5,634	27,911	35,069	0.286	0.286
rs2121875	C	A	G	0.0488	7,650	30,562	30,394	0.334	0.334
rs2273669	G	A	G	0.0677	1,554	17,506	49,157	0.151	0.151
rs2292884	G	A	G	0.131	2,452	14,894	22,256	0.25	0.25
rs2405942	A	G	A	0.131	54,578	5	14,030	0.795	0.205
rs2427345	G	A	G	0.0583	26,916	32,147	9,502	0.627	0.373

SNP*	EA	OA	GWAS EA	GWAS effect	EE	EO	OO	EAF	MAF
rs2430386	A	G	T	0.131	19,337	33,435	14,985	0.532	0.468
rs2660753	A	G	T	0.1655	916	13,452	54,258	0.111	0.111
rs3096702	A	G	A	0.0677	6,306	18,454	14,830	0.392	0.392
rs339331	A	G	T	0.1989	34,425	28,148	6,045	0.707	0.293
rs3771570	A	G	A	0.1133	1,759	18,341	48,530	0.159	0.159
rs3850699	A	G	A	0.0953	34,854	28,011	5,751	0.712	0.288
rs4242382	A	G	A	0.5068	1,098	14,594	52,895	0.122	0.122
rs4242384	C	A	C	0.6313	1,087	14,535	52,990	0.122	0.122
rs4245739	A	C	A	0.0953	37,809	26,160	4,647	0.742	0.258
rs445114	A	G	T	0.1989	29,097	31,127	8,392	0.651	0.349
rs4793529	A	G	T	0.28	16,691	34,013	17,047	0.497	0.497
rs4962416	G	A	C	0.157	5,495	27,491	35,617	0.28	0.28
rs56232506	A	G	A	0.0583	0	0	68,447	0	0
rs5759167	C	A	G	0.1655	18,673	34,399	15,527	0.523	0.477
rs5919432	A	G	A	0.0583	55,764	1	12,841	0.813	0.187
rs5945572	A	G	A	0.207	15,337	7	24,137	0.389	0.389
rs5945619	G	A	C	0.174	15,409	0	24,175	0.389	0.389
rs6062509	A	C	A	0.1133	34,147	28,424	6,046	0.705	0.295
rs636291	A	G	A	0.1655	32,240	29,705	6,671	0.686	0.314
rs6465657	G	A	C	0.1133	15,621	34,235	18,751	0.477	0.477
rs651164	G	A	G	0.1398	19,829	16,255	3,520	0.706	0.294
rs684232	G	A	G	0.0953	9,196	31,847	27,526	0.366	0.366
rs6869841	A	G	A	0.0677	3,158	23,175	42,294	0.215	0.215
rs6983267	C	A	G	0.2311	20,282	33,875	14,455	0.542	0.458
rs6983561	C	A	C	0.6259	105	5,487	62,896	0.042	0.042
rs71277158	A	C	T	0.1989	48,959	17,866	1,648	0.845	0.155
rs7130881	G	A	G	0.27	2,345	20,415	45,868	0.183	0.183
rs7141529	G	A	G	0.0862	17,702	34,148	16,752	0.507	0.493
rs7153648	C	G	C	0.1044	57,670	10,459	481	0.917	0.083
rs721048	A	G	A	0.1398	1,579	12,495	25,495	0.198	0.198
rs7241993	G	A	G	0.0862	33,928	28,703	5,979	0.704	0.296
rs742134	G	A	G	0.1484	53,810	13,837	965	0.885	0.115
rs7501939	G	A	C	0.3436	27,104	31,978	9,537	0.628	0.372
rs7584330	G	A	C	0.0583	3,974	25,007	39,640	0.24	0.24
rs7611694	A	C	A	0.0953	24,684	32,625	11,308	0.597	0.403
rs7629490	A	G	T	0.0583	4,887	18,053	16,655	0.351	0.351
rs7679673	C	A	C	0.1133	26,493	31,831	10,072	0.62	0.38
rs76934034	A	G	T	0.1222	58,309	9,814	459	0.922	0.078
rs7725218	G	A	G	0.1398	30,112	29,965	7,846	0.664	0.336
rs7758229	A	C	T	0.1398	8,405	30,383	29,763	0.344	0.344
rs7929962	A	G	T	0.1398	18,995	34,036	15,576	0.525	0.475
rs7931342	C	A	G	0.174	19,378	34,044	15,192	0.531	0.469
rs8008270	G	A	G	0.1133	46,379	20,064	2,169	0.822	0.178
rs80130819	A	C	A	0.131	57,555	10,540	526	0.916	0.084
rs8014671	G	A	G	0.0583	23,770	33,214	11,633	0.588	0.412
rs8064454	C	A	C	0.2151	21,237	33,614	13,709	0.555	0.445
rs8102476	G	A	C	0.1133	12,175	19,569	7,846	0.555	0.445
rs817826	G	A	C	0.3436	1,484	16,713	50,430	0.143	0.143
rs902774	A	G	A	0.157	1,793	18,465	48,351	0.161	0.161
rs9287719	G	A	C	0.0583	16,188	33,635	18,549	0.483	0.483
rs9364554	A	G	T	0.157	6,429	28,218	33,973	0.299	0.299
rs9600079	A	C	T	0.1655	13,618	33,977	20,992	0.446	0.446
rs9623117	G	A	C	0.1655	3,512	23,803	41,307	0.225	0.225

*SNP = single nucleotide polymorphism, EA = effect allele, OA = other allele, GWAS EA = genome-wide association study effect allele, EE = number of men with homozygous effect alleles, EO = number of men with heterozygous effect alleles, OO = number of men with no effect alleles, EAF = effect allele frequency, MAF = minor allele frequency

The albatross plot: A novel graphical tool for presenting results of diversely reported studies in a systematic review

Sean Harrison^{1,2}  | Hayley E. Jones¹ | Richard M. Martin^{1,2} | Sarah J. Lewis^{1,2} | Julian P.T. Higgins^{1,2}

¹ School of Social and Community Medicine, University of Bristol, Bristol, UK

² MRC Integrative Epidemiology Unit (IEU) at the University of Bristol, Bristol, UK

Correspondence

Sean Harrison, School of Social and Community Medicine, University of Bristol, Bristol, UK.

Email: sean.harrison@bristol.ac.uk

Funding Information

Wellcome Trust, Grant/Award Number: 102432/Z/13/Z; Medical Research Council, Grant/Award Number: MR/M014533/1; World Cancer Research Fund, Grant/Award Number: RFA 2012/620; University of Bristol, Grant/Award Number: MC_UU_12013/9; Cancer Research UK, Grant/Award Number: C18281/A19169.

Abstract

Meta-analyses combine the results of multiple studies of a common question. Approaches based on effect size estimates from each study are generally regarded as the most informative. However, these methods can only be used if comparable effect sizes can be computed from each study, and this may not be the case due to variation in how the studies were done or limitations in how their results were reported. Other methods, such as vote counting, are then used to summarize the results of these studies, but most of these methods are limited in that they do not provide any indication of the magnitude of effect.

We propose a novel plot, the albatross plot, which requires only a 1-sided *P* value and a total sample size from each study (or equivalently a 2-sided *P* value, direction of effect and total sample size). The plot allows an approximate examination of underlying effect sizes and the potential to identify sources of heterogeneity across studies. This is achieved by drawing contours showing the range of effect sizes that might lead to each *P* value for given sample sizes, under simple study designs. We provide examples of albatross plots using data from previous meta-analyses, allowing for comparison of results, and an example from when a meta-analysis was not possible.

KEYWORDS

evidence synthesis, graphical tool, methodology, systematic review

1 | INTRODUCTION

Meta-analyses combine the results of multiple studies of a common research question. They typically focus on estimation of an underlying (average) effect size across studies, and often illustrate the individual study results and the pooled result using a forest plot. Meta-analyses of this type are, however, not always possible. This is especially the case when data are collected, analysed, or reported in different

ways in different studies. Sometimes statistical analysis results are presented in ways that do not facilitate estimation of a comparable effect size for every study. One possibility in this situation is to perform a narrative synthesis of the findings across studies. This can be cumbersome to digest and there is a risk that conscious or unconscious bias may affect the way in which the results are presented.

Where statistical test results are available from each study, vote counting might be used, in which the numbers of studies

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

Copyright © 2017 The Authors. *Research Synthesis Methods* Published by John Wiley & Sons Ltd.

reporting a positive, negative or null association using a predefined P value threshold are counted. Harvest plots have been proposed as an extension of vote counting, providing a graphical tool for displaying the results from each study.¹ In a harvest plot, each study is represented by a bar whose height and appearance convey information related to confidence in the result (eg, study design), and the bars are grouped by whether the study found a positive, negative, or null association. The practice of distinguishing between “significant” and “nonsignificant” findings has lost favour in recent years.² Some specific limitations of drawing this distinction using vote counting approaches to meta-analysis are that they do not account for differences in the relative sizes of the studies and do not provide measures of the magnitude of any effects. They are also problematic when studies are small, since statistically significant results can be difficult to obtain even when effect sizes are reasonably large. The vote counting approach has been widely criticized for these reasons.^{3,4}

More attractive alternatives to vote counting are available. One possibility is a sign test, in which directions of effect are counted rather than conclusions around statistical significance. A second possibility is the statistical combination of exact precise P values across studies, for example, using methods of Fisher⁵ or Stouffer.⁶ These approaches both produce an overall P value for testing the null hypothesis of no effect in every study. However, again, they take no account of the relative sizes of the studies and do not provide an estimate of effect magnitude.

In this paper, we propose a novel plot to illustrate findings from quantitative studies when insufficient information is available to present results in a forest plot. Our albatross plot is based on minimal statistical information that is usually available from each study, namely, a precise P value and a total sample size. However, unlike the simple methods described above, albatross plots allow an approximate examination of underlying effect sizes and the potential to identify sources of heterogeneity across studies.

We introduce the albatross plot in Section 2, and explain how we illustrate approximate effect sizes using superimposed contours in Section 3. In Section 4, we illustrate the albatross plot using four example data sets: the first is a meta-analysis of randomized trials of exercise training after acute myocardial infarction (MI), which had originally been analysed using effect sizes; the second is a meta-analysis of correlations between student ratings of college professors and their achievement levels, which has been used to illustrate meta-analysis of P values; the third is from a systematic review of the association between milk intake and insulin-like growth factor-I (IGF-I), where a meta-analysis was not possible because of the diverse reporting of studies; and the fourth is from a review of the association between body mass index (BMI) and prostate specific antigen (PSA), where a meta-analysis was possible but not for all studies.

2 | THE ALBATROSS PLOT

Appropriate interpretation of P values requires information about the sizes of the study from which they come. For example, a P value of 0.2 may arise from a large study with a small underlying effect or from a small study with a large underlying effect. Our basic albatross plot is a scatter plot of study sample sizes against 1-sided P values. Equivalently, this is a scatter plot of study sample sizes against 2-sided P values, with results separated according to the observed direction of effect. The albatross plot allows the P values to be interpreted in the context of the study sample size. Small studies appear towards the bottom of the plot and large studies towards the top. Throughout this paper, we plot 2-sided versions of the P values, separated by direction of effect, because these are much more commonly reported. One-sided P values can readily be transformed to 2-sided P values.

Small 2-sided P values from strong negative results (eg, corresponding to 1-sided P values near 0) appear at the left of the plot and small 2-sided P values from strong positive results (corresponding to 1-sided P values near 1) appear at the right of the plot, with studies with null results towards the middle. We plot both the sample size axis and the P value axis on the log scale for improved visual interpretation.

Two types of enhancement to the basic albatross allow approximate examination of effect sizes and their heterogeneity. First, we superimpose contours on the plot to reflect different hypothetical effect sizes that would have given rise to particular P values. These contours will be specific to the type of data (and statistical methods) used to calculate the P values and are to be interpreted very approximately. The contours typically resemble large flying birds, giving rise to our proposed name of an albatross plot. We describe some simple derivations for contours in the following section and derive a variety of other possibilities in an online supplement. Second, different subgroups of studies can be drawn using different colours or symbols to facilitate identification of subgroup effects.

Figures 1–4 provide examples of albatross plots, and we discuss these in more detail in Section 4. It is visually clear if the studies generally convey a positive or a negative effect size, since the studies will cluster on one side of the plot. If the studies have generally a similar effect size, the points will fall around an effect size contour; if there is heterogeneity of effect size, then the points will be scattered across contours reflecting heterogeneous effect sizes. If there is no association, points will fall evenly around the null in the centre of the plot. Smaller studies will have larger variation around the true value and will have points that are more clustered around the null than larger studies. However, many small studies can still point towards a single effect size contour if the underlying effects are homogeneous. Furthermore, outlying studies can be identified with ease, and a brief narrative synthesis conducted alongside the albatross plot might propose possible explanations for the findings in these studies.

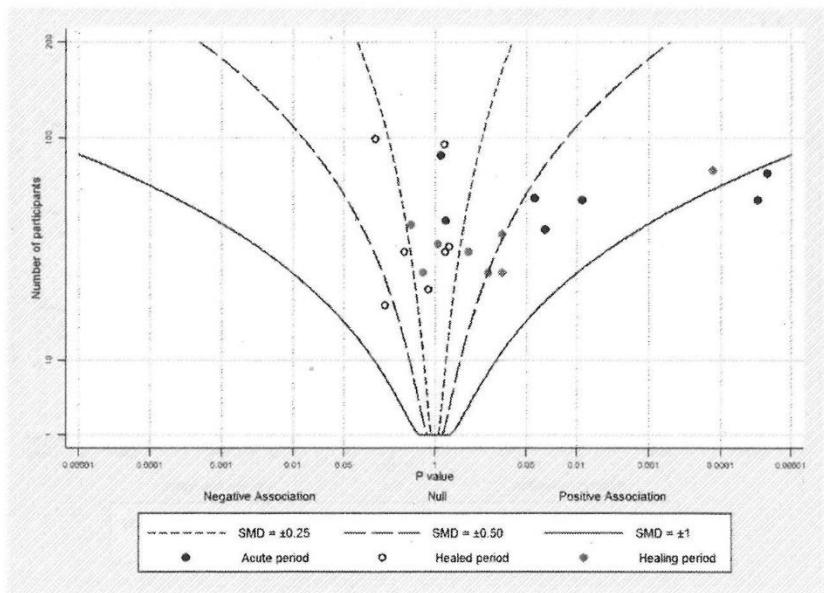


FIGURE 1 Albatross plot for studies of the effect of exercise training on left ventricular fraction after acute myocardial infarction, with contours for standardized mean differences (SMDs), using data from Zhang et al¹¹.

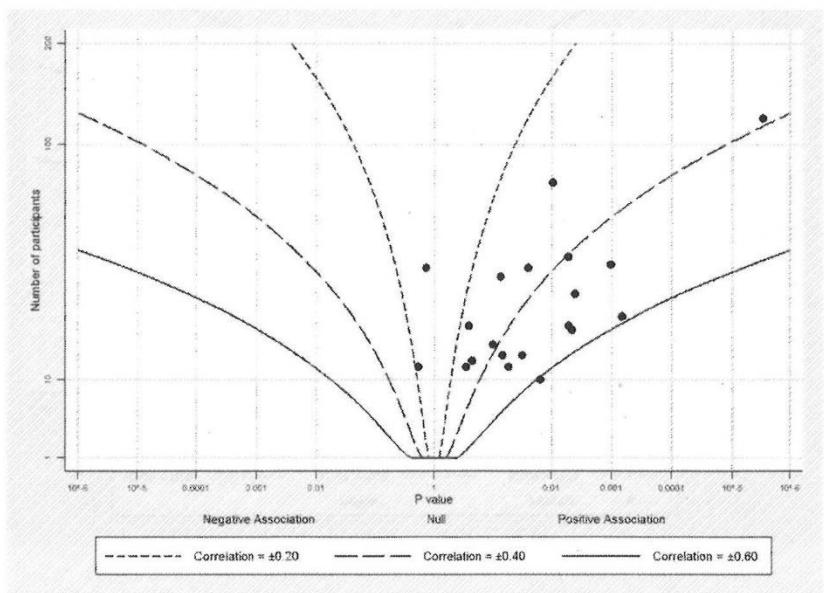


FIGURE 2 Albatross plot for studies of student ratings of their college instructors and student achievement levels with contours for correlation coefficients, using data from Becker¹⁴.

3 | EFFECT SIZE CONTOURS

Our approximate effect size contour lines are based on the general assumption that the P values were derived from Wald tests. A Wald test involves division of the effect size estimate (b) by its standard error (SE) to calculate a Z-statistic

$$Z_P = \frac{b}{SE}. \quad (1)$$

This statistic is compared to a standard normal distribution to obtain the P value. Conversely, a Z-statistic can be

obtained from a (reported) P value using the same distribution, so we write $Z_p \equiv \Phi^{-1}(P)$, where Φ^{-1} denotes the inverse of the standard normal distribution function.

In general, the SE is proportional to the inverse of the square root of the total number of participants (N) in the study, so that we can write

$$SE = \frac{\phi}{\sqrt{N}}. \quad (2)$$

The quantity ϕ may be a fixed number, or it may involve the effect size itself, and it may additionally involve other

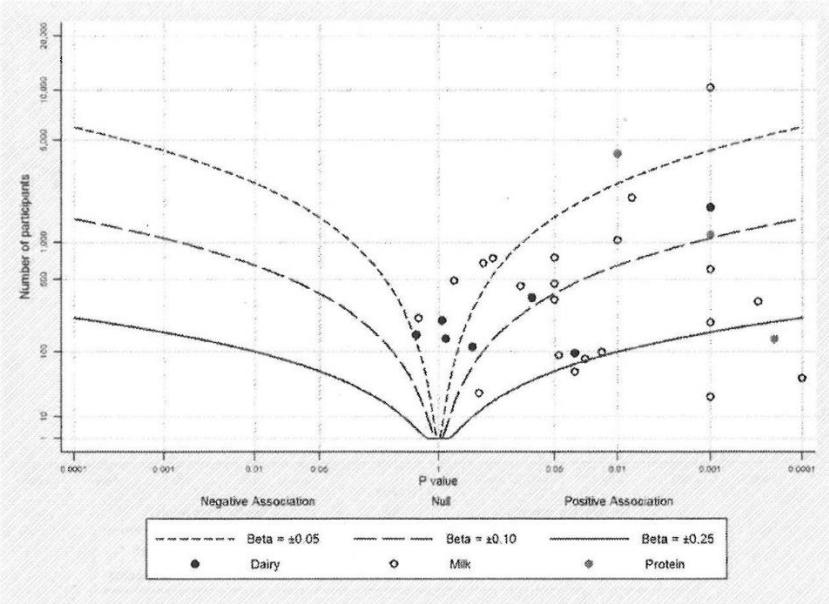


FIGURE 3 Albatross plot for studies of the association between milk intake and insulin-like growth factor-I, using data from Harrison et al (Harrison et al, In press meta-analysis, 2016)¹⁵

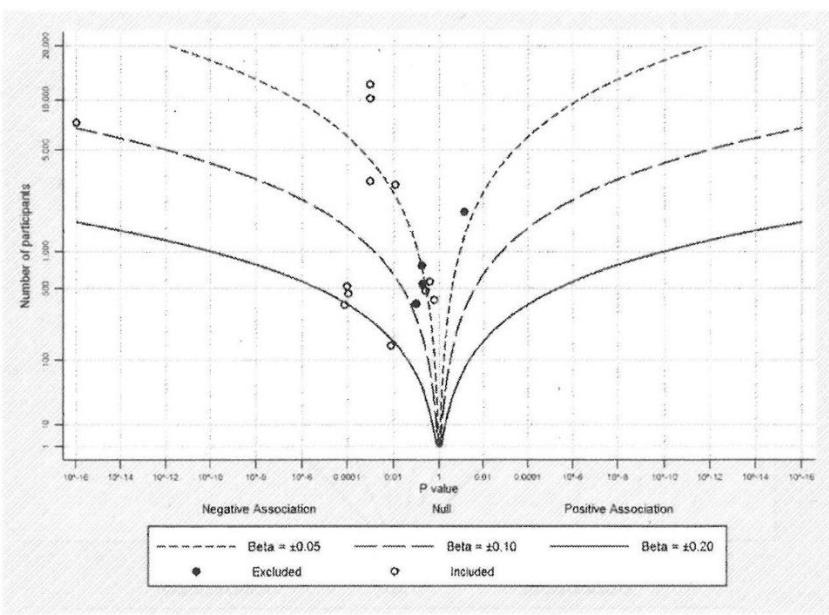


FIGURE 4 Albatross plot for the association between body mass index and prostate specific antigen, using data from Harrison et al (Harrison et al, unpublished meta-analysis, 2016)¹⁶

quantities that need to be specified to define an effect size contour uniquely.

Rearranging Equations 1 and 2, we can express the sample size in the form

$$N = \frac{\phi^2}{b^2} Z_p^2. \quad (3)$$

Since Z_p has a one-to-one correspondence with the horizontal axis in the albatross plot, to obtain a contour corresponding to a hypothetical effect size b , we need only determine the quantity ϕ appropriate for the choice of effect size, and plot N as a function of ϕ , b , and P .

In Section 3.1, we illustrate the derivation of contours for standardized mean differences (SMDs). We then use simple relationships between SMDs, odds ratios (ORs), and correlation coefficients to produce crude contours for these other two measures. In Data S1, we show how effect size contour lines can be derived more exactly for ORs and correlation coefficients, as well as for a variety of other effect measures that are commonly encountered in meta-analysis, including mean differences, risk ratios, and regression coefficients. We provide the formulae that define contours for these effect measures in Table 1. Stata code to generate an albatross plot, with a help file and examples, is available to download from the SSC using the Stata

TABLE 1 Formulae for calculating effect size contours for different effect measures

Effect measure	Equation	Additional variables requiring values
Mean difference (MD) equal sized groups	$N = \frac{4SD^2}{MD^2} Z_p^2$	Standard deviation (SD)
Mean difference (MD) unequal sized groups	$N = \frac{SD^2(r+1)^2}{r \times MD^2} Z_p^2$	Standard deviation (SD) Ratio of group sizes ($r = n_1/n_2$)
Standardized mean difference (SMD) equal sized groups	$N = \frac{8+SMD^2}{2SMD^2} Z_p^2$	(none)
Standardized mean difference (SMD) unequal sized groups	$N = \frac{2(r+1)^2+r \times SMD^2}{2r \times SMD^2} Z_p^2$	Ratio of group sizes ($r = n_1/n_2$)
Correlation coefficient (ρ)	$N = \frac{1-\rho^2}{\rho^2} Z_p^2$	(none)
Standardized beta coefficient (β_s) from univariable linear regression	$N = \frac{1-\beta_s^2}{\beta_s^2} Z_p^2$	(none)
Odds ratio (OR) equal sized groups	$N = \frac{2[(1-\pi_2+\pi_2 \times OR)^2+OR]}{\pi_2(1-\pi_2) \times OR \times (\ln OR)^2} Z_p^2$	Control group risk (π_2)
Odds ratio (OR) unequal sized groups	$N = \frac{(r+1)[(1-\pi_2+\pi_2 \times OR)^2+r \times OR]}{r\pi_2(1-\pi_2) \times OR \times (\ln OR)^2} Z_p^2$	Control group risk (π_2) Ratio of group sizes ($r = n_1/n_2$)
Risk ratio (RR) equal sized groups	$N = \frac{2(1+RR-2RR \times \pi_2)}{\pi_2 \times RR \times (\ln RR)^2} Z_p^2$	Control group risk (π_2)
Risk ratio (RR) unequal sized groups	$N = (r+1) \left(\frac{r+RR(r-\pi_2-\pi_2)}{r\pi_2 \times RR \times (\ln RR)^2} \right) Z_p^2$	Control group risk (π_2) Ratio of group sizes ($r = n_1/n_2$)

N = total number of participants ($N = n_1 + n_2$)
in two-group studies.

Z_p = Z value for the associated 2-sided P value; $\Phi^{-1}(P) \equiv Z_p$.

package name albatross (type “ssc install albatross” in Stata to download).

Contours for all effect measures require specification, by the user, of the effect size to which the contour relates. For some effect sizes, including the SMD for a study with equal group sizes, this effect size is all that needs to be specified to determine ϕ . In other cases, further variables must be specified, such as the ratio of group sizes or the baseline risk. Values for the additional variables might be chosen using the most common values in the included studies; for instance, for trials the ratio of group sizes is often 1, so there are equal number of participants in the intervention and control arms of the trial.

Some studies do not report explicit P values but state that P is more than or less than a threshold value, for example, $P < 0.05$. Where possible, a precise P value should be calculated from data. If this is not possible, for $P <$ threshold, we suggest either assuming $P =$ threshold (eg, $P = 0.05$) or drawing a line instead of a point on the albatross plot to show the range of P values compatible with the reported finding (eg, a line between $P = 0.05$ and $P = 0.01$). For $P >$ threshold, we suggest that either omitting the study or again using a line to show the range of P values compatible with the reported finding (eg, a line between 1 and 0.05).

3.1 | Effect size contours for SMDs

A simple and commonly used effect size is the SMDs, which compares mean responses between two groups, such as an intervention group and a control group in a randomized trial. We will use this effect measure as an example of how we create effect contours for the albatross plot.

The SMD is defined as

$$SMD = \frac{\mu_1 - \mu_2}{SD}, \quad (4)$$

where μ_1 and μ_2 are the mean responses in the two groups of the study and SD is the standard deviation of responses. A simple estimate (Cohen's d) of the SMD is obtained by substituting estimates of the means and the pooled SD into Equation 4. An approximate standard error for this estimated SMD is

$$SE = \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

where n_1 and n_2 are the sample sizes in the two groups, such that $N = n_1 + n_2$. If the two groups are the same size, then $SE = \frac{2}{\sqrt{N}}$. Hence ϕ is simply equal to 2, and the effect size contour for each specific value of SMD is obtained as

$$N = \frac{4}{SMD^2} Z_P^2.$$

A better approximation³ to the SE of the SMD is

$$SE = \sqrt{\frac{1}{n_1} + \frac{1}{n_2} + \frac{SMD^2}{2(n_1 + n_2)}}. \quad (5)$$

Again assuming equally sized groups, we obtain

$$\phi = \sqrt{\frac{8 + SMD^2}{2}},$$

such that the contours are defined by

$$N = \frac{8 + SMD^2}{2SMD^2} Z_P^2. \quad (6)$$

Thus, the effect contour for a given SMD, which shows the number of participants required for a particular P value, does not require any additional information to create. For example, for an SMD of 0.1, a P value of 0.05 ($Z = 1.96$) would arise from a study of 1539 participants.

The assumption of equal sample size in both groups may be a reasonable approximation for experimental studies such as randomized trials but may not be appropriate for observational studies. To account for unequal group sizes when the ratio varies across studies, it is possible to adjust the plotting position for individual points to reflect an effective sample size rather than the observed sample size. Such adjustments to the sample size may require assumptions to be made about the magnitude of the effect size. In our experience, the effect of this adjustment is typically minimal.

While it is possible to use the contour generation equations to estimate the effect size and SE of each study, this is not the purpose of the equations, nor is it advisable for all studies. Rather, the equations define hypothetical effect contours that can aid in interpretation of the magnitude of effect across many studies and are useful for when there is insufficient information to calculate effect sizes and SEs for all included studies.

3.2 | Crude effect size contours for ORs and correlation coefficients

Under particular assumptions, ORs and correlation coefficients can be transformed to and from an SMD.^{3,7–9} Thus, the contours described above for SMDs can be used to provide contours for these measures. To obtain crude contours for the OR, we can substitute the approximation

$$SMD = \frac{\sqrt{3} \times \ln OR}{\pi}$$

into Equation 6 to obtain

$$N = \frac{8\pi^2 + 3 \ln OR^2}{6 \ln OR^2} Z_P^2,$$

where $\ln OR$ is the (natural) logarithm of OR.

Similarly, using the approximation for correlation coefficients (ρ),

$$SMD = \frac{2\rho}{\sqrt{1-\rho^2}},$$

we obtain

$$N = \frac{2-\rho^2}{2\rho^2} Z_P^2. \quad (7)$$

Correlation coefficients are equivalent to standardized regression coefficients (sometimes referred to as betas) from univariable linear regression,¹⁰ so Equation 7 can be used for standardized regression coefficients, with ρ substituted by beta.

4 | APPLICATIONS

4.1 | Example 1: randomized trials of exercise training

To demonstrate the connections between an albatross plot and established techniques, we first illustrate the albatross plot using a data set that was originally analysed using effect sizes and which could, therefore, be illustrated in a forest plot. Zhang et al performed a meta-analysis of randomized trials evaluating the effect of exercise training after an acute MI.¹¹ They examined the endpoint of left ventricular function (LVEF), subgrouping the trials by the period after the MI during which exercise training was initiated. We reproduce the results in a forest plot in Figure 5. In standard random-effects meta-analyses, 7 studies initiating training in the “acute” period (6 hours–7 days) gave mean $SMD = 0.60$ (95% confidence interval 0.28 to 0.93), 8 studies initiating during the “healing” period (7–28 days) gave mean $SMD = 0.33$ (0.03 to 0.63), and 7 studies initiating during the “healed” period (29 days and beyond) gave mean $SMD = -0.10$ (-0.32 to 0.10). There is evidence of heterogeneity across studies for the first subgroup and to a lesser extent also in the second subgroup but not in the third subgroup. It is, however, evident from these results that exercise training has greater benefit, on average, the earlier it is started.

The albatross plot for these studies is presented in Figure 1, based on P values and total sample sizes from the contributing studies. The SMD effect size contours are generated using Equation 6, assuming equal numbers of

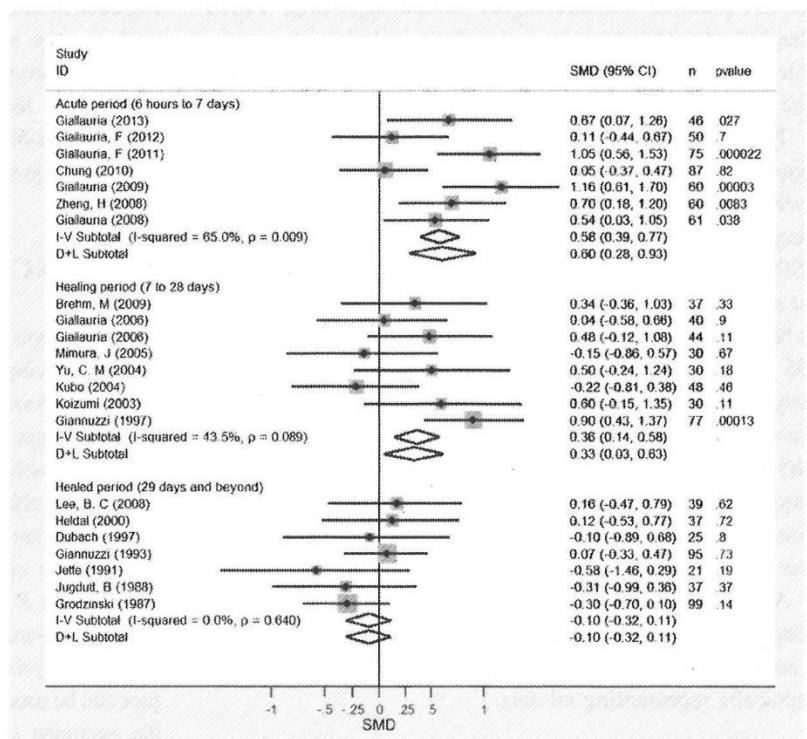


FIGURE 5 Forest plot of studies of the effect of exercise training on left ventricular fraction after acute myocardial infarction, data from Zhang et al.¹¹ I-V subtotals represent fixed effect meta-analyses; D+L subtotals represent random effect meta-analyses. I-squared is a relative measure of heterogeneity in relation to total variability within each subgroup. SMD, standardized mean difference

participants in the exercise training and control group. Different plotting symbols correspond to the three subgroups. For studies in which exercise training started in the acute period, results are clustered to the right side of the plot, showing an improvement in LVEF. The points are centred to the right of the effect size contour with magnitude 0.50, close to the value of 0.60 obtained in the random-effects meta-analysis of SMDs. However, heterogeneity among these studies is clear: although each study has a similar sample size, the *P* values are spread horizontally along the graph.

For studies in which exercise training started during the healing period, the points are mostly clustered around an SMD of 0.25. The exception¹² is separated from the other studies, corresponding to the large SMD of 0.9 observed in this study. For the third subgroup of studies in the healed period group, points are clustered around the null or a little to the left side of the graph, showing no improvement or a little detriment in LVEF, and reflecting the meta-analysis summary SMD of -0.11.

4.2 | Example 2: correlation between student ratings of their college instructors and student achievement levels

Our second example uses data from Cohen,¹³ as presented and discussed by Becker in her text about methods for combining *P* values.¹⁴ Cohen examined the correlation between student ratings of their college instructors and student achievement levels.

achievement levels. Becker implemented different methods for combining *P* values, producing combined *P* values between 1.99×10^{-4} (using a “minimum *P* value” approach) and 1.25×10^{-16} (using Stouffer’s *Z*). These tests all demonstrated that student ratings of instructors and achievement levels were correlated. Indeed, a meta-analysis of the correlation coefficients gave an average correlation of 0.36 (obtained via the weighted average of Fisher *Z*-transformed correlation coefficients), with no evidence of heterogeneity.

We provide the albatross plot for these *P* values and corresponding study sample sizes in Figure 2. Effect size contours are based on Equation 7. The majority of studies fall between the 0.2 and 0.6 correlation contours, consistent with the observed average of 0.36. Most studies have similar sample sizes, with between 10 and 100 participants. The points, however, have some noticeable scatter across different contour lines. For large sample sizes, this would reflect heterogeneity, whereas for these relatively small sample sizes, it likely reflects mainly sampling error.

4.3 | Example 3: association between milk intake and IGF-I

Our third example uses data from a review of studies investigating associations between intake of milk products and IGF-I, a protein found in blood (S. Harrison et al., In press meta-analysis, 2016)¹⁵. In total, 28 studies (with 31 data points) examined this association, but meta-analysis

was not possible because highly diverse reporting, which often lacked sufficient information for calculation of comparable effect sizes and standard errors. The exposure groups have been subdivided into milk, dairy, and dairy protein.

The albatross plot for these studies is presented in Figure 3. Effect size contours are drawn corresponding to several hypothetical standardized regression coefficients using Equation 7, where standardised beta = 1 would be a 1 SD increase in outcome for a 1 SD increase in exposure. The majority of studies show a positive effect size; most studies fall between the standardised beta coefficient contours of 0.05 and 0.25, with the average magnitude of association likely around standardised beta = 0.1. We interpreted this as a small positive association between milk and IGF-I. The dairy subgroup showed a slightly smaller magnitude of association (around 0.05 SD) than milk; the protein subgroup showed a similar association as milk, but there were only three studies in this group.

As a meta-analysis was not possible for this review, the albatross plot was useful in determining the likely magnitude of association between milk intake and IGF-I, as well as in graphically representing all data.

4.4 | Example 4: association between BMI and PSA

Our fourth example uses data from a review of studies investigating the association between BMI and PSA, a protein found in blood often used as a screening test for prostate cancer (S. Harrison et al., unpublished meta-analysis, 2016)¹⁶. In total, 10 studies were included in the meta-analysis, and 4 additional studies did not have sufficient data for inclusion. An albatross plot was created as a sensitivity analysis to determine if the excluded studies were consistent with the included studies: if the excluded studies fall around or near the included studies, then there is no reason to suspect the inclusion of these studies would materially change the outcome of the meta-analysis.

The albatross plot for these studies is presented in Figure 4, showing which studies were included in the meta-analysis and which were not. Effect size contours are drawn corresponding to several hypothetical standardized regression coefficients using Equation 7. All included studies showed a negative effect, with an average magnitude of association likely around standardised beta = -0.05, representing a small negative association between BMI and PSA. Three of the four excluded studies are consistent with the included studies; one excluded study is not as it shows a small positive association. One explanation for this may be that while this study's population was predominantly African men, all other studies had Caucasian and Asian populations.

In this example, the albatross plot was useful in determining that the excluded studies were broadly consistent with the

included studies so that, had they been included in the meta-analysis, the effect estimate would have been unlikely to change. However, the plot also identified one inconsistent study. This identification allowed us to consider possible explanations of the difference and to qualify the results, which may not be generalizable to all populations, appropriately.

5 | DISCUSSION

Albatross plots provide a versatile and simple way of graphically displaying data from multiple studies when meta-analysis is not feasible. The basic requirements of the plot are minimal, since it needs only a 1-sided *P* value and a sample size from each study (or equivalently, a 2-sided *P* value, a direction of effect, and a sample size). A single plot can show a large amount of information, including multiple results per study. Since most studies will report the number of participants and a *P* value, albatross plots can be more inclusive than a meta-analysis. Should some studies be excluded from a meta-analysis due to insufficient information, an albatross plot can be used as a sensitivity analysis to determine whether the excluded studies were consistent with the studies with sufficient information.

Heterogeneity can be seen in the spread of points across the plot, and outliers easily identified. Our proposed effect size contours give a broad indication of the magnitude of an association, while different plotting colours and symbols allow for subgroups to be compared informally. As albatross plots are not designed to estimate the magnitude of an association precisely, the effect size contours and any association seen must be interpreted as approximate; precise combined effect estimates can only be obtained through meta-analysis. A narrative synthesis may be needed to supplement the plot to fully explore the data, especially for outlying studies.

ACKNOWLEDGEMENTS

SH is funded by the Wellcome Trust (grant 102432/Z/13/Z). The work was also funded in part by the World Cancer Research Fund (grant RFA 2012/620) and Cancer Research UK Programme Grant C18281/A19169 (Integrative Cancer Epidemiology Programme). HEJ is funded by a Medical Research Council (MRC) career development award in biostatistics (MR/M014533/1). SH, RMM, SL, and JPTH are members of the MRC Integrative Epidemiology Unit at the University of Bristol, which is supported by the Medical Research Council and the University of Bristol (grant MC_UU_12013/9).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Ogilvie D, Fayter D, Petticrew M, et al. The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Med Res Methodol.* 2008;8
2. Sterne JA, Davey Smith G. Sifting the evidence-what's wrong with significance tests? *BMJ.* 2001;322:226-231.
3. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. *Introduction to Meta-Analysis.* Wiley; 2008.
4. Friedman L. Why vote-count reviews don't count. *Biol Psychiatry.* 2001;49:161-162.
5. Fisher RA. *Statistical Methods for Research Workers.* Edinburgh: Oliver and Boyd; 1925.
6. Stouffer SA, Suchman EA, Devinney LC, Star SA, Williams RMJ. *The American Soldier. I: Adjustment during Army Life* Princeton University Press: Princeton; 1949.
7. Anzures-Cabrera J, Sarpatwari A, Higgins JP. Expressing findings from meta-analyses of continuous outcomes in terms of risks. *Stat Med.* 2011;30:2967-2985.
8. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med.* 2000;19:3127-3131.
9. Hasselblad V, Hedges LV. Meta-analysis of screening and diagnostic tests. *Psychol Bull.* 1995;117:167-178.
10. Kirkwood BR, Sterne JAC. *Essential Medical Statistics.* Blackwell Publishing; 2003.
11. Zhang YM, Lu Y, Tang Y, et al. The effects of different initiation time of exercise training on left ventricular remodeling and cardiopulmonary rehabilitation in patients with left ventricular dysfunction after myocardial infarction. *Disabil Rehabil.* 2015;1-9.
12. Giannuzzi P, Temporelli PL, Corra U, Gattone M, Giordano A, Tavazzi L. Attenuation of unfavorable remodeling by exercise training in postinfarction patients with left ventricular dysfunction: results of the exercise in left ventricular dysfunction (ELVD) trial. *Circulation.* 1997;96:1790-1797.
13. Cohen PA. A selective review of the validity of student-ratings of teaching - comment. *Journal of Higher Education.* 1983;54:448-458.
14. Cooper H, Hedges LV. *The Handbook of Research Synthesis.* New York: The Russel Sage Foundation; 1994.
15. Harrison S, Lennon R, Holly J, et al. Does milk intake promote prostate cancer initiation or progression via effects on insulin-like growth factors (IGFs)? A systematic review and meta-analysis. *Cancer Causes & Control.* 2017;28:497. <https://doi.org/10.1007/s10552-017-0883-1>
16. Harrison S, Jones HE, Turner E, Tilling K. Associations between body-mass index, prostate cancer and prostate-specific antigen. Unpublished

How to cite this article: Harrison S, Jones HE, Martin RM, Lewis SJ, Higgins JPT. The albatross plot: A novel graphical tool for presenting results of diversely reported studies in a systematic review. *Res Syn Meth.* 2017;8:281–289. <https://doi.org/10.1002/jrsm.1239>

Investigating the prostate specific antigen, body mass index and age relationship: is an age–BMI-adjusted PSA model clinically useful?

Sean Harrison¹ · Kate Tilling¹ · Emma L. Turner¹ · J. Athene Lane¹ · Andrew Simpkin¹ · Michael Davis¹ · Jenny Donovan¹ · Freddie C. Hamdy² · David E. Neal³ · Richard M. Martin¹

Received: 19 January 2016 / Accepted: 26 October 2016 / Published online: 9 November 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Purpose Previous studies indicate a possible inverse relationship between prostate-specific antigen (PSA) and body mass index (BMI), and a positive relationship between PSA and age. We investigated the associations between age, BMI, PSA, and screen-detected prostate cancer to determine whether an age–BMI-adjusted PSA model would be clinically useful for detecting prostate cancer.

Methods Cross-sectional analysis nested within the UK ProtecT trial of treatments for localized cancer.

Of 18,238 men aged 50–69 years, 9,457 men without screen-detected prostate cancer (controls) and 1,836 men with prostate cancer (cases) met inclusion criteria: no history of prostate cancer or diabetes; PSA < 10 ng/ml; BMI between 15 and 50 kg/m². Multivariable linear regression models were used to investigate the relationship between log-PSA, age, and BMI in all men, controlling for prostate cancer status.

Results In the 11,293 included men, the median PSA was 1.2 ng/ml (IQR: 0.7–2.6); mean age 61.7 years (SD 4.9); and mean BMI 26.8 kg/m² (SD 3.7). There were a 5.1% decrease in PSA per 5 kg/m² increase in BMI (95% CI 3.4–6.8) and a 13.6% increase in PSA per 5-year increase in age (95% CI 12.0–15.1). Interaction tests showed no evidence for different associations between age, BMI, and

PSA in men above and below 3.0 ng/ml (all p for interaction >0.2). The age–BMI-adjusted PSA model performed as well as an age-adjusted model based on National Institute for Health and Care Excellence (NICE) guidelines at detecting prostate cancer.

Conclusions Age and BMI were associated with small changes in PSA. An age–BMI-adjusted PSA model is no more clinically useful for detecting prostate cancer than current NICE guidelines. Future studies looking at the effect of different variables on PSA, independent of their effect on prostate cancer, may improve the discrimination of PSA for prostate cancer.

Keywords Prostate cancer · PSA · BMI · Age · Prostate cancer screening · PSA–BMI equation

Background

Prostate cancer is the second most common cancer in men worldwide, with 1.1 million new cases diagnosed in 2012 [1]. Although prostate cancer deaths are considerably fewer in number than incident cancers (307,000 deaths worldwide in 2012 [1]), prostate cancer is the fifth leading cause of death from cancer in men. Developed countries tend to have a higher incidence of prostate cancer than others, in part due to increasing and widespread testing for serum prostate-specific antigen (PSA) [1].

There is considerable controversy over the effectiveness of PSA screening for prostate cancer [2–6]. One issue affecting the accuracy of the PSA test is that it is influenced by many variables other than the presence of cancer, for example diet [7], ethnicity [8], genetic variation [9, 10], certain drugs [11–14], and non-malignant disease [15–17]. Several studies report an inverse relationship between PSA

✉ Sean Harrison
sean.harrison@bristol.ac.uk

¹ School of Social and Community Medicine, University of Bristol, Canyne Hall, 39 Whatley Road, Bristol BS8 2PS, UK

² Nuffield Department of Surgery, John Radcliffe Hospital, University of Oxford, Oxford OX3 9DU, UK

³ University Department of Oncology, Addenbrooke's Hospital, Hills Road, Cambridge CB2 0QQ, UK

and body mass index (BMI) [18–23], which may explain the observed inverse relationship between BMI and incident prostate cancer [24]. However, studies suggest the observed inverse relationship is a result of confounders and is eliminated in multivariable models (adjusted for age, current statin, aspirin and other NSAID use, diabetes, and benign prostatic hyperplasia) [25]. Most studies also report an increase in PSA with age [18].

The explanation for the positive relationship between age and PSA is well understood; with age, the prostate enlarges and contains more PSA-producing tissue. Older prostates tend to leak more PSA because the normal physiologic barriers breakdown, which allows PSA to escape into capillaries leading to a slight increase in serum PSA concentration [26]. Here, the degeneration of prostatic cells is independent of prostate cancer; although degeneration may increase cancer risk over time, the increase in PSA caused by the degeneration is therefore not directly caused by cancer.

The explanation for the inverse relationship between BMI and PSA levels is more uncertain; one suggestion is that obesity causes hemodilution due to an increased plasma volume [22, 27, 28]; another is that reduced androgen levels and increased estrogen in overweight men cause lower circulating PSA levels [23]. The reduction in PSA caused by an increased BMI may lead to men not receiving a biopsy when a smaller man would, which may help explain the observed paradoxical inverse relationship of BMI with prostate cancer detection, but the positive relationship between BMI and increased prostate cancer mortality [24]. However, it is difficult to establish whether obesity affects prostate cancer directly or whether its effect on PSA means obese men are diagnosed later with an associated worse prognosis.

In current UK practice, PSA value thresholds are used when screening for prostate cancer to indicate further investigation by prostate biopsy. National Institute for Health and Care Excellence (NICE) guidelines advise using age-specific cutoff PSA measurements: for men aged 50–59 years ≥ 3.0 ng/ml; 60–69 years ≥ 4.0 ng/ml; 70 years and older ≥ 5.0 ng/ml [29]. However, BMI is not taken into account when considering whether to send a man for biopsy.

Prostate cancer risk calculators (ERSPC risk calculator 6 [30], the Prostate Cancer Prevention Trial (PCPT) risk calculator [31], and PSA-AV developed by Patel et al. [32]) take age into account when considering prostate cancer risk and are available in the literature and on the Internet. The PCPT risk calculator can also include BMI category [33]; the BMI-adjusted PSA for a man is calculated by multiplying his PSA by the ratio of the geometric mean of PSA for BMI <25 to the geometric mean of PSA for his BMI category.

The majority of men undergoing PSA tests in the UK are likely to be overweight or obese [34]. If the inverse PSA-BMI relationship is considerable, having a PSA threshold which decreases with increasing BMI (adjusting the PSA for BMI) may improve the accuracy of the test for detecting prostate cancer. The aim of this study was to examine the relationship between PSA and BMI in a large population-based study of men undergoing PSA tests, to derive a model to adjust the observed PSA for the relationship between BMI, age, and PSA and investigate whether an age–BMI-adjusted model for PSA would be clinically useful for detecting prostate cancer.

Materials and methods

We conducted a cross-sectional study nested within ProtecT (Prostate testing for cancer and Treatment), a population-based randomized controlled trial which compares treatments for clinically localized prostate cancer. Study details are published elsewhere [35]. In brief, 225,000 men aged 50–69 in 9 centers across the UK were invited for PSA testing. Of the 111,000 men who attended a PSA test, 10,000 had a PSA ≥ 3.0 ng/ml: Those with a PSA <20 ng/ml were invited for a 10-core transrectal ultrasound-guided biopsy, a repeat PSA test, and a digital rectal exam, while men with a PSA over 20 ng/ml were referred to usual care. Those with clinically localized PCa were invited into the ProtecT randomized treatment trial, comparing radical surgery, radical conformal radiotherapy and active monitoring [35].

We selected all 3,096 men diagnosed with prostate cancer and a random sample of 18,231 men without prostate cancer and with full information on covariates. The random sample was generated prior to this study by matching cases with 6 men without prostate cancer from within the same 5-year age band and GP practice. All potential matches were ordered by computer-generated random numbers, and the first 6 controls were chosen as a match. Men without prostate cancer were defined as having received a PSA test with no subsequent histological confirmation of prostate cancer, either because they were not indicated for biopsy (PSA < 3.0 ng/ml) or because a 10-core biopsy was negative (PSA ≥ 3.0 ng/ml). All men provided written informed consent prior to inclusion in the study. Trent Multicentre Research Ethics Committee (MREC) approved the ProtecT study (MREC/01/4/025) and the associated Prostate Mechanisms of Progression and Treatment (ProMPT) study which collected height data (MREC/01/4/061) as part of a diet, health and lifestyle questionnaire. All data were anonymized prior to analysis.

Inclusion and exclusion criteria

For men without prostate cancer, the inclusion criteria for this analysis were: age between 50 and 69 years with no previous history of prostate cancer or diabetes, a BMI between 15 and 50 kg/m², and PSA ≤ 10.0 ng/ml. Of the 18,231 men without prostate cancer, 9,457 (52%) satisfied the inclusion criteria. Most excluded men lacked data on height ($n = 6,916$, 38%) to compute BMI, as height data were collected as part of a separate questionnaire not filled in by all ProtecT participants and thus were likely missing at random.

Men with diabetes and men with no information on diabetes status ($n = 1,572$, 8.6%) were excluded, because diabetes influences PSA levels, prostate cancer risk and is associated with BMI [20, 36]. Men with a PSA above 10.0 ng/ml were excluded ($n = 33$, 0.2%), as high PSA levels can be associated with increased risk of false negatives at prostate biopsy [37]. A value of 10.0 ng/ml was chosen as the threshold as it compromises between excluding those most at risk of having a false negative at prostate biopsy and keeping as many men as possible in the analysis.

For men with prostate cancer, the inclusion criteria for this analysis were: age between 50 and 69 years, no previous history of diabetes, and a BMI between 15 and 50 kg/m². Of the 3,096 men with prostate cancer, 1,830 (59%) satisfied the inclusion criteria; most excluded men lacked height data to compute BMI ($n = 931$, 30%), and 280 men with diabetes were excluded (9%).

For all men without height data ($n = 7,847$), the average age and weight were 61.1 years and 86.4 kg, slight differences to men with height data ($n = 13,412$), 62.0 years and 85.0 kg (both $p < 0.01$). Men with prostate cancer were more likely to have height data, 70 versus 62% of men without prostate cancer. The mean PSA for men without prostate cancer but with height data was 1.39 ng/ml, slightly more than for men without prostate cancer and height data, 1.28 ng/ml ($p < 0.01$). Conversely, the mean PSA for men with prostate cancer and height data was 9.47 ng/ml, slightly less than for men with prostate cancer but without height data, 10.78 ng/ml ($p = 0.21$).

For men without prostate cancer only, men with diabetes ($n = 1,029$) had an average age of 63.2 years, BMI of 29.9 kg/m², and PSA of 1.23 ng/ml, whereas men without diabetes ($n = 11,772$) had a lower average age (61.8 years), BMI (27.2 kg/m²), and higher PSA (1.39 ng/ml). Men with missing diabetes status ($n = 8,526$) had an average age of 61.3 years, BMI of 27.6 kg/m², and PSA of 1.30 ng/ml.

Statistical analysis

The BMI of each man was calculated by dividing their weight (in kg) by their height squared (in meters squared).

94% of the men were weighed at clinic by a nurse, but 6% of men only had self-reported weights in stones and pounds as part of a diet, health and lifestyle questionnaire. Height was self-reported in feet and inches. Self-reported weight and height measurements were converted from imperial to metric units when calculating BMI.

The data from the 9,457 men without prostate cancer and 1830 men with prostate cancer were used to derive a model associating age and BMI with PSA. A multiplicative model for PSA was assumed, where the relationships between PSA and age and BMI were dependent on an initial level of PSA; a change in age or BMI leads to a proportional change in PSA.

A multiplicative model is intuitively more appropriate than an additive model, as a man with a high PSA would be expected to have a larger change in PSA than a man with a low PSA for the same change in age or BMI. Multivariable linear regression of the natural logarithm of PSA against age, BMI, and case-control status, separately and together in univariable and multivariable models, was used to estimate the coefficients for the model. Case-control status was included as a covariate to account for any associations between age and BMI with prostate cancer, which would otherwise bias the results.

The model was used to derive “adjusted” PSAs, removing the effects of age and BMI on PSA separately and together. The adjustment changes the man’s observed PSA by an amount depending on the man’s observed PSA and the difference between the man’s age or BMI and the mean age and BMI in this study. The larger the difference between the man’s age or BMI from the study mean, and the larger the observed PSA, the more the observed PSA is altered. The adjusted PSA can be interpreted as what the man’s PSA would have been, if they had been of average age and BMI.

The equation to adjust PSA for age and BMI is shown here:

$$\text{Age/BMI adjusted PSA} = \frac{\text{PSA}}{e^{(a \times \text{age}_{\text{coef}} + b \times \text{BMI}_{\text{coef}})}}$$

where age-BMI-adjusted PSA is PSA adjusted for age and BMI, PSA is a man’s observed prostate-specific antigen in ng/ml, a is the difference between the man’s age and the population mean age in years, b is the difference between the man’s BMI and the population mean BMI in kg/m², age_{coef} is the coefficient of age from our linear regression model, BMI_{coef} is the coefficient of BMI from our linear regression model, and e is the exponential function. This model assumes that the relationship between PSA and BMI and age is the same for men with and without prostate cancer, in this population.

The age-BMI-adjusted PSA was used to determine whether the adjustment of PSA for BMI and age was

clinically useful for detecting prostate cancer. Sensitivity and specificity estimates were calculated for the use of PSA to detect prostate cancer at biopsy (see Box 1). These used thresholds of 3.0 and 4.0 ng/ml (commonly used thresholds in clinical practice in the UK) for observed PSA values, age-adjusted PSA, BMI-adjusted PSA, age- and BMI-adjusted PSA, and were also compared to the sensitivity and specificity of the UK NICE guideline thresholds for PSA testing [29].

Sensitivity and specificity

The sensitivity was calculated as the number of men with diagnosed prostate cancer who had a PSA above the threshold level divided by the total number of men with diagnosed prostate cancer (PSA-positive cases/total cases). The specificity was calculated as the number of men without diagnosed prostate cancer who had a PSA below the threshold level divided by the total number of men without prostate cancer (PSA-negative controls/total controls).

Men with a PSA below 3 ng/ml may have undiagnosed prostate cancer, as they were not biopsied. In the Prostate Cancer Prevention Trial (PCPT), 17% ($n = 759$) of men with a PSA below 3 ng/ml had prostate cancer on biopsy [31]. As the ages and BMIs of men in the PCPT were different from ProtecT, it is difficult to estimate the number of men in ProtecT who had undiagnosed prostate cancer.

Because some men will not have been biopsied but will have undiagnosed prostate cancer, the calculated sensitivities in this study will be higher than the true sensitivities. Additionally, even men who received a biopsy may also have had undiagnosed prostate cancer due to the sensitivity of 10-core biopsy; Haas demonstrated that a 12-core biopsy of cadavers showed a sensitivity for all prostate cancers of between 36 and 53%, depending on sampling location within the prostate [38]. The sensitivity of 12-core biopsy rises to 80 and 85% for “clinically significant” and large ($\geq 0.5 \text{ cm}^3$) cancers, respectively. The specificity of biopsy was 99%, indicating there should be few incorrect diagnoses of prostate cancer.

Therefore, the total number of men with prostate cancer is underestimated in this study, both by men not receiving a biopsy and by the biopsy not detecting all cancers, so the sensitivity of each PSA test will be overestimated (see Box 2). The total number of men without prostate cancer is overestimated in this study by the same amount, but this is unlikely to affect the specificity as there are far more men without than with prostate cancer. However, the main analysis will focus on the 4.0 ng/ml threshold; as the

biopsy threshold in ProtecT was 3 ng/ml almost all men who might change over a 4.0 ng/ml threshold will have been biopsied. Assuming there is no strong relationship between PSA and missing prostate cancer at biopsy, this means any change in sensitivity or specificity seen in the PSA models is unlikely to be affected by men with undiagnosed prostate cancer. Therefore, all PSA models in this study can be directly compared, even though they do not represent the true sensitivity of PSA as a test for prostate cancer.

NICE guidelines use different thresholds for different age groups, making it difficult to directly compare the sensitivity and specificity to the other models. To show clinical utility, our PSA models would require both a higher sensitivity and specificity; otherwise, there would be a trade of specificity for sensitivity, or vice versa. Therefore, the sensitivity of all models was compared at the same specificity seen when using the NICE guidelines; any model with a higher sensitivity would necessarily be more clinically useful. McNemar's test was used to determine whether any model was preferred over the NICE guidelines [39] when the specificities of all models were equal.

As the sensitivities and specificities of the models are likely to be inaccurate, ROC curves (receiver operating characteristics) [40] and area under the curves (AUCs) were not generated.

Tenfold cross-validation [41] was used to determine whether the sensitivities and specificities of the adjusted PSAs were consistently better or worse than the NICE guideline thresholds for PSA testing. In tenfold validation, the dataset is split into 10 equal parts and each part of the dataset is considered the “validation” dataset, with the other 9 parts used to calculate the model that will be validated. This is repeated 10 times, until each part of the dataset has acted as the “validation” dataset. The sensitivities and specificities of each model were averaged across the 10 validations, with the mean and standard deviation recorded. These were then compared across models to give a robust indication of the performance of each model. The advantage of using tenfold cross-validation as opposed to split-cohort validation is that the training data can be as large as possible without compromising the robustness of the model performance in the testing data.

An interaction test [42] was performed to determine whether there was a difference in the associations between age, BMI, and PSA for men with PSA values above and below 3 ng/ml; this was to test whether the undiagnosed prostate cancers from men not being biopsied were causing any bias in the results. Multivariable regressions were performed as above, restricted to men with a PSA above

and below 3.0 ng/ml separately. For both age and BMI, the difference in coefficients was divided by the combined standard error to give a Z-score, which was converted to a *p* value.

Gleason score was used as an additional outcome in men with diagnosed prostate cancer, and ordered logistic regression was used to determine whether age and BMI were associated with Gleason score in this population. In order to examine the sensitivity of conclusions to the relative proportions of cases and controls, we re-ran all primary analyses using only controls to derive the age- and BMI-adjusted PSA, and then examined the sensitivity and specificity of this model. A further sensitivity analysis used multiple imputations by the MICE system of chained equations to estimate missing BMI and diabetes data from weight, height, age, case-control status, diabetes status, and log-PSA to determine whether missing height/diabetes would likely have caused bias.

All analyses were performed using Stata 13.1 (Stata-Corp, TX).

Results

Summary demographics are presented in Table 1. The mean age for all men ($n = 11,293$), cases ($n = 9,457$), and controls ($n = 1,836$) was 61.7 (SD 4.9), 61.6 (SD 4.9), and 61.8 (SD 4.9) years, respectively. The mean BMI was 27.2 (SD 3.7), 27.2 (SD 3.8), and 27.1 (SD 3.6) kg/m², respectively ($p = 0.21$). The median PSA was 5.0 ng/ml (IQR 3.7–8.0) in cases and 1.0 ng/ml (IQR 0.6–1.7) in controls ($p < 0.0001$). Ordered logistic regression of 1,830 men with a Gleason score with age and BMI showed both were weakly associated with Gleason score: age: coef = 0.03, *p* = 0.002; BMI: coef = 0.03, *p* = 0.04.

In univariable models ($n = 11,293$), where log-PSA was regressed separately against age and BMI (with case-control status as a covariate), PSA increased by 13.55% (95% CI 12.01–15.11) per 5-year increase in age and decreased by 5.58% (95% CI 3.83–7.29) per 5 kg/m² increase in BMI. When case-control status was omitted from the regression, PSA increased by 14.30% (95% CI

Box 1 Definitions of sensitivity and specificity

Sensitivity—the true positive rate; the number of people *with* prostate cancer who had a PSA *above* the threshold level divided by the total number of people *with* diagnosed prostate cancer

Specificity—the true negative rate; the number of people *without* prostate cancer who had a PSA *below* the threshold level divided by the total number of people *without* prostate cancer

Box 2 Effect of undiagnosed prostate cancer on sensitivity and specificity of PSA testing for prostate cancer. Bold letters are the true number of men in each cell, and italic letters are the study number of men in each cell

	Prostate cancer	No prostate cancer
PSA \geq 3.0 ng/ml	A	B
	<i>A</i> – <i>x</i>	<i>B</i> + <i>x</i>
PSA < 3.0 ng/ml	C	D
	<i>0</i>	<i>D</i> + <i>C</i>
Number of men with (left) and without (right) prostate cancer	A + C	B + D
	<i>A</i> – <i>x</i>	<i>B</i> + <i>x</i> + <i>D</i> + <i>C</i>
A and C are the number of men truly with prostate cancer, B and D are the number of men truly without prostate cancer, and x is the number of men with prostate cancer but the biopsy missed the cancer.		
Men with a PSA above 3.0 ng/ml were biopsied, so any underestimation in A and overestimation in B is from biopsies that miss the cancer [38]. If there are <i>x</i> men with missed cancers, then A becomes <i>A</i> – <i>x</i> and B becomes <i>B</i> + <i>x</i> .		
Men were not biopsied if their PSA was less than 3.0 ng/ml, so C becomes 0. Some men will have prostate cancer and a PSA less than 3.0 ng/ml [31], so C is underestimated and D becomes <i>D</i> + <i>C</i> .		
As C is 0, all calculated sensitivities [<i>A</i> / <i>(A</i> + <i>C)</i>] with a PSA threshold of 3.0 ng/ml or less will be 1. In this study, the sensitivity will always be overestimated, as the total number of men with prostate cancer (<i>A</i> + <i>C</i>) will always be underestimated.		
The specificity [<i>D</i> / <i>(B</i> + <i>D)</i>] may be over- or underestimated as both B and D are overestimated by different amounts, but overall the calculated specificity should not differ too much from the true specificity.		
Example: 2,500 men truly had prostate cancer (<i>A</i> + <i>C</i>), but 100 were missed at biopsy (<i>x</i>) and 500 had a PSA less than 3.0 ng/ml (<i>C</i>) so were not biopsied, and 7,500 men truly did not have prostate cancer (<i>B</i> + <i>D</i>); 1,000 of these men had a PSA greater than 3.0 ng/ml (<i>B</i>).		
The calculated sensitivity at 3.0 ng/ml would be (<i>Ax</i>) / (<i>A</i> – <i>x</i>) = (2,000 – 100) / (2,000 – 100) = 1, but the true sensitivity would be <i>A</i> / (<i>A</i> + <i>C</i>) = 2000 / 2500 = 0.8; the sensitivity is overestimated.		
The calculated specificity at 3.0 ng/ml would be (<i>D</i> + <i>C</i>) / (<i>B</i> + <i>x</i> + <i>D</i> + <i>C</i>) = (6,500 + 500) / (1,000 + 100 + 6,500 + 500) = 0.86, and the true specificity would be <i>D</i> / (<i>B</i> + <i>D</i>) = 6,500 / 7,500 = 0.87; the specificity is very close to the true value.		

Table 1 Summary demographics of included participants; cases have a diagnosis of prostate cancer

	All	Cases	Controls	<i>p</i> value for difference
<i>n</i>	11,293	1,836	9,457	NA
Age (SD)	61.7 (4.94)	61.8 (4.90)	61.6 (4.95)	0.29
BMI (SD)	27.2 (3.73)	27.1 (3.57)	27.2 (3.76)	0.21
PSA (IQR)	1.2 (0.7–2.6)	5 (3.7–8.0)	1 (0.6–1.7)	<0.0001
BMI categories in kg/m ² [<i>n</i> (%)]				
<25	3,274 (29)	522 (28.4)	2,752 (29.1)	<i>p</i> for trend = 0.65
25–29.9	5,805 (51.4)	979 (53.3)	4,826 (51)	
>30	2,214 (19.6)	335 (18.3)	1,879 (19.9)	

Table 2 Results of linear regression of age and BMI against PSA separately (univariable) and together (multivariable)

Variable	<i>n</i>	Change in PSA (%)	Log(PSA) change per 5 unit increase in covariate		
			Coefficient	<i>p</i> value	95% CI
Univariable					
Age	11,293	13.55	0.127	<0.00001	0.113 to 0.141
BMI	11,293	-5.58	-0.057	<0.00001	-0.076 to -0.039
Multivariable					
Age	11,293	13.43	0.126	<0.00001	0.112 to 0.140
BMI	11,293	-5.14	-0.053	<0.00001	-0.071 to -0.035

The beta coefficients for change in log-PSA are equivalent to the logarithm of the multiplicative change in PSA per 5 unit increase in age–BMI, which has been expressed as a percentage change in the table. Small change in log-PSA is broadly interpretable as the percentage changes in PSA

PSA prostate-specific antigen, BMI body mass index, SE standard error

Table 3 Sensitivities (proportion of men diagnosed with prostate cancer (cases) above the threshold PSA) and specificities (proportion of men not diagnosed with prostate cancer (controls) below the threshold PSA) for prostate cancer detection at biopsy for different models

Model	Threshold 3.0 ng/ml		Threshold 4.0 ng/ml	
	Proportion of cases above threshold	Proportion of controls below threshold	Proportion of cases above threshold	Proportion of controls below threshold
Observed	1	0.931	0.687	0.964
Age–PSA	0.947	0.932	0.699	0.968
BMI–PSA	0.978	0.932	0.682	0.965
Age–BMI–PSA	0.942	0.932	0.708	0.967
NICE guidelines ^a	—	—	0.797	0.958

PSA prostate-specific antigen, BMI body mass index

^a NICE guidelines—ages 50–59: 3.0 ng/ml, ages 60–69: 4.0 ng/ml

12.22–16.41) per 5-year increase in age and decreased by 6.55% (95% CI 4.24–8.81) per 5 kg/m² increase in BMI.

In the multivariable model (*n* = 11,293), where log-PSA was regressed against BMI, adjusting for age and case–control status together, there was a 5.14% decrease in PSA per 5 kg/m² increase in BMI (95% CI 3.41–6.84, *p* < 0.001) (Table 2). The BMI-adjusted results for age are not presented, as there is no plausible mechanism by which BMI can confound the age–PSA association.

The sensitivities and specificities (as defined by the proportion of cases above the threshold PSA and the proportion of controls below the threshold PSA, respectively), of the different PSA models at 3.0 and 4.0 ng/ml thresholds, and the NICE guidelines (threshold dependent on age) are presented in Table 3. All sensitivities and specificities were calculated assuming no undiagnosed prostate cancers among men in the control group or misdiagnosis by biopsy and are thus inaccurate (the sensitivities are overestimates,

and the specificities may be over- or underestimates, as explained in Box 2).

Compared to PSA alone, sensitivity was improved when adjusting PSA for BMI [by about 0.01 (1%)], but specificity worsened at both 3.0 and 4.0 ng/ml thresholds [by 0.003 (0.3%) and 0.001 (0.1%), respectively]. In comparison, adjusting PSA for age improved both the sensitivity and specificity at 4.0 ng/ml [by 0.026 (2.6%) and 0.002 (0.2%), respectively], but worsened both sensitivity and specificity at 3.0 ng/ml [by 0.021 (2.1%) and 0.003 (0.3%), respectively]. Adjusting PSA for both age and BMI showed similar results to adjusting solely for age, with the exception of improved sensitivity at 4.0 ng/ml, implying that adjusting for BMI as well as age did not materially improve the discrimination of PSA for prostate cancer.

When the specificity of the model PSAs was set to the same as the NICE guidelines (0.958), the sensitivities were all below that of the NICE guidelines: NICE: 0.797; PSA alone: 0.768; age-adjusted PSA: 0.794; BMI-adjusted PSA: 0.758; and age–BMI-adjusted PSA: 0.796. However, when each model was compared with the NICE guidelines using McNemar's test, there was no evidence to say any model was better or worse at detecting prostate cancer: PSA alone p value 0.65; age-adjusted PSA p value: 1; BMI-adjusted PSA p value: 0.53; age–BMI-adjusted PSA p value: 1.

The tenfold cross-validation showed that the averaged sensitivities and specificities were very close to the sensitivities and specificities of the results from the main analysis. The standard deviations were low, indicating the models performed robustly across the ten validation sets (Table 4). Therefore, it is unlikely that the results from the main analysis are the product of validating the models in the same dataset used to develop the models.

The interaction test examining whether men with a PSA at or above 3.0 ng/ml ($n = 2,489$) had different estimates of the associations between age, BMI, and PSA to men with a PSA less than 3.0 ng/ml ($n = 8,804$) showed no

evidence of an effect of not receiving a biopsy; the p value for the interactions with age and BMI, respectively, was 0.22 and 0.24.

In the sensitivity analysis using only men without prostate cancer ($n = 9,457$), multivariable linear regression showed a 5.51% decrease in PSA per 5 kg/m² increase in BMI (95% CI 3.62–7.36) and univariable regression showed a 14.25% increase in PSA per 5-year increase in age (95% CI 12.54–15.98). Sensitivity and specificity for a threshold of 3 ng/ml were 0.900 and 0.898 and for a threshold of 4 ng/ml were 0.673 and 0.930, respectively. These results indicate that while the sensitivities and specificities were slightly less than the full model (3 ng/ml: 0.942, 0.932; 4 ng/ml: 0.708, 0.967), the association between PSA and age and BMI in men without prostate cancer in this population is similar to that in the entire population.

In the sensitivity analysis where BMI and diabetes data were imputed ($n = 19,524$), multivariable linear regression showed a 5.22% decrease in PSA per 5 kg/m² increase in BMI (95% CI 3.67–6.80) and univariable regression showed a 13.33% increase in PSA per 5-year increase in age (95% CI 12.14–14.54). Averaged sensitivity and specificity over ten imputations for a threshold of 3 ng/ml were 0.941 and 0.938 and for a threshold of 4 ng/ml were 0.701 and 0.970, respectively. These values are very similar to the main analysis indicating missing height and diabetes data were unlikely to have biased the result.

Discussion

This study has shown that in men aged 50–69 in the UK there is an inverse relationship between PSA and BMI, and a positive relationship between age and PSA. The magnitude and direction of effect of these relationships is consistent with previous research [15, 18–23]. In previous research,

Table 4 Averaged sensitivities and specificities for prostate cancer detection at biopsy for different models from tenfold cross-validation

Model	Threshold 3.0 ng/ml		Threshold 4.0 ng/ml	
	Proportion of cases above threshold (SD)	Proportion of controls below threshold (SD)	Proportion of cases above threshold (SD)	Proportion of controls below threshold (SD)
PSA	1 (0)	0.931 (0.011)	0.687 (0.022)	0.964 (0.005)
Age–PSA	0.946 (0.019)	0.933 (0.008)	0.699 (0.028)	0.968 (0.005)
BMI–PSA	0.978 (0.008)	0.932 (0.010)	0.683 (0.020)	0.965 (0.005)
Age–BMI–PSA	0.977 (0.008)	0.932 (0.010)	0.706 (0.028)	0.967 (0.005)
NICE guidelines ^a	—	—	0.798 (0.017)	0.958 (0.006)

PSA prostate-specific antigen, BMI body mass index

^a NICE guidelines—ages 50–59: 3.0 ng/ml, ages 60–69: 4.0 ng/ml

however, the relationship between BMI and PSA may have been assumed to be additive; this study considered the relationship to likely be multiplicative and has presented the results as such. This interpretation fits better with the proposed theories of why BMI may be associated with PSA.

Overall, adjusting for BMI did not materially improve the discrimination of PSA for detecting prostate cancer, as there was no evidence the NICE guidelines for screening for prostate cancer using age-band-specific PSA thresholds performed worse than our age-, BMI-, and age-BMI-adjusted PSA models. It is unlikely that the associations seen between age, BMI, and PSA are solely the result of associations with prostate cancer, as the likelihood tests showed no evidence that the results from men with a PSA less than 3.0 ng/ml (not biopsied) were different from men with a PSA more than 3.0 ng/ml (biopsied).

It is not clear from this study whether there is a causal relationship between BMI and PSA—the relationships between BMI, PSA, and prostate cancer are complex, and as not all men were biopsied it is impossible to disentangle relationships with underlying prostate cancer. Even if all men were biopsied, there would still be a risk of false negatives from missing the cancer on the biopsy [37, 38]. This, combined with the use of PSA testing as a means of screening for prostate cancer in general practice (as well as any variables which affect the overall risk of receiving a PSA test), makes it difficult for any study to examine the relationships surrounding PSA and prostate cancer risk.

The limitations of this study are recognized. The study population is large and taken from multiple centers across the UK, but will not necessarily be demographically diverse or applicable to other populations as almost all men classified themselves as “white” ethnicity. Although the tenfold cross-validation show very similar results to the main analysis, these results were not replicated in an external dataset. However, as internal validation usually shows some measure of overfitting, the conclusion that the NICE guidelines show better discrimination for prostate cancer should be robust.

PSA values were only taken on one day, but PSA levels are influenced by factors such as biological [43] or laboratory variation [44], inflammation [15] or infection [45]. Some evidence exists for seasonal variation in PSA [46]; however, a study using ProtecT data showed no variation in PSA due to time of year or amount of sunlight per day [47] and thus time of year should not have affected these results. As not all men were biopsied, the calculated specificities and sensitivities are likely overestimates as there were likely to be men with undiagnosed prostate cancer (Box 2). ROC curves and AUCs could not be generated as a result, and these results should not be compared with the true sensitivity and specificity of PSA as a test for detecting prostate cancer.

Height data were self-reported; this may have slightly biased the results if taller or shorter men were more likely to misreport their height. Additionally, a large number of men did not have a recorded height as they did not participate in the additional study that recorded height data; data were likely to be missing at random (i.e., it is unlikely that taller men were more likely to have missing height data). Multiple imputations showed that missing height data were unlikely to have biased the analysis as the models were very similar.

Men with diabetes were not included in the model, as BMI is a recognized risk factor for diabetes, and diabetes is associated with a lower PSA [48] and prostate cancer [36], so men with diabetes may obscure the relationship between BMI and PSA. Although there is conflicting evidence, smoking [49], exercise [50], and a low-fat diet [51] have all been associated with decreased PSA and BMI, and high alcohol intake and benign prostatic hypertrophy [52] have been associated with increased PSA and BMI [53]. These variables were not considered in this study, but these associations would indicate a positive relationship between BMI and PSA which is not observed in this study; thus, it is unlikely that the observed BMI-PSA relationship was biased away from the null by any of these variables.

Conclusions

This study has described the relationship between age and BMI and PSA in men without diabetes. These relationships were used to adjust PSA to examine the potential clinical utility of an age-BMI-adjusted PSA, but this did not perform better than current NICE guidelines. More studies examining the effects of variables on PSA, independent of the effect on prostate cancer, could help to improve PSA testing for prostate cancer.

Acknowledgements The ProtecT trial is funded by the UK National Institute for Health Research (NIHR) Health Technology Assessment Programme (projects 96/20/06, 96/20/99) with the University of Oxford (Oxford, UK) as sponsor. The views and opinions expressed herein are our own and do not necessarily reflect those of the Department of Health. We acknowledge the tremendous contribution of all the ProtecT study participants, investigators, researchers, data monitoring committee, and trial steering committee (Chair: Michael Baum). We acknowledge the support from the Oxford NIHR Biomedical Research Centre through the Surgical Innovation and Evaluation Theme and the Surgical Interventional Trials Unit, and Cancer Research UK through the Oxford Cancer Research Centre. This work was supported by Cancer Research UK project Grant Grants C11043/A4286, C18281/A8145, C18281/A11326 and C18281/A15064. The ProtecT study is supported by the UK National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme, HTA 96/20/99; ISRCTN20141297. The authors would like to acknowledge the support of the National Cancer Research Institute (NCRI) formed by the Department of Health, the Medical Research Council (MRC) and Cancer Research UK. The

NCRI provided funding through ProMPT (Prostate Mechanisms of Progression and Treatment), and this support is gratefully acknowledged. SH is a Wellcome Trust-funded PhD student with Grant Code 102432/Z/13/Z. The funding source had no role in the design, conduct of the study, collection, management, analysis and interpretation or preparation, review, or approval of the article.

Conflict of interest The authors declare no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr/old/FactSheets/cancers/prostate-new.asp>. Accessed on 7 Nov 2016
2. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M et al (2009) Screening and prostate-cancer mortality in a randomized European study. *New Engl J Med* 360(13):1320–1328
3. Andriole GL, Crawford ED, Grubb RL 3rd, Buys SS, Chia D, Church TR, Fouad MN, Gelmann EP, Kvale PA, Reding DJ et al (2009) Mortality results from a randomized prostate-cancer screening trial. *New Engl J Med* 360(13):1310–1319
4. Moyer VA, Force USPST (2012) Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 157(2):120–134
5. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, Kwiatkowski M, Lujan M, Lilja H, Zappa M et al (2012) Prostate-cancer mortality at 11 years of follow-up. *New Engl J Med* 366(11):981–990
6. Croswell JM, Kramer BS, Crawford ED (2011) Screening for prostate cancer with PSA testing: current status and future directions. *Oncology* 25(6):452–460, 463
7. Burton AJ, Martin RM, Donovan JL, Lane JA, Davis M, Hamdy FC, Neal DE, Tilling K (2012) Associations of lifestyle factors and anthropometric measures with repeat PSA levels during active surveillance/monitoring. *Cancer Epidemiol Biomark Prev* 21(10):1877–1885
8. Hosain GM, Sanderson M, Du XL, Chan W, Strom SS (2011) Racial/ethnic differences in predictors of PSA screening in a tri-ethnic population. *Cent Eur J Public Health* 19(1):30–34
9. Rodriguez S, Al-Ghamdi OA, Burrows K, Guthrie PA, Lane JA, Davis M, Marsden G, Alharbi KK, Cox A, Hamdy FC et al (2013) Very low PSA concentrations and deletions of the KLK3 gene. *Clin Chem* 59(1):234–244
10. Safarinejad MR, Asgari SA, Farshi A, Iravani S, Khoshdel A, Shekarchi B (2013) Opium consumption is negatively associated with serum prostate-specific antigen (PSA), free PSA, and percentage of free PSA levels. *J Addict Med* 7(1):58–65
11. Helfand BT, Loeb S, Hu Q, Cooper PR, Roehl KA, McGuire BB, Baumann NA, Catalona WJ (2013) Personalized prostate specific antigen testing using genetic variants may reduce unnecessary prostate biopsies. *J Urol* 189(5):1697–1701
12. Marberger M, Freedland SJ, Andriole GL, Emberton M, Pettaway C, Montorsi F, Teloken C, Rittmaster RS, Somerville MC, Castro R (2012) Usefulness of prostate-specific antigen (PSA) rise as a marker of prostate cancer in men treated with dutasteride: lessons from the REDUCE study. *BJU Int* 109(8):1162–1169
13. Kim YJ, Kim SO, Ryu KH, Hwang IS, Hwang EC, Oh KJ, Jung SI, Kang TW, Kwon DD, Park K et al (2011) Prostate cancer can be detected even in patients with decreased PSA less than 2.5 ng/ml after treatment of chronic prostatitis. *Korean J Urol* 52(7):457–460
14. Yang L, Egger M, Plattner R, Klocker H, Eder IE (2011) Lovastatin causes diminished PSA secretion by inhibiting AR expression and function in LNCaP prostate cancer cells. *Urology* 77(6):1508 e1501–1508 e1507
15. Man LB, Li GZ, Huang GL, Wang JW, Liu BY (2012) Aggressiveness and extent of prostatic inflammation relates with serum PSA levels in type IV prostatitis. *Natl J Androl* 18(8):710–714
16. Muller H, Raum E, Rothenbacher D, Stegmaier C, Brenner H (2009) Association of diabetes and body mass index with levels of prostate-specific antigen: implications for correction of prostate-specific antigen cutoff values? *Cancer Epidemiol Biomark Prev* 18(5):1350–1356
17. Lin Y, Mao Q, Zheng X, Yang K, Chen H, Zhou C, Xie L (2011) Human papillomavirus 16 or 18 infection and prostate cancer risk: a meta-analysis. *Int J Med Sci* 180(2):497–503
18. Pater LE, Hart KW, Blonigen BJ, Lindsell CJ, Barrett WL (2012) Relationship between prostate-specific antigen, age, and body mass index in a prostate cancer screening population. *Am J Clin Oncol* 35(5):490–492
19. Werny DM, Thompson T, Saraiya M, Freedman D, Kottiri BJ, German RR, Wener M (2007) Obesity is negatively associated with prostate-specific antigen in U.S. men, 2001–2004. *Cancer Epidemiol Biomark Prev* 16(1):70–76
20. Naito M, Asai Y, Mori A, Fukuda Y, Kuwabara M, Katase S, Hishida A, Morita E, Kawai S, Okada R et al (2012) Association of obesity and diabetes with serum prostate-specific antigen levels in Japanese males. *Nagoya J Med Sci* 74(3–4):285–292
21. Culp S, Porter M (2009) The effect of obesity and lower serum prostate-specific antigen levels on prostate-cancer screening results in American men. *BJU Int* 104(10):1457–1461
22. Banez LL, Hamilton RJ, Partin AW, Vollmer RT, Sun L, Rodriguez C, Wang Y, Terris MK, Aronson WJ, Presti JC Jr et al (2007) Obesity-related plasma hemodilution and PSA concentration among men with prostate cancer. *JAMA* 298(19):2275–2280
23. Baillargeon J, Pollock BH, Kristal AR, Bradshaw P, Hernandez J, Basler J, Higgins B, Lynch S, Rozanski T, Troyer D et al (2005) The association of body mass index and prostate-specific antigen in a population-based study. *Cancer* 103(5):1092–1095
24. Rodriguez C, Freedland SJ, Deka A, Jacobs EJ, McCullough ML, Patel AV, Thun MJ, Calle EE (2007) Body mass index, weight change, and risk of prostate cancer in the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomark Prev* 16(1):63–69
25. Wright JL, Lin DW, Stanford JL (2011) The effect of demographic and clinical factors on the relationship between BMI and PSA levels. *Prostate* 71(15):1631–1637
26. Oesterling JE (1996) Age-specific reference ranges for serum PSA. *New Engl J Med* 335(5):345–346
27. Rundle AG, Neugut AI (2009) Modeling the effects of obesity and weight gain on PSA velocity. *Prostate* 69(14):1573–1578
28. Loeb S, Carter HB, Schaeffer EM, Ferrucci L, Kettermann A, Metter EJ (2009) Should prostate specific antigen be adjusted for body mass index? Data from the Baltimore Longitudinal Study of Aging. *Journal Urol* 182(6):2646–2651

29. National Institute for Health and Care Excellence 2015 guidelines for suspected cancer, part 1.8.6. Available from: <https://www.nice.org.uk/guidance/NG12/chapter/1-Recommendations-organised-by-site-of-cancer#urological-cancers>. Accessed on 7 Nov 2016
30. Roobol MJ, Zhu X, Schroder FH, van Leenders GJ, van Schaik RH, Bangma CH, Steyerberg EW (2013) A calculator for prostate cancer risk 4 years after an initially negative screen: findings from ERSPC Rotterdam. *Eur Urol* 63(4):627–633
31. Thompson IM, Ankerst DP, Chi C, Goodman PJ, Tangen CM, Lucia MS, Feng Z, Parnes HL, Coltman CA Jr (2006) Assessing prostate cancer risk: results from the Prostate Cancer Prevention Trial. *J Natl Cancer Inst* 98(8):529–534
32. Patel S, Issa MM, El-Galley R (2013) Evaluation of novel formula of PSA, age, prostate volume, and race in predicting positive prostate biopsy findings. *Urology* 81(3):602–606
33. Liang Y, Ankerst DP, Sanchez M, Leach RJ, Thompson IM (2010) Body mass index adjusted prostate-specific antigen and its application for prostate cancer screening. *Urology* 76(5):1268 e1261–1266
34. National Statistics, The NHS Information Centre for Health and Social Care, Lifestyle Statistics. Statistics on Obesity, Physical Activity and Diet: England, 2012. <http://www.hscic.gov.uk/article/2021/Website-Search?productid=10152&q=obesity&sort=Relevance&size=10&page=1&area=both#top>
35. Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, Turner EL, Mason MD, Metcalfe C, Peters TJ et al (2014) Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol* 15(10):1109–1118
36. Turner EL, Lane JA, Donovan JL, Davis MJ, Metcalfe C, Neal DE, Hamdy FC, Martin RM (2011) Association of diabetes mellitus with prostate cancer: nested case-control study (prostate testing for cancer and treatment study). *Int J Cancer* 128(2):440–446
37. Keetch DW, Catalona WJ, Smith DS (1994) Serial prostatic biopsies in men with persistently elevated serum prostate specific antigen values. *J Urol* 151(6):1571–1574
38. Haas GP, Delongchamps NB, Jones RF, Chandan V, Serio AM, Vickers AJ, Jumbelic M, Threlate G, Korets R, Lilja H et al (2007) Needle biopsies on autopsy prostates: sensitivity of cancer detection based on true prevalence. *J Natl Cancer Inst* 99(19):1484–1489
39. McNemar Q (1947) Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 12(2):153–157
40. Egan JP (1975) Signal detection theory and roc analysis, series in cognition and perception. Academic Press, New York
41. Mosteller F, Tukey JW (1968) Data analysis, including statistics. In: Lindzey G, Aronson E (eds) *Handbook of Social Psychology*, vol 2, Addison-Wesley
42. Altman DG, Bland JM (2003) Interaction revisited: the difference between two estimates. *BMJ* 326(7382):219
43. Soletormos G, Semjonow A, Sibley PE, Lamerz R, Petersen PH, Albrecht W, Bialk P, Gion M, Junker F, Schmid HP et al (2005) Biological variation of total prostate-specific antigen: a survey of published estimates and consequences for clinical practice. *Clin Chem* 51(8):1342–1351
44. Forde JC, Marignol L, Blake O, McDermott T, Grainger R, Crowley VE, Lynch TH (2012) Standardization of assay methods reduces variability of total PSA measurements: an Irish study. *BJU Int* 110(5):644–650
45. Ulleryd P, Zackrisson B, Aus G, Bergdahl S, Hugosson J, Sandberg T (1999) Prostatic involvement in men with febrile urinary tract infection as measured by serum prostate-specific antigen and transrectal ultrasonography. *BJU Int* 84(4):470–474
46. Naselli A, Fontana V, Introini C, Andreatta R, Puppo P (2011) Effect of age, family history of prostate cancer, prostate enlargement and seasonality on PSA levels in a contemporary cohort of healthy Italian subjects. *Int J Biol Markers* 26(2):102–107
47. Down L, Metcalfe C, Martin RM, Neal DE, Hamdy FC, Donovan JL, Lane JA (2011) Seasonal variation in prostate-specific antigen levels: a large cross-sectional study of men in the UK. *BJU Int* 108(9):1409–1414
48. Fiukui M, Tanaka M, Kadono M, Imai S, Hasegawa G, Yoshi-kawa T, Nakamura N (2008) Serum prostate-specific antigen levels in men with type 2 diabetes. *Diabetes Care* 31(5):930–931
49. Li J, Thompson T, Joseph DA, Master VA (2012) Association between smoking status and free, total, and percent free prostate-specific antigen. *Am J Epidemiol* 175:S24–S24
50. Oremek GM, Seiffert UB (1996) Physical activity releases prostate-specific antigen (PSA) from the prostate gland into blood and increases serum PSA concentrations. *Clin Chem* 42(5):691–695
51. Demark-Wahnefried W, Robertson CN, Walther PJ, Polascik TJ, Paulson DF, Vollmer RT (2004) Pilot study to explore effects of low-fat, flaxseed-supplemented diet on proliferation of benign prostatic epithelium and prostate-specific antigen. *Urology* 63(5):900–904
52. Wang S, Mao Q, Lin Y, Wu J, Wang X, Zheng X, Xie L (2012) Body mass index and risk of BPH: a meta-analysis. *Prostate Cancer Prostatic Dis* 15(3):265–272
53. Nackaui JDE, Colla RH, Ravazzani GR, Gaido MI, Bertolotto P, Actis AB (2012) Prostate-specific antigen: its relationship with alcohol intake and tobacco. *Med Oncol* 29(2):823–826