UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Application of surrogate outcomes for reimbursement decision-making in prostate cancer

by Rafael Emilio de Feria Cardet

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

under the supervision of

Stephen Goodall
Richard De Abreu Lourenço
Tracey-Lea Laba
Marion Haas
Ian Davis.

University of Technology Sydney
Faculty of Health
Centre for Health Economics Research and Evaluation

January 2024

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Rafael Emilio de Feria Cardet declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Centre for Health Economics Research and Evaluation\ Faculty of Health at the University of Technology Sydney. This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Production Note:
Signature: Signature removed prior to publication.

Date: 4/4/2023

On my brave heart is engraved

The sorrow hidden from all eyes:

The son of a land enslaved

Lives for it, suffers and dies.

José Martí

Acknowledgments

First and foremost, my deepest gratitude to my wife. Coni, your unwavering love, support, and patience have been the cornerstones of my PhD. Your presence during both the highs and lows, have made this challenging endeavour infinitely more meaningful. Juggling parenthood and a PhD would not have been feasible without your encouragement and assistance. Thank you for being my constant source of strength, I could not have accomplished this without you.

To my *familia*, particularly *Mami* and *Papi*. Your support, love, and belief in me have been a constant inspiration. Your sacrifices and encouragement have profoundly influenced the person I am today. I am deeply thankful for the values you instilled in me. To my children, Rafael, Elena and Camilo, I am immensely thankful for the joy and motivation you brought to my life. I hope to have made you all proud.

I am extremely grateful to my supervisors, for their guidance and expertise over these years. Thank you for the encouragement and for pushing me to think more critically. I appreciate you all taking me under your wing and sharing your wealth of knowledge in the field. Your mentorship has helped me develop as a researcher.

Finally, I am grateful for the financial support provided by ANZUP.

Contents

Acknowledgments iv
Table headings xiv
Figures heading xvi
Abbreviation list xviii
Abstract xxiii
Chapter 1 The use of surrogate outcomes in economic evaluations 25
1.1 Abstract 25
1.2 Introduction 25
1.3 Assessing benefit 26
1.3.1. Surrogate outcomes 27
1.3.1.1. Preferences in the use of surrogate outcomes 28
1.3.1.2. Surrogate outcomes in cancer 29
1.3.2. Validation of surrogate outcomes 29
1.3.3. Surrogate outcomes in reimbursement decision-making 31
1.3.4. The decision-maker 32
1.4 Economic evaluation 32
1.4.1.1. Economic Evaluation and Health Technology Assessment 34
1.5 Prostate Cancer 36
1.5.1. The health problem 37
1.5.2. Diagnosis and risk classification of prostate cancer 38
1.5.3. Treatment of prostate cancer 40
1.5.4. Motivating example 41
1.5.5. Surrogate outcomes and prostate cancer 42
1.5.5.1. Chain surrogacy 43
1.6 Aims and Objectives 43
1.7 Overview of the research 44
Chapter 2 Systematic review of economic evaluations incorporating surrogate outcome measures 46
2.1 Abstract 46

2.2 Introduction 46
2.3 Objectives 47
2.4 Methods 48
2.4.1. Search strategy 48
2.4.2. Exclusion criteria 49
2.4.3. Review process 50
2.4.4. Data Extraction 50
2.4.5. Quality assessment 51
2.5 Results 52
2.5.1. Excluded records 55
2.5.2. Surrogate outcome used 56
2.5.3. Quality assessment 58
2.5.4. The case of metastatic hormone refractory prostate cancer 59
2.6 Discussion 59
2.7 Limitations of the systematic review 61
2.8 Conclusions 62
Chapter 3 Evaluation of evidence on the use of MFS as a surrogate outcome in LPC for economic evaluations 63
3.1 Abstract 63
3.2 Introduction 63
3.2.1. Surrogate outcomes in healthcare decision-making 63
3.2.1.1. Assessment of surrogate outcomes for use in economic evaluations 65
3.3 Objectives 66
3.4 Methods 66
3.4.1. The ICECaP database 67
3.4.1.1. ICECaP Methods 68
3.4.1.1.1. Statistical analysis within ICECaP 69
3.4.1.1.2. ICECaP outcome definition 70
3.4.1.2. ICECaP results 71
3.4.1.2.1. The surrogacy relationship 72
3.4.1.2.2. Refining the ICECaP equation for health economics 72

3.4.2. The PBAC Surrogate Framework 74
3.4.3. Survival analysis 75
3.4.3.1. Mean survival time 77
3.4.3.2. Proportional hazard assumption for the intervention 77
3.4.3.3. Adjustment of survival curves 79
3.4.4. Recurrent disease 80
3.5 Results 80
3.5.1. The surrogacy relationship for health economics 80
3.5.2. Assessment of ICECaP against the PBAC Surrogate Framework 82
3.5.2.1. Definition, selection, and measurement. 82
3.5.2.1.1. Proposed surrogate measure. 82
3.5.2.1.2. Final outcome and relationship with the proposed surrogate 83
3.5.2.2. Biological reasoning and randomised trial evidence 84
3.5.2.2.1.1. Biological reasoning 84
3.5.2.3. Mechanism of actions of interventions in the ICECaP study 85
3.5.2.4. Statistical analysis 87
3.5.2.4.1. The surrogacy relationship 88
3.5.3. Survival analysis 89
3.5.3.1. Overall survival 89
3.5.3.2. Metastasis-free survival 90
3.5.3.3. Biochemical recurrence-free survival 91
3.5.3.4. Adjustment of the overall survival curve for population mortality 92
3.5.3.5. Mean survival time 93
3.6 Discussion 93
3.6.1. The ICECaP trials and biological reasoning 94
3.6.2. Statistical analysis of the surrogacy relationship 95
3.6.3. Results of the survival analysis 96
3.7 Conclusions 97
Chapter 4 Incorporating a surrogate outcome in an economic evaluation. 98
4.1 Abstract 98
4.2 Introduction 99
4.2.1.1. Economic evaluation and surrogate outcomes 100

4.2.2. Treatment of prostate cancer 100
4.2.3. Local treatment 100
4.2.3.1. Hormone-focused therapy and advanced disease 102
4.2.3.2. Chemotherapy 105
4.2.4. Imaging technique and monitoring 105
4.2.5. Health-related quality of life in patients with prostate cancer 106
4.2.6. Costs 107
4.2.7. Time horizon and the need for extrapolation 108
4.2.8. A timely decision for society 108
4.2.8.1. Opportunity cost 109
4.3 Objectives 109
4.4 Pre-methodological assessment to determine model structure 110
4.5 Methods 113
4.5.1. Structure of the economic model: partitioned survival model 114
4.5.2. Patient population, intervention, and comparator 114
4.5.3. Treatment pathways in the model 115
4.5.4. Survival analysis 118
4.5.5. Treatment effect 119
4.5.5.1. The surrogate threshold effect 120
4.5.6. Perspective 120
4.5.7. Costs 121
4.5.7.1. Follow-up costs 121
4.5.7.2. Intervention and comparator cost 123
4.5.7.3. Systemic therapy 124
4.5.7.4. Adverse events 127
4.5.8. Utilities 130
4.5.9. Sensitivity analysis 133
4.5.9.1. Scenario analysis 134
4.5.9.1.1. Testing the surrogacy relationship 134
4.5.10. Benefits to society 135
4.5.10.1. Flow of patients 135

4.5.10.2. Timely decision....136
4.5.11. Value of information 138
4.5.12. Applicability beyond high-risk LPC patients 140
4.6 Results....140
4.6.1. Model of choice 141
4.6.2. Estimating Survival Probabilities using the ICECaP database 142
4.6.3. Costs 142
4.6.4. Utilities 143
4.6.5. Incremental analysis 144
4.6.6. Sensitivity analysis 144
4.6.6.1. The surrogacy relationship 146
4.6.6.2. Scenario analysis 146
4.6.6.3. Waning effect 148
4.6.7. Validation 149
4.6.8. Benefits to society 149
4.6.8.1. Flow of patients 149
4.6.8.2. Delayed decision 150
4.6.8.3. Value of information 151
4.7 Discussion 152
4.7.1. Model of choice 153
4.7.1.1. Limitations of the model 154
4.7.2. The cost of uncertainty 156
4.7.3. Health-related quality of life 158
4.7.4. Applicability issues 159
4.7.5. Delayed decision 160
4.8 Conclusions 161
Chapter 5 Discrete choice experiment 162
5.1 Abstract 162
5.2 Introduction 162
5.2.1. Background 162
5.2.2. Measuring preferences 163

5.2.2.1. Stated preferences 164
5.3 Objective 164
5.4 Methods 165
5.4.1. Theoretical framework 165
5.4.2. Development of a DCE 166
5.4.2.1. Identification of relevant attributes and levels 166
5.4.2.1.1. Focus group 167
5.4.2.1.2. Literature review 169
5.4.2.2. Discussion with the ICECaP Steering Committee 176
5.4.2.3. Design 177
5.4.2.4. Survey conduct 179
5.4.2.5. Choice set development 179
5.4.2.6. Statistical analysis 180
5.4.2.6.1. Multinomial logit model 181
5.4.2.6.2. Heterogeneity 182
5.4.2.7. Willingness to accept analysis 184
5.5 Results 185
5.5.1. Follow-up choice question 187
5.5.2. Conditional logit model 188
5.5.3. Heterogeneity 190
5.5.3.1. Mixed logit model 190
5.5.3.2. Latent class model 192
5.5.3.3. Willingness-to-accept analysis 198
5.5.4. Subgroup analysis 199
5.6 Discussion 200
5.6.1. Surrogate outcomes to support decisions 201
5.6.2. Limitations and future research 202
5.6.2.1. Population 202
5.6.2.2. Eliciting preferences 203
5.6.2.3. Surrogate outcome as an attribute 204
5.7 Conclusion 204
Chapter 6 Discussion 205

6.1 Abstract 205
6.2 The decision-making process 206
6.3 MFS as a surrogate of OS 207
6.3.1. The ICECaP data and validation 207
6.3.2. Chain surrogacy 209
6.3.3. Suggestions for embedding the use of surrogate outcomes in reimbursement decisions 210
6.4 Incorporating MFS in an economic model for LPC 212
6.4.1. The model 212
6.4.2. The role of additional research 214
6.4.3. Ethical reimbursement decisions 215
6.5 Preference for the use of surrogates in reimbursement decisions 216
6.6 The need for collaboration 217
6.7 Plausibility beyond equipoise 218
6.8 Concluding remarks 219
Appendix 1: Search Strategy for the Systematic Review 221
Appendix 2: Quality assessment of included studies 223
Appendix 3. Detailed assessment of MFS as a surrogate of OS according to the PBAC Guidelines 228
Appendix 4. Extrapolation and model parameters 253
A. Extrapolation and model parameters for the overall population developing biochemical recurrence. 253
B. Extrapolation and model parameters for the overall population developing metastasis. 262
C. Extrapolation and model parameters for OS in the overall population. 269
Appendix 5 Survey conducted among steering committee members to validate the assumptions of resources use in the treatment of localised prostate cancer. 275
Appendix 6. Metanalysis of utilities found in the literature 299
A. Metanalysis results of utilities found for RP 299
B. Metanalysis results of utilities found for RT 300
C. Metanalysis results of utilities found for adverse events 301

D. Metanalysis results of selected studies reporting utilities in metastatic patients 304

Appendix 7. Model manual. 309

Table of Contents 310

6.9 Index of figures and tables 311

Figures 311

Tables 311

1 Manual overview 312

1.1 Context....312

1.2 The model 312

1.3 Workbook structure 313

2 Population.... 317

2.1 Time Horizon and discount rate 318

2.2 Perspective....319

3 Economic model.... 320

3.1 Structure of the economic model: partitioned survival model....320

3.2 Clinical input....323

3.2.1 Overall survival 326

3.2.2 Metastasis free survival.... 328

3.2.3 Biochemical recurrence free survival 330

3.2.4 General population mortality 331

3.3 The surrogacy relationship 332

3.4 Economic inputs....335

3.4.1 Primary treatment.... 340

3.4.1.1 Comparator 340

3.4.1.2 Intervention 341

3.4.2 Subsequent Treatment.... 341

3.4.3 Adverse events costs 343

3.4.4 Follow up.... 345

3.4.5 Other costs 345

3.5 Quality of life inputs 346

3.6 Model outputs....348
3.6.1 Benefits to society....349
3.7 Sensitivity analysis....350
3.7.1 Deterministic sensitivity analysis....351
3.7.2 Probabilistic sensitivity analysis....352
Attachment 1....353
Proportional hazard assumption for RP....353
Attachment 2....356
Appendix 7 Bibliography....365
Appendix 8. Systematic review search strategy and PRISMA diagram. 369
Appendix 9. DCE vignette 371
Appendix 10. Modifications to DCE after steering committee discussion 382
Bibliography 386

Table headings

Table 2.1. PICO framework 48

Table 2.2. Data extracted from studies found. 51

Table 2.3. Study characteristics 54

Table 3.1. Characteristics of ICECaP studies included and metastasis definition. 87

Table 4.1. Comparison of population among ICECaP and PCOR. 115

Table 4.2. Probability of treatment's adverse event occurrence assumed in the model. 118

Table 4.3. Treatment effect of the outcomes of interest reported by Attard et.al. used in the model. 120

Table 4.4. Imaging cost. 122

Table 4.5. Cost of follow-up patients. 123

Table 4.6. Cost of the intervention. 123

Table 4.7. Costs associated with RT. 124

Table 4.8. Cost of ADT. 125

Table 4.9. Cost of medication added to ADT for patients with M0CRPC. 126

Table 4.10. Cost of chemotherapy for metastatic patients. 126

Table 4.11. Cost of treatment for metastatic resistant to docetaxel and ADT. 127

Table 4.12. Cost prostate cancer treatment adverse events. 128

Table 4.13. Cost of ADT-related adverse events. 130

Table 4.14. Pooled estimates of utility values used in the model. 133

Table 4.15. Total cost by health state and cost category in the model. 143

Table 4.16. QALYs per patient estimated in the model. 144

Table 4.17. Incremental analysis. 144

Table 4.18. Sensitivity analysis of the surrogacy relationship. 146

Table 4.19. Lifetime benefits to society of an earlier decision. 151

Table 4.20. EVPI at individual and population level at different effective lifetime of use AAP. 152

Table 5.1. Attribute domains from focus group 169

Table 5.2. Attributes identified in the literature review for prostate cancer. 170

Table 5.3. Attributes and levels for the health decision-making process identified by the systematic review of Trapero-Bertran et al.2019 and the focus group. 174

Table 5.4. Attributes and levels included in the survey. 177
Table 5.5. Demographic characteristics of the surveyed population. 186
Table 5.6. Other characteristics of respondents. 187
Table 5.7. Results of the conditional logit model. 189
Table 5.8. Results of mixed logit model. 191
Table 5.9. Information criterion for the latent class models using different number of classes. 193
Table 5.10. Demographic characteristics of respondents by class membership. 196
Table 5.11. Results of LCM with two classes. 197
Table 5.12. Wald test of linearty for tue time and cost coefficients. 198
Table 5.13. Results of the mixed logit model for the subgroup of decision-makers population in the survey. 200

Figures heading

Figure 2.1. Prisma diagram of literature search results. 53

Figure 3.1. Correlation between MFS and OS at individual and trial level for all the studies included in ICECaP. 72

Figure 3.2. PRISMA diagram of studies identified. 73

Figure 3.3. Difference in the surrogacy relationship due to the data used. 82

Figure 3.4. Survival estimates for all patients in ICECaP database 85

Figure 3.5. Overall survival from ICECaP and extrapolation. 89

Figure 3.6. Biochemical failure-free survival from ICECaP and extrapolation. 90

Figure 3.7. Metastasis-free survival from ICECaP and extrapolation. 91

Figure 3.8. Adjustment of survival using the overall population mortality records in the model. 92

Figure 3.9. Mean survival time of all health states in the model by population. 93

Figure 4.1. Clinical algorithm of localised prostate cancer treatment. 101

Figure 4.2. A Markov model for localised prostate cancer. 111

Figure 4.3. Partitioned survival model for localised prostate cancer. 112

Figure 4.4. Study duration required to observe the number of events according to HRs. 137

Figure 4.5. Diagram of the model comparing timely vs delay decision in LPC. 138

Figure 4.6. State membership for the overall population in ICECaP. 141

Figure 4.7. Adjusted survival curves for AAP and SoC with extrapolation. 142

Figure 4.8 . Total cost per patient for SoC and AAP in the overall population. 143

Figure 4.9. Tornado plot summarising multiple one-way sensitivity analyses for the overall population. 145

Figure 4.10. Probabilistic sensitivity analysis. 145

Figure 4.11. Variations in total cost by parametric model for survival analysis. 147

Figure 4.12. Variations in total QALYS gained by parametric model for survival analysis. 148

Figure 4.13. Impact of the waning effect of the intervention on results. 149

Figure 4.14. Flow of patients in the timely and delayed decision arms in the model. 150

Figure 5.1. Example of choice set. 180

Figure 5.2. Number of respondents who would recommend the chosen option or not. 188

Figure 5.3. Relative importance of attributes and levels in the conditional logit model 1. 190

Figure 5.4. Relative importance of attributes in the mixed logit model.....192
Figure 5.5. Utility estimates when the levels of the coefficients of time and cost are used.....198
Figure 5.6. Willingness to accept analysis using the results of the mixed logit model.....199

Abbreviation list

3D-CRT	Three-dimensional conformal radiotherapy
AAP	Abiraterone plus prednisolone
AE	Adverse events
AADT	Adjuvant androgen deprivation therapy
AIC	Akaike information criterion
ADT	Androgen deprivation therapy
AUC	Area under the curve
ABS	Australian Bureau of Statistics
BIC	Bayesian information criterion
BCR	Biochemical relapse
Caps	Capsules
CRPC	Castration resistant prostate cancer
CT	Computed tomography
CAIC	Conditional Akaike information criterion
CI	Confidence interval
CHEERS	Consolidated Health Economic Evaluation Reporting Standards
CVM	Contingent valuation method
CBA	Cost-benefit analysis
CEA	Cost-effectiveness analysis
CMA	Cost-minimisation analysis
CUA	Cost-utility analysis
CREA	Creatinine
DNA	Deoxyribonucleic acid

DRE	Digital rectal examination
DCE	Discrete choice experiment
DFS	Disease-free survival
DM	Distant metastasis
EBRT	External beam radiotherapy
ECOG	Eastern Cooperative Oncology Group
ED	Erectile dysfunction
EMA	European Medicines Agency
EORTC	European Organisation for Research and Treatment of Cancer
EFS	Event-free survival
EVPI	Expected value of perfect information
FN	False negative
FP	False positive
FDA	Food and Drug Administration
Ga	Gallium
GI	Gastrointestinal toxicity
GP	General practitioner
HbA1c	Glycosylated haemoglobin
HR	Hazard ratio
HRQoL	Health-related quality of life
HTA	Health technology assessment
HUI	Health Utilities Index
HDBT	High dose rate brachytherapy
HIV	Human immunodeficiency virus

HREC	Human research ethics committee
ICER	Incremental cost-effectiveness ratio
IPD	Individual patient data
IMRT	Intensity-modulated radiotherapy
ICECaP	Intermediate Clinical Endpoints in Cancer of the Prostate
KM	Kaplan–Meier
LRP	Laparoscopic radical prostatectomy
LCM	Latent class models
LPC	Localised prostate cancer
LT	Long term
LDBT	Low dose rate brachytherapy
LHRH	Luteinising hormone-releasing hormone
MRI	Magnetic resonance imaging
MSAC	Medical Services Advisory Committee
MeSH	Medical Subject Headings
MBS	Medicare Benefits Schedule
MET	Metastasis
MFS	Metastasis-free survival
mo	Month
AMI	Acute myocardial infarction
NICE	National Institute for Health and Care Excellence
NHS	National Health Service
NHCDC	National Hospital Cost Data Collection
NADT	Neoadjuvant androgen deprivation therapy

NA	Not applicable
ORP	Open radical prostatectomy
OS	Overall survival
PROs	Patient-reported outcomes
PSM	Partitioned survival model
PBAC	Pharmaceutical Benefits Advisory Committee
PBS	Pharmaceutical Benefits Scheme
PET	Positron emission tomography
PFS	Progression-free survival
PS	Proposed surrogate
PCOR	Prostate Cancer Outcomes Registry
PSA	Prostate specific antigen
PSMA	Prostate-specific membrane antigen
PLAC	Prostheses List Advisory Committee
QALY	Quality adjusted life year
QHES	Quality of Health Economic Studies
QoL	Quality of life
RP	Radical prostatectomy
RT	Radiotherapy
RR	Risk ratio
RARP	Robotic-assisted radical prostatectomy
RCT	Randomised clinical trial
Scht	Sachet
SF-36	Short Form 36 Health Survey Questionnaire

ST	Short term
STFO	Surrogate to final outcome
Tbl	Tablets
TCO	Target clinical outcome
TDR	Time to disease recurrence
TTM	Time to metastasis
TTO	Time trade-off
Tx	Treatment
TNM	Tumour node metastasis classification system
UG	Urological toxicities
WTP	Willingness to pay
WTW	Willingness to wait

Abstract

This PhD thesis examines the use of surrogate outcomes to support reimbursement decision-making for patients with localised prostate cancer (LPC). The selection of LPC was motivated by its high prevalence and its natural history necessitating lengthy clinical trials, most often powered on detecting differences in overall survival (OS). The thesis begins by outlining the difficulties of conducting clinical trials due to the extended natural history of LPC. To address this challenge, surrogate outcomes can enable earlier observations before final outcomes become available. Within prostate cancer, the ICECaP study demonstrated that 5-year metastases free survival (MFS) was a strong surrogate for 8-year OS. This implies a potential reduction of three years in the reporting time of LPC trial data, thereby enabling earlier access to new treatments if MFS informs an economic evaluation as the foundation for reimbursement decisions.

A systematic review indicated a lack of robust published evidence supporting the use of surrogate outcomes in economic evaluations of LPC. The objectives of this thesis were threefold: first, to evaluate the viability of applying surrogate outcomes in economic evaluations for LPC; second, to construct an economic model using a surrogacy relationship to assess the societal benefits and costs associated with the use of surrogate outcomes in reimbursement decisions for LPC; and third, to examine the trade-offs between data maturity and time in forming advice for reimbursement considerations.

The nature of the surrogacy relationship impacts the data requirements for the model which in turn influences how can be constructed. Therefore, the model selection was influenced by the available data and the underlying disease pattern of LPC. The model drew upon clinical guidelines and input from clinical experts to reflect practice in Australia. However, its flexibility enables it to be adjusted for use in other settings and clinical contexts.

The analysis of preferences indicated that the availability of data from longer-term studies was a significant factor in positively supporting reimbursement decisions. The results also showed that the potential uncertainty associated with shorter term trials, perhaps reflected in the use of surrogate outcomes, could be mitigated through interventions with lower (better) incremental cost-effectiveness ratios (ICERs).

The findings of this thesis indicate that constructing a model using a surrogate outcome is feasible and could potentially expedite a reimbursement decision-making process. However, for this

approach to ultimately enhance societal welfare hinges on whether the expected benefits materialise in relation to the opportunity cost of making an earlier decision.

Chapter 1 The use of surrogate outcomes in economic evaluations

1.1 Abstract

This thesis explores the use of surrogate outcomes to inform economic models used in reimbursement decisions of new interventions in healthcare, examining the challenges and benefits of, and preferences for, making an earlier reimbursement decision based on evidence from surrogate outcomes. In the context of healthcare, the use of clinical trials to demonstrate the safety and effectiveness of new interventions is a critical component of the regulatory and reimbursement process. However, clinical trials can take several years to report on overall survival (OS) outcomes, leading to an increase in the time required to make new treatments available to patients.

To address this issue, surrogate outcomes are being used as part of reimbursement decision-making processes to make potentially effective treatments available sooner than might occur if waiting for final outcomes like OS. However, there is a risk of introducing uncertainty into the reimbursement decision-making process when using surrogate outcomes to estimate long-term benefits as the treatment effect on the final outcome is not known. To minimise this risk, surrogate outcomes need to be validated and their clinical relevance needs to be established.

Despite the limitations of surrogate outcomes, they have the potential to benefit society and patients by allowing for the earlier introduction of potentially more effective interventions. The use of surrogate outcomes is particularly relevant in the case of localised prostate cancer, which has a high incidence rate among the Australian population. Additionally, the lengthy time frame required for randomised trials to report on OS outcomes can delay the introduction of effective treatments, can lead to decreased quality of life due to diseases progression and more aggressive treatments and ultimately, increase mortality. Therefore, the use of surrogate outcomes in economic evaluations of prostate cancer treatments can be a valuable alternative approach to evaluating treatment efficacy in a timely manner.

1.2 Introduction

As the world grapples with ever-rising healthcare costs and persistent disparities in access and outcomes, health economics has emerged as a critical tool for policymakers and researchers alike. With the aging of populations, the increasing burden of chronic disease like cancer, and rapid advances in medical technology, there is a growing need for evidence-based decision-making to guide efficient allocation of healthcare resources. Health economics offers a unique perspective on how to

optimise healthcare delivery, improve health outcomes, and ensure value for money. From cost-effectiveness analysis and health technology assessment (HTA) to health insurance and payment systems, health economics is the key to unlocking solutions for some of the most pressing healthcare challenges of our time.

Clinical trials are an integral part of the drug development process. However, phase 3 clinical trials are designed to establish the safety and efficacy of a drug, not necessarily to assess its cost-effectiveness or its impact on public health. Historically, clinical trials have been designed to meet regulatory requirements rather than reimbursement requirements. This has led to some challenges in the reimbursement process, as the evidence generated by clinical trials may not be sufficient to demonstrate the value of the intervention from a reimbursement perspective.

1.3 Assessing benefit

Choosing an endpoint or outcome to assess the benefit of new interventions in healthcare is one of the most critical factors in the design of clinical trials. The endpoint chosen must be clinically relevant and sufficiently sensitive to detect the desired treatment effect (52).

Clinical trials evaluating treatments can use endpoints tailored to the underlying disease and aim of the investigational therapy. For example, trials in respiratory conditions may examine lung function or exacerbation rates, while cardiovascular trials may assess outcomes like time to revascularization, angina episodes, or functional capacity measures. However, clinical trials often prioritise survival outcomes, particularly overall survival, as primary endpoints since they represent objective measures of treatment efficacy that are critical for regulatory approval. Overall survival is particularly relevant for evaluating the efficacy of treatments for life-threatening conditions, such as cancer. A literature review of the use of OS in cancer clinical trials found that 46% of the clinical effectiveness papers reported OS to be the primary endpoint of the trial and 49% reported OS as a secondary endpoint (53). However, survival outcomes do not account for the quality-of-life aspects of treatment. Benefits of treatments may be offset by the detrimental impact of side effects and reduced quality of life. Thus, the incorporation of quality-of-life outcomes can better inform clinical decision-making and provide a more complete picture of a treatment's overall impact on patients (54). Another important limitation of survival outcomes is that they may require a large sample size or long follow-up period for meaningful outcomes to be observed, making the conduct of clinical trials costly.

As a result of the limitations of survival outcomes, patient-reported (PROs) and disease specific outcomes have increasingly been used as primary outcomes in clinical trials, often relegating survival

to be a secondary outcome (55). Disease specific outcomes and PROs are valuable in understanding clinical pathways of new treatments and evaluating their impacts on patients' daily lives (56). They can be beneficial in assessing the immediate impact of a treatment and are often accepted to support approvals of new interventions by regulatory bodies and the clinical community. However, disease-specific outcomes and PROs are often seen as more subjective than survival outcomes, and importantly, they are not directly comparable across health issues/interventions (56). These outcomes thus present challenges in terms of their use to inform reimbursement decision-making which requires a comprehensive assessment of the benefits and costs of a treatment, including its long-term impact on the combined impact on patient survival and quality of life and the economic impact it has on the healthcare system (57).

1.3.1. Surrogate outcomes

When the final outcome can only be observed beyond the normal trial period, or when dealing with high mortality and morbidity diseases such as some types of cancer or human immunodeficiency virus (HIV), avoiding the delay of a potential lifesaving intervention has justified the use of intermediate outcomes without the evidence supporting their prediction ability (58). Regulatory agencies across the world have tried to minimise the uncertainty around the use of non-validated surrogated outcomes in clinical trials by assigning conditional approvals for new interventions. Intermediate outcomes likely to predict clinical benefits are widely accepted in this accelerated regulatory submission process on condition that further studies are carried out to demonstrate the impact of the intervention on the final outcome (59).

Given the potential advantage that surrogate outcomes offer in clinical trials, they are often used in this setting (60). The use of surrogate outcomes such as CD4 cell count in HIV trials, or glycosylated haemoglobin (HbA1c) in diabetes mellitus, has demonstrated their potential in providing evidence enabling earlier treatment of patients and promoted their use in other therapeutic areas based solely on epidemiological studies in which a correlation was shown between surrogate and final outcome. Nevertheless, the reliability of surrogate outcomes depends upon scientific evidence that supports the relationship with and predictive ability in relation to the final outcome (61).

Prior to delving into surrogate outcomes, it is important to note that the term 'surrogate outcome' has been used interchangeably with other terms such as 'biomarker' and 'biological marker', leading to potential confusion regarding its definition. However, according to the Biomarkers Definitions Working

Group[1] of the United States National Institutes of Health, a surrogate outcome must have two distinct characteristics (62):

- a surrogate outcome is intended to substitute a final outcome that is clinically relevant to patients; and
- a surrogate outcome has to predict clinical benefits based on scientific evidence.

Although intermediate endpoints or biomarkers have the potential to become surrogates, they do not always predict changes in a final outcome. A surrogate outcome has the connotation of replacement of the true outcome in a clinical study by another outcome. Intermediate outcomes or biomarkers, on the other hand, are medical signs, characteristics and/or measurements or a set of measurements indicative of symptoms, health or general wellbeing of patients (62). This definition of surrogate outcomes has been adopted by regulatory agencies and in the literature and the HTA process in Australia and around the world (55, 63-65). Throughout this document, the term 'surrogate outcome' will be used to refer to intermediate outcomes or biomarkers that are able to predict a treatment effect on a final outcome, but that always substitute for the final outcome in a clinically relevant manner to patients.

1.3.1.1. Preferences in the use of surrogate outcomes

Importantly, the use of surrogate outcomes to support making reimbursement decisions in healthcare raises questions about the trade-off between speed and reliability of evidence. On the one hand, using surrogate outcomes may reduce the time required to make new treatments available, enabling patients to access potentially effective treatments sooner. On the other hand, there are concerns that the use of surrogate outcomes can lead to approval of unsafe, ineffective or cost-ineffective treatments, which can have negative consequences for patients and society (55).

Society's preference for the use of surrogate outcomes is influenced by several factors, including the severity and urgency of the condition being treated, the availability of alternative treatments, and the potential risks and benefits of those treatments (66). In some cases, where the condition is life-threatening and there are no alternative treatments available, there may be a greater willingness to

[1] The Biomarkers Definitions Working Group of the United States National Institutes of Health was formed in 1998 to address the need for clear and consistent definitions of biomarkers in biomedical research. The working group consisted of experts in the fields of laboratory medicine, clinical trials, and regulatory affairs tasked with developing a framework for defining biomarkers and establishing criteria for their validation and clinical use. The group has developed guidelines and definitions for the use of surrogate outcomes, important for ensuring consistency and clarity in the use of surrogate outcomes in clinical research and decision-making. Its recommendations have had a significant impact on the field of biomarker research and development.

accept the use of surrogate outcomes to speed up access to treatments. However, in other cases, where alternative treatments are available, society may be more cautious about approving treatments based on surrogate outcomes, preferring to wait for more reliable evidence of their effectiveness and safety. This is particularly true for chronic conditions, where the long-term benefits and risks of treatment are uncertain, and where patients may be exposed to potential harm from ineffective treatments (67).

Ultimately, society's preference for the use of surrogate outcomes depends on the balance between the potential benefits of early access to treatments and the potential risks of approving ineffective treatments. Consequently, it is crucial that reimbursement decisions consider the reliability of surrogate outcomes in estimating the treatment effect on the final outcome (68).

1.3.1.2. Surrogate outcomes in cancer

The use of surrogate endpoints is increasing, particularly as the primary outcome measure in oncology trials. A review of randomised clinical trials (RCTs) in oncology conducted between 1995 and 2009 reveals that while OS was the primary endpoint in 49% of trials between 1995 and 2004, this proportion declined to 36% in the period 2005–2009 (55). Although response rates decreased from 14% to 6%, the use of time-to-event endpoints such as progression-free survival (PFS) gained ground, rising from 26% to 43% of primary outcomes.

Surrogate outcomes have been widely used in cancer clinical trials to evaluate treatment efficacy. In cancer research, surrogate endpoints can be used to predict a clinical outcome such as OS or disease-specific survival. These surrogate outcomes are often used to estimate the effects of a new treatment or intervention, especially when measuring the final outcome such as OS or PFS would take a long time or be too expensive (68).

1.3.2. Validation of surrogate outcomes

The risk of using intermediate outcomes that do not meet the scientific criteria defining a reliable surrogate outcome to inform regulatory and clinical decisions can be illustrated by the example of encainide and flecainide. These antiarrhythmic drugs were approved by the US Food and Drug Administration (FDA) based on epidemiological studies that found ventricular arrhythmia to be associated with a higher risk of death after myocardial infarction (69). Two years after the drugs came onto the market, the results from the Cardiac Arrhythmia Suppression Trial showed that their use was

associated with 2.5 times higher mortality in patients with asymptomatic ventricular arrhythmia after myocardial infarction (70).

Prentice first described the four criteria[2] a validated surrogate outcome needs to satisfy (71). Since Prentice's approach, which is based on hypothesis testing, more practical methods that allow the use of surrogate outcomes as prediction tools have been developed. Freedman (72) and, later, Buyse and Molenberghs (73) stress the importance of estimating and predicting by proposing different concepts: one is the 'proportion explained' concept, which examines what proportion of the surrogate outcome explains the final outcome; and the other is the 'relative effect' concept, which aims to explain how the treatment effect observed in the surrogate is reflected in the final outcome in relative effect terms. Buyse and Molenberghs' method provides a feasible approach for validating the surrogacy relationship, but it has limitations as it relies on data from a single trial, making it susceptible to bias, lacking generalisability and difficult to interpret. In response, the meta-analytical validation method (64) was developed, which incorporates the ideas of Buyse and Molenberghs by considering multiple randomised clinical trials on different interventions. This new method offers several methodological and technical advantages over previous methods. The results obtained through the meta-analytical method implicitly take into account the heterogeneity observed when multiple trials in different patient populations, clinical settings and study designs are combined in order to establish a common result (64). In this approach, the association between surrogate and final outcome is measured at the trial level while the relative effect of the overall treatment effect on the final outcome over the surrogate outcome is measured among trials at a population level. For validation purposes, it is crucial that the treatment effect observed on the surrogate outcome can accurately predict the treatment effect on the final outcome, particularly at the trial level, where the link between surrogate and final outcome is agnostic of the type of therapy received. In the meta-analytical method, a large coefficient of determination at the individual level, represented by R^2_{indiv}, close to one, shows that the surrogate and final outcome are likely to be linked. On the other hand, a high population level determination coefficient, represented by R^2_{popul} close to one, implies that the surrogate explains a

[2]Prentice criteria
f(S|Z) ≠ f(S), treatment has a significant impact on the surrogate outcome,
f(T|Z) ≠ f(T), treatment has a significant impact on the final outcome,
f(T|S) ≠ f(T), the surrogate outcome has a significant impact on the final outcome,
f(T|S,Z) = f(T|S). the full effect of treatment upon the final outcome is captured by the surrogate.
Where T and S are random variables that denote the true and surrogate endpoints, respectively, and Z is a binary indicator variable for treatment.

large proportion of the treatment effect on the final outcome. The meta-analytical method is explored further in Chapter 3.

1.3.3. Surrogate outcomes in reimbursement decision-making

The Pharmaceutical Benefits Advisory Committee (PBAC) is an Australian Government body responsible for providing recommendations to the Minister of Health on the listing and pricing of medicines on the Pharmaceutical Benefits Scheme (PBS). The PBS provides subsidised access to essential medicines for Australian residents, and the role of the PBAC is to ensure that medicines listed on the scheme are safe, effective, and provide value for money to the Government and taxpayers.

The PBAC bases its recommendations on evidence-based assessments of the clinical effectiveness and cost-effectiveness of medicines. Once a medicine is listed on the PBS, patients can access it at a reduced cost, with the Government covering the bulk of the cost of the medicine.

In Australia, the PBAC has published guidelines on the conduct of HTA and economic evaluations for reimbursement purposes. Within these guidelines, the PBAC has outlined how to assess the use of surrogate outcomes in HTA for those interventions seeking subsidy when data on the comparative treatment effect on a patient relevant final outcome from randomised clinical trials is not available. The PBAC has provided a detailed process – or validation framework – enabling the appropriate incorporation of surrogate outcomes into economic modelling (74) (Chapter 3). The PBAC validation framework aims to ensure that all aspects regarding the use of surrogate outcomes in economic evaluations are considered. It not only covers statistical considerations, but also takes into account the consequences and uncertainties generated by the use of a surrogate outcome in HTA.

Using surrogate outcomes in economic evaluation requires estimating the extent of long-term benefits relevant to patients through statistical transformations from the observed effect on surrogate outcomes to the effect on final outcomes. Surrogate outcomes may not capture the full spectrum of the safety profile associated with the intervention, and important long-term effects may be missed or underestimated. For example, a treatment that improves a surrogate outcome in the short term may have negative long-term effects that are not captured by the surrogate outcome, such as increased risk of side effects, drug resistance, or disease progression. This results in inherent uncertainty that could potentially impact the reliability of an economic model. Hence, it is important to evaluate how sensitive the results from an economic model are to the estimated benefit over the final outcome. Sensitivity analyses can help to ensure that the results of the economic evaluation are robust and reliable (see Chapter 4).

1.3.4. The decision-maker

Clinical decisions, registration decisions, and reimbursement decisions are critical components of healthcare delivery. Each of these decisions is made at different levels of the healthcare system, often by different individuals with distinct perspectives and goals.

At the individual level, clinical decisions are made by healthcare professionals and patients, who use their medical expertise and experience to diagnose, treat and manage patients' health problems. These decisions are based on evidence-based medicine, the most up-to-date research or clinical experience treating similar patients, and the individual patient's needs and preferences. The ultimate goal of clinical decisions is to improve patient health outcomes and quality of life (75)

On the other hand, registration decisions are made at the population level. Regulatory agencies and public health organisations review and evaluate new interventions in healthcare before deciding whether they can be made available to treat patients. The goal of these decisions is to ensure the safety and efficacy of healthcare interventions before they can be used by healthcare professionals to treat patients. These decisions are based on the results of clinical trials, which provide evidence of the effectiveness and safety of the treatment in question.

Reimbursement decisions, meanwhile, are also made at the population level. Government agencies, insurance companies, and healthcare organisations determine the cost of healthcare interventions and decide which treatments will be reimbursed. These decisions are based on the treatment's cost-effectiveness, its impact on population health, and on the healthcare budget, among other considerations (76). Reimbursement decisions must balance the need for affordable and accessible healthcare.

This thesis is firmly focused on reimbursement decisions.

1.4 Economic evaluation

Economic evaluations provide a comparative analysis of interventions in a particular clinical setting considering all the associated costs incurred and benefits realised. The primary objective of undertaking an economic evaluation within HTA, is to provide decision-makers with relevant information to support an efficient and equitable decision-making process. To that end, economic evaluation often relies on models to simulate the course of a disease, the effects of an intervention, and the costs associated with different health outcomes over time.

A model is a simplified representation of a complex reality which aims to provide reimbursement decision-makers with relevant information to support an efficient decision-making process. When used in health, a model can illustrate the natural history of a disease and show how the disease is affected by treatments, and can provide a comparative analysis of all interventions of interest by considering all the associated costs incurred and benefits accrued (76). Improvements in computational tools have allowed the use of increasingly complex models over time (for example, multistate and microsimulation models) (77); however, the models most widely used are decision trees, state transition and partition survival models (PSM); the latter is commonly used in economic evaluation of cancer interventions (78) (79).

The choice of model depends on the data available and the type of events to be captured. Decision trees do not usually incorporate a time variable and are therefore useful for modelling acute conditions or unique events in a patient's life. On the other hand, state transition models and PSM allow time to be included in the form of regular cycles, thus better representing health conditions that appear as repeated episodes or are chronic in nature. Chronic conditions can be explained by a series of events across time where individuals move between different health states at a certain speed. State transition models are able to represent these conditions by allowing model state membership or transition between the health states at a certain rate. The rate of movement between health states is known as the transition probability (76).

By inputting relevant data, such as transition probabilities and cost of treatment, economic models in healthcare can provide insights into the potential outcomes of new interventions. This approach can help to better understand the economic impact of diseases, allowing policymakers and healthcare professionals to make informed decisions and allocate resources effectively.

The main result of an economic model is expressed as an incremental cost-effectiveness ratio (ICER), which illustrates the incremental differences between the alternative interventions in terms of costs (in the numerator) and health benefits (in the denominator). The ICER is interpreted as the additional cost per patient that must be spent to obtain an additional unit of benefit (76). Furthermore, sensitivity analyses can assess the robustness of the model by testing the uncertainty within the structure of the model and parameters inputs. This includes analyses that can provide an estimate of the probability associated with a result being cost-effective (80).

1.4.1.1. Economic Evaluation and Health Technology Assessment

An increase in healthcare spending is a common factor in most countries around the world. The large number of new technologies available, the aging of the population and the high prices associated with some emerging therapeutic options are three of the main factors that have driven the rise in healthcare expenditure. In this context, health systems should support their decisions by basing them on tools that provide information regarding the efficient allocation of scarce resources. This implies ensuring not only that new technologies will produce health in those groups with the capacity to benefit (patients in need of these technologies), but also that the benefits they provide outweigh the benefits that would be obtained if the resources were allocated to other activities in the health sector. In economics, this is defined as the opportunity cost, a core element in economic evaluation in healthcare (76).

Furthermore, the results of an economic evaluation focus not only on the average estimate of cost-effectiveness, but also on the sources of uncertainty around the parameters used and how this uncertainty affects the results. Cost-effectiveness analyses, in this context, can provide an estimation of the probability of cost-effective results when compared with information regarding what society is willing to pay for an extra unit of benefit (80).

There is a broad consensus that health outcomes in economic evaluations should integrate both quantity and quality of life using a metric capable of comparing the result across interventions (76). One of the reasons for the popularity of the generic outcome measure quality adjusted life years (QALYs) in economic evaluations is the fact that it integrates both key elements needed to compare interventions across different diseases. In the QALY, the amount of time spent in a health state is weighted by the quality of life associated with that particular health state, valued using utility weights (81). Decision-makers are then able to compare across a wide range of health technologies by comparing the results of economic evaluations (ICER) expressed in terms of cost per additional QALY gained.

HTA is a multidisciplinary process used to evaluate a broad spectrum of evidence including clinical, social, economic, organisational, and ethical consequences derived from the adoption of health interventions in a health system. Its main goal is to inform decisions about resource allocation and coverage of new health technologies given the actual constraints of the system and taking into consideration a broad set of social values. The rationale behind HTA is that the use of the limited resources of the health system requires complex judgements to achieve the best possible allocation.

Hence, in the context of budgetary restrictions, economic evaluation in HTA has become the tool to ensure efficiency (76).

There are three HTA expert advisory committees in Australia providing the federal government with advice on the reimbursement of medical services (MSAC), protheses (PLAC) and pharmaceuticals (PBAC)[3] respectively (82). The PBAC is an independent body of experts including clinicians, health professionals, health economists, and consumer representatives that evaluates the medical conditions, clinical effectiveness, safety, and cost-effectiveness of pharmaceuticals compared with other treatment options for the same population. The main role of the PBAC is to advise the Health Minister on the suitability of medications for public funding in Australia (83). Since the use of surrogate outcomes may influence the results of an economic model, the PBAC has published guidelines providing detailed instructions on how to best use surrogate outcomes to inform an economic evaluation (74). The methodology and results of the validation of metastasis-free survival (MFS) conducted by the Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP) team, contextualised by the Australian framework for the use of surrogate outcomes in economic evaluations, are discussed in detail in Chapter 3.

Within the constraints of limited financial resources, HTA uses economic evaluations as a practical tool to ensure efficiency. These evaluations rely heavily on the clinical outcomes reported by clinical trials, which must be both reliable and promptly available. However, as evidenced by the ICECaP study (section 1.5.5), the time required for clinical trials to demonstrate treatment effectiveness can be long and has an impact over the timelines for the regulatory and reimbursement decision.

Approving a therapy based on early evidence that later proves the treatment effect was inaccurately estimated, can have negative consequences. For example, as noted above, encainide and flecainide were approved by the FDA for the treatment of myocardial infarction based on surrogate outcomes but were later withdrawn from the market due to increased mortality. Similarly, bevacizumab was initially approved for HER2-negative breast cancer based on improved progression-free survival (PFS),

[3] The MSAC (Medical Services Advisory Committee) is an independent committee that advises the Australian government on the safety, effectiveness, and cost-effectiveness of medical services and procedures, and whether they should be publicly funded.
The PLAC (Prostheses List Advisory Committee) is an independent committee that advises the Minister for Health on the reimbursement of prostheses under the Prostheses List.
The PBAC (Pharmaceutical Benefits Advisory Committee) is an independent committee responsible for recommending which medicines should be subsidised by the Australian Government under the Pharmaceutical Benefits Scheme (PBS). It assesses the clinical and cost-effectiveness of medicines and makes recommendations to the Australian Government about their suitability for listing on the PBS.

but subsequent trials showed it did not improve overall survival and was associated with increased toxicity, leading to revocation of its approval for that indication. However, delaying access to a potentially effective and efficient health technology could also have a negative impact on society, incurring an opportunity cost by depriving patients of timely and appropriate treatment, the magnitude of which remains uncertain.

An alternative to waiting for the final outcome of a clinical trial to become available is to replace it by another outcome, which might be measured earlier, more conveniently or more frequently. This will certainly have a clinical impact on patients but is also likely to have an impact in terms of costs and benefits, which will ultimately impact on the efficiency of the allocation of public resources. If a valid replacement outcome is used to inform economic evaluations in HTA, patients and society will obtain benefits from a potentially effective intervention earlier, thus potentially decreasing the morbidity and mortality associated with later stages of a diseases.

Although surrogate outcomes may be useful in assessing treatment effectiveness, there is still uncertainty about the level of reliability required in relation to their prediction to fully incorporate them into economic models with confidence. The validation and assessment of surrogate outcomes for economic evaluations is critical in this context. Nevertheless, careful evaluation and consideration on a case-by-case basis are necessary before their implementation in economic models to inform reimbursement decisions.

1.5 Prostate Cancer

In prostate cancer clinical trials, the clinical endpoint of choice has traditionally been survival, but the natural history of localised prostate cancer (LPC) is usually long and indolent. In this population, death caused by prostate cancer is difficult to observe and most patients die of other causes (84). A review of published randomised prostate cancer trials estimated that 11.5 years were needed before publication to demonstrate significant survival results (85). Considering the prolonged natural history of the disease and the aging demography of Australia, it is estimated an increase in the number of men[4] with prostate cancer over the coming decade (86). Moreover, considering the various treatment options at all stages of the disease and the survival benefits of early diagnosis, the economic burden of prostate cancer is also expected to increase (87). Thus, early-stage prostate cancer is a disease where

[4] Throughout this thesis, men refer to people with a prostate.

a surrogate outcome may have an essential role in ensuring patients are provided with more effective treatments earlier.

1.5.1. The health problem

The prostate is an exocrine gland that surrounds the initial portion of the male urethra. Its principal function is to secrete a fluid that contributes to semen production and helps with the proper function of the sperm cells. This gland has a high degree of oncogenic potential that may result in abnormal growth of neoplastic tissue causing prostate cancer, which, in most cases, is driven by endocrine androgen hormones (88).

Prostate cancer is the most common incident cancer in men globally, with 1.3 million new cases diagnosed each year, and a leading cause of death among males causing nearly 400,000 deaths in 2018 (89). The region of Australia and New Zealand shows the highest incidence in the world (90) and Australia was expected to reach 24,217 new cases in 2022, with 3,507 deaths in the same year (91).

Prostate cancer is more common in men at older ages, with a median age at diagnosis in Australia of approximately 68 years (92). When diagnosed, prostate cancer can be localised to within the gland, invade regional tissues like seminal vesicles or regional lymph nodes, or metastasise to distant organs (generally the bone) (93). Of those newly diagnosed with prostate cancer in Australia, more than 90% have cancer confined to the prostate gland (localised disease). Metastatic and loco regional disease are less frequent at 6% and 4% respectively at the time of diagnosis (92).

Localised prostate cancer has typically a slow growth rate. Men who receive no treatment for their primary tumour have a 49% risk of local progression 15 years after diagnosis, while those who have their prostate removed have a 22% risk of progression over the same period (94). When the disease progresses, haematuria (the presence of blood in urine), obstructive voiding symptoms (urinary hesitancy, diminished stream, straining, incomplete emptying, interruption of the urinary stream, and dribbling), bowel problems, pain, erectile dysfunction and incontinence are common symptoms. The morbidity of local progression together with the risk of metastasis and cancer specific death negatively impacts quality of life while new treatment pathways aim to delay and avoid diseases progression (95, 96). Although prostate cancer has a low mortality rate compared with other types of cancer (90), metastatic disease has a significant impact on cancer specific survival and most deaths from prostate cancer occur at this stage of the disease (97).

As with the incidence rate, prostate cancer mortality increases with age; nearly 90% of all cancer-related deaths occur after age 70 (91). There are several factors that influence the prognosis of prostate cancer. Age, progression and the presence of metastasis are of most relevance when assessing proper treatment and prognosis (98); however, early detection and treatment of asymptomatic men has the potential to improve life expectancy by reducing the probability of progression (99).

In summary, prostate cancer is a common and potentially lethal cancer in men, with a high incidence in Australia. The cancer is driven by endocrine androgen hormones and can lead to symptoms disturbing quality of life such as haematuria, urinary problems and erectile dysfunction. The risk of local progression increases over time with cancer-specific deaths occurring most frequently in older men with metastasis. Early detection and treatment of asymptomatic men can improve life expectancy and reduce the probability of progression.

Understanding treatment pathways is important in the context of HTA because it allows for a comprehensive evaluation of the impact of a new technology or intervention on patient outcomes and healthcare resource utilisation. Treatment pathways help to identify the sequence of interventions or procedures that patients typically receive during their care, including the use of different drugs, diagnostic tests and medical procedures. By understanding treatment pathways, health technology assessors can assess the potential impact of a new technology or intervention on the existing treatment pathway and how it may alter patient outcomes and resource utilisation. This information is crucial in determining the clinical and economic benefits of a new technology or intervention and can inform decision-making about its adoption and reimbursement.

1.5.2. Diagnosis and risk classification of prostate cancer

Digital rectal examination and a blood test to measure the PSA level are commonly used to detect early stages of prostate cancer. The Australian PSA testing guideline recommends that testing should commence at 45 years of age for men with a family history of prostate cancer, and otherwise every 2 years from age 50 to age 69 (100). While many prostate cancer cases are diagnosed based on high levels of PSA (PSA > 4 ng/mL), the test's reliability is low (100). Despite being synthesised specifically in the prostate, PSA is a protein produced under normal circumstances as well as in the presence of prostate cancer. In addition, the test can be abnormal due to benign hypertrophy of the prostate, urinary infections or biopsy procedures and the acceptable level of PSA in healthy patients increases with age (101). At the current threshold level of PSA > 4 ng/mL, the PSA test can miss nearly 82% of

young patients who have prostate cancer (102). Clinical guidelines deal with the limitations in sensitivity and specificity of the PSA test in clinical practice by recommending prostate biopsies to confirm abnormal results (100).

Risk assessment of patients diagnosed with prostate cancer is a critical step before recommending a specific course of treatment. Biopsy results, the differentiation status of the prostate gland cells (Gleason Score), clinical staging (in terms of size and location of the tumour, lymph node involvement and metastasis assessment [TNM]) and the level of PSA in the blood are all used to determine the best treatment option and stratify patients into risk groups (103). Altogether, the results of these diagnostic tests may predict the probability of biochemical failure[5] after administering definitive local therapy and, based on this, patients can be classified as (104):

- Low risk:
    - PSA < 10 ng/mL and,
    - Gleason score < 7 and,
    - the tumour may have evolved to cover one-half of one side of the prostate.

- Intermediate risk
    - PSA is between 10 and 20ng/mL or,
    - Gleason score = 7 or,
    - the tumour may has evolved to cover more than one-half of one side of the prostate but not both sides.

- High risk
    - PSA > 20 or,
    - Gleason score > 7 or,
    - the tumour has grown into both sides of the prostate.

Risk group stratification is an important step in the diagnosis of prostate cancer as it guides the design of clinical trials, helps reporting comparable outcomes and ultimately, assists physicians and patients in the choice of treatment (105). Patients diagnosed with high-risk diseases have an increased risk of recurrence and, therefore, more aggressive therapies are recommended. Potent hormonal drugs such as abiraterone, usually reserved for advanced stages of the disease, can play an important role as adjuvant treatments to improve the long-term prognosis of these patients (106).

[5] Biochemical failure is defined as a rise in the level of prostate-specific antigen (PSA) in the blood after primary treatment for prostate cancer. Is generally defined as a rise in PSA level of a certain amount or percentage above the lowest post-treatment level.

1.5.3. Treatment of prostate cancer

The therapeutic approach for prostate cancer patients is documented in Chapter 4; however, for the purpose of this chapter, a succinct overview will be provided.

Treatment options for prostate cancer patients depend on various factors, including disease assessment, morbidity, patient preferences and resource availability. For low-risk LPC patients, active surveillance is the most common management option and has shown benefits for those with a low probability of developing advanced stages of the disease (107). Intermediate and high-risk LPC patients, on the other hand, typically receive active treatment. The available options include prostatectomy with/without radiotherapy and radiotherapy with/without androgen deprivation therapy (ADT) (108).

In terms of radiotherapy, recent advances have led to more effective delivery methods, such as the integration of imaging and highly conformal radiotherapy techniques (109). The current clinical guidelines recommend external beam radiation therapy (EBRT) or brachytherapy for treatment of the primary tumour. EBRT utilises an external X-ray source and techniques such as intensity-modulated radiotherapy (IMRT) to deliver high doses of radiation to the target while minimising toxicity. Brachytherapy, on the other hand, involves the implantation of radioactive seeds or capsules into the prostate, allowing for extreme dose escalation while minimising radiation exposure to surrounding tissue. Surgical techniques have also evolved, with the advent of laparoscopic and robot-assisted surgery, but with less clear improvements in clinical outcomes compared to radiotherapy (108).

Patients diagnosed with high-risk LPC have an increased risk of recurrence and are typically recommended hormonal therapies added to primary treatment, such as abiraterone. ADT, also known as hormone therapy, is a commonly used systemic treatment for advanced prostate cancer. This therapy aims to reduce the level of androgen hormones by either preventing androgen hormone secretion or blocking the androgen hormone receptors in cancer cells (110). To alleviate the adverse effects of low androgen levels associated with hormonal therapy, intermittent ADT, which involves alternating periods of treatment and off-treatment, is often prescribed. However, the evidence regarding the effectiveness of this approach is variable, and there is no consensus on the preferred approach (111). Most patients with hormone-sensitive disease will respond well to ADT but eventually develop resistance to castration levels of androgen hormones, known as castration-resistant or hormone-insensitive prostate cancer (112). In such cases, ADT is usually maintained continuously (113).

New antiandrogen treatments (abiraterone, apalutamide, darolutamide, and enzalutamide) have been tested and recently recommended in clinical guidelines (114). These treatments have shown promising results in treating advanced prostate cancer; however, they have not yet been listed on the PBS for non-metastatic patients in Australia. In the economic model developed here, abiraterone was employed as an example in the primary treatment regimen in combination with radiotherapy.

The management of advanced prostate cancer has undergone significant changes in recent years, with new systemic treatments being recommended for patients with distant organ involvement. The primary treatment for such patients has been secondary hormonal manipulation with antiandrogens and corticosteroids (115). However, the addition of docetaxel to ADT has produced significant improvements in OS and failure-free survival (116). Despite this, patients may develop resistance to docetaxel and ADT, necessitating more aggressive treatment options. These include additional chemotherapy, second-generation antiandrogens, radium-223 for symptomatic palliation, immunotherapy with Sipuleucel-T, inhibitors of the antiandrogen precursor synthesis and bone-targeted therapy.

1.5.4. Motivating example

The choice of prostate cancer as a focus was driven by the prolonged natural history of the disease and the aging population in Australia, with projections indicating an increase in the number of men with prostate cancer over the coming decade (86). Moreover, considering the various treatment options at all stages of the disease and the survival benefits of early diagnosis due to the PSA surveillance program, the economic burden of prostate cancer is expected to increase (87). Given that LPC is the most prevalent stage of the disease, an economic evaluation focusing on these patients will potentially encompass the most relevant population in terms of overall health spending on this condition.

The treatment alternatives for patients with LPC have been studied widely (117-122). In general, treatment will depend on the risk assessment (staging) of each patient, life expectancy and performance status. In low-risk patients or in patients who are not willing to receive active treatment, evidence supports the use of active surveillance. For those diagnosed as high risk and who are willing to receive treatment, treatment options are limited and usually compromise quality of life (123).

Current clinical guidelines reserve the use of abiraterone for metastatic stages of the diseases (108) (124); however, recent studies have found that AAP added to RT can improve quality and PFS (125, 126) The use of treatments traditionally reserved for more advanced stages of disease in patients with

non-metastatic conditions is a topic of debate among healthcare professionals and can partly be addressed through the implementation of an economic evaluation. Therefore, the model built here compared AAP added to RT to the standard of care (SoC, consisting of RT plus ADT) for patients newly diagnosed with high-risk LPC and who are willing to receive an active treatment for their primary tumour according to Australian settings.

1.5.5. Surrogate outcomes and prostate cancer

The long natural history of prostate cancer has motivated research into surrogate outcomes that could predict survival and bring new products to market early. Biochemical relapse using different threshold values of PSA, event-free survival, MFS, and disease-free survival are among the surrogates most frequently explored (127-131). While most of the outcomes mentioned above have scientific evidence supporting their relationship with the final outcome for assessing surrogacy, the majority of them rely on data from single studies, which hinders interpretation and applicability.

The lack of a validated and robust surrogate outcome in LPC motivated the work of the Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP) working group. The objectives of this group are to reduce the time needed to conduct clinical trials and to accelerate the development of new treatments to be able to treat patients at the earliest stage possible of the disease (85). The work of the ICECaP working group has involved the use of individual patient data (IPD) from randomised controlled clinical trials conducted in countries with highly regulated health authorities (Canada, the United Kingdom, Europe, Australia/New Zealand, and the United States). The ICECaP working group employed a meta-analytic approach to examine the suitability of MFS as a surrogate outcome for OS in LPC, and the findings showed good results (132). Chapter 3 covers the details of the ICECaP study and how those data have informed the research explored in this thesis.

Considering the significance of health economics in the context of reimbursement approvals, a key aim of the ICECaP working group was to evaluate how surrogate outcomes could be used to support the health economic evidence. While surrogate outcomes can expedite the timeframe within which a technology is submitted for reimbursement consideration, and potentially facilitate early access to novel interventions, it is crucial to consider the trade-offs between speed and accuracy of approval decisions. This thesis will help to shed light on the aforementioned conundrum.

1.5.5.1. Chain surrogacy

Chain surrogacy involves using multiple surrogate outcomes sequentially to predict a clinical outcome. This could provide a more accurate assessment of a treatment's effect by capturing different aspects of disease progression and response.

For example, a treatment's impact on prostate-specific antigen (PSA) levels could be used to predict its effect on metastasis-free survival (MFS), which in turn predicts overall survival (OS). By modelling the relationships between the successive surrogates in a statistical "chain", the thresholds each surrogate must meet based on the prior surrogate's effect can be determined. This chained approach might allow for uncertainty in one surrogate's prediction to be reduced by incorporating information from another surrogate (133). In situations where a single surrogate is not ideal, this chain surrogacy approach can offer a way to strengthen predictions by synthesising information from complementary outcomes measured at different stages of disease and response to treatment. Additionally, considering multiple links in the chain may give a fuller picture of how a treatment works at various stages (134).

However, chain surrogacy can also increase the complexity of evaluation. Incorporating multiple surrogate relationships requires more data and advanced statistical modelling. This places greater demands on data collection and analysis. Additionally, inconsistencies could arise as different surrogate outcomes may respond varying amounts or speeds to a treatment. This could lead to delays between the surrogates and accurately predicting the final clinical outcome.

While chain surrogacy has potential benefits for understanding disease mechanisms, limited evidence currently exists for its application to prostate cancer. Further research is still needed to refine methods for incorporating chain surrogacy into economic evaluations and assessing its impact on uncertainty. Statisticians and health economists continue to explore this challenging area.

1.6 Aims and Objectives

This research aims to investigate the potential to use MFS as a surrogate outcome for estimating the economic value of a new intervention to treat localised disease in prostate cancer.

The specific objectives of the project are:

1. To analyse and summarise the published evidence for surrogate outcomes in economic evaluations of localised prostate cancer.

2. To assess the use of MFS as a surrogate outcome for OS in economic evaluations of HTA according to the PBAC guidelines.

3. To assess the feasibility and challenges of incorporating MFS as a surrogate outcome of OS in an economic model for LPC.

4. To assess the benefits to society of bringing a product to clinical practice early based on a validated surrogate outcome compared to waiting for mature OS data.

5. To determine the value associated with the use of surrogate outcomes and the trade-offs between the ability to treat patients sooner through reimbursement decisions based on the surrogate outcome and waiting for longer-term evidence on survival.

This thesis will provide healthcare decision-makers with the means to evaluate the use of surrogate outcomes in economic evaluations to introduce new treatments into practice for localised disease in prostate cancer. As the use of surrogate outcomes impact on the duration of the studies and research-related costs, the results will also be of significance to those involved in the design of clinical trials who may consider designs of clinical trials that involve outcomes aligned with those relevant to reimbursement decision-makers. Likewise, health economics researchers and academics will have the opportunity to use a novel approach when dealing with economic models that need to be informed by outcomes that imply a long time of follow-up in LPC or find motivation to investigate on similar population.

1.7 Overview of the research

This thesis will evaluate the use of MFS as a surrogate outcome in economic evaluations when introducing new interventions into clinical practice for LPC. The objectives proposed in this research will be developed through six chapters:

Chapter 2 will summarise the findings of a systematic literature review of the use of surrogate outcomes in economic evaluations of localised prostate cancer. The analysis will be focused on the identification of the surrogate and final outcome used in each study, the methods used to correlate both outcomes and the impact of the use of the surrogate on the results.

Chapter 3 implements the assessment of the proposed surrogate outcome based on the method suggested in the PBAC guidelines. The PBAC guidelines for the conduct of economic evaluations have

emphasised the need for a validation process for the use of surrogate outcomes before incorporating them in economic modelling when the comparative treatment effect of a new intervention seeking subsidy on a final outcome relevant to patients measured in randomised clinical trials is not available. As MFS will be used as a surrogate for OS for patients diagnosed with LPC in the proposed economic evaluation of new interventions submitted to PBAC, an assessment according to the guidelines is conducted.

Chapter 4 builds an economic evaluation incorporating the surrogacy relationship found by ICECaP to compare new intervention versus a relevant comparator to treat LPC. The results of the economic evaluation will be used to assess the potential benefits to society of making a new intervention available for reimbursement early based on the validated surrogate endpoint (MFS) compared to waiting for mature OS data. This chapter will also focus on the consequences to the healthcare system, and society as a whole, of the cost and benefits of bringing a product to clinical practice early based on the use of surrogate outcomes.

Chapter 5 reports the results of an analysis aimed at understanding societal values and preferences associated with evaluating new interventions which rely on evidence of surrogate outcomes. Within the chapter, a discrete choice experiment (DCE) is used to evaluate the trade-offs between uncertainty and time associated with a medicine becoming available, in the context where members of the general population were asked to take the role of advisors to a reimbursement decision-maker.

Chapter 6 summarises the results of the thesis and sets out the health and policy implications of the implementation of surrogate outcomes to inform economic evaluations supporting reimbursement decision-making, not just for interventions in LPC but beyond.

Chapter 2 Systematic review of economic evaluations incorporating surrogate outcome measures

2.1 Abstract

Health Technology Assessment (HTA) is widely used to inform reimbursement decisions about which interventions should be made available to the general public. Economic evaluations are an integral part of HTAs and facilitate comparisons between interventions in terms of cost and outcomes. However, using final outcomes in economic evaluations for localised prostate cancer (LPC) treatments can be challenging due to the long natural history of the disease. Therefore, surrogate outcomes are commonly used instead to expedite reimbursement decisions. While the use of surrogate outcomes can facilitate faster decision-making, it can also increase uncertainty in the results of economic evaluations. This systematic review aims to assess the use of surrogate outcomes in economic evaluations of LPC treatments.

For that purpose, a systematic literature search was conducted in March 2019 and updated in June 2022 using the PICO methodology to frame the research question and guide the literature search strategy. The search included PubMed, EMBASE, EconLit, NHSEDD, and Cochrane databases. From the 987 publications identified in March 2019 and 110 new publications identified in June 2022, 52 publications were assessed for full text review. Finally, only one study met the inclusion criteria (economic evaluation incorporating a surrogate outcome in LPC to estimate the treatment effect on a final outcome).

The study conducted by Lundkvist et al., used a surrogate outcome (PSA free survival) to predict changes in overall survival and inform an economic evaluation in LPC. However, the limitations of this study (mainly, regarding differences in how the surrogate was defined and lack of validation and sensitivity analysis testing the surrogacy relationship) make it difficult to draw firm conclusions on the use of surrogate outcomes in LPC. The findings of this review indicate a lack of strong published evidence supporting the use of surrogate outcomes in economic evaluations of LPC.

2.2 Introduction

Health technology assessment is commonly used to inform decisions about which intervention could be reimbursed in any health system, especially within the context of systems financed by public health insurance (135) . Around 80% of countries have implemented HTA processes to help with the decision-making process in healthcare and more than 60% have established institutions or organisations to

support recommendations regarding the use and subsidisation of health technologies to the corresponding Ministry of Health (136). To facilitate efficient resource allocation in healthcare for reimbursement decisions, economic evaluations are usually incorporated as part of HTAs. Through the use of models that simulate specific health conditions, economic evaluations provide the means to facilitates the comparison of different interventions for various diseases competing for scarce resources (76).

Ideally, models used in economic evaluations for HTA in the context of reimbursement decision-making should incorporate final outcomes relevant to patients. However, this poses a challenge in LPC. Prostate cancer has a long natural history; thus, trials focused on this disease need a long period of follow-up to report significant results in terms of overall survival (OS). The long lead time to report on OS in clinical trials in LPC has led to the use of surrogate outcomes (i.e., the ICECaP initiative discussed in Chapter 3), aiming to expedite the reimbursement and regulatory decision-making process regarding new interventions for LPC (137) (138).

The use of surrogate outcomes to inform economic models in reimbursement decision-making has the potential to allow interventions to be considered for reimbursement sooner, as the comparative results of surrogate outcomes in clinical trials are usually available before survival outcomes. However, the use of surrogate outcomes may increase uncertainty regarding the results of an economic evaluation and consequently in the reimbursement decision. Their use in HTA can result in two types of errors made when deciding for reimbursement of new interventions: a false positive conclusion when a new treatment has greater health benefits than risks when the opposite is true, and an overestimate of the actual benefit of a new treatment (bias). These errors could potentially lead to harmful consequences and, therefore, policy-makers and HTA analysts should aim to avoid them when possible (139).

2.3 Objectives

The primary objective of this systematic review was to assess the use of surrogate outcomes in economic evaluations of LPC treatments. The specific aims were to:

1) Identify the surrogate and final outcomes used in economic evaluations of LPC treatments
2) Identify and report the methods used to correlate surrogate and final outcomes
3) Identify and report the methods for handling the uncertainty inherent with surrogate outcomes in economic evaluations

2.4 Methods

2.4.1. Search strategy

A search of the published literature on economic evaluations incorporating the use of surrogate outcomes in prostate cancer was performed following the recommendations of Mathes et al. for systematic reviews of economic evaluations (140). The systematic review was conducted in March 2019 and updated in June 2022. Publications were searched for in the following databases:

- PubMed
- EMBASE
- EconLit
- NHSEDD
- Cochrane

The PICO methodology was used to frame the research question and guide the literature search strategy (Table 2.1).

Table 2.1. PICO framework

Component	Description
Patients	Patients diagnosed with localised or locally advanced prostate cancer.
Intervention	All interventions used to treat localised or locally advanced prostate cancer were eligible for inclusion (surgical procedures, chemotherapy, chemical treatment, and radiation therapy).
Comparator	Not defined in the search strategy as identifying the comparator was not the focus of the research question.
Outcome	Efficacy outcome: Any well documented and supported intermediate outcome used as a surrogate intended to substitute for a meaningful or final clinical endpoint.
	Economic evaluation: Any type of cost per benefit outcome (cost per QALY/DALY or unit of effectiveness that allows for comparison across different interventions.
Types of studies	Economic evaluations, (i.e., CEA, CBA, CUA and CMA [if the equal efficacy statement was based on the use of a documented surrogate outcome]).

Abbreviations: CEA= Cost-effectiveness analysis; CBA= cost-benefit analysis; CUA= cost-utility analysis; CMA= cost-minimisation analysis; DALY= disability-adjusted life year.

A set of keywords related to the PICO framework was used to develop the search strategy for each database. For this purpose, search terms were obtained by scrutinising a previous systematic review focused on surrogate outcomes (141) and by adapting the Pharmaceutical Benefits Advisory Committee (PBAC) framework for the use of surrogate endpoints. The search strategy developed for the PBAC framework (which contains the terms used in the systematic review conducted to inform the guideline) is found in Appendix C of the framework (74). It was used to define all terms related to 'surrogate outcomes' in the search strategy (see Appendix 1).

The core term in the search strategy merits specific consideration. The phrase 'surrogate outcome' has been widely used in the literature but given various definitions. Terms such as 'intermediate endpoint' or 'biomarker' are often used to refer to 'surrogate outcomes', leading to ambiguities and misinterpretations of what constitutes a surrogate outcome. Therefore, a broader list of terms which are often used interchangeably to describe the concept of a surrogate outcomes was employed in the search criteria to ensure it captured all possibly relevant literature:

- Surrogate end-point
- Intermediate clinical end-point
- Intermediate end-point
- Intermediate outcome
- Biomarker
- Surrogate outcome

To identify potentially relevant studies not captured by the search strategy, pearl referencing of potentially relevant articles was also conducted. In addition, a specific search of the grey literature was performed in Google and Google Scholar.

Studies were limited to those conducted in humans and all publications in languages other than English were excluded. No limits were placed on publication dates.

2.4.2. Exclusion criteria

The following type of publications were excluded from the review:

- Abstracts and posters, as there is insufficient information available for meaningful data extraction.
- Studies that only reported quality of life (QoL), patient-reported outcomes (PROs) or utilities without assessing the cost component of an economic evaluation.
- Reviews (non-systematic), editorials or commentaries.
- Studies focusing on benign hyperplasia of the prostate or a neoplasia other than prostate cancer.
- Studies where a relationship between the surrogate outcome and the final outcome was not identified within the publication.

- Economic evaluations using methodologies other than cost-effectiveness analysis, cost-benefit analysis, cost-utility analysis, or cost-minimization analysis as defined in Table 2.1. This may include studies utilizing alternative approaches to evaluate the costs and outcomes of healthcare interventions.

2.4.3. Review process

After removal of duplicates and checking for inclusion and exclusion criteria, screening using title and abstract was conducted by a single reviewer. If the use of surrogate outcomes was not reflected within the title or abstract of a publication, the full text was retrieved for further assessment. A consistency check[6] of the generated library of included and excluded studies was performed by a second reviewer (RDL).

2.4.4. Data Extraction

Data extraction was performed by a single reviewer based on the framework suggested by Mathes et al. (140). As the main objective of this review was to evaluate the use of surrogate outcomes in economic evaluations of LPC treatments, all elements within a publication related to the use of surrogate outcomes were of great importance and were extracted from included publications. This was facilitated by modifying the data extraction table suggested by Mathes et al. to include items associated with the role of surrogate outcomes in economic evaluations that were identified in the 'Surrogate to Final Outcome Framework' proposed in the PBAC guidelines (74). The data extracted from the identified study is described in Table 2.2.

[6] RDL conducted a consistency check to ensure that the inclusion/exclusion criteria were correctly applied.

Table 2.2. Data extracted from studies found.

Data Extracted	Based on Mathes et al.	Based on PBAC framework
General characteristics of methods and results reported	X	
Identification of the final outcome		X
Characterisation of surrogate outcome used in economic evaluation		X
Source of surrogate outcome evidence used in the economic evaluation (e.g., observational or randomised studies)		X
Evidence of validation of surrogate outcome		X
Methods used to quantify the link between surrogate outcome and final outcome (e.g., Regression-based approach)		X
Discussion of the uncertainty associated with the use of surrogate outcomes in the results or conclusions or elsewhere in the report		X

2.4.5. Quality assessment

A quality assessment was conducted to identify possible publication bias and limitations of the included study. The quality was assessed using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) (142, 143) and the Quality of Health Economic Studies tool (QHES) (142, 143). The quality assessment using both tools was conducted by the same reviewer.

The CHEERS checklist is a qualitative instrument with open-ended items recommended by researchers and peer review journals to assess the reporting practices of economic evaluations. As a qualitative tool, it does not provide a score to enable a quantitative comparison of the quality of studies assessed. Furthermore, it does not use weights or assign differential rankings based on the relative importance of each item. By not distinguishing these implicit differences, CHEERS assumes that all criteria are of equal importance (142), making it difficult to discriminate between economic evaluations of varying quality.

Accordingly, a quantitative checklist, the QHES, that provides an overall score (100-point scale) of the quality of economic evaluations identified was also used. The QHES consists of a set of 16 criteria informed by a literature review and selected by a steering committee comprising health economists, HTA experts, and healthcare decision-makers. Each criterion is assigned a weight according to their importance that ranges from 0 to 9 adding to 100 points for the 16 criteria. The instrument has a yes/no response for each criterion and assigns the full weight associated with the question when the response is positive or 0 when the response is negative, rather than a continuous scale for each criterion. Although the QHES tool has no interpretation standard for the scores obtained by each study, other users of the tool have set the interpretation scores as follow: 75–100 as good, 51–74 as

fair, 25–49 as poor, and 0–24 as extremely poor (144-146). These categories were used to interpret the scores of the studies found in this systematic review.

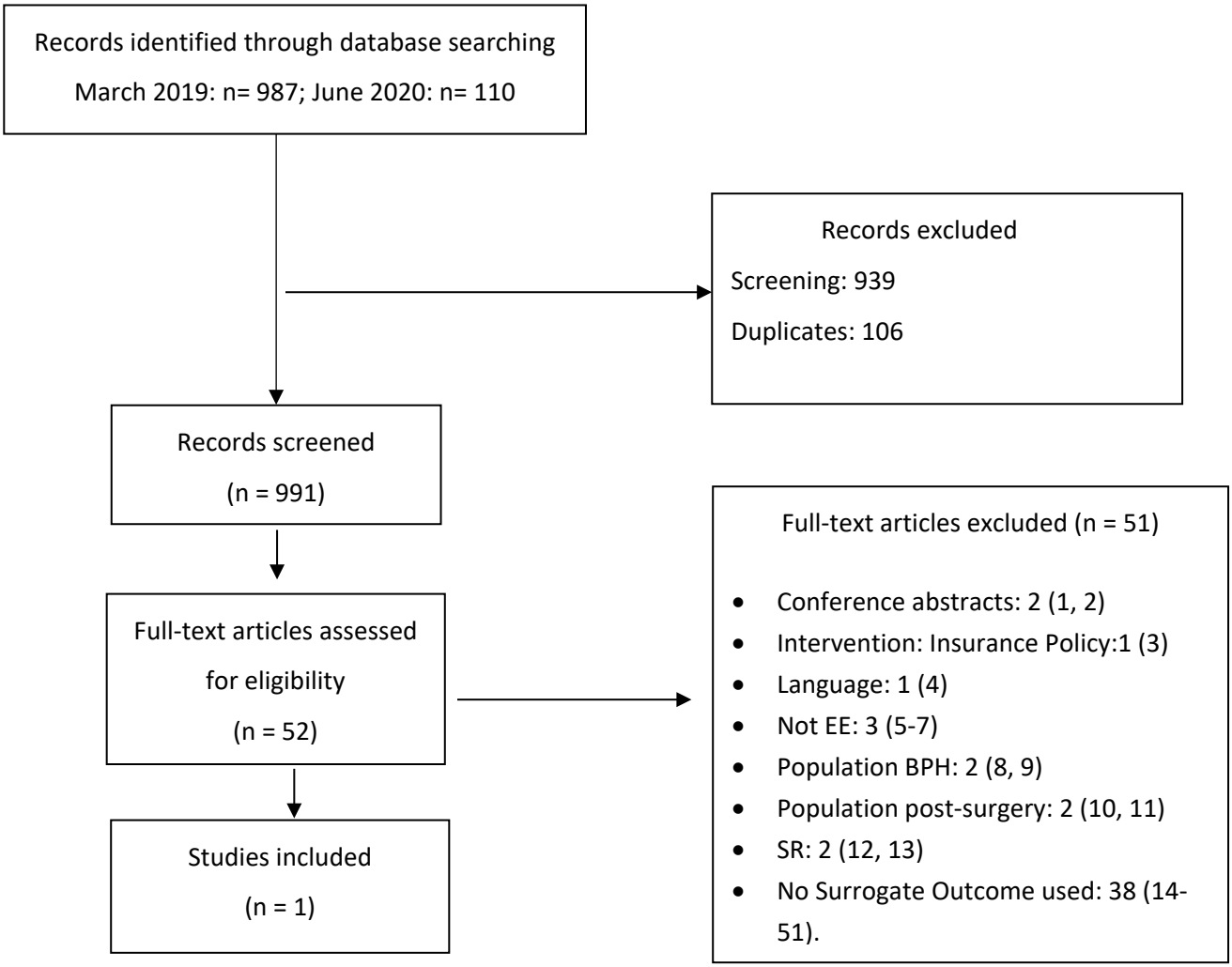
2.5 Results

Based on the inclusion and exclusion criteria, 987 publications were identified in March 2019 and 110 additional publications were identified during the update conducted in June 2022 (details of the results of the replicated literature search are not provided in this document). All results were imported into EndNote® for review. The full text of 52 publications was retrieved after abstracts and titles were screened and from these a total of 51 articles were excluded. Among these, 2 were conference abstracts, 1 study was related to insurance policy, and 1 was in a language other than English. Three studies were not economic evaluations, 2 studies were focused on benign prostate hyperplasia, and 2 studies were focused on post-surgery patients. Additionally, 2 studies were systematic reviews, and 38 studies did not use surrogate outcomes to predict treatment effect on the final outcome in their economic evaluation.

In eight studies, constituting 16% of the total, the presentation of cost-effectiveness analysis results was focused on intermediate outcomes. Within this subset of studies, the measurement of treatment effectiveness in terms of intermediate outcomes was primarily observed under two circumstances: either the assumption of equal efficacy between the intervention and comparator was made (16), or when the study was conducted within a relatively short time horizon (24, 33, 50).

One study, Lundkvist et al. (147) was deemed appropriate for inclusion. The PRISMA diagram shown in Figure 2.1 summarises the selection process.

Figure 2.1. Prisma diagram of literature search results.



Abbreviations: BPH= Benign prostate hyperplasia; EE= Economic Evaluation; SR= Systematic review.

The main objective of Lundkvist was to assess the cost-effectiveness of proton therapy compared to conventional radiation therapy in common types of tumours (prostate, breast, head and neck, and medulloblastoma). The characteristics of the included study, focusing on what was reported regarding prostate cancer, are shown in Table 2.3.

Table 2.3. Study characteristics

Section	Description
	Methods
Patient characteristics	65-year old men diagnosed with prostate cancer.
Intervention or comparison	Proton radiotherapy versus conventional radiation therapy.
Type of economic evaluation	CUA.
Modelling (modelling technique, structural, assumptions)	Markov model consisting of 3 health states (healthy, chronic adverse event conditions and death).
Perspective	Not clearly specified in the study; a societal perspective could be assumed based on the types of costs included.
Setting, country, or jurisdiction	Sweden.
Time horizon	From diagnosis until death or 100 years of age.
Included costs (cost type, cost categories) and resource items	Direct costs associated with the delivery of radiation; management and treatment of AEs and new facilities for proton therapy. Indirect costs associated with travel and accommodation (only for proton therapy).
Data sources for costs and resource use	Cost of different radiotherapy delivery methods was based on costing studies published in Sweden and valued using local sources. The cost of AEs was based on the literature.
The data sources for outcomes and benefits	Literature.
The methods of measuring or valuing outcomes and benefit	Not reported.
Currency	EUR (€) 2002.
An analysis of sensitivity or uncertainty (the type of analysis, parameter variations, structural variations)	Deterministic sensitivity analysis was performed on a few selected parameters using arbitrary ranges.
	Results
Costs and resource use	Only incremental cost was reported, which in the base case scenario was estimated to be €7,952.6. The total costs for the intervention and comparator were not reported.
Outcomes and benefits	The incremental QALYs gained were reported, estimated to be 0.297 QALYs in the base case scenario. The total accrued QALYs gained for the intervention and comparator were not separately reported.
The incremental cost-effectiveness ratio	€26,776 /QALY gained.
Assumptions	Proton therapy delivered a higher target radiation dose. 20% reduction in cancer recurrence with proton therapy. Lower risk of AEs with proton therapy (RR=0.6).
An analysis of sensitivity and uncertainty	ICER ranged between €13,346/QALY gained (when an RR= 0.6 for prostate cancer specific death was assumed) and €662,571/QALY gained when no difference in terms of OS was assumed (RR= 1).
	Surrogate Outcomes
Characterisation of surrogate outcome used in CEA and identification of derived final outcome	The surrogate outcome was cancer recurrence measured as PSA-free survival. A direct relationship between cancer recurrence and cancer mortality was assumed – explained as being that a 20% reduction in cancer recurrence was equal to 20% reduction in cancer mortality.
Source of surrogate outcome evidence used in CEA (e.g., Systematic review of clinical trials)	Individual patient data from 950 patients treated in a single centre between 1987 and 1998 with EBRT alone or EBRT plus short course (≤6 months) ADT (observational data).

Table 2.3. Study characteristics

Section	Description
Methods used in the report to quantify the link between surrogate outcome and final outcome (e.g., Regression-based approach)	Stepwise multivariate analysis to determine whether biochemical failure was a significant determinant of cancer specific mortality.
Consideration of the uncertainty associated with using surrogate outcomes in the results or conclusions or elsewhere in the report	Not reported.

Abbreviations: ADT= Androgen deprivation therapy; CEA= Cost-effectiveness analysis; EBRT= External beam radiotherapy; EUR= Euro; ICER= Incremental cost-effectiveness ratio; PSA= Prostate specific antigen; QALY= Quality adjusted life years; RR= Risk ratio.

As proton therapy was not a common therapy option for prostate cancer at the time the study was conducted, limited information was available on treatment effect and safety. The authors assumed that proton therapy would be delivered at a higher target radiation dose compared to conventional radiation, thereby increasing the health benefits derived. Based on a report by the Swedish Council on Technology Assessment in Health Care, they concluded that patients with localised, intermediate and high-risk prostate cancer would benefit from higher radiation doses in terms of freedom from failure or metastases.

2.5.1. Excluded records

As noted previously, the term surrogate outcome has been used in the literature with multiple definitions which has led to incorrect classification of the published literature in the databases. As this systematic review was able to identify only one study, it was important to understand how the excluded studies modelled their final outcomes. For this purpose, for all those studies excluded using full text review (n=51), a subsequent full text review was conducted to identify the outcome measure used to inform the economic evaluation.

Of the excluded studies reviewed (n=51), 33 (65%) did not use surrogate or intermediate outcomes and were not reviewed further as they used a final outcome to inform their economic model. From those 33 studies, the most common outcome used was direct observation of OS or cancer-specific survival, with 32 of the 33 excluded publications incorporating them into the economic model (14, 17-20, 22, 23, 25-27, 29-31, 34-49, 148-150). However, in the majority of the excluded studies, the source of survival data ranged from medical opinion and assumptions to observational studies and randomised trials reporting survival, often immature and not significant. Consequently, the application

of surrogate outcomes for predicting survival benefits was not employed, resulting in the exclusion of these studies from the systematic review conducted in this thesis.

The remaining studies (n=18) were not economic evaluations or used intermediate outcomes as the basis of the cost-effectiveness outcomes specifically:

- successful surgical treatment (24);
- absence of PSA failure (PSA <0.4 ng/ml) (10);
- hormone-unresponsive disease progression (31);
- Health-related quality of life (HRQoL) (16, 33);
- testosterone levels < 50 ng/dL (50);
- disease progression (29).
- Not economic evaluation ((1-7, 12, 13, 32))

Half of the studies that used intermediate outcomes compared surgical techniques, two compared different radiotherapy delivery methods and one focused on ADT. In contrast, the excluded studies that used observed survival outcomes compared different methods of radiation delivery (16 studies), active treatment vs active surveillance or watchful waiting (6 studies) and the rest focused on ADT and surgical techniques.

2.5.2. Surrogate outcome used

Lundkvist et al. proposed PSA-free survival as a surrogate outcome of prostate cancer specific survival. The difference in treatment effect between proton therapy and conventional radiotherapy, measured as PSA-free survival, was assumed to translate into a difference in cancer specific mortality. The correlation between the surrogate and final outcome was derived from Pollack et al. who proposed biochemical failure as a determinant of distant metastasis and death in prostate cancer patients treated with radiotherapy (151). The methodology used by Pollack consisted of Kaplan-Meier estimates of observational studies that measured PSA-free survival from approximately 1000 patients receiving radiotherapy as the primary treatment with curative intent in a single centre. Cox proportional hazard regressions were investigated using a multivariate analysis of factors associated with distant metastasis and cancer-specific death (151) Biochemical failure, as a time-dependent covariate, was found to be the strongest determinant of distant metastasis and significantly related to prostate cancer death.

The definition of the surrogate and final outcome is one of the elements identified by the PBAC validation framework that facilitates the comparison of different trials based on a consistent measure, thus allowing the use and generalisability of their results (74) The most fundamental assumption in the model proposed by Lundkvist et al. was the correlation between cancer recurrence measured as PSA rise-free survival (the surrogate outcome) and prostate cancer specific mortality (the final outcome). The analysis estimated that the treatment effect of proton therapy led to a 20% reduction in cancer recurrence. This point estimate represents the outcome PSA-free survival which was obtained from a non-systematic literature search (152, 153) A direct relationship between cancer recurrence and cancer specific mortality was then assumed which was interpreted as a 20% reduction in prostate cancer mortality when proton therapy was employed.

The surrogacy relationship in Lundkvist et al. is dependent on the definition of recurrence measured as the level of PSA. Pollack et al. (151) define biochemical failure based on the American Society for Radiation Oncology (ASTRO) consensus statement(154),[7] which considers that a patient experiences a biochemical failure after three consecutive rises in PSA level observed following RT treatment. However, the treatment effect assessed by Lundkvist et al. was not based on the ASTRO definition of biochemical failure. Lundkvist et al. consider the treatment effect reported by Hanks et al. (153) and Slater et al. (152), which defined biochemical failure as 'three consecutive rises in PSA of greater than 10%, or a single dramatic rise that necessitated initiation of androgen therapy'.

Although Pollack et al. conducted a statistical analysis to determine the correlation between biochemical failure and cancer specific survival, the validation of this outcome in this population was not their objective. Lundkvist et al. did not report the use of any source of validation for biochemical failure in the study and the predictive capability of this outcome is not supported in this population. A meta-analytical approach to assess the surrogacy relationship of PSA with prostate cancer specific survival, as proposed in the PBAC guidelines, would ensure the estimates of the final outcome were reliable enough to be used in an economic model. Therefore, according to the definition of surrogate outcomes from the Biomarker Definitions Working Group mentioned in Chapter 1, PSA-free survival should not be used as a surrogate for overall survival in LPC as proposed by Lundkvist et al.

In addition, the study by Lundkvist et al. assumes a direct relationship between recurrence and mortality from prostate cancer. although the results of the analysis conducted by Pollack et al. were

[7] The ASTRO consensus statements are a series of expert opinions and recommendations on various aspects of radiation therapy for cancer treatment. These statements are developed through a process of literature review, expert panel discussions, and consensus-building among ASTRO members.

robust and statistically significant, methods that assign a predictive value to a surrogate outcome based on the availability of data sourced from a single trial are often arbitrary and prone to bias (155) (64) (156) (157). Other methods such as the meta-analytic validation approach, currently suggested by the PBAC in their 'Surrogate to Final Outcome Framework' and discussed in Chapter 3, can overcome this limitation (74)

A predictive relationship between the surrogate and final outcome proposed by Lundkvist et al. has been tested before in the literature in prostate cancer with unfavourable results (158, 159). Lundkvist et al. did not directly assess the uncertainty generated by the surrogate outcome on predicting the treatment effect of the intervention over OS. Moreover, there was no external validation of the model structure and assumptions. However, it was possible to assess the direction of the impact on the results through the results of the deterministic analysis conducted. When an equal risk of death from prostate cancer between the intervention and comparator was tested, the ICER increased by 2,474% compared to the base case scenario (base case scenario RR=0.8 and ICER= €26,776 per QALY vs ICER= €662,571 per QALY when RR=1). Although this analysis tested several assumptions at the same time (surrogate to final outcome relation, a 20% reduction in cancer recurrence for proton therapy and equal treatment effect among proton and conventional therapy), it is likely that the assumption of a direct relationship between the surrogate and final outcome accounts for a high proportion of the variation observed in the ICER.

2.5.3. Quality assessment

When the CHEERS checklist for report quality assessment was applied, 14 of the 24 possible items assessed by the tool were included in the study. From these, at least 3 presented incomplete information or did not address the item's objective properly. Similar results were obtained when the QHES scale. The study of Lundkvist et al. scored 49 out of a possible 100 points, classifying it as poor quality (144-146).

The main limitations of the study were due to the lack of justification provided for key assumptions (i.e., the relationship between surrogate and final outcome), the choice of the technique used for the economic modelling, the source used to derive utilities and the sensitivity analysis. A detailed analysis of each criterion in the CHEERs and QHES checklist along with their score is provided in Appendix 2.

2.5.4. The case of metastatic hormone refractory prostate cancer

During the literature update conducted in June 2022, a study in which a Bayesian multiparameter synthesis was employed to inform decision-making in the context of metastatic hormone-refractory prostate cancer was found (160). As the study focused on a different population to LPC, it was excluded from this analysis; however, it is noteworthy to acknowledge the novel method proposed by Tan and colleagues for estimating treatment effect to improve the granularity of an economic model's result.

Tan et al. expanded the 2-health (dead and alive) state Markov model used in an HTA submitted to the national institute for health and care excellence (NICE) in England (161) to evaluate docetaxel in combination with prednisolone in metastatic hormone-refractory prostate cancer. By employing a bivariate meta-analysis, Tan et al. predicted the treatment effect of docetaxel on progression-free survival (PFS) and were able to populate a 3-state Markov model (stable disease, progressive disease, and dead) which improved the granularity of the results.

The objectives of the study by Tan et al. were different from Lundkvist et. al., as a surrogate outcome was not used to predict the final outcome. Instead, the bivariate meta-analysis used the results of clinical trials including a final and an intermediate outcome to predict the treatment effect of the intervention over a surrogate outcome that was not measured in the clinical trial. This facilitated the generation of a novel health state, thereby augmenting the granularity of the findings of an economic model.

2.6 Discussion

This systematic review of the literature identified one study that used a surrogate outcome to inform an economic model of localised prostate cancer. Lundkvist et al., assumed a direct relationship between cancer recurrence and prostate cancer specific survival. Several limitations of the study were identified, mainly regarding how the surrogate was defined, and a lack of validation and sensitivity analysis testing of the proposed surrogate relationship, which limited the reliability and interpretation of its results.

Overall, this review has highlighted a paucity of good quality publications addressing the use of surrogate outcomes in economic evaluations of LPC. Prostate cancer specific survival and OS were the most frequent main outcomes used to inform economic models in this setting (65%). Only few studies (16%) utilised intermediate outcomes in their cost-effectiveness analyses, primarily in cases of

assumed equal efficacy between the intervention and comparator or when the model was conducted with a short time horizon.

There are two key issues raised by the results of this literature review: (1) the small evidence base addressing the research question i.e., the use of surrogate outcomes in economic evaluations in prostate cancer, and; (2) the absence of evidence addressing the intrinsic uncertainty of the use of a surrogate measure in a way that is informative for reimbursement decision-makers as seen in Lundkvist et al. The latter should reflect the impact that the predicted final outcome, measured in terms of changes in treatment effect on the surrogate outcome, has on the results of economic evaluations. This should be informed by an uncertainty analysis, designed to test the correlation between the surrogate and final outcome, as well as assessing the uncertainty around variations in treatment effect. As the use of surrogate outcomes in an economic evaluation may translate to a higher degree of uncertainty in the results, rigorous and specific uncertainty analysis that addresses the relationship is needed to inform and facilitate the decision-making process.

The optimum level of PSA after treatment may vary depending on the type of primary therapy received; therefore, comparing results among trials based on biochemical failure or biochemical recurrence may be challenging (162). In addition, as noted previously, there was an inconsistency between how PSA failure was defined in the research informing the development of the surrogate used by Lundkvist et al. and how it was applied in the resulting economic model. This disconnect between the treatment effect and the surrogacy relationship can lead to biased results in terms of outcomes prediction and the economic evaluation. As illustrated by Kestin et al., different measures of biochemical failure produced variations of at least 1.6 years in 5-year predicted PC specific survival (163). Moreover, endpoints based on PSA are subject to a variety of confounding factors such as adjuvant treatment received (85), hence the need for a validation process becomes more imperative.

The protracted duration of clinical trials in LPC required to report survival outcomes has prompted a surge of interest in identifying surrogate endpoints that could serve as proxies for overall survival (OS). The ICECaP initiative, discussed in Chapter 3, has been launched to explore the potential of surrogate outcomes in LPC. This endeavour aims to identify reliable and valid surrogate endpoints that can estimate OS and improve the efficiency of clinical trials and economic evaluations. The lack of validated surrogate outcomes in LPC partially explains the identified gap in the literature assessing surrogate outcomes in economic evaluations. This is echoed by a recent review of economic evaluations published in peer reviewed journals which focused on LPC; the results showed that surrogate outcomes were not explicitly used when modelling the natural history of the disease (164).

This scarcity of surrogate outcomes to inform economic evaluations in LPC is not exceptional. The use of surrogate outcomes in HTA submitted to health authorities is also limited. A review of HTA in the UK showed that, between 2005 and 2006, only 10% of economic evaluations based their economic model on surrogate outcomes. Of these analyses, none were related to prostate cancer (139, 164)

2.7 Limitations of the systematic review

The aim of this systematic review was to identify all possible published economic evaluations for interventions in LPC using surrogate outcomes in their model. While the search complies with the recommendations of Mathes et al. about best practice of systematic reviews of health economics publications to include Medline and Embase and at least one health economic database (NHS EED) (140), the search strategy was complicated by the variation in terminology and the lack of a Medical Subject Headings (MeSH) or a standard term to address the use of surrogate outcomes in the published literature. While the search strategy implemented was broad (yielding almost 1.000 titles), and a manual search of the references of the relevant study was also performed, it is possible that not all relevant publications were identified.

Economic evaluations may have different purposes according to the authors' objectives. Although several journals are dedicated to publishing economic evaluations, they are often conducted for the purpose of reimbursement and associated decision-making. This may not always result in publication of a peer reviewed manuscript, particularly where reimbursement applications are considered commercial-in-confidence, and thus may not have been captured in this review. Although the impact of this situation can be minimised by searching the NHSEED database, the information available might be limited and not always available. Similarly, the limitations identified in terms of the analysis undertaken and the lack of detail in the study identified made it difficult to understand the impact generated by the use of surrogate outcomes and how uncertainty should best be handled.

A methodological limitation is the absence of a dual-reviewer approach for independently screening the entries and extracting data. Typically, in systematic reviews, this practice involves having two or more reviewers separately evaluate and extract information from each included study to enhance accuracy and minimise bias. In our review, this dual-reviewer process was not employed, potentially leaving room for variations in interpretation and oversight of relevant details in the screened studies. To minimise potential for bias in study selection, a second reviewer independently checked studies screened against the eligibility criteria (consistency check).

2.8 Conclusions

This systematic review identified one study that used a surrogate outcome to predict changes in OS and inform an economic evaluation of prostate cancer. Given the paucity of available studies, it is difficult to draw firm conclusions on the use of surrogate outcomes for economic evaluations of localised prostate cancer. The results of this systematic review reveal a lack of strong published evidence supporting the use of surrogate outcomes in localised prostate cancer. This finding reaffirms the idea emphasised by the ICECaP working group (85) that a well-defined and validated surrogate outcome in LCP is a key element to inform economic evaluations for HTA to support reimbursement decisions. Rigorous research into the use and usefulness of surrogate outcomes in economic evaluations of treatment for LPC is needed and will be undertaken and reported in the following chapters.

Chapter 3 Evaluation of evidence on the use of MFS as a surrogate outcome in LPC for economic evaluations

3.1 Abstract

This chapter examines and refines the evidence requirements for the use of surrogate outcomes to inform an economic model in a health technology assessment setting. To achieve this, the ICECaP study is presented as a source of evidence, together with its associated analysis validating the use of metastasis-free survival (MFS) as a surrogate outcome for overall survival (OS). Adjustments made to the ICECaP data to enable their use in economic evaluations are also discussed.

The reliability of the treatment effect estimated via the surrogate relationship is of utmost importance for reimbursement decisions. Therefore, the use of MFS as a surrogate outcome for OS in economic evaluations of novel interventions for patients with localised prostate cancer was evaluated in accordance with the PBAC guidelines.

In conclusion, the ICECaP study provided robust evidence to support the use of MFS as a surrogate outcome for OS in localised prostate cancer. The results of the study and assessment against the PBAC framework endorses the use of MFS as a surrogate outcome for OS in economic evaluations aimed at informing reimbursement decision-making in LPC.

3.2 Introduction

3.2.1. Surrogate outcomes in healthcare decision-making

Regulatory approvals of new drugs are traditionally granted when the benefits of a new intervention outweigh its potential risk. To measure the benefit and risk, an established clinical outcome is often measured in clinical trials. Clinical outcomes provide objective and relevant measures of treatment effectiveness, including improvements in symptoms, quality of life, and survival rates. The use of established clinical outcomes is crucial in comparing the effectiveness of different medical products consistently and ensuring that they meet regulatory standards for safety and efficacy (165).

In many cases, the actual desired outcome, such as survival or an improvement in quality of life, can take a long time to occur or be difficult to measure directly. In such cases, surrogate outcomes play an important role in providing evidence of the effectiveness of a treatment or intervention (166). Hence, the use of surrogate outcomes has been increasing as a means of supporting regulatory decisions, particularly to expedite approvals of novel interventions aimed at addressing unmet medical needs.

Regulatory agencies such as the Therapeutic Goods Administration (TGA) in Australia or the US Food and Drug Administration (FDA), offer an accelerated approval process for interventions for rare diseases or serious conditions that fill an unmet medical need (167) In this expedited process, clinical trials results in terms of final outcomes are commonly not available and new drugs are registered on the basis of early data on efficacy and safety, conditional on the provision of updated evidence within an agreed timeframe. In Europe, the majority of new interventions submitted to the European Medicines Agency (EMA) through the conditional marketing authorisation and accelerated assessment pathways were based on intermediate outcomes (168) Marketing authorisation is granted for a new intervention based on an intermediate outcome that is 'reasonably likely' to predict clinical benefit in terms of a final. However, the use of intermediate outcomes as a basis for healthcare decision-making can introduce uncertainty into the process, as the impact of the new intervention on the final outcome remains uncertain. The degree of confidence in regulatory approval decisions based on surrogate outcomes depends on the strength of the correlation between the surrogate and final outcome. Therefore, it is crucial that the treatment effect observed on the surrogate outcome can accurately predict the treatment effect on the final outcome, particularly at the trial level in meta-analysis, where the link between surrogate and final outcome is agnostic to the type of therapy received. To this end, validation studies examining the relationship between the surrogate and final outcomes should be conducted (58).

However, expedited regulatory approvals do not always rely on evidence from validated surrogate outcomes, but lower order, intermediate outcomes. The likely delay in access to potentially lifesaving interventions has been used to justify the use of intermediate outcomes in regulatory/reimbursement decision-making processes, even when there is incomplete evidence supporting their use (58). Regulatory agencies such as the FDA and the EMA have attempted to minimise the uncertainty around the use of non-validated surrogate outcomes in clinical trials for registration purposes by granting conditional approvals. Consequently, intermediate outcomes that are likely to predict clinical benefits, but are without supporting validation studies, are widely accepted for registration purposes (59). A recent review of regulatory approvals granted by the FDA showed that of 25 drugs approved through the accelerated pathway, 84% (21/25) had no trial-level validation studies (55).

The use of non-validated outcomes to grant marketing authorisation may result in disagreement between payers and regulators regarding access to drugs. Although both regulators and payers value clinical benefits, regulators do not usually consider the economic aspect for granting approvals to drugs based on the available evidence and payers react to the uncertainty generated by intermediate

outcomes by demanding additional evidence of benefit. Additionally, a recent examination of the methodological guidance for the utilisation of surrogate outcomes across various international health technology assessment (HTA) agencies revealed that the majority of guidelines primarily concentrate on the application of surrogate outcomes from a clinical standpoint, resulting in the under-development of other areas within health technology assessment, such as safety and economic evaluation (65)

3.2.1.1. Assessment of surrogate outcomes for use in economic evaluations

Regulatory approval and reimbursement decisions regarding new technologies should ideally be based on patient relevant outcomes and evidence from randomised clinical trials (63) However, these outcomes can be complicated to measure or take a long time to become observable, potentially delaying more effective interventions being considered for reimbursement. Surrogate outcomes can be a useful tool in the decision-making process when evidence of the effect on outcomes relevant to patients is not yet available. Nevertheless, the use of surrogate outcomes in the decision-making process is limited by the extent to which confidence can be placed on the predicted treatment effect in terms of the final outcome, as the opportunity cost of an incorrect decision is high.

The validity of a surrogate outcome is assessed by its ability to precisely predict a treatment effect on a final outcome (58) In this regard, biological and epidemiological evidence establishing a connection between the surrogate and the final outcome is essential. When the surrogate outcome lies on the clinical pathway leading to the final outcome, they are likely correlated (58), and there is a good chance that the surrogate will predict a meaningful treatment effect on the final outcome regardless of the treatment received.

Nevertheless, evidence of individual level surrogacy provided only by biological and epidemiological information should not be used to confirm correlation as there are multiple factors that may influence the treatment effect observed in the final outcome (169) Evidence from randomised clinical trials comparing different interventions is needed to quantify the association of treatment effect on both the surrogate and the final outcome (69). Individual and trial level surrogacy occurs when both surrogate and final outcome are correlated, and the treatment effect quantified, which is usually demonstrated using linear regression (73).

The validation of surrogate outcomes involves the analysis of individual patient data from randomised trials using statistical methods suited to assess the acceptability of surrogate outcomes. For the assessment of the correlation between outcomes and the estimated treatment effect on the final

outcome, decision-makers usually outline the requirements the validation process has to comply with before incorporating the surrogacy relationship into economic evaluations. However, the guidelines published by the majority of HTA agencies tend to be unclear about explicit criteria to assess the validity of the surrogate outcome under evaluation (65). A recent review of guidelines for the use of surrogate outcomes in HTA found that only five HTA agencies provided specific methodological advice required to assess the acceptability of a surrogate outcome through a validation process; the PBAC in Australia has developed one of the most detailed guidelines on the inclusion of evidence from surrogate outcomes (65). Moreover, there are inconsistencies in the use and consideration of surrogate outcomes by HTA agencies internationally, with no standardised approach to assessing factors like the level of statistical validation between the surrogate and final clinical outcomes (170). In this thesis, the Australian guidelines for the use of surrogate outcomes in economic evaluations are used to assess the validation process of a surrogacy relationship.

3.3 Objectives

This chapter critically examines the evidence requirements for the use of a surrogate outcome as a means of informing an economic model within the context of health technology assessment. The specific objectives are:

- To elaborate on the ICECaP study and to present a detailed examination of the analysis performed by the ICECaP team to validate the use of MFS as a surrogate outcome of OS.
- To examine the adjustments made to the ICECaP data that are necessary to enable its application in economic evaluations.
- To evaluate the suitability of MFS as a surrogate outcome for OS in economic evaluations of novel interventions for patient diagnosed with localised prostate cancer, against the surrogate framework provided in the PBAC guidelines.

3.4 Methods

A validation assessment of the relationship between the ICECaP surrogacy variables, namely MFS and OS, was carried out using the PBAC framework. The assessment of the ICECaP surrogacy relationship involved the analysis of each criterion in the PBAC framework ([1] definition, selection, and measurement, [2] biological reasoning and epidemiological evidence, [3] evidence from RCT coming from other interventions, [4] comparative treatment effect and [5] any relevant considerations), to

ensure the quality of the evidence, the relevance of the estimates from surrogate to final outcomes, and the potential limitations of the data.

Furthermore, individual patient data (IPD) were obtained from the ICECaP study and subject to a survival analysis in order to support the development of a subsequent economic model (see Chapter 4). The survival analysis focused on the suitability of the individual trials for the proposed economic model and subsequently to model the natural history of patients undergoing treatment for LPC. The survival analysis focused on three crucial outcomes of interest that were informed by the ICECaP study:

- Biochemical recurrence-free survival
- Metastasis-free survival
- Overall survival.

As noted in Section 3.5.2.3, the ICECaP largely reflects treatment experience with RT, one of the mainstays of therapy in LPC. As such, RT was chosen as the comparator arm for this demonstration model. The selection of the intervention was based on the disease burden of high-risk LPC patients, as well as the contrast between current clinical guidelines for abiraterone (which is limited to metastatic stages) (108) (124). The choice of the intervention is discussed Chapter 4. Briefly, it was based on the burden of disease of patients with high-risk LPC and the contrast of what is currently recommended in clinical guidelines for abiraterone (limited to metastatic stages) and the emerging evidence from recent studies demonstrating improved quality of life and progression-free survival (PFS) in these patients (125, 126).

In Chapter 4, the overall structure of the partitioned survival model built using data from the ICECaP study is discussed. Briefly, the model partitions the ICECaP study population into a four-health state model comprising BCR-free, biochemical recurrence (BCR), metastasis, and death. The movements of individuals between these four health states over time is modelled based on estimates of OS, MFS, and BCR-free survival derived from the ICECaP data.

3.4.1. The ICECaP database

The lack of a validated and robust surrogate outcome in localised prostate cancer motivated the work of the Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP) working group. The main objective of this group was to reduce the time needed to conduct clinical trials in patients diagnosed

with localised prostate cancer and to accelerate the development of new interventions for these patients (85).

Given the long time it takes for a new intervention to show statistically significant results in OS for patients diagnosed with localised prostate cancer, the use of a surrogate outcome has been proposed for regulatory and reimbursement decisions in this population (132). Previous attempts at identifying intermediate clinical endpoints (ICEs) as surrogates for overall survival (OS) in localised prostate cancer used single studies. Validation using a single trial usually does not provide sufficient evidence to support the validity of the surrogate outcome. A single trial may have limitations such as small sample size, short follow-up period, or lack of generalisability, and may be subject to various sources of bias and confounding which can affect the accuracy of the results (58).

To ensure that a surrogate outcome is valid, it is necessary to include data from multiple trials, ideally with large sample sizes and a long follow-up period, and to consider various sources of bias and confounding. This will provide a more robust evidence base for the validation of the surrogate outcome (58) Hence, the ICECaP hypothesis was that DFS and/or MFS could be considered surrogates for OS as they have a strong correlation with death from prostate cancer as demonstrated in multiple trials. The study also evaluated other intermediate outcomes such as time to metastasis (TTM) or time to disease recurrence (TDR) using data sourced from various trials (132).

While the design and conduct of the ICECaP study were not part of the work involved in this PhD, its description is essential to understand the source, strengths and limitations of the data accessed and applied to this analysis. Accordingly, the following section provides a general description of the methods and results of the ICECaP project. The specifics of how data were sourced from ICECaP and modified for the purposes of this PhD are subsequently addressed.

3.4.1.1. ICECaP Methods

The ICECaP Working Group conducted a systematic literature review and meta-analysis to identify studies in LPC patients reporting outcomes that could potentially be used as surrogates for OS (85). For this propose, the following sources were searched:

- PubMed
- Clinical Trials registries
    - NIH (http://www.clinicaltrials.gov)
    - ISRCTN registry
    - EU CTR:

    - ANZ CTR
- Cooperative Groups' websites: ACOSOG, CALGB, CUOG, ECOG, EORTC, GETUG, MRC CTU, NCIC-CTG, OCOG, RTOG, SPCG, SWOG, TROG, UK MRC CTSU, UK NIHR CRN/NCRI PCSG, UK ICR/CTSU etc.
- Research organisation meeting websites:
    - ASCO
    - AUA
    - ESMO
- Review of citations in publications including meta-analysis papers

Trials were considered eligible if[8]:

- They were randomised, controlled trials for localised prostate cancer.
- They were conducted in Canada, UK, Europe, Australia/New Zealand, US.
- Accrual was completed /terminated.

An invitation letter was sent to the study chairs of each identified study and IPD was collected from all trials that agreed to participate.

3.4.1.1.1. Statistical analysis within ICECaP

IPD from the trials was pooled to perform surrogacy analysis of intermediate clinical endpoints for overall survival. The surrogacy criterion was evaluated using a two-stage validation meta-analytic approach where two conditions were tested:

- Condition 1: required that the potential surrogate outcome and OS be correlated (both at the patient level and trial level) independent of the treatment received. The correlation was assessed by means of a regressions analysis of the 5-year rate of the surrogate outcome over the 8-year rate of OS [9].
- Condition 2: required the correlation of treatment effects (performed by means of a Cox proportional hazard regression) observed on both outcomes.

[8] The systematic review excluded trials that had a primary endpoint such as safety, toxicity, QOL, feasibility, dosimetry, patient decision-making or without systematic long-term follow-up.

[9] A copula is a function which joins a multivariate distribution function to its one-dimensional marginal distribution function. They are used in statistical analysis to describe the correlation between random variables. In this case, a bivariate copula model fitted to the IPD was used to correlate the potential surrogate outcome with OS at the patient level.

For Condition 1, the bivariate Clayton's copula model method was used to analyse the relationship between the surrogate and the final outcome by modelling their dependence structure (58). The bivariate Clayton's copula model was used to determine how likely the surrogate outcome influenced the final outcome and the degree of their dependence.

To quantify the association or correlation between outcomes a Kendall's Tau model was used (171). This model is a non-parametric method that measures the strength of the relationship between two variables based on the number of concordant and discordant pairs. The number of discordant and concordant pairs is then used to calculate a value, known as Kendall's Tau, that ranges from -1 to 1, with 1 indicating a perfect positive relationship and -1 indicating a perfect negative relationship.

In the context of the bivariate Clayton's copula model quantified by Kendall's Tau method, the Clayton's copula is used to model the dependence structure between two variables, while the Kendall's Tau method is used to quantify the strength of the relationship between those variables.

For the trial level surrogacy or Condition 2, the correlation between treatment effects on both MFS and OS was analysed in an error-in-variables regression model to account for estimation errors in the treatment effects. For this purpose, a Cox proportional hazards regression model was estimated to obtain the treatment effect measured as the HR on MFS and OS. A weighted linear regression was then fitted using the treatment effects HRs of each trial to find the correlation equation linking both outcomes. The validation of the surrogate was reflected by the strength of the correlations of the regression analysis defined a priori by an R^2 value ≥ 0.7.

3.4.1.1.2. ICECaP outcome definition

The ICECaP database measured different outcomes in a population stratified by risk (intermediate/low or high risk[10]) and age (64 or younger, 65-74, 75 or older). The outcomes of interest identified as potentially relevant surrogate and patient relevant events were defined as follows:

- Overall survival (OS): measured from the date of randomisation to death from any cause, censored at the date of last follow-up.
- Disease-free survival (DFS): measured from the date of randomisation to the first evidence of recorded clinical recurrence (local/regional progression and/or distant metastases confirmed

[10] High risk was defined if patient had one of these features: high risk by NCCN or D'Amico criteria, high risk by pathological criteria (pathological Gleason>7, or seminal vesicles involvement, or ≥pT3b stage) or Pathology N1.

by imaging or histological evidence) or death from any cause; or censored at the date of last follow-up.

- Biochemical recurrence (BCR): defined according to 2006 RTOG-ASTRO Phoenix (172). A PSA rise by 2 ng/mL or more above the nadir was considered biochemical failure; patients not meeting these PSA criteria for failure who underwent salvage therapies were also declared as failures at the time of salvage initiation.
- Metastasis-free survival (MFS): measured from the date of randomisation to the first evidence of recorded distant metastases (excluding pelvic lymph nodes metastasis) or death from any cause; or censored at the date of last follow-up.
- Time to metastasis (TTM): defined analogously to MFS but non-prostate cancer deaths without metastasis were censored at date of death.

3.4.1.2. ICECaP results

The study analysed data from 21,140 patients enrolled in 24 trials and 12,712 patients enrolled in 19 trials with documented data on DFS and MFS, respectively. Five trials were divided based on the type of primary therapy received, resulting in 31 and 21 study units for DFS and MFS analysis. The trials enrolled patients between 1987 and 2011 with a median follow-up of 10 years, hence, the survival analysis of ICECaP was truncated at 10 years. Over 80% of patients were over 75 years old. The 5-year rates for DFS, MFS, and OS were 76%, 79%, and 84%, respectively.

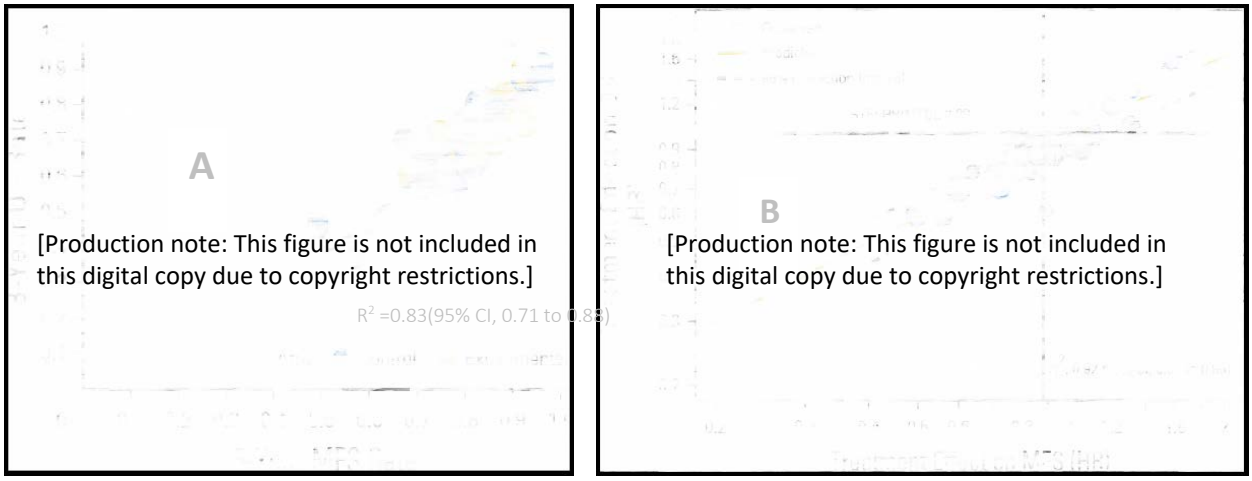
The study found a strong correlation between DFS and MFS and the final outcome, OS, at the individual patient level (R^2 of 0.85 for DFS and 0.91 for MFS). The correlation between the 8-year OS rates and the 5-year DFS and MFS rates was also found to be high, with R^2 of 0.86 and 0.83 respectively. When considering the correlation between treatment effects at the trial level, the study found a high correlation between log(HR)-OS and log(HR)-MFS across trials (R^2 of 0.92), with a similar result even when non prostate cancer deaths were censored (R^2 of 0.89).

The validation of MFS as a surrogate outcome of OS performed by the ICECaP working group, indicates that MFS is a strong surrogate for OS in localised prostate cancer and provides an opportunity to use the surrogacy relationship in an economic evaluation (132). The assessment of the methodology used to test the surrogacy relationship conducted by the ICECaP group as well as the results and issues faced will be addressed in section 3.5 where the PBAC framework for the use of surrogate outcomes in economic evaluations is applied to MFS as a surrogate of OS.

3.4.1.2.1. The surrogacy relationship

At the individual level, the results of the bivariate copula model correlating MFS and OS yielded an R^2 of 0.91 (95% CI, 0.91 to 0.91)[11] when regressed against OS. When the MFS rate at 8 years was regressed against the OS rate at 5 years, the model returned a R^2=0.83 (95% CI, 0.71 to 0.88) (Figure 3.1). At the trial level, when the HRs of MFS were regressed against the HRs of OS observed in each trial, the R^2 was 0.92 (95% CI, 0.81 to 0.95) and the resulting equation correlating both outcomes in terms of relative treatment effect was *Log (HR)-OS= -0.021 + 0.740*Log (HR)-MFS.*

Figure 3.1. Correlation between MFS and OS at individual and trial level for all the studies included in ICECaP.



Source: Xie 2017.
Abbreviations: HR= Hazard ratio, MFS= Metastasis-free survival, OS= Overall survival. A) Correlation of outcomes, MFS rate and OS rate independent of treatment received by trial and arm. B) Correlation of treatment effect between MFS and OS.

3.4.1.2.2. Refining the ICECaP equation for health economics

The ICECaP study aimed to identify the most appropriate surrogate outcome measure for OS in localised prostate cancer (LPC) and therefore included studies utilising a variety of outcome measures. For health economics purposes, the ICECaP study served a crucial role in shaping the model; specifically, the study provided vital information for two key aspects:

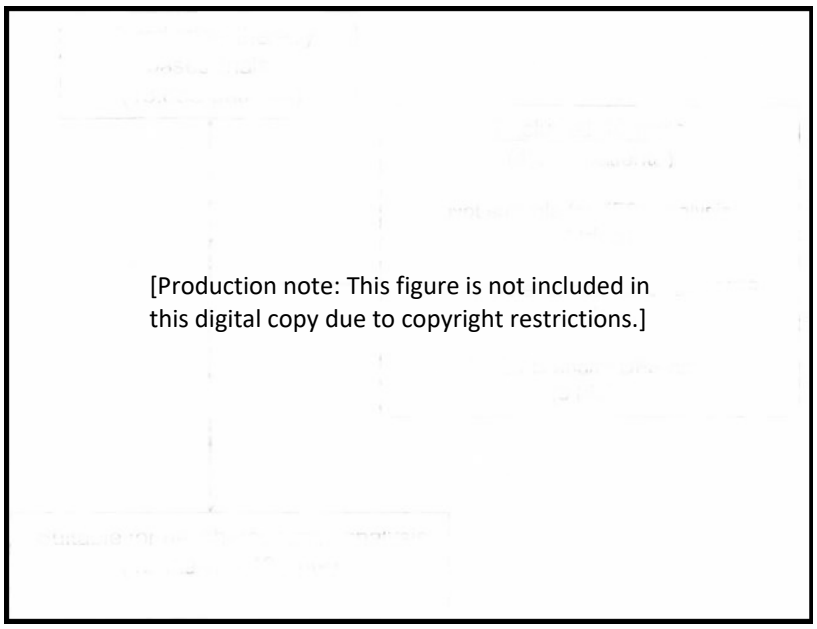
[11] R-squared (R^2) is a statistical metric that reflects the proportion of variance in the dependent variable that is explained by one or more independent variables in a regression model, in this case, it reflects the proportion of variation in the treatment effect of OS that is explained by variation in the treatment effect on MFS. It differs from the correlation in that it focuses on the variance of the variables instead of the strength of the correlation between the variables. For example, an R^2 value of 0.91 indicates that 91% of the variation observed in the treatment effect of OS can be explained by the variations of the treatment effect over MFS.

1. The treatment effect of a new intervention on overall survival (OS) was estimated based on the treatment effect reported on metastasis-free survival (MFS), taking into consideration the surrogacy relationship between these two outcomes.
2. The model's representation of the natural progression of LPC was built based on the survival analysis conducted with the individual patient data obtained from the ICECaP study.

In order to construct an economic model that accurately represents the natural history of the disease for health economic purposes, studies that included data on metastasis and recurrence after primary treatment were deemed necessary. Therefore, studies that did not include MFS and DFS were excluded from the analysis for health economic purposes. From the 22 studies in ICECaP, only 12 reported outcomes related to recurrence and metastatic events and were requested for the purposes of investigating the use of the surrogate in an economic evaluation.

From the ICECaP database of trials, 10 studies (4,357 patients) were excluded that did not collect metastasis and event-free survival data, had a median follow-up period of less than 6 years or had fewer than 3 prostate cancer-related deaths. The Prisma diagram describing the inclusion process of the studies identified and how they differ from the analysis performed in the ICECaP study is shown in Figure 3.2.

Figure 3.2. PRISMA diagram of studies identified.



Source: Figure provided by Wanling Xie.
DFS: Disease-free survival; EFS: Event-free survival, MFS: Metastasis-free survival

The IPD from 12 randomised clinical trials which reported their results in terms of MFS were used for the economic evaluation. The studies, which started randomisation in 1992 and finished in 2010, included 9,481 patients in total. The mean age of patients ranged between 67 and 72 years and the majority were classified as high risk (61%). While radiotherapy was the main treatment received by all patients, a variety of oral and localised treatments were given as adjuvant therapy. Cryoablation was used in one study, neoadjuvant ADT was used in 9 studies and adjuvant ADT in 5 studies. Treatment with neoadjuvant ADT ranged between 3 to 8 months whereas adjuvant ADT treatment varied between 4 months and 3 years.

The data custodians from the ICECaP study were supportive of strengthening the research by sharing relevant analysis of their work for health economic evaluation purposes.

3.4.2. The PBAC Surrogate Framework

The use of biomarkers and intermediate outcomes in regulatory and reimbursement evaluations is becoming increasingly prevalent. In order to standardise their use in the decision-making process, several regulatory agencies have established guidelines for the evaluation and validation of surrogate outcomes (170). Despite the frequent use of surrogate outcomes in regulatory approvals, there is a scarcity of comprehensive guidance regarding the statistical methods to be employed for the validation of such outcomes across health technology assessment agencies. This absence of clarity increases the uncertainty surrounding the use of surrogate outcomes as a basis for economic evaluations in the context of reimbursement decision-making. A recent review of methodological guidelines of the use of surrogate outcomes across 73 international HTA agencies found that only 5 agencies provided detailed advice on the statistical methods that should be used for the validation of surrogate outcomes (65). In the study of Grigore et al., the majority of the methodological guidelines were non-specific, with only a few agencies providing detailed criteria for the acceptance of surrogate endpoints.

One of the few agencies with detailed information on how to assess the validity of surrogate outcomes to inform economic evaluations for HTA is the PBAC. The PBAC has formulated a comprehensive methodology for the evaluation of surrogate outcomes to be used in submissions which incorporate a surrogate outcome. This framework outlines a systematic approach for considering the inherent uncertainties involved in translating the quantitative and qualitative effects derived from surrogate outcomes to the final outcomes, where that information is to be integrated into economic evaluations (74).

The framework defines the information required for a surrogate outcome to be used to predict the effectiveness of a new intervention in terms of a final outcome for use in economic evaluations required for the decision-making process. The framework is divided into five major topics:

1. Definition, selection, and measurement.
2. Biological reasoning and epidemiological evidence.
3. Evidence from RCT coming from other interventions.
4. Comparative treatment effect.
5. Relevant considerations (such as the uncertainty brough by the use of surrogate outcomes in the model).

The aim of the framework is to provide guidance on the use of surrogate outcomes within economic evaluations for applicants preparing submissions to the PBAC. While it is not the aim of this PhD to produce an economic evaluation for submission to the PBAC, the intent is to demonstrate whether the proposed ICECaP surrogate (MFS) is consistent with the conditions outlined by the PBAC for the use of a surrogate for inclusion in an economic evaluation.

Consequently, in this Chapter, an overview is presented that compares the proposed ICECaP surrogate (5-year MFS) with overall survival using the PBAC framework. A comprehensive examination of how all individual components of the framework are addressed with respect to the use of the ICECaP surrogate (5-year MFS) can be found in Appendix 3.

It is important to note that the SoC (radiotherapy) for patients with LPC where the surrogate relationship is applied is a non-drug intervention and therefore, the PBAC framework is not strictly applicable. Although MSAC provides a general guideline on the 'Use of surrogate outcomes to estimate final outcomes' (173), the PBAC provides a more detailed framework to evaluate the relationship and, in practice, it can be applied across different interventions regardless of the mechanism of action. Additionally, as the MSAC guidelines for the use of surrogate outcomes are based on the PBAC framework, for the purpose of this analysis, it was decided to use the latter to assess the validity of MFS as a surrogate of OS.

3.4.3. Survival analysis

Quantification of costs and benefits are required in economic evaluations over the lifetime of the cohort of patients analysed. To account for all costs and benefits accrued beyond the 10-year truncated analysis in ICECaP, extrapolation of the outcomes of interest was required.

Different parametric models (exponential, Weibull, generalised gamma, gompertz, log-logistic, log-normal) were evaluated to determine the best fit of the ICECaP data considering all possible outcomes and over the whole population. The adequacy of the parametric model was assessed through visual inspection of the parametric curves obtained through overlapping with the Kaplan Meier curve and assessing three statistical parameters: the Akaike information criteria (AIC), the Bayesian information criteria (BIC)[12] and the log-cumulative hazard plots[13] (174).

While the AIC and BIC are commonly used to evaluate the fit between modelled and actual data over a given follow-up period, they do not provide any insight into the performance of a model beyond the follow-up period (175) In order to assess the performance of a model beyond the follow-up period, it is crucial to visually inspect the hazard function and consider the clinical viability of the survival function resulting from extrapolation. Accordingly, the results of the survival analysis were discussed with health economics experts[14] to determine the appropriate survival probabilities and parametric distribution to use for extrapolating the outcomes of interest. These experts provided valuable information based on prior knowledge and expectations about the disease modelling that helped to elucidate plausible clinical outcomes in the extrapolated portion of the survival curves.

In the survival analysis conducted, the outcomes used to inform the economic model were considered independent of other outcomes being modelled. Therefore, there was a possibility that the survival curves crossed over the extrapolation period. This could lead to an incorrect interpretation of the survival i.e., the hazard of progressed disease was higher than the hazard of death. To avoid this scenario, survival curves in such models are frequently arbitrarily adjusted which has led to criticism by HTA agencies (176).

---

[12] AIC and BIC were used to compare different parametric models and to select the best fitting model. AIC measures the trade-off between the goodness of fit of the model and its complexity, i.e., the number of parameters in the model. A lower AIC value was indicative of better fitting model. The formula for AIC is:
AIC = 2k - 2ln(L), where k is the number of parameters in the model and L is the maximum likelihood of the model.
BIC was use like AIC. The formula for BIC is:
BIC = ln(n)k - 2ln(L), where n is the sample size and k and L are as defined for AIC.
The model selected was the one with the lowest AIC and/or BIC value, as this was indicative of the best fitting model with the optimal balance between goodness of fit and model complexity.

[13] The log cumulative hazard plot was used to help the identification of the best model to fit the data. A visual examination of the graph of lg[-lg[survival probability]] vs lg[time] can help to identify if the hazard function is monotone, constant or has different shapes over time as changes on the slope of the cumulative hazard function can indicate a variation of the underlying hazard function. When the log cumulative hazard plot shows different gradients, it implies the hazard function changes overtime and therefore, more complex methods of modelling are required to accurately capture the survival function.

[14] The members of the panel of health economists were Richard De Abreu Lourenço, Tracey-Lea Laba and Stephen Goodall.

In this model, the choice of extrapolation curves was hierarchical to OS. The parametric model for OS was first assessed and chosen according to the approach described above. Subsequently, the choice of parametric models for MFS and BCR-free survival was guided by the best fit for those curves which lay below the best fitting OS curve, avoiding curves crossing among different populations and between outcomes evaluated for each population. All parametric models were evaluated for all subsets of the population in ICECaP (64 or younger, 65 to 74 years of age, 75 or older, high risk and low/intermediate risk). As the mechanism of action of the interventions compared were similar, the outcomes modelled (BCR-free, MFS and OS), were assumed to exhibit the same functional form in both the comparator and intervention arms. Therefore, the parametric distribution selected was applied to both, the intervention (RP) and comparator (RT). The survival analysis to obtain the Kaplan-Meier curves and extrapolations was conducted using STATA® (version 15.1).

3.4.3.1. Mean survival time

Mean survival time or the time patients spend in metastasis or recurrence health states can help in understanding the benefits of the intervention, where the gains in quality of life come from, or differences in treatment effect in terms of survival (177) This analysis of mean survival was determined by the area under the curve (AUC) for each outcome modelled from time 0 until the last time of follow-up in the ICECap data plus the AUC of the extrapolated portion (178). The trapezoidal[15] rule was used to calculate the AUC in Microsoft Excel®.

3.4.3.2. Proportional hazard assumption for the intervention

The construction of MFS and BCR survival curves for AAP was based on the HR comparing RT versus AAP under the assumption of Cox's proportional hazards given by equation 3.4.1 (179):

$$\ln\left(\frac{h(t,X)}{h_0(t)}\right) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots \qquad \text{(Eq.3.4.1)}$$

In this model, *h(t,X)* is the hazard (instant rate of failure, that is, failure between *t* and *Δt*, of a patient who has survived up to time *t*, divided by *Δt*, as *Δt* approaches zero) at time t for a patient with covariates $X_1, X_2, .., X_n$. The hazard of a patient when all covariates = 0 at time t is *h0(t),* often called the baseline hazard at time *t*. Where $X_1, X_2, ..X_n$, stand for variables that may influence the failure

[15] In the trapezoidal rule, the AUC is estimated as the sum of the area of all trapezoids formed between 2 consecutive time points: $AUC_0^{t_n} = \sum_{i=1}^{n-1} \frac{1}{2}(t_{i+1} - t_i)(S_{(t)i} + S_{(t)i+1})$ where t = time and $S_{(t)}$ = survival probability.

probability such as patient's age, presence of co-morbidities or treatment received. By exponentiating this equation, it can be rewritten in terms of the hazard:

$$h(t,X) = h_0(\mathrm{t})\exp[\beta_1 X_1 + \beta_2 X_2 + \ldots \beta_n X_n]. \quad \text{(Eq.3.4.2)}$$

As an example, the hazard function for distant metastases developing in patients undergoing radiotherapy (i.e., X_1 = 1), is given by

$$h_{RT}(\mathrm{t,X}) = h_0(\mathrm{t})\exp[\beta_1 X_1 + \ldots \beta_n X_n], \quad \text{(Eq.3.4.3)}$$

while for patients undergoing radical prostatectomy (i.e., X_1 = 0), the hazard function for distant metastasis is given by

$$h_{RP}(\mathrm{t,X}) = h_0(\mathrm{t})\exp[\beta_2 X_2 + \ldots \beta_n X_n], \quad \text{(Eq.3.4.4)}$$

Then the ratio of these hazard functions (HR):

$$HR_{DM} = \frac{h_{RT}(\mathrm{t,X})}{h_{RP}(\mathrm{t,X})} = \frac{h_0(\mathrm{t})\exp[\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n]}{h_0(\mathrm{t})\exp[\beta_2 X_2 + \ldots \beta_n X_n]} \quad \text{(Eq.3.4.5)}$$

$$= \frac{h_0(t)\exp[\beta 1]\exp[\beta_2 X_2 + \ldots \beta_n X_n]}{h_0(t)\exp[\beta_2 X_2 + \ldots \beta_n X_n]} = \exp[\beta_1].$$

As the hazard function can be understood in terms of the survival probability (and vice versa), survival at time t (which is denoted by *S(t)*) can be written in terms of the hazard function:

$$S(t,X) = \exp\left[-\int_0^t h(u,X)\mathrm{du}\right] = \exp\left(-\exp(\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n)\left[-\int_0^t h_0(u)\mathrm{du}\right]\right. \quad \text{(Eq.3.4.6)}$$

$$S(t,X) = \left(\exp\left[-\int_0^t h(u,X)\mathrm{du}\right]\right)^{\exp(\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n)}$$

$$S(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n)}.$$

Thus, the survival function for patients who receive radiotherapy is given by

$$S_{RT}(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n)}, \quad \text{(Eq.3.4.7)}$$

and the survival function for patients who receive AAP is given by

$$S_{AAP}(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \ldots \beta_n X_n)} = S_0(t)^{\exp[\beta 1] exp[\beta_2 X_2 + \ldots \beta_n X_n)} \quad \text{(Eq.3.4.8)}$$

$$= (S_{RT}(t))^{\exp[\beta 1]}$$

$$= (S_{RT}(t))^{[\mathrm{HR}]}.$$

Thus, in this analysis, the probability of dying, developing distant metastases or BCR for patients undergoing AAP was calculated from the survival function of patients who received RT raised to the

power of the hazard ratio when both alternatives are compared under the Cox's proportional hazard assumption.

3.4.3.3. Adjustment of survival curves

Statistical tests can help validate parametric models that have been fit to the Kaplan Meier curves. As seen in section 4.5.4 log cumulative hazard plots, as well as the AIC and BIC, were used as indicators of how closely the parametric models fit the within-trial data. However, these methods are not appropriate to validate the estimated survival probabilities beyond the available follow-up period as they lack the capacity to estimate the plausibility of extrapolations (180). Biological plausibility, expert opinion, and comparison with different data sets from the same population, are required to adjust the extrapolated portion of the survival curves (174).

Different methods to adjust the OS probabilities have been developed which use external data from life tables to adjust the extrapolated portion (181) (182). The long-term survival probabilities should be estimated or adjusted by the overall population mortality records depending on the relationship between the disease and the population hazard [16]. By adjusting the survival curve to the life tables in the Australian population, the effects of differences in age distributions between populations or subgroups can be incorporated, making the survival estimates more plausible.

An important factor to consider in this population is that the mortality rate observed during trials usually represents the cancer specific mortality rate. When this occurs, extrapolations beyond those data may be driven by disease specific mortality, which may be problematic because the hazard of death after the clinical trial period may be driven by other causes. In LPC, the late onset of the disease and the relative high survival rate may result in a long-term hazard of death similar to that of the general population (183) This situation may produce biased survival estimates in the extrapolation portion of the curve (fewer deaths than would otherwise occur) as they reflect the hazard function of a younger population under treatment observed in the trial. Therefore, the overall population mortality records in Australia (184) were used to reflect the hazard function in the long term. In particular, the extrapolated sections of the overall survival curves for the intervention and comparator

[16] The risk of death between the patients under study and the general population can be:

- Same at all times: In this case, the long-term hazard could be estimated from the trail or directly from the overall population mortality records.
- Different in the short term but same in the long term: the short-term hazard should be informed by the trial data and the long-term hazard could be estimated from the overall population mortality records.
- Different in the short and long term: In this case, the long-term hazard estimated from the trial data should be adjusted by the hazard observed in the overall population mortality records.

were compared to the observed survival probabilities in the general population. The curves converged with the general population's survival curve when the extrapolated hazard function exceeded the overall mortality rate in the general population.

3.4.4. Recurrent disease

The ICECaP database defined biochemical failure (BCR) according to the 2006 RTOG-ASTRO Phoenix based on elevated PSA after primary treatment which is accordant with the literature in the field (172). This definition is based on PSA elevation, a widely used test in clinical practice for patients with prostate cancer; however, it does not make a distinction between patients with elevated PSA due to BCR or metastasis.

Progression without metastasis is an important state to consider as it leads to early adoption of ADT and therefore the adverse events and loss of quality of life associated with this treatment. It is also associated with castration resistant prostate cancer after failure with ADT and the use of more complex therapies for patients who develop metastasis subsequent to treatment. However, recurrent disease without metastasis was not an outcome analysed in the ICECaP study.

To extract BCR without metastasis, other outcomes analysed in the ICECaP study were used. As TTM recorded the number of patients that developed distant metastasis and DFS was defined as the number of patients who developed local or regional progression and distant metastasis, the subtraction from the total cohort of failures among each of these populations resulted in the number of patients who developed recurrent disease but were free of metastasis at each time interval.

3.5 Results

3.5.1. The surrogacy relationship for health economics

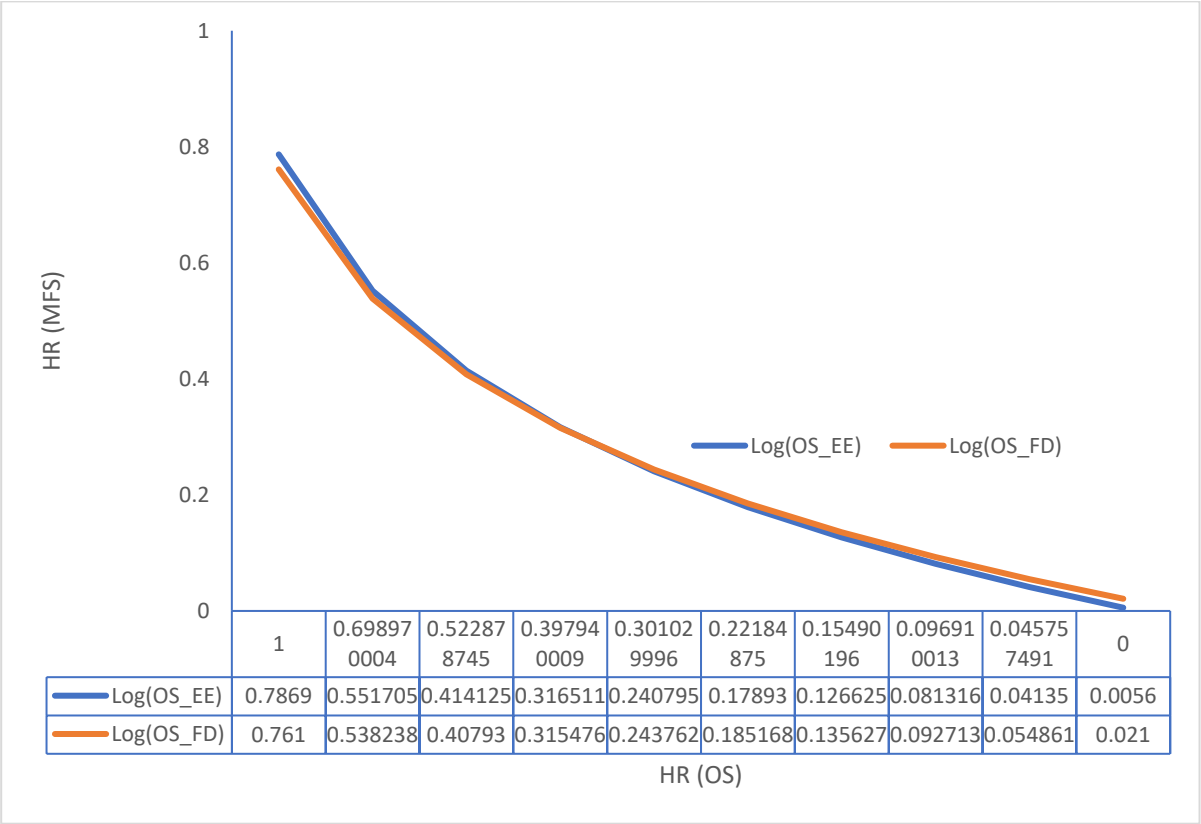
The results described above were produced using the full set of trials included in the ICECaP database. When restricted to those trials suitable for the health economics analysis, the regression equation found was: Log (HR)-OS= -0.0056 + 0.7813* Log (HR)-MFS. The performance of the two regression equations was compared with respect to the predicted HR when applying the IPD from the full dataset and for the health economic analysis. Figure 3.3 shows the predicted values of HR obtained when the regression equation, using both the full data set and the data from the health economic analysis, is applied to predict OS using a broad range of HR values for MFS.

The results of the regression analysis indicate that in practical terms, the results of both regression equations are similar [17], with the exception of the estimates of OS-HR that were calculated using values of MFS-HR closer to 1. This outcome was anticipated due to two factors: (1) When the comparative treatment effect measure, as indicated by the HR, is equal to 1, the efficacy of an intervention is not statistically significant; (2) for MFS HR values greater than 0.88, which is defined as the STE[18], the surrogacy relationship do not predict a significant effect on OS.

[17] A log rank test was conducted to evaluate differences in survival times between the full data set and the economic evaluation data set. The test was not significant, $\chi 2(1) = 0.2$, $p = 0.6$, indicating survival probabilities estimated with the full data set and the economic evaluation data set are similar.

[18] The STE is the minimum treatment effect on the surrogate outcome that can predict a meaningful treatment effect on the final outcome. It can be defined in terms of the 'amount' or magnitude of the treatment effect that must be observed in the surrogate outcome to anticipate a significant treatment effect on the final outcome.

Figure 3.3. Difference in the surrogacy relationship due to the data used.



	1	0.698970004	0.522878745	0.397940009	0.301029996	0.22184875	0.15490196	0.096910013	0.045757491	0
Log(OS_EE)	0.7869	0.551705	0.414125	0.316511	0.240795	0.17893	0.126625	0.081316	0.04135	0.0056
Log(OS_FD)	0.761	0.538238	0.40793	0.315476	0.243762	0.185168	0.135627	0.092713	0.054861	0.021

Abbreviations: FD= Full data; EE= economic evaluation, HR= Hazard ratio, MFS= Metastasis-free survival, OS= Overall survival.

3.5.2. Assessment of ICECaP against the PBAC Surrogate Framework

3.5.2.1. Definition, selection, and measurement.

3.5.2.1.1. Proposed surrogate measure.

The definition of the outcome is a key element of the use of surrogate outcomes to inform an economic evaluation. The consistency of the outcome measure and definition across all studies used to assess the relationship between surrogate and final outcome will allow for consistent prediction of the final outcome to inform economic evaluations.

In prostate cancer, metastasis is usually investigated using imaging techniques in patients with signs or symptoms of failure after initial treatment. Although new methods such as PSMA-PET/CT with improved accuracy have been found to be efficient in identifying metastasis in certain groups of patients, (185), international clinical guidelines still recommend CT + bone scans for the general population of prostate cancer patients (108, 186).

MFS was measured from the time of diagnosis until death or metastasis as confirmed by CT and bone scan. As in other neoplastic diseases, metastasis is the final state of the natural course of the disease and it is discussed that it has been shown to causes 90% of cancer specific deaths (187).

The ICECaP analysis included studies that consistently applied the recommended imaging techniques for diagnosis. With the exception of one study (188) which did not specify the imaging technology used, all other studies utilised CT and bone scan to assess patients for potential distant metastasis. Of the 12 studies included in the meta-analysis, 10 used bone scan, one did not specify the imaging used (188) and one provided details on the range of different imaging technologies used, including CT and bone scans (189). Therefore, in terms of the method of measurement, the surrogate outcome was defined consistently in all studies included in ICECaP; this definition is also in line with recommendations in clinical practice and international guidelines.

3.5.2.1.2. Final outcome and relationship with the proposed surrogate

Overall survival, the final outcome for which MFS is proposed as a surrogate, is, by definition, patient centred, irreversible and certain. The relationship between OS and MFS has been explored in the ICECaP analysis using data from RCTs over a long period of time (the enrolment of patients in all studies occurred between 1992 and 2010. Despite the fact that the RCTs used in the meta-analysis may not be deemed contemporary, the imaging methods employed for patient staging and the radiotherapy are still recommended in clinical guidelines.

RCTs included in the ICECaP database recorded the time to metastasis and survival from all patients whose data was pooled in the meta-analysis. The median time at risk of developing metastasis for these patients was 7.3 years when OS was measured; for patients who developed metastasis it was 7.0 years. The longer time at risk of metastasis in relation to OS and the proximity of both outcomes is likely indicative of the chronology of localised disease where metastasis occurs first, followed by death.

The surrogate relationship could also be explained through epidemiological evidence, but this type of evidence is more uncertain compared with evidence from RCTs. The PBAC guidelines states that ‘the closer the evidence is to level I (multi-trial meta-regression possible with one or more drugs) the less important epidemiological data become’ (74). Consequently, as the relationship between MFS to OS was built using data from multi-trial meta-regression analysis conducted with different interventions, epidemiological data linking these outcomes is not discussed in this analysis.

3.5.2.2. Biological reasoning and randomised trial evidence

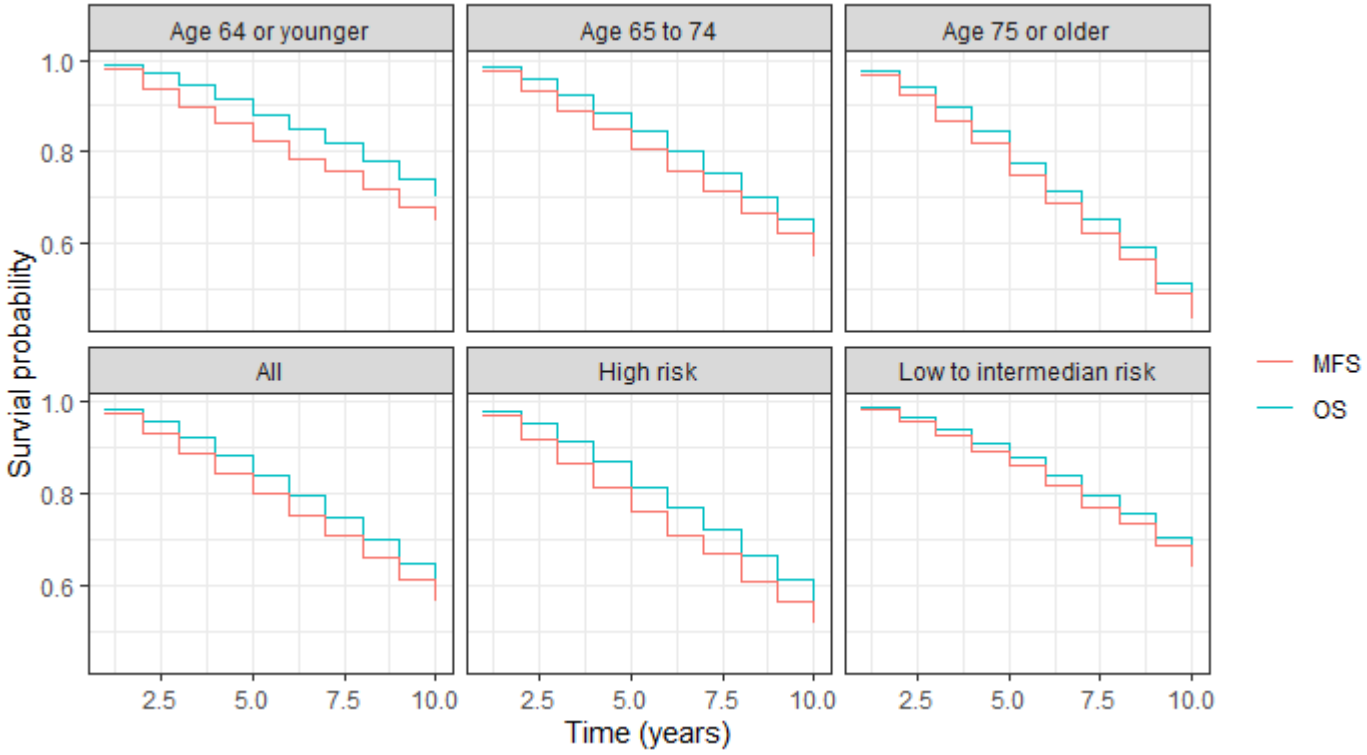
3.5.2.2.1.1. Biological reasoning

The development of metastasis is an inherent aspect of the progression of prostate cancer. Patients who suffer from recurrent disease and have become resistant to second line treatments are known to have a poor prognosis and frequently experience metastasis. This stage of the disease is generally considered the last stage and is associated with a reduction in life expectancy, poor quality of life, and poor prognosis (190).

The precise mechanisms underlying progression and metastasis in prostate cancer are not fully understood. Nonetheless, it is believed that different pathways may play a role. The disease starts with the proliferation of luminal cells confined within the prostate, which eventually evolves into an invasive tumour due to the neoplastic cells' lack of apoptotic ability. The natural course of the disease includes local prostate adenocarcinoma as the cancer cells invade neighbouring structures such as the seminal vesicle. The increasing demand for nutrients by the growing tumour leads to the formation of new vascular networks, which in turn helps the neoplastic cells to drain into adjacent lymph nodes, and eventually, to colonise distant organs, leading to the occurrence of metastasis. The progression of the disease in all colonised organs, if left untreated, eventually leads to death (191).

This relationship between MFS and OS can be seen in the ICECaP study, as demonstrated in Figure 3.4. The biological plausibility previously mentioned implies that patients who develop metastasis are at a higher risk of death, and remaining free of metastasis is therefore associated with a lower risk of death and improved OS.

Figure 3.4. Survival estimates for all patients in ICECaP database



MFS= metastasis free survival, OS= Overall survival

3.5.2.3. Mechanism of actions of interventions in the ICECaP study

The radiation delivery method used in the included studies was similar. Three-dimensional conformal radiotherapy (3D-CRT) was used in eight studies, one study used 3D CRT or IMRT and three used conventional radiotherapy techniques. All but one study (that used 66Gy compared with cryoablation) used doses of RT above 70Gy either in control or intervention arms, in line with recommended clinical guidelines (124) (108). Table 3.1 provides a summary of the mechanisms of action of the treatments evaluated and the population characteristics identified in the studies that were included in the ICECaP analysis for the economic evaluation. The table summarises the key features of the studies that were analysed.

Additionally, the majority of the studies included ADT, an add-on hormonal therapy to RT (Table 3.1) that shares the same mechanism of action of the proposed intervention AAP. The mechanism of action of ADT in prostate cancer has been related to the effect of testosterone on the cancer cells expressing hormone receptors for testosterone. By depriving patients of this male hormone, ADT inhibits the growth of the tumour.

Abiraterone is an androgen that works by blocking the production of androgens, by inhibiting CYP17A1. The enzyme CYP17A1 is involved in the synthesis of androgens from cholesterol, thereby abiraterone reduces the production of androgens in the body, leading to a decrease in the levels of testosterone and other androgens. This reduction in androgen levels has been shown to slow the progression of prostate cancer in some patients.

In addition to its direct inhibitory effect on CYP17A1, abiraterone has been shown to have anti-tumour effects through a variety of other mechanisms, including reducing oxidative stress and inflammation, inhibiting angiogenesis, and inducing apoptosis in prostate cancer cells.

In the case of RT, high-energy radiation is employed to eradicate cancer cells, while androgen deprivation therapy (ADT) restricts the growth of remaining cancer cells. Although these treatment modalities employ distinct approaches, both have the common objective of controlling the spread of cancer cells by destroying the neoplasm and suppressing its growth, thereby improving overall survival (OS). However, in the course of RT and to mitigate the toxic effects on surrounding healthy tissue, some marginal cancer tissue may receive reduced radiation doses and remain alive. This residual tissue can then potentially grow and colonise both local and distant tissues, leading to disease recurrence and the eventual emergence of metastasis. As a result, combining ADT with radiotherapy is expected to defer the onset of recurrence by inhibiting the growth of residual cancer tissue.

The similar mechanism of actions of AAP added to RT (the intervention) compared with the population of ICECaP (RT + ADT) suggests that the impact of the possible intervention on the progression of the disease and the outcome of the patient will be similar regardless of the treatment received. Therefore, the relationship between metastasis and OS is not expected to vary based on the type of therapy received, whether it be RT+ADT or RT+AAP.

In conclusion, the link between metastasis and OS does not differ based on the type of therapy received. Both radiotherapy and surgery share the same mechanism of action, which is to destroy the metastasis and improve OS. This finding highlights the importance of considering the impact of therapy on the progression of the disease and the outcome of the patient when evaluating the effectiveness of different treatments for prostate cancer.

Table 3.1. Characteristics of ICECaP studies included and metastasis definition.

Study	Control Arm	Experimental Arm	Stage	Mechanism of action
Cryo vs RT Ontario trials(192)	NADT 6mo + RT	NADT 6mo + Cryoablation	T2c-3b	Cryoablation: Apoptosis and necrosis of cells caused by intracellular ice crystal formation leading to failure of microcirculation in the ablated tissue following thawing. ADT: inhibits the action of circulating androgens at the receptor level. RT: cellular death caused by damaging the DNA after delivering ionising radiation.
GETUG06(193)	RT 70 Gy	RT 80 Gy	T1b-3a	
MRCRT01(194)	NADT 3-6mo + RT 64 Gy	NADT 3-6mo + RT 74 Gy	T1b-3a	As mentioned above for ADT and RT
EORTC22961(195)	NADT 6mo + RT	NADT 6mo + RT + ADT 2.5yr	T1c-4	As mentioned above for ADT and RT
EORTC22991(189)	RT	RT + AADT 6mo	T1b-2a	As mentioned above for ADT and RT
Mottet et al (196)	ADT 3yr	ADT 3yr + RT	T3-4	As mentioned above for ADT and RT
GICOR-DART01/05(197)	NADT 4mo + RT	NADT 4mo + RT + AADT 2yr	T1c-3b	As mentioned above for ADT and RT
ICORG9701(198)	NADT 4mo + RT	NADT 8mo + RT	T1-4	As mentioned above for ADT and RT
RTOG9202(188)	NADT 4mo + RT	NADT 4mo + RT + AADT 2yr	T2c-4	As mentioned above for ADT and RT
RTOG9408(199)	RT	NADT 4mo + RT	T1b-2b	As mentioned above for ADT and RT
RTOG9413(200)	RT + AADT 4mo	NADT 4mo + RT	T1c-4	As mentioned above for ADT and RT
TROG9601(201)	RT	NADT 3/6 mo + RT	T2b-4	As mentioned above for ADT and RT

Abbreviations: AADT= adjuvant androgen deprivation therapy; ADT= androgen deprivation therapy; mo= month; NADT= neoadjuvant androgen deprivation therapy; RT= radiotherapy.

3.5.2.4. Statistical analysis

Within the overall ICECaP dataset, the potential surrogacy relationships were evaluated using the meta-analytic approach which is considered the gold standard to test surrogacy with IPD from multiple trials, (202) (64) (74). This method requires the fulfilment of two separate conditions, the 'individual level' or 'outcome' and the 'trial level' or 'effect' surrogacy.

Surrogacy at the individual level was explored using bivariate Clayton's copula model using IPD to compare MFS with OS. The fitted model was then quantified by Kendall's Tau estimated from the copula model. Surrogacy at trial level was evaluated using weighted linear regression analysis comparing the survival function estimates of OS at 8 years with MFS at 5 years. For this purpose, a Cox proportional hazards regression model was estimated to obtain the treatment effect measured as the HR on MFS and OS. These methods are explained in Section 3.4.1.1.1.

All analyses described in this section were conducted by the ICECaP data on the full data from the 19 trials included in the ICECaP study. For economic evaluation purposes, all the analysis leading to the validation of MFS as a surrogate outcome of OS were conducted again on the IPD from the 12 trials used for the economic evaluation purposes (see Section 3.4.1.2.2 for the explanation of how this sub population was defined).

3.5.2.4.1. The surrogacy relationship

At the individual level, the results of the bivariate copula model correlating MFS and OS produced an R^2 of 0.91 (95% CI, 0.91 to 0.91) when regressed against OS. As explained in Section 3.4.1.1.1, the result of the regression analysis represents a strong positive relation between outcomes. When the MFS rate at 8 years was regressed against the OS rate at 5 years, the model returned an R^2 of 0.83 (95% CI, 0.71 to 0.88) (Figure 3.3). At the trial level, when the HRs of MFS were regressed against the HRs of OS observed in each trial, the R^2 was 0.92 (95% CI, 0.81 to 0.95) and the resulting equation correlating both outcomes in terms of relative treatment effect was *Log (HR)-OS= -0.021 + 0.740*Log (HR)-MFS.*

The results described above were produced using the full set of trials included in the ICECaP database. When restricted to those trials suitable for the health economics analysis (see section 3.4.1.2.2 for more information) the regression equation found was Log (HR)-OS= -0.0056 + 0.7813* Log (HR)-MFS. Figure 3.3 shows the difference observed in the estimated HR for OS when the full data set of ICECaP studies is used compared with the data for health economic analysis.

Although a variety of oral and localised treatments were given as an adjuvant therapy (see Table 3.1), RT was the main treatment received by all patients. The probabilities of developing recurrent disease, metastasis or die used in the economic model (Chapter 4) were obtained from the results of the meta-analysis.

Finally, the ICECaP study provided strong evidence supporting the use of MFS as a surrogate outcome for OS in LPC, showing a high correlation between changes in MFS and changes in OS. Based on the evaluation of the validation and methodology of the ICECaP study according to the PBAC framework, it can be concluded that the use of MFS as a surrogate outcome in economic evaluations for the reimbursement decision-making process in LPC is compliant with the framework

3.5.3. Survival analysis

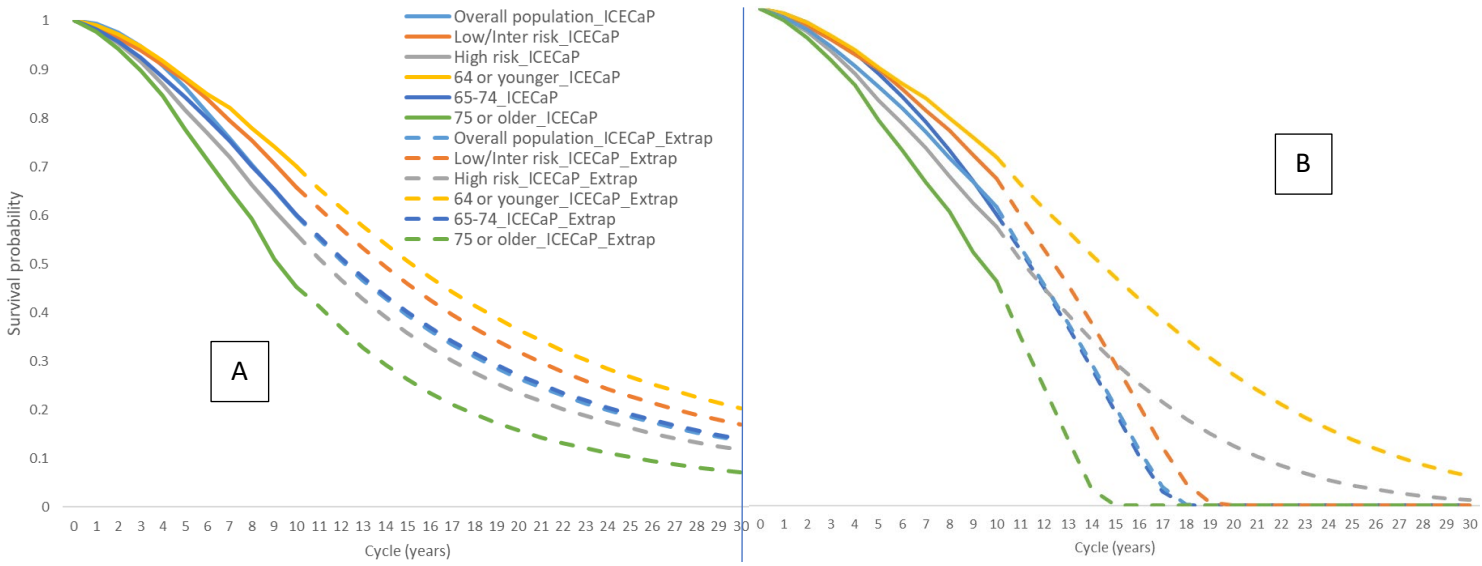
In Chapter 4, the surrogate outcome MFS will be incorporated into an economic model to estimate the lifetime costs and benefits of a new intervention for LPC. Cost-benefit analyses typically utilise lifetime outcomes and costs to comprehensively inform healthcare decision-making. The ICECaP study reported results based on a truncated analysis at 10 years; thus, to estimate costs and benefits over the lifetime of patients, outcomes from the ICECaP were extrapolated. Visual inspection of the log cumulative hazard plots for BCR-free, MFS and OS, showed a monotone hazard function as well as a consistent shape and gradient during the follow-up time (Appendix 4). Therefore, conventional parametric distribution models were evaluated to extrapolate all outcomes of interest in the ICECaP data (203) (180).

The next sections present the extrapolation of all sub-populations described in the ICECaP analysis to exemplify the steps involved in the choice of extrapolation model.

3.5.3.1. Overall survival

In nearly all the populations analysed, the parametric model chosen was the generalised gamma except for the high-risk sub-population and those aged 64 or younger where the Weibull was the best fit. Figure 3.5 (B) shows the ICECaP data in solid lines and the corresponding extrapolations in dashed lines following the AIC and BIC criteria.

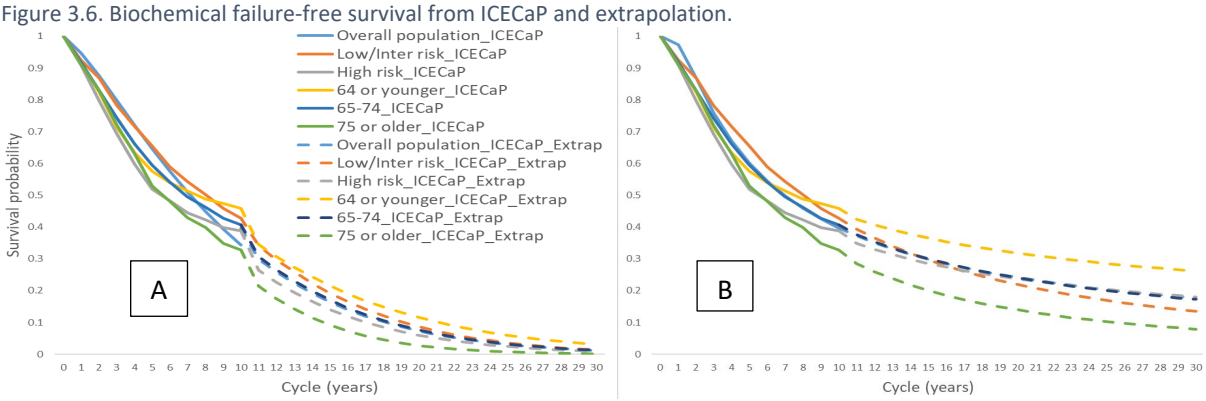
Figure 3.5. Overall survival from ICECaP and extrapolation.



A: Extrapolation of OS following AIC and BIC criteria (all population analysed were parametrised in the extrapolation portion following the generalised gamma except for the high risk and the 64 or younger population where the Weibull fit the data the best. B: Extrapolation of OS where all population analysed were parametrised in the extrapolation portion following the Loglogistic distribution.

The distribution chosen in most of the populations evaluated showed a rapid decline in the survival function which was not expected, where the state membership in BCR-free accounts for a high proportion of patients at all times (Figure 3.6). Moreover, crossing of the curves in the extrapolation portion translated into inconsistencies where patients in low to intermediate risk groups (orange dashed line in Figure 3.5B) had a higher hazard of death than high-risk patients (grey dashed line in Figure 3.5B). Therefore, the log-logistic distribution was chosen for all populations as it showed consistency with the survival probabilities estimated among all subgroups analysed and a clinically expected shape (see Figure 3.5A).

Figure 3.6. Biochemical failure-free survival from ICECaP and extrapolation.



A: Extrapolation of BCR-free where all population analysed were parametrised in the extrapolation portion following the Loglogistic distribution. B: Extrapolation of BCR-free following AIC and BIC criteria (all population analysed were parametrised in the extrapolation portion following the Weibull distribution.

3.5.3.2. Metastasis-free survival

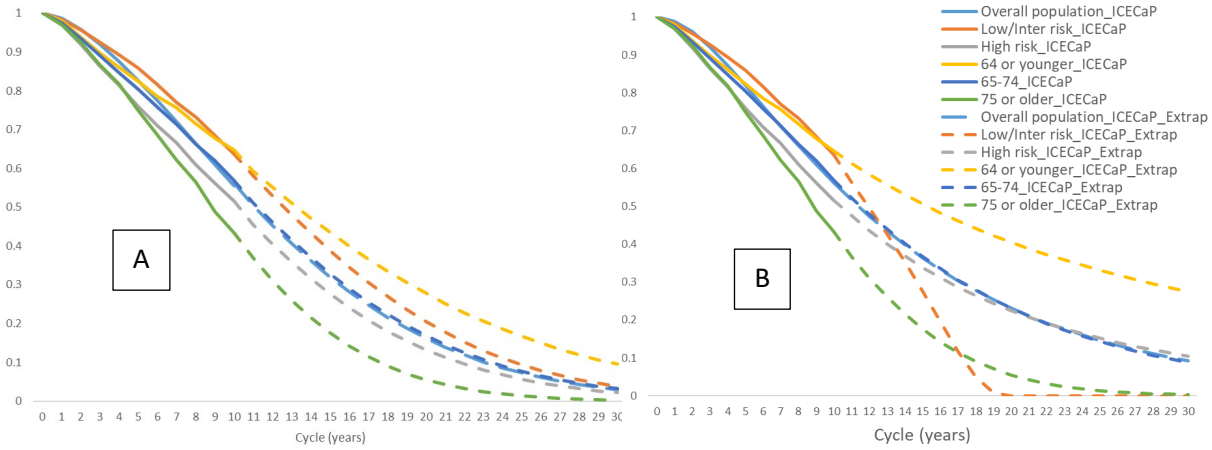
Generalised gamma showed the best fit of the ICECaP data for MFS in most of the populations analysed, log-normal and Weibull were the best fit for groups aged 64 or younger and 75 or older, respectively. Figure 3.7B shows the ICECaP data and the corresponding extrapolation for all populations analysed following the AIC and BIC.

The distributions chosen showed inconsistencies in most of the populations evaluated when compared to OS curves. This was also observed when MFS curves were compared among different populations in the analysis, where patients with low to intermediate risk (orange dashed line in Figure 3.7B), who are expected to have better comparative MFS probabilities, showed a sharp decline in survival function with higher hazard compared with high-risk patients (grey dashed line in Figure 3.7B),

15 years after diagnosis. On the other hand, the long tail observed for patients aged 64 years or younger is clinically implausible as it predicts a 30% chance of being free of metastasis 30 years after diagnosis.

Therefore, the Weibull distribution, which estimated clinically acceptable probabilities and comparable shapes among all populations analysed, was chosen to extrapolate MFS. Figure 3.7A shows the results of the distributions chosen for MFS.

Figure 3.7. Metastasis-free survival from ICECaP and extrapolation.



A: Extrapolation of MFS where all population analysed were parametrised in the extrapolation portion following the Weibull distribution. B: Extrapolation of MFS following AIC and BIC criteria (all population analysed were parametrised in the extrapolation portion following the generalised gamma except for the 75 or older and the 64 or younger population where the Lognormal and Weibull were the best fit to the ICECaP data.

3.5.3.3. Biochemical recurrence-free survival

For patients in the BCR-free state, regardless of the risk level or age group, the generalised gamma distribution showed the best fit for the data. However, the extrapolations following the generalised gamma distribution in this population resulted in abnormal tails, which were clinically implausible when compared to other outcomes analysed (Figure 3.6B). Therefore, the Weibull distribution was chosen for the base case analysis as it showed consistent comparative results with the other outcomes analysed and more conservative hazards that shortened the length of the tail. Figure 3.6B shows the ICECaP and corresponding extrapolations following the AIC and BIC criteria while (A) shows the results using the Weibull distribution for all the populations analysed.

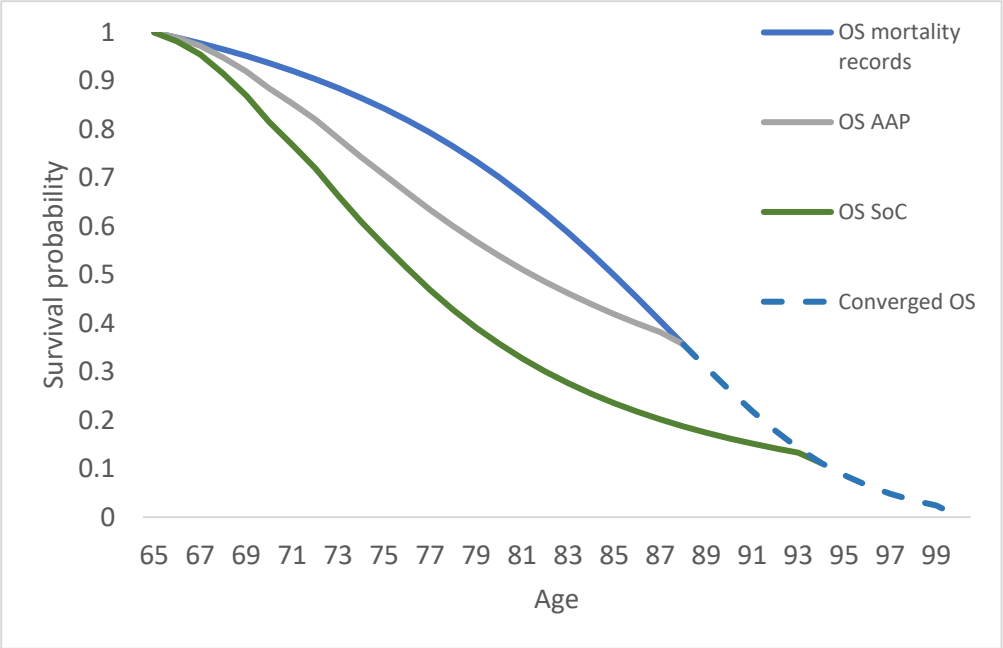
The results showing the log cumulative hazard plots, parametric models evaluated and their respective information criteria for the overall population are provided in Appendix 4.

3.5.3.4. Adjustment of the overall survival curve for population mortality

Although the ICECaP data can be considered mature, it has been postulated that the hazard function carried over the extrapolated period is influenced by disease specific mortality observed during the within-trial period (181) (182). The extrapolation of the hazard function observed within-trial does not reflect the increasing hazard of death observed in the overall population at advanced ages and therefore, the OS curves needs to be adjusted to reflect general population mortality(182). The adjustment of the survival curve using external data from population life tables is expected to reduce the possible inconsistencies related to the structural independence of outcomes in the PSM and increase the robustness of the results.

Figure 3.8 shows the survival curves of the AAP and SoC patients when the overall population mortality records in Australia (184) were used to adjust the hazard function of the population in the model in the long term. The survival curves of both AAP and SoC converge towards the survival curve of the mortality records in Australia in the tail of the extrapolation portion of the curves, with convergence occurring at different times. As the surrogacy relationship estimates significant survival benefits for patients under AAP treatment, the hazard of this population is lower compared with patients receiving SoC only and convergence occurs 6 years before the comparator arm.

Figure 3.8. Adjustment of survival using the overall population mortality records in the model.

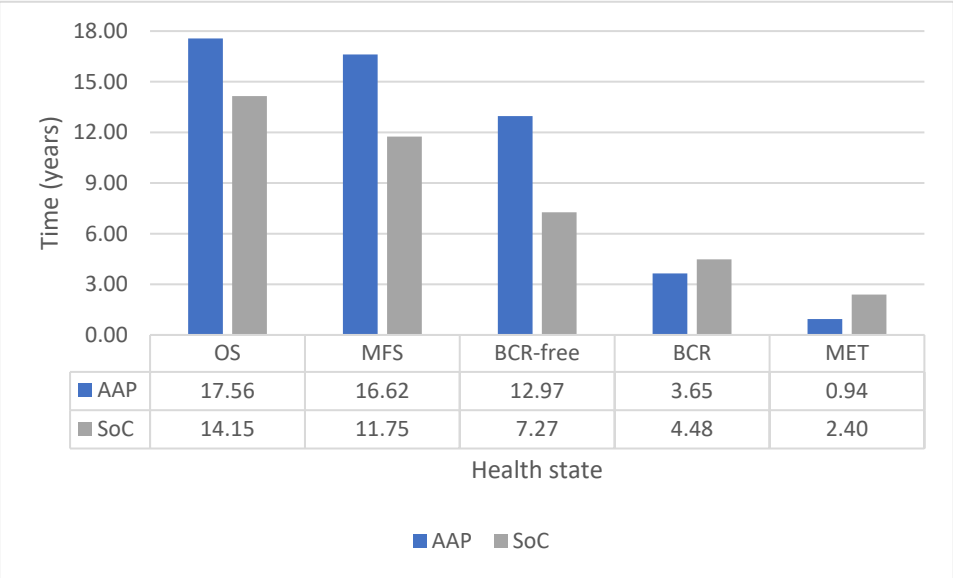


Abbreviations: AAP= abiraterone + prednisolone; OS= overall survival; SoC= standard of care.

3.5.3.5. Mean survival time

The time spent in each health state estimated by the survival curves is shown in Figure 3.9. Patients spent most time free of metastasis followed by the BCR-free state in both interventions compared. Patients who received SoC spent approximately seven additional years in progressed states (BCR and metastasis) when compared to patients who received AAP. The largest difference in time spent in a heath state in the model between AAP and SoC was for patients free of recurrence; those receiving AAP spent on average an additional six years in the BCR-free state compared with those treated with SoC.

Figure 3.9. Mean survival time of all health states in the model by population.



Abbreviations: AAP= abiraterone plus prednisolone; BCR= biochemical failure; MET= metastasis; MFS= metastasis-free survival; OS= overall survival; SoC= standard of care.

3.6 Discussion

The results of the ICECaP study regarding the correlation between MFS and OS have been presented and an assessment of MFS as a surrogate for OS for patients with localised prostate cancer conducted according to the PBAC guideless. The current PBAC framework for evaluating the use of surrogate outcomes does not include a quantitative threshold analysis for the outputs of linear regression, instead relying on a qualitative approach. However, the results of the surrogacy relationship analysis of MFS demonstrated a strong correlation, both at the individual and trial levels, with R-squared values exceeding 0.9. Such a high level of association, where R-squared values are above 0.9, have been deemed acceptable according to other agencies and studies within the field (65). The results of the assessment of MFS as a surrogate outcome of OS to be used to support economic evaluation

according to the PBAC guidelines, are also in accordance with the guidance from the FDA which has approved the surrogacy relationship of MFS with OS for regulatory purposes in a similar population (patients with nonmetastatic castrate-resistant prostate cancer) (204).

As expected, the specific surrogate equation is influenced by the data used to derive the surrogacy relationship between MFS and OS. While there are slight variations in the results when a subset of the data for economic evaluation and the values predicted when the full dataset is used to estimate OS, the results are marginally different mainly for values of MFS-HR above the STE. However, these variations become negligible and quite similar when a clinically reasonable range of HR for MFS is applied. Further details can be found in Appendix 1 (mainly for values above the STE), they are similar when a clinically plausible range of HR for MFS is used (more information in Appendix 1). This shows that the estimates of OS by means of the surrogacy relationship are robust, as similar comparative results were observed when a subset of patients who were mainly given RT as their primary treatment (comprising the health economic analysis set described in Section 3.4.1.2.2), was used to estimate the surrogacy relationship.

3.6.1. The ICECaP trials and biological reasoning

Although the RCTs included in the ICECaP meta-analysis were mainly radiotherapy-based, the data set included results from other treatments with different mechanisms of action. ADT was used in some form by all except two studies, one that compared different RT schedules and one that compared cryosurgery techniques. If the effect of the surrogate outcome compared with the final outcome is consistent over multiple interventions with different mechanism of actions, it is likely that the surrogate predicts the final outcome regardless of the intervention studied. A similar result is predicted when a new intervention to be evaluated has a similar mechanism of action compared to the interventions used in the studies where the surrogate relationship was assessed.

In addition, the long follow-up time of the trials used in the ICECaP analysis provides a robust and accurate evaluation of the surrogate outcome. The maximisation of follow-up time for OS in clinical trials has a significant impact on the use of surrogate outcomes in modelled economic evaluations. This approach can lead to reliable predictions of treatment effects and ultimately better decisions regarding the reimbursement of treatments and interventions that rely less on survival estimates from parametric extrapolations. By maximising the follow-up time for OS, the use of surrogate outcomes can be greatly enhanced in modelled economic evaluations.

3.6.2. Statistical analysis of the surrogacy relationship

The ICECap study used a bivariate meta-analytical method to predict the treatment effect on OS based on the pooled treatment effect from multiple trials. Although this meta-analytical approach is recommended by the PBAC guidelines and in general by HTA agencies that provide detailed information about the validation framework (65), new variations of the meta-analytical approach have been developed (160). In a recent update of the framework for evaluation of surrogate outcomes published by the National Institute for Health and Care Excellence (NICE), the Bayesian multivariate approach has been suggested (65).

The Bayesian multivariate meta-analysis and bivariate meta-analytical methods are similar as they are both rooted in the use of a meta-analytical approach but differ in the way they analyse and integrate data from multiple studies. Bivariate meta-analysis combines data from multiple studies to estimate the correlation between a surrogate endpoint and a clinical endpoint of interest. It provides a summary of the strength and direction of the association between the two variables but does not account for the uncertainty or variability across studies (160).

Bayesian multivariate meta-analysis, on the other hand, extends the bivariate meta-analysis method to account for multiple surrogates and multiple clinical endpoints (160). It uses Bayesian statistical methods to incorporate prior information and to provide robust and flexible estimates of the associations between surrogates and clinical endpoints. The approach allows for the modelling of complex relationships between multiple variables and for the quantification of uncertainty and variability across studies. However, the Bayesian method is still not widely recommended and has limitations when compared to bivariate meta-analysis:

- Complexity: The Bayesian multivariate meta-analysis approach can be more complex compared to traditional meta-analysis methods, requiring a more advanced understanding of Bayesian statistics and specialised software for implementation. The bivariate meta-analytical method is relatively simple compared to other meta-analysis methods, making it easier to understand and implement for those with limited statistical background.
- Prior information: The Bayesian approach requires the specification of prior information, which can be challenging, particularly if little prior information is available. This can impact the validity of the results and the estimation of uncertainty. The bivariate meta-analytical method is robust to small sample sizes and is not sensitive to the distributional assumptions underlying the data.

- Model misspecification: The Bayesian approach is sensitive to model misspecification, which can lead to biased results if the model is not correctly specified. The bivariate meta-analytical method provides a straightforward summary of the strength and direction of the association between two variables, which can be easily interpreted and communicated.

3.6.3. Results of the survival analysis

The results of the survival analysis produced in this chapter indicate that the use of MFS as a surrogate outcome for OS has limitations when proportional hazards are assumed using extreme values. This could be explained by the boundaries of the data used in the ICECaP analysis, which included studies reporting HRs for MFS and overall survival between 0.3 and 1.6 and between 0.5 and 1.4 respectively. Additionally, trials have been historically designed with an OS HR that ranges from 0.71 to 0.75 (132) therefore, it is unlikely that the surrogacy relationship would be applied on results from trials with extreme values of HR.

The survival analysis conducted was based on a proportional hazards assumption to estimate the survival probabilities of the intervention. Although survival analysis based on proportional hazards can provide valuable insights into the treatment effects on survival outcomes, it is important to carefully consider its limitations when interpreting the results (203):

- Proportional hazards assumption: The proportional hazards assumption states that the hazard ratio between the treatment groups remains constant over time. This assumption may not hold in practice, especially in the presence of time-varying confounding factors.
- Censoring: Censoring, or missing data, can lead to bias in the estimation of the hazard functions and the hazard ratios, especially if the censoring is not random.
- Model specification: The choice of parametric or semi-parametric model for the hazard functions may have an impact on the results. Different models may lead to different conclusions about the relative effectiveness of the treatments.
- Small sample size: In small sample sizes, the estimation of the hazard functions and hazard ratios may be imprecise, leading to unreliable results.
- Time-fixed covariates: If the covariates are time-fixed, i.e., they do not change over time, they cannot be used to adjust for time-varying confounding factors, which may lead to bias in the results.

In the context of survival analysis, it is crucial to consider the parametric distributions used to model the outcomes. These distributions are typically estimated based on the hazards observed over the within-trial period. In some cases, it may be reasonable to justify the use of different parametric

distributions for the same outcome for an intervention and comparator, for instance, when the underlying treatment effect of the intervention changes the natural history of the disease. However, care should be taken when choosing different distributions, as this could lead to an incorrect conclusion that one intervention is favoured over the other. In the survival analysis conducted, the extrapolations of the ICECaP outcomes for intervention and comparator were assumed to follow the same distribution, as the underlying mechanism of action was similar.

In survival analysis, it is widely accepted that the hazard function carried over the extrapolated period may be influenced by disease-specific mortality observed during the within-trial period (182). As such, relying on the extrapolation estimates may not fully capture the potential hazard of other causes of death observed in the general population, which could lead to an overestimate of survival associated with new interventions (181) (182). To address this issue, the model adjusted the survival curves using external data from Australia population life tables (182). This adjustment of the survival probabilities in the extrapolated portion is expected to address any inconsistencies arising from the structural independence of outcomes in model and increase the robustness of the results while providing a more comprehensive assessment of the disease-specific hazard, and thereby improve the accuracy of the extrapolated survival probabilities.

3.7 Conclusions

In conclusion, the ICECaP study provides strong evidence supporting the use of MFS as a surrogate outcome for OS in localised prostate cancer. The results of the study demonstrate a high correlation between changes in MFS and changes in OS. The validation and its methodology were evaluated according to the PBAC framework, a well-established framework providing detailed information to assess the use of surrogate outcome in economic evaluations to support reimbursement decisions. The results of the ICECaP study and the assessment conducted following the PBAC framework support the use of MFS as a surrogate outcome of OS in economic evaluations for the reimbursement decision-making process in LPC. The study not only validates the surrogacy relationship between MFS and OS but also provides crucial evidence of treatment effect, which is critical for conducting an economic evaluation.

Chapter 4 Incorporating a surrogate outcome in an economic evaluation.

4.1 Abstract

A partitioned survival model incorporating the surrogacy relationship in ICECaP assessed in the previous chapter was built. The model compared the costs and benefits of AAP and SoC for the treatment of high-risk LPC as a case study. The model predicted a QALY gain of 8.42 for AAP, and 7.09 for SoC, with an ICER of $39,640 per QALY. The sensitivity analysis indicated that the model was most sensitive to the cost of AAP and the time on treatment, rather than the comparative treatment effect estimated though the surrogacy relationship. This result implies that the use of a validated surrogate outcome can inform an economic model used in the reimbursement decision-making process for new interventions in LPC.

Additionally, the model was adapted to analyse the benefits to society arising from an early reimbursement decision based on MFS as a surrogate outcome of OS. The benefit of an earlier decision implied that more patients will be BCR-free. Delaying the decision by 3 years to allow for the observation of OS would lead to differences in costs and benefits as more patients would commence treatment in an advanced stage of disease, rather than in the BCR-free state. The average cost per patient over a lifetime is $11,558 higher, and 4.48 life years and 0.47 QALYs are lost when the reimbursement-decision is delayed for 3 years. The results of the analysis of benefits to society show that most of the benefits gained from an earlier reimbursement decision occur due to patients being kept free from biochemical relapse. The largest cost impact is derived from the BCR state, while the greatest impact on QALYs comes from the BCR-free state.

The analysis shows that AAP produces short-term benefits in the first 3 years of treatment for LPC, and an earlier reimbursement decision can result in long-term benefits. The study also shows that AAP is a cost-effective treatment for LPC. The expected value of information for the cohort population is high, implying that conducting further studies to confirm the surrogacy relationship is not economically viable.

Importantly, MFS as a surrogate for OS was incorporated into a partitioned survival model to potentially facilitate reimbursement decisions for novel interventions in high-risk LPC patients in Australia. Incorporating the surrogate endpoint into an economic model can allow for earlier reimbursement decisions, leading to improved patient outcomes through timely access to effective treatments. This could potentially translate into a greater number of patients experiencing prolonged

progression-free survival, reduced costs associated with treating recurrence or metastasis, improved quality of life and, ultimately, better survival outcomes. The anticipated societal value presented in this case study outweighs the costs associated with additional research needed to address the surrogacy relationship's uncertainty.

4.2 Introduction

The systematic review of the literature presented in Chapter 2 revealed a lack of evidence supporting the use of surrogate outcomes in economic evaluations for new interventions for LPC. Although one publication was identified, it had multiple limitations that made it difficult to draw conclusions regarding how surrogate outcomes (as defined by The Biomarkers Definitions Working Group (62) and discussed in Chapter 3) are being applied in economic evaluations in prostate cancer, particularly those that could be used to support reimbursement decisions. The absence of validated surrogate outcomes supported the need for further research in the area.

The results of the ICECaP study demonstrated that 5yr-MFS is a strong surrogate outcome of OS in patients with LPC (132). The potential impact on the timing for interventions to become available for patients when reimbursement decisions are based on surrogate outcomes, together with the clinical need for new interventions in LPC, motivated the development of an economic model incorporating the surrogacy relationship to assess the use of MFS as a surrogate outcome in this population to support reimbursement decisions (see Chapter 3). The evidence for the surrogacy relationship of MFS for OS was assessed according to the PBAC requirements for the use of surrogate outcomes in economic evaluations in Australia. The results showed that the use of MFS satisfied the requirements of the surrogate framework established by PBAC for informing economic evaluations undertaken to support reimbursement decision for new interventions for LPC in Australia.

In this chapter, a flexible economic model will be built using MFS to predict OS to compare current treatment options for high-risk patients diagnosed with LPC – radiotherapy (RT) added to androgen deprivation therapy (ADT) and abiraterone plus prednisolone (AAP) – as a case study. The model was intended to reflect the Australian clinical practice as a case study but was designed to be sufficiently flexible to incorporate other interventions and clinical settings. The results of the model will identify the possible benefits to society of using surrogate outcomes to support reimbursement decisions that allow treatment to commence earlier than decisions based on final outcome.

4.2.1.1. Economic evaluation and surrogate outcomes

The use of surrogate outcomes in economic models in healthcare is typically not related to the ability of the surrogate outcome to accurately predict the final outcome. When used, economic evaluations have historically focused on models that incorporate these surrogate outcomes as the primary clinical outcome metric – for example, cost-effectiveness models that report results in terms of the cost per unit of reduction in blood pressure achieved through intervention for heart disease, rather than OS (205). Less attention has been paid to the predictive ability of surrogate outcomes over the final outcomes to inform the conduct of economic evaluations. When the final outcome of interest is overall survival, the use of surrogate outcomes enables cost utility analysis, which can incorporate the use of generic outcomes such as quality of life, thereby facilitating earlier comparison of new interventions across different healthcare areas.

The use of surrogate outcomes in economic evaluations performed to inform reimbursement decisions has several benefits. Surrogate outcomes can be measured more easily and quickly than the final outcome, reducing the time and cost associated with conducting a study. Surrogate outcomes can also reduce the costs associated with conducting clinical trials, as the time required to collect data can be significantly reduced. This can result in a significant reduction in the overall costs as it eliminates the need for extended patient follow-up, reduces the need for expensive imaging or laboratory tests, and limits the costs of personnel and supplies (141).

4.2.2. Treatment of prostate cancer

4.2.3. Local treatment

As in many chronic diseases, treatment options for patients diagnosed with prostate cancer depend on disease assessment, morbidity, treatment preferences and the availability of resources. For patients categorised as low risk, active surveillance or watchful waiting (no active treatment) is generally the most common course. This treatment option has shown benefits in a selected group of patients with a low probability of developing advanced stages of the disease and who are willing to accept living with cancer (107).

On the other hand, in patients categorised as intermediate or high risk, active treatment is the usual standard of care given the higher probability of disease progression. If the patient and physician opt for active treatment, there are two main options currently available in Australia: (1) prostatectomy, with or without radiotherapy; and (2) radiotherapy, with or without androgen deprivation therapy

(ADT) (108). The clinical algorithm for the treatment of localised prostate cancer is depicted in Figure 4.1.

Figure 4.1. Clinical algorithm of localised prostate cancer treatment.

Source: Philippou et al. 2014 (206)
Abbreviations: CT= computed tomography; EBRT= external bean radiotherapy; HIFU= high-intensity focused ultrasound; MDT= multidisciplinary team; MRI= Magnetic resonance imaging; PSA= prostate specific antigen; TNM= tumour, node, metastasis staging system.

Methods of delivering ionising radiation to a target area have changed progressively over the past two decades. The integration of imaging and the use of highly conformal radiotherapy techniques have resulted in more effective delivery of the radiation through escalating doses directed to the target while minimising toxicity on healthy tissue and other organs (109). Currently, clinical guidelines recommend radiotherapy to the primary tumour through the use of external beam radiation therapy (EBRT) or internal radiation therapy (brachytherapy) (104, 108).

In EBRT, radiation is generated by an external X-ray source and delivered to the primary tumour. Highly conformal radiotherapy techniques are recommended in order to concentrate higher doses of radiation in the tumour and spare surrounding tissue. To achieve this, intensity-modulated radiotherapy (IMRT) techniques have been developed whereby the tumour is approached from additional angles and radiation beams are shaped based on individual patient's requirements. This allows a high-dose area around the target, leading to higher probabilities of cure with similar or slightly higher toxicity.

Brachytherapy can be delivered by the implant of permanent radioactive seeds into the prostate, known as low dose brachytherapy or by the temporary insertion of highly radioactive seeds into the prostate for a limited period, known as high dose brachytherapy.

Surgical techniques have also evolved; however, the impact of these changes on relevant clinical outcomes is less clear compared to the improvements observed with radiotherapy. Radical prostatectomy can be performed as open surgery, open radical prostatectomy (ORP), through a laparoscopic approach (LRP), or by means of robotic-assisted surgery (RARP) (108). Although the advent of laparoscopic and robot-assisted radical prostatectomy promised better short- and long-term outcomes, fewer surgery-related adverse events and better quality of life compared to open prostatectomy, the results reported from clinical trials have yielded similar results to ORP (207).

4.2.3.1. Hormone-focused therapy and advanced disease

In Australia, around 26% of patients are at high- risk at diagnosis (208). Similar to low and intermediate risk stages, primary treatment with RT or RP is recommended for these patients;

however, due to the extension of their disease or elevated PSA level, adjuvant treatment[19] in the form of hormone therapy, also known as androgen deprivation therapy (ADT), is usually prescribed.

The deprivation of androgen hormones in a patient can be achieved by surgery (removing the testicles, where androgens are synthesised) or by the use of hormonal drug treatment. Although surgical castration has been identified as an inexpensive and effective method of depriving the prostate of androgen hormones (209), it is the less popular choice due to the psychological impact of surgical castration and the irreversible nature of, and adverse events associated with, this procedure (210). Therefore, surgical castration was not considered in this analysis and ADT is only considered as medical castration.

Medical castration using an agonist or an antagonist of the luteinising hormone-releasing hormone (LHRH), has been the standard practice in developed countries (211). Drug treatment targets different biological processes that ultimately impact on the effect of hormones on the cancer cells. In hormone-sensitive prostate cancers, ADT lowers the amount of androgen hormones available to the tumour cells by preventing androgen hormone secretion or preventing them from binding to its receptor (110). In hormone-sensitive LPC, tumour growth has a strong relationship with the activation of androgen receptors in the tumour cells. Medical castration can be achieved with a gonadotropin-releasing hormone (GnRH) agonist or a GnRH antagonist. GnRH antagonists inhibit the production of luteinising hormone (LH) at the pituitary level and, consequently, the synthesis of testicular testosterone. On the other hand, LH agonists stimulate the release of LH at the pituitary level and, therefore, an immediate synthesis of testicular testosterone. The high affinity of the GnRH agonist for its receptor causes a downregulation of GnRH receptors, leading to a decline in the pituitary production of LH. Testicular testosterone synthesis is then reduced to castrate levels within weeks of the start of treatment and is maintained while the drug is administered. Due to the initial impact of GnRH agonists on testosterone levels (known as flare phenomenon), an antiandrogen therapy, which blocks the effect of the increased testosterone, is usually given during the first month of GnRH agonist therapy. Increasingly, an alternative approach involves the use of a GnRH antagonist, such as degarelix, which is favoured due to its lack of an initial flare.

[19] Adjuvant treatment, in the context of patients receiving primary treatment for localised prostate cancer, refers to additional therapy given in combination with the initial treatment (such as surgery or radiation) in order to improve the chance of curing the cancer or preventing its recurrence. Adjuvant treatment is usually recommended for patients with a high risk of disease recurrence based on certain pathological and clinical features of their cancer.

While ADT has been shown to be effective when applied in adjuvant settings together with RT or RP in locally advanced diseases, it has a significant impact on quality of life. Hot flashes, loss of libido, erectile dysfunction, depression, loss of muscle mass, anaemia and osteoporosis are among the most frequent adverse events observed in patients receiving ADT (212).

Medical androgen deprivation is often prescribed as an intermittent treatment that helps to ameliorate the occurrence of adverse events related to hormonal treatment. This approach involves either a fixed period of time on-and-off treatment, or a PSA response-driven treatment duration, in which ADT is withdrawn once PSA reaches a required nadir, only to be reinitiated after the PSA level starts to rise. The evidence in the literature regarding intermittent ADT regimens is diverse and there is no consensus (111). In general, intermittent ADT regimens consist of an induction period where castration levels of testosterone are achieved followed by an off-treatment time during which testosterone levels recover. The adverse effects associated with the ADT and low levels of androgen tend to decline when the treatment is withdrawn, and intermittent treatment is therefore associated with gains in quality of life. In a recent review of ADT treatment, the time-of treatment varied from 3 to 8 months (median 6 months) whereas the time-off treatment varied among cycles from 15 to 33 weeks (median 24 weeks) (111).

The response to ADT for patients who develop recurrent diseases after initial treatment is generally good. Virtually all patients with hormone-sensitive disease will decrease their androgen hormones to castration levels following ADT therapy. Eventually, irrespective of hormone treatment, most men with advanced disease will ultimately have elevated PSA levels, develop new symptoms or clinical evidence of disease progression, and be diagnosed with castrate-resistant prostate cancer (213). Due to the poor prognosis after progression on ADT, patients' age, physical and psychological status and quality of life under systemic and more aggressive therapies, the treatment of patients for whom primary treatment fails represents a challenge in clinical practice (214).

For patients that develop resistance to castrate levels of androgen hormones, ADT is generally maintained and given in a continuous schedule (113). New treatments with anti-androgens previously reserved for more advanced states of the diseases (abiraterone, apalutamide, darolutamide and enzalutamide) have been tested, approved and recently recommended in clinical guidelines for recurrent patients (114) and PBS-listed in Australia for use in the castration-resistant setting for non-metastatic patients.

4.2.3.2. Chemotherapy

Colonisation of distant organs will occur in most men, and metastases develop. Secondary hormonal manipulation with first-generation antiandrogens, such as bicalutamide and nilutamide, or corticosteroids in combination with chemotherapy have been the main treatment options for these patients. Over the last two decades treatment recommendations for those with distant organ involvement have evolved, and these patients can now be offered different systemic treatment options (115). Taxanes are the only cytotoxic chemotherapy agents shown to have survival benefits in patients with metastatic disease (116). In particular, docetaxel has been added to ADT for newly diagnosed metastatic patients, with significant improvements in OS and failure-free survival (215).

However, patients can develop resistance to docetaxel and ADT and will then require more aggressive treatment. In this setting, patients can receive additional chemotherapy, with a different taxane, second-generation antiandrogens, radium-223 for symptomatic palliation in men with bone metastases, inhibitors of the synthesis of antiandrogen precursor and bone-targeted therapy all recommended for this stage of the disease (108).

4.2.4. Imaging technique and monitoring

The staging of patients with high-risk disease or those with clinical signs of progression relies on imaging techniques to identify disease spread. Given the affinity of prostate cancer to metastasise in the bones, CT and bone scans are the gold standard imaging techniques to identify spread of disease. Although new imaging techniques (like GaPSMA-PT/CT) have proved more specific and sensitive than traditional CT + bone scans (185), conventional imaging is still recommended in clinical guidelines to assess and stage patients with prostate cancer (124) (108).

Different stages of the disease require different monitoring schedules. More advanced stages, where the risk of progression or metastasis is higher, require a close follow-up of clinical symptoms and imaging that enables timely treatment to prevent worsening of quality of life or a potential life-threatening condition.

After primary treatment, low-risk patients do not require routine imaging as follow-up; rather, they undergo regular PSA and digital rectal examination as the recommended follow-up (104). Where the risk of metastasis is high, follow-up imaging is needed to assess possible dissemination to distant organs. For example, testosterone levels should be monitored during routine follow-up in high-risk

patients and in patients with metastatic disease, while bone scans are recommended as routine follow-up to identify possible bone lesions at risk of fracture.

Metastatic patients require not only the use of imaging during routine follow-up, but also general tests such as assessment of renal and hepatic function. Adverse effects associated with the use of ADT plus concomitant treatment with cytotoxic agents also require assessment of renal and hepatic function during the follow-up routine (104).

As mentioned above, the objective of ADT is to deprive patients, and consequently the prostate tumour, of androgen hormones. Nevertheless, there are patients who do not respond to treatment and develop hormone-insensitive prostate cancer. Additionally, some patients may experience a temporary surge in testosterone during treatment with LHRH agonist. Therefore, testosterone levels should be monitored during routine follow-up (104).

4.2.5. Health-related quality of life in patients with prostate cancer

Decision-makers prefer the benefits of new interventions to be measured in meaningful outcomes for patients. Life expectancy and quality of life are arguably the most important outcomes for patients and clinicians alike. The use of generic measures in the reimbursement decision-making process allows the comparison of different interventions in different therapeutic areas. In this regard, the QALY, a composite outcome of the length of life and quality of life, has become popular in economic evaluations in HTA due to its versatility in quantifying benefits across multiple healthcare areas (76).

For the purposes of QALY assessment, quality of life is measured as a 'utility' value in which preferences for one health state over another are quantified. Conventionally, utility values are measured on a scale between 0 for death and 1 for full health, although some authors have established that for some individuals, there are some health states that are perceived to be worse than death (76). The measurement of quality of life is crucial for the evaluation of treatment outcomes. Generic health state instruments, such as EQ-5D, SF-36, and HUI,[20] are frequently used in

[20] Generic health state instruments are standardised questionnaires or surveys used to measure the health-related quality of life of individuals. These instruments are designed to capture the impact of health conditions and interventions across different diseases or health conditions, without focusing on specific symptoms. Some commonly used generic health state instruments include:

- EQ-5D (EuroQol 5 dimensions): This instrument assesses five dimensions of health: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. It also includes a visual analogue scale (VAS) where respondents rate their overall health status.

economic evaluations. However, these instruments have limitations in fully capturing the impact of prostate cancer treatment on urinary, bowel and sexual function, which are significant factors that define the quality of life for these patients (216). To address these limitations, specific utility instruments for prostate cancer, such as PORPUS, have been designed; they are more widely used than generic measures (217). The results obtained from a specific quality-of-life instrument can be converted into a generic preference-based measure though the use of mapping methodologies (218). This allows for a common and comparable measure of quality of life to be used for decision-making purposes. Moreover, the validation of the specific measures in prostate cancer has enabled the translation of results across different methodologies used to capture a patient's health status accurately (219).

4.2.6. Costs

Cost analysis is an essential part of economic evaluations. The perspective of the analysis determines which costs should be measured in an economic evaluation. Which perspective to choose will depend on the scope of the analysis but also on the approach recommended by the relevant HTA agency. When defining the perspective, it is important to consider the importance of different cost components associated with the intervention assessed and the decision problem (76). The most common perspectives are the societal and the healthcare system; of these, the broadest is the societal, where direct and indirect costs incurred by the overall population are identified and measured in the analysis (76). Direct costs reflect healthcare resources used in the delivery of an intervention and can also include costs borne by the patient, such as out-of-pocket expenses. Indirect costs include the cost of time of patients (or their families/carers) or productivity loss due to illness (or due to caring responsibilities).

The first step in analysing the cost is to identify the resources used for the delivery of the intervention and the comparator. The next step is to measure and determine the quantity of each resource needed and its frequency of use. Finally, each resource must be valued (a price attached per unit of resource)

- SF-36 (Short-Form 36): This questionnaire measures eight domains of health: physical functioning, role limitations due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitations due to emotional problems, and mental health.
- HUI (Health Utilities Index): This instrument measures eight attributes of health: vision, hearing, speech, ambulation, dexterity, emotion, cognition, and pain. It also includes a single summary score that represents overall HRQoL.

using relevant sources that are specific to the context in which the reimbursement decision is being made.

The analysis approach depends on available data and time constraints. The most precise method is micro-costing, individually tracking and valuing every applicable resource. In contrast, macro-costing uses averages to account for entire categories. Other options include process-based costing that breaks the intervention down into component steps. The approach should flexibly accommodate the full scope of potential costs rather than focusing solely on specific categories of cost.

4.2.7. Time horizon and the need for extrapolation

Financial constraints often guide the design of clinical trials to obtain results in the shortest possible time. On the other hand, decision-makers – and, therefore, economic evaluations to support the process – need to account for the costs and benefits in the long term. Therefore, extrapolation of outcomes and healthcare exposure beyond the time horizon evaluated in a clinical trial is common practice in HTA (174).

Although the ICECaP data contains information on outcomes over a 10-year follow-up period, an economic evaluation for the purposes of decision-making requires the quantification of costs and benefits over the lifetime of the patient. In Australia, the median age of diagnosis is 67 years (92); as the data extracted from ICECaP contains 10 years of follow-up, to account for all cost and benefit accrued during the lifetime of a patients, extrapolation of the outcomes of interest was needed. However, given than the median age at diagnosis for LPC in Australia is 67 years, and life expectancy is 83 years, extrapolation estimates only contribute to small proportion of the results. The analysis involved in the extrapolation of the ICECaP data is presented in Chapter 3.

4.2.8. A timely decision for society

In Australia, the universal public health insurance system, Medicare, which provides subsidies for healthcare provided to the general population, is funded through general tax revenue. Economic evaluations within HTA play a crucial role in helping reimbursement decision-makers allocate limited healthcare budget resources effectively and efficiently, in order to maximise benefits for society. Hence, a societal perspective is most desirable as it reflects the interests of all those affected, however, decision-making is often restricted to a health system perspective given the challenges in adopting a societal perspective when deciding whether to subsidise new healthcare interventions.

When decisions are based on a final outcome for health conditions that require long follow-up, there is a potential delay in the reimbursement of the new interventions. Delaying reimbursement decisions on potentially effective and efficient health technologies may have a detrimental effect on patients and society. There is a potential opportunity cost to patients arising from delayed treatment access, such as experiencing disease progression sooner than might have otherwise occurred, or other disease-related complications. This is particularly important in clinical trials of long and potentially indolent health conditions such as prostate cancer. In this context, the length of clinical trials based on final outcomes motivated the ICECaP working group to look for alternative outcomes that could be measured earlier, more conveniently or more frequently to facilitate earlier trial outcomes/findings (132). Patients may benefit from a potentially effective intervention earlier if a valid surrogate outcome is used to inform economic evaluations in HTA. In theory, this may potentially decrease the morbidity and mortality associated with later stages of a disease.

4.2.8.1. Opportunity cost

In economics, opportunity cost refers to the loss of benefits from the next-best alternative. This can be seen in healthcare as the potential benefits forgone if resources had been invested in an intervention that could generate a greater health improvement. In this case, the opportunity cost of delaying an effective decision is the benefit of an intervention that could have been obtained by patients who received a less effective treatment. A point of comparison to understand the extent of the opportunity cost in delaying reimbursement decisions for a new intervention in prostate cancer can be guided by evidence of delay in cancer care. A recent meta-analysis shows that delaying cancer treatment by one month can raise the risk of death by around 10% (220). The consequences of deferring effective treatment can be burdensome for both patients and society when considering the increased financial strain from more intensive therapy, the reduction in productivity as a result of early mortality, and the diminished quality of life due to the morbidity associated with a metastatic state.

Conversely, making a premature decision based on surrogate outcomes to reimburse an ineffective intervention may also lead to opportunity costs – for example, the costs of providing and administering an ineffective or potentially harmful intervention, as well as potential harm to patients who receive the ineffective intervention rather than the more effective one, leading to decreased quality of life and increased morbidity and mortality.

4.3 Objectives

The primary objectives of this chapter are:

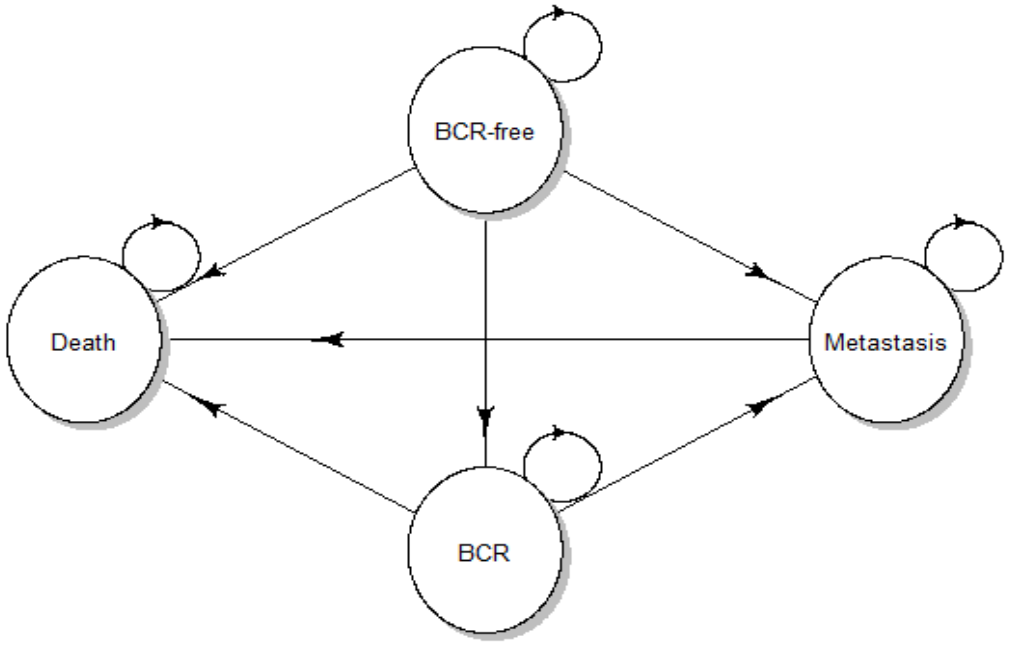
- To assess the feasibility and challenges of incorporating MFS as a surrogate outcome of OS by building an economic model for LPC to estimate the cost and benefits of a new intervention for LPC incorporating the surrogacy relationship found in the ICECaP.
- To assess the benefits to society of bringing a product to clinical practice early based on a validated surrogate outcome compared to waiting for mature OS data.

4.4 Pre-methodological assessment to determine model structure

The pathway of neoplastic diseases such as prostate cancer can be depicted and understood through a series of health states throughout the lifetime of patients. This particular characteristic makes the use of states transition models, like Markov and partitioned survival models (PSM), attractive. Widely used in this area, these models have advantages and disadvantages, necessitating thorough justification of model choice. This choice should be based on methodologies and a structure that allow the modeller to represent the natural history of the disease with certainty and that, at the same time, result in a model structure that can represent the decision problem for which it was built (76).

Markov models use probability matrices to determine the transition from one state to another; therefore, state membership is formed by an explicit link between states given by the transition probability associated with the potential move between mutually exclusive health states. Values for costs and benefits are assigned to each health state and, together with the time spent in each state, determine the total costs and benefits of the disease process (or choice) being depicted within the model. A candidate Markov model for naïve localised prostate cancer patients is shown in Figure 4.2.

Figure 4.2. A Markov model for localised prostate cancer.

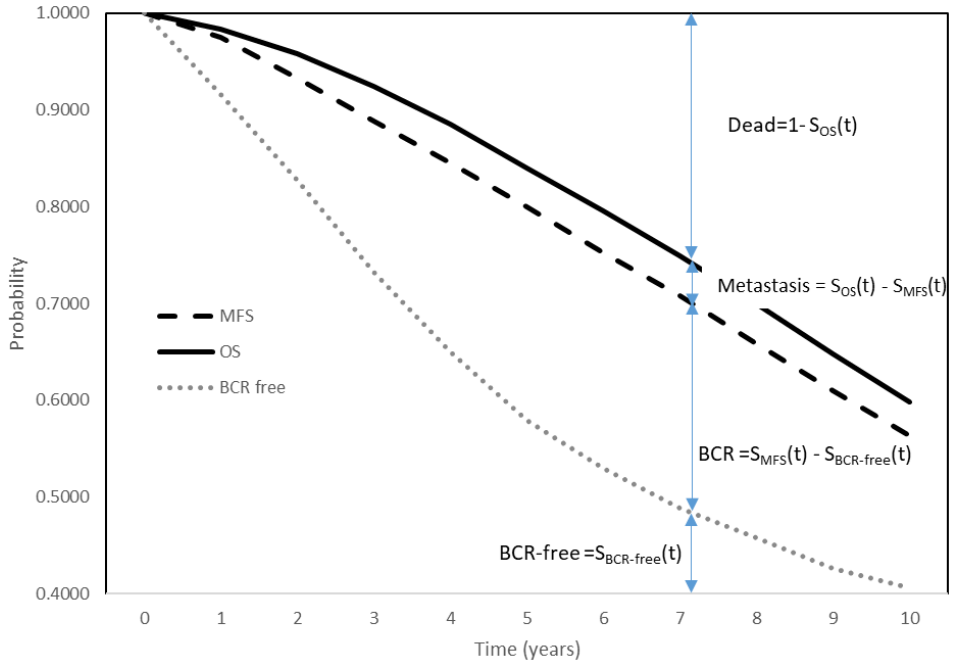


BCR=Biochemical recurrence.

Similarly, PSMs assign values of costs and benefits to health states. In both state transition models and PSMs, health outcomes and costs of the intervention analysed are estimated for each health state, aggregated over the entire time horizon, and then compared in an incremental analysis against the comparator. The difference between these approaches (PSM and state transition models) lies in how they calculate the proportion of patients in each heath state in time.

PSM were first introduced into healthcare analysis to evaluate the impact of censoring in a within-trial setting of quality of life and were found to be useful for evaluating the costs and benefits of new interventions (221) (222). In these models, individual patient data (IPD) is commonly used in a survival analysis of the outcomes of interest to determine the proportion of patients at each health state at all times. For example, the proportion of patients alive at any time is estimated by the survival probabilities from the OS curve. This curve is later disaggregated or 'partitioned' into different health states that can coexist. As they are non-mutually exclusive, state membership must be derived from the survival analysis rather than through transition probabilities matrices as in a state transition model. Therefore, in PSM, state membership is dictated by a set of independent survival curves, and the OS curve is 'partitioned' among different health states, each with a unique survival function. A candidate partitioned survival model for naïve localised prostate cancer patients is shown in Figure 4.3

Figure 4.3. Partitioned survival model for localised prostate cancer.



Abbreviations: BCR= biochemical recurrence, MFS= metastasis-free survival, OS= overall survival.

Economic models for decision-making purposes are usually built to capture the costs and benefits of an intervention over the lifetime of patients with a disease (76). Clinical trials, on the other hand, are often focused on particular outcomes; usually, secondary outcomes such as post-progression probabilities are not measured or not reported. This point becomes more important when the economic evaluation for the intervention to be analysed relies on extrapolation of benefits and surrogate outcomes, as the uncertainty carried by the use of surrogates is potentially carried over the extrapolation period as well.

Clinical trials with immature data often rely on clinically based intermediate outcomes or biomarkers that are more straightforward and quicker to measure than long-term outcomes such as OS. Post-progression probabilities that are more closely related to the final outcome, such as survival after metastasis, are time-consuming to identify and measure as they require longer follow-up time compared with intermediate outcomes. Consequently, there is a potential risk of disrupting the connection between the surrogate and the final outcome when attempting to incorporate any intermediate health state in the modelling process.

As the surrogacy relationship is only informative about MFS in state transition models, the results of the model will be explicitly affected by post-progression transition probabilities, usually sourced from the literature in a different data set. This is a source of uncertainty, as the full treatment effect of the intervention analysed over OS will not be translated completely into the model but, rather, will be

modulated by post-progression transition probabilities (if those probabilities have not been defined within the same dataset from which the surrogacy relationship emerged). As a consequence, the use of Markov models, which rely on transition probabilities, is likely to introduce external confounding into the application of a surrogate relationship in this particular case.

In contrast, in PSM, outcomes are modelled separately; the state membership in PSMs is informed only by the survival analysis of each outcome. As a result, these models do not need an explicit relationship between OS and post-progression probabilities (i.e., transition probabilities between BCR and death). Consequently, the surrogacy relationship based on MFS to predict OS in a PSM will depend only on the treatment effect observed in MFS. The independence of outcomes in PSMs preserves the initial surrogacy link (in this case between MFS and OS) and avoids the potential biases that may arise if external transition probabilities are required (as might be the case if implementing a Markov model). As a result, the independence of outcomes in PSMs make them the model of choice for this analysis using a surrogate outcome.

4.5 Methods

An economic model was developed to investigate how MFS could be used as a surrogate outcome for OS for estimating the value of new interventions to treat patients with LPC. The natural history of the disease – described in Chapter 3 and summarised in section 4.2.2 – formed the basis for the structure of the model and the identification of resources in section 4.5.7. The survival analysis – in Chapter 3 and summarised in section 4.5.4 – conducted with IPD from the ICECaP study was used to inform the state membership probabilities of the comparator arm of a PSM. Comparative treatment effects (section 4.5.5) incorporating the surrogacy relationship between MFS and OS were used to estimate the survival curves of the intervention arm in the model assuming proportional hazards.

Utilities were derived from the literature (section 4.5.8) and the identification of resources and intensity of their use were extracted from clinical guidelines and modified according to Australian settings through a survey administered to local clinicians. The payer perspective was used to define prices in the Australian setting for the cost analysis (Section 4.5.7). Sensitivity analyses were used to test the assumptions of the model, and the results of the economic evaluation were used in section 4.5.10 to estimate the possible benefits to society when a reimbursement decision regarding a new intervention for LPC is made earlier based on the use of MFS as a surrogate outcome of OS, as compared with delaying such a decision until data on OS are available.

4.5.1. Structure of the economic model: partitioned survival model

To understand model structures, a targeted literature search was conducted to identify economic models used in prostate cancer. A PSM model was used to represent the course of LPC. In a PSM of n-states, n−1 survival curves are required where the cumulative survival function S(t), represents the probability of survival up to time t. State membership in each alive health state can be calculated as the difference between $S_n(t)$ and the previous health state $S_{(n-1)}$ (t). Finally, the probability of being in the absorbing or final health state is calculated as $1- S_{OS}(t)$, or the area above the OS survival curve (176).

The number of health states in PSMs are determined by the key biological or clinical processes being studied and how interventions impact these processes (176). In a simple pathway, prostate cancer patients can move from being free of recurrence to progression of disease and to death. Development of metastases is critical to capture as a health state given the importance it has for both patients and clinicians, and its subsequent use in this instance to estimate OS in the model.

In prostate cancer, progression includes BCR; hence, this health state was based on data from patients who failed primary treatment for localised disease. Another important factor to consider is the potential for overtreatment that BCR patients may experience. If a patient fails primary treatment, hormonal treatment (ADT) can be initiated to delay metastasis and potential mortality. However, the use of ADT increases the risk of decreased quality of life. Patients can remain in surveillance to avoid treatment for a condition that may never really impact their wellbeing. Given the impact of ADT on quality of life of patients with recurrent disease, it was important to capture this health state in the model to reflect the natural history of the disease.

4.5.2. Patient population, intervention, and comparator.

The model is designed to facilitate the evaluation of new interventions in LPC in the subpopulations identified in Table 4.1 of the ICECaP study. However, for the purpose of model validation and testing, high-risk patients were selected as the relevant case study. Therefore, the intervention and the comparator were based on the treatment of high-risk LPC patients with RT added to ADT (SoC received in the ICECaP study) compared to AAP added to SoC as primary treatment.

In the base case, the model assumes a mix of patients' characteristics to reflect the cohort of patients newly diagnosed with LPC in the Australian clinical setting. Table 4.1 compares the characteristics of the population in the ICECaP analysis and those captured in clinical practice across Australian and New

Zealand by the Prostate Cancer Registry[21] (PCOR) database (92). The patients' age distribution is comparable among the two groups, but this is not the case for risk stage.

The main difference in risk groups between patients included in ICECaP and those observed in Australian clinical practice might be explained by the interventions compared in the studies included in ICECaP. As reflected in the PCOR database, active surveillance/watchful waiting is the main treatment option in low-risk patients (66%). In contrast, ICECaP included studies that recruited patients who underwent active treatment. As most low-risk patients avoid receiving active treatment, this population will be underrepresented in the ICECaP trials.

Table 4.1. Comparison of population among ICECaP and PCOR.

	ICECaP (n)	**PCOR (n)**
Age		
64 or younger	25% (3,223)	37% (9,323)
65–74	58% (7,310)	44% (11,293)
75 or older	17% (2,174)	20% (4,979)
Risk group		
Low/Intermediate	34% (4,691)	66% (16,140)
High	66% (7,586)	34% (5,891)

Abbreviations: ICECaP= intermediate clinical endpoint in cancer of prostate, PCOR= Prostate Cancer Outcome Registry.

4.5.3. Treatment pathways in the model

The treatment options for patients with prostate cancer have been discussed previously (Section 4.2.2). All treatment pathways, as well as treatment schedules, were validated by consultation with clinicians using a survey specifically designed for this purpose (Appendix 5). The survey was distributed online to a set of clinicians identified by the steering committee of the project. A total of 8 responses were received and these were employed to outline and adjust the model parameters used to develop the treatment pathway in the model.

A summary of the assumed treatment options in the model is presented below.

Primary treatment was assumed to be delivered after diagnosis during the first cycle of the model when all patients in the cohort are in the BCR-free state. The RT in the comparator arm, SoC, was

[21] The PCOR (Prostate Cancer Outcomes Registry) is a joint registry established in Australia and New Zealand to collect and collate data on men diagnosed with prostate cancer. The registry is aimed at improving the understanding of the treatment and outcomes of prostate cancer patients. The registry contains data on patients' demographics, diagnosis, treatment, management, and long-term health outcomes, and it produces publicly available reports on https://prostatecancerregistry.org/.

assumed to be delivered as IMRT according to the Australian & New Zealand consensus guidelines for definitive external beam radiotherapy for prostate carcinoma (223) added to intermittent ADT. The intervention, AAP, was assumed to be given orally corresponding to abiraterone acetate 1000 mg with prednisolone 5mg once daily added to the SoC and was to continue for 2 years or until progression, whichever came first.

ADT (as adjunct therapy in the comparator arm and subsequent lines of therapy) was assumed to be GnRH agonists. Although evidence suggests that GnRH antagonists might have a better safety profile than agonists, including a lack of the testosterone flare phenomenon, the experience gained over the years with GnRH agonists and their proven safety make them the preferred option for ADT (224). Due to GnRH agonist in testosterone levels, an antiandrogen therapy, which blocks the effect of the increased testosterone, was assumed to be given during the first month of GnRH agonist therapy.

In the model, patients who develop BCR were assumed to receive intermittent ADT at intervals of 26 weeks on and 24 weeks off treatment. Based on the findings of the Australian and New Zealand Timing of Androgen Deprivation (TOAD) trial (225), patients who are treated with ADT continue ADT with a different agent immediately after recurrence. At any cycle, for patients in the BCR state it was assumed that 34% were hormone-insensitive and required a different approach to ADT (226).

Patients diagnosed with metastases were assumed to receive the treatment protocol published in the STAMPEDE trial (227), that is ADT as either GnRH agonists or antagonists, docetaxel, and RT regardless of the initial treatment received (assumed to follow the same schedule as the primary treatment RT).

The follow-up schedule, as well as the test performed at each visit, varied by model health state. BCR-free patients were assumed to visit a specialist (i.e., an oncologist/urologist) and have a PSA test every 5 months, whereas for patients in BCR and metastasis states the visits were more frequent, once every 3 months. Every patient in BCR required a PSA at each visit; 80% of patients also required testosterone level measurement every 5 months. In addition, 26% of patients in each cycle received a bone scan, which is required if bone pain or other symptoms of progression are present. For all metastatic patients, pathology tests to evaluate renal and liver function were assumed to occur, with 50% of metastatic patients receiving bone scans due to bone pain or symptomatic progression of their disease.

The model included adverse events requiring treatment that impacted on costs and quality of life. The probabilities of adverse events after primary treatment were extracted from an economic evaluation of LPC (228). The probability of adverse events after ADT was extracted from Hershman et al 2016

(229). For metastatic patients, adverse events associated with ADT were considered throughout the whole treatment duration. To account for the cytotoxic effect, the model accounted for the incidence of adverse events that entail treatment costs or hospitalisation as reported in the STAMPEDE trial (215).[22]

The probability of occurrence of adverse events for patients under treatment assumed in the model is shown in Table 4.2.

[22] The only adverse event with these characteristics was febrile neutropaenia.

Table 4.2. Probability of treatment's adverse event occurrence assumed in the model.

Adverse event	**Probability**		
Short term AAP	Value	Range	
Major complications after AAP	0.167	0.134	0.2
Minor complications after AAP	0.833	0.799	0.866
Urinary toxicity	0.47	0.235	0.94
Erectile dysfunction	0.77	0.385	1
Urethral stricture	0.034	0.0172	0.068
Short term RT	Value	Range	
Urinary toxicity	0.3	0.15	0.6
Gastrointestinal toxicity	0.18	0.09	0.36
Chemotherapy	Value	Range	
Febrile neutropenia after chemotherapy	0.15	0.12	0.27
Long term AAP	Value	Range	
Prob. Uri. Tox. AAP long-term	0.127	0.0635	0.254
Prob. ED long-term for AAP	0.453	0.2265	0.906
Long term RT	Value	Range	
Prob. Uri. Tox. RT long-term	0.134	0.075	0.192
Prob. Gastro. Tox. RT long-term	0.02	0.01	0.04
Prob. ED long-term for RT	0.064	0.032	0.128
ADT	Value	Range	
Erectile dysfunction short term	0.800	0.7	0.9
Erectile dysfunction long term	0.066	0.066	0.248
Probability of skeletal fractures with ADT	0.015	0.012	0.022
Probability of acute MI with ADT	0.009	0.006	0.01
Probability of thrombosis with ADT	0.019	0.013	0.02
Probability of ischaemic heart disease ADT	0.008	0.006	0.012

Abbreviations: ADT= androgen deprivation therapy, MI= myocardial infraction, AAP= abiraterone plus prednisolone added to standard of care, RT= radiotherapy.

4.5.4. Survival analysis

The method used for the survival analysis to build the model is described in Chapter 3. In summary, to quantify costs and benefits over the lifetime of a patient, and to account for all costs and benefits beyond the 10-year truncated analysis of the ICECaP study, it was necessary to extrapolate the outcomes used in the model. Different parametric models were evaluated to determine the best fit for the data, considering all possible outcomes and the entire population. The adequacy of the model was assessed using statistical parameters (AIC, BIC) and by visually inspecting the hazard function. The survival analysis was discussed with health economics experts to determine appropriate survival probabilities and parametric distributions for extrapolation. The choice of extrapolation curves was

hierarchical to OS and all parametric models were evaluated for all subsets of the population in ICECaP. The same parametric distribution was applied to both the intervention (AAP) and comparator (RT) as the mechanism of action was comparable. The analysis was conducted using STATA®.

4.5.5. Treatment effect

The treatment effect in terms of MFS after RT and AAP was obtained from the survival analysis undertaken as part of the ICECaP project, as previously explained in Chapter 3. For OS, the ICECaP analysis reflects the relationship between MFS and OS in terms of the logarithm of the hazard ratio of OS and the hazard ratio of metastasis-free survival through the regression:

$$\text{Log(HR)-OS} = -0.0056 + 0.7813 * \text{Log(HR)-MFS} \qquad \text{(Eq.0.9)}$$

As previously explained in Chapter 3, equation 0.9 was updated for the economic analysis based on the publication of Xie et al. with an R-Square of 0.93 (95% CI: 0.76–0.96) (132). Therefore, the comparative treatment effect used in the model for AAP in terms of OS was translated from the comparative treatment effect observed for MFS.

To identify studies comparing AAP versus RT and reporting results in terms of HR, a targeted literature review was conducted in PubMed. The most recent meta-analysis was used as a source of relative treatment effect of RT compared to AAP for MFS and BCR-free (125).

Attard et al. meta-analysed data from two phase 3 trials assessing the efficacy of adding AAP to RT in patients with high-risk non-metastatic prostate cancer. Eligible patients had non-metastatic prostate cancer and were classified as high-risk, had a WHO performance status of 0-2 and underwent local radiotherapy. The open-label trials randomly assigned patients to either ADT alone or AAP. The primary endpoint was MFS, which was significantly longer in the combination-therapy groups compared to the control group (median not reached vs not reached, HR=0.53, 95% CI 0.44–0.64). Other secondary endpoints such as OS, prostate cancer-specific survival, biochemical failure-free survival and PFS were also significantly longer in the combination-therapy groups. The results suggested that the addition of abiraterone and prednisolone to RT improves outcomes in high-risk non-metastatic prostate cancer patients. The outcomes of interest used in the model reported by Attard et.al. are presented in Table 4.3.

Table 4.3. Treatment effect of the outcomes of interest reported by Attard et.al. used in the model.

Measure	HR (95% CI)	p-value
MFS	0.53 (0.44 – 0.64)	<0.0001
OS*	0.60 (0.48 – 0.73)	<0.0001
BCR-free	0.44 (0.36 – 0.54)	<0.0001

Abbreviations: BCR= biochemical failure; MFS= metastasis-free survival; OS= overall survival.
Note: The OS HR reported by Attard et.al. was used to validate the results of the model. OS HR in the model was estimated by the surrogacy relationship in ICECaP.

4.5.5.1. The surrogate threshold effect

The STE is the minimum treatment effect on the surrogate outcome that can predict a meaningful treatment effect on the final outcome. It can be defined in terms of the 'amount' or magnitude of the treatment effect that must be observed in the surrogate outcome to anticipate a significant treatment effect on the final outcome. Aside from giving practical information[23] on the application of a surrogate outcome, one important element of the STE is its interpretation beyond statistical validation. A low STE value in terms of the HR would indicate that a considerable treatment effect on the surrogate outcome (i.e., MFS) is required in order for an intervention to establish a meaningful effect on the final outcome. On the other hand, a high value of the STE (closer to 1 in terms of HR) would make impractical the use of the surrogacy relationship, as even if the surrogate was found to be valid (with individual and trial R^2 close to 1), it would predict a non-significant treatment effect on the final outcome.

The ICECaP analysis found that the minimum treatment effect that can be measured in terms of the HR for MFS in order for surrogate relationship to predict a nonzero treatment effect (that is, an HR different from 1 on OS) was 0.88. A sensitivity analysis was conducted of the economic model which applies an HR for MFS of 0.88 to assess the robustness of the model results when attempting to use a surrogate at the limits of the STE.

4.5.6. Perspective

The model allows both societal and healthcare system perspectives to be used, incorporating direct costs across three main categories: diagnosis, treatment and monitoring of the disease. From a

[23] From a practical perspective, a high STE value suggests that a significant treatment effect on the surrogate endpoint would need to be observed in order to conclude a positive effect on the true endpoint. In such circumstances, relying on the surrogate as a means of evaluation would not be appropriate, even if the surrogate was considered 'potentially' valid. The STE can provide a practical and fast approach to inform about the suitability of the surrogate in a specific context.

healthcare system perspective, the analysis only accounts for the direct costs incurred by the healthcare system. The societal perspective, on the other hand, included out-of-pocket expenses incurred by patients for treatments and medical services, usually in the form of co-payments. As the analysis presented in this thesis focused on HTA from a healthcare system perspective, only costs from that perspective were presented; indirect costs such as productivity costs and intangible costs were not included.

4.5.7. Costs

The identification and measurement of resource use to inform the economic model was based on a review of published clinical guidelines (124) (108) coupled with clinical guidance on how the resulting treatment algorithm reflected local practice in Australia. Resource use was described as per each of the three health states in the model.

The valuation of the resources used in the model was mainly based on publicly available prices as published via MBS for medical/diagnostic services and the PBS for pharmaceuticals[24]. The average list prices of at least 3 retail pharmaceuticals were used for drugs not subsidised via the PBS. For all the resources listed in the PBS and MBS, the unit cost used was the scheduled price/fees as applicable.

To account for time preference, costs were discounted at a 5% rate as per Australian HTA guidelines. The unit prices used in the model were based on the year 2021. For those unit prices that were published prior to this date, the general inflation rates reported by the Reserve Bank of Australia was used to estimate 2021 prices.

4.5.7.1. Follow-up costs

In this analysis, it was assumed all patients moving to BCR and metastasis required an imaging test as part of follow-up to confirm disease stage regardless of their initial state. The list of cost components associated with the imaging technique is presented in Table 4.4.

[24] In Australia, medical practitioners have the option to charge patients above the scheduled Medicare fee. As such, the out-of-pocket costs estimated in this analysis provide only an indication of potential patient costs.

Table 4.4. Imaging cost.

Code	Description	Total cost payer**	Source
56507	Computed tomography	$ 414.20	MBS
61441	Bone marrow study	$ 416.25	MBS

Abbreviations: MBS= medical benefit scheme.
Note: ** Unit costs used in the Payer's perspective analysis

As well as imaging, patients with metastasis also require additional general tests during follow-up (Table 4.5). The ADT adverse effect profile plus the use of cytotoxic agents means that patients also require assessment for renal and hepatic function during follow-up (104). As some patients may experience treatment failure or a temporary surge in testosterone during treatment with LHRH agonist, testosterone levels were assumed to be monitored during follow-up for BCR patients receiving ADT (104). The use of routine imaging is not recommended for patients on ADT with no symptoms.

For the purposes of the model, it was assumed that at least 26% of patients will experience some symptoms and therefore be prescribed a bone scan (104). Table 4.5 lists the follow-up schedule assumed in the model according to disease stage.

Table 4.5. Cost of follow-up patients.

Code	Description	Dose	Quantity	Proportion applied in model	Total cost payer**	Source
BCR-free patients						
66660	PSA : Prostate specific antigen – quantitation	1 every 5 months	2	100%	$ 63.50	MBS
105	Professional consultation + DRE	1 every 5 months	2	100%	$ 76.50	MBS
BCR						
66660	PSA : Prostate specific antigen – quantitation	1 every 3 months	4	100%	$ 127.00	MBS
105	Professional consultation + DRE	1 every 3 months	4	100%	$ 153.00	MBS
66695	Testosterone	1 every 5 months	2	80%	$ 41.52	MBS
61441	Bone scan		1	26%	$ 124.88	MBS
Metastasis						
66660	PSA : Prostate specific antigen – quantitation	1 every 5 months	2	100%	$ 63.50	MBS
105	Professional consultation + DRE	1 every 5 months	2	100%	$ 76.50	MBS
66500	Blood test	1 every 5 months	2	80%	$ 66.00	MBS
66500	CREA	1 every 5 months	2	80%	$ 13.20	MBS
66500	Alkaline Ph	1 every 5 months	2	80%	$ 13.20	MBS
66695	Testosterone	1 every 5 months	2	80%	$ 41.52	MBS
61441	Bone scan	Upon symptoms	1	55%	$ 228.94	MBS

Abbreviations: CREA= creatinine, DRE= digital rectal examination, MBS= medicine benefits schedule, PSA= prostate specific antigen.
Note: ** Unit costs used in the Payer's perspective analysis

4.5.7.2. Intervention and comparator cost

Treatment with abiraterone requires concomitant use of low dose prednisolone and the quantification of aminotransferase given the potential hepatotoxicity of the treatment (123). Table 4.6 lists the cost of the intervention used in the model.

Table 4.6. Cost of the intervention.

Code	Description	Quantity	Total cost payer**	Source
L05A	Abiraterone 500mg	730	$ 41,379.81	PBS
L05B	Prednisolone 5mg	365	$ 92.95	PBS
66500	Quantitation in serum of aminotransferase	15	$ 145.50	MBS

Abbreviations: MBS=medical benefit scheme, PBS= pharmaceutical benefit scheme.
Note: ** Unit costs used in the Payer's perspective analysis

The cost associated with SoC was defined as the cost of delivery of RT for these patients. RT to the prostate can be achieved using different techniques. For the purposes of the model, it has been assumed that RT is delivered via IMRT as this facilitates the delivery of higher radiation doses, reduces the incidence of gastrointestinal toxicities and it is widely used in Australia in this patient population (223). The Australian & New Zealand consensus guidelines for definitive external beam radiotherapy for prostate carcinoma (223) was used to identify the use of healthcare resources involved in the delivery of RT. In brief, the treatment course comprises simulation of the IMRT with or without the use of contrast medium, preparation of a dosimetry plan (which involves an imaging technique), treatment verification and the final delivery of the RT (see Table 4.7).

Table 4.7. Costs associated with RT.

Code	Description	Quantity	Total cost payer**	Source
15565	IMRT dosimetry plan	1	$3,332.65	MBS
15555	Simulation IMRT	1	$648.05	MBS
15248	RO treatment 1	1	$52.30	MBS
15263	RO treatment, 2+	5	$195.75	MBS
15715	Treatment verification	6	$402.90	MBS
104	RO 1st visit	1	$76.15	MBS
105	RO visit during treatment	6	$229.50	MBS

Abbreviations: IMRT= intensity-modulated radiation therapy, MBS=medical benefit scheme, RO= radiation oncology.
Note: ** Unit costs used in the Payer's perspective analysis

For high-risk patients, clinical guidelines recommend more aggressive primary treatment (124), (108). Therefore, it was assumed that in the SoC arm patients receive adjuvant ADT for 2 years after RT.

4.5.7.3. Systemic therapy

Systemic treatment is usually given for a short period of time after progression is established. In order to calculate the costs after failure of primary treatment and developing recurrence or metastasis, the proportion of patients failing at each cycle was estimated according to the survival analysis of the ICECaP data for SoC and AAP. However, as mentioned before, the ICECaP data does not include information about post-progression therapy. Therefore, the use of subsequent lines of treatment after the second recurrence without metastasis and after the failure of first metastasis treatment was determined by an epidemiological model that calculated the occurrence and yearly advancement of various stages of prostate cancer (226). It is important to note that this analysis uses epidemiological data on disease progression probabilities following initial treatment to estimate associated healthcare costs, rather than to directly evaluate treatment efficacy. While this provides insight into potential

economic burden, it does not on its own constitute evidence of the clinical or therapeutic benefits to estimate transition probabilities in the model.

ADT

In this model, intermittent ADT was assumed to be provided with an induction time of 6 months followed by 24 weeks off-treatment. Patients receiving ADT require close monitoring of their PSA serum level. When on treatment, monitoring of PSA levels helps to assess the treatment response and the maximal response time to ADT. In the off-treatment period, patients are exposed to increasing levels of testosterone and, therefore, are at risk of recurring or worsening disease. A follow-up time of 3 to 6 months is recommended for these patients to ensure PSA levels remain constant and that there is no sign of clinical progression (104).

Within the model, a follow-up time of 4.5 months was assumed for patients receiving intermittent ADT. Based on the use of these drugs in Australia, goserelin and leuprorelin were assumed to be used by 70% and 30% of patients respectively. Table 4.8 lists the treatment schedule and ADT costs used in the model.

Table 4.8. Cost of ADT.

Code	Description	Dose	Proportion applied in model	Quantity	Total cost payer**	Source
8094B	Bicalutamide 50 mg	50 mg /day	100%	28	$ 16.82	PBS
8093Y	Goserelin 10.8 mg implant	1 implant/ 12 weeks	50%	2	$ 1,893.34	PBS
8859G	Leuprorelin acetate 45 mg implant	1 implant/6 months	50%	1	$ 875.25	PBS
105*	Professional consultation	NA	100%	3	$ 83.00	MBS

Abbreviations: MBS= Medicare Benefits Schedule, PBS= Pharmaceutical Benefits Scheme.
Note: *This code corresponds to the cost associated with the management of GnRH agonist implants.

For patients who develop castration-resistant prostate cancer (CRPC), ADT is generally maintained and given in a continuous schedule (113). Treatment with new antiandrogens (apalutamide, darolutamide and enzalutamide) previously reserved for more advanced states of disease has been tested, approved, and recently recommended in clinical guidelines for this population (114). However, none has been listed on the PBS for use in non-metastatic CRPC. As enzalutamide is publicly funded for patients with metastatic prostate cancer, the unit price for this drug was used as a proxy in the model for the cost of new antiandrogens (Table 4.9).

Table 4.9. Cost of medication added to ADT for patients with M0CRPC.

Code	Description	Dose	Quantity	Total cost payer**	Source
10174L	Enzalutamide 40 mg	160 mg /day	1460	$45,776.21	PBS

PBS: Pharmaceutical Benefits Scheme.
Note: ** Unit costs used in the Payer's perspective analysis

Chemotherapy

Docetaxel was used in combination with ADT for newly diagnosed metastatic patients following the treatment schedule of the STAMPEDE trial (215) (see Table 4.10).

Table 4.10. Cost of chemotherapy for metastatic patients.

Code	Description	Dose	Quantity	Total cost payer**	Source
10148D	Docetaxel 160 mg/16 ml injection, 16 ml vial	¥six 3-weekly cycles of 75 mg/m²	18	,542.68	PBS
1917X	Prednisolone 5 mg tablet	10mg/day per 6 weeks	84	$ 21.03	PBS
1291Y	Dexamethasone 8 mg	8 mg i.v. every 21 days	10	212.70	PBS

Abbreviations: PBS= pharmaceutical benefits scheme.
¥The mean height (1.72 m) and wight (85 kg) for males aged between 65-74 was used to calculate the appropriate dose of docetaxel for these patients (230).
Note: ** Unit costs used in the Payer's perspective analysis

It was assumed that patients who develop resistance to docetaxel and ADT and require more aggressive treatment would require systemic treatment with other cytotoxic agents. The identification of resources used in these settings was based on the drugs used when patients relapsed in the STAMPEDE trial (215) (see Table 4.11).

Table 4.11. Cost of treatment for metastatic resistant to docetaxel and ADT.

Code	Description	Dose	Proportion applied in model	Quantity	Total cost payer**	Source
Primary medication no chemotherapy						
8094B	Enzalutamide 40 mg	160 mg/day	8%	1460	$ 3,662.10	PBS
11206T	Abiraterone 500 mg	1gr/day	28%	730	$ 11,639.97	PBS
Primary medication chemotherapy						
4376H	Cabazitaxel 60 mg/1.5 ml injection	20 mg/m2 i.v. Every 21 days x 10 cycles	7%	10	$ 5,210.60	PBS
10148D	Docetaxel 160 mg/16 ml injection	Six 3-weekly cycles of 75 mg/m² x 18 cycles	14%	18	$ 2,542.68	PBS
1917X	Prednisolone 5 mg tablet	10mg/day per 6 weeks	14%	84	$ 21.03	PBS
Chemotherapy-induced side effects medication						
1291Y	Dexamethasone 8 mg	8 mg i.v. every 21 days	12%	10	$ 212.70	PBS
Bone-targeted therapy						
10061M	Denosumab 120 mg/1.7ml	120 mg s.c. every 4 weeks	13%	13	$ 794.47	PBS

Abbreviations: ADT= androgen deprivation therapy; i.v= intravenous; PBS= pharmaceutical benefits scheme.
Note: ** Unit costs used in the Payer's perspective analysis

4.5.7.4. Adverse events

Patients were classified according to the time and duration of adverse events after primary treatment. This analysis focused on serious adverse events requiring treatment as these are most likely to impact on costs and quality of life. Adverse events after primary treatment have been categorised into those occurring immediately after primary treatment and lasting no more than 90 days (short-term adverse events) and those starting 90 days after primary treatment with a long-term effect on patients (long-term adverse events).

Primary treatment for prostate cancer may have a negative impact on sexual, urinary and bowel function (231). Consistent with the approach in other economic evaluations, adverse events after primary treatment were described as urinary toxicities, gastrointestinal toxicities or erectile dysfunction (228)[25] (Table 4.12).

[25] Urinary toxicity includes treatment for irritative voiding symptoms and incontinence. Gastrointestinal toxicity includes treatment for rectal bleeding, pain and diarrhoea.

Table 4.12. Cost prostate cancer treatment adverse events.

Code	Description	Dose	Quantity	Proportion applied in model	Total cost payer**	Source
ST-Urinary toxicity						
105	Professional		1	100%	$ 38.25	MBS
23	GP visits		3	100%	$ 116.25	MBS
11900	Urine flow study		1	100%	$ 24.15	MBS
10582Y	Paracetamol 500 mg	2tbl/8hr/ 3 months	540	100%	$ -	PBS
4049D	Ural	1scht/8hrs/3 months	270	100%	$ 63.64	PBS
ST-Gastrointestinal toxicity						
G61A	Hospital admission for rectal bleeding:		1	30%	$ 1,937.42	NHCDC round 22
105	Professional consultation		3	100%	$ 114.75	MBS
23	GP visits		2	100%	$ 77.50	MBS
65070	Blood test		2	100%	$ 82.50	MBS
4039N	Oinment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg		3	100%	$ -	PBS
4040P	Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg	2/day/3 months	180	100%	$ -	PBS
1215Y	Paracetamol 500 mg + codeine 30 mg tablet	1tbl/12hrs/3m	180	100%	$ -	PBS
1571Q	Loperamide 2 mg capsule	1caps/12 hrs/3m	180	100%	$ -	PBS
LT- Urinary toxicity						
105	Professional consultation		3	100%	$ 114.75	MBS
23	GP visits		5	100%	$ 193.75	MBS
65070	Blood test:		2	100%	$ 82.50	MBS
11900	Urine flow study		2	100%	$ 48.30	MBS
NA	Incontinence pads	2/day	730	11%	$ -	Market
NA	Incontinence pads	4/day	1460	2%	$ -	Market
1215Y	Paracetamol 500 mg + codeine 30 mg tablet	1tbl/12hr /6months	365	100%	$ -	PBS
4049D	Ural (3 sachets a day)	1scht/8hrs/year	1095	100%	$ 258.11	PBS

LT- Gastrointestinal toxicity					
105	Professional consultation		3	100%	MBS
23	GP visits		5	100%	MBS
4039N	Oinment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg		2	100%	PBS
4040P	Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg	4/day/6 weeks	168	100%	PBS
10582Y	Paracetamol 500 mg	2tbl/8hr/3 months	540	100%	PBS
1571Q	Loperamide 2 mg capsule	1caps/12hrs/ 6months	365	100%	PBS
Erectile dysfunction					
105	Professional consultation		1	100%	MBS
23	GP visits		2	100%	MBS
4585H	Sildenafil 50 mg tablet	2/month	24	90%	PBS
4579B	Alprostadil 10 microgram injection	1/month	12	10%	PBS

Abbreviations: Caps= Capsules; GP= general practitioner; LT= long term; MBS= Medicare Benefits Schedule; NHCDC= national hospital cost data collection; NA= not applicable; PBS= pharmaceutical Benefits Scheme; Scht= Sachet, ST= short term; Tbl= tablets.
Note: ** Unit costs used in the Payer's perspective analysis

Patients undergoing hormonal treatment with ADT have a different adverse event profile from those receiving primary treatment. A recent study that assessed adverse events after continuous and intermittent ADT finds that the most common events were endocrine (41%), bone-related events (14% having a fracture), and ischaemic and thrombotic events (10% having ischaemic heart disease) (229). As this model focuses on serious adverse events requiring treatment, the adverse events considered were acute myocardial infarction, thrombosis, ischaemic heart disease and fractures. The aggregated cost of these adverse events was extracted from the NHCDC (see Table 4.13). Coronary atherosclerosis was used as a proxy for ischaemic heart disease and the cost of distal femoral fractures was used as a proxy for skeletal fractures (unspecified in the study).

Table 4.13. Cost of ADT-related adverse events.

Code	Description	Total cost payer**	Source
F76B	Acute MI	$1,892	NHCDC.r22
F66B	Coronary atherosclerosis	$1,816	NHCDC.r22
F63B	Thrombosis	$1,894	NHCDC.r22
I61Z	Distal femoral fractures	$7,968	NHCDC.r22

Abbreviations: MI= myocardial infraction, NHCDC= national hospital cost data collection.
Note: ** Unit costs used in the Payer's perspective analysis

For metastatic patients, ADT-related adverse events were incorporated throughout the treatment duration, with the cost of treating febrile neutropaenia added during the period of treatment with cytotoxic agents where relevant. The cost of febrile neutropaenia was extracted from Lingaratnam et al (232), who evaluated the cost of febrile neutropaenia management in Australia. A total aggregate cost of $4,304.85 was used in the model, corresponding to the cost for the intermediate strategy for treating febrile neutropaenia as reported by Lingaratnam et al and updated to 2021 prices based on the general inflation rate since 2010.

4.5.8. Utilities

The quality of life impacts incorporated in the model were derived from a systematic review and meta-analysis conducted of studies reporting health-related quality of life outcomes using utility values in patients with prostate cancer. A literature search was conducted in MEDLINE to find relevant studies. Relevant systematic reviews and meta-analyses were also reviewed to provide synthesised evidence (233). A meta-analysis was performed using a random-effects model in STATA software and heterogeneity across studies was assessed using I2 statistics.

Studies reporting utility values for specific health states were included, such as post-treatment, recurrence, metastatic disease, and treatment side effects. Study characteristics and population details were extracted. When multiple methods were used, EQ-5D utility values were prioritised for inclusion in the meta-analysis.

A built-in function in STATA, meta suite commands, was used to perform a meta-analysis of the results of individual studies included with similar characteristics into one pooled result. The meta-analysis used a random-effects[26] model and restricted maximum likelihood approach, as recommended by the

[26] The random-effects meta-analysis approach assumes the difference in utility observed in each study are related and follows the same distribution as opposed to the fixed effect methods where the differences between study

Cochrane handbook (233). The heterogeneity found in the pooled results was assessed using the I^2 statistic.[27]

Based on I^2 statistics for each pooled result, a more detailed review of the included studies was performed. Excluding a result based on a study's findings is not recommended as it can create biased results. Nevertheless, if there is a strong explanation for the outlier finding, the analysis result can be excluded with confidence, as heterogeneity could be caused by studies with contradictory results due to different methodologies, settings or populations (233).

All studies that analysed AAP reported quality of life outcomes for patients with metastases or after progression. As the model includes SoC as a comparator, and no estimates of quality of life were found for patients with LPC receiving AAP, utility values reported for radical prostatectomy, another primary treatment for LPC, were used as a proxy.

In this study, the most recent systematic review and meta-analysis by Magnus and colleagues was used to identify studies reporting health utilities for patients with prostate cancer (217). The meta-analysis included 29 studies that reported utilities for patients in different health states, such as following primary treatment and experiencing recurrence and adverse events of treatments (218, 219, 234-260). While nine studies reported utilities for patients with metastases, one included utilities of partners of patients living with prostate cancer (237) and two reported the combined utilities of patients with metastases together with patients in other health states (219, 239), and as a result were not included in the current meta-analysis.

Six studies included utility values for patients after surgery (240, 247, 248, 251-253). Of these, one reported the utility for patients receiving adjuvant therapy (251) and was excluded and one reported only short-term utility values and was therefore excluded (248). Four studies reported the utilities of patients after SoC (247, 248, 251, 253). Of those, one reported the utility for patients on adjuvant

---

results are assumed to be caused by chance. The random-effect approach addresses the heterogeneity in results that is not explained by obvious difference in population, settings or method stated in the study by treating variations as if they were random.

[27] I^2 denotes the variability in the pooled utility value, which in this case is caused by heterogeneity. This describes the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), and is calculated using the following equation:

$I^2 = \left(\frac{Q-df}{Q}\right) x\ 100$, where Q is the Chi^2 statistics test for heterogeneity and df is the degree of freedom.

The Cochrane handbook recommends interpreting the results of the test as follows:
0–40%: might not be important.
30–60%: may represent moderate heterogeneity;
50–90%: may represent substantial heterogeneity;
75–100%: considerable heterogeneity.

therapy (251) and was excluded and one reported only short-term utility values and was therefore excluded (248). Only one study reported utilities of patients after ADT (251) and was used to populate the model. The characteristics of the studies included in the current meta-analysis are described in Appendix 6.

Although different methods were used to elicit the preference of patients, the majority used the indirect method EQ-5D utility index (218, 240, 244, 247, 253, 254, 257, 259) followed by the time trade off (236, 243, 256), standard gamble (251) and PORPUS (an indirect method to elicit patients' preferences that is specific for prostate cancer)(252). When more than one method was used in the same study for the population of interest, values of the EQ-5D were extracted and used in the meta-analysis.

Application in the economic model

The pooled values were used to represent different health states depending on treatment received and adverse events experienced. Adjustments were made for declining quality of life with age. Assumptions were made around long-term impacts and changes over time. Pooled utility values after SoC were used to populate the BCR-free state for patients who did not develop recurrence after initial treatment and had no adverse events. The utility values of urinary and bowel/rectal toxicities and erectile dysfunction were used for patents in the BCR-free state presenting at least one of these events after initial treatment (as reported in section 4.5.3). Although there is a long-term probability of developing these adverse events for all patients independent of initial treatment failure, it was assumed that utility values in subsequent health states capture the detrimental effect caused by the long-term adverse events after primary therapy.

The heterogeneity (measured as I^2 test) for the pooled results of the utility values for metastasis, urinary incontinence and erectile disfunction was high. For metastatic patients, one study (254) included only chemotherapy-naive patients and the utility value reported was therefore higher than those reported in other studies that included patients treated with chemotherapy. After the removal of this study, the pooled results showed moderate heterogeneity.

On the other hand, for erectile dysfunction and urinary toxicity, although the I^2 indicates high variability of the results due to heterogeneity, the population characteristics and methods used in preference elicitation were similar among all the studies included. Therefore, the pooled results of all studies reporting these adverse events were used to populate the model. The pooled utility values

obtained from the meta-analysis and used in the model are summarised in Table 4.14 and a forest plot of those results is presented in Appendix 6.

Table 4.14. Pooled estimates of utility values used in the model.

Health state	Utility value	CI-U	CI-L	I^2
BCR-free with no AE	0.880	0.933	0.843	39.23%
Metastatic cancer	0.671	0.701	0.640	71.64%
Erectile dysfunction	0.799	0.923	0.675	97.29%
BCR	0.790	0.986	0.594	NA
BCR-free with gastrointestinal side effects	0.859	0.885	0.833	NA
BCR-free with genitourinary side effects from treatment	0.763	0.906	0.620	97.77%
Male population from 71 and older	0.717	0.731	0.703	NA*

Abbreviations: BCR: Biochemical recurrence; CI-L: Low confidence interval; CI-U: Upper confidence interval; I2: variability of pooled utility value due to heterogeneity; AAP: Radical prostatectomy; SoC: Radiotherapy.
Note: *Source: Norman et al. 2013.

The impact of treatment on quality of life is analysed in the model by assigning the utility values based on the event type, with the probability of those events varying between treatment groups (AAP + SoC and SoC). Additionally, to adjust for a declining quality of life at older ages, the model assigns differential utility values for patients alive and free of recurrence, according to their age as reported for the Australian population (261). Only patients free of recurrence were considered, as those in recurrence or metastatic stages are assigned utility values corresponding to the treatment received, which were lower than the utility values at older ages.

The model assumed that 5 years after receiving primary treatment, the quality of life of patients living in BCR-free with erectile dysfunction equals that of the general population at the same age.

4.5.9. Sensitivity analysis

The challenge in performing and interpreting sensitivity analyses in a PSM stems from the structural independence of the endpoints (176). In order to estimate the outcomes for the extrapolation period, various parametrisation techniques were evaluated, yet no formal approach was found to assess the structural validity of the survival curves during the trial period. Deterministic and probabilistic sensitivity analysis using a tornado diagram and a scatter plot to illustrate the uncertainty caused by varying all the parameters used in the model was conducted.

Deterministic sensitivity analysis: For the deterministic analyses, parameter inputs for costs and utilities were tested within their confidence intervals (CI), when available. When CIs were not

published, sensitivity analyses were performed by varying the parameter estimate applied in the analysis by 20% under and over the point estimate.

Probabilistic sensitivity analysis: For the probabilistic sensitivity analysis, adverse event probabilities and costs were assumed to follow a beta distribution, ratios a normal distribution and cost- and time-dependent variables a gamma distribution. A Monte Carlo simulation was conducted using 5,000 random samples to create a probability cloud for the resulting incremental cost-effectiveness ratios depicting the uncertainty of the parameters used.

Although Australia does not have an explicit threshold, published studies have set an implicit cost-effectiveness threshold of $50,000 per QALY (262). Additionally, in a report to the Australian Parliament, the Department of Health has indicated that 'the ICERs of the middle 50 per cent of recommended cancer drug applications were in the range of $42 000–$61 000' (263). Therefore, in this analysis a cost-effectiveness threshold of $50,000 per QALY was assumed to estimate the incremental net benefits.

4.5.9.1. Scenario analysis

The tornado diagram arising from a deterministic sensitivity analysis is normally used to illustrate the impact of reasonable changes on individual variables compared to the base-case assumptions. However, its interpretation can be influenced by the choice of base case values and the range of values used for sensitivity analysis. To better reflect the impact of the surrogacy relationship and the choice of distribution for the survival analysis, a set of scenario analyses were performed.

4.5.9.1.1. Testing the surrogacy relationship

When MFS is used as a surrogate for OS in an economic model, changes to HR MFS not only impact the membership of patients in metastatic health states in the model, but also affect the overall survival curve. The interpretation of univariate sensitivity analysis when multiple variables are responsible for variation on results is complex and could lead to biased conclusions.

Accordingly, a sensitivity analysis testing the use of the surrogate outcome in the model was performed using the range of possible values of HR MFS reported in the meta-analysis of abiraterone in LPC as inputs to the regression equation correlating MFS and OS (125). Additionally, the most informative results of the model would be given by testing extreme values of the MFS-HR used to estimate OS. This translates into values where the surrogacy relationship predicts a change of

direction of the treatment effect (HR > 0.88) to extreme values of HR that were reported in the ICECaP trial (HR = 0.202) (132).

The surrogacy relationship is based on the construction of comparative MFS for the new intervention. The MFS curve, as with other time to events analysis, shows an increase in the uncertainty around the survival probability measure with time, as fewer metastasis events are observed in the cohort being followed as time increases. High censoring at the end of the follow-up period, which is common in long-term trials, as well as the low probabilities observed in the tail (asymptotic curves) can lead to bias in the Kaplan–Meier estimator in the tail of the distribution (264). Therefore, the comparative efficacy predicted, based on the last portion of the curve, may be less certain. This is of particular interest when an extrapolated survival function is used in the model. To test the impact of an uncertain treatment effect after the ICECaP analysis period, a scenario analysis was performed where a waning effect on the surrogacy relation was used. In this case, the treatment effect of both SoC and AAP was assumed to be equal after the truncated analysis in ICECaP, resulting in the model attributing the same OS probabilities extrapolated from the ICECaP data to AAP.

4.5.10. Benefits to society

4.5.10.1. Flow of patients

In general, the results of an economic model are presented in terms of mean cost and QALYs gained per patient. Therefore, the size of the population plays a minor role in the development of economic models; it is of most importance when calculating the budget impact of new interventions. However, population size may be relevant when estimating the aggregate impact of a new intervention. In this context, the results of the model were used to estimate the benefits to society arising from making an earlier decision regarding reimbursement of an intervention (as opposed to delaying that decision while awaiting final outcomes) and, therefore, absolute patient numbers are required.

The size of the cohort of patients receiving primary treatment after diagnosis was estimated using the incidence of prostate cancer reported by the Australian Institute of Health and Welfare (86). Information from PCOR (92) was used to identify the demographic characteristics of patients, such as their stage of disease at diagnosis, likelihood of recurrence, available treatment options, and their willingness to receive active treatment. Based on this information, the likely number of patients diagnosed with LPC who would receive AAP in combination with RT for primary treatment was estimated.

The Australian Institute of Health and Welfare estimated the incidence of prostate cancer patients in 2019 to be 19,508 (86). However, not all patients with a recent diagnosis will receive treatment for the primary tumour. It is estimated that 8% of these patients will present with regional or metastatic disease and, therefore, receive treatment specific for their condition (208). On the other hand, not all patients with localised prostate cancer will be willing or able to receive active treatment. The proportion of patients undergoing active surveillance/watchful waiting ranged from 66% in low-risk patients to 4% in high-risk patients, whereas patients who received ADT as a primary treatment ranged from 1% in intermediate risk to 6% in high-risk patients (208). Patients receiving SoC for their high-risk disease were assumed to be treated with adjuvant AAP if reimbursed.

Based on the pattern of care mentioned above, the benefits to society were calculated using a starting population of 1,545 patients receiving AAP as primary treatment after initial diagnosis of LPC.

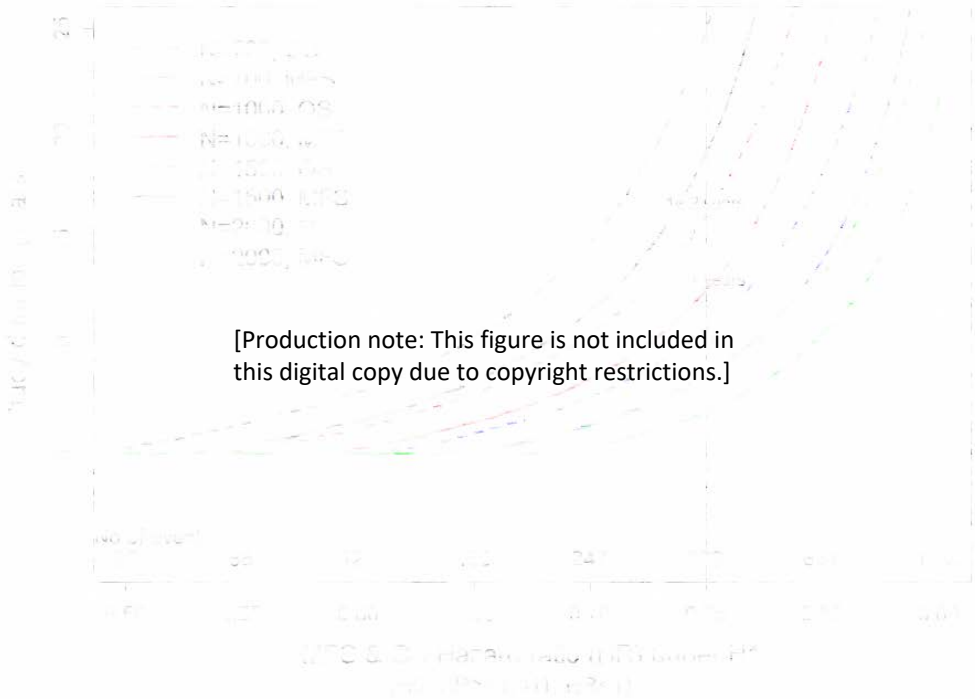
As AAP is not a current treatment for these patients, the model assumed the proportion of patients receiving AAP was the same as the proportion of high-risk patients receiving RT reported in the PCOR (92) (208).

4.5.10.2. Timely decision

The potential time saved when a new intervention for LPC is evaluated based on an economic evaluation informed by MFS as a surrogate of OS was estimated based on the results of the ICECaP study. The ICECaP study found a correlation between the OS rate at 8 years and the MFS rate at 5 years (132), suggesting a potential reduction of 3 years in the decision-making process. Furthermore, a simulation conducted as part of the ICECaP study of trial duration, taking into account different scenarios of possible HRs of MFS and population size typically used in study designs, estimated a reduction of study duration by slightly over 3 years when using MFS as a surrogate for OS (Figure 4.4).

Therefore, the benefits to society were calculated by comparing the results of the model in the base case scenario with those assuming a 3-year delay in decision-making (that is, what outcomes would be observed if reimbursement decisions were made on the basis of 5yr-MFS rather than waiting for 8yr-OS). As the use of surrogates implies that a decision is made to provide the intervention earlier, it was assumed that the potential loss from a delay in the short term would be given by the difference in the results between the intervention and the comparator in the first 3 years.

Figure 4.4. Study duration required to observe the number of events according to HRs.



Source: Xie et.at. 2017 (supplementary material).
Abbreviations: HR= hazard ratios; MFS= metastasis-free survival; OS= overall survival.

The potential short-term benefits to society were estimated by using the difference in absolute values that would result from a 3-year delayed decision for the entire patient cohort. The long-term benefits and costs were then calculated by comparing a 'timely decision' with a 'delayed decision' using the difference produced in state membership between the SoC and AAP over the initial 3 years.

Short term: The population starting the model consisted of patients diagnosed with localised disease who received either AAP or SoC as their primary treatment option. The model was run over a time horizon of 3 years, during which the status of patients who remained disease-free, experienced recurrence, developed metastases or passed away was recorded. The model not only distinguished between state membership, but also provided an analysis of the cost-benefit outcomes when comparing AAP to SoC over a period of 3 years.

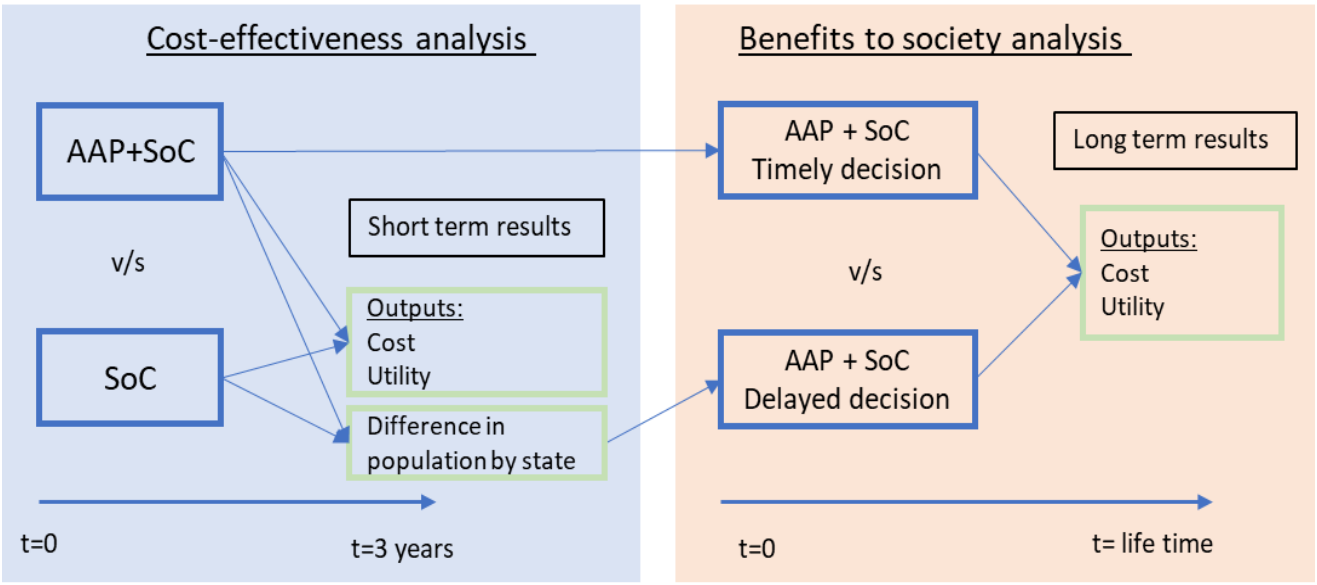
Long term: The population starting the model is informed by the results of the short-term analysis. The long-term analysis compared the difference in cost and benefits arising when a 'timely decision' is compared with a 'delayed decision'. Therefore, for the 'delayed arm', it was assumed that only a proportion of the initial cohort would be able to receive AAP as their primary treatment, as some patients had progressed to an advanced stage of the disease due to treatment delay (only patients

remaining free of diseases in the short-term analysis). The model was modified to compare the costs and benefits over the long term for the 'timely decision arm' and the 'delay arm'.

The difference in the first 3 years for the short-term analysis between SoC and AAP was used to populate the 'delay arm' of the model. It was assumed that patients who progressed while receiving SoC within the time that it takes clinical trials to report OS (defined previously as 3 years) in the 'delay decision arm' do not receive primary treatment (AAP) and enter the model while on BCR or in the metastatic state.

On the other hand, when a timely decision is made (which constitutes the 'timely decision' arm), all patients of the cohort are assumed to enter the model while in the BCR-free state and receive AAP as primary treatment. The model is run to evaluate the costs and benefits arising from the difference in state membership among the populations in the 'delay decision' arm and the 'timely decision' arms in the long term (see Figure 4.5). Costs and benefits to society were estimated using the resource and utilities values identified for each health state for the economic model in sections 4.6.3 and 4.6.4.

Figure 4.5. Diagram of the model comparing timely vs delay decision in LPC.



Abbreviations: AAP= abiraterone plus prednisolone, SoC= standard of care, t= time.

4.5.11. Value of information

The use of surrogate outcomes in economic evaluations can result in uncertainty in reimbursement decisions, as the treatment effect on the final outcome is estimated through a correlation with the treatment effect observed in the surrogate outcome. As discussed in Chapter 3 this is particularly

useful when the time it takes clinical trials to show a significant treatment effect results in terms of final outcomes (i.e., OS) is exceptionally long, as observed in trials of LPC.

When new interventions are supported by surrogate outcomes, decision-makers have the choice of making a decision using the information at hand or delaying the decision until additional information on the final outcome is available, which will reduce the uncertainty. In such cases, value of information analysis can assess the value of conducting additional trials to reduce uncertainty (265). In this approach, the cost of delaying a reimbursement decision based on a surrogate outcome is determined by the cost of the additional trial plus the expected forgone reduction in cost compared to SoC had the decision been made immediately. On the other hand, the expected value of a delayed decision is determined by the reduction in uncertainty associated with reducing the probability of making the wrong reimbursement decision. The assessment is then made by assessing whether the reduction in uncertainty provided by the information available as the result of additional trials out-weighs the cost of conducting those trials.

The expected value of additional information in the current model can be estimated by the opportunity loss when a wrong decision is made given the uncertainty of all parameters in the model. This analysis can also be interpreted as the expected value of perfect information (EVPI) to remove uncertainty from the decision problem (76). The EVPI for an individual patient can be calculated as the difference between the expected net benefit made with perfect information and the expected net benefit arising from the existing evidence. The net benefit arising from a reimbursement decision is defined by the opportunity cost expected to be forgone as a consequence of the incremental cost of the new intervention; therefore, to estimate the incremental health benefits it is necessary to define the cost-effectiveness threshold within the setting in which the decision is been made.

The EVPI is usually estimated using the outputs of the probabilistic sensitivity analysis, which take into account all distributions of possible values of each parameter in the model. In this study, the EVPI was conducted with the results of the probabilistic sensitivity analysis using a threshold of $50,000 per QALY (discussed in section 4.5.9). The EVPI can be calculated as per equation 0.10:

$$EVPI = E_\theta max_j\ NB(j,\theta) - \ max_j E_\theta NB(j,\theta) \qquad \text{(Eq.0.10)}$$

where $E_\theta max_j\ NB(j,\theta)$ is the average of the maximum benefit (NB) of all possible results that would provide perfect information for a certain level of known uncertainty (θ) when perfect information is used to choose alternative *j; and* $max_j E_\theta NB(j,\theta)$ is the maximum net benefit generated when the optimal decision is taken with the existing evidence and unknown uncertainty.

The results of equation 0.10 provide the EVPI for an individual patient; however, as we are concerned with how the decision impacts society, the EVPI should be estimated for all possible patients in the population that would benefit from the intervention (266):

$$EVPI_{population} = EVPI \sum_{t=1,2,...T} I_t/(1+r)^t \quad \text{(Eq.0.11)}$$

where *T* is the time expected the intervention to be given to patients, *I* is the incidence of LPC and r is the discount rate of benefits. The incident population was the same estimated for the benefit to society analysis in section 4.5.10 and the expected life time of the use of the intervention was estimated to be between 10 and 20 years (267).

4.5.12. Applicability beyond high-risk LPC patients

The model was built on the basis of the ICECaP analysis, explained in Chapter 3. This analysis stratified outcomes by risk and age, and a sensitivity analysis demonstrated the consistent relationship between MFS and OS as surrogates among all stratified subpopulations. Therefore, a flexible model was created, integrating the relationship between MFS as a surrogate of OS. This model provides a comprehensive framework for assessing the effectiveness and efficiency of various interventions for the treatment of LPC patients, making it a useful tool for both researchers and decision-makers alike. To ensure that users fully comprehend the model, a manual was developed to provide instructions on how to use it effectively and is provided as an Appendix 7.

4.6 Results

The results of the survival analysis from the ICECaP database, which was used to construct the partitioned survival model for estimating the treatment effect on OS by AAP, were outlined in Chapter 3. In this chapter, a brief overview of those survival results is provided as they are essential to contextualising the development of the PSM. Cost and utilities characterising each health state in the model are subsequently presented, followed by the cost-effectiveness and sensitivity analysis results. Subsequently, the results of the benefits to society analysis are presented.

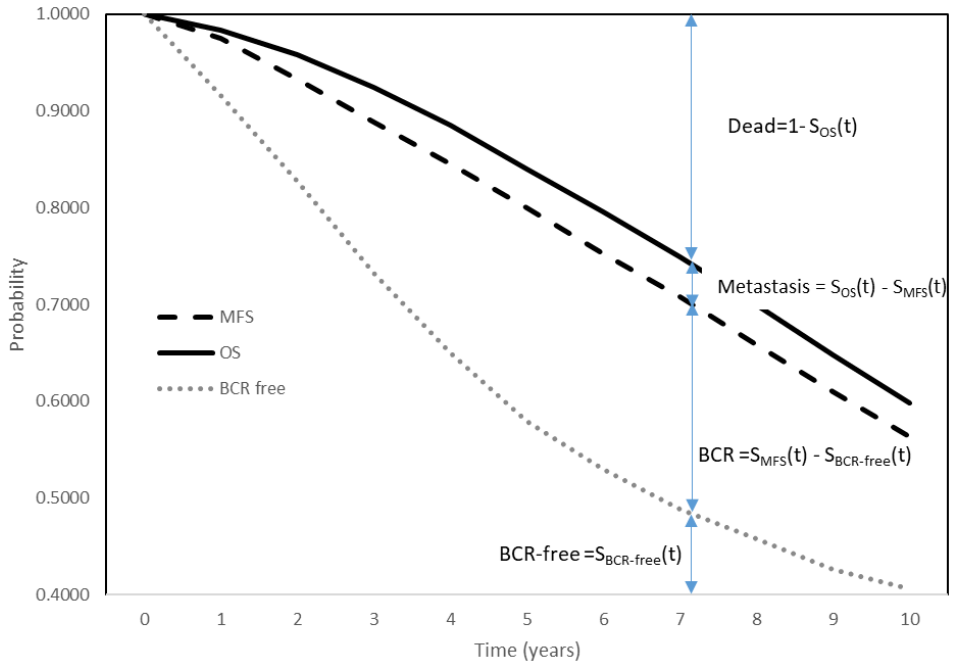
4.6.1. Model of choice

From 151 articles found in the literature, six systematic and scoping reviews of economic evaluations investigating the cost-effectiveness of primary treatments in LPC were identified and reviewed to assess the suitability of different modelling approaches in prostate cancer (12, 268-272).

The most popular models used to reflect the natural history of LPC after primary treatment were Markov models, representing 73% of all models used, followed by PSM. In general, models consisted of 3 to 5 health states including: 1) no evidence of metastatic prostate cancer; 2) metastatic cancer for which hormonal therapy provides effective control of symptoms and progression of disease; 3) progressive metastatic cancer that is refractory to hormonal therapy; 4) death from progressive metastatic cancer; and 5) death due to other causes.

Due to the argument in Section 4.5, a PSM was chosen. State membership of a PSM can be represented graphically as the difference of the survival curves. The state membership of the ICECaP study in a four-health state model (BCR-free, BCR, Metastasis and Death) for the overall population using OS, MFS and BCR-free survival curves is shown in Figure 4.6.

Figure 4.6. State membership for the overall population in ICECaP.

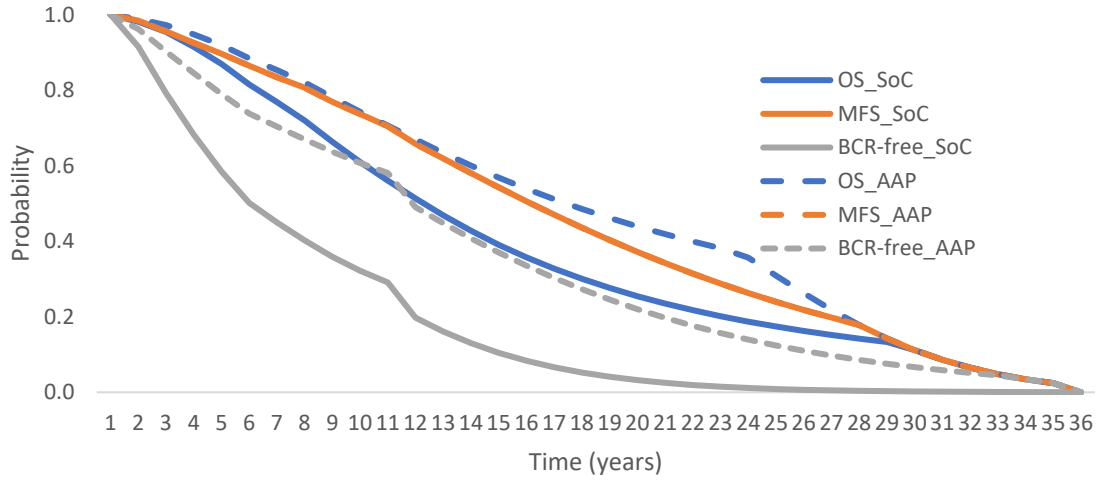


Abbreviations: BCR= biochemical recurrence, MFS= metastasis-free survival, OS= overall survival.

4.6.2. Estimating Survival Probabilities using the ICECaP database

The data collected in the ICECaP study that correlated OS and MFS was used to populate the comparator arm (SoC) in the model. Survival probabilities of BCR-free, MFS and OS in the 10-year follow-up period after primary treatment with SoC were informed by individual patient data collected by ICECaP. To estimate costs and benefits over the lifetime of patients, the outcomes based on a truncated analysis at 10 years from the ICECaP study were extrapolated. Conventional parametric distribution models were evaluated to extrapolate all outcomes of interest in the ICECaP data. The graph in Figure 4.7. shows the extrapolated survival probabilities for the high-risk population used in the model for the outcomes of MFS, BCR-free and OS, taking into account the adjustment for other causes of death in the Australian general population.

Figure 4.7. Adjusted survival curves for AAP and SoC with extrapolation.



Abbreviations: AAP= abiraterone plus prednisolone; ADT= androgen deprivation therapy; BCR= biochemical failure; MFS= metastasis-free survival; OS= overall survival; SoC= standard of care.

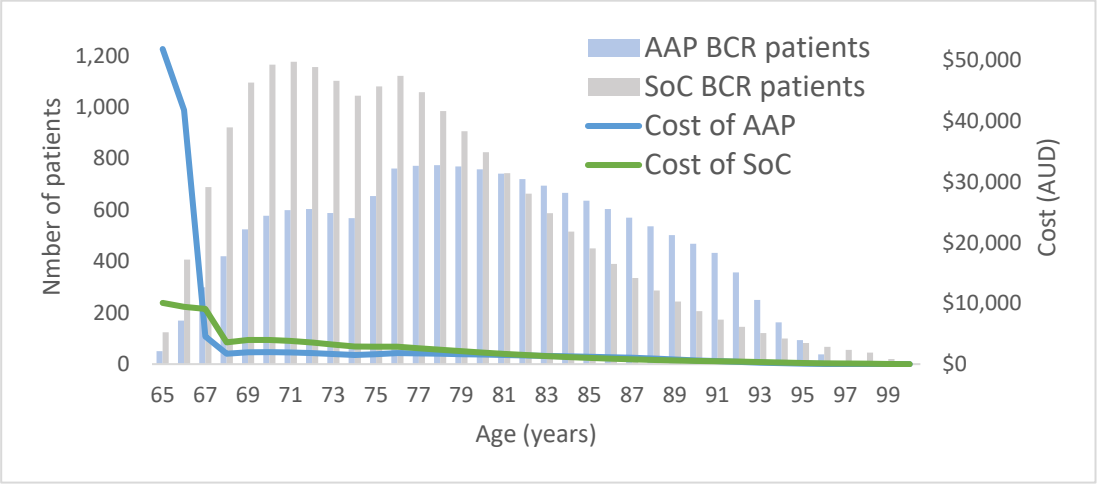
The mean survival time is reported in the survival analysis conducted as part of Chapter 3. Briefly, the mean survival time for OS was 17.56 years for AAP and 14.15 years for SoC. For MFS, AAP had a mean survival time of 16.62 years, while SoC had a mean survival time of 11.75 years. Similarly, AAP had a higher mean survival time for BCR-free (12.97 years) compared to SoC (7.27 years). However, for BCR and MET, SoC had a higher mean survival time of 4.48 years and 2.40 years respectively, compared to AAP mean survival time of 3.65 years and 0.94 years.

4.6.3. Costs

The costs associated with SoC and AAP were higher during the first year, when both interventions are delivered to all patients in the cohort. The assumed length of treatment can explain the sharp drop in

cost after the second year; costs subsequently rise with the appearance of failure following primary treatment in subsequent cycles. Figure 4.8 illustrates the relationship between the total cost per patient in the model and the distribution of patients failing primary treatment throughout the time horizon of the model.

Figure 4.8 . Total cost per patient for SoC and AAP in the overall population.



Abbreviations: AAP= abiraterone plus prednisolone, BCR= biochemical recurrence; SoC= standard of care.

The total costs arising in the model are driven by the use of resources and prices of treatments for newly diagnosed patients and for those failing primary treatment. Although the cost of systemic therapies used to treat metastatic patients is considerably higher than primary treatment, the higher probability of recurrence compared to the probability of metastasis development influences the total costs estimated by the model (Table 4.15)

Table 4.15. Total cost by health state and cost category in the model.

	By health state			**Cost category**				
	BCR-free	BCR	Metastasis	Initial treatment	AE	Follow-up	Imaging	Subsequent treatment
SoC	$ 24,594	$ 45,815	$ 10,248	$ 22,886	$ 1,801	$ 2,430	$ 465	$ 53,075
AAP	$ 97,732	$ 31,598	$ 4,162	$ 95,350	$ 1,660	$ 2,251	$ 274	$ 33,957

Abbreviations: AAP= abiraterone plus prednisolone, BCR= biochemical recurrence; SoC= standard of care.

4.6.4. Utilities

The total QALYs per patient associated with SoC and AAP were driven by the quality of life of patients in the BCR-free state, which was the health state with the higher quality of life (Table 4.16) (Figure 4.8). In contrast, the model estimated the lowest quality of life for the metastatic state.

Table 4.16. QALYs per patient estimated in the model.

	BCR-free	BCR	Metastasis	Total
SoC	4.24	2.13	0.72	7.09
AAP	6.64	1.49	0.29	8.42

Abbreviations: AAP= abiraterone plus prednisolone, BCR= biochemical recurrence; SoC= standard of care.

4.6.5. Incremental analysis

Patients with LPC treated with AAP accrued more QALYs than those treated with SoC alone during their lifetime (the average patient treated with AAP accumulated 1.33 additional QALYs compared with the SoC). However, the higher cost of abiraterone for primary treatment makes the cost of the average patient receiving AAP $52,835 higher than a patient treated with SoC. Table 4.17 presents the results of the incremental analysis in the model when MFS is used as a surrogate for OS.

Table 4.17. Incremental analysis.

AAP Total cost	SoC Total cost	Incremental cost	AAP total QALY	SoC total QALY	Incremental QALY	ICER
$133,490	$80,655	$52,835	8.42	7.09	1.33	$39,640 /QALY

Abbreviations: AAP= abiraterone plus prednisolone, ICER= incremental cost effectiveness ratio, QALY= quality adjusted life years, SoC= standard of care.

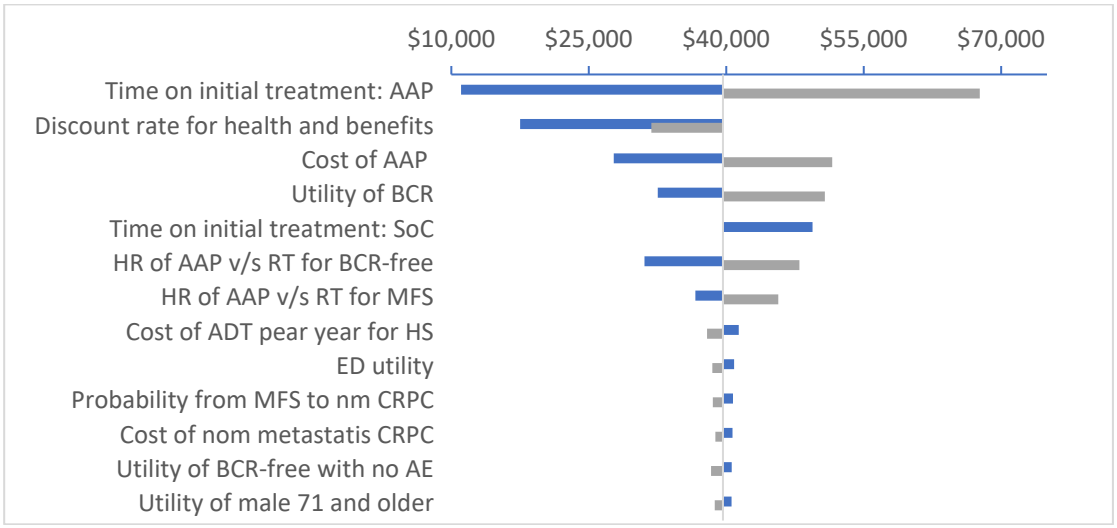
4.6.6. Sensitivity analysis

The results of univariate sensitivity analyses are shown in Figure 4.9. The variable with the greatest impact on the results was the time on treatment for AAP[28]. As the initial cost of AAP is one of the drivers of the model, varying the time on treatment (time on treatment for AAP was varied from 1 to 3 years) –and thus the quantity and, ultimately, the total costs of treatment – impacts on the results of the model.

As shown in Chapter 3, the model predicts that patients spend most time in the BCR-free state, which is assigned the highest utility values in the model. As the BCR-free curve determines the survival probability of patients in this state, variations in the utility assigned to patients spending time in BCR-free and its comparative treatment effect (HR of BCR-free) also have a large influence on the results. Lower HR for the BCR-free state translated into more patients failing to initiate treatment (missing the QALYs gained in this state) and receiving further treatment (increasing the cost and decreasing QALYs compared to being BCR-free) following SoC compared to AAP.

[28] While time on treatment has the potential to influence MFS, this analysis used it solely as a input parameter for estimating potential cost implications rather than as a measure of any clinical benefit derived from prolonged ADT.

Figure 4.9. Tornado plot summarising multiple one-way sensitivity analyses for the overall population.



Abbreviations: ADT= androgen deprivation therapy; AAP= abiraterone plus prednisolone, ED= erectile dysfunction; HR= hazard ratio; MFS= metastasis-free survival; SoC= standard of care. Note: Variables not shown in the figure have low impact on the results. Blue bars indicate variation in results when the lower range of the variable is used in the analysis, grey bars, the opposite. Variation towards the left of the figure favours AAP.

The probabilistic sensitivity analysis provided in Figure 4.10 shows a concentrated cloud around the base case with the majority of simulations in the upper right quadrant of the cost-effectiveness plane. The total QALYs gained by SoC were greater than the QALYs for AAP in 7.7% of the 5,000 simulations. Likewise, fewer than 0.03% of the simulations resulted in SoC being more costly than AAP. The broad distribution of the scatter plot along the cost axis suggests a higher uncertainty regarding the cost parameters compared to QALYs.

Figure 4.10. Probabilistic sensitivity analysis.

4.6.6.1. The surrogacy relationship

The impact on the ICER of varying the surrogacy relationship was tested in the sensitivity analysis by varying the values of HR MFS among a range of possible values. The estimated results of the model were recorded when values of HR MFS varied from where the surrogacy relationship predicts a change of direction of the treatment effect (HR > 0.88), to the extreme values of HR MFS reported in the ICECaP trial (132) (the lower inferior confidence interval value of HR MFS= 0.202). Table 4.18 shows the relationship of HR MFS with HR OS, cost and QALYs when a range of possible values were tested in a sensitivity analysis of the surrogacy relationship.

Table 4.18. Sensitivity analysis of the surrogacy relationship.

HR-OS	HR-MFS	Cost AAP	Diff. cost	QALY AAP	Diff. QALY	ICER
0.28	0.20	$156,037	$75,382	9.39	2.30	$32,708
0.39	0.30	$151,812	$71,157	9.20	2.11	$33,717
0.48	0.40	$144,272	$63,618	8.87	1.78	$35,674
0.58	0.50	$135,851	$55,196	8.52	1.43	$38,576
0.66	0.60	$127,582	$46,928	8.17	1.09	$43,155
0.75	0.70	$119,895	$39,240	7.85	0.76	$51,300
0.83	0.80	$113,430	$32,775	7.55	0.47	$69,916
0.91	0.90	$110,500	$29,845	7.29	0.21	$143,130
0.99	1.00	$109,785	$29,130	7.06	−0.02	AAP dominated
1.07	1.10	$109,464	$28,810	6.85	−0.23	AAP dominated

Abbreviations: AAP= abiraterone plus prednisolone, HR= hazard ratio; MFS= metastasis-free survival; OS= overall survival; SoC= standard of care. Note: Cost and QALYs of SoC do not change when the HR MFS varies and therefore, are not presented.

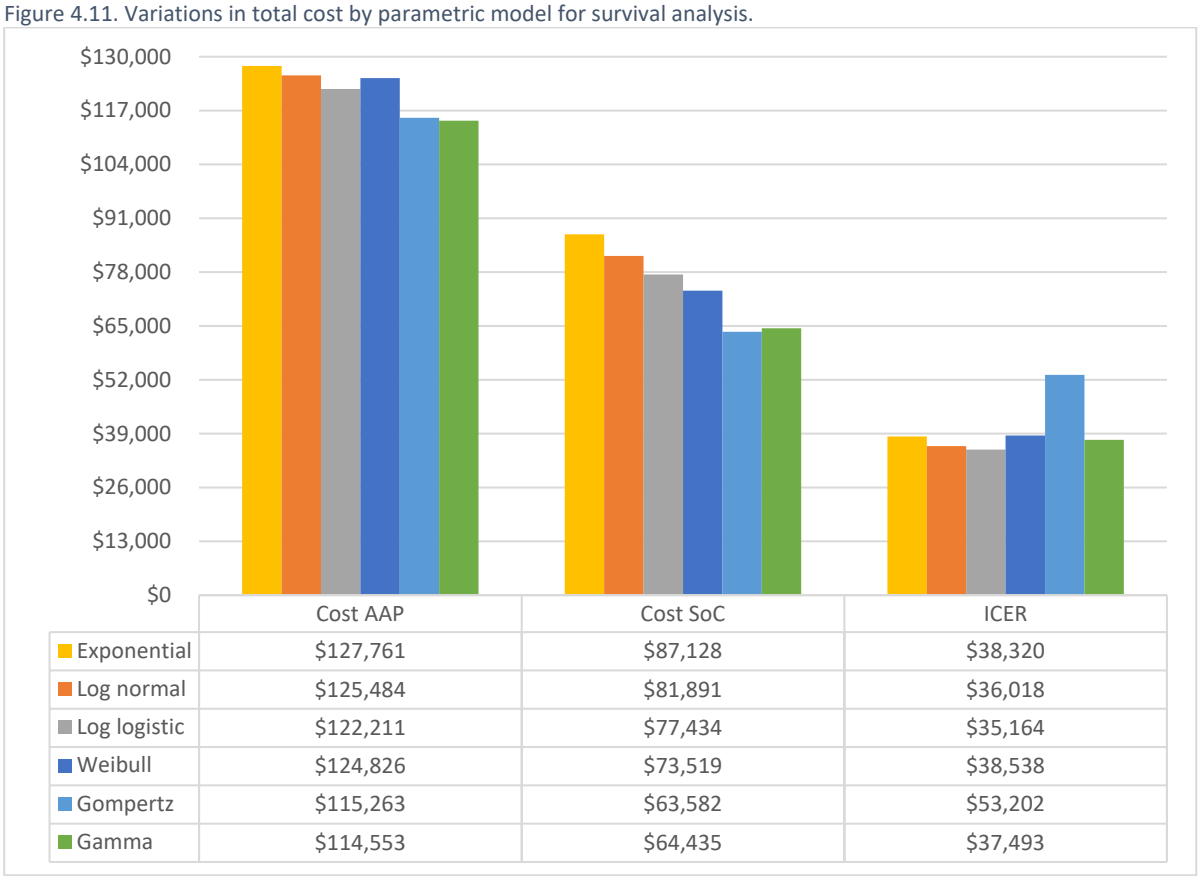
As expected, when the HR MFS predicts an HR OS closer to 1, AAP is dominated by SoC. HR MFS values closer to the STE (the surrogate threshold effect of MFS is 0.88) produce higher ICER values, which is in line with the uncertainty produced by the surrogacy relationship. When HR MFS is greater than the STE, the surrogacy relationship predicts a non-significant treatment effect on OS and this is translated into higher ICER values in the model (Chapter 3).

4.6.6.2. Scenario analysis

When all outcomes were modelled with the exponential distribution, it consistently estimated the highest costs and gains in QALYs (Figure 4.11 and Figure 4.12). In this scenario, the estimated hazard of failure is constant and independent of time elapsed[29] (203).

[29] The hazard in the exponential distribution is given by the equation $h(t) = \lambda$, where λ is a positive constant independent of the time elapsed, calculated by fitting the model to the observed data.

Figure 4.11. Variations in total cost by parametric model for survival analysis.



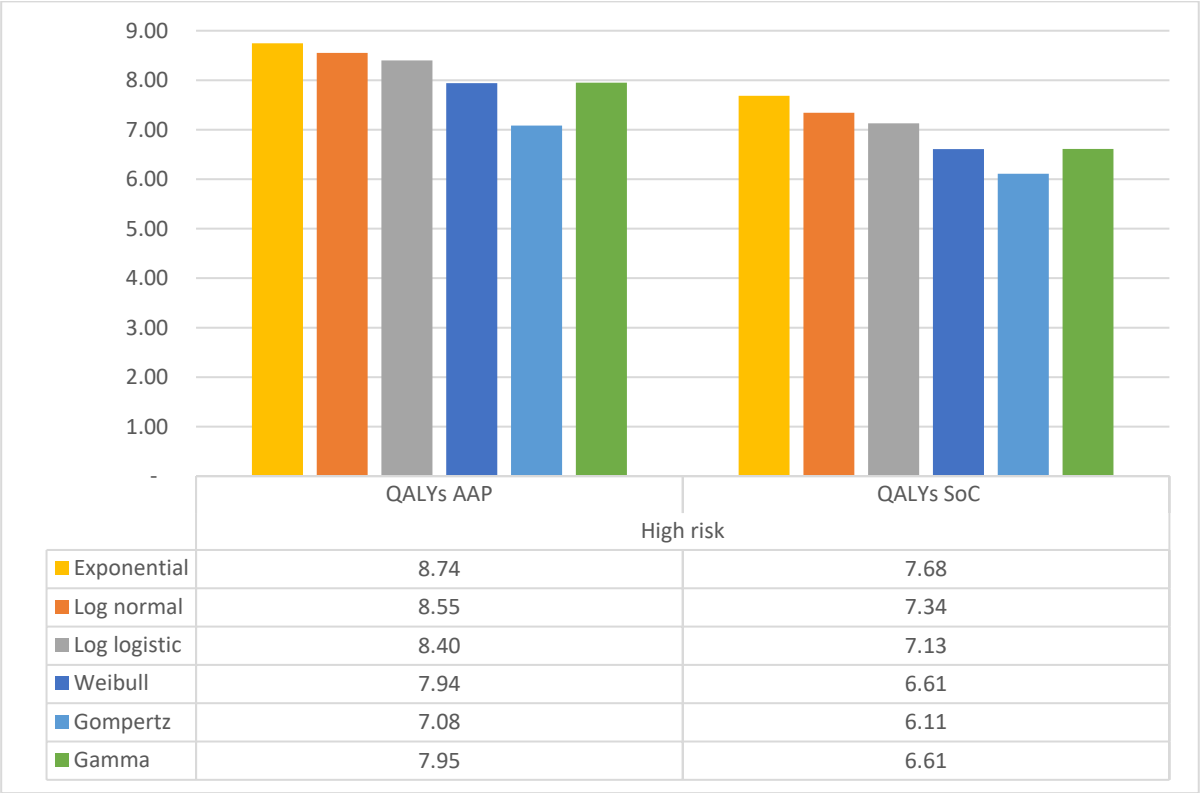
	Cost AAP	Cost SoC	ICER
Exponential	$127,761	$87,128	$38,320
Log normal	$125,484	$81,891	$36,018
Log logistic	$122,211	$77,434	$35,164
Weibull	$124,826	$73,519	$38,538
Gompertz	$115,263	$63,582	$53,202
Gamma	$114,553	$64,435	$37,493

Abbreviations: MFS= metastasis-free survival; AAP= radical prostatectomy; SoC= radiotherapy.

Parametric distributions where the hazard of failure is a function of multiple parameters (e.g., Weibull, gamma or loglogistic) showed the lowest values of AIC despite being penalised by the statistical method[30] (Chapter 3). As an example, in Weibull and loglogistic (the distributions selected for the extrapolations of outcomes in the model) the failure rate can be proportional to a power of time (unlike the exponential distribution where the hazard is constant and independent of time elapsed) (203). Increasing values of time, such as those observed in time to event data, led to a hazard that increased with time, which is concordant with models trying to depict the survival function of older patients.

[30] Comparisons between models can be made using the AIC: $AIC = -2\log L + \alpha q$, where L is the log-likelihood function, q is the number of parameters in the model and α is a constant. In general, more parameters will give a better fit (reduce log-likelihood function) and so there needs to be a penalty for fitting a model with more parameters. Values of the AIC will increase when unnecessary parameters are added to the model.

Figure 4.12. Variations in total QALYS gained by parametric model for survival analysis.



	QALYs AAP	QALYs SoC
	High risk	
Exponential	8.74	7.68
Log normal	8.55	7.34
Log logistic	8.40	7.13
Weibull	7.94	6.61
Gompertz	7.08	6.11
Gamma	7.95	6.61

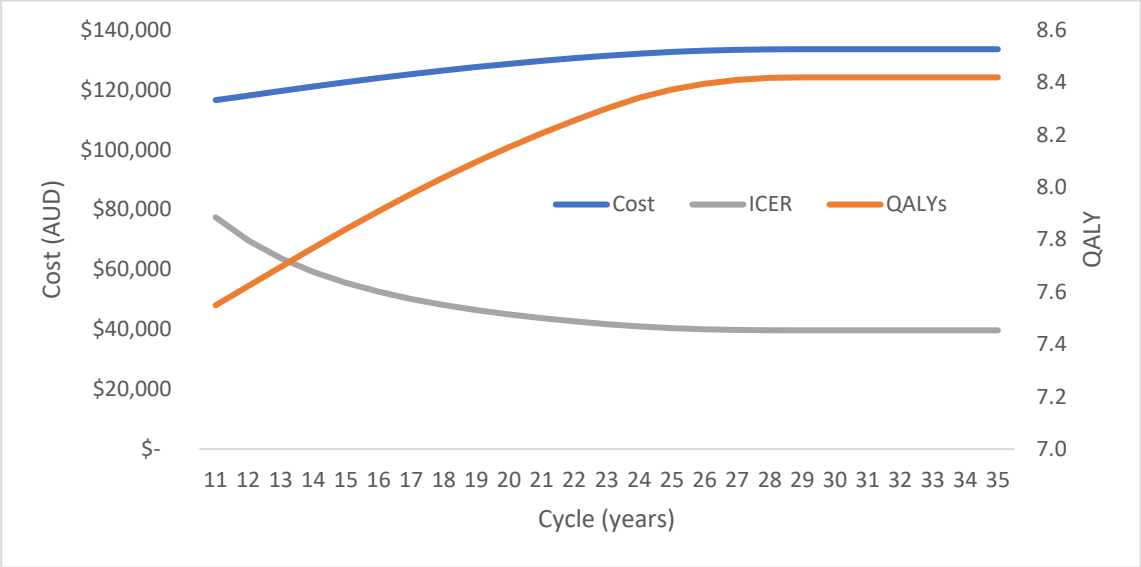
Abbreviations: AAP= abiraterone + prednisolone; MFS= metastasis-free survival; QALY= quality adjusted life years, AAP= radical prostatectomy; SoC= radiotherapy.

As shown in Figure 4.11 and Figure 4.12, the results of the model are influenced by the choice of the parametric model used to extrapolate the survival probabilities of the outcomes after the follow-up period. A relevant finding of the scenario analysis is the relative position of the parametric distributions used to extrapolate the outcomes in the base case analysis (loglogistic and Weibull distribution) across all distributions evaluated. In particular, the impact on cost and benefits when these distributions are used can be considered moderate when compare with the exponential and Gompertz, both sitting on the extremes of the scenario analysis.

4.6.6.3. Waning effect

The waning effect of the intervention was tested at different cycles. Figure 4.13 shows how the results of the model, in terms of total cost and QALYs gained per patient as well as the ICER, change when the effect of AAP over OS is assumed to decline and becomes equivalent to the treatment effect of SoC after the follow-up period. As expected, results are most impacted when the treatment effect of AAP, and therefore OS, is assumed to be the same as observed with SoC in the first year after the follow-up period (cycle 11 in this case).

Figure 4.13. Impact of the waning effect of the intervention on results.



Abbreviations: ICER= incremental cost effectiveness ratio; QALY= quality adjusted life year.

4.6.7. Validation

The results of OS reported by Attard et al. (125) were used to compare the results of the model when the survival outcomes in the intervention arm were inputted to the model directly from the clinical trial instead of being estimated by the surrogacy relationship. The surrogacy relationship estimates a HR OS of 0.601 whereas the HR OS reported by Attard et al. was 0.6. As expected, the results of the model did not change significantly when the HR OS is sourced directly from clinical trials. The ICER was $39,639 per QALY when the surrogacy relationship was used to estimate OS and $39,573 per QALY when the OS results form Attard et al. were used directly.

4.6.8. Benefits to society

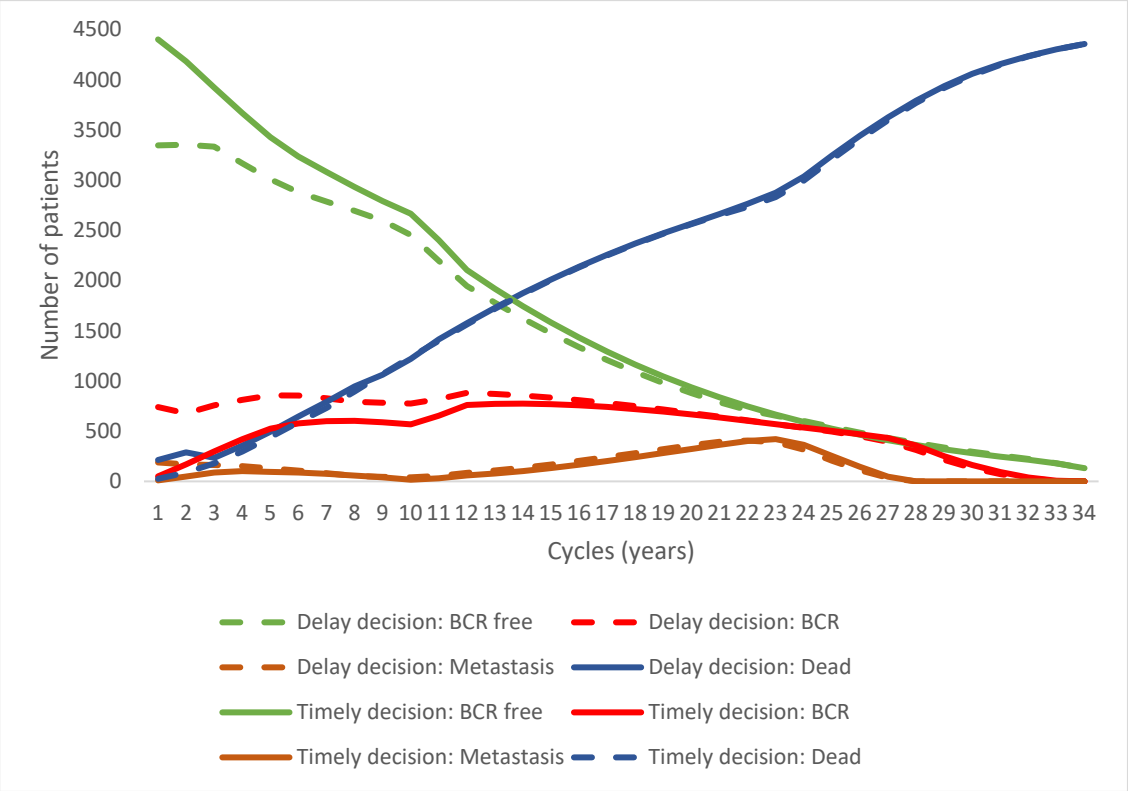
4.6.8.1. Flow of patients

The results of the model indicate that during the first 3 years, compared to the AAP group, 117 more patients in the SoC group develop BCR, 30 more patients develop metastasis and there are 63 additional deaths. These results illustrate the short-term benefits to society during the first 3 years of treatment for LPC.

To estimate the long-term benefits to society when the surrogate outcome is used in the decision-making process, the model is calibrated in the first cycle to reflect the short-term benefits produced in the first 3 years mentioned above. That is, if a timely decision is made, it is assumed that society

benefits from 210 fewer patients who are BCR-free during the first year (117 patients in BCR, 30 metastatic and 63 deaths). The flow of patients through health states between both the 'timely decision' arm and the 'delay decision' arm in the model is illustrated in Figure 4.14. The main difference in the flow of patient between AAP and SoC is the area between the BCR-free curves. Using MFS as a surrogate for OS to bring forward a decision to reimburse AAP provides more time in the BCR-free state for more men compared to waiting 3 years for OS to become available.

Figure 4.14. Flow of patients in the timely and delayed decision arms in the model.



Abbreviations: BCR= biochemical recurrence.

Importantly, when the model is modified to compare the 'timely decision' versus 'delay decision' the results tell us not only the cost-effectiveness ratio of AAP when compared with SoC but also the potential benefits arising from an earlier decision when surrogate outcomes are used in an economic evaluation to support the reimbursement decision-making process.

4.6.8.2. Delayed decision

To simulate the difference in costs and benefits to society arising from 3 years of missing the most effective treatment of the two being compared, the model was set to start with the patient distribution resulting from the 3-year delay and, therefore, starting in advanced stages of the disease rather than in the BCR-free state. When the treatment is not received in a timely manner, the average

cost per patient accrued in a lifetime is $11,560 higher, and 4.48 life years and 0.47 QALYs are lost (Table 4.19).

Table 4.19. Lifetime benefits to society of an earlier decision.

	AAP 3 years delayed	AAP timely decision	Difference	Avg/patient
Cost BCR	$73,727,341	$48,818,924	−$24,908,417	−$33,814
Cost MET	$10,716,647	$6,429,581	−$4,287,066	−$5,820
Cost BCR-free	$130,315,778	$150,995,480	$20,679,702	$28,073
QALYs MET	582.62	444.58	−138.04	−0.19
QALYs BCR	3,206.68	2,308.70	−897.98	−1.22
QALYs BCR-free	8,872.89	10,252.82	1,379.93	1.87
Life years			3,302.73	4.48
Total QALYS	12,662.18	13,006.10	343.92	0.47
Total Cost	$214,759,766	$206,243,985	−$8,515,781	−$11,560

Abbreviations: AAP=abiraterone + prednisolone; BCR= biochemical recurrence, MET= metastasis.

The results of the analysis of benefits to society show that, in terms of state membership, the largest cost impact is derived from the BCR state, while the greatest impact on QALYs arises from the BCR-free state. This highlights that most of the benefits gained from an earlier reimbursement decision result from keeping patients free from biochemical relapse.

4.6.8.3. Value of information

The EVPI was estimated for the LPC cohort using a Monte Carlo simulation to obtain the joint distribution of costs and benefits when AAP is compared to SoC. For this analysis, a cost-effectiveness threshold of $50,000 per QALY was used for the probabilistic sensitivity analysis. New interventions are normally replaced or displaced to a later line of therapy by more efficient interventions in the future, therefore, the value of information in the decision will depend on, among other factors, the expected effective lifetime of the intervention under analysis. Table 4.20 presents the EVPI at both an individual and population level, taking into account different effective durations of AAP being the dominant treatment in the market. As expected, given the impact of uncertainty on a decision that extends over a long period of time, the value of the EVPI is higher when a longer effective duration of AAP market dominance is assumed.

As the estimation of the EVPI is based on the difference in expected net benefit with perfect and current information, this value can be interpreted as the costs of uncertainty. For the cohort of patients that might receive treatment in Australia, the cost of additional information needed to offset

the uncertainty in the reimbursement decision ranges from almost $7.5 million to $12 million depending on the effective time of AAP considered. (273)

Table 4.20. EVPI at individual and population level at different effective lifetime of use AAP.

EVPI	Value
Individual	$620.30
Population-20y market dominance	$11,988,925.27
Population-15y market dominance	$9,993,089.58
Population-10y market dominance	$7,445,841.29

Abbreviations: EVPI= expected value of information; y= years.

If the cost of conducting additional studies is lower than the cost of the uncertainty estimated in the model, the results of the EVPI analysis should promote future research.

4.7 Discussion

MFS as a surrogate outcome for OS was used in a flexible economic model to compare AAP and SoC in patients diagnosed with LPC. A previous systematic review, conducted as part of this research, concluded that there were no published economic evaluations using validated surrogate outcomes in this patient population. This study is the first to use a validated surrogate outcome to predict OS in patients with LPC as part of an economic evaluation.

The study's model predicts that the average cost per patient in the high-risk population with prostate cancer is $133,491 for AAP and $88,657 for SoC. In terms of QALYs, the model estimates an average gain of 8.42 and 7.09 for AAP and SoC respectively, with an ICER of $39,640 per QALY for AAP compared to SoC. While there is no other cost-effectiveness analysis for AAP in LPC, the model's estimates can be compared to the results of other studies with regards to SoC. Comparison with the results of a recent cost-effectiveness analysis by Weng et al. (274) revealed similar results for the SoC arm, with an average cost per patient of $78,575 (2019 US Dollars), and an average gain of 6.86 QALYs (no ICER was reported for high-risk patients in this study). However, the study by Weng et al. did not specify whether ADT was added to RT as recommended by guidelines for high-risk patients. Another study by Dorth et al. estimated a cost of $75,600 (2013 US dollars) when ADT was added to RT in high-risk LPC patients (275), which is similar to the results of the present study (an ICER of $32,300 USD/quality adjusted life expectancy was reported in this study).

The results of the sensitivity analysis indicated that the model was more sensitive to variables such as time on treatment and the cost of AAP, rather than the comparative treatment effect of HR-MFS used

to estimate the treatment effect over OS. This implies that the model's results remain relatively stable even when extreme values of the comparative effect of MFS were used. (276)

The use of surrogate outcomes in economic evaluations can lead to uncertainty in reimbursement decisions due to the estimation of treatment effects on the final outcome through correlation with the observed treatment effect on the surrogate outcome. The sensitivity analyses conducted in this study revealed that the model predicted ICERs of approximately $140,000 per QALY when HR MFS values were above the STE. This uncertainty was reflected not only in the higher ICER values predicted by the model, but also in the non-significant predicted treatment effect. As discussed in Chapter 3, when HR MFS is near the surrogate threshold effect, the predicted treatment effect on HR OS is close to one, resulting in non-significance. In such cases, the uncertainty introduced by the use of MFS as a surrogate of OS, coupled with the high ICER value, does not justify the use of MFS as a surrogate outcome of OS to support the reimbursement decision-making. In this situation, the results of this research would recommend support waiting for OS data before making a reimbursement decision.

This uncertainty brought about by the use of surrogate outcomes to support reimbursement decisions could be perceived as too high for manufacturers that do not want to invest in new technologies with weak evidence or for reimbursement decision-makers (276). However, the validation process conducted by ICECaP and the assessment conducted in this thesis according to the PBAC framework provides confidence that the use of a validated MFS as surrogate for OS in economic evaluations could provide patients with LPC new interventions earlier. Therefore, in cases where there is ongoing uncertainty, a managed entry scheme that collects health outcome data can be crucial in promoting the use of surrogate outcomes in the decision-making process. This approach could potentially mitigate the risk associated with the uncertainty by involving both the decision-maker and the manufacturer in sharing the risk, through the implementation of risk-sharing agreements that are contingent upon the use of surrogate outcomes. This could allow for a reimbursement decision to be made based on the surrogate outcome while additional data is gathered to validate the surrogacy relationship.

4.7.1. Model of choice

The most popular models in economic evaluations in cancer were evaluated for use with surrogate outcomes. The structural basis of Markov models uses transition probabilities, which means that the surrogacy relationship can be modulated by post-progression probabilities needed to estimate OS, metastasis and BCR. If those post-progression transition probabilities are not available from the same

data source that provides the MFS data as a surrogate for OS, they need to be extracted from external sources. The latter introduces an external dependency on OS probabilities that may alter the underlying MFS to OS surrogacy link.

PSMs overcome this issue by allowing OS to be independent of other outcomes; therefore, the surrogacy relationship among both outcomes is not disturbed by uncertainties introduced by external sources of post-progression probability data. The use of surrogate outcomes in other types of models, like microsimulations, was not explored in this analysis. More complex models, such as multistate models, calculate state membership based on survival functions similarly to PSMs (277); however, further research is needed to assess potential challenges and to evaluate their feasibility.

4.7.1.1. Limitations of the model

The use of PSM has the advantage of being able to incorporate commonly reported clinical endpoints without access to IPD, which is generally not available for public use. However, it also has important limitations. The structural independence of the outcomes modelled can be considered a limitation. Outcomes in a PSM are implicitly dependent, which is reflected in the model within-trial period; however, when extrapolation is needed, this implicit dependency of the outcomes is ignored. This may result in the curves crossing, generating implausible results. Moreover, the implicit relationship among outcomes holds during the within trial follow-up period, but it is difficult to ensure that this relationship also holds throughout the extrapolated period. This issue becomes more important with shorter follow-up studies, as the full impact and waning of the treatment effect might not be captured in the study.

In contrast, in state transition models the transition probabilities from one state to another account for that relationship between outcomes and unrealistic results are avoided. However, it is unclear how these structural differences impact the cost-effectiveness evidence produced by the models and their ability to inform the decision-making process (79).

Another factor to consider is the parametric distributions used, which were estimated based on the hazards observed over the timeframe of the analysis. Different parametric distributions of the same outcome could be justified when the underlying treatment effect changes the natural history of the disease. However, choosing different distributions could incorrectly mean that one intervention is favoured over the other. Regarding this potential issue, evaluation agencies such as NICE recommend presenting the outcomes of different distributions in sensitivity analysis. The extrapolations of the

ICECaP outcomes used in this model were evaluated for each population and the same distribution was applied for SoC and AAP for each outcome.

Figure 4.11 and Figure 4.12 showed a high degree of variability when different parametric models were used to extrapolate the ICECaP data, but the difference is higher when the comparison includes the exponential distribution, which is unlikely to represent the cohort being modelled (203). The average difference between the highest and lowest estimates of cost when all six parametric models are compared among intervention and comparator is $64,179. The same analysis yields an average difference in QALYs of 2.63. These numbers, in contrast with the average difference in costs and QALY for patients in the base case analysis (difference in cost: $52,834 and difference in QALYS: 1.33), imply that the results are dependent on the parametric model of choice. This is accentuated in PSMs as they rely entirely on survival analysis to determine state membership.

It is important to consider that MFS as a surrogate outcome for OS estimated by the ICECaP meta-analysis has limitations when proportional hazards were assumed using extreme values. As seen in Chapter 3, the surrogacy relationship predicts abnormal results at extreme values of HRs. This is due to the fact that the correlation equation is bounded by the data used to analyse the relationship between MFS and OS. However, if the survival probabilities for MFS and OS are similar, there would be no need for the use of the surrogacy equation, as they would be considered equal. On the other hand, most studies are designed to produce a treatment effect in terms of HR of around 0.7 (132); therefore, a HR equal to 0.3 is very unlikely to be observed.

The surrogacy relationship applied in the preceding analysis correlates the comparative treatment effect for MFS with OS in terms of HR. The use of HRs as a measure of treatment effect implies the need for a common comparator to estimate the treatment effect on the intervention arm. As the relationship between outcomes has proved to be independent of the treatment provided (discussed in Chapter 3), a correlation in terms of the absolute impact of MFS on OS would allow greater flexibility in applying a surrogacy relationship for the purposes of constructing an economic evaluation. Absolute values of treatment effect would avoid the need to assume proportional hazards and a study comparing both interventions and reporting outcomes in terms of HRs.

The model developed was able to calculate the possible benefits to society arising from the use of MFS as a surrogate outcome for OS in a decision-making process that can potentially reduce the time to commencement of treatment by 3 years. However, decision-making dilemmas can arise because errors may result either from deciding too soon or from delaying decisions too long. This analysis

focused on the opportunity cost implied in delaying decisions; however, making a decision based on surrogate outcomes may introduce additional unmeasured uncertainties. This analysis did not evaluate the opportunity cost that could arise from making a premature decision, such as the costs to society of providing and administrating an ineffective intervention, or of potential harm to patients who receive the ineffective intervention rather than the more effective one, leading to decreased quality of life and increased morbidity and mortality. These issues represent an interesting area for further research.

4.7.2. The cost of uncertainty

The EVPI analysis showed that the cost of uncertainty for the cohort of patients modelled ranged from $7.5 million to almost $12 million depending on the assumed effective time over which AAP was the dominant therapy. This value represents the cost of uncertainty as using perfect information for every parameter in the model can eliminate the probability of making a wrong decision. Theory dictates that when the value of perfect information exceeds the cost of conducting future research, there is a need to conduct trials in order to confirm or overturn the decision made under uncertainty (80); however, it is difficult to put a cost on studies to reduce decision uncertainty.

A review of the cost of conducting oncology clinical trials for new drugs submitted to FDA between 2015 and 2016 found that oncology clinical trials cost between $49.7 to $122.9 million (in 2021 AUD) (273). These figures relate to the cost of de novo clinical trials; however, the cost of extension or rollover clinical trials would be the appropriate consideration in this instance. A report submitted to the FDA in 2014 estimated that the cost of Phase 4 trials (a proxy to rollover studies) in oncology was $38.9 million (in 2014 USD) (278). Compared to the cost of uncertainty estimated in the model (ranging from almost $7.5 million to $12 million for the population, depending on the effective time of AAP considered), investing on conducting further studies would not be economically viable for the reimbursement decision-making process.

The results of the EVPI analysis imply that using MFS as a surrogate outcome of OS provides sufficient information to make informed decisions without the need for additional clinical studies to explore this issue. However, there are other factors that need to be considered in this case. First, the use of MFS as a surrogate outcome to estimate OS may be considered as the primary source of uncertainty in making reimbursement decisions. Nevertheless, this uncertainty is not the only factor captured by the EVPI. The EVPI takes into account the results of the probabilistic sensitivity analysis, which considers the combined distribution of costs and benefits when all parameters are evaluated under uncertainty

(i.e., all parameter inputs are assumed to be variable). However, it does not accurately represent the opportunity cost due to a delayed decision, as other inputs considered in the probabilistic sensitivity analysis also contribute to the uncertainty measure used in the EVPI estimates.

Other analyses can help to elucidate the cost of uncertainty arising from the distribution of single or a selected group of parameters. The expected value of perfect information for parameters estimates the value of reducing the uncertainty surrounding a particular parameter in the model. As with the EVPI, the EVPI for parameters is calculated as the difference between the expected value with perfect and current information and not, as in this case, when a single parameter or a group of parameters varies. The computational burden of this analysis is extensive as it requires several samples of the results when the probabilistic sensitivity analysis is dependent on the distribution of the single parameter under study or the joint distribution of a group of parameters investigated. As the results of the EVPI for all parameters were low and indicated that no further studies were required, single parameter variation of the results is expected to have low or no impact over which alternative has the maximum expected net benefit (80). Therefore, the expected value of perfect information for single parameters (in our case, the translation of treatment effect of OS given MFS as per the surrogacy relationship) was not estimated.

The EVPI analysis results implied that the cost of conducting further studies is not justified given the low-cost uncertainty that would be displaced by the information that could be acquired from additional trials. This result is not surprising given the ICER predicted by the model when AAP is compared with SoC and given the relation of EVPI and the cost-effectiveness threshold. As the ICER estimated by the model is lower than the cost-effectiveness threshold assumed in the EVPI analysis, it is expected that AAP would be cost-effective and, therefore, the reimbursement decision based on this ICER is less likely to be modified by further research. However, it is important to consider that the recommendations interpreted from an EVPI analysis are driven by the threshold assumed. A lower cost-effectiveness threshold might lead to the opposite recommendations.

Finally, and most critically, the employment of MFS as a surrogate outcome for OS in economic evaluations provides information about the treatment effect, but not specifically about the comparative safety profile of the new intervention. Although the safety profile of interventions to be compared are considered in the model, by using the surrogacy relationship to estimate the treatment effect on the final outcome assumes that all safety signals have been observed within the timeframe of the MFS collection. Typically, regulatory bodies grant early approvals for interventions based on surrogate outcomes under the condition that the sponsor agrees to conduct follow-up trials to confirm

both the treatment effect and safety of the new intervention (279). As discussed before, such agreements are also used for reimbursement decisions for interventions relying on surrogate outcomes, which may lead to pricing arrangements tied to the production of additional evidence, not just to verify the surrogacy relationship, but also to confirm the comparative safety profile of the new intervention.

4.7.3. Health-related quality of life

Treatment-specific quality of life outcomes after AAP therapy in patients with high-risk LPC are needed to be able to compare it accurately with SoC. The difference in utility values used in this model is derived from patients receiving surgery as a proxy for the quality of life after primary treatment. Given the invasive nature of surgical interventions, it is probable that the quality of life for patients following surgery would be lower than for patients receiving AAP. Conversely, it is expected that patients receiving AAP would face an increase in adverse events due to hormonal treatment in the long-term and, thus, lower quality of life compared to surgery. The results of the sensitivity analysis conducted to assess the impact of varying utility values on the model did not produce any substantial effects on the model outcomes.

The results of the meta-analysis used to pool utility values in the model were consistent both with those reported in the literature and with the natural history of the disease. For example, metastasis had the lowest utility value among all health states in the model with a pooled result of 0.705 (95% CI: 0.638–0.771).

The meta-analysis conducted by Magnus et al. shows that several of the economic evaluations relied on utilities estimated pre-2007. Given the introduction of new treatments in the last decade targeting early and advanced stages of prostate cancer, and improvements in health literacy, it is possible that the utility values used in the economic model are outdated. Up-to-date analysis of patients' preferences is likely to better reflect current practice and health-related quality of life. Results from the ENZAMET trial are expected to bring that level of certainty to the model but are not yet available for inclusion (280).

Although the model allows for patients to experience more than one adverse event at a time, the utility values reported in the meta-analysis did not fully capture the difference in quality of life of

patients who experience multiple adverse events. Different methods exist to estimate joint-state[31] utilities; however, this is beyond the objectives of this analysis. Overall, it is not expected that the incorporation of joint-state utilities will substantially impact the cost-effectiveness; however, these may be relevant when high prevalence of events is anticipated (243).

4.7.4. Applicability issues

The model presented in this analysis compared AAP versus SoC, specifically considering Australian practice and costs in a context of high-risk LPC. However, it is flexible enough to incorporate parameters, assumptions and algorithms to suit specific needs in other settings as well as other interventions. As the model was based on the analysis of ICECaP, it can also incorporate the analysis of other populations of localised disease stratified by age and risk, as conducted in ICECaP.

It is important to note that the ICECaP included studies that recruited patients who underwent active treatment, as opposed to active surveillance or watchful waiting, current treatment alternatives for LPC (124). As most low-risk patients avoid receiving active treatment, this population will be underrepresented in the ICECaP trials. Nevertheless, the absence of low-risk patients opting for surveillance in the ICECaP study would not affect the validity of the model's results when used to evaluate intervention in low/intermediate-risk patients, as the model is specifically designed to simulate the treatment decisions made by patients who elect to receive active therapy.

The studies included in the ICECaP database were conducted between 1987 and 2011, where conventional imaging (CT + Bone Scan) was used as the main technique to stage and diagnose metastasise. This may raise some questions regarding the applicability of the finding in the longer term. With the arrival of Ga-PSA/PET-CT, a more specific and sensitive diagnostic technique (185), it is possible that the MFS population within the trials in the ICECaP database comprised a mix of patients free of metastatic disease and patients with metastasis that could not yet be identified through CT + bone scan. This may lead to potential applicability issues for future research, as the treatment effect may reflect the outcome for a mixed population.

Ultimately, it is important to consider that the results of this economic evaluation are limited by the current definition of metastasis under conventional imaging techniques. A new surrogacy relationship will be required if a new imaging technology replaces CT + bone scans as the gold standard in trials and

[31] A joint health state represents a health condition where an additional or multiple conditions or comorbidities exist concurrently.

clinical practice. At the moment, the prognostic value of Ga-PSA/PET-CT to detect metastatic disease is not known (281), so it is not yet possible to incorporate it as the basis for determining the surrogacy relationship or the clinical course of treatment.

This model provides reimbursement authorities with a means to evaluate the cost effectiveness of new interventions for localised prostate cancer when MFS is used as a surrogate outcome. As the use of surrogate outcomes impacts the duration of clinical trials and related costs, the results are also significant to clinicians involved in the design of clinical trials who may consider the use of outcomes aligned with those relevant to decision-makers. Likewise, health economics researchers and academics may have the opportunity to use a different approach when dealing with economic models that need to be informed by outcomes that imply a long duration of follow-up in localised prostate cancer.

4.7.5. Delayed decision

Time-wise, decision-making dilemmas emerge when a decision is postponed for too long or made too quickly. The consequences of a delayed decision can be measured in terms of the opportunity cost compared to a timely decision. The model estimated the possible benefits to society arising from the use of MFS as a surrogate outcome for OS in a decision-making process that can potentially reduce the time to commencement of treatment by 3 years. However, decision-making dilemmas can arise because errors may result either from deciding too soon or from delaying decisions too long. While this analysis focused on the opportunity cost implied in delaying decisions, deciding too soon based on surrogate outcomes may introduce additional uncertainties that are worth exploring in future research:

- Clinical: The use of surrogate outcomes may not fully reflect the actual clinical benefits and harms of the intervention. There is a risk that adopting a new intervention based on surrogate outcomes alone may result in unexpected clinical outcomes, either positive or negative.
- Safety: Surrogate outcomes may not reflect the safety profile of the intervention. There may be unknown or unexpected safety risks associated with the intervention that are not captured by the surrogate outcomes.
- Long-term uncertainty: The use of surrogate outcomes may not capture the long-term effects of the intervention, such as late-onset side effects or the development of resistance to the treatment.

- Economic: The adoption of a new intervention based on surrogate outcomes may result in additional healthcare costs that are not justified by the actual clinical benefits. This is particularly relevant when the intervention is expensive or has limited efficacy.
- Patient-centred: Adopting an intervention too soon based on surrogate outcomes may not reflect patients' perspectives or preferences and may not align with patients' values and priorities.

4.8 Conclusions

In conclusion, MFS as a surrogate outcome for OS can be used in a partitioned survival model to inform reimbursement decisions regarding new interventions for high-risk LPC patients in Australia. By implementing the surrogacy relationship into an economic model, it is possible to make a reimbursement decision earlier, which benefits patients through access to more effective interventions. When the reimbursement decision is brought forward by 3 years, society benefits from more patients remaining free of progression for longer under the new treatment, which translates into lower costs to treat recurrence or metastasis, higher quality of life and, ultimately, better survival outcomes when compared to waiting for information on OS to become available. The anticipated value to society demonstrated in this case study outweighs the costs associated with the additional research warranted to address the uncertainty associated with the surrogacy relationship.

Chapter 5 Discrete choice experiment

5.1 Abstract

A validated surrogate outcome to inform economic evaluations in localised prostate cancer (LPC) can potentially help bring new products into a timely, evidence-based decision-making process. To this end, a discrete choice experiment (DCE) was undertaken to elicit preferences about the use of surrogate outcomes and understand the trade-off between time and uncertainty among general public taking the role of decision-makers advisors.

Seven attributes (value for money, the extra cost to the Government, adverse events, proportion of men expected to benefit from treatment, the nature of those benefits, the completeness of the data available and the follow-up time required by the evidence) with 2 to 4 levels were defined. An online survey was conducted using a sample of the general population who were asked to assume the role of a decision-maker. Preferences of respondents were analysed using logistic regression to investigate the likelihood of a respondent supporting reimbursement of a treatment options for LPC.

The survey was completed by 1,003 participants. As expected, analysis of respondents' preferences showed an inclination to favour interventions for LPC that benefit a high proportion of patients (OR=2.10 CI 95%=1.91, 2.32), are better value for money (OR=0.81 CI 95%= 0.78, 0.88) and are low cost (OR= 0.61 CI 95%= 1.91, 2.32). Respondents also preferred interventions based on long follow-up time (OR= 1.20 CI 95%= 1.09, 1.31) and did not favour interventions supported by evidence from ongoing studies (OR= 0.83 CI 95%= 0.78, 0.88).

The analysis of responses when the role of decision-maker was adopted showed a preference for an intervention with less financial impact on the health system while providing maximum benefits to society. Respondents preferred the use of studies with longer follow-up time to support reimbursement decisions regarding new interventions, despite the delay in time to treatment. However, the uncertainty associated with the use of surrogate outcomes to support reimbursement decisions can be compensated for by interventions with a better cost-effectiveness ratio (ICER).

5.2 Introduction

5.2.1. Background

Reimbursement decision-makers are often faced with reconciling service demand and restricted resources when deciding on the reimbursement of innovative and high-cost health technologies

subsidised through public insurance programs. To maximise the efficiency of resource allocation in the societal context and to enhance accountability of the decision-making process in a utility-maximising economic framework, decision-makers usually establish what factors are considered and the relative importance of those factors in the decision-making process (282).

In a previous chapter, MFS was used as a surrogate outcome of OS in an economic evaluation, which resulted in benefits to society in the short and long term. The results presented in Chapter 4 provide encouragement for the use of surrogate outcomes for reimbursement decisions in LPC. However, decisions not based on clinical evidence reporting the final outcome can introduce uncertainty into the process and has generated scepticism (Chapter 1). Hence, the use of surrogate outcomes in the decision-making process has been regarded as controversial (55).

Deciding on funding for a new intervention based upon surrogate outcomes can increase uncertainty and the opportunity cost associated with their use is not always clear. Although surrogate outcomes can reduce the time it takes for new products to become available to patients, the benefits may be less than the possible harm when the impact on the final outcome has not been validated by empirical evidence (55). Therefore, when surrogate outcomes are used to inform economic evaluations used as part of reimbursement decision-making, another layer of information is added that policymakers need to consider and balance with all the other factors involved in the decision-making process.

A DCE was used to understand how trade-offs are made between data maturity and time to make treatments available. The methods, results and interpretation of the DCE are reported in this chapter.

5.2.2. Measuring preferences

For surrogate outcomes to be accepted as input into in the reimbursement decision-making process, it is necessary to understand the preferences of decision-makers towards their use relative to other factors affecting decision-making (e.g., cost, cost-effectiveness, AEs etc.). Reimbursement decision-makers need to balance the uncertainty arising from studies that are not supported by a final outcome against the time it takes to report significant OS results and, therefore, the delay in potential benefits provided by efficient interventions for newly diagnosed patients.

The preferences of decision-makers and the factors considered in the decision-making process can be revealed by analysing the choices made in a market where the alternatives of interest are traded. The process involved in this type of analysis is known as revealed preference measurement. However, the healthcare market is not an ideal market in which supply and demand determine an efficient trade of

goods, services and price. Unlike other markets – where the balance of demand and supply that occurs when sellers and buyers interact determines efficient prices of goods and services traded – the healthcare market generally fails to allocate resources efficiently and the equilibrium of supply and demand is normally not attained (283). It is usually the case that the consumers of health (i.e., patients) know relatively little about the process of care and their certainty is related to the services provided by the supplier (i.e., the physician), who can induce demand for healthcare services or goods (76).

Additionally, the cost of healthcare services is usually subsidised by insurance. As the level of subsidy impacts how costs affect individuals' choices, it follows that choices are not usually driven by patient preferences. In turn, understanding patients' preferences cannot be achieved by analysing their choices. Moreover, when a new good or service is introduced, the true market for that product often does not exist.

An alternative approach to analysing revealed preferences is to analyse choice behaviour in hypothetical but realistic scenarios. The techniques used to understand these stated preferences are based on a construct of virtual markets in which the researcher asks individuals to state their preferences over the alternatives under study. Hypothetical markets can be designed to facilitate the identification of all effects of interest associated with the choice of one alternative over another.

5.2.2.1. Stated preferences

There are two well-known approaches to collecting stated preference in hypothetical markets, contingent valuation models (CVMs) and discrete choice experiments (DCE) (284). In CVM, individuals provide monetary valuations of the impact on health by stating their willingness to pay for the interventions analysed under a hypothetical market (285). In a DCE, stated preferences are collected through surveys that present respondents with different alternatives available within a hypothetical but realistic market. By presenting multiple scenarios consisting of a range of values to respondents, DCEs are more sensitive than CVMs. In this Chapter, a DCE was used to elicit the preferences of the general population.

5.3 Objective

Elicitation of the preference of decision-makers can help to identify important characteristics of interventions that could be driving choices of a particular alternative over others and the marginal rate of substitution among them. Clinical evidence supported by surrogate outcomes to inform

reimbursement decisions can expedite the availability of treatments at the expense of introducing uncertainty into the decision-making process. In Chapter 4, the estimates of benefits to society when a surrogate outcome (MFS) was used in an economic evaluation comparing different interventions for LPC were presented. When MFS was used to inform the reimbursement decision, in the Australian context, there was a potential to prevent 1,834 patients from requiring early ADT and 128 patients from developing metastases and to save 96 lives by bringing reimbursement decision forward by 3 years. This translates to an average cost per patient saved of $13,765, 3.59 life years and 1.33 QALYs gained in a lifetime. Although these results are an incentive to adopt the use of MFS to evaluate new interventions in LPC, there is still hesitancy regarding the use of surrogate outcomes in the decision-making process (58) (286). In this chapter, a DCE was used to assess societal preferences associated with the factors affecting the reimbursement of new interventions for LPC, including but not limited to data maturity and time, the extent of clinical benefit and cost-effectiveness.

5.4 Methods

5.4.1. Theoretical framework

The theoretical framework of DCEs shares most elements of the standard economic theory of consumer behaviour. As with consumer theory, in DCEs there is an underlying assumption that individuals make rational decisions to maximise their utility within a context of budget constraints. However, there are some elements where DCEs are supported by extensions of the classic consumer theory. For example, where the latter explains individual choices based on a deterministic approach, DCEs assume probabilistic choice behaviour, thus random, which derives from the random utility maximisation framework (287). Another important difference in DCEs is that the utility assigned to a good by individuals can be explained by a series of attributes characterising that good (288). In contrast, in classic economic theory, the value of goods is homogeneous, and goods have an intrinsic value – it is understood as a property of a good. Finally, in DCEs individuals make choices based on a discrete and mutually exclusive number of alternatives, as opposed to the classic theory where the space for goods is continuous.

Grounded in random utility theory (RUT) proposed by Lancaster, DCEs assume that the choices individuals make over one particular alternative are driven by the benefits obtained compared to other alternatives available (288). It is also assumed that the benefits of an alternative can be defined by the properties or attributes characterising that alternative (the systematic and explainable component of the utility) plus a random and unexplainable component that accounts for preference

variation due to unobserved attributes or heterogeneity in taste and errors in measurement. Taking into account the choices individuals make, the comparative utility of an alternative j when individual i makes a choice is given by equation (Eq. 5.1):

$$U_{ij} = V_{ij}, + \varepsilon_{ij} \quad \text{(Eq. 5.1)}$$

where U_{ij} is the latent utility of individual i associated with an alternative j, V_{ij} is the systematic component of the utility specified by the attributes and ε_{ij} is the random component representing unmeasured variation of respondents' preferences.

The systematic component of the utility is usually assumed to have a linear relationship with the utility as a function of the attributes and levels characterising the alternatives being evaluated (284). The estimated coefficient associated with each attribute represents the preference weight or relative contribution of each attribute to the utility assigned by respondents to the alternatives. The random component of the utility function can explain heterogeneity in preferences among individuals, or preferences driven by explanatory variables not included in the model and other factors influencing the choice such as non-rational choices or random error (289).

5.4.2. Development of a DCE

A series of steps are involved in the design and execution of a DCE:

1. Definition of the choice decision problem
2. Identification of relevant attributes and levels
3. Design and choice set development
4. Experiment and data extraction
5. Statistical analysis
6. Interpretation of results

The aim of the current DCE is to elicit the preferences of the general public, who, for the purposes of the research, are asked to assume the role of advisors to reimbursement decision-makers. Therefore, in this DCE, respondents were asked to trade-off between time and uncertainty when facing reimbursement decisions regarding the choice of alternatives supported by surrogate outcomes.

5.4.2.1. Identification of relevant attributes and levels

The choice of attributes and levels is one of the most important steps in the development of a DCE and should be supported by evidence (290) as each attribute and its corresponding levels determine the

utility associated with the good or service under evaluation. All attributes that characterise an alternative and are important in affecting choice should be included in the survey design. Failure to include all relevant attributes may lead to biased results as the choice or the trade-offs captured in the analysis might be driven by absent attributes. On the other hand, too many attributes may contribute to the cognitive burden of respondents associated with greater information and thus increase the variances of the error term (291).

There are multiple approaches available to identify attributes relevant to respondents and those important in making a choice. To support the inclusion or exclusion of attributes, literature reviews, consultation with clinical experts, qualitative research or preliminary studies that provide evidence for relevant attributes in different settings can be employed (292). In this study, a 3-stage approach was used to identify the attributes to be included in the survey:

1. A focus group with stakeholders was formed to discuss the elements of the decision-making process in the setting of localised prostate cancer.
2. A literature review was undertaken of published studies investigating stated preference of decision-makers in prostate cancer setting.
3. Discussion with the ICECaP Steering Committee.

5.4.2.1.1. Focus group

Focus groups to identify attributes and levels have been used widely in the DCE literature (293) (294). In this methodology, the researcher adopts a secondary role and facilitates the discussion among the participants to obtain their opinions, experiences, and attitudes towards the research topic (295).

A focus group was held with stakeholders to discuss the factors that influence decision-making in the treatment of localised prostate cancer. The focus group was structured and guided by Professor Richard de Abreu Lourenço, who has previous experience in the dynamics of focus group to support the development attributes for DCEs (296) and particularly in the prostate cancer setting (297). The facilitator introduced the purpose of the project, outlining the definition of surrogate outcomes and how their use might impact the decision-making process and encourage the discussion of factors influencing clinical, patient and reimbursement decisions.

Focus group participants were chosen based on the relevant information they could provide to describe factors important to the decision-making process. Therefore, it was important that

participants had experience in the decision-making process in healthcare. For this purpose, the focus group was formed by 6 participant members of the following groups:

- Clinician (1)
- Reimbursement agency (1)
- Industry (2)
- Patient with previous diagnosis of LPC (1)
- Researcher (1)

The meeting was recorded with the consent of all participants. The transcribed proceedings were analysed to identify key themes that would support candidate attributes.

The question of interest for the focus group was whether a treatment for LPC should be recommended for funding. Thus, the review of the group feedback considered those aspects of the discussion which were relevant to this topic; how public reimbursement decisions are made. The resulting themes that informed the candidate attributes and supporting quotes are shown in Table 5.1.

Table 5.1. Attribute domains from focus group

Theme – Attribute Domain	Comments
Context/condition	'decision making is a spectrum…it's not going to be one size fits all'. '(localised) compared to de novo metastatic disease, compared to castration resistant disease…' 'understanding the treatment and the treatment algorithm… is really important'
Life Expectancy/prognosis	'vastly different stages and symptoms and you're looking at the difference between somebody whose life expectancy might be 10 to 15 years versus somebody whose life expectancy is five versus 18 to 24 months.'
Basis for change (Endpoint to affect change)	'we've kind of got two schools of thought ..one school of thought is well, around survival and that drives their decision, and another school of thought is well, there is a quality of life benefit being shown and sort of is that intermediate time point.' 'the prognostic factor doesn't help me unless I can say how will it change treatment.' 'maybe the quality of life on its own could be enough to ..change practice um, and then the survival is cream on top'
Confidence/familiarity	'from a consumer view point, ..I don't care what it is, but as soon as you're confident that it might work, I want to know about it' 'that once you know that something is [a] good end-point, reliable end-point that you can then recommend the drug'
Timing	'If the science says that something's a surrogate for waiting another 15 years, and enough people ..agree, that that is correct, then that should be sufficient in itself. ' 'the information that you have in 15 years' time might be relevant for today, but it might not be relevant in 15 years' time. '
Recommendation	'But there are patients who like a plan made out of for them. You know, a bit like, this is what you could do. These are the options and this how I'll stage you through it.' 'I get the gradation of the patient who comes to me with 200 pages from the internet and we go through, or from the patient who says, 'I am scared. You are the doctor. I came to you as the expert, I want you to make the decision.' And then there's everything in between. And as a clinician you have to cater for all of those decision making um, pathways, because your job is to make the patient comfortable with whatever decision finally happens.'

5.4.2.1.2. Literature review

After the list of candidate attributes was developed from the results of the focus group discussion, a systematic review of the literature was conducted to identify studies analysing preference of decision-makers for reimbursement of new interventions in prostate cancer. The main objective was to compare the candidate attributes identified in the focus group and to assess whether there were other potentially relevant attributes not identified by the focus group.

Cochrane, CINAHL, EconLit, MEDLINE, Psychology and Behavioral Sciences Collection APA, PsycInfo APA and PsycArticles were searched. However, no studies were found which met the inclusion criteria (preferences of reimbursement decision-makers for interventions in prostate cancer). As there were no studies found when these two search topics were used together, a separate targeted literature search was used to identify attributes important when choosing treatment for prostate cancer patients (regardless of the decision-maker) and attributes used for health technology assessment

(regardless of the intervention or healthcare problem) in the decision-making process. The search was limited to literature published in English.

The following two sections describe the results of the literature search. Section 5.4.2.1.2.1 presents the relevant attributes found in the literature when decision-makers (patients, physicians and general population, as there were no studies found describing the preferences of reimbursement decision-makers) choose a treatment for prostate cancer. The results in section 5.4.2.1.2.2 describe the relevant attributes used to characterise treatment in HTA.

The results of the literature search formed the basis for the development of the attributes' levels. The levels found in the literature were discussed among experts from the steering committee to reflect possible values considering the Australian clinical setting.

5.4.2.1.2.1. Relevant attributes when choosing treatment for prostate cancer

A total of 190 studies were screened for title and abstract after removing duplicates, from which 60 full text studies were assessed; 22 studies (298-319) met the inclusion criteria and attributes identified in them were extracted for further analysis. There were no studies found which measured the preferences of reimbursement decision-makers in prostate cancer. The studies which reported the choice of treatment for patients diagnosed with prostate cancer presented in this section focus on preferences from the general population, patients and physicians. A total of 12 themes were defined from which 99 different possible attributes were identified (Table 5.2).

Table 5.2. Attributes identified in the literature review for prostate cancer.

Theme	Attributes	Source
Cost		Jayadevappa et al. (2019)
	Out of pocket expenses	Lloyd et al. (2008)
		Rentz et al. (2021)
	Cost	Richman et al. (2005)
	Out of pocket expenses	Sculpher et al. (2004)
Effect on ability to work	Effect on ability to work	Feldman-Stewart et al. (2004)
	Work loss due to treatment	Uemura et al. (2016)
Others	Taking action	Antonio et al. (2018)
	Others' support (doctor and family recommended)	Antonio et al. (2018)
	Frequency of PSA testing with a risk of new biopsies	de Bekker-Grob et al. (2013)
	Steroid use, Food restrictions	Eliasson et al. (2017)
	How to know if treatment is working, What other patients choose, What doctor would choose, Effect on family, Doctor's experience, Getting a second opinion, My age.	Feldman-Stewart et al. (2004)
	Family was very supportive, Anxious of living with cancer	Gorin et al. (2011)
	Care giver burden	Jayadevappa et al. (2019)
	Immediate action, Others approve	Kaplan et al. (2019)

Theme	Attributes	Source
	Control of cancer	Richman et al. (2005)
Quality of life	Effect on quality of life	Feldman-Stewart et al. (2004)
	Effect on usual activities	Feldman-Stewart et al. (2004)
	Social interaction	Jayadevappa et al. (2019)
	Expected QOL	Nakayama et al. (2018)
	Life quality	Richman et al. (2005)
	Time efficient or interfere with daily life	Underwood et al. (2010)
	Return to normal activities	Watson et al. (2020)
Recommended by doctor	Doctor's advice	Feldman-Stewart et al. (2004)
	Doctor recommended	Gorin et al. (2011)
	Physician recommendation	Volk et al. (2014)
Recovery after treatment	Time to recover	Feldman-Stewart et al. (2004)
	Recovery time	Jayadevappa et al. (2019)
Recurrence treatment options	Options if cancer doesn't disappear, Options if cancer gets worse, Options if cancer comes back, other options (alternatives).	Feldman-Stewart et al. (2004)
Side effects	Bowel issues, Urinary incontinence, Sexual functioning.	Antonio et al. (2018)
	Risk of urinary incontinence, Risk of erection problems, Risk of other permanent side effects.	De Bekker-Grob et al. (2013)
	Tiredness and fatigue, Bone pain, Nausea/vomiting or diarrhoea, Risk of infection caused by treatment affecting your immune system	de Freitas et al. (2019)
	Effectiveness, Fogginess, Bone pain, Drug-drug interactions, Tiredness.	Eliasson et al. (2017)
	Effect on bladder functioning, Effect on sexual functioning, Effect on bowel functioning (colostomy), Other side-effects, Effect on bowel functioning (meds), Effect on discomfort, Chances of dying from treatment.	Feldman-Stewart et al. (2004)
	Erection, incontinence	Gorin et al. (2011)
	Urinary function, Bowel function, Side effects, Sexual function, Psychological distress.	Jayadevappa et al. (2019)
	Bowel function, Urinary functioning, Sexual functioning.	Kaplan et al. (2019)
	Erectile dysfunction, Urinary leakage, Urinary blockage, Bowel symptoms, Fatigue, Hormonal effects, Loss of libido.	King et al. (2012)
	Risk of Haematuria, Risk of Diarrhoea, Risk of Gynecomastia	Lloyd et al. (2008)
	Side effects	Nakayama et al. (2018)
	Pain, fatigue	Rentz et al. (2021)
	Major complications	Richman et al. (2005)
	Blood product transfusion risk	Richman et al. (2005)
	Ability to get or maintain erection, Diarrhoea, Sex drive or libido, Physical energy, Hot flushes, Breast swelling or tenderness.	Sculpher et al. (2004)
	Fatigue, Skin rash, Serious fall, Serious fracture, Time to Pain Progression, Fatigue, Skin rash, Serious fall, Serious fracture, Cognitive problems,	Srinivas et al. (2019)
	Fatigue, Falls or fractures, Time until cancer-related pain, Rashes, Cognitive impairment, Hypertension.	Suzuki et al. (2021)
	Bone pain, Fatigue, Time until pain, Falls or fractures, Rashes, Fatigue, Cognitive impairment, Hypertension	Uemura et al. (2016)
	Side effects	Underwood et al. (2010)

Theme	Attributes	Source
	Side effects	Volk et al. (2014)
	Erectile function, Urinary function	Watson et al. (2020)
Survival	Lifespan	Antonio et al. (2018)
	Chances of dying of cancer, Effect on length of life	Feldman-Stewart et al. (2004)
	Survival (survival)	Jayadevappa et al. (2019)
	Life span	Kaplan et al. (2019)
	Additional Length of Life	Lloyd et al. (2008)
	Survival	Richman et al. (2005)
	Life expectancy	Sculpher et al. (2004)
	Survival, Overall survival	Srinivas et al. (2019)
	Overall survival	Uemura et al. (2016)
	Cancer specific survival	Watson et al. (2020)
Treatment convenience	Cutting (surgery)	Antonio et al. (2018)
	Mode of administration	de Freitas et al. (2019)
	Procedure, How treatment works, When treatment can start	Feldman-Stewart et al. (2004)
	Cutting (surgery), Radiotherapy	Gorin et al. (2011)
	Treatment duration, Need of cutting, Radiation or seeds, Cancer control (follow-up)	Jayadevappa et al. (2019)
	No cutting	Kaplan et al. (2019)
	Convenience (delivery of treatment)	Lloyd et al. (2008)
	Convenience of treatment	Nakayama et al. (2018)
	Dosing	Rentz et al. (2021)
	Administration	Uemura et al. (2016)
	Treatment modality	Watson et al. (2020)
Treatment effectiveness	Main aim is cure	de Bekker-Grob et al. (2013)
	Cure	Gorin et al. (2011)
	Cure	Richman et al. (2005)
	Cancer cure	Underwood et al. (2010)
	Cure	Volk et al. (2014)
	Treatment benefit or effectiveness	de Freitas et al. (2019)
	Chances of symptoms from spread	Feldman-Stewart et al. (2004)
	Treatment effectiveness	Nakayama et al. (2018)
	Time until cancer metastasises	Rentz et al. (2021)
	Not needing more cancer treatment	Watson et al. (2020)
Severity of the diseases	Chances of PSA rising, Rate tumour grows	Feldman-Stewart et al. (2004)
	Cancer recurrence	Jayadevappa et al. (2019)
	Time to progression	Srinivas et al. (2019)
	Time until a complication of bone metastases	Uemura et al. (2016)

5.4.2.1.2.2. Relevant attributes used in health technology assessment

The targeted literature search for attributes used in health technology assessments found a systematic review of DCEs that helped identify attributes and levels used to evaluate health technologies that

could potentially be implemented in a DCE for the decision-making process (320). This systematic review and meta-analysis focused on DCEs to reflect preferences of patients, policymakers, providers, and the general public that were part of a health technology assessment (HTA). As HTAs are used to inform the reimbursement decision-making process, the attributes pooled in this study are assumed to be of importance for the decision-makers in the reimbursement process of new technologies.

As showed by Trapero-Bertran et al. (320),there are some attributes common to a variety of health technologies. In effect, all categories of attributes identified in section 5.4.2.1.2.1 in the specific setting of prostate cancer were covered in the list of essential attributes to be used in priority-setting decisions proposed by Trapero-Bertran et al.

5.4.2.1.2.3. Comparison of attributes

A cross-reference of the attributes identified for the treatment choice of prostate cancer with those found in HTAs in general and those that were identified from the focus group is provided in Table 5.3.

The attributes found in the literature differed from those identified by the focus group. The population, settings and the topics discussed in the focus group compared to the survey design used in the literature may explain the differences. The focus group included a heterogenous population from different spectra of the decision-making process, whereas the population surveyed in the studies were usually homogeneous and structured. Additionally, the aim of the studies differed from the ones outlined in this chapter and therefore the attributes included differed as well. When comparing the attributes identified by the focus group with those found in the literature, the following is noted for each of the themes identified from the focus group:

- Context (the decision been made): appeared with multiple definitions such as the avoidance of surgery or radiation, mechanism of action of treatment, options of alternative treatment and treatment if cancer recurs, or follow-up schedule required (298), (302), (303), (304), (305), (310).
- Basis for change (endpoint to affect change): the type and magnitude of outcomes used as the basis for change focused on clinical definitions of outcomes, such as cure (including survival), time to progression, and avoidance of recurrence or bone pain ((298) , (302), (304), (305), (307), (310), (311), (312, 313), (315) , (319), (299), (300), (301), (303), (308), (309), (317), (318), (300).
- Confidence (familiarity with the available treatments under consideration): no papers noted the inclusion of confidence or familiarity with the treatment.

- Recommendation (advice received regarding the available treatments): appeared in studies found for attributes in prostate cancer ((298), (302), (303), (318).
- Treatment delay (time to treatment): appeared only as an attribute for HTAs (320) and in one paper found in the literature review of prostate cancer attributes (302).

Table 5.3. Attributes and levels for the health decision-making process identified by the systematic review of Trapero-Bertran et al.2019 and the focus group.

Attributes found by Trapero-Bertran	Cross references attributes found for PC	Context for the decision
Attributes identified in the focus group		
Improvement in health	As time to progression (309) (312) (315) As cure (299), (303), (310), (310), (317) Avoid recurrence (302) (304). As survival: (298), (302), (304) (305) (307), (310), (311), (313), (312), (313), (315) As quality of life (302), (304), (308), (310), (317), (319) Others (301), (308) (300)	Physician (312). Patients (309), (315), (303), (317), (318), (302), (304), (307), (311), (319), (301), (301). Patients and physicians (299), (310), (308) Adult men with negative biopsy (305)
Waiting time for the treatment or treatment duration	Time to treatment (302)	Patients
Attributes not identified in the focus group		
Severity of the disease	Chances of PSA rising (302) Rate tumour grows (302) Cancer recurrence (304) Time to progression (312) Time until a complication of bone metastasis (315)	Patients (315), (302), (304). Physician (313).
Value for money	NI	
Side effects	Sexual function (298), (299), (302), (303), (304), (305), (306), (311), (319) Urinary toxicity (303), (298), (299), (304), (305), (306), (319) Bowel toxicity (298), (302), (304), (305), (306), (307) Pain (300), (301), (309), (312), (314), (315, 316), (321), (300) Tiredness and fatigue (306), (309), (313), (312), (314), (315, 316), (301), (300) Others (310), (311), (302), (312), (314), (316), (312, 313), (301), (306), (300), (304), (307), (308), (317), (318), (321)	Patients (309), (315), (301), (302), (303), (304), (307), (311), (319), (300), (298), (306) (315, 316), (321), (311), (317) . Patients and physicians (299), (310), (308). Adult men with negative biopsy (305). Physician (322), (314).
Cost (price) of treatment	Out of pocket cost (304), (307), (309), (310), (311), (321)	Patients (309), (304), (307), (321), (311). Patients and physicians (310).

Abbreviations: NI= Not identified; PC= Prostate cancer; PSA= prostate specific antigen.

No studies focusing on prostate cancer described the attribute cost-effectiveness or value for money. This was expected given the decision-context in the studies is from the perspective of patients and prescribers. Also, only one study (302) included the time to treatment as an attribute for patients

when deciding for treatment. However, that study was focused on the difference of active surveillance, RT or RP, not on the delay to treatment using a new intervention due to the decision-making process. No studies on decision-making in prostate cancer included delay in treatment delivery due to time needed for the final study results to become available. However, in the systematic review of Trapero-Bertran et al. for attributes to be used in HTA, 'waiting time for the treatment or treatment duration' was one of the six most common attributes found (320).

While decision delay due to data availability and cost-effectiveness or value for money were not included in the studies found in the literature for prostate cancer, these are central questions to this project: how potential delays in the receipt of 'final data' and the value for money of new interventions might be traded off against other elements of reimbursement decision-making. Also, consideration of cost-effectiveness is a key input to public reimbursement decisions in Australia, and consequently they were retained as attributes in the final survey.

In addition, most of the studies found (in the literature review of prostate cancer and the systematic review of HTAs) reflected the potential for side effects associated with prostate cancer treatment. This suggested the appropriate inclusion of the likelihood for side effects as a candidate attribute for consideration. Likewise, the attribute health benefits was common to the focus group and both reviews and therefore, was included in the final survey.

Four studies (298), (302), (303), (323) found in the prostate cancer literature included the attribute recommendation (either by physician, family, or other patients). Physician's recommendation of treatment was also a relevant topic arising from the focus group analysis. This is an important attribute to consider when comparing treatment alternatives available for patients, particularly when those options involve active surveillance (318). However, the interest of this project is on the decision-making process for reimbursement purposes and therefore this attribute was not included in the final survey, as it did not focus on clinical decision-making.

Five studies (300) (307), (308), (309) (315, 316) also included an attribute to reflect the mode of treatment administration or potential convenience of treatment delivery. The inclusion of such an attribute would be relevant if a particular treatment was being considered. In this instance, the decision context is not treatment specific, so mode of administration was not included as an attribute.

From the systematic literature search of attributes used in DCEs of prostate cancer, only three studies included quality of life attributes (302, 308, 310). On the other hand, the systematic review of attributes used in HTA (320) included an attribute called 'improvement in health'; however, it did not

specify measures of improvement such as quality of life. Although primary treatments (such as radiotherapy and radical prostatectomy) for LPC impact the quality of life of patients in the short term, the negative impacts tend to diminish over time and the diagnosis by itself does not seem to affect general domains of health-related quality of life (324). The benefits of interventions measured in terms of survival and symptoms of the disease were more commonly used to characterise benefits of interventions in LPC (survival or expected life expectancy was included in 10 of the studies found in the literature review of attributes in LPC (298, 302, 304, 305, 307, 310-313, 319) and symptoms of the disease were included in virtually all studies found). As the measure of benefits described before was a factor in the focus group and is closely related to the natural history of LPC, it was included as measure of benefit from the intervention in the DCE.

5.4.2.2. Discussion with the ICECaP Steering Committee.

The assessment of the candidate attributes discussed in the focus group with those found in the literature informed the list of attributes and levels piloted among a sample of 104 individuals of Australia general population. Those completing the pilot were asked to provide feedback regarding the clarity of the concepts and question and the level of comprehension of complex tasks and attributes in the survey (see section “Follow-up questions in the pilot survey” in Appendix 9). The feedback provided was discussed among the committee members and changes were made to the description of attributes and the survey text. Appendix 10 contain a summary of results obtained in the pilot survey and changes proposed. The final list of attributes and levels included in the survey as discussed with the ICECaP Steering Committee is presented in Table 5.4.

Table 5.4. Attributes and levels included in the survey.

Description	Levels	Expected direction
The benefit to men from treatment is:	The disease does not progress to their bones or other parts of the body. There is no evidence of disease in their blood. They have no symptoms of the disease. The disease does not impact on how long they live.	NA
The proportion of men who are likely to benefit from treatment:	10 in 100; 20 in 100; 30 in 100; 40 in 100	(+)
The proportion of men who would experience side effects from treatment	15 in 100; 30 in 100; 45 in 100; 60 in 100	(-)
The cost to the government required to produce an additional year of full health compared with usual care is:	$25,000; $50,000; $75,000; $100,000	(-)
The additional cost per year to the national health budget from funding this drug is:	$10,000,000 $50,000,000 $100,000,000 $200,000,000	(-)
The evidence for the treatment is from a:	Completed study An ongoing study	NA
The evidence available from the study reflects:	4 years of experience. 6 years of experience. 8 years of experience. 10 years of experience.	(-)

Abbreviations: (+) = positive relation with utility, (-) = negative relation with utility, NA= not applicable.

It was hypothesised that attributes related to benefits had a positive relationship with the utility (meaning that utility values predicted by levels with higher values are higher than those predicted when levels with lower values are used). The opposite was hypothesised for attributes related to cost and AEs. For the time attribute, a negative relationship was expected, meaning that levels implying a longer waiting time would be associated with lower utility values when compared to the utility derived from interventions supported by studies with a shorter time (expected direction of attributes is shown in the last column of Table 5.4).

5.4.2.3. Design

The DCE survey presents respondents with different scenarios where the levels of the attributes characterising the available alternatives to choose from have been varied. Known as full factorial designs, ideally the estimation of preferences can be obtained by presenting respondents with all attributes and levels identified. (284)

In this analysis, the six attributes identified with four levels and one attribute with two levels translate into a full factorial design containing 8,192 possible combinations of attributes and levels for each choice in the survey (2^1*4^6). For a pairwise design, all possible combination of attributes and levels

yield the unmanageable number of over 33 million choice sets (8,192*8,191/2). Yet, not all possible combinations contribute new information about the marginal rate of substitution or relative utility of the intervention under analysis, as many of the scenarios are redundant (284).

Due to the complexity of the implementation of full factorial designs, fractional factorial designs are normally used. Fractional factorial designs exploit the redundancy characterising full factorial designs, thus allowing the reduction of the burden involved in conducting the full factorial design experiment by reducing the number of choice sets to those providing information about the most important features (284).

Generally, the utility derived from an alternative can be estimated by analysing the proportional utility (part-worth utility) that could be obtained with each level of an attribute. Designs developed to estimate this independent preference, focus on the main effects over preference in a choice alternative derived from each level thus reducing the number of choice sets presented to respondents by limiting them to those providing information about main effects. Nevertheless, interactions between attributes may be observed, suggesting that preferences for one particular attribute could be explained by the levels of other attributes in the experiment (284). However, it has been argued that interaction effects can only explain a small variation in preferences and main effects designs are mostly used for DCE in the health economics literature (325).

The fractional factorial design used in the DCE was developed under the guidance of Professor Deborah Street. The ideal fractional factorial design has to be able to estimate the utility function ensuring statistically efficient experiments that allow for the precise estimation of parameters. It is also important to have a balanced design in which respondents are not exposed to excessive cognitive burden and, at the same time, the choice set presented to respondents reflects real market situations (326). The design approach proposed by Street and Burgess (327) achieved these objectives while providing properties of optimal choice design (328).[32]

Additionally, the choice sets used in the DCE were subdivided to reduce the cognitive effort for respondents, which promotes the efficiency of the experiment (329). The fractional factorial design

[32] Huber and Zwerina (320) propose that efficient choice designs are characterised by four properties:

1) Level balance: when all levels of the attributes appear with equal frequency in the survey.
2) Orthogonality: when the joint occurrence of any two levels of different attributes occurs in proportional frequencies.
3) Minimal overlap: when the design reduces the probability of same level repetition for attributes appearing in each choice set to a minimum.
4) Utility balance: When respondents are equally attracted to the option given in a choice set.

developed containing 64 choice sets was distributed into 8 blocks, each containing 8 different choice sets and each respondent in the survey was randomly allocated to one block.

5.4.2.4. Survey conduct

The survey was loaded, hosted and managed using the PureProfile (https://www.pureprofile.com/) online platform specialising in surveys and choice experiments. Through its online survey engine, PureProfile recruited members representative of the Australian general population, coded the scenarios presented, and collected the survey data.

To characterise respondents according to demographic characteristics, a set of follow-up questions were asked. These were related to self-report of health status, confidence with medical forms and health material, previous cancer diagnosis, sex, age, place of residence, language at home, education level, employment status, income and relationship status.

Additionally, to assess the level of complexity of concepts and how well they were understood, a set of follow-up questions were included in a pilot survey implemented among 104 respondents. Appendix 9 describes the follow-up questions in detail.

5.4.2.5. Choice set development

With attributes and levels selected, choice sets were developed to allow respondents to choose between hypothetical alternatives formed by a combination of attribute levels as informed by the mathematical design. Figure 5.1 shows an example of a choice set. As the main objective of the DCE was to assess the values and preferences of the general population when asked to take the role of decision-makers evaluating a new intervention for LPC, a vignette was used to assist respondents to imagine the healthcare reimbursement setting. The vignette asked respondents to assume the perspective of an advisor to an HTA agency (see the text box below) while explaining characteristics of LPC and less familiar concepts such as value for money.

> Imagine that you are on a panel providing advice to the Government on which of the following drugs it should subsidise through the Pharmaceutical Benefits Scheme for the treatment of localised prostate cancer. Consider these two options:

An example of a choice set presented to the respondents and the vignette used to situate them in the decision-making context is given in Appendix 9.

Figure 5.1. Example of choice set.

Description	Option A	Option B
*The cost to the Government required to produce an additional **year of full health** compared with usual care is:*	$75,000	$100,000
The benefit to men from treatment is:	The disease does not **progress to their bones** or other parts of the body.	They have **no symptoms** of the disease.
The proportion of men who are likely to benefit from treatment:	**20 in 100**	**30 in 100**
The proportion of men who would experience side effects from treatment:	**60 in 100**	**15 in 100**
The additional cost per year to the national health budget from funding this drug is:	$50,000,000	$10,000,000
The evidence for the treatment is from a:	**Completed study**	**Completed study**
The evidence available from the study reflects:	8 years of experience.	4 years of experience.

Which option would you recommend funding:

Please select ONE.

- ○ Option A
- ○ Option B

The choice task appeared as two profile choice sets using a forced choice. Each choice task was preceded with the following question: Which option would you recommend funding?

It is common for DCEs to include an opt-out option that allows respondents to not choose any of the options in a choice set. An opt-out alternative is usually needed when the research question focuses on predicting the adoption of a new intervention or treatment, as, in reality, respondents might choose the status quo (330). As mentioned above, the objective of this DCE was to understand how potential delays in the availability of data might be traded off against other elements of reimbursement decision-making. Therefore, the opt-out alternative was not considered necessary, and a forced choice approach was adopted.

5.4.2.6. Statistical analysis

Consistent with the random utility theory mentioned in section 5.4.1, the analysis of the data collected in the survey was based on the assumption that when facing multiple alternatives, respondents make rational choices that maximise their utility. By changing the level of attributes characterising the alternatives under evaluation through the design of the survey, the probability of selecting one

alternative can be related to the combination of characteristics providing higher utility. Therefore, respondents' preferences are revealed by analysing their choices.

In the analysis described below, the attributes were coded as categorical with the exception of the delay and cost variables, which were treated in some models as continuous variables to facilitate the estimation of the marginal rate of substitution in monetary terms (willingness to pay per unit of change in the attribute level) and in terms of time (willingness to wait for certainty).

The statistical models were conducted using the results from all respondents. However, as this experiment asked members of the general population to assume the role of reimbursement decision-makers, a subgroup analysis was also conducted in which the answers of respondents who reported having experience in budget allocation decisions were analysed and compared to the results obtained for the overall sample.

5.4.2.6.1. Multinomial logit model

The nature of the experiment usually determines the statistical methods required to analyse the data collected. In DCEs, it is assumed that individuals derive maximum utility from the choices they make when deciding among a bundle of goods. Statistical models try to emulate and predict individual choice behaviour based on the attributes and levels characterising the interventions under analysis. Multinomial logit models (MNL) or conditional logit models are the natural starting point for the analysis of discrete choice data (331) as they are based on the random utility theory proposed by McFadden (287) and can deal with multiple alternatives. They are simple, easy to estimate and interpret and are not computationally demanding (284). As expressed in the RUT (section 5.4.1), in MNL models, the utility obtained by individual choices is a function of the attributes characterising the alternative chosen and a variation associated with unmeasured preferences not explained by the attributes (Equation 5.2):

$$U_{ijk} = V\left(X_{ijk}, B_{ijk}\right) + \varepsilon_{ijk} \quad \text{(Equation 5.2)}$$

where the utility (U_{ijk}) gained by individual i from alternative j in a choice set k, is a function of the attributes of the alternative chosen (X_{ijk}), the coefficient B_{ijk} represents the weighted contribution made by the attribute to the utility gained and (ε_{ijk}) is the error associated with unmeasured variation in preferences and is assumed to be independently distributed.

Despite its popularity and advantages, MNL has been criticised due to the restrictive assumptions it relies on, specifically the independence of irrelevant alternatives (IIA). This restriction assumes that

individual preferences between alternatives is independent of other new alternatives that become available. This can produce biased results as it may be unrealistic to assume that alternatives do not compete asymmetrically with one another (284).

The conditional logit model was conducted in Stata (version 17) using the inbuilt Stata command clogit2. To test the direction and magnitude of the cost and delay coefficients and how these attributes impacted the fit of the model, two conditional logit models were conducted where the cost and delay attributes were introduced in the model as categorical and continuous variables. The selection of the categorical and continuous model was based on the goodness of fit informed by Akaike and Bayesian information criteria (AIC and BIC). Coefficients of the model with the best goodness of fit are also presented to measure the relative importance of attributes.

5.4.2.6.2. Heterogeneity

Differences in preferences among respondents to the survey can create a heterogeneous sample that makes estimation of realistic substitution patterns more difficult and might lead to biased elicitation of marginal utility. Individual characteristics (e.g., education level, health literacy, health status or income) could be associated with behavioural variation and are possible sources of variability in preferences. To account for preference variability within individuals surveyed, additional analyses using latent class and mixed logit models were conducted.

5.4.2.6.2.1. Latent class model

The main limitation found in the conditional logit model (IIA) can be relaxed in latent class models (LCM). LCM are an extension of the logit model but assume that individual behaviour depends on observable attributes and on latent heterogeneity that varies with factors that are unobserved by the analyst. Equation 5.3 describes the equation to estimate the utility derived from an alternative in a choice experiment using an LCM.

$$U_{ijk} = \beta' \chi_{ijk} + \varepsilon_{ijk} \qquad \text{Equation 5.3}$$

where the utility derived from individual i in the alternative jth in a choice set k is given by U_{ijk}, β' is a random vector of unknown coefficients associated with respondents' preferences, χ_i is a vector of attribute levels based on individual choice $i's$. In the LCM, β' varies randomly allowing preference to change across individuals.

Thus, it is assumed that individuals are implicitly grouped into a set of classes in which members share similar tastes or derive the same utility function from an alternative (332). The LCM extends the

conditional logit model to account for heterogeneity of preferences that could be grouped into clusters or classes within the sample of individuals surveyed. Being a semiparametric model, LCM relaxes the assumption of the distribution of individual heterogeneity while retaining probabilistic characteristics to estimate the probability of class membership for each respondent (333). Therefore, the number of classes or the class membership of any particular individual is not an a priori parameter known that could be used in the model.

Different measures of relative fit of the model to the data can be used to assess the optimal number of classes (334). Although there is no consensus about the best method to compare models using different numbers of classes, statistical information criteria (i.e., Akaike, Bayesian and the Conditional Akaike information criteria) and theoretical interpretation should guide the decision (334).

The LCM was estimated in Stata (version 17) using the user-written command lclogit2 (335). To estimate the impact of covariates, an LCM including a set of demographic covariates that could explain heterogeneity in preferences was conducted. The information criterion commonly used in the literature (AIC, BIC and CAIC) was used to estimate the estimate the model with the best goodness of fit according to the number of classes included.

5.4.2.6.2.2. Mixed logit

Mixed logit models are flexible models developed from MNL where the restricted substitution patterns arising from the IIA are relaxed (336). In mixed logit models, the coefficients (the weighted contribution of each attribute to the total utility gained) are assumed to be randomly distributed, thus allowing for variability of preferences among respondents. This is opposed to conditional logit models, where the coefficients are assumed to be fixed and common to all respondents. Equation 5.4 describes the equation used to estimate the utility derived from an alternative in a choice experiment using the mixed logit model.

$$U_{ijs} = \beta_i X'_{ijs} + \left(\eta_i X'_{ijs} + \varepsilon_{ijs}\right) \qquad \text{Equation 5.4}$$

where β_i is a vector of coefficients, X'_{ijs} is a vector of explanatory variables, and η_i is random term with zero mean depending on respondent's specific deviations from the mean. The choice probability is estimated by the mean over a number of simulations that draws from the probability density

function of the coefficients.[33] The results of the mixed logit model are not independent of the number of draws. Generally, a higher number of draws tend to stabilise the estimated values of the coefficients; however, this can become a computationally intensive process (336). The analysis conducted in this study used 1,000 Halton draws.

In mixed logit models, the coefficients are not estimated directly, but through the parameters that define their distribution, thus obtaining the mean and variance of the coefficients (336). Using the standard deviations of the mean of the coefficients it is possible to identify heterogeneity, as attributes with standard deviations that are significantly different from 0 indicate that the preferences of respondents are not homogeneous.

The mixed logit model in this analysis was estimated in Stata (version 17) using the mixlogit command (337). For the willingness to accept analysis, 'cost' and 'delay time' variables were introduced to the model as continuous variables and their coefficients was used as the denominator relative to other coefficients (i.e., ratio of coefficients) and their coefficients were assessed against those of other attributes.

5.4.2.7. Willingness to accept analysis

Within the context of a new intervention for LPC, the inclusion of cost and time attributes in the DCE allows the estimation of monetary and time values of attributes characterising interventions in LPC. When the cost attribute is used to measure the marginal rate of substitution against other attributes, the analysis is called willingness to pay (WTP). In this case, when the delay attribute was used instead, the analysis yielded the willingness to wait (WTW) or the time respondents are willing to wait for a marginal change in the levels of other attributes.

In a willingness to accept analysis, the attribute of interest is usually coded as a continuous variable to allow a straightforward interpretation of the results. However, by coding a variable as continuous in the model, a linear relationship of the effect of the attribute over the utility is assumed. If the utility derived from a choice is more sensitive to respondents facing higher values of cost or time compared to their base level, assuming a linear relation would produce biased results. Therefore, levels of the

[33] Since η_i is not given and can take different distributional forms, the choice probability is integrated over all values of η: $P_i = \int L_i(\eta) f(\eta|\Omega) d\eta$, where $L_i(\eta)$ is the choice probability and $f(\eta|\Omega)$ is the probability density function of η and Ω are the parameters characterising the distributions. This integral is defined in open-form and therefore, cannot be calculated. Simulation methods are used to approximate the integral where the parameters of the distributions are draw from the density function and then averaged over the number of draws to estimate the choice probability.

cost and time attributes were tested for linearity before the willingness to accept analysis was conducted. Visual inspection of the attribute levels for cost and time as well as a Wald test were used to test for linearity.

Willingness to pay (WTP) was estimated as the ratio of the value of the coefficients of the attributes of interest to the negative of the 'cost' attribute. Similarly, willingness to wait (WTW) was estimated as the ratio of the value of the coefficients of attributes of interest to the negative of the 'delay time' attribute. The results of the mixed logit model were used to estimate WTP and WTW and their confidence intervals. The WTP and WTW analysis were estimated in Stata (version 17) using the Stata user written command wtp[34] (338).

5.5 Results

The general characteristics of the surveyed population are described in Table 5.5. The most recent demographic data from the Australian Bureau of Statistics, published in 2021, was used to compare the sample with the general population. Respondents in the survey were generally representative of the Australian population. However, there were disparities in some demographic characteristics – specifically, the difference observed in education and the employment status, where the unemployment rate and education level among respondents were higher relative to the Australian population. This could be attributable to age, given that respondents were older than the broader population across all age brackets surveyed, except for those in the 16–19 years bracket.

[34] The STATA command wtp used the logit model and maximum likelihood assuming the coefficients in the model are asymptotically normally distributed. This means that as the sample size increases, the distribution of the estimated coefficients approaches a normal distribution. Assuming that the model is estimated using a large sample, it is reasonable to assume that WTP is given by the ratio of two normally distributed variables. Confidence intervals were estimated using the bootstrap method. As the bootstrap method generates a simulated distribution for the coefficients by drawing a large number of samples, it makes no assumptions about the distribution of the coefficients in the model. Therefore, it does not rely on the assumption that WTP is symmetrically distributed, nor does it require that the coefficients themselves are joint normally distributed. The bootstrap method may perform better than the other methods (such as the Fieller and Krinsky Robb methods), but it is the most computationally demanding method as it requires re-estimating the model for each bootstrap sample.

Table 5.5. Demographic characteristics of the surveyed population.

State	Freq.	Percent	ABS %
NSW	322	32.1	31.82
VIC	258	25.72	25.83
QLD	201	20.04	20.29
SA	72	7.18	6.89
WA	103	10.27	10.42
ACT	16	1.6	1.68
TAS	21	2.09	2.10
NT	10	1	0.96
Gender	Freq.	Percent	ABS %
Male	500	49.85	49.55
Female	502	50.05	50.45
Non-Binary	1	0.1	NR
Family size	Freq.	Percent	ABS %.
0	2	0.2	NA
1	175	17.45	25.56
2	376	37.49	33.52
3	173	17.25	15.89
4	169	16.85	15.43
5	72	7.18	6.34
6	24	2.39	3.26
7	5	0.5	NA
8	5	0.5	NA
9	1	0.1	NA
11	1	0.1	NA
Education	Freq.	Percent	ABS %
Finished school before Year 10	32	3.19	6.6
Year 10 or equivalent	79	7.88	10
Year 12 or equivalent	167	16.65	18
Trade/apprenticeship	267	26.62	27.3
Undergraduate university degree	312	31.11	23.2
Postgraduate university degree	146	14.56	8.1

Employ status	Freq.	Percent	ABS %
Full time	397	39.58	44.0
Part time	138	13.76	19.5
Casual	82	8.18	NA
Home duties	143	14.26	33.0*
Studying	36	3.59	NA
Unemployed	203	20.24	3.5
Parental leave	4	0.4	NA
Born in Australia	Freq.	Percent	ABS %
No	314	31.31	29.8
Yes	689	68.69	70.2
Age	Freq.	Percent	ABS %
16-19	22	2.19	5.76
20-24	95	9.47	6.33
25-29	88	8.77	7.18
30-34	97	9.67	7.45
35-39	102	10.17	7.24
40-44	74	7.38	6.41
45-49	96	9.57	6.39
50-54	73	7.28	6.24
55-59	81	8.08	5.98
60-64	72	7.18	5.67
65-69	71	7.08	4.97
70-74	45	4.49	4.42
75-99	87	8.67	7.43
Relationship status	Freq.	Percent	ABS %
Married	491	48.95	48.7
Long-term relationship	160	15.95	NA
Separated	13	1.3	NA
Divorced	58	5.78	11.4
Widowed	30	2.99	5.5
Single	251	25.02	34.3

Abbreviations: ABS= Australian bureau of statistics, Freq= Frequency, NA= Not applicable.
* The 'home duties' and 'studying' status were assumed to be included in the category 'not in labour force' reported by the ABS.

Over 23% of respondents had previously been diagnosed with prostate cancer and 77% had a close friend or relative who had been diagnosed with cancer in the past. In general, all respondents rated their health status as good, with over 80% rating it as good to excellent. The majority never or rarely needed help with health literacy and were confident with medical forms (Table 5.6). When asked about familiarity with a decision-making role, approximately 38% of respondents had previous experience making decisions about allocating budget, funds or resources to competing uses.

Table 5.6. Other characteristics of respondents.

Decision-maker	Freq.	Percent
Yes	380	37.89
No	623	62.11
Health literacy	**Freq.**	**Percent**
Never	567	56.53
Rarely	213	21.24
Sometimes	140	13.96
Often	64	6.38
Always	19	1.89
Health status	**Freq.**	**Percent**
Poor	29	2.89
Fair	139	13.86
Good	418	41.67
Very good	339	33.8
Excellent	78	7.78

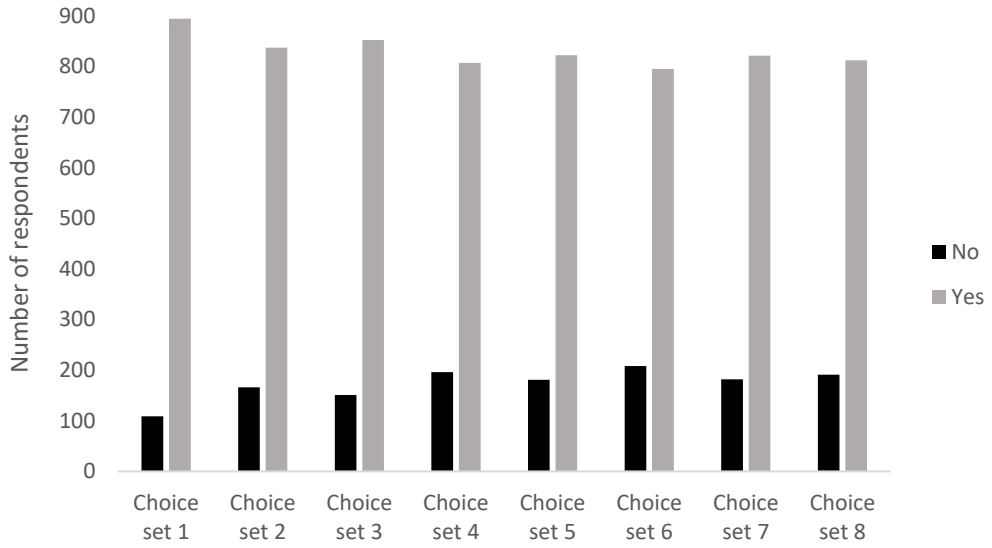
Confident re forms	Freq.	Percent
Not at all confident	16	1.6
A little bit confident	42	4.19
Somewhat confident	180	17.95
Quite a bit confident	313	31.21
Extremely confident	452	45.06
Close Ca diagnosed	**Freq.**	**Percent**
No	232	23.13
Yes	771	76.87
PC diagnosed	**Freq.**	**Percent**
No	772	76.97
Yes	231	23.03

Abbreviations: Ca= cancer, PC= prostate cancer, Freq= Frequency.

5.5.1. Follow-up choice question

All respondents (n=1,003) answered the follow-up questions from all choice sets presented. The number of respondents who would recommend/not recommend any of the options given in each choice set is presented in Figure 5.2. The results of the follow-up choice question showed that in general, the levels and attributes presented in the survey were acceptable alternatives within the context of choice task. From the 8 choice sets presented to 1,003 respondents, 82.75% of the recommendations were consistent with the choice and 582 respondents (58.03%) always recommended the choice selected. Only 33 (3.29%) of respondents always chose not to recommend the selected alternative.

Figure 5.2. Number of respondents who would recommend the chosen option or not.



5.5.2. Conditional logit model

Results of the two conditional logit model are presented in Table 5.7. In both models (where the cost and delay attributes were introduced in the model as categorical and continuous variables), respondents preferred interventions to treat LPC where most patients gained survival benefits as observed in the general population and experienced a low rate of adverse events. Respondents also preferred interventions that were low cost to the health system and represented by lower values of the cost-effectiveness ratio compared to usual care (i.e., represented better value for money). Evidence to support the decision-making process was preferred to be sourced from complete studies with a long-term period of follow-up as opposed to ongoing studies with a shorter period of follow-up.

The attribute with the highest relative importance was the proportion of men that would benefit from the treatment, with a strong preference for a higher proportion of patients who could benefit from treatment. When cost was coded as a categorical variable, it was the second most important attribute in model 1 with a strong preference for lower treatment costs (Figure 5.3).

Table 5.7. Results of the conditional logit model.

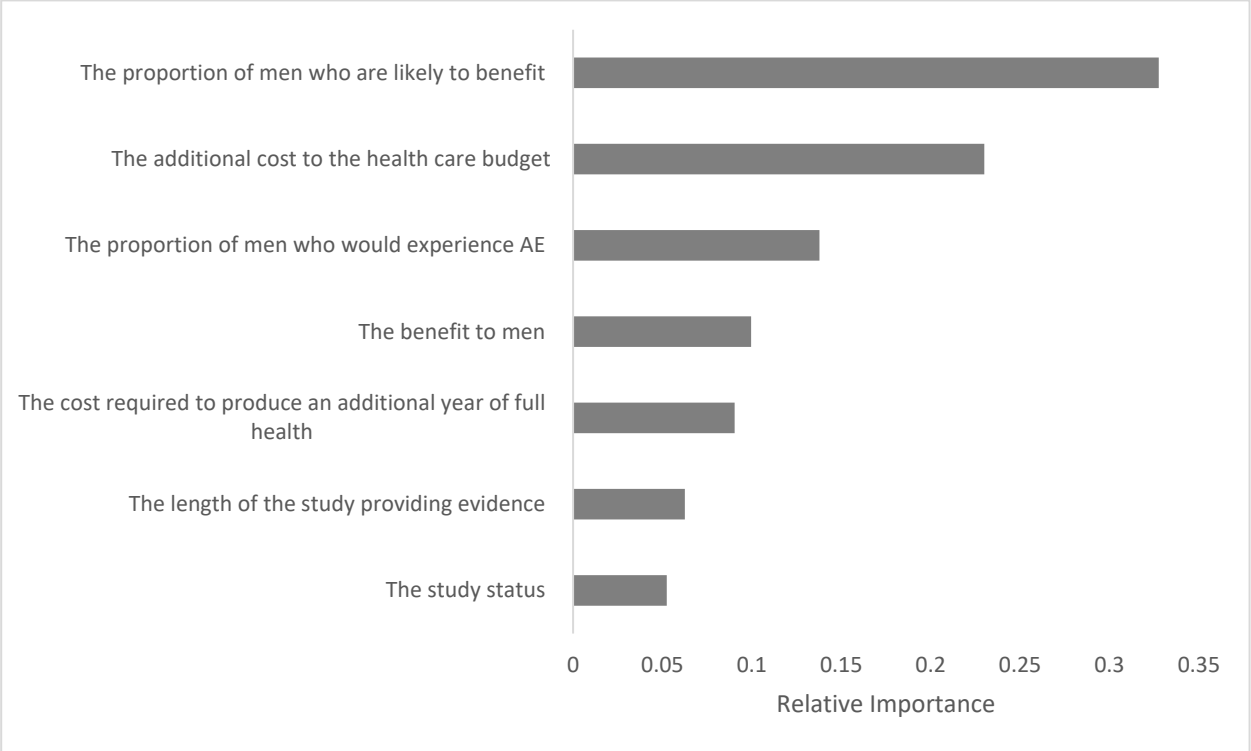
Attributes	Levels	Model 1		Model 2	
		Coefficient	Std. err.	Coefficient	Std. err.
The benefit to men (base: No bone Met)	No circulating Ca	0.109	0.046	0.107*	0.046
	No symptoms	0.134*	0.040	0.132**	0.04
	Normal survival	0.234**	0.047	0.233***	0.047
The proportion of men who are likely to benefit (base: 10%)	20%	0.284**	0.039	0.285***	0.039
	30%	0.566**	0.052	0.567***	0.052
	40%	0.743**	0.051	0.744***	0.051
The proportion of men who would experience AE (base: 15%)	30%	−0.060	0.043	−0.060	0.043
	45%	−0.223**	0.047	−0.224***	0.047
	60%	−0.349**	0.045	−.351***	0.045
The cost required to produce an additional year of full health (base: $25,000)	$50,000	−0.096*	0.040	−0.094*	0.039
	$75,000	−0.129*	0.051	−0.129*	0.051
	$100,000	−0.210**	0.043	−0.209***	0.043
The additional cost per year to the national health budget (base: $10,000)	$50,000	−0.261**	0.041		-
	$100,000	−0.382**	0.039		-
	$150,000	−0.499**	0.045		-
The additional cost per year to the national health budget (continuous)	Continuous	-	-	−0.003***	0
The length of the study providing evidence (base: 4 years)	6 years	0.005	0.044		-
	8 years	0.083*	0.040		-
	10 years	0.179**	0.048		-
The length of the study providing evidence (continuous)	Continuous	-	-	0.031***	0.007
The study status (base: completed)	Ongoing	−0.187**	0.034	−0.187**	0.034
Observations	Model 1			Model 2	
LL	−5225			−5233	
AIC	10489			10496	
BIC	10635			10612	

Abbreviations: Ca= cancer. *** p<0.01, ** p<0.05, * p<0.1.

The goodness of fit changed marginally in model 2 when the cost and delay attributes were included as continuous variables. The pseudo log likelihood was better in model 1 as well as the AIC while the opposite was observed for model 2, where BIC showed a lower value compared to model 1.

All attributes in the model significantly contributed to the utility obtained from treatment for LPC except for 'the proportion of men who would experience AE' at 30%, compared to 15%, and the 'length of the study' at 6 years, compared to 4 years.

Figure 5.3. Relative importance of attributes and levels in the conditional logit model 1.



5.5.3. Heterogeneity

5.5.3.1. Mixed logit model

A mixed logit model was estimated to account for individual preference heterogeneity. The model produced similar levels of attribute importance (Table 5.8) to that obtained by the conditional logit model (Table 5.7). Overall, respondents preferred interventions used to treat LPC which produced benefits in terms of survival at similar levels to that observed in the general population, had low rates of adverse events, were low cost to the health system and were good value for money comparted to usual care. Evidence to support the decision-making process was preferred to be sourced from complete studies with a long term period of follow-up as opposed to ongoing studies with a shorter period of follow (as would be the case if surrogate outcomes were used). Additionally, the standard deviation coefficients from the mixed logit model indicate that there is some heterogeneity in all attributes except for 'the proportion of men who would experience AE', 'the ICER', and 'the study status'. The variability observed for the time attribute implies that respondents' preferences were not homogeneous when considering the time it takes a study to report results. This may indicate the presence of a subgroup among respondents that may prefer to trade off the maturity of the data to expedite the availability of interventions to treat LPC.

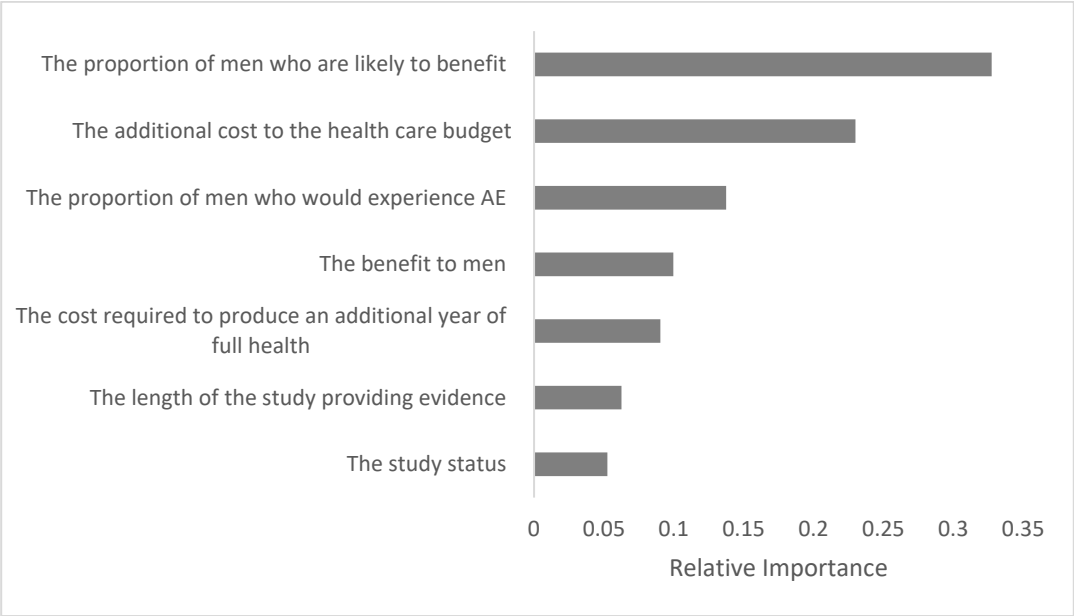
Table 5.8. Results of mixed logit model.

Attributes	Levels	Mixlogit coef	Std.err	Std Dev coef	Std.err
The benefit to men (base: No bone Met)	No circulating Ca	0.118*	0.054	0.493***	0.106
	No symptoms	0.149**	0.046	0.005	0.053
	Normal survival	0.290***	0.058	0.684***	0.085
The proportion of men who are likely to benefit (base: 10%)	20%	0.336***	0.045	0	0.007
	30%	0.670***	0.063	−0.01	0.01
	40%	0.893***	0.064	0.613***	0.09
The proportion of men who would experience AE (base: 15%)	30%	−0.064	0.05	0.251	0.164
	45%	−0.257***	0.055	−0.014	0.017
	60%	−0.413***	0.054	0.297	0.171
The cost required to produce an additional year of full health (base: $25,000)	$50,000	−0.113*	0.047	0.012	0.017
	$75,000	−0.126*	0.061	−0.029	0.023
	$100,000	−0.219***	0.051	0.017	0.037
The additional cost per year to the national health budget (from $10,000-$150,000)	Continuous	−0.004***	0	0.005***	0.001
The length of the study providing evidence (from 4 to 10years)	Continuous	0.039***	0.008	0.117***	0.016
The study status (base: completed)	Ongoing	−0.209***	0.041	0.056	0.147
Observations					
LL			−5171		
AIC			10403		
BIC			10633		

Abbreviations: AIC= Akaike information criterion; BIC= Bayesian information criterion; Ca= Cancer; Coef= Coefficient; LL= Log-likelihood; Std.err= standard error; *** p<0.01, ** p<0.05, * p<0.1.

The mixed logit model and the conditional logit model were observed to exhibit a similar relative importance of attributes (Figure 5.4). The proportion of men who benefit from the intervention and cost were the most important attributes with a strong preference for a higher proportion of men acquiring benefits and lower treatment costs.

Figure 5.4. Relative importance of attributes in the mixed logit model.



5.5.3.2. Latent class model

The LCM evaluates the probability of membership for individuals with different preferences within the sample. All the utility coefficients of attributes coded categorically were allowed to vary from class to class in the model.

The information criterion commonly used in the literature (AIC, BIC and CAIC) estimated for LCM was explored when a range of class membership (2 to 10) was used in the model. For models with 5 classes or more, convergence was not achieved, indicating that these models were not identified empirically (see Table 5.9). When the information criterion was compared, the opposite results were observed for some of the statistics. Specifically, it was observed that the AIC criteria favoured models that were penalised by BIC and CAIC. Given the conflicting results and the objectives of the model (the explanatory nature of the DCE), the choice of class size was based on BIC and CAIC as they emphasise parsimonious models over models with good performance in some predictive sense (339).

Table 5.9. Information criterion for the latent class models using different number of classes.

Classes	AIC	CAIC	BIC
2	10347.06	10583.49	10543.49
3	10291.81	10652.36	10591.36
4	10259.2	10743.88	10661.88
5≥	Models did not converge		

Abbreviations: AIC= Akaike information criterion; BIC= Bayesian information criterion; CAIC= Conditional Akaike information criterion.

Demographic characteristics of respondents regarding their experience with decision-making, health literacy and health status were used to determine class membership. However, the information criterion did not differ from the model without covariates, therefore, was not informative regarding the optimum number of classes.

Based on CAIC/BIC and the theoretical interpretation of results of the model, an LCM with two classes was used for the analysis. From all respondents, 67% (672 respondents) were in class 1 and 33% (331 respondents) in class 2. The demographic characteristics of the two classes are presented in

Table 5.10 by class membership. In general, compared to members of class 1, members of class 2 were older, had higher incomes, higher health literacy and usually did not need help to understand health literature, were more likely to have assumed a decision-making role and less likely to have had a cancer diagnosis or have a family or friend diagnosed with cancer.

The code proposed by Yoo (335) was used to estimate the probability that the model will differentiate between the preferences of respondents among classes. The mean posterior probability of the LCM using two classes is 0.86. Values closer to 1 usually indicate that the model is able to classify respondents into class membership considering the underlying difference in preferences among classes with a cut-off of 0.8 proposed in the literature as an acceptable posterior probability threshold (334).

The results of the LCM using two classes are described in Table 5.11. Compared to class 1 where respondents preferred the benefits of the intervention to be translated into normal survival, respondents in class 2 preferred to have no symptoms of the disease and had a higher aversion for AE and higher expenditure. Class membership was impacted by demographic characteristics of respondents. Health status, income, health literacy and close cancer diagnosis were significantly different among members of each class. Higher income and less experience with cancer diagnosis defined respondents in class 2 as a group with a lower tolerance of risk. This was reflected in their preferences regarding how the benefit of the intervention was measured (as no symptoms of the diseases) and a higher aversion for adverse events and higher expenditure. In contrast, members of class 1 were more tolerant of risk and indifferent to other measures of benefit other than for treatments which produce comparable survival to the overall population.

Table 5.10. Demographic characteristics of respondents by class membership.

	Class 1. (%) (n=672)	Class 2. (%) (n=331)
Gender		
Male	51.34	47
Female	48.51	53
Non-binary	0.15	
Age		
16-19	2.53	1.51
20-24	11.16	6.04
25-29	10.12	6.04
30-34	11.16	6.65
35-39	10.86	8.76
40-44	7.29	7.55
45-49	9.08	10.57
50-54	7.14	7.55
55-59	7.59	9.06
60-64	5.95	9.67
65-69	5.36	10.57
70-74	4.02	5.44
75-99	7.74	10.57
Education		
Finished school before Year 10	2.83	3.93
Year 10 or equivalent	8.18	7.25
Year 12 or equivalent	15.03	19.94
Trade/apprenticeship	27.83	24.17
Undergraduate university degree	31.7	29.91
Postgraduate university degree	14.43	14.8
Cancer diagnosis		
No	74.85	81.27
Yes	25.15	18.73
Need help with health literacy		
Never	50.45	68.88
Rarely	20.83	22.05
Sometimes	16.37	9.06
Often	9.52	-
Always	2.83	-
Class share	33	67

	Class 1. (%) (n=672)	Class 2. (%) (n=331)
Prior experience with budget allocation decisions		
Yes	35	42.9
No	65	57.1
Income		
Negative or zero Income	1.34	1.21
$1 - $9,999 per year	3.42	2.42
$10,000 - $19,999 per year	4.61	6.34
$20,000 - $29,999 per year	7.14	8.16
$30,000 - $39,999 per year	8.78	7.55
$40,000 - $49,999 per year	8.04	7.55
$50,000 - $59,999 per year	7.59	12.08
$60,000 - $79,999 per year	11.01	13.29
$80,000 - $99,999 per year	8.33	10.88
$100,000 - $124,999 per year	9.52	11.48
$125,000 - $149,999 per year	8.78	6.34
$150,000 - $199,999 per year	5.36	6.65
$200,000 or more per year	4.91	5.14
Not Reported	7.44	0.91
Don't know	3.72	0
Confidence with medical forms		
Not at all confident	2.38	-
A little bit confident	4.91	2.72
Somewhat confident	21.88	9.97
Quite a bit confident	31.1	31.42
Extremely confident	39.73	55.89
Close cancer diagnosis		
No	17.26	35.05
Yes	82.74	64.95
Health status		
Poor	2.53	3.63
Fair	11.76	18.13
Good	43.45	38.07
Very good	34.08	33.23
Excellent	8.18	6.95

Table 5.11. Results of LCM with two classes.

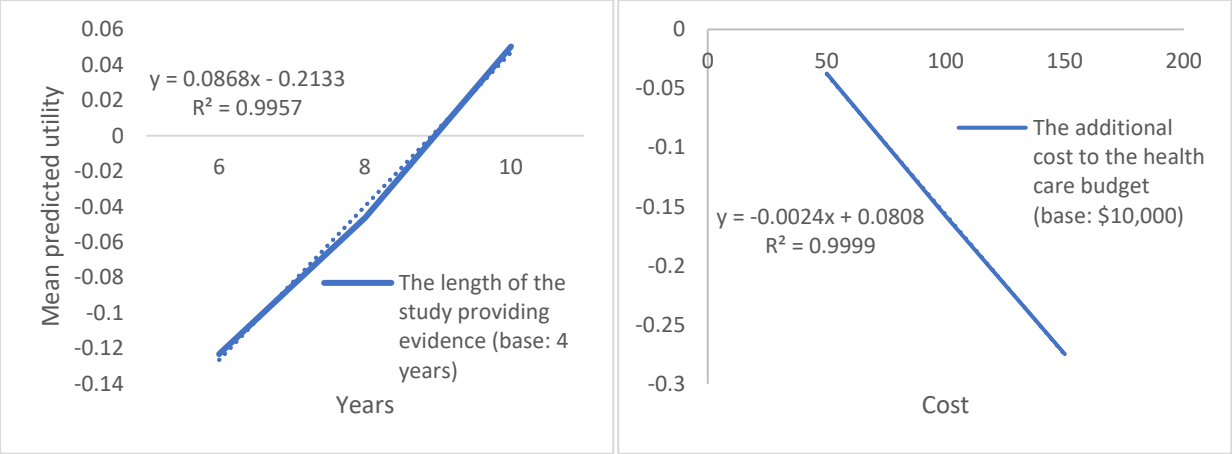
Attributes	Levels	Class 1		Class 2	
		Coefficient	Std. err.	Coefficient	Std. err.
The benefit to men (base: No bone Met)	No circulating Ca	−0.007	0.062	0.395**	0.151
	No symptoms	−0.018	0.056	0.797***	0.198
	Normal survival	0.146*	0.066	0.645***	0.18
The proportion of men who are likely to benefit (base: 10%)	20%	0.085	0.059	1.092***	0.205
	30%	0.224**	0.075	2.122***	0.354
	40%	0.317***	0.077	2.661***	0.398
The proportion of men who would experience AE (base: 15%)	30%	−0.109	0.063	−0.162	0.162
	45%	−0.101	0.069	−0.977***	0.223
	60%	−0.051	0.07	−1.702***	0.279
The cost required to produce an additional year of full health (base: $25,000)	$50,000	−0.113*	0.057	0.008	0.123
	$75,000	−0.024	0.075	−0.744***	0.2
	$100,000	−0.086	0.064	−0.825***	0.175
The additional cost per year the national health budget (base: $10,000)	$50,000	−0.176**	0.059	−0.513**	0.158
	$100,000	−0.212***	0.058	−1.114***	0.201
	$150,000	−0.256***	0.074	−1.457***	0.199
The length of the study providing evidence (base: 4 years)	6 years	0.064	0.06	0.006	0.165
	8 years	−0.004	0.055	0.404*	0.179
	10 years	0.139*	0.065	0.558**	0.189
The study status (base: completed)	Ongoing	−0.193***	0.047	−0.363**	0.131
Class Share:	**Class 1: n=672** **Class 2: n=331**			**Coefficient (Std. err.)**	
Gender (Male)				206 (−0.222)	
CA diagnosis (Yes)				−0.289 (−0.273)	
Close CA diagnosis (No)				0.609** (-0.25)	
Decision-maker (Yes)				−0.386 (−0.239)	
Health status (Good to excellent)				0.657** (−0.281)	
Health literacy (Good)				−3.412 (−3.747)	
Health confidence (Confident)				−1.029*** (−0.323)	
Income (Above average)				-0.497** (−0.22)	
Education (Highest)				0.355 (−0.249)	
Constant				4.077 (−3.968)	
		Observations			
LL		−5106			
AIC		10308			
BIC		10544			
CAIC		10592			

Abbreviations: AIC= Akaike information criterion; BIC= Bayesian information criterion; CAIC= Conditional Akaike information criterion; LL= Log-likelihood; SE= standard error; *** p<0.01, ** p<0.05, * p<0.1.

5.5.3.3. Willingness-to-accept analysis

Visual inspection of the coefficient levels in Figure 5.5 shows a linear relationship between the levels of coefficients of the attributes of time and cost and their mean predicted utility. Moreover, results of the Wald test for linearity (Table 5.12) confirm the linear relationship between levels of the cost and time attributes.

Figure 5.5. Utility estimates when the levels of the coefficients of time and cost are used.



The relative marginal rate of substitution for attributes coded as categorical values was estimated as the ratio of the coefficients obtained in the mixed logit model. The attribute of the 'length of the study' was used to calculate the willingness to wait while the 'total cost to the government' attribute was used to estimate the willingness to pay when both were coded as a continuous variable.

Table 5.12. Wald test of linearty for tue time and cost coefficients.

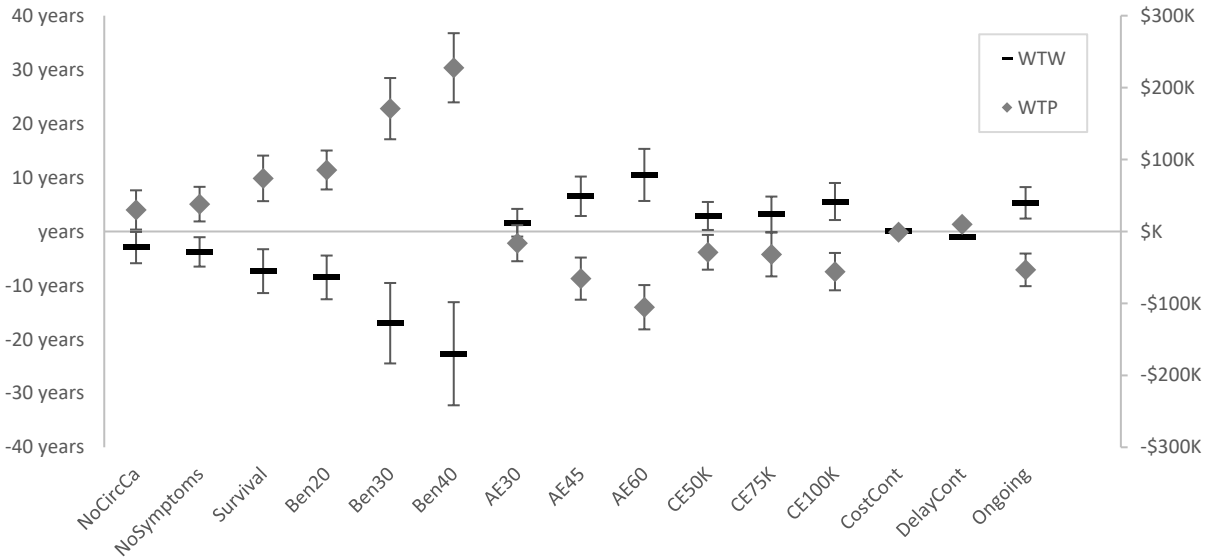
Attribute	Wald test
Cost	7.44 (p<0.006)
Time	19.95 (p<0.000)

Note: The Wald test was estimated using the following linear combinations: [Delay6]-[Delay8] – [Delay10] =0 for the time attribute and [Cost150M] – [Cost100M] – [Cost50M] =0 for the cost attribute.

Respondents were willing to pay $227,500 or wait 23 years for an intervention that benefitted 40% of patients with LPC. The opposite was observed for interventions with a high probability of AEs where respondents need to be compensated $105,300 or with 10 years of time. The preference of respondents for completed studies was also reflected in the willingness to accept analysis. When the evidence to support a new intervention came from an ongoing study, respondents needed

compensation of 5 years or $53,300. Additionally, respondents were willing to pay $10,000 for evidence with an extra year of follow-up and be compensated by 1.2 months for every extra $1,000 of the cost of the intervention. Figure 5.6 presents the willingness to wait and pay for all levels of attributes with the exception of 'cost' and 'delay' that were coded as continuous.

Figure 5.6. Willingness to accept analysis using the results of the mixed logit model.



Abbreviations: WTW= willingness to wait; WTP= willingness to pay. Note: Bars represent the 95% confidence interval. Note: left axis relates to the time for the WTW analysis and right axis to monetary values for the cost attribute.

5.5.4. Subgroup analysis

The analysis of respondents with self-reported experience (n= 380) in budget allocation decisions found similar results to those for the overall population. The results of the mixed logit model for this subpopulation are shown in Table 5.13.

Table 5.13. Results of the mixed logit model for the subgroup of decision-makers population in the survey.

Attributes	Levels	Coefficient	std. err.	SD coefficient	std. err.
The benefit to men (base: No bone Met)	No circulating Ca	0.107	0.086	0.428	0.29
	No symptoms	0.166**	0.074	−0.075	0.215
	Normal survival	0.345***	0.093	0.628***	0.157
The proportion of men who are likely to benefit (base: 10%)	20%	0.258***	0.071	0.023	0.052
	30%	0.622***	0.099	−0.053	0.173
	40%	0.951***	0.098	0.512***	0.149
The proportion of men who would experience AE (base: 15%)	30%	−0.100	0.078	−0.216	0.255
	45%	−0.294***	0.090	0.147	0.13
	60%	−0.475***	0.089	−0.470***	0.159
The cost required to produce an additional year of full health (base: $25,000)	$50,000	−0.141**	0.072	0.026	0.065
	$75,000	−0.251***	0.095	0.003	0.136
	$100,000	−0.325***	0.082	0.201	0.183
The additional cost per year the national health budget (base: $10,000)	$50,000	−0.248***	0.075	−0.035	0.339
	$100,000	−0.386***	0.072	0.206	0.19
	$150,000	−0.510***	0.083	0.489***	0.139
The length of the study providing evidence (base: 4 years)	6 years	0.0847	0.084	−0.028	0.154
	8 years	0.0518	0.074	0.141	0.181
	10 years	0.289***	0.088	0.200	0.165
The study status (base: completed)	Ongoing	−0.332***	0.064	−0.130	0.183

Abbreviations: Ca=cancer, std.err= standard error, SD= standard deviation. *** p<0.01, ** p<0.05, * p<0.1.

The direction and magnitude of the attributes and levels in the decision-maker subgroup are aligned with those observed in the overall population. Decision-makers preferred most patients to benefit from an intervention with low rates of adverse events and lower cost compared to usual care, as observed in the general population. However, the preferences of decision-makers were generally less affected by the cost attribute, and they were less likely to prefer interventions with high rates of AEs.

5.6 Discussion

Through the use of a DCE, the study elicited the preference of general population when taking the role of advisors of reimbursement decision-makers about the use of surrogate outcomes supporting evidence to inform the decision-making process. Regardless of the model of choice used to analyse the survey data, the direction of the coefficients was consistent and as expected. In general, the results of this analysis revealed a strong preference for benefit attributes and a dislike for high cost

and rates of adverse events. This is consistent with the literature indicating that attributes related to cost and benefits gained have a high relative importance to decision-makers (66, 320, 340).

An important finding of the study was the relative importance of certainty over waiting time. The willingness to accept analysis showed respondents were willing to pay $10,000 for evidence with an extra year of follow-up, implying a preference for studies supported by more mature outcomes. The results of the WTP analysis can be interpreted as an aversion to uncertainty. The benefit to society analysis reported in Chapter 4 showed that the use of MFS as surrogate of OS can expedite the reimbursement of new interventions for LPC by 3 years compared to waiting for the certainty of OS to take a decision. For the respondents in this survey, the uncertainty associated with a reimbursement decision supported by the use of shorter or ongoing trials could be offset by interventions with better ICERs.

The relative importance of the benefits of treatment in terms of survival and aversion to adverse events is comparable to other studies conducted in LPC (341).

5.6.1. Surrogate outcomes to support decisions

The DCE found that evidence to support the reimbursement decision-making process was preferred to be sourced from complete studies with a long-term period of follow-up as opposed to ongoing studies with a shorter period of follow-up (as would be the case if surrogate outcomes were used). This was despite the fact that longer follow-up periods could potentially delay the reimbursement decision.

In LPC, the follow-up required to observe meaningful changes in OS is particularly long, leading to a search for alternatives to expedite the treatment of patients with new and potentially better interventions (132). The need for more effective interventions earlier is driving clinical trial design to promote the role of surrogate outcomes, relegating the final outcomes to a secondary outcome reported in continuing studies during the post-market setting. (342) (343). Several regulatory agencies, such as the TGA and FDA, have implemented multiple expedited drug-review pathways that support intermediate and surrogate outcomes to demonstrate clinical benefit. However, respondents to this survey preferred interventions supported by completed studies.

Using surrogate outcomes to support decision-making is part of our daily life, and as a society we use them in a wide variety of fields. As an example, good academic performance is assumed to lead to a successful professional career, and reduction of fossil fuel consumption is expected to save our planet from climatological collapse. As individual members of a society and society as whole, we are used to

replacing the final outcome with a surrogate when there is no other feasible alternative that allows for a timely decision. However, the choice set used in this survey involved life-and-death decisions and therefore is unlike daily decision-making situations. The concept of surrogate outcomes in healthcare and their analytical considerations in clinical trials can be daunting, encompassing several elements such as the validation of such outcomes to ascertain the reliability of the estimated results. Communicating these nuances to the lay population can be challenging. Therefore, the attributes presented in the DCE did not explicitly refer to the concept of surrogate outcomes in healthcare but, rather, focused on the length and status of studies providing clinical evidence, to simulate the use of surrogate outcomes in the reimbursement decision-making process.

Importantly, the heterogeneity found among respondents in the LCM and mixed logit models can help to identify variations in preferences. The mixed logit model found significant heterogeneity among respondents' preferences regarding the length of the studies providing evidence and the LCM identified 2 subgroups with distinct preference profiles. These models' outcomes suggest that a subset of individuals may prefer to make decisions earlier and, as such, may be more inclined to endorse reimbursement decisions supported by evidence derived from surrogate outcomes. For this subgroup, risk-sharing agreements could compensate for the uncertainty associated with the use of surrogate outcomes.

5.6.2. Limitations and future research

5.6.2.1. Population

Resource allocation in healthcare is a complex task that should involve rational thinking to 'maximise the overall health and wellbeing of a society' (344), a difficult task that we, as society, usually rely on a committee of experts to undertake. This analysis focused on the preferences of the general population when asked to assume the role of reimbursement decision-makers for a new intervention in the LPC setting when certainty and delay attributes were used to characterise a treatment alternative. Although members of the general public are not used to facing budgetary decisions, or expected to be familiar with value-for-money concepts on the scale of a government, approximately 38% of respondents reported having previous experience with making decisions regarding allocation of budgets, funds or resources to competing uses. The results presented in section 5.5.4 showed the same direction and comparable magnitude of coefficients when the overall population was compared with the subgroup of respondents involved in budgetary decisions.

Despite the difficulties mentioned above, other studies have successfully elicited preferences when respondents adopt the role of decision-makers in healthcare settings. For example, Hampson et al. developed a DCE to elicit preferences in healthcare with members of general public taking the role of decision-makers responsible for funding decisions (345). Hampson et al. included similar attributes to the ones used in this analysis but decided to omit cost-related attributes to reduce the burden of the survey, as respondents were not typically exposed to healthcare costs. However, cost-related attributes have been used before in this population and other authors have found that the preferences of the general public differed from those of decision-makers. Whitty et al., in an analysis comparing the preferences of Australian reimbursement decision-makers with those of the general public, found that members of society are less sensitive to cost attributes than decision-makers in the context of decisions regarding pharmaceutical subsidies (346). The relative importance of attributes in the conditional and mixed logit models showed that the cost attribute ranked as the second most important, which is comparable to the results found by Whitty et al. for the decision-maker subgroup. The results demonstrate that members of society understood and were able to trade-off cost-related attributes for others that characterised interventions for LPC.

The unsuccessful use of non-validated surrogate outcomes in the regulatory setting, and their overuse for regulatory and clinical decisions, have led to scepticism and sparked controversies about the value of surrogate outcomes (55). This has presumably led to negative reactions among decision-makers when they are used as evidence to support new interventions. Future research could clarify whether healthcare reimbursement decision-makers trade off the time when a decision can be made against the uncertainty of an ongoing study with limited follow-up time, as observed in respondents to this study.

5.6.2.2. Eliciting preferences

The systematic review conducted by Trapero-Bertran et al. did not identify any studies where cost and value for money were used as attributes in a DCE (320). When these attributes are combined in the same survey as in the analysis conducted here, there is a potential for double counting as respondents can perceive them as similar. However, these attributes are well characterised in the survey and have distinct definitions that are generally included in HTA and reimbursement decision discussions.

Other studies focusing on the preferences of decision-makers have found these attributes are of great importance for policy makers as they are an essential part of HTA (320, 340). Moreover, the follow-up questions in the pilot study revealed that respondents had a good understanding of the concept of

value for money. When respondents were asked how they interpreted the factor related to the cost required to produce an additional year of full health, 63% reported that they understood that the cost would accrue to the government. This shows a clear understanding of the definition of cost-effectiveness and, therefore, the difference between the cost-effectiveness and cost attributes.

5.6.2.3. Surrogate outcome as an attribute

In a DCE, attributes are selected to represent different characteristics of an intervention that respondents may consider when making decisions. Although surrogate outcomes may be included as one of these attributes, misinterpretation or misunderstanding of the attribute's meaning may result in biased responses, as the general population may not have the healthcare knowledge necessary to understand the concept of a surrogate outcome. A previous survey aimed to understand agencies' preferences regarding the use of surrogates in decision making for new health interventions, referring to surrogate outcomes explicitly (347). The survey was distributed among HTA agencies with potential experience evaluating surrogate outcomes. In contrast, the stated preference experiment conducted for this thesis involved eliciting perspectives from the general population asked to take the role of decision makers, rather than surveying experienced HTA agencies directly. To address this issue, researchers usually consider whether the target population possesses the necessary knowledge to understand complex concepts or explore the use of alternative attributes or additional information to clarify the concept. In the present analysis, trial duration and the status of the studies providing clinical evidence were used as a proxy for the use of surrogate outcomes in informing reimbursement decisions. However, future research may explore better methods to explain this concept to the general population.

5.7 Conclusion

The analysis of respondents sampled from the public who were asked to adopt the role of decision-makers' advisors showed that they preferred an intervention which provided maximum benefits to society while minimising the financial impact on the health system. Respondents preferred the use of studies with longer follow-up time to support reimbursement decisions regarding new interventions despite the delay in time to treatment. However, the uncertainty associated with the use of surrogate outcomes to support reimbursement decisions can be compensated for by interventions with better ICER.

Chapter 6 Discussion

6.1 Abstract

This chapter discusses the conclusions drawn from the research described in the previous chapters and the policy implications of the results. The overall aim of the thesis was to investigate the use of surrogate outcomes to support decisions about reimbursing new interventions for LPC. The research conducted for the thesis investigated the surrogacy relationship between MFS and OS (Chapter 3) using an economic model that was built to estimate the cost-effectiveness of new interventions for LPC (Chapter 4).

Chapter 4 explored the use of a surrogate outcome within a partitioned survival model and the potential effects of incorporating post-progression survival probabilities obtained from the literature into state transition models. The results showed that MFS as a surrogate outcome for OS could be used in an economic model to compare abiraterone plus prednisolone (AAP) with standard of care (SoC) in patients diagnosed with LPC. The results of the analysis of the expected value of perfect information (EVPI) shows that no further clinical trials were needed to overcome the cost of uncertainty arising from the use of surrogate outcomes to support reimbursement decisions.

However, basing reimbursement decisions on surrogate outcomes rather than final outcomes can lead to uncertainty in the decision-making process. An analysis of preferences among members of the general population who assumed the role of advisors to reimbursement decision makers showed that respondents preferred more certain outcomes such as OS even though this meant waiting longer for the results of trials (Chapter 5), thus delaying treatment for LPC. However, they were also willing to accept interventions with better cost-effectiveness ratios (ICER). In turn, the remaining uncertainty could be shared with the manufacturer through the implementation of risk-sharing agreements including a requirement to conduct further research, conditioning the price on the confirmation of the benefits of the surrogate over the final outcome.

These types of agreements have the potential to promote collaboration between regulatory and reimbursement agencies, which may lead to pricing arrangements tied to the production of additional evidence, not just to verify the surrogacy relationship used in economic evaluations, but also to confirm the comparative safety profile of the new intervention needed for regulatory approvals. However, the ethical implications associated with conducting longer-term studies when the surrogacy relationship consistently predicts the treatment effect of new interventions needs further research.

While this thesis focused on the application of surrogacy in the case of LPC, it is important to note that further research is necessary to determine the utility of surrogacy in other diseases.

6.2 The decision-making process

The evidence supporting the clinical claims of new interventions has increasingly become more challenging to assess. Unmet medical needs are often complex and difficult to address, and the magnitude of benefit and clinical significance from new interventions is usually smaller (348). Additionally, the rapid pace of advancements in disease management has resulted in new treatments being developed and approved for use. While this occurs, the appropriate comparators are chosen based on the best available evidence at the time. However, by the time new treatments are evaluated, the comparators used in the clinical trials may no longer be appropriate. This can pose a challenge for reimbursement decision-makers to determine the true value of the new treatments, as they often need to rely on indirect comparisons that could lead to potential inaccuracies. In this context, the design of clinical trials to support reimbursement decisions has become increasingly important. The inclusion of surrogate outcomes measures in clinical trials may play an important role in supporting reimbursement decisions. As decision-makers rely on available data to support decisions, the use of validated surrogate outcomes to inform economic evaluations can facilitate an earlier decision-making process.

This thesis paves the way for discussions regarding the use of surrogate outcomes in reimbursement decisions and the methods that can be used to test surrogacy relationships. The long survival times characterising LPC was one of the motivating factors for ICECaP to find a suitable outcome to replace OS in clinical trials in this population. The advances in modern medicine would gradually expand survival outcomes across an increasing number and range of health disorders. Further, the onset of serious comorbidities and deaths of patients enrolled in clinical trials across multiple health conditions would increase the follow-up time and cost of conducting clinical trials focussing on OS, which, in turn, would delay access to interventions and increase the cost of effective interventions which meet patients' needs. Implementing surrogate outcomes in clinical trials can enable the use of smaller sample sizes, as the surrogate endpoint may be a more sensitive measure of treatment efficacy. A reduction in sample size can also lead to cost savings in terms of recruitment, monitoring, and data management.

Surrogate outcomes are a ubiquitous component of our daily lives and play an important role in guiding our decision-making as individuals and as a society. For example, test scores are used as

surrogate outcomes to measure student learning and school performance and the use of renewable energy is expected to decrease carbon dioxide emissions and slow down global warming. As individuals and within society as whole, we are used to replacing the final outcome with a surrogate as a means of expediting decision-making. To achieve an efficient decision-making process, the level of uncertainty inherent in using surrogate outcomes in place of final outcomes must be balanced against the opportunity cost. In the present case, the uncertainty resulting from the replacement of OS with MFS should be balanced against the possible benefits arising when a reimbursement decision for new interventions in LPC is taken earlier (e.g., fewer patients progressing in the short term which may be associated with lower costs due to reduced need for further lines of treatments and better quality of life in patients who remain free from BCR).

6.3 MFS as a surrogate of OS

6.3.1. The ICECaP data and validation

The analyses presented in this thesis utilised a surrogacy relationship between MFS and OS developed by the ICECaP working group (Chapter 3). For the purposes of the current research, the surrogacy relationship (and further survival analysis for the construction of the comparator arm in the model) was revised using a subset of the ICECaP data that included metastasis and recurrence after primary treatment and was deemed relevant for health economic purposes. The resulting surrogacy relationship was assessed against the PBAC framework, which supported the use of MFS as a surrogate outcome in economic evaluations for reimbursement decisions (Chapter 3).

One of the key issues found in the literature review reported in Chapter 3 was the absence of evidence addressing the intrinsic uncertainty of the use of a surrogate measure in a way that is informative for reimbursement decision-makers. As the use of surrogate outcomes in an economic evaluation may translate to a higher degree of uncertainty in the results, rigorous and specific uncertainty analyses that address the relationship are needed to inform and facilitate the decision-making process. In this thesis, the uncertainty relating to the surrogacy relationship was assessed by multiple sensitivity analyses. First, the results of the model were assessed when the range incorporating the uncertainty on the HR of the treatment effect of the intervention on MFS was used. Second, the surrogate threshold effect was used to test the results of the model when the surrogacy predicted nonsignificant treatment effects on OS. Lastly, a time dependency surrogacy was assumed by introducing a waning effect after the truncated analysis in ICECaP.

The sensitivity analyses performed reflected the impact that the predicted final outcome had on the results of the economic evaluation and tested the correlation between the surrogate and final outcome, as well as assessing the uncertainty around variations in treatment effect. The results of the sensitivity analyses were aligned with the qualitative assessment of MFS conducted according to the PBAC framework and support the use of the surrogacy relationship in ICECaP to inform economic evaluations for the reimbursement decision-making process in LPC.

It is important to mention that the high quality of the ICECaP data and analysis (a large cohort of patients consisting of 12,712 patients from 19 trials for health economics purposes with a median time to follow-up of 10 years and the use of meta-analytical validation approach) allowed for the assessment of long-term outcomes, increasing the statistical power of the analysis and improving the accuracy of the results. Furthermore, the ICECaP results have clinical significance, advocating for the adoption of MFS as a valid surrogate for OS in patients with LPC. This recommendation is relevant not only for its use in economic evaluations (as in this thesis) but also for other analyses in LPC (125).

An important characteristic of ICECaP is that it based the surrogacy relationship analysis on relative or comparative measures of treatment effect. This approach entailed incorporating proportional hazards assumptions into the survival analysis to estimate the intervention's treatment effect over OS (Chapter 3). This approach had several advantages in that it provided a standardised measure of association that can be compared across different datasets and contexts (the hazard ratio [HR]), it was able to indicate the direction of the surrogacy relationship, and it can be easily understood and interpreted by clinicians who are familiar with comparative measures, such as HRs. However, it also has limitations, as relative measures may confound the actual size of the effect or magnitude of the relationship. In practice this translated into additional computational effort in building the economic model (Chapter 4). An alternative approach to establish the surrogacy relationship would be to use an absolute correlation between outcomes. This would provide a direct measure of the strength of the relationship to be used to estimate the value of OS given an absolute value of MFS without the need for complex modelling or assumptions of proportional hazards to build the intervention's survival probabilities. Absolute measures to inform the surrogacy relationship in economic evaluations can play an essential role to support reimbursement decisions in diseases that rely on absolute prediction derived from an intermediate outcome.

6.3.2. Chain surrogacy

In LPC, the primary outcomes of interest usually include OS, MFS and recurrence or disease-free survival (85). The latter two involve prostate-specific antigen (PSA) measures often used to screen for prostate cancer and monitor patients with the disease. From a timing perspective, using PSA measures to inform reimbursement decisions may offer a significant advantage as those measures can be used for early detection of prostate cancer and recurrence before it becomes clinically evident (349). If a surrogacy relationship between PSA and OS in LPC could be established, the benefits associated with reducing the time where reimbursement decisions are taken could potentially exceed the assumed three-year time frame for MFS used in the model in this thesis. Although PSA testing is a widely used and easily measured clinical tool that detects deviations from the normal range early in the course of the disease, multiple attempts to validate its use as a surrogate outcome of OS have failed or produced inconsistent results (350) (159).

An interesting approach which has the potential to overcome this issue is using PSA in combination with a validated surrogate outcome such as MFS in a chain surrogacy. Chain surrogacy refers to a situation where multiple surrogate outcomes are used to predict a clinical outcome. In practical terms, this could mean that a treatment effect on PSA could be used to predict another treatment effect on MFS, which in turn predicts OS. This chain surrogacy could potentially improve the accuracy of PSA as multiple surrogates may capture different aspects of the disease process or treatment response. Additionally, the use of chain surrogacy may offer insight into the underlying disease mechanism and how treatment works at various stages of the disease.

However, the use of chain surrogacy might potentially increase the complexity of the development and evaluation process, requiring more data and advanced modelling techniques to incorporate and test multiple surrogacy relationships, which in turn will increase the burden of data collection and analysis. The use of multiple surrogate outcomes may also lead to inconsistencies (such as variability in treatment response due to difference in the predicted clinical outcome may respond differently to the same treatment and delayed effect of treatment, as different surrogate outcomes may respond more quickly to treatment than others, leading to a delay in the predicted clinical outcome). Despite the lack of published evidence for chain surrogacy in prostate cancer, incorporating chain surrogacy estimates of treatment effects into economic models and the assessment of its effect on uncertainty remains a challenging and interesting area for further research for statisticians and health economists alike.

6.3.3. Suggestions for embedding the use of surrogate outcomes in reimbursement decisions

As described in Chapter 1, the term surrogate outcome has been described in the literature using multiple definitions, often leading to confusion or misuse. The terms intermediate end point, biomarker, biological marker and others have been used in the literature as analogous terms, but they lack two distinctive characteristics of a surrogate outcome: (1) a surrogate outcome is intended to substitute for a final outcome clinically relevant to patients and (2) a surrogate outcome must predict clinical benefits based on scientific evidence.

While intermediate outcomes may provide some indication of the impact of a treatment on a final outcome, the absence of sufficient scientific evidence to accurately predict the treatment effect on the final outcome precludes such intermediate outcomes from being considered as surrogates. Despite the uncertainty in the ability to accurately predict the treatment effect on a final outcome, intermediate outcomes have been used by HTA agencies as the basis for reimbursement decisions (170). The risk of using non-validated intermediate outcomes to inform reimbursement and regulatory decisions can be illustrated by the case of encainide and flecainide. These antiarrhythmics drugs were approved by the US Food and Drug Administration (FDA) using ventricular arrhythmia as a surrogate outcome of death after myocardial infarction based on epidemiological studies (69). The Cardiac Arrhythmia Suppression Trial findings, reported two years after the drugs were introduced to the market, indicated that a surrogacy relationship between ventricular arrhythmia and death after myocardial infarction could not be adequately supported by epidemiological studies due to their failure in assessing treatment effect. Moreover, the use of these anti-arrhythmic drugs was associated with 2.5 times higher mortality in patients with asymptomatic ventricular arrhythmia after myocardial infarction (70).

Strengthening the definition and selection of surrogate outcomes in government and healthcare literature with clarity, transparency, and consistency is key to promoting their rational use in healthcare decisions. Currently, the PBAC provides a framework that defines the requirements for a surrogate outcome that can be used to predict the treatment effect of new interventions in terms of a final outcome for use in economic evaluations to support reimbursement decisions. The framework offers a tool for a qualitative assessment of the uncertainties associated with the translation of quantitative and qualitative effects from surrogate to final outcome; however, there is no clarity on how submissions for new interventions are using the framework to justify the inclusion of a surrogate outcome in an economic evaluation.

The existing PBAC guidelines (in Appendix 3) offer a clear framework for evaluating surrogate outcomes as per the PBAC guidelines, but the public summary documents (PSDs) arising from PBAC decision-making do not consistently report on the use of surrogate outcomes. This is evidences by the results from a review of a recent study which aimed to map the methodological approaches for the validation of surrogate endpoints applied and reported across a sample of international HTA agencies (including PBAC) showed that the PSDs available in Australia often do not report either the metrics of association or any quantified association (170). Furthermore, for over 50% of the 11 interventions that were subject to review, and for which surrogate outcomes were endorsed by other agencies, there was insufficient information provided in PSDs as to the extent that surrogate outcomes had been considered by the PBAC. This lack of consistent reporting represents a challenge in assessing the use of surrogate outcomes for reimbursement decisions in Australia.

The recommendations underlying the assessment and use of surrogate outcomes for economic evaluations are as follows:

- An evaluation of any proposed surrogate endpoint against each of the criteria established in the framework should be available in the PSD for consideration by the PBAC. Importantly, the evaluation should make clear the threshold of treatment effect in the surrogate considered sufficient to translate to significant clinical improvement in terms of the final outcome (Chapter 3). The dissemination of an assessment of the surrogacy association would facilitate subsequent amendments to the PBAC recommendation, especially when recommending the conduct of trials pot-listing to collect evidence to support risk sharing agreements.

- The PBAC framework assessment, being qualitative in nature, lacks the ability to generate scorecards or scaled measures of strength of the surrogate outcome, hence there is a need for a structured assessment of the surrogacy relationship and its suitability for use in economic evaluations, which should be publicly available. This would allow for standardisation of the use of surrogate outcomes in health decision-making.

- Importantly, the assessment report of surrogate outcomes could include suggestions for future actions. Particularly when reimbursement decisions hinge on additional evidence, the report could outline the outcomes and measures necessary to support or revise any conditional decision that might emerge on the basis of using a surrogate outcome.

- Furthermore, the report should stipulate the timeline for revising the decision and articulate a clear course of action that comprises both immediate and long-term

measures in future investigations of the association between surrogate and final outcomes in continuation studies.

These recommendations have the potential to promote harmonisation in the use of surrogate outcomes in economic evaluations across HTA jurisdictions. Harmonisation could lead to greater uniformity in decision-making, which may enhance the predictability and transparency of the decision-making process, improve the comparability of findings, facilitate cross-jurisdictional learning and knowledge-sharing, and promote international collaboration. Moreover, the implementation of consistent methodologies and outcomes across different HTA jurisdictions has the potential to reduce the cost associated with the development and execution of economic evaluations as the same methodologies and outcomes can be used across different HTA jurisdictions.

6.4 Incorporating MFS in an economic model for LPC

6.4.1. The model

In this thesis, MFS as a surrogate outcome of OS was used in an economic model to inform reimbursement decisions regarding new interventions for high-risk LPC patients in Australia. The implementation of this surrogacy relationship into an economic model has the potential to facilitate earlier reimbursement decision-making, which benefits patients by providing access to more effective interventions. When the reimbursement decision is brought forward by 3 years, society also benefits from more patients remaining free of progression for longer under the new treatment, which translates into lower costs to treat recurrence or metastases, higher quality of life and ultimately, better survival outcomes when compared to waiting for information on OS to become available. The anticipated value to society demonstrated in this case study outweighs the costs associated with the additional research warranted to address the uncertainty associated with the surrogacy relationship (Chapter 4).

The economic model developed incorporating MFS as a surrogate outcome of OS in this thesis was a flexible PSM that allowed interventions for LPC to be investigated in different populations based on risk assessment and clinical settings. This approach extended the current literature that primarily used Markov models. Markov models rely on transition probabilities to establish their structural basis which have the potential to affect the surrogacy relationship by incorporating post-progression probabilities. However, for new interventions with MFS as the primary outcome in clinical trials, it is likely that post-

progression transition probabilities will not be available, requiring their extraction from external sources. This reliance on external sources for post-progression transition probabilities introduces a dependence on OS probabilities that, in other cases, would be contingent on observed MFS probabilities and the surrogacy relationship. Moreover, by incorporating an intermediate health state in the modelling process, the OS estimated through the surrogacy relationship is uncoupled from the data used to form the surrogacy relationship.

PSMs overcome this issue by allowing OS to be independent of other outcomes and therefore, the surrogacy relationship among both outcomes is not affected by uncertainties introduced by external sources of post progression probabilities. The model was developed to calculate the possible benefits to society arising from the use of MFS as a surrogate outcome for OS in a decision-making process that can potentially reduce the time to commencement of treatment by three years.

However, decision-making dilemmas arise because errors may result either from deciding too soon or from delaying decisions too long. While this analysis focused on the opportunity cost implied in delaying decisions, deciding too soon based on surrogate outcomes may introduce additional uncertainties that are worth exploring in future research:

- Efficacy: The use of surrogate outcomes may not fully reflect the actual clinical benefits and harms of the intervention. There is a risk that adopting a new intervention based on surrogate outcomes alone may result in unexpected clinical outcomes, either positive or negative.
- Safety: Surrogate outcomes may not reflect the safety profile of the intervention. There may be unknown or unexpected safety risks associated with the intervention that are not captured by the surrogate outcomes.
- Long-term uncertainty: The use of surrogate outcomes may not capture the long-term effects of the intervention, such as late-onset side effects or the development of resistance to the treatment.
- Economic: The adoption of a new intervention based on surrogate outcomes may result in additional healthcare costs that are not justified by the actual clinical benefits. This is particularly relevant when the intervention is expensive or has limited efficacy.
- Patient-centred: Adopting an intervention too soon based on surrogate outcomes may not reflect the patients' perspective or preferences and may not align with the patient's values and priorities.

Additionally, the use of surrogate outcomes in other types of models such as microsimulations, or more complex models such as multistate models in which state membership is based on survival functions similar to PSMs (277) was not explored in this analysis. Further research is needed to assess

potential challenges and evaluate the practicalities of surrogate outcomes implemented in other economic models and importantly, in other diseases.

6.4.2. The role of additional research

Chapter 4 reported the results of an Expected Value of Perfect Information (EVPI) analysis to estimate the cost of uncertainty regarding the use of median progression-free survival (MFS) as a surrogate outcome for overall survival (OS) in the case of abiraterone acetate plus prednisone (AAP) versus standard of care (SoC) for the treatment of metastatic castration-resistant prostate cancer. The results showed that the cost of uncertainty ranged from $7.5 million to almost $12 million and therefore, investing on conducting further studies would not be economically viable for the reimbursement decision-making process.

The analysis implied that using MFS as a surrogate outcome of OS provides sufficient information to make informed reimbursement decisions. However, the use of MFS as a surrogate outcome to estimate OS may be considered as the primary source of uncertainty in making reimbursement decisions, and the EVPI considers the results of the probabilistic sensitivity analysis. Additionally, it is important to recognise the limitations of using MFS as a surrogate outcome for OS. The surrogacy relationship in ICECaP provides information on treatment effect, but it does not provide information regarding the comparative safety profile of the new intervention, which may require additional studies. Regulatory bodies may grant early approvals based on surrogate outcomes, but additional evidence may be needed to confirm both treatment effect and safety.

Regulatory bodies typically grant early approvals for interventions based on surrogate outcomes on condition that the sponsor agrees to conduct follow-up trials to confirm both the treatment effect and safety of the new intervention, and reimbursement decisions for interventions relying on surrogate outcomes may lead to pricing arrangements tied to the production of additional evidence, not just to verify the surrogacy relationship, but also to confirm the comparative safety profile of the new intervention.

The possible loss of efficiency arising from the uncertainty of deciding without knowledge of the final outcome is the trade-off made for the time gained and all the associated benefits in treating patients earlier. However, the risk associated with the uncertainty around the decision could potentially be shared by the decision maker and the manufacturer, leading to the introduction of managed entry arrangements subject to the use of surrogate outcomes and ensure that further data are collected to

assess the surrogacy relationship and safety profile. Pricing arrangements may be tied to the production of such evidence.

6.4.3. Ethical reimbursement decisions

Based on the value of information theory, when the cost of uncertainty is greater than the cost of conducting further clinical trials that help dissipate the decision uncertainty, the reimbursement decision should be made once the data become available. Nevertheless, the use of surrogate outcomes can have implications beyond the pragmatism of supporting healthcare decisions through their use in economic evaluations. As such, whether further clinical trials should be conducted to resolve uncertainty in reimbursement decisions when surrogate outcomes are used to support them should be carefully examined.

The ethical conundrum involved in treating patients with a less effective intervention in a clinical trial can be extrapolated to the reimbursement decision making process. Delaying an effective intervention that improves health based on a surrogate outcome by waiting for the results of OS would inevitably produce worse outcomes for newly diagnosed patients who will receive a less effective intervention prior to information on OS becoming available. Evidence-based justifications are typically required to support decisions in the decision-making process. Decision makers are accountable for the choices they make to the public and therefore, they need to justify if they decide to delay treatment for patients with a more efficacious intervention when the evidence is dependent on surrogate outcomes.

For reimbursement decisions however, this ethical issue can be ameliorated to some extent by focusing on the efficient trial design rather than the conduct of new studies. Platform trial designs offer a means to systematically decrease evidentiary uncertainty for reimbursement decisions compared to discrete single-drug trials by employing a master protocol that allows ongoing evaluation without requiring repetitive trial initiation activities (351). This adaptive, continuously evolving approach aligns well with the concept of basing reimbursement evaluations on mature arms within living evidence platforms, rather than static standalone studies. The flexible, amendable framework of platform trials can build on existing knowledge infrastructure by facilitating successive consideration without repetitive lags. By integrating these efficient trial concepts, decision uncertainty around surrogate endpoints may be incrementally reduced through stepwise refinement within an adaptive model, better informing priority-setting and maximizing the health system benefits of ongoing research investments.

The evidence supporting the surrogate outcome usually derives from phase 3 trials where both outcomes, surrogate and final, are measured. Follow up studies, where patients have been already randomised, can produce the evidence needed by decisions makers to confirm the relationship between surrogate and final outcome previously used to make a conditional decision[35]. Reimbursement decisions conditional on evidence supporting the surrogacy relationship in the population of interest, could be made avoiding the need to conduct new clinical trials thereby ameliorating ethical dilemmas arising from the predictive ability of the surrogate outcome and the principle of equipoise.

6.5 Preference for the use of surrogates in reimbursement decisions

Through the use of a DCE, the preferences of the general public, who were asked to adopt the role of advisors to decision-makers for new interventions for LPC, were elicited. In general, the results of this analysis revealed a strong preference for benefits and a dislike of high costs and increased rates of adverse events. This is consistent with the literature indicating that attributes related to costs incurred and benefits gained have a high relative importance to decision makers (66, 320, 340). Regardless of the model of choice used to analyse the survey data, the directions of the coefficients were consistent and as expected.

An important finding of the study was the relative importance of certainty over waiting time. The willingness to accept analysis showed that respondents were willing to pay $10,000 for evidence with an extra year of follow-up implying a preference for studies supported by more mature outcomes. The results of the WTP analysis can be interpreted as an aversion to uncertainty, as might typically be associated with the use of surrogate outcomes. The benefit to society analysis (Chapter 4) showed that the use of MFS as surrogate for OS can expedite the reimbursement of new interventions for LPC by three years compared to waiting for the certainty of OS to take a decision. For the respondents in

[35] Conditional decisions in the context of reimbursement of healthcare interventions refer to decisions made by healthcare payers, to approve the use of a particular intervention under certain conditions, based on the available evidence for the intervention, such as its efficacy, safety, and cost-effectiveness. These conditions are usually designed to manage uncertainty or reduce the financial impact of the intervention and may be accompanied by requirements for additional data collection, monitoring, or evaluation, which can be used to inform future decisions about the intervention. If the evidence supporting the conditional decision does not comply with the requirements, the healthcare payer may reconsider or revoke the conditional decision, resulting in a change in the coverage or reimbursement status of the intervention.

this survey, the uncertainty regarding a reimbursement decision supported by evidence from shorter follow-up or ongoing trials could be offset by interventions with better (i.e., lower) ICERs.

The literature has reported on the unsuccessful use of non-validated surrogate outcomes in regulatory settings and their overuse informing regulatory and clinical decisions, which has led to scepticism and sparked controversies about the value of surrogate outcomes (55). This, presumably, has fostered a negative opinion among decision makers regarding their use in supporting evidence about new interventions (55). Future research could seek to clarify whether healthcare reimbursement decision makers trade off the time during which a decision can be made against the uncertainty of an ongoing study with limited follow-up time, as was observed in the present study.

There are a number of policy implications arising from the findings of the DCE. The results suggest that the increase in uncertainty associated with using surrogates should be considered and their use compensated for with lower ICERs. If lower ICERs are not forthcoming, decision-makers may prefer to rely on data from longer-term studies to provide more robust evidence for reimbursement decisions. These implications are particularly important for healthcare policymakers who must balance the need to provide timely access to effective treatments with the requirement to ensure that public funds are spent wisely and based on robust evidence. Further research is needed to explore how these findings can be translated into policy recommendations and to determine the most effective ways to communicate these recommendations to decision-makers.

6.6 The need for collaboration

Healthcare decision makers in the regulatory and reimbursement area can benefit from mutual collaboration. The most common practice used to make new healthcare interventions available to patients involves a stepwise process where reimbursement decisions are taken after regulatory approval. However, it is often the case that the decisions of regulatory and reimbursement agencies are disconnected (352).

Clinical trials are an integral part of the drug development process and are primarily aimed at regulatory approval. However, the clinical endpoint used to demonstrate efficacy for regulatory decisions may not always be appropriate as the basis for reimbursement decisions, particularly if there is a need to estimate the relative benefit of new interventions. In other words, clinical trials are designed to establish the safety and efficacy of a drug, and not necessarily to assess its cost-effectiveness or its impact on public health.

The disconnect between decisions based on efficacy and those based on cost-effectiveness is evident when the number of new interventions authorized by regulatory bodies based on intermediate outcomes is compared with the number of new treatments funded by payers, even when final outcomes are used (352). Although both regulators and payers value the clinical benefits, regulatory agencies usually do not consider economic cost of the therapies approved, while payers focus on the cost and value for money factors in making reimbursement decisions. Thus, as discussed in the existing literature (352), there is a need for a more balanced approach that takes into account both the clinical and economic aspects while approving and reimbursing new therapies.

The results reported in Chapter 5 revealed a preference for the use of more comprehensive evidence emanating from clinical trials with prolonged follow-up periods. Nevertheless, the introduction of uncertainty resulting from the use of surrogate outcomes could be mitigated if such outcomes are used to support a clinical claim during regulatory decisions that involve conditional approval processes contingent upon ongoing clinical trials and evaluations of health economic outcomes for reimbursement decisions. This type of process would allow healthcare decision makers, in general, to propose an approval strategy where the revision of evidence based on further research to confirm the surrogacy relationship can impact both the approval decision and price negotiation schemes.

Additionally, as adverse events are often underrepresented in clinical trials (353), a collaborative approach could improve the accuracy of the evaluation of safety, (353) providing decision-makers with a more complete understanding of the treatment's risk-benefit profile, ultimately leading to more informed reimbursement decisions. This type of flexible healthcare decision making structure would facilitate the availability of more efficient interventions for patients in need while reducing the uncertainty for decision makers.

6.7 Plausibility beyond equipoise

Clinical trials raise ethical concerns as participants are given experimental drugs, and their safety and well-being must be prioritised (354). Experimentation in humans is based on the ethical concept of equipoise or the “uncertainty principle”. Applying this concept, studies can enrol patients into randomised trials only if the uncertainty of benefits gained using the interventions under comparison are considerable. In other words, there must be an equal balance of benefits and risks in both treatment groups. Under the strict interpretation of this principle, most industry driven trials can be considered unethical, as only those research projects that are likely to have positive results are sponsored. A harmonised view suggests that an ethical study is one where “the results of a study

cannot be predicted consistently in advance" (355). Thus, when surrogate outcomes reliably predict the final outcome, it may be unnecessary or unethical to continue the trial or conduct new ones to confirm the long-term outcomes.

The predictive ability of potential surrogate outcomes to inform prior probabilities of gains in benefits from new interventions has led to ethical controversies in the conduct of clinical trials. The literature has reported that when the benefits of a new intervention have been proven using potential surrogate outcomes prior to the conduct of phase 3 RCT, patients are likely to think that randomisation to the less effective intervention is not ethical (356).

The use of surrogate outcomes in clinical trials has raised ethical questions about the validity and reliability of these outcomes in assessing the effectiveness of new treatments. One such question is whether it is ethical to conduct further studies based on surrogate outcomes if they are not supported by the expected value of perfect information (EVPI). On the one hand, conducting further studies to validate surrogate outcomes may be necessary to ensure that patients are receiving the best possible treatments. On the other hand, conducting studies that are not supported by the EVPI (considering the EVPI analysis would not be the only determinant since EVPI relies on the estimates of effect) may be a waste of resources that could be better allocated to other areas of healthcare.

Another question is whether it is ethical to not conduct longer-term studies when members of society prefer data from such studies (Chapter 4). Long-term studies are often preferred as they provide more comprehensive information about the effectiveness and safety of treatments. However, such studies can be costly and time-consuming, and delaying the availability of new treatments may be detrimental to patients who need them urgently.

Ultimately, the balance between using surrogate outcomes and conducting longer-term studies lies in ensuring that patients receive the best possible care while minimising harm and resource wastage. It may be necessary to accept that surrogate outcomes are the 'final' outcomes in their own right, as long as they are reliable and valid indicators of treatment effectiveness. However, further research and evaluation is needed to ensure that the use of surrogate outcomes in reimbursement decisions is both ethical and effective in assessing the value of new treatments.

6.8 Concluding remarks

The research in this thesis has focused on examining how surrogate outcomes can be used in economic models to facilitate the reimbursement decisions of new interventions in patients with LPC.

The findings indicate that MFS can serve as a surrogate outcome to estimate the treatment effect over OS with sufficient certainty to support reimbursement decisions that benefit society. When members of the general public acted as advisors to reimbursement decision-makers, they indicated that the remaining uncertainty associated with the use of surrogate outcomes in the reimbursement decision-making process could be offset by the adoption of interventions that offer better value for money. These results suggest that the incorporation of surrogate outcomes in economic models can be valuable in supporting timely and informed reimbursement decisions that promote societal welfare.

From a policy making perspective, the uncertainty caused by basing reimbursement decisions on outcomes other than OS or final outcomes can hamper the use of surrogate outcomes. However, surrogate outcomes can be used to accelerate the regulatory and reimbursement decision making process. Healthcare decision makers should prioritise the needs of patients and be open to methods that allow them to utilise the information at their disposal as effectively as possible. As advancements in interventions continue apace, it will be more crucial than ever to decide on the most appropriate means of including surrogate endpoints in choices in healthcare and particularly in reimbursement decisions.

Appendix 1: Search Strategy for the Systematic Review

The search strategy used is on Table 14.

Table 14.Systematic review strategy

Number	Search strategy	Data base
1	((("Surrogate outcome*") OR ("biomarkers"[MeSH Major Topic]) OR ("Metasta* Free Survival") OR ("diseas* free survival") OR ("Surrogate end-point*") OR ("Surrogate endpoint*") OR ("Surrogate end point*") OR ("Intermediate outcom*") OR ("Intermediate clinical end-point*") OR ("Intermediate clinical end point*")) AND (((prostat*) AND ((neoplas*) OR (cancer*) OR (*carcinom*) OR (tumor) OR (tumour) OR (malignan*) OR (adencarcinoma))) OR (neoplasm, prostatic[MeSH Terms])) AND (((Economic*) OR ("cost analysis") OR ("Economic evaluation") OR ("cost-utility analysis") OR ("cost-minimization analysis")) OR (analysis, cost benefit[MeSH Terms])))	Pubmed
2	#1 ("Surrogate outcome*") (Word variations have been searched) 461 #2 MeSH descriptor: [Biomarkers] explode all trees 18797 #3 ("metastases free survival") 55 #4 ("disease free survival") 11096 #5 ("Surrogate end-point"):ti,ab,kw 84 #6 "Intermediate outcomes" OR "Intermediate outcome" 283 #7 "Intermediate clinical end-point*" or "Intermediate clinical endpoint*" or "Intermediate clinical end point*" 3 #8 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 29601 #9 prostat* 18188 #10 (neoplas*) 67789 #11 (cancer) 132932 #12 carcin* OR adenocarcinom* 38121 #13 (tum*r) or tumor* or tumour* 52141 #14 MeSH descriptor: [Neoplasms] explode all trees 69032 #15 (malignan*):ti,ab,kw 14771 #16 #10 OR #11 OR #12 OR #13 OR #14 OR #15 172640 #17 #9 AND #16 12320 #18 MeSH descriptor: [Prostatic Neoplasms] explode all trees 4851 #19 #17 OR #18 12320 #20 Economic* 23933 #21 "cost analysis" 2440 #22 "Economic evaluation" 3134 #23 ("cost-utility analysis") 1102 #24 cost-minimization analysis OR cost-minimisation analysis OR cost minimisation analysis OR cost-benefit analysis 8766 #25 MeSH descriptor: [Costs and Cost Analysis] explode all trees 9696 #26 #20 OR #21 OR #22 OR #23 OR #24 OR #25 27493 #27 #8 AND #19 AND #26 48	Cochrane

3	('cost benefit' OR 'cost-minimization' OR 'cost-minimisation' OR 'cost-utility' OR 'cost utility analysis' OR 'economic evaluation' OR 'cost analysis' OR economic* OR 'cost effectiveness') AND prostat* AND (adencarcinom* OR malignan* OR tumour OR tumor OR carcinom* OR cance* OR neoplas*)	NHSEDD
4	("Economic evaluation" OR "cost-minimization analysis" OR "cost-minimisation analysis" OR "cost-utility analysis" OR "cost utility analysis"OR "cost analysis" OR "cost effectiveness" OR Economic*) AND ("prostat*) AND (adencarcinom* OR malignan* OR tumour OR tumor OR carcinom* OR cance* OR neoplas*)	Econlit
5	('cost benefit' OR 'cost-minimization' OR 'cost-minimisation' OR 'cost-utility' OR 'cost utility analysis' OR 'economic evaluation' OR 'cost analysis' OR economic* OR 'cost effectiveness') AND prostat* AND (adencarcinom* OR malignan* OR tumour OR tumor OR carcinom* OR cance* OR neoplas*) AND ('intermediate clinical end-point*' OR 'intermediate clinical endpoint*' OR 'intermediate clinical end point*' OR 'intermediate outcom*' OR 'surrogate end-point*' OR 'surrogate endpoint*' OR 'surrogate end point*' OR 'diseas* free survival' OR 'metasta* free survival' OR 'biomarker*' OR 'surrogate* outcome*')	EMBASE

Appendix 2: Quality assessment of included studies

CHEERS Checklist (357)

Section	Item No	Recommendation	Reported on page No/line No
Title and Abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	1/2
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.	1
Introduction			
Background and objectives	3	Provide an explicit statement of the broader context for the study. Present the study question and its relevance for health policy or practice decisions.	1-2
Methods			
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.	No
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	No
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	No
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	3-4
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	4
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	4
Choice of health outcomes	10	Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed.	4
Measurement of effectiveness	11a	Single study-based estimates: Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data.	NA
	11b	*Synthesis-based estimates*: Describe fully the methods used for identification of included studies and synthesis of clinical effectiveness data.	No
Measurement and valuation of preference based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	No
Estimating resources and costs	13a	*Single study-based economic evaluation*: Describe approaches used to estimate resource use associated with the alternative interventions. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	NA
	13b	*Model-based economic evaluation*: Describe approaches and data sources used to estimate resource use associated with model health states. Describe primary or secondary research methods for valuing each resource item in terms of its unit cost. Describe any adjustments made to approximate to opportunity costs.	No

Currency, price date, and conversion	14	Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. Describe methods for converting costs into a common currency base and the exchange rate.	4
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.	No
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.	No
Analytical methods	17	Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	No
Results			
Study parameters	18	Report the values, ranges, references, and, if used, probability distributions for all parameters. Report reasons or sources for distributions used to represent uncertainty where appropriate. Providing a table to show the input values is strongly recommended.	8
Incremental costs and outcomes	19	For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios.	9-10
Characterizing uncertainty	20a	*Single study-based economic evaluation*: Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact of methodological assumptions (such as discount rate, study perspective).	NA
	20b	*Model-based economic evaluation*: Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions.	9
Characterizing heterogeneity	21	If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information.	NA
Discussion			
Study findings, limitations, generalizability, and current knowledge	22	Summarise key study findings and describe how they support the conclusions reached. Discuss limitations and the generalisability of the findings and how the findings fit with current knowledge.	10
Other			
Source of funding	23	Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support.	11
Conflicts of interest	24	Describe any potential for conflict of interest of study contributors in accordance with journal policy. In the absence of a journal policy, we recommend authors comply with International Committee of Medical Journal Editors recommendations.	No

Quality of Health Economic Studies (QHES) (143)

1. Was the study objective presented in a clear, specific, and measurable manner?

The study objective was presented clear at the beginning of the reading: to evaluate whether the medical benefits of proton therapy are large enough to motivate the higher costs. Score: 7 out of 7.

2. Were the perspective of the analysis (societal, third-party payer, etc.) and reasons for its selection stated?

- Although one can assume a societal perspective was used, is was not specified in the analysis. Score: 0 out of 4

3. Were variable estimates used in the analysis from the best available source (i.e., randomised control trial - best, expert opinion - worst)?

- Variables in the study were obtained from clinical trials and epidemiologic analysis of Sweden population. Score 8 out of 8.

4. If estimates came from a subgroup analysis, were the groups prespecified at the beginning of the study?

- No subgroup analysis was made. Score 1 out of 1.

5. Was uncertainty handled by (1) statistical analysis to address random events, (2) sensitivity analysis to cover a range of assumptions?

- A deterministic sensitivity analysis was performed with selected variables although the method for the variables selected to perform the sensitivity analysis was not specified. The range used in the sensitivity analysis along with the point estimates of the relative risk for adverse events were chosen arbitrarily. Score 9 out of 9.

6. Was incremental analysis performed between alternatives for resources and costs?

- Incremental analysis of resources and health benefits measured in QALYs was performed. Score 6 out of 6.

7. Was the methodology for data abstraction (including the value of health states and other benefits) stated?

- Although findings from the literature were used to populate de model, there is no information explaining the literature review methods, inclusion or exclusion criteria, data abstraction or population on the studies found. Score 0 out of 5.

8. Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3% to 5%) and justification given for the discount rate?

- The time horizon chosen was the lifetime of the patient, which in a chronic condition like prostate cancer accounts for all possible health benefits and cost. Both benefits and cost were discounted at a 3% annual rate. Although it is assumed the 3% discount rate is required by Swedish Health Authorities, no justification was given. Score 7 out of 7.

9. Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?

- Specifically, for prostate cancer, the study fails to clearly describe the methodology used to estimate all resources and cost used. Score 0 out of 8.

10. Were the primary outcome measure(s) for the economic evaluation clearly stated and did they include the major short-term was justification given for the measures/scales used?

- It was clearly stated that QALY was the primary effect outcome of the economic evaluation, nevertheless, although prostate cancer involves treatments like radiotherapy and surgery with major short-term negative outcomes, there was not mention whether these outcomes were considered in the analysis.
- In addition, the treatment effect was measured as cancer recurrence based on PSA free survival, nevertheless, the study fails to provide correctly, all references that were used to support the 20% reduction when proton therapy is used compared to conventional treatment. Score 0 out of 7.

11. Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?

- Health outcome in prostate cancer was overall survival. The authors of the study stated that there was inconclusive evidence on the overall survival effect of radiation therapy at different doses. They assumed a 20% reduction in cancer recurrence for proton therapy, based upon PSA free survival reported in studies by Slater et al., Hanks, and Valicenti et al. [62/64]. They further assumed a direct relation between cancer recurrences and cancer mortality, which means a 20% reduction in mortality from prostate cancer with proton therapy was assumed. Although these assumptions are an evident source of uncertainty, there was no justification for its use nor uncertainty analysis around it.
- The analysis gives no information about the source of utilities of prostate cancer patients. Although they assumed patients with GI and UG events would have a 7% reduction in health utility, there is no method stated for the collection of these utilities. Score: 0 out of 7.

12. Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear, transparent manner?

- A general model structure is found in the report nevertheless, the study does not report all components of the numerator and denominator in a clear and transparent manner. Baseline utility for prostate cancer patients are not found in the report not their source. Cost and QALYs difference are reported in scenario analysis bur the total cost and QALYs of each alternative compared is missing. Score 0 out of 8.

13. Were the choice of economic model, main assumptions, and limitations of the study stated and justified?

- The choice of a Markov model technic for modelling a chronic disease like prostate cancer is logical, however, it was not justified. Key assumptions of the study like the relative risk of adverse events or the difference in radiation dose delivered were not properly justified. Score 0 out of 7

14. Did the author(s) explicitly discuss direction and magnitude of potential biases?

- No, there was no discussion of the magnitude or direction of the potential bias of the study. Score 0 out of 6.

15. Were the conclusions/recommendations of the study justified and based on the study results?

- Yes, the results of the study indicate that proton therapy may be a cost-effective treatment if appropriate risk groups are chosen as targets for proton therapy. Score 8 out of 8.

16. Was there a statement disclosing the source of funding for the study?

- Yes, analysis was part of the Swedish Proton Therapy Centre investigation and financial support was received from Uppsala University, Uppsala County Council and the Swedish Cancer society. Score: 3 out of 3.

Appendix 3. **Detailed assessment of MFS as a surrogate of OS according to the PBAC Guidelines**

To assess the use of the surrogate outcome in the decision-making process, PBAC developed a framework for evaluating proposed surrogate measures and their use in submissions to PBAC (74).

Although this framework covers qualitative and quantitative aspects of the use of a surrogate outcome to inform an economic evaluation, given the complexity involved in the assessment of a surrogate outcome and the lack of guidance to perform a quantitative analysis, the overall assessment is intended to be qualitative rather than quantitative. This framework aims to identify the uncertainties when surrogate outcomes are used in economic evaluations and provide a qualitative conclusion associated with the prediction of treatment effects in terms of the final outcome.

In the following section, the use of MFS to predict a treatment effect in terms of OS for its use in an economic evaluation that compares radical prostatectomy against RT will be assessed following the steps provided in the framework.

Part 1 Definition, selection, and measurement.

1A) Proposed surrogate measure (PS): Metastasis-Free Survival (MFS).

1A.1) A definition of the PS/comparative treatment effect on the PS in (a) for each and every different epidemiological study and randomised trial identified to support the assessment of the transformation of the PS to the TCO and (b) for each and every relevant randomised trial available for the proposed drug. (For example whether it is a continuous or dichotomous variable) [Note: some of the information which could address this need would rely on information for trials of the proposed drug which would be reported in response to requests in Subsection B.5 of the PBAC Guidelines.]

The definition of the outcome is a key element in the correct use of surrogate outcomes to inform an economic evaluation. Consistency in terms of both the outcome measure and its definition across all studies used to assess the relationship between surrogate and final outcome will enable a consistent prediction of the final outcome to inform further economic evaluations.

Metastasis is usually explored through imaging technics in prostate cancer patients with signs or symptoms of failure after initial treatment. Although new methods with improved accuracy have been found to be efficient in identifying metastasis in certain groups of patients, (185),

international clinical guidelines still recommend CT + bone scans for the overall population of patients diagnosed with prostate cancer (108, 186).

All the studies included in this analysis consistently used the recommended imaging technology for diagnosis (see Table 6.15). All expect one study (188) (did not specify the imaging technology, but probably used CT and bone scan) used CT and bone scan to assess patients for possible DM development. From the 12 studies included in the meta-analysis, 10 used bone scan, one did not specify (188) and 1 specifically stated the used of range of different imaging technologies (189).

Table 6.15. Characteristics of ICECaP studies included and Distant Metastasis definition.

Study	Total number of patients	Median Follow-up (years)	Imaging technique for DM definition
Cryo vs RT Ontario trials(192)	64	10.7	CT + Bone Scan
GETUG06(193)	306	5.1	bone scintigraphy and CT[1]
MRCRT01(194)	843	10	Technetium-99m bone scans, CT [2]
EORTC22961(195)	970	6	Bone scanning, chest radiography and CT[2]
EORTC22991(189)	819	7.2	Chest x-ray and CT or MRI of the abdomen and pelvis[1]
Mottet et al (196)	264	8.4	CT and bone scans
GICOR-DART01/05(197)	352	4.5	CT and bone scan
ICORG9701(198)	276	12	CT and bone scan
RTOG9202(188)	1520	19.6	DM was defined as the clinical evidence of distant disease (by any method)
RTOG9408(199)	1979	9.9	Bone scan, CT and chest X-ray
RTOG9413(200)	1270	17.1	Bone scan and chest X-ray
TROG9601(201)	818	11.4	Bone scans, CT or MRI

CT: Computed tomography, DM: Distant metastasis, MRI: Magnetic resonance imaging, Yr: year. Notes:
[1] Bone scanning was performed when PSA >10 ng/mL.
[2] MRI of the abdomen and pelvis was also performed if indicated.

Metastasis is a dichotomous health state, diagnosed using imaging techniques and able to be confirmed by histopathological samples of tissues. Difference in imaging techniques used to detect metastasis will have a potential impact on the relationship between surrogate and final outcome. More accurate and sensitive diagnostic techniques are less prone to produce false negative results and thus, the proportion of patients with metastasis among the whole population scanned may potentially increase.

The treatment approach for patients with micro metastasis and the implications for usual clinical practice are yet to be clarified as more accurate and sensitive imaging techniques are adopted (185). Although advanced diagnostic methods are mentioned in international clinical guidelines, metastasis assessment with widely available and well-used imaging techniques is still recommended (108, 186). The potential impact on the surrogate to final outcome relationship with more accurate imaging techniques will not be discussed in this analysis.

1A.2) A measurement of the PS/comparative treatment effect on the PS in (a) for each and every different epidemiological study and randomised trial identified to support the assessment of the transformation of the PS to the TCO and (b) for each and every relevant randomised trial available for the proposed drug. [For example, the tool used to measure the PS, the timing of measurement both in terms of chronology (e.g., with respect to duration of disease and/or age) and, in randomised trials, with respect to the commencement of therapy (and completion of therapy if applicable) and the units of measurement. If the PS is measured more than once to determine a more precise estimate (e.g., repeating the blood pressure and taking the average of the two readings when reporting the PS), a description of the approach taken is needed.]

Most studies examined patients showing signs of progression, as clinically indicated and recommended in clinical guidelines (108, 186). Table 6.16 summarises the imaging technique, criteria, and frequency of measures that patients underwent to confirm metastasis.

Table 6.16. Imaging technique and timing of DM measure in ICECaP studies

Study	Imaging technique for DM definition	Frequency of imaging
Cryo vs RT Ontario trials(192)	CT + Bone Scan	Bone scan as clinically indicated
GETUG06(193)	bone scintigraphy and CT[1*]	NR
MRCRT01(194)	Technetium-99m bone scans, CT [2]	Bone scans, CT, or MRI were done as clinically indicated
EORTC22961(195)	Bone scanning, chest radiography and CT[2]	Chest X-ray, Technetium bone scan, CT scan of pelvis and abdomen are not mandatory every year, but are required, should there be a clinical and/or biochemical (PSA) suspicion of progression. CT scan or MRI should be performed when the interpretation of bone scan is difficult.
EORTC22991(189)	Chest x-ray and CT or MRI[1]	Imaging studies (Bone Scan,CT-of abdomen +pelvis, MRI or Chest X-ray) will be done in case of suspicion of clinical and/or biochemical progression.
Mottet et al (196)	CT and bone scans	CT and bone scans were systematically performed in case of clinical or biologic progression

Study	Imaging technique for DM definition	Frequency of imaging
GICOR-DART01/05(197)	CT and bone scan	Imaging (abdominal-pelvic CT and bone scan) was repeated in cases in which clinical or biochemical progression was suspected.
ICORG9701(198)	CT and bone scan	The follow-up examinations were monthly for the first year and every 3 months thereafter.
RTOG9202(188)	DM was defined as the clinical evidence of distant disease (by any method)	Bone scan as indicated. A bone scan will be performed on any patient who presents with complaints of bone pain that cannot be attributed to any inter-current disease. Discretionary plain films may be needed to evaluate lesions seen on bone scan to confirm the diagnosis of metastatic disease.
RTOG9408(199)	Bone scan, CT and chest X-ray	Bone scan as indicated. A bone scan will be performed on any patient who presents with complaints of bone pain that cannot be attributed to any inter-current disease. Discretionary plain films may be needed to evaluate lesions seen on bone scan to confirm the diagnosis of metastatic disease.
RTOG9413(200)	Bone scan and chest X-ray	As clinically indicated, e.g., rising PSA or bone pain. A bone scan will be performed on any patient who presents with complaints of bone pain that cannot be attributed to any intercurrent disease. Discretionary plain films may be needed to evaluate lesions seen on bone scan to confirm the diagnosis of metastatic disease.
TROG9601(201)	Bone scans, CT or MRI	Investigations—including biopsy, CT scan, chest radiograph, and isotope bone scan—were mandated if symptoms suggested a need, or if PSA reached 20 μg/L without signs of recurrence.

Abbreviations: CT= Computed tomography, DM= Distant metastasis, MRI= Magnetic resonance imaging, Yr= year. Notes: *Not reported within the study. Information taken from the registries at https://clinicaltrials.gov/ct2/show/record/NCT00010244

[1] Bone scanning was performed when PSA >10 ng/mL.

[2] MRI of the abdomen and pelvis was also performed if indicated.

1A.3) Ranges of values (both baseline values and comparative treatment effect values from baseline) over which the PS/comparative treatment effect on the PS have been observed, and the ranges over which their relationships with the TCO/comparative treatment effect on the TCO have been studied.

MFS as a surrogate outcome for OS estimated by the ICECaP meta-analysis included studies reporting HRs for MFS and overall survival between 0.3 and 1.6 and between 0.5 and 1.4 respectively. This could be seen as a limitation when proportional hazards are assumed using extreme values. For example, assuming a survival probability of 0.95 for MFS and 0.960 for OS for the comparator, would yield abnormal survival probabilities in the intervention arm resulting in an HR-MFS below 0.3. The surrogacy equation predicts a HR of 0.385 for OS when the HR-MFS is

equal to 0.3. Under the proportional hazard assumption, the survival probability of the intervention for MFS was equal to 0.985 ($0.951^{0.3}$) and for OS equal to 0.984 ($0.960^{0.385}$) which are implausible results. Nevertheless, the treatment effect observed in MFS in terms of HR is more commonly observed in clinical trials at values close to 0.7 (132).

1A.4) An assessment of the possibility of a waning treatment effect over time (e.g., tolerance or resistance for anti-infective or oncological drugs) in the timing of the measurement of the comparative treatment effect on the PS.

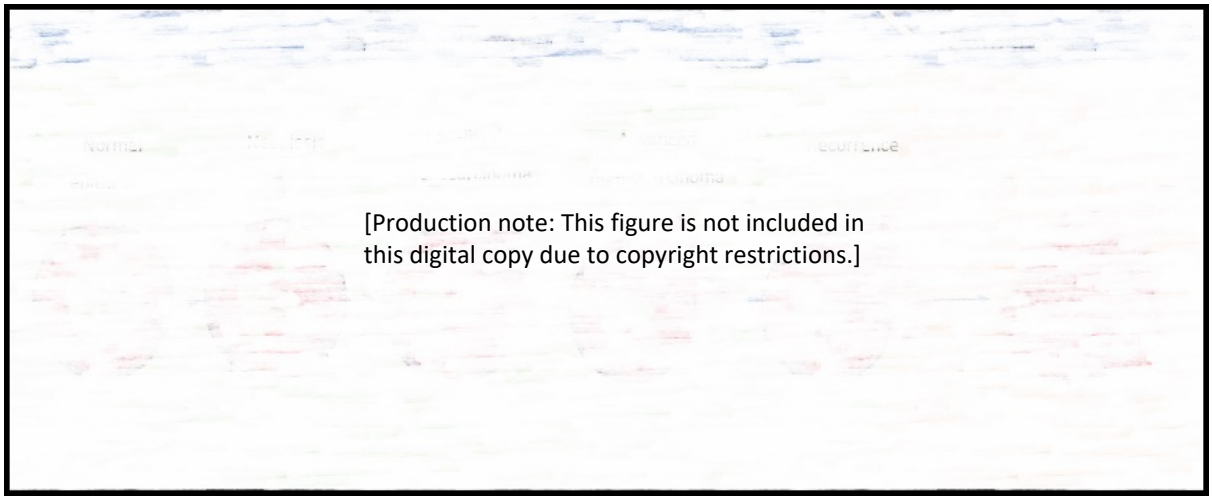
The waning of the treatment effects can be assessed via information regarding the recurrence of prostate cancer. Naïve prostate cancer patients undergoing treatment may eventually find different pathways through which to develop metastasis over time.

The waning effect of the primary treatment for prostate cancer could be understood as the development of recurrent disease. These treatments have been associated with distinctive timing and proportion of patients across time that potentially develop recurrent diseases after initial treatment (358).

The time period during which the measures of recurrent disease were estimated is an important characteristic that can potentially impact the surrogate outcome measure given the natural history of the disease. If the waning effect were to happen after or close to the time the surrogate outcome is measured, it is likely that the surrogate outcome would not be able to capture the waning effect that may be potentially observed after the intervention and therefore, the transformation to FO would be subjective.

Although the exact mechanism of action of recurrent disease is still unknown, it may be explained by different pathways. Prostate cancer involves the proliferation of luminal cells of the prostate that eventually causes a dysplasia along the ducts. The natural history of the disease involves localised prostate adenocarcinoma when the cancer cells invade through the basal lamina which eventually leads to cancer draining into the adjacent lymph nodes and from there, the colonisation of distant organs and metastasis occurs. Recurrent disease in prostate cancer is a step that lags behind metastasis in the disease pathway (Figure 6.1). (191)

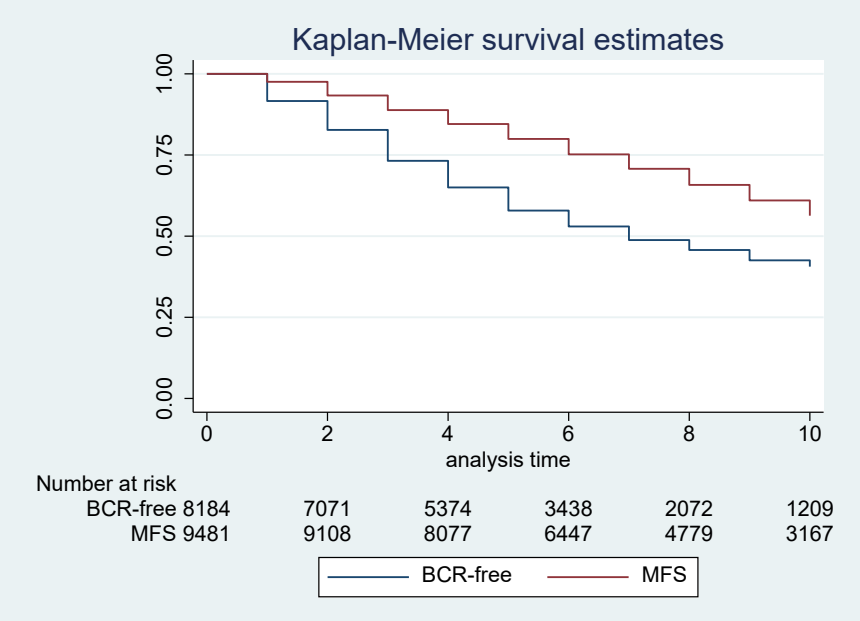
Figure 6.1. Natural history of prostate cancer.



Source: adapted from Wang G, Zhao D, Spring DJ, DePinho RA. Genetics and biology of prostate cancer. Genes Dev. 2018;32(17-18):1105-1140.

An analysis of BCR-free survival and MFS using ICECaP data for the overall population shows that within the follow-up time needed to capture the surrogate outcome MFS, patients with BCR are consistently captured. In other words, recurrence is an event with a faster rate of occurrence than metastasis and therefore, the possible waning effect from the treatment is captured during the time period in which the surrogate outcome MFS is measured. As seen in Figure 6.2, the median time to failure for MFS is not reached in the follow period whereas for BCR, its reached after 7 years.

Figure 6.2. Survival estimates for MFS and BRR-free on the overall population from ICECaP.



1A.5) An assessment of the comparability of these definitions and methods of measurement across all the available sources of information. In particular, the comparability of the definitions and methods across (a) the evidence identified to support the assessment of the transformation of the PS to the TCO and (b) the evidence identified for the PS in the randomised trials available for the proposed drug which would be used to generate the predictive transformation to the TCO analysis.

AS seen in Table 6.15, all trials defined MFS in the same way. At the time where all trials were conducted, the diagnostics techniques used in clinical practice today were available and to allow MFS to be defined consistently.

1A.6) An assessment of the validity, reliability, and responsiveness of the method for measuring the PS.

As the surrogate outcome to be evaluated is MFS, the methods used to diagnose metastasis will be assessed at this point. Metastasis is defined as the colonisation of distant organs from a primary site and in prostate cancer patients, this usually occurs in the bones, with the majority of patients with advanced disease having some degree of skeletal involvement (359). As progression is usually asymptomatic, guidelines recommend a close follow-up after the initial diagnosis of prostate cancer and treatment or staging high-risk patients using PSA, digital rectal examination, and imaging. Metastasis is only suspected and evaluated when the results of these examinations produce results that are outside of the normal ranges. Due to the world-wide availability of CT and bone scans, they have been the gold standard used to stage and diagnose metastasis in patients with prostate cancer (REF) with a pooled meta-analysis sensitivity and specificity of 79% and 82% respectively (360).

Nevertheless, this technique has limitations as it is only able to detect metastases when the radioisotope is absorbed into the bone surface depending on the blood flow and osteoblastic activity reaction in the presence of metastatic lesions. However, it is not able to distinguish between metastatic lesions and other conditions such as bone fracture or conditions prone to induce osteoblastic activity. (361) Recent improvements in nuclear medicine assisted by more sensitive and specific imaging techniques can more accurately detect micro metastases of prostate cancer in distant organs. The use of Ga-PSMA [36] assisted by PET/CT or MRI has been able to detect metastatic tumours outside the prostate of significantly smaller size than those detected by conventional imaging techniques. (185)

[36] A radioligand specifically designed to detect prostate cancer cells expressing the receptor of the PSMA.

Despite the existence of more specific and sensitive imaging technique to detect metastases in prostate cancer, CT and bone scan are 'the most thoroughly validated tests in clinical trials and in clinical practice' and remain the gold standard (361). The clinical relevance of micro metastatic findings using new techniques is still a topic of debate (281). Also, clinical trials are still using conventional imaging as opposed to the use of more advanced and sensible techniques as to assess and diagnose progression (360).

1B) Target clinical outcome (TCO)

1B.1) A definition of the TCO/comparative treatment effect on the TCO in each and every different epidemiological study and randomised trial available to support the analysis. (For example, whether it is a continuous or dichotomous variable.).

Overall survival, the target clinical outcome, is by definition, patient centred and irreversible. In addition, the follow-up time of the OS in the clinical trials is considered of long duration maximising the use of the surrogate outcome in modelled economic evaluations. OS was measured as the time from enrolment of any causes with a median follow-up of 10 years.

1B.2) A measurement of the TCO/comparative treatment effect on the TCO in each and every different epidemiological study and randomised trial available to support the analysis. [For example, the tool used to measure the TCO, when the TCO is measured both in terms of chronology from the TCO and, in randomised trials, with respect to the commencement of therapy (and completion of therapy if applicable) and the units of measurement.]

Overall survival is a dichotomous outcome with a universal and consistent definition that has been applied in each epidemiological study and randomised trial available to support the analysis.

1B.3) Ranges of values (both baseline values and values of subsequent differences from baseline) over which the TCO/comparative treatment effect on the TCO have been observed, and the ranges over which their relationships with the PS/comparative treatment effect on the PS have been studied.

Figure 6.5 shows the values of the HRs over which MFS was evaluated against OS.

1B.4) An assessment of the possibility of a waning treatment effect over time (e.g., tolerance or resistance for antiinfective or oncological drugs) in the choice of TCO and the timing of measurement of the TCO/comparative treatment effect on the TCO.

As the final outcome is survival, there is no possibility that the waning effect of the treatment is not captured within the timing of the FO measure.

1B.5) An assessment of the comparability of these definitions and methods of measurement across the available sources of information, and for the version of the comparative treatment effect on the TCO generated by the proposed transformation for the proposed drug.

Prostate cancer patients who received treatment for their primary tumour will eventually become resistance to treatment and may die of prostate cancer after recurrent disease is diagnosed. The mechanism of action of this resistance is still unknown but there are different pathways that may explain the emergence of treatment resistant disease (358).

1B.6) An assessment of whether the TCO is 'unequivocally relevant and important to patients':

i) Whether the TCO is patient centred. Whether it reflects an outcome which by its nature is relevant to patients.

Survival is patient centred and arguably the most relevant outcome to patients.

ii) Whether the extent of change in the TCO reflects an improvement that is unequivocally important clinically. For example, the extent of improvement in severity of burden of disease or organ morbidity represented by the TCO.

An extension in OS is the most relevant outcome for patients and clinicians alike.

iii) Whether the TCO is reversible.

Not reversible.

iv) *Whether the TCO is also substituting for a more clinically relevant outcome.*

Overall survival is patient centred by definition and represents the most relevant outcome for patients and clinicians alike.

1B.7) A definition and measurement of the TCO currently used in clinical practice, any difference to its measurement in the studies and trials identified, and the implications of any difference in terms of the underlying biological rationale for the proposed transformation.

Not applicable.

1B.8) An assessment of the appropriateness of the TCO for the intervention.

Overall survival is the gold standard and preferred outcome for clinical trials of new interventions. Measuring the benefits of a new intervention in terms of OS is the ideal scenario to demonstrate clinical efficacy and support the decision-making process.

2A) Biological reasoning

2A.1) An assessment of the biologic reasoning, independent of any drug effects, behind the association between the PS and the TCO. In other words, whether the rational basis for a disease pathway causally linking the PS to the TCO.

Metastasis development is a part of the natural course of prostate cancer. Patients who develop recurrent disease and became resistant to second line therapies often have a poor prognosis and experience metastasis. Metastasis is usually the last stage of disease and therefore it is associated with a poor prognosis, low quality of life and a reduction in life expectancy (190).

Although the mechanism is still unknown there are different pathways that may explain progression and metastasis. Prostate cancer begins with the proliferation of luminal cells confined within the prostate. As these neoplastic cells lack apoptosis ability, they evolve and eventually causes a dysplasia along the ducts forming an invasive tumour. The natural history of the disease includes localised prostate adenocarcinoma when the cancer cells invade through the basal lamina into nearby structures such seminal vesicle. The increased demand for nutrients by the growing tumour leads to new vascular networks around the lesion which eventually helps neoplastic cells to drain into the adjacent lymph nodes and from there, the colonisation of distant organs when metastasis occurs. Metastasis then follows and from this health state, the disease progresses in all colonised organs which, left untreated, eventually leads to death (191).

This can also be highlighted in ICECaP which shows a close relationship between MFS and OS (Figure 3.4). The biological plausibility mentioned before translates into patients developing metastases being at higher risk of death, and therefore remaining free of metastasis is associated with a lower risk of death and hence improved OS.

2A.2) An assessment of the proximity of the PS and the TCO to each other in terms of chronological time.

As previously discussed, patients who die from prostate cancer often do so after metastases are diagnosed. The RCTs included in the ICECaP database recorded the time to metastasis and survival for all patients pooled. The median survival time and median time to metastasis was not captured in the 10 year analysis of the overall population. As the median survival time was not reached, the

median time at risk is provided for comparison. The median time at risk for these patients was 7.3 years when OS was measured whereas for patients developing metastasis it was 7.0 years. Figure 3.4 in Chapter 3, shows the Kaplan-Meier curves for both outcomes of interest.

Longer time at risk when OS was measured compared to metastasis indicates the chronology of the disease where metastasis occurs first. The proximity of both outcomes reduces the risk of confounding.

2A.3) An assessment of the proximity of the PS and the TCO to each other in terms of the pathological mechanism of the disease.

As discussed in point 1A.4, OS and MFS are closely related as death is closely followed by metastasis development. In fact, MFS is a composite outcome which involves survival; therefore, it is expected they behave in similar manner.

2A.4) An assessment of any multiple pathways identified for the disease beyond the disease pathway involving the PS identified for.

The mechanisms by which metastasis leads to death are multiple and englobe systemic organ failure, thus encompassing multiple pathways that could lead to death. No other pathways were identified for prostate cancer after metastasis development which could lead to death different from those discussed previously.

2A.5) *Also whether these pathways operate independently of the disease pathway involving the PS and are likely to affect the TCO.*

The results of the surrogacy relationship from ICECaP published in the literature were different from those used in the model. Studies analysed for the purpose of economic evaluations were exclusively RT trials whereas multiples treatments were used in the trials included in the full dataset. This potential issue however, showed little influence on the results. As seen in Figure 3.1 both estimates are similar, showing that different diseases pathways, mechanism of actions of the interventions analysed involving the surrogate outcome, are not likely to affect the treatment effect seen on the OS

2A.6) *An assessment of other possible impacts on the TCO which are independent of the disease. This will vary with different TCOs, e.g., compare disease-related mortality with all-cause mortality, so this assessment will be affected by the definition of the TCO.*

Prostate cancer has a long disease course. In Australia, the average age at diagnosis is 67 years (92). The late onset of the disease means that patients often have multiple morbidities and coexisting diseases. The multiple risks associated with these older patients together with the long duration of localised prostate cancer results in patients dying of causes other than prostate cancer (362). This situation of competing risk could overestimate the overall probability of prostate cancer survival as patients who die of other causes could be counted as censored and therefore, as patient at risk of dying of prostate cancer (363). However, the ICECaP analysis overcomes this potential issue by relating the MFS to OS rather than prostate cancer specific survival.

A different final outcome would yield a different result if it were used as a final outcome to be substitute for by a surrogate. When the ICECaP analysis evaluated different final and surrogate outcomes, disease specific survival was found to have a low correlation with metastasis when compared with OS. The universal definition of the final outcome as chosen in ICECaP (OS) makes the relationship more robust and generalisable to other jurisdictions.

2B) Epidemiological evidence

The surrogate relationship could also be explained through epidemiological evidence, nevertheless, this type of evidence is more uncertain than that from RCTs. The closer the evidence is to level I (multi-trial meta-regression possible (with one or more drugs) the less important epidemiological data becomes (74). Consequently, as the MFS to OS relationship was built using data from multi trial meta-regression analysis conducted with different interventions, epidemiological data linking both outcomes is not discussed in this analysis.

3B) Part Three: Randomised trial data for other drugs. Mechanisms of drug action studied

3A.1) The different mechanisms of drug action for which the relationship between comparative treatment effects on this PS and this TCO have been studied in randomised trials. Where appropriate, arrange similar mechanisms of action into drug classes.

Randomised clinical trials in the ICECaP meta-analysis were manly radiotherapy-based trials. Nevertheless, the strength of the surrogate relationship between MFS and OS for other treatments with different mechanisms of action has also been tested. ADT was used in some form by all studies included except for one that compared different RT schedules and one used a cryosurgery technique in the intervention arm.

RT and surgery in prostate cancer have one objective, to remove or kill cancer cells. Nevertheless, when RT is performed there is a minimum of safety marginal tissue that receives less than the full

dose planned to diminish toxicities in surrounding healthy tissue. Likewise, the resection of the prostate cancer by surgical means has also marginal tissue spare to protect nerves and the urethra the prostate is surrounding. This tissue that is left behind has the potential to grow and colonise distant and local tissues generating the recurrence of the diseases and eventually developing of metastasis.

The development and progression of prostate cancer have been related to the effect of testosterone on the cells expressing receptors. By depriving patients of this hormone, ADT inhibits the growth of the tumour.

Table 6.17 describe the different mechanism of action found in the included studies. As discussed in section 2A.5, when a different subset of the full data used to correlate MFS and OS were used, there was no impact in terms of the difference in comparative treatment effect estimated by MFS on both groups. Therefore, there are no reasons to believe that the proposed surrogate relationship would not hold when interventions with different mechanisms of actions are used.

Table 6.17. Characteristics of ICECaP studies included and distant metastases definition.

Study	Control Arm	Experimental Arm	Mechanism of action
Cryo vs RT Ontario trials(192)	NADT 6mo + RT	NADT 6mo + Cryoablation	Cryoablation: Apoptosis and necrosis of cells caused by intracellular ice crystal formation and failure of the microcirculation following thawing. ADT: inhibits the action of circulating androgens at the receptor level.
GETUG06(193)	RT 70 Gy	RT 80 Gy	RT: cellular death caused by damaging the DNA after delivering ionising radiation.
MRCRT01(194)	NADT 3-6mo + RT 64 Gy	NADT 3-6mo + RT 74 Gy	ADT: inhibits the action of circulating androgens at the receptor level. RT: cellular death caused by damaging the DNA after delivering ionising radiation.
EORTC22961(195)	NADT 6mo + RT	NADT 6mo + RT + AADT 2.5yr	As above where ADT is used
EORTC22991(189)	RT	RT + AADT 6mo	As above where ADT is used
Mottet et al (196)	ADT 3yr	ADT 3yr + RT	As above where ADT is used
GICOR-DART01/05(197)	NADT 4mo + RT	NADT 4mo + RT + AADT 2yr	As above where ADT is used
ICORG9701(198)	NADT 4mo + RT	NADT 8mo + RT	As above where ADT is used
RTOG9202(188)	NADT 4mo + RT	NADT 4mo + RT + AADT 2yr	As above where ADT is used
RTOG9408(199)	RT	NADT 4mo + RT	As above where ADT is used
RTOG9413(200)	RT + AADT 4mo	NADT 4mo + RT	As above where ADT is used

Study	Control Arm	Experimental Arm	Mechanism of action
TROG9601(201)	RT	NADT 3/6 mo + RT	As above where ADT is used

Abbreviations: AADT= Adyuvant androgen deprivation therapy, ADT= Androgen deprivation therapy, mo= Month, NADT= Neoadyuvant androgen deprivation therapy, RT= Radiotherapy.

3A.2) The biologic reasoning for each of the mechanisms of action of the drug(s) on the PS and thus on the TCO (including any rationale for a lag in onset of treatment effect and/or a lag in offset of treatment effect and whether this has implications for the PS, the TCO or both outcomes). An assessment of how close the mechanisms of action of the studied medicines are to the pathological mechanism(s) of the disease.

By directing ionising radiation beams to the cancer tissue, RT damages the DNA of cancerous cells. Healthy cells respond using DNA repair mechanisms, but cancer cells are often less differentiated, reproduce more and often lack the ability to repair. These characteristics make cancer cells more likely to be affected by RT as damage caused by RT accumulates and those cells with sub-lethal damage pass on the defective DNA when they reproduce, causing a lower reproduction rate or death.

Cryoablation refers to the destruction of tissue by freezing. When ice is formed in the intracellular space, it is thought that the ice crystals cause physical damage to the membranes of organelle and the cell itself. Even though not all malignant cells are frozen in the process, the underlying tissue receiving non-lethal temperatures is subjected to an adverse environment for cellular survival as the result of cold-induced changes to the tissue. After the cryoablation the effect will last for weeks resulting in necrosis of the frozen tissue and its immediate surroundings. (364)

Prostate cancer is as an androgen responsive disease sensitive to the presence of the main active androgen hormones in males, testosterone and dihydrotestosterone. Androgen function in the body is mediated through the interaction with its specific receptor which is found in different tissues but manly in sex organs. An activated androgen receptor is capable of inducing growth of the prostate cancer by targeting genes involved in cell regulation and proliferation and also inhibiting the apoptosis process (365). In this context ADT plays an important role, reducing the availability of androgens hormones and consequently blocking the cascade of action that produces an activated androgen receptor in the neoplastic tissue.

As seen in the previous section, the disease pathway occurs in a series of progression states that ultimately results in dysplasia of the tumour cells due to abnormal growth and the colonisation of

distant organs. By killing or inducing cellular apoptosis of neoplastic tissue, RT and cryoablation reduce the probability of growth and therefore, the progression states. Similar results could be achieved with ADT treatment by cutting one of the signals that promote internal cell growth. By decreasing the size and progression of the disease, these treatments have the potential to stop metastasis development and consequently, death due to prostate cancer. There are no reasons to believe that avoiding metastases after receiving treatment would have any different effect on OS conditional of the mechanism of action of the intervention received.

3A.3) The drugs within each different mechanism of action for which the relationship between comparative treatment effects on this PS and this TCO have been studied in randomised trials:

Table 6.18 shows the drugs and RT techniques as well as the dose used in each of the RCT included in the ICECaP meta-analysis.

Table 6.18. RT technique, drugs and doses used in the ICECaP study.

Study	Control Arm	Experimental Arm	Mechanism of action
Cryo vs RT Ontario trials(192)	3 mo of goserelin previous EBRT (4-field box technique: 66 Gy in 33 fractions administered at 2 Gy per day, 5 days a week) followed by 3 mo of goserelin.	3 mo of goserelin previous the Cryoablation followed by 3 mo of goserelin.	Cryoablation: Apoptosis and necrosis of cells caused by intracellular ice crystal formation and failure of the microcirculation following thawing. ADT: inhibits the action of circulating androgens at the receptor level. RT: cellular death caused by damaging the DNA after delivering ionising radiation.
GETUG06(193)	Three-dimensional conformal RT 70 Gy in 2 Gy/fraction, 5 times/wk	Three-dimensional conformal RT80 Gy in 2 Gy/fraction, 5 times/wk.	

Study	Control Arm	Experimental Arm	Mechanism of action
MRCRT01(194)	Goserelin 3.6 mg/4 weeks or leuprorelin 3.75 mg/4 weeks, accompanied by cyproterone acetate 100 mg 3 times daily for 3 weeks + 3D CRT: 64 Gy in 32 fractions.	Goserelin 3.6 mg/4 weeks or leuprorelin 3.75 mg/4 weeks, accompanied by cyproterone acetate 100 mg 3 times daily for 3 weeks + 3D CRT: 74 Gy in 32 fractions.	ADT: inhibits the action of circulating androgens at the receptor level. RT: cellular death caused by damaging the DNA after delivering ionising radiation.
EORTC22961(195)	3D CRT: 70 Gy 5 time/week x 7 weeks + 6mo of: triptorelin + 750 mg of flutamide per day or 50 mg of bicalutamide per day), initiated 1 week before triptorelin	3D CRT: 70 Gy 5 time/week x 7 weeks + 2.5 years of: triptorelin + 750 mg of flutamide per day or 50 mg of bicalutamide per day), initiated 1 week before triptorelin.	As above
EORTC22991(189)	3D CRT or IMRT 70, 74, or 78 Gy.	78 Gy + 6mo of: goserelin + 50 mg of bicalutamide per day), started 1 week before goserelin.	As above
Mottet et al (196)	leuprorelin 11.25 mg every 3 mo for 3 yr + flutamide 750 mg/d during month 1.	leuprorelin 11.25 mg every 3 mo for 3 yr + flutamide 750 mg/d during month 1 + 3D CRT 70 +- 4 Gy.	As above
GICOR-DART01/05(197)	4mo of goserelin + flutamide 750 mg per day or bicalutamide 50 mg per day during the first 2 mo + 3D CRT: 76–82 Gy + 24 mo of goserelin.	4mo of goserelin + flutamide 750 mg per day or bicalutamide 50 mg per day during the first 2 mo + 3D CRT: 76–82 Gy.	As above
ICORG9701(198)	4mo of Triptorelin + flutamide 750 mg/day + 3D CRT: 70Gy	8mo of Triptorelin + flutamide 750 mg/day + 3D CRT: 70Gy	As above
RTOG9202(188)	4mo of Triptorelin + flutamide 750 mg/day + 3D CRT: 70Gy.	4mo of Triptorelin + flutamide 750 mg/day + 3D CRT: 70Gy + 24mo of triptorelin.	As above
RTOG9408(199)	RT 66.6 Gy	4mo of goserelinm or leuprolide+ flutamide 750 mg/day + RT 66.6 Gy	As above
RTOG9413(200)	4mo goserelin acetate 3.6 mg/month or leuprolide 7.5 mg/month + flutamide 750mg/day. + Conventional four-field 'box': 70.2 Gy.	Conventional four-field 'box': 70.2 Gy + 4mo goserelin acetate 3.6 mg/month or leuprolide 7.5 mg/month + flutamide 750mg/day.	As above

Study	Control Arm	Experimental Arm	Mechanism of action
TROG9601(201)	3mo goserelin acetate 3.6 mg/month or leuprolide 7.5 mg/month + flutamide 750mg/day. + 66 Gy,RT	6mo goserelin acetate 3.6 mg/month or leuprolide 7.5 mg/month + flutamide 750mg/day. + 66 Gy,RT	As above

AADT: Adjuvant androgen deprivation therapy, ADT: Androgen deprivation therapy, mo: Month, NADT: Neoadjuvant androgen deprivation therapy, RT: Radiotherapy.

3A.4) The randomised trials in which the relationship between comparative treatment effects on this PS and this TCO been studied (categorised into trials for each main mechanism of action and then into trials for each drug with similar mechanisms of action).

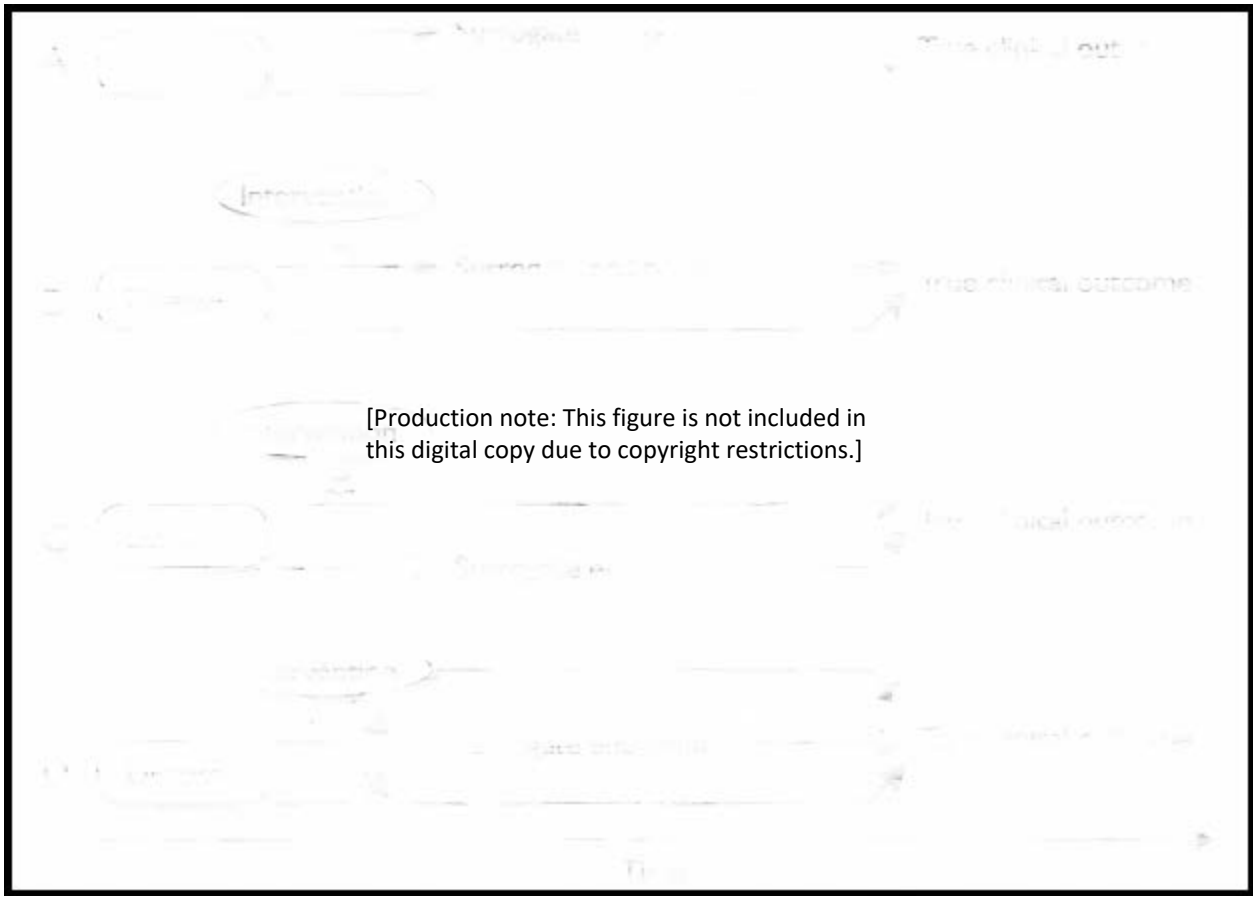
Table 6.18 shows the RCTs included in the ICECaP meta-analysis in which the relationship between MFS and OS has been studied.

3A.5) A biological justification for how the PS-TCO comparative treatment effect relationships could be extrapolated across any different mechanisms of action and different drugs where randomised trial evidence is also available for review.

Although there are multiple interventions with different mechanisms of action, metastasis has been linked with poor prognosis and death. Thus, interventions which improve MFS have the potential to decrease mortality associated with this state of the disease. The surrogacy relationship was evaluated in different primary treatments and adjuvant therapies.

For the surrogate not to be able to capture the treatment effect observed in the final outcome, the intervention evaluated must have an effect on OS which is not mediated by MFS (Figure 6.3). Because MFS is a composite outcome that indirectly involves the final outcome, survival, the treatment effect on MFS is expected to be related to the treatment effect observed in OS.

Figure 6.3. Possible pathways of treatment effect perturbing the surrogacy relationship.



Source: Taylor RS, Elston J. The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports. Health Technol Assess 2009;13(8).

It also important to note that in the sensitivity analysis reported in the study that correlated MFS with OS, (132) the correlation was maintained independently of the type of local treatment received or adjuvant therapy.

3B) Different patient populations and circumstances that the PS has been studied in

3B.1) Different stages of disease in which the relationship between comparative treatment effects on this PS and this TCO been studied in the randomised trials.

Patients diagnosed with prostate cancer are usually classified according to their risk of progression and metastasis development and also according to their tumour characteristics (366). The studies included in the ICECaP database were classified according the TNM[37]

[37] The TNM classification:

classification system for malignant tumours which has been widely used to stage this disease (367). Patients diagnosed in different stages of the disease where the correlation between MFS and OS has been studied are detailed in Table 3.1 by each study included in the ICECaP database. In general, studies included patients with localised or locally advanced disease from T1b to T4 where the cancer has not yet reached lymph nodes or distant organs (T1: 23%, T2: 31%, T3: 29% and T4: 17%). The study did not classify for risk; however, a sensitivity analysis conducted in high-risk population had consistent results compared to the overall population.

3B.2) *Different patient populations (e.g., age, gender, comorbidities) and circumstances (e.g., hospital versus community setting, geographical locations, dates of trials, groups randomised to different dose regimens in the same or across different randomised trials) that the relationship between comparative treatment effects on this PS and this TCO been studied in the randomised trials.*

General characteristics of patients in all studies included in ICECaP are comparable. Patients were of a similar age range and were all diagnosed with localised or locally advanced disease. This is an important aspect of the analysis as the time to event for surrogate and final outcome is subject to these baseline characteristics independent of the treatment received (58).

Prostate cancer is a late onset disease, with a median age at diagnosis in Australia around 67 years of age (92). Comorbidities increase with age and the number of concomitant comorbidities is higher in those aged 65 or older with cardiovascular and metabolic diseases

---

The TNM classification has been used worldwide to evaluate the extent of the disease, how much the tumour has grown and possible spread to other organs and tissues. In this cancer staging system T refers to the size and extent of the main tumour, N refers to the number of lymph nodes involved, and M refers to whether the cancer has metastasised.
T1: cancer is not detectable on a scan or felt during DRE. It is subdivided into T1a-c where a) refers to a tumour less than 5% of the resected tissue; b) refers when tumours correspond to 5% or more of the resected tissue and c) when the cancer is found by biopsy
T2: cancer is localised inside the prostate gland. It is subdivided into T2a-c. a) refers to tumours localised is in only one half of one side of the prostate gland; b) refers to a tumour found in more than one half of one side of the prostate gland, but not both sides and c) when the neoplastic tissue is found in both sides.
T3: cancer has spread through the capsule of the prostate gland. It is subdivided in T3a-b. a) refers to tumour spread through the capsule of the prostate gland and b) when the tumour has spread into the seminal vesicles.
T4: cancer has spread into nearby organs other than the seminal vesicles, such as the back passage, bladder, or the pelvic wall.
N is subdivided into N0-1 where 0) refers to nearby lymph nodes free of cancer cells and 1) when cancer cells are found in lymph nodes near the prostate
There are 2 stages of metastasis, M0 when the cancer has not spread to distant organs and M1 when the cancer has spread to organs outside the pelvis. M1 is subdivided in M1a-c where a) refers to cancer cells in distant lymph nodes; b) refers to bone colonisation by cancer cells and c) refers to cancer cells in distant organs other than bone.

among the most prevalent (368). Metabolic comorbidities are an important factor to consider in this population of patients due to the possible sides effect of ADT. In particular, cardiovascular diseases could be exacerbated by prolonged ADT which may affect adherence to treatment or the dropout rate which eventually might impact the treatment effect observed (201). Although the majority of the studies used ADT in some form, only three provided information about baseline comorbidities (189, 193, 195), and only two provided information about baseline cardiovascular disease (189, 195). No study stratified outcomes by baseline cardiovascular disease. Thus, whether there is an impact on the surrogate or final outcome due to baseline comorbidities remain uncertain. Table 6.19 describe the population characteristics of the studies included.

Table 6.19. Population characteristics of the ICECaP studies.

Study	Stage	Year Enrolled	Median age	Risk	Median Follow-up (years)
Cryo vs RT Ontario trials(192)	T2c-3b	1999-2002	70	NR	10.7
GETUG06(193)	T1b-3a	1999-2002	67	I: 46% H:54%	5.1
MRCRT01(194)	T1b-3a	1998-2002	67	L:19% I:37% H: 44%	10
EORTC22961(195)	T1c-4	1998-2002	69	NR	6
EORTC22991(189)	T1b-2a	1997-2001	70	I: 75% H: 25%	7.2
Mottet et al (196)	T3-4	2001-2008	71	NR	8.4
GICOR-DART01/05(197)	T1c-3b	2000-2003	71	I: 8% H: 92%	4.5
ICORG9701(198)	T1-4	2005-2010	67	I: 20% H: 80%	12
RTOG9202(188)	T2c-4	1997-2001	70	NR	19.6
RTOG9408(199)	T1b-2b	1992-1995	71	L: 35% I: 54% H: 11%	9.9
RTOG9413(200)	T1c-4	1994-2001	70	NR	17.1
TROG9601(201)	T2b-4	1995-1999	68	I: 16% H: 84%	11.4

3C) The quality of the randomised trials (e.g., as assessed according to the standard approach outlined in Section B of the PBAC Guidelines).

3C.1) The quality of the randomised trials (e.g., as assessed according to the standard approach outlined in Section B of the PBAC Guidelines).

Table 6.5 summarise the risk of bias as outlined in section 2 of the PBAC guidelines (369).

3C.2) Whether relevant randomised trials have been excluded from any systematic overviews or meta-analyses and, if so, the reasons for each exclusion.

From the 102 trials identified as potentially eligible, 43 (42%) had both data suitable for use and a study group that agreed to participate. This resulted in possible IPD from 28,905 patients. As not all trials collected MFS, the meta-analysis used data from 12,712 patients in 19 (44%) of 43 trials. Those trials that did not document data on MFS were excluded. The detailed information can be accessed from the ICECaP study and their supplementary material (132).

3C.3) Whether the analysis of the PS in each randomised trial was designed retrospectively or prospectively.

All trials included in the ICECaP were retrospective.

3C.4) *Whether each trial was designed to be included in the meta-analysis prospectively.*

The ICECaP database was an initiative where information from randomised clinical trials conducted in localised prostate cancer was analysed. The objective was to find a surrogate outcome capable of replacing OS given the characteristics of this disease. A systematic review was conducted to identify the potential trials published at the time. The study groups of relevant publications found were invited to participate in the initiative and to provide IPD from their study. Therefore, all studies used in the meta-analysis were designed, conducted, and published before the systematic review was conducted.

3C.5) *The presence of extension studies of these randomised trials which are continuing to measure the comparative treatment effects on the PS and/or the TCO.*

Data from extended follow-up trials identified and used in the meta-analysis conducted by ICECaP were captured in the information provided to them.

3D) Statistical analysis

3D.1) The results of any meta-analyses (meta-regressions) that have been performed to predict the effect of treatment on the TCO based on the observed effect of treatment on the PS.

The results of the meta-regression of all studies included in ICECaP:

i. the intercept and coefficient (and their 95% CI):

Log(HR)-OS= -0.0056 + 0.7813* Log(HR)-MFS

ii. the coefficient of determination (R^2 trial)

R^2 trial: 0.93 C.I = 0.76 to 0.96

iii. R^2 individual:

Correlation at the patient level measured by Kendall's τ from a copula model

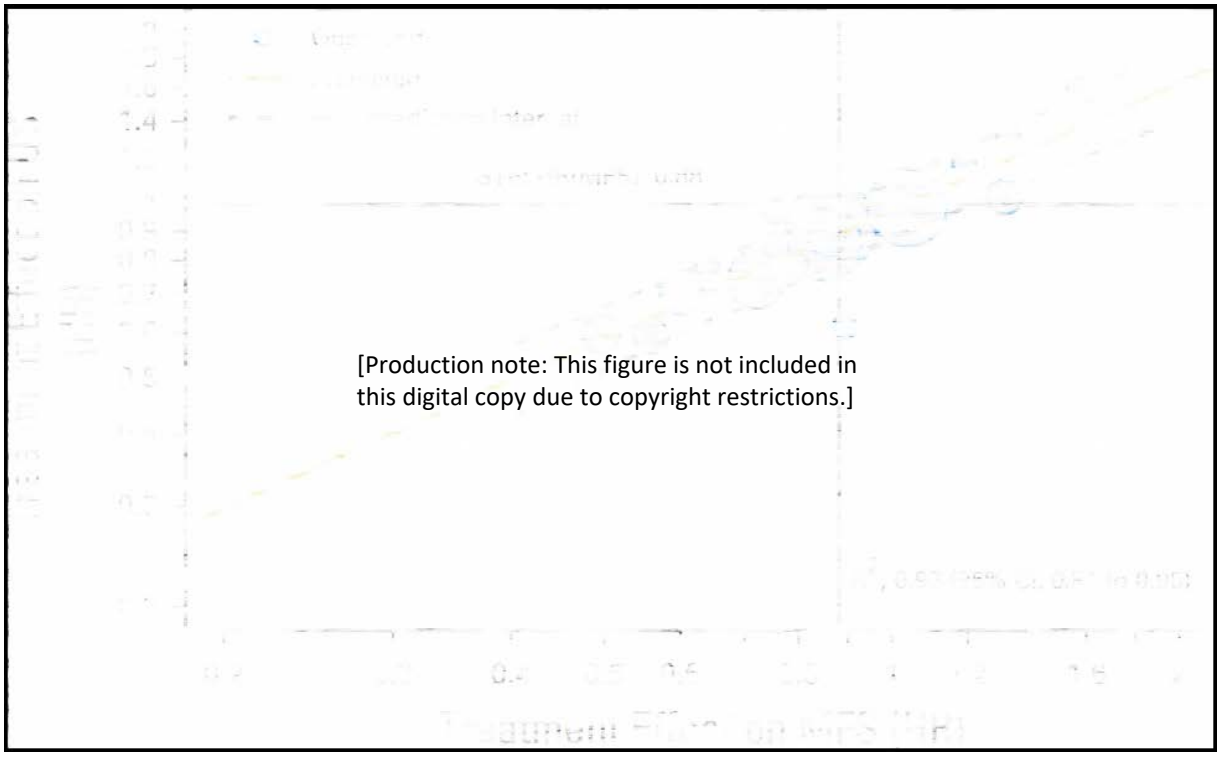
R^2 individual: 0.91 C.I = 0.91 to 0.91

iv. the surrogate threshold effect (STE) as determined by prediction bands.

Figure 6.4 shows the surrogate threshold effect (STE) for MFS as a surrogate of OS along with the prediction bands. The STE in this case is of 0.88.

Nevertheless, as previoulsy discussed, a diferent dataset was used for the economic evaluation. Figure 3.3 describes the difference observed in the estimated HR for OS when the full data set of ICECaP is used compared with the data for health economic analysis. Nevertheless, although some degree of variability is patent, results are similar when a 'practical' range of HR for MFS is used (see section 3.5). This shows robustness of the results as similar comparative results were observed when a subset of patients which mainly used RT as their primary treatment, was used to estimate the surrogacy relationship.

Figure 6.4. Surrogate threshold effect and predictor bands.

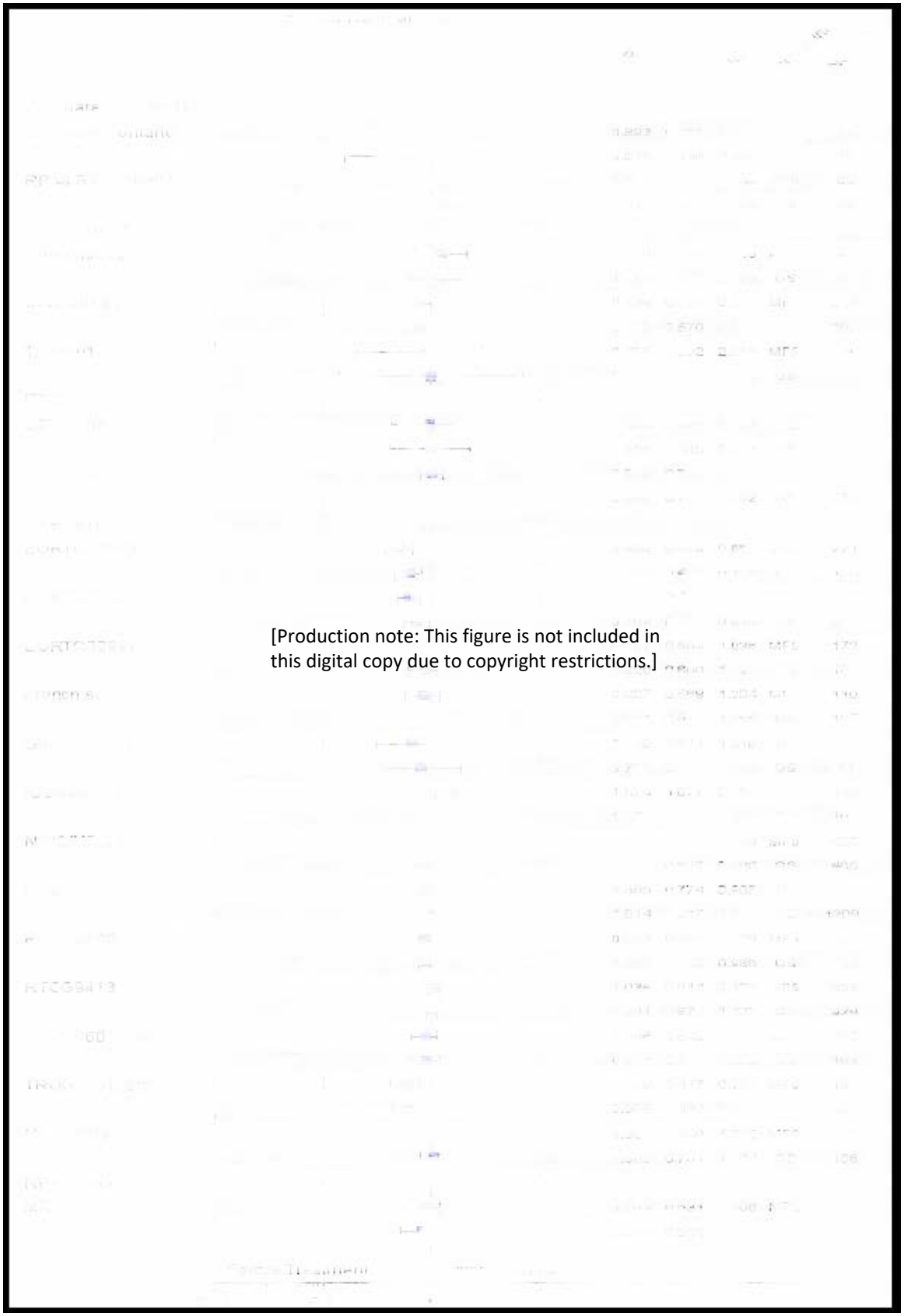


Source: Xie et.al 2017.

3E) Consistency of results across the randomised trials

3E.1) A multidimensional assessment of MFS and OS across populations and trials is given in the form of a forest plot (Figure 6.5). Not all of the studies included in Figure 6.5 were used to calculate the relationship between MFS and OS for the health economic analysis. As noted Chapter 3, only the studies reporting MFS, EFS and OS were used to inform the surrogacy relationship to be applied in the economic model. All of the trials included in the meta-analysis reporting the outcomes of interest have similar qualitative results and the direction of treatment effect is the same.

Figure 6.5. Impact of treatment on MFS and OS



Source: Xie et.al 2017.

3F) Multiplicity of disease pathways, treatment effects, PSs and TCOs

3F.1) An assessment of the consistency of the results across different mechanisms of action and different drugs identified for 3A.4 above and whether the biological reasoning available for 3A.5 above plausibly explains any inconsistencies:

The surrogacy relationship is independent of the primary treatment received or type of adjuvant therapy (Table 6.20).

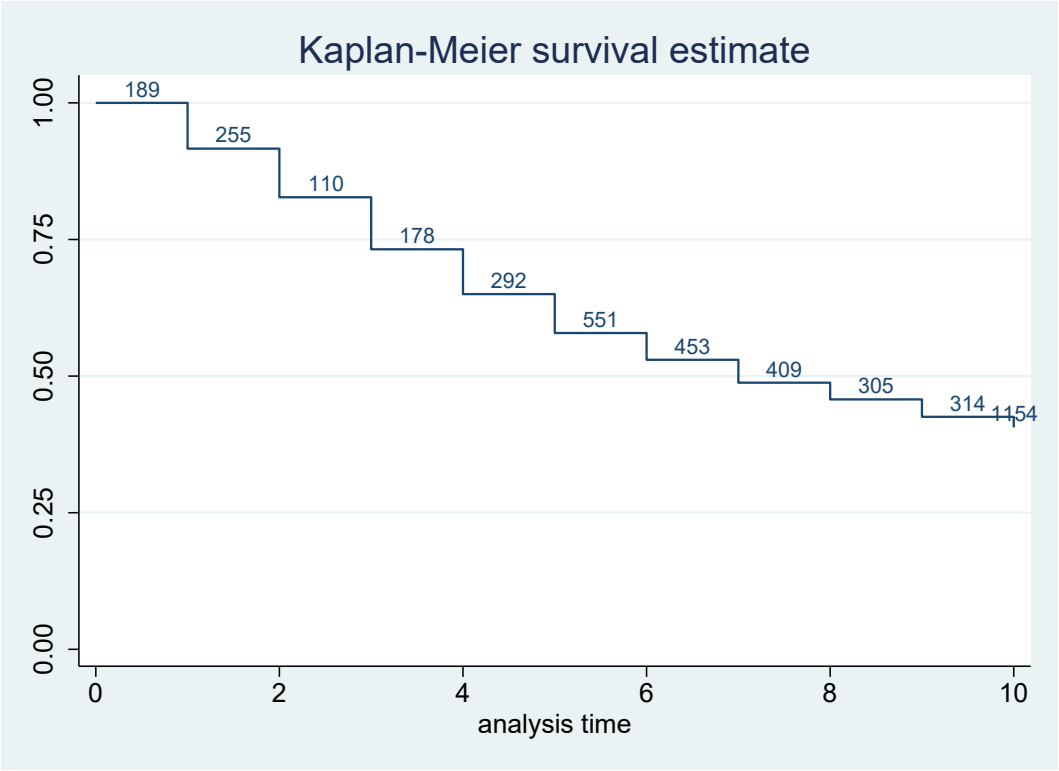
Table 6.20. Subgroup analysis of correlation between the surrogates and true endpoints at the patient level.

	Correlation between MFS and OS	
	No. of patients (trials)	Kendall's Tau (95%CI)
All patients	12,712(21)	0.91 (0.91-0.91)
By type of therapy		
RP based		
RT based*	11,599(17)	0.91 (0.90-0.91)
No local therapy		
Within RT-based trials		
No use of ADT**	2,544(8)	0.91 (0.90- 0.92)
Use of ADT	9,055(15)	0.91 (0.91 0.91)
Duration of ADT		
≤6months	5,674(11)	0.90 (0.89-0.90)
>6months	3,381(8)	0.93 (0.93-0.94)
High risk patients only	8,394(21)	0.88 (0.88-0.89)

Appendix 4. Extrapolation and model parameters

A. Extrapolation and model parameters for the overall population developing biochemical recurrence.

Figure 6.2. KM BCR in overall population.

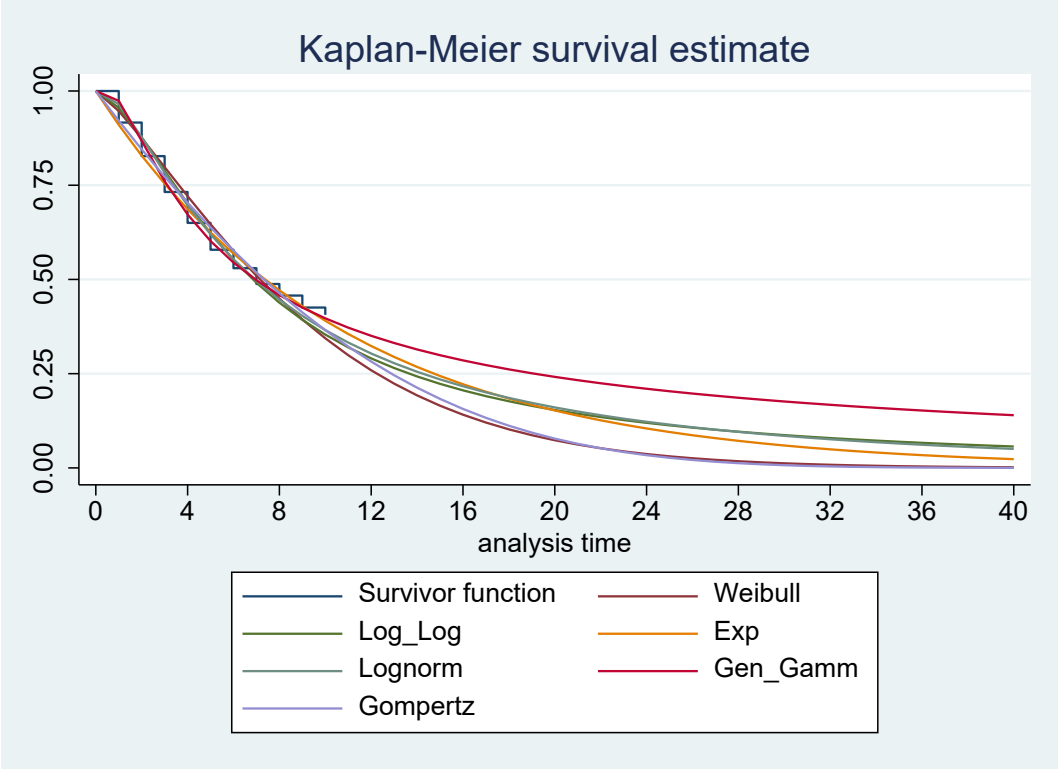


Source: produced in this thesis with the ICECaP data.

Table 21. Survivor function of BCR for overall population.

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
0	8184	0	189	1	.	.	.
1	7995	669	255	0.9163	0.0031	0.91	0.9222
2	7071	687	110	0.8273	0.0043	0.8187	0.8355
3	6274	722	178	0.7321	0.005	0.7221	0.7418
4	5374	602	292	0.6501	0.0055	0.6392	0.6607
5	4480	491	551	0.5788	0.0057	0.5675	0.59
6	3438	291	453	0.5298	0.0059	0.5181	0.5414
7	2694	213	409	0.4879	0.0061	0.4759	0.4999
8	2072	130	305	0.4573	0.0063	0.445	0.4696
9	1637	114	314	0.4255	0.0065	0.4127	0.4382
10	1209	55	1154	0.4061	0.0067	0.3929	0.4193

Figure 6.3. Extrapolation of different parametric functions of the survival probabilities for BCR in the overall population.



Source: produced in this thesis with the ICECaP data.

The STATA output for survival analysis of BCR is shown on Table 22.

Table 22. Akaike's information criterion and Bayesian information criterion of parametrised extrapolations for BCR in the overall population.

```
. streg, distribution(loglogistic)

         failure _d:  BCR
   analysis time _t:  year
             weight:  [fweight=w]

Iteration 0:   log likelihood = -14932.676  (not concave)
Iteration 1:   log likelihood = -8781.0325
Iteration 2:   log likelihood = -8663.1825
Iteration 3:   log likelihood = -8600.4581
Iteration 4:   log likelihood =  -8600.203
Iteration 5:   log likelihood =  -8600.203

Loglogistic regression -- accelerated failure-time form

No. of subjects =        8,184                 Number of obs     =        8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               Wald chi2(0)      =            .
Log likelihood  =    -8600.203                 Prob > chi2       =            .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	1.924003	.0140652	136.79	0.000	1.896435	1.95157
/ln_gam	-.4647739	.0133568	-34.80	0.000	-.4909527	-.4385952
gamma	.6282771	.0083917			.612043	.6449418

```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	.	-8600.203	2	17204.41	17218.43

Note: N=Obs used in calculating BIC; see [R] BIC note.

```
. streg, distribution(exp) nohr

        failure _d:  BCR
  analysis time _t:  year
            weight:  [fweight=w]

Iteration 0:   log likelihood = -8889.9712
Iteration 1:   log likelihood = -8889.9712

Exponential regression -- log relative-hazard form

No. of subjects =        8,184                 Number of obs     =        8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               LR chi2(0)        =         0.00
Log likelihood  =   -8889.9712                 Prob > chi2       =            .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	-2.363734	.015863	-149.01	0.000	-2.394825	-2.332643

```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	-8889.971	-8889.971	1	17781.94	17788.95

Note: N=Obs used in calculating BIC; see [R] BIC note.

```
. streg, distribution(ggamma)

        failure _d:  BCR
  analysis time _t:  year
            weight:  [fweight=w]

Iteration 0:   log likelihood = -11682.549
Iteration 1:   log likelihood =  -9099.473
Iteration 2:   log likelihood =  -8557.007
Iteration 3:   log likelihood =  -8460.206
Iteration 4:   log likelihood = -8425.5454
Iteration 5:   log likelihood = -8424.8032
Iteration 6:   log likelihood = -8424.7981
Iteration 7:   log likelihood = -8424.7981

Generalized gamma regression -- accelerated failure-time form

No. of subjects =        8,184                 Number of obs    =       8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               Wald chi2(0)     =           .
Log likelihood  =   -8424.7981                 Prob > chi2      =           .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	1.52859	.0410039	37.28	0.000	1.448224	1.608956
/ln_sig	.1544511	.0119136	12.96	0.000	.1311009	.1778012
/kappa	-.9645067	.0829032	-11.63	0.000	-1.126994	-.8020193
sigma	1.167017	.0139033			1.140083	1.194588

```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	.	-8424.798	3	16855.6	16876.63

Note: N=Obs used in calculating BIC; see [R] BIC note.

```
. streg, distribution(gompertz) nohr

         failure _d:  BCR
   analysis time _t:  year
             weight:  [fweight=w]

Iteration 0:   log likelihood = -37768.383
Iteration 1:   log likelihood =  -8984.287
Iteration 2:   log likelihood =  -8869.212
Iteration 3:   log likelihood = -8865.2958
Iteration 4:   log likelihood = -8865.2919
Iteration 5:   log likelihood = -8865.2919

Gompertz regression -- log relative-hazard form

No. of subjects =        8,184                 Number of obs     =        8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               Wald chi2(0)      =            .
Log likelihood  =   -8865.2919                 Prob > chi2       =            .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	-2.519379	.0277749	-90.71	0.000	-2.573817	-2.464941
/gamma	.0430963	.0060707	7.10	0.000	.0311981	.0549946

```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	.	-8865.292	2	17734.58	17748.6

Note: N=Obs used in calculating BIC; see [R] BIC note.

```
. streg, distribution(lognormal)

         failure _d:  BCR
   analysis time _t:  year
             weight:  [fweight=w]

Iteration 0:   log likelihood = -20852.561  (not concave)
Iteration 1:   log likelihood = -8633.1562
Iteration 2:   log likelihood = -8495.4195
Iteration 3:   log likelihood =  -8494.549
Iteration 4:   log likelihood = -8494.5489

Lognormal regression -- accelerated failure-time form

No. of subjects =        8,184                 Number of obs     =        8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               Wald chi2(0)      =            .
Log likelihood  =   -8494.5489                 Prob > chi2       =            .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	1.935543	.0147694	131.05	0.000	1.906595	1.96449
/ln_sig	.0661249	.0121862	5.43	0.000	.0422404	.0900095
sigma	1.06836	.0130193			1.043145	1.094185

```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	.	-8494.549	2	16993.1	17007.12

Note: N=Obs used in calculating BIC; see [R] BIC note.

```
. streg, distribution(weibull) time nohr

         failure _d:  BCR
   analysis time _t:  year
             weight:  [fweight=w]

Fitting constant-only model:

Iteration 0:   log likelihood = -8889.9712
Iteration 1:   log likelihood = -8733.7613
Iteration 2:   log likelihood =  -8732.421
Iteration 3:   log likelihood = -8732.4209

Fitting full model:
Iteration 0:   log likelihood = -8732.4209

Weibull regression -- accelerated failure-time form

No. of subjects =        8,184                 Number of obs     =        8,184
No. of failures =        3,974
Time at risk    =     42245.89
                                               LR chi2(0)        =         0.00
Log likelihood  =   -8732.4209                 Prob > chi2       =            .
```

_t	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	2.252083	.0131859	170.80	0.000	2.226239	2.277927
/ln_p	.2552762	.0136294	18.73	0.000	.2285631	.2819894
p	1.290818	.0175931			1.256793	1.325765
1/p	.7747025	.0105587			.7542817	.7956761

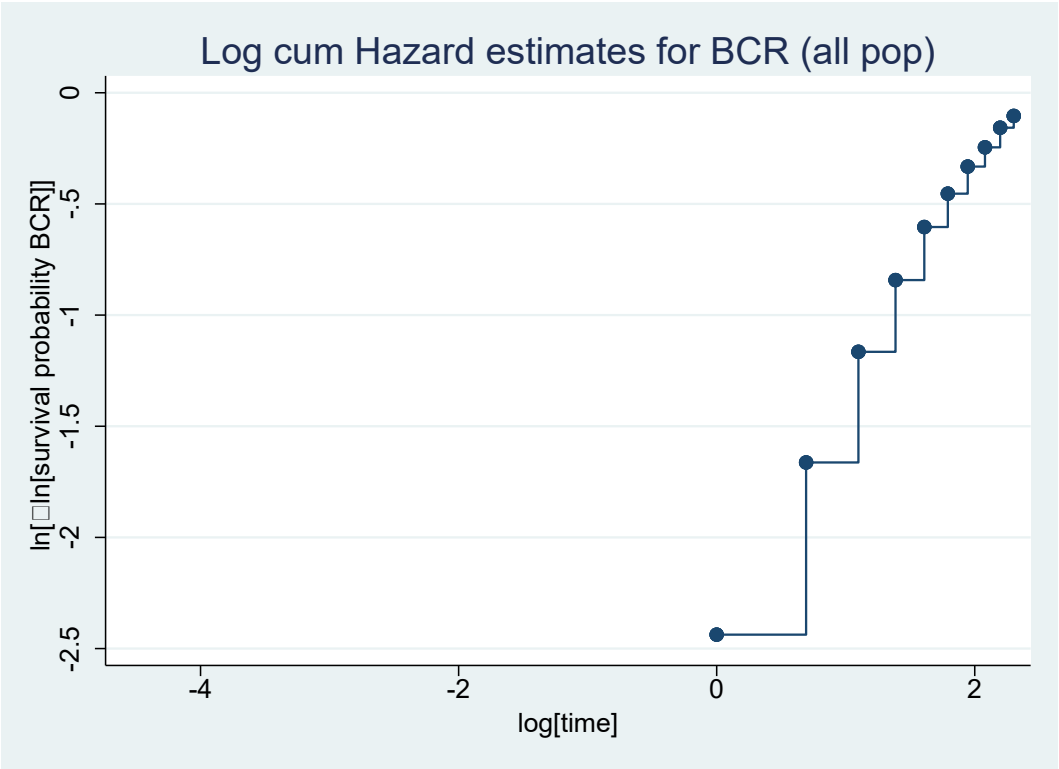
```
.
. estat ic

Akaike's information criterion and Bayesian information criterion
```

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	8,184	-8732.421	-8732.421	2	17468.84	17482.86

Note: N=Obs used in calculating BIC; see [R] BIC note.

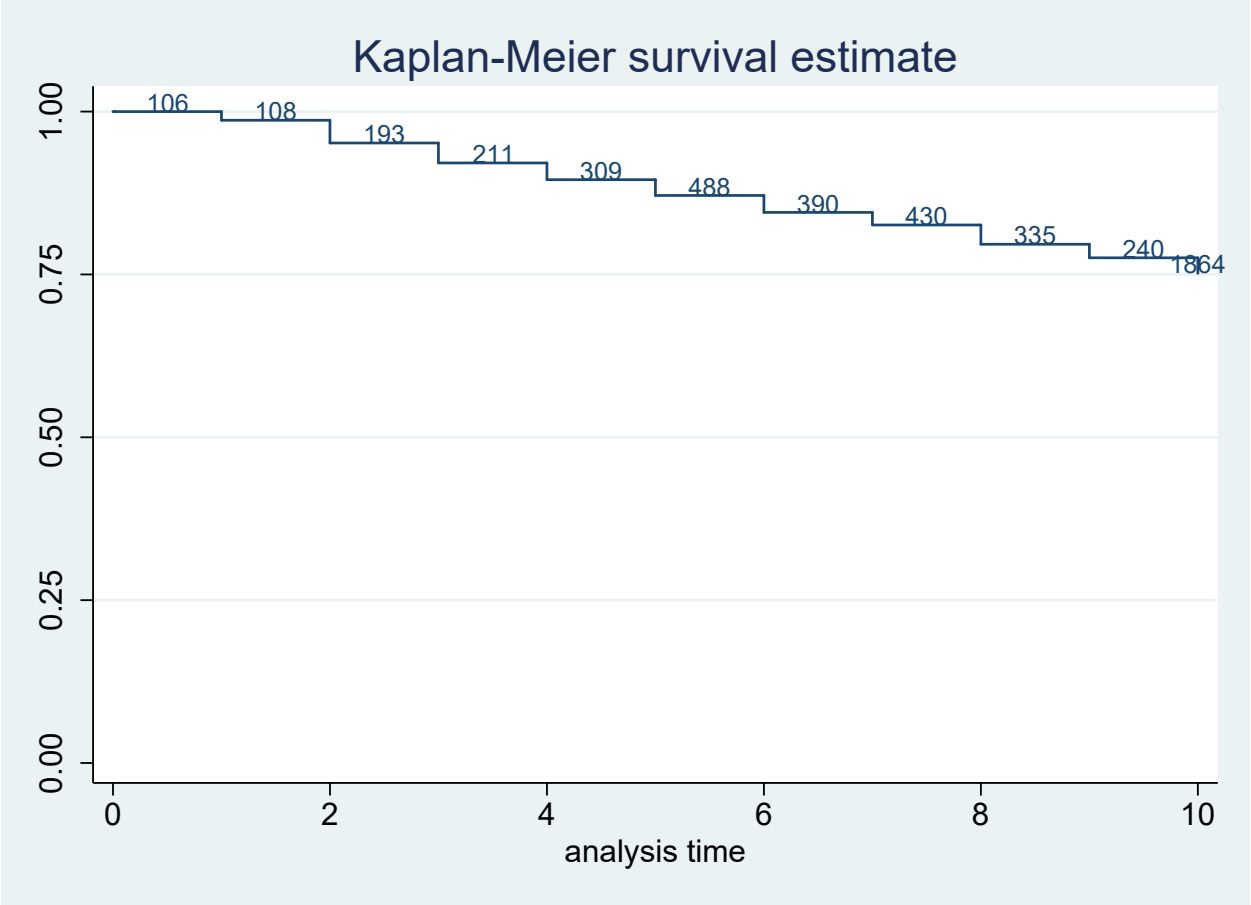
Figure 6.4. Log cumulative hazard of BCR survival probabilities in the overall population.



Source: produced in this thesis with the ICECaP data.

B. Extrapolation and model parameters for the overall population developing metastasis.

Figure 6.5. KM Metastasis in overall population.

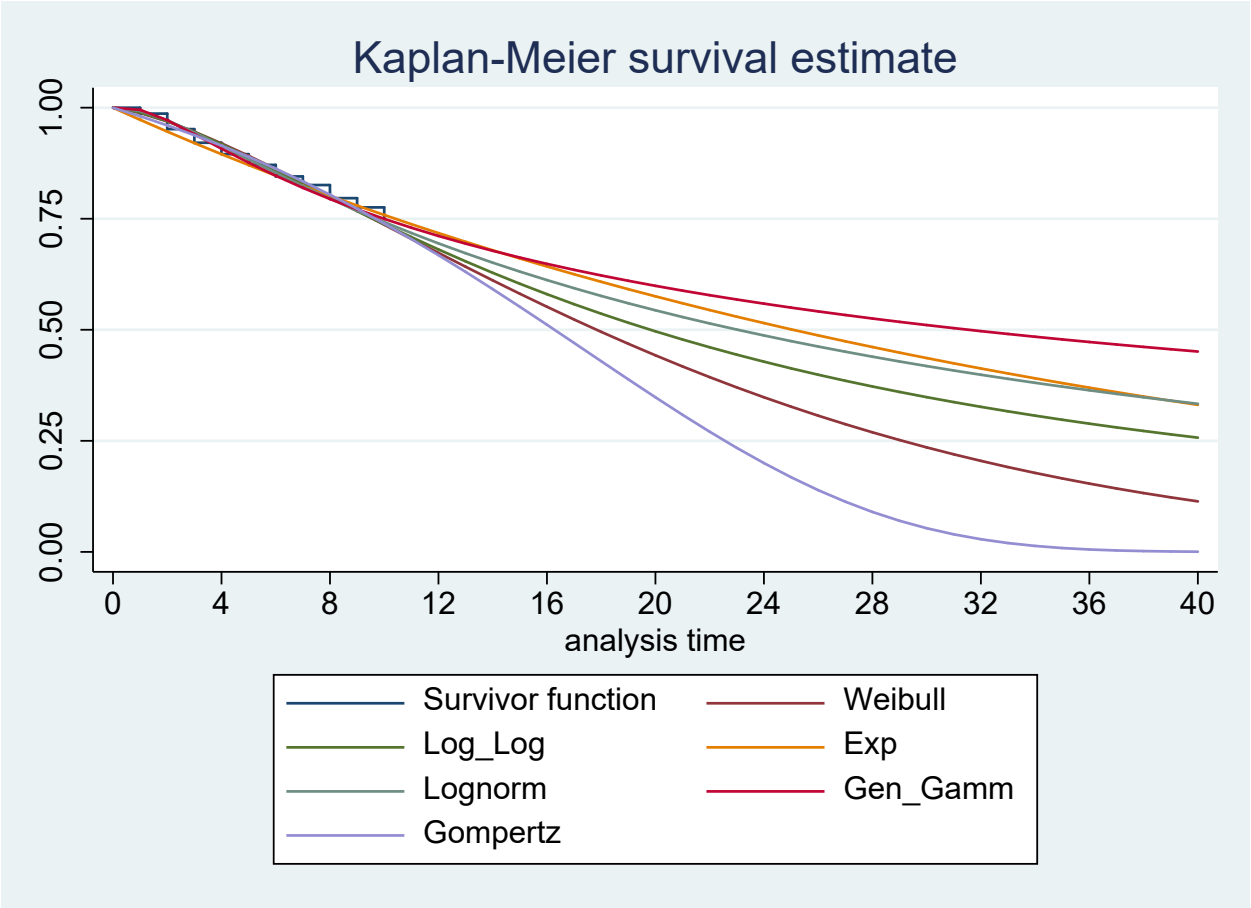


Source: produced in this thesis with the ICECaP data.

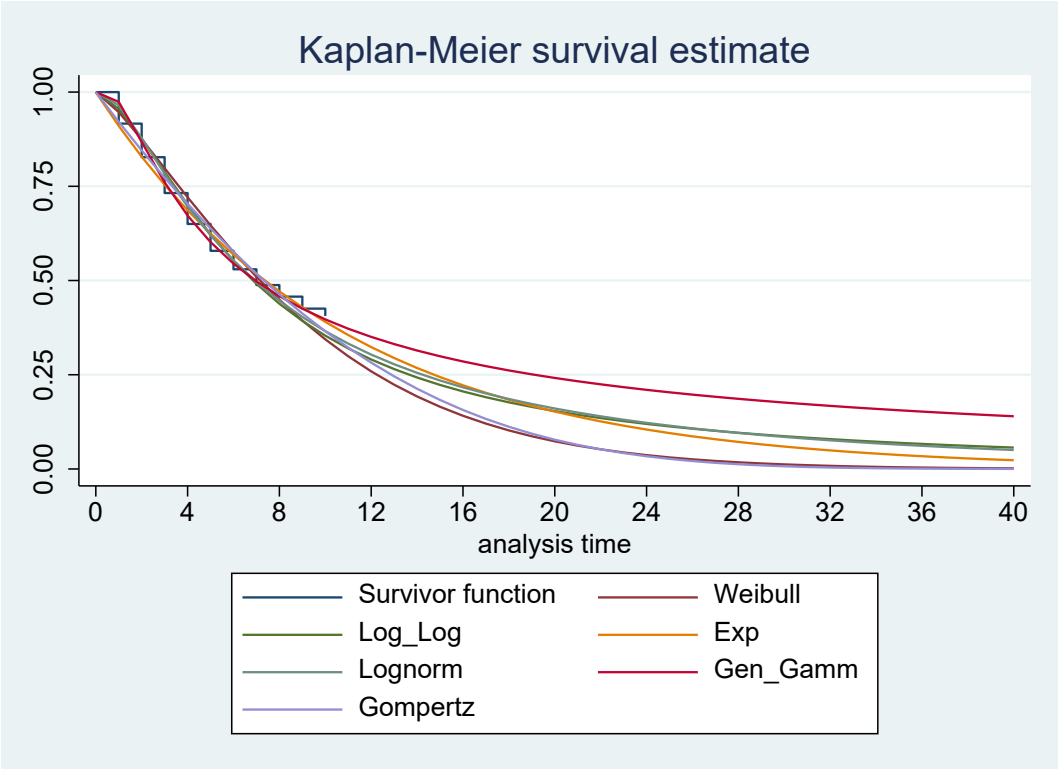
Table 23. Survivor function of metastasis for overall patients diagnosed with prostate cancer.

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
0	5744	0	106	1	NA	NA	NA
1	5638	75	108	0.9867	0.0015	0.9833	0.9894
2	5455	193	193	0.9518	0.0029	0.9458	0.9571
3	5069	163	211	0.9212	0.0036	0.9137	0.928
4	4695	131	309	0.8955	0.0042	0.887	0.9034
5	4255	115	488	0.8713	0.0046	0.8619	0.8801
6	3652	109	390	0.8453	0.0051	0.8349	0.855
7	3153	72	430	0.826	0.0055	0.8149	0.8364
8	2651	95	335	0.7964	0.0061	0.7842	0.808
9	2221	58	240	0.7756	0.0065	0.7625	0.788
10	1923	59	1864	0.7518	0.007	0.7377	0.7652

Figure 6.6. Extrapolation of different parametric functions of the survival probabilities for metastasis in the overall population.



Source: produced in this thesis with the ICECaP data.



Source: produced in this thesis with the ICECaP data.

The STATA outputs for the survival analysis of MFS is shown in Table 24.

Table 24. Akaike's information criterion and Bayesian information criterion of parametrised extrapolations for MFS in the overall population.

Weibull AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	3.14	0.036	86.75	0	3.069	3.211	***
ln_p	0.347	0.028	12.6	0	0.293	0.401	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Chi-square	0
Prob > chi2	.	Akaike crit. (AIC)	6733.078

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	-3364.539	-3364.539	2	6733.078	6746.39

Note: N=Obs used in calculating BIC; see [R] BIC note.

Exponential PH regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	-3.589	0.031	-117.38	0	-3.648	-3.529	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Chi-square	0
Prob > chi2	.	Akaike crit. (AIC)	6869.855

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	-3433.928	-3433.928	1	6869.855	6876.511

Note: N=Obs used in calculating BIC; see [R] BIC note.

Loglogistic AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.987	0.035	85.77	0	2.918	3.055	***
lngamma	-0.412	0.027	-15.26	0	-0.465	-0.359	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's informatio Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-3358.139	2	6720.278	6733.59

Note: N=Obs used in calculating BIC; see [R] BIC note.

Lognormal AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	3.137	0.041	76.47	0	3.057	3.217	***
lnsigma	0.248	0.025	10.1	0	0.2	0.297	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-3333.049	2	6670.099	6683.411

Note: N=Obs used in calculating BIC; see [R] BIC note.

Gompertz PH regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	-3.99	0.06	-66.24	0	-4.108	-3.872	***
gamma	0.091	0.011	8.35	0	0.07	0.113	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-3399.633	2	6803.266	6816.578

Note: N=Obs used in calculating BIC; see [R] BIC note.

Generalized gamma AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.893	0.102	28.33	0	2.693	3.093	***
lnsigma	0.519	0.049	10.5	0	0.422	0.615	***
kappa	-0.919	0.242	-3.79	0	-1.394	-0.444	***

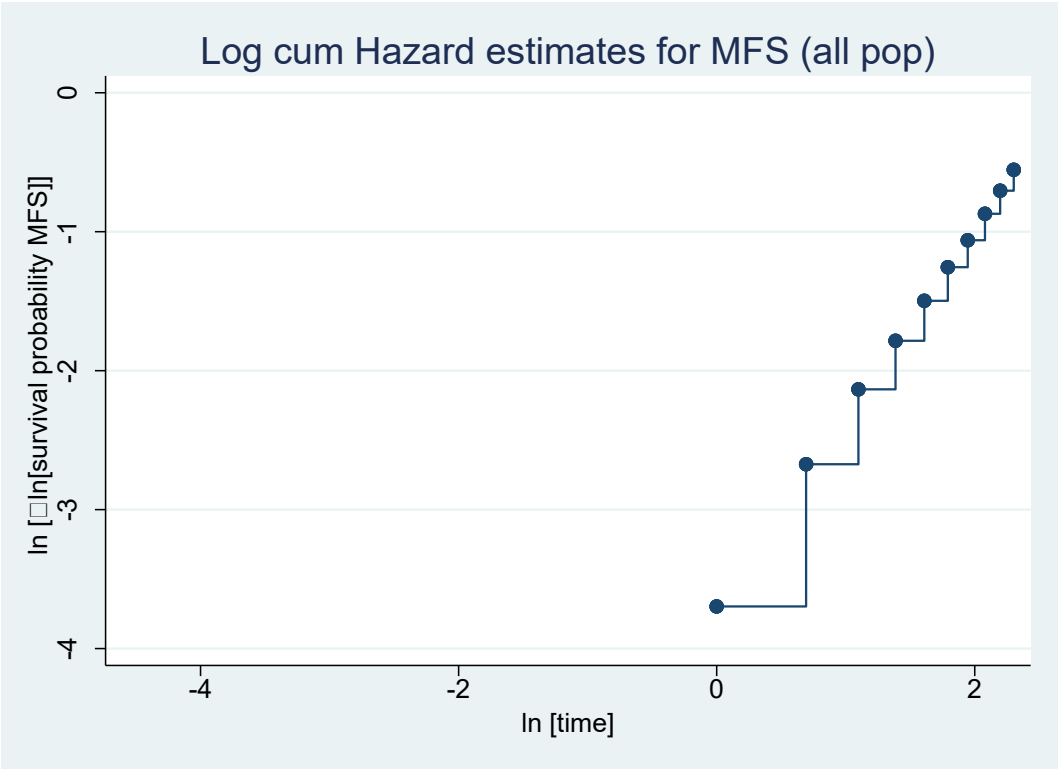
Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-3324.658	3	6655.315	6675.283

Note: N=Obs used in calculating BIC; see [R] BIC note.

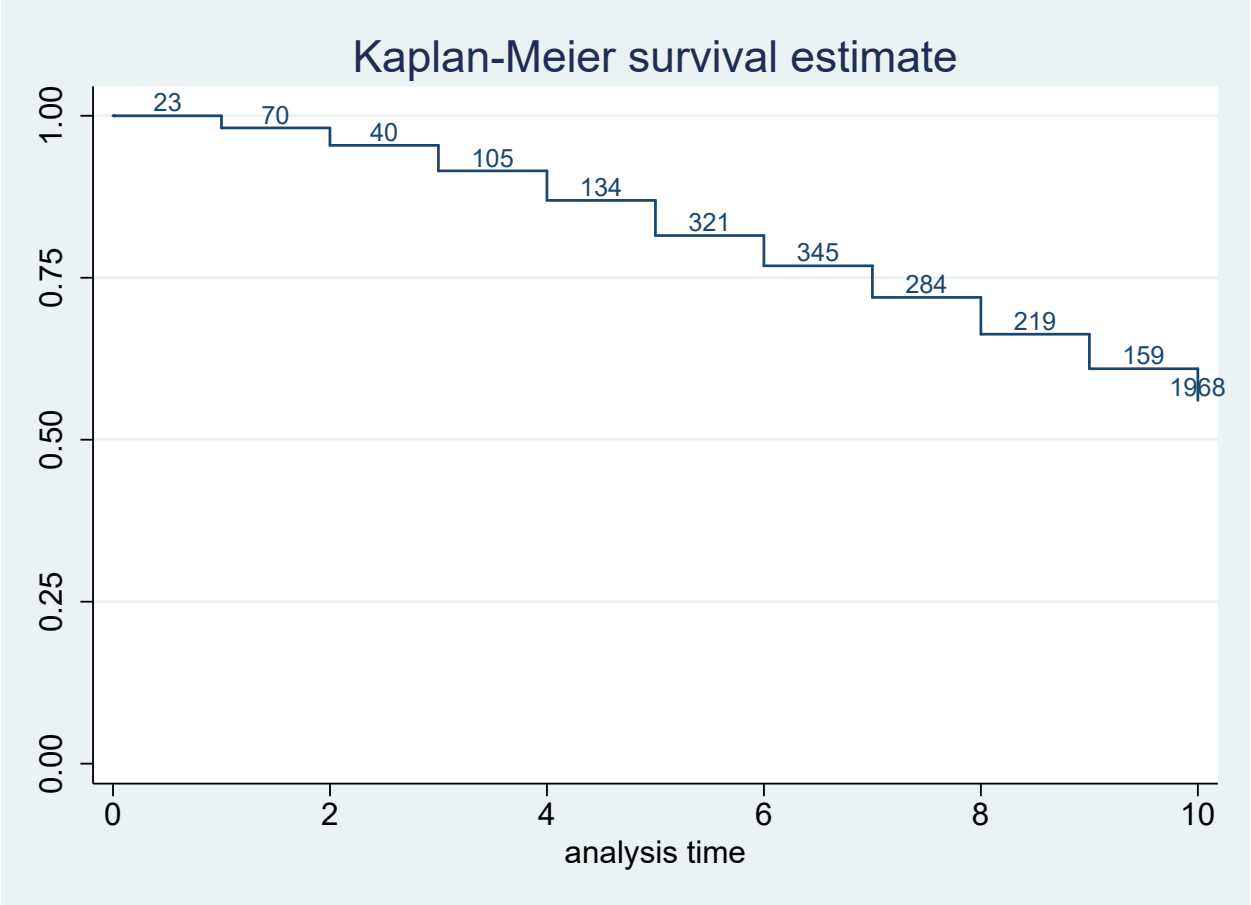
Figure 6.7. Log cumulative hazard of MFS survival probabilities in the overall population.



Source: produced in this thesis with the ICECaP data.

C. Extrapolation and model parameters for OS in the overall population.

Figure 6.8. KM Overall Survival in the overall population.

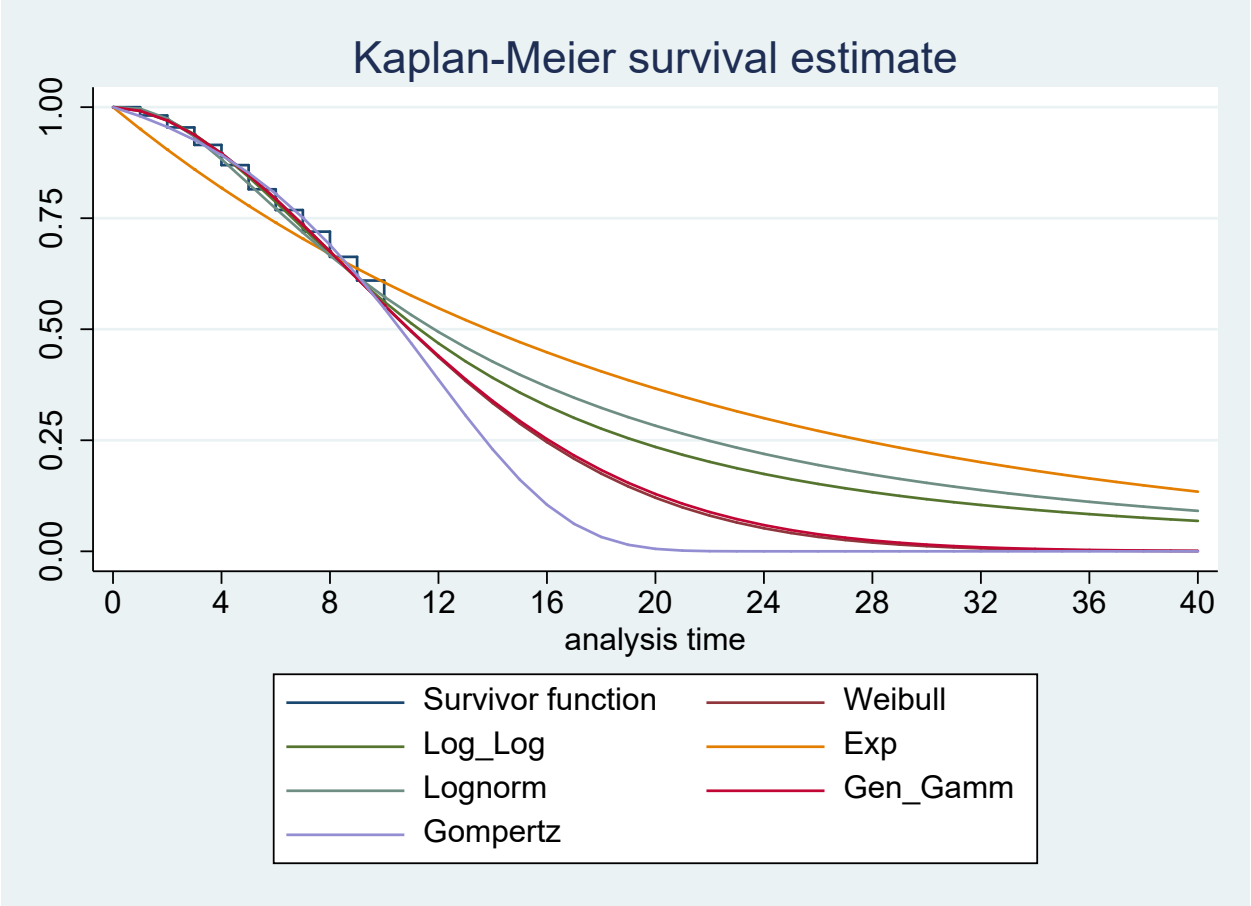


Source: produced in this thesis with the ICECaP data.

Table 25. Survivor function for OS for the overall population.

Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf. Int.]	
0.01	5744	0	23	1	.	.	.
1	5721	107	70	0.9813	0.0018	0.9774	0.9845
2	5544	152	40	0.9544	0.0028	0.9486	0.9595
3	5352	221	105	0.915	0.0037	0.9074	0.922
4	5026	250	134	0.8695	0.0045	0.8604	0.878
5	4642	290	321	0.8152	0.0052	0.8046	0.8252
6	4031	231	345	0.7684	0.0058	0.7569	0.7795
7	3455	219	284	0.7197	0.0063	0.7072	0.7318
8	2952	233	219	0.6629	0.0068	0.6494	0.676
9	2500	201	159	0.6096	0.0072	0.5953	0.6236
10	2140	172	1968	0.5606	0.0075	0.5457	0.5753

Figure 6.9. Extrapolation of survival curves for OS in the overall population.



Source: produced in this thesis with the ICECaP data.

The STATA outputs for the survival analysis of OS is shown in Table 24.

Table 26. Akaike's information criterion and Bayesian information criterion of parametrised extrapolations of OS in the overall population.

Weibull AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.589	0.015	175.1	0	2.56	2.617	***
ln_p	0.609	0.02	30.99	0	0.571	0.648	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Chi-square	0
Prob > chi2	.	Akaike crit. (AIC)	9167.124

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	-4581.562	-4581.562	2	9167.124	9180.436

Note: N=Obs used in calculating BIC; see [R] BIC note.

Loglogistic AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.424	0.015	160.13	0	2.394	2.453	***
lngamma	-0.724	0.019	-37.8	0	-0.762	-0.687	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-4591.816	2	9187.633	9200.945

Note: N=Obs used in calculating BIC; see [R] BIC note.

Exponential PH regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	-2.992	0.022	-136.32	0	-3.035	-2.949	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Chi-square	0
Prob > chi2	.	Akaike crit. (AIC)	9929.154

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	-4963.577	-4963.577	1	9929.154	9935.81

Note: N=Obs used in calculating BIC; see [R] BIC note.

Lognormal AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.471	0.018	137.42	0	2.436	2.507	***
lnsigma	-0.091	0.017	-5.37	0	-0.125	-0.058	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-4618.368	2	9240.737	9254.048

Note: N=Obs used in calculating BIC; see [R] BIC note.

Generalized gamma AFT regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	2.585	0.018	140.73	0	2.549	2.621	***
lnsigma	-0.58	0.098	-5.93	0	-0.771	-0.388	***
kappa	0.952	0.157	6.04	0	0.643	1.26	***

Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-4581.518	3	9169.036	9189.004

Note: N=Obs used in calculating BIC; see [R] BIC note.

Gompertz PH regression

_t	Coef.	St.Err.	t-value	p-value	[95% Conf	Interval]	Sig
Constant	-3.991	0.051	-78.93	0	-4.09	-3.892	***
gamma	0.203	0.008	25.42	0	0.187	0.219	***

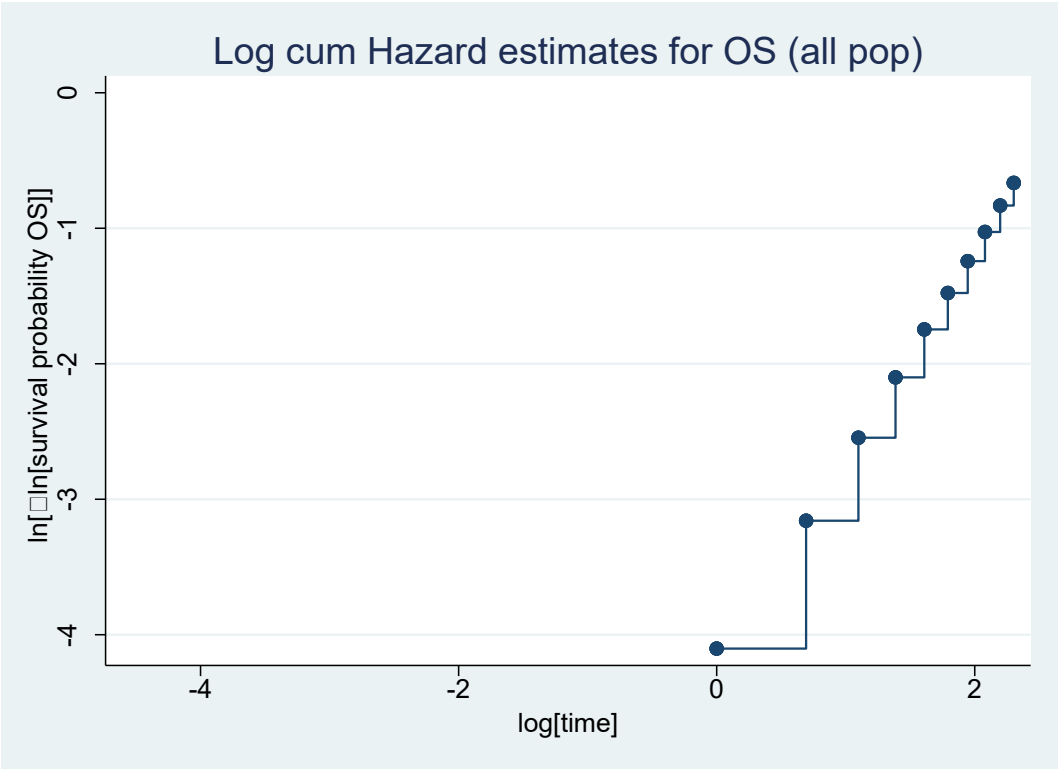
Mean dependent var	5.239	SD dependent var	3.112
Number of obs	5744	Akaike crit. (AIC)	.

*** p<0.01, ** p<0.05, * p<0.1

Akaike's information criterion and Bayesian information criterion Obs	ll(null)	ll(model)	df	AIC	BIC
5,744	.	-4633.644	2	9271.288	9284.6

Note: N=Obs used in calculating BIC; see [R] BIC note.

Figure 6.10. Log cumulative hazard of OS survival probabilities in the overall population.



Source: produced in this thesis with the ICECaP data.

Appendix 5 Survey conducted among steering committee members to validate the assumptions of resources use in the treatment of localised prostate cancer.

A survey was carried out to collect insights about resources used to populate the model. The survey participants were selected through a process facilitated by the project's steering group, which identified clinicians with expertise in the treatment of localised prostate cancer involving experienced clinicians practicing in Australia. This group of clinicians, totalling eight in number, provided their professional perspectives.

To ensure the efficient and distribution and collection of data, an online platform was used. Specifically, the Qualtrics Platform was employed to design and administer the survey. This online approach not only streamlined the data collection process but also provided a convenient means for participants to respond.

It is important to note that the survey was designed with a strict commitment to ethical practices and data privacy. In this regard, no personal or sensitive information was collected from the respondents, aligning with the principles of respect for privacy and confidentiality. Moreover, prior to engaging in the survey, respondents were provided with a clear and transparent overview of the survey's purpose and objectives, as well as the individuals and organizations involved in the collection and use of the data. This transparency was maintained throughout the data collection process, ensuring informed participation from all respondents.

Assumptions of ICECaP economic analysis

Start of Block: Introduction

At the last steering committee meeting where the updates of the ICECaP economic analysis were presented, feedback was provided for model assumptions.

As a follow-up, we would like to seek your feedback regarding a number of treatment related assumptions included in the model. The information collected in this survey will be used to adjust the model parameters used to develop the treatment pathway.

For each of the following questions, please review the components of treatment and provide any variations you think are relevant to clinical practice in Australia.

Thank you for participating in this survey.

End of Block: Introduction

Start of Block: Short term adverse events after primary treatment

The model assumes patients undergoing radiotherapy or radical prostatectomy as primary treatment develop adverse events. The following are the adverse events assumed to happen within 90 days after treatment and the probability of occurrence.

Within 90 days of treatment

Probabilities	RP
Urinary toxicity*	47/ 100 pts
Erectile dysfunction	77/ 100 pts
Urethral stricture	4/ 100 pts
	RT
Urinary toxicity*	30/ 100 pts
Gastrointestinal toxicity**	18/ 100 pts

*Urinary toxicity includes irritative voiding symptoms and incontinence.

**Gastrointestinal toxicity includes rectal bleeding, pain and diarrhea.

Am I missing any specific adverse event happening within 90 days after primary treatment?

- ○ No (1)
- ○ Yes (2)

Please indicate any missing adverse event happening within 90 days after primary treatment.

	Probability of occurrence (1)
Other (specify) (5)	
Other (specify) (7)	
Other (specify) (8)	
Other (specify) (9)	
Other (specify) (10)	

Do probabilities of occurrence sound reasonable?

- ○ No (1)
- ○ Yes (2)

How would you modify the probability of occurrence of the short term adverse events?

	Probability of occurrence (1)
Urinary toxicity after RP (1)	
Urinary toxicity after RT (2)	
Erectile dysfunction after RP (3)	
Urethral stricture after RP (4)	
Gastrointestinal toxicity after RT (6)	
Other (specify) (5)	

Are there any other variations you would make to the assumptions in the model about short-term adverse events after primary treatment?

_____

_____

_____

_____

_____

End of Block: Short term adverse events after primary treatment

Start of Block: Long term adverse events after primary treatment

Long term adverse events are assumed to happen post 90 days of treatment. The following is the list of adverse events assumed to happen in the long term after primary treatment and the probability of occurrence.

Post 90 days of treatment	Probabilities
RP	
Urinary toxicity*	13/ 100 pts
Erectile dysfunction	46/ 100 pts
RT	
Urinary toxicity*	4/ 100 pts
Gastrointestinal toxicity**	2/ 100 pts
Erectile dysfunction	7/ 100 pts

*Urinary toxicity includes irritative voiding symptoms and incontinence.

**Gastrointestinal toxicity includes rectal bleeding, pain and diarrhea.

Am I missing any specific adverse event happening post 90 days of primary treatment?

- ○ No (1)
- ○ Yes (2)

How would you modify the long term adverse events?

	Probability of occurrence (1)
Other (specify) (6)	
Other (specify) (5)	
Other (specify) (7)	
Other (specify) (8)	

Do probabilities of occurrence sound reasonable?

- ○ No (1)
- ○ Yes (2)

How would you modify the long term adverse events?

	Probability of occurrence (1)
Urinary toxicity after RP (1)	
Urinary toxicity after RT (2)	
Erectile dysfunction after RP (3)	
Erectile dysfunction after RT (4)	
Gastrointestinal toxicity after RT (6)	
Other (specify) (5)	

Are there any other variations you would make to the assumptions in the model about long term adverse events after primary treatment?

_____

_____

_____

_____

_____

End of Block: Long term adverse events after primary treatment

Start of Block: Treatment of adverse events: urinary toxicity

The model assumes that short-term treatment of urinary toxicity requires the following regardless of the primary treatment received:

Urinary toxicity	Dose	Quantity	Proportion of Men Treated
Professional consultation	3 in 3 months	3	100%
GP visits	3 in 3 months	5	100%
Urine flow study		1	100%
Paracetamol 500 mg	2tbl/8hr/3 months	540	100%
Ural	1scht/8hrs/3 months	270	100%

Are these treatments according to Australian settings?

No (1)

○ Yes (2)

How would you modify the short-term treatment of urinary toxicity?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Professional consultation (1)			
GP visits (2)			
Urine flow study (3)			
Paracetamol 500 mg (4)			
Ural (5)			
Other (specify) (6)			
Other (specify) (7)			
Other (specify) (8)			
Other (specify) (9)			

Are there any other variations you would make to the assumptions in the model about treatment of short-term urinary toxicity?

_____

_____

_____

_____

_____

The model assumes that long-term treatment of urinary toxicity requires the following regardless of the primary treatment received:

Description	Dose	Quantity	Proportion of Men Treated
Professional consultation	3 in 12 months	3	100%
GP visits	5 in 12 months	5	100%
Blood test		2	100%
Urine flow study		2	100%
Incontinence pads	2/day	730	11%
Incontinence pads	4/day	1460	2%
Paracetamol + codeine	1tbl/12hr/6months	365	100%
Ural (3 sachets a day)	1scht/8hrs/year	1095	100%

Are these treatments according to Australian settings?

- No (1)
- Yes (2)

How would you modify the long-term treatment of urinary toxicity?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Professional consultation (1)			
GP visits (2)			
Blood test (6)			
Urine flow study (7)			
Incontinence pads 2/day (8)			
Incontinence pads 4/day (9)			
Paracetamol 500 mg + codeine 30 mg tablet (10)			
Ural (3 sachets a day) (11)			

Other (specify) (3)			
Other (specify) (4)			
Other (specify) (5)			

Are there any other variations you would make to the assumptions in the model about treatment of long-term urinary toxicity?

_____

_____

_____

_____

_____

End of Block: Treatment of adverse events: urinary toxicity

Start of Block: Treatment of AE: gastrointestinal toxicity

The model assumes that short-term treatment of gastrointestinal toxicity requires the following regardless of the primary treatment received:

Description	Dose	Quantity	Proportion of Men Treated

Hospital admission for rectal bleeding		1		30%	Professional consultation 3 in 3 months 3
			100%		
GP visits	3 in 3 months	2		100%	
Blood test		2		100%	
Oinment: zinc oxide+ peru balsam		3		100%	Suppository: zinc + peru balsam 2/day/3 months 180
			100%		
Paracetamol + codeine 1tbl/12hrs/3m		180		100%	Loperamide 2 mg capsule 1cap/12hrs/3m 180
			100%		

Are these treatments according to Australian settings?

- ○ No (1)
- ○ Yes (2)

How would you modify the short-term gastrointestinal toxicity treatment pattern?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Hospital admission for rectal bleeding (3)			
Professional consultation (1)			
GP visits (2)			
Blood test (5)			
Oinment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg (9)			
Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg (10)			
paracetamol 500 mg + codeine 30 mg tablet (4)			

Loperamide hydrochloride 2 mg capsule (11)			
Other (specify) (6)			
Other (specify) (7)			
Other (specify) (8)			

Are there any other variations you would make to the assumptions in the model about treatment of short-term gastrointestinal toxicity?

_____

_____

_____

_____

_____

Page Break

The model assumes that long-term treatment of gastrointestinal toxicity requires the following regardless of the primary treatment received:

Description	Dose	Quantity	Proportion of Men Treated
Professional consultation	3 in 12 months	3	100%
GP visits	5 in 12 months	5	100%
Ointment: zinc + peru balsam		2	100%
Suppository: zinc + peru balsam	4/day/6 weeks	168	100%
Paracetamol 500 mg	2tbl/8hr/3 months	540	100%
Loperamide 2 mg capsule	1caps/12hrs/6months	365	100%

Are these treatments according to Australian settings?

- ○ No (1)
- ○ Yes (2)

How would you modify the long-term treatment of gastrointestinal toxicity?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Professional consultation (1)			
GP visits (2)			
Ointment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg (6)			
Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg (7)			
Paracetamol 500 mg (8)			
Loperamide hydrochloride 2 mg capsule (9)			
Other (specify) (3)			
Other (specify) (4)			

Other (specify) (5)			

Are there any other variations you would make to the assumptions in the model about treatment of long-term gastrointestinal toxicity?

_____
_____
_____
_____
_____

End of Block: Treatment of AE: gastrointestinal toxicity

Start of Block: Treatment of AE: Erectile dysfunction

The model assumes that treatment of erectile dysfunction requires the following regardless of the primary treatment received:

Description	Dose	Quantity	Proportion of Men Treated
Professional consultation	1 in 3 months	1	100%
GP visits	2 in 3 months	2	100%
Sildenafil 50 mg tablet	2/month	24	90%
Alprostadil injection	1/month	12	10%

Are these treatments according to Australian settings?

- ○ No (1)
- ○ Yes (2)

How would you modify the erectile dysfunction treatment pattern?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Hospital admission for rectal bleeding (3)			
Professional consultation (1)			
GP visits (2)			
Sildenafil 50 mg tablet (5)			
alprostadil 10 microgram injection (9)			
Other (specify) (6)			
Other (specify) (7)			
Other (specify) (8)			

Are there any other variations you would make to the assumptions in the model about treatment of erectile dysfunction?

_____

_____

_____

_____

_____

End of Block: Treatment of AE: Erectile dysfunction

Start of Block: Chemotherapy Block

It is assumed that within the metastatic state, there are patients that are hormone responsive (receiving STAMPEDE treatment schedule) and hormone insensitive patients who have failed following previous ADT.

These patients are treated as described below: they start with docetaxcel as primary medication, and of those who failed to respond to this treatment, 18% receive cabacitaxel, 28% abiraterone and 8% enzalutamide as primary medication. For medical castration 57% receive Goserelin and as bone targeted therapy, 13% receive denosumab.

Description	Dose	Quantity	Proportion of Men Treated
Primary medication no QTP			
Enzalutamide 40 mg	160mg/day	1460	8% Abiraterone 500mg 1gr/day 730 28%
Primary medication QTP			
Cabazitaxel 60mg/1.5mL injection	20mg/m2/21days	10	18%
Docetaxel 160mg/16mL	75 mg/m²	18	100%
Prednisolone 5 mg tablet	10mg/day	84	100%
Chemotherapy-induced side effects medication			
Dexamethasone 8mg	8 mg i.v	10	100%
Medication for medical castration in M1CRPC			
Goserelin 10.8 mg implant	1 /12weeks	4.3	57%
Bone-targeted therapy			
Denosumab 120mg/1.7mL	120 mg s.c.	13	13%

Are these treatments according to Australian settings?

- ○ Yes (1)
- ○ No (2)

How would you modify the treatment pattern for chemotherapy for hormone insensitive metastatic prostate cancer?

	Description of Treatment (1)	Dose or Frequency of Use (2)	Proportion of Men Treated (3)
Enzalutamide 40 mg (1)			
Abiraterone 500mg (2)			
Cabazitaxel 60 mg/1.5 mL injection (3)			
Docetaxel 160 mg/16 mL injection, 16 mL vial (4)			
Prednisolone 5 mg tablet (5)			
Dexamethasone 8mg (6)			
Goserelin 10.8 mg implant (7)			
Denosumab 120mg/1.7mL (8)			
Other (specify) (9)			

Other (specify) (10)			
Other (specify) (11)			
Other (specify) (12)			

End of Block: Chemotherapy Block

Start of Block: Block 11

Your Initials:

_____

End of Block: Block 11

Appendix 6. Metanalysis of utilities found in the literature

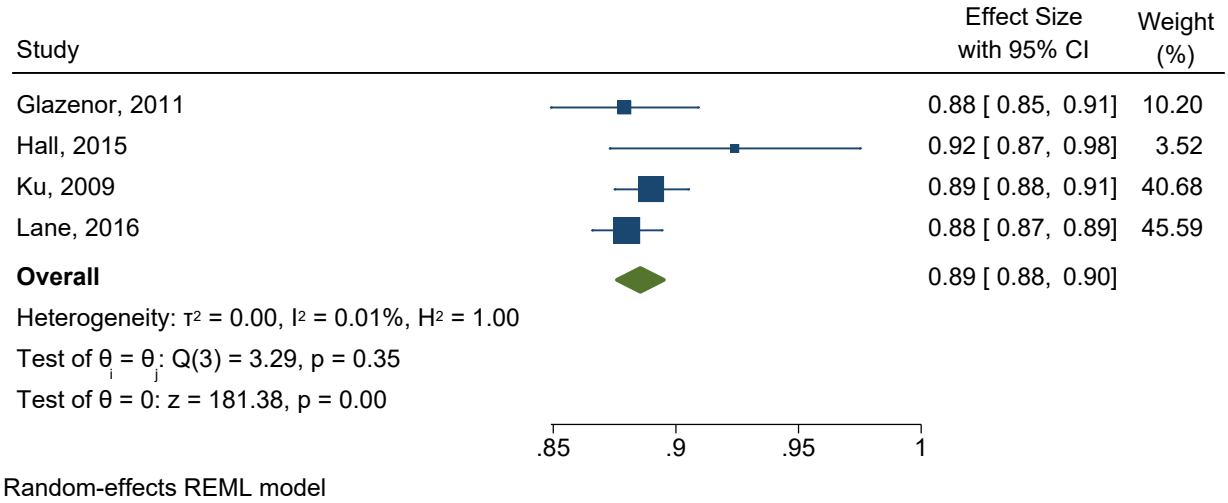
Metanalysis of the utilities values found in the literature for the different health states in the treatment of prostate cancer.

A. Metanalysis results of utilities found for RP

```
Meta-analysis summary                    Number of studies =      4
Random-effects model                     Heterogeneity:
Method: REML                                          tau2 =  0.0000
                                                   I2 (%) =    0.01
                                                       H2 =    1.00
```

Study	Effect Size	[95% Conf.	Interval]	% Weight
Glazenor, 2011	0.879	0.849	0.909	10.20
Hall, 2015	0.924	0.873	0.975	3.52
Ku, 2009	0.890	0.875	0.905	40.68
Lane, 2016	0.880	0.866	0.894	45.59
theta	0.886	0.876	0.895	

```
Test of theta = 0: z = 181.38                Prob > |z| = 0.0000
Test of homogeneity: Q = chi2(3) = 3.29        Prob > Q = 0.3484
```

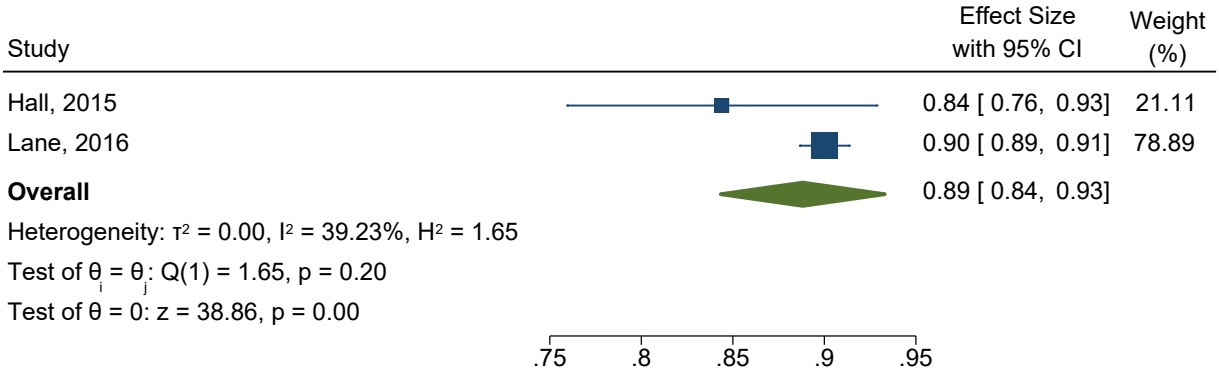


B. Metanalysis results of utilities found for RT

Meta-analysis summary — Number of studies = 2
Random-effects model — Heterogeneity:
Method: REML — tau2 = **0.0006**
I2 (%) = **39.23**
H2 = **1.65**

Study	Effect Size	[95% Conf.	Interval]	% Weight
Hall, 2015	**0.844**	**0.759**	**0.928**	**21.11**
Lane, 2016	**0.900**	**0.887**	**0.913**	**78.89**
theta	**0.888**	**0.843**	**0.933**	

Test of theta = 0: z = **38.86** — Prob > |z| = **0.0000**
Test of homogeneity: Q = chi2(**1**) = **1.65** — Prob > Q = **0.1996**

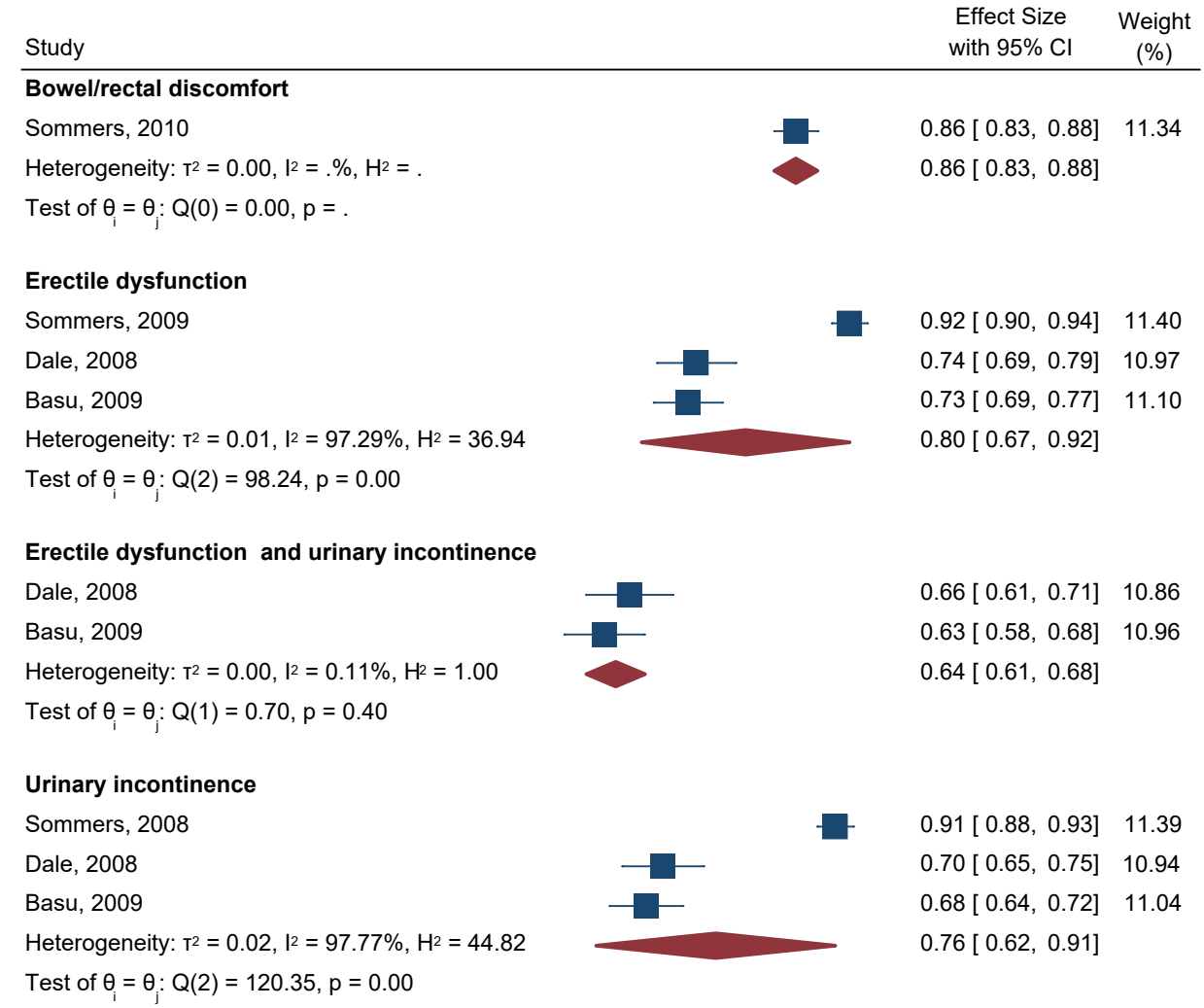


Random-effects REML model

C. Metanalysis results of utilities found for adverse events

Subgroup meta-analysis summary Number of studies = 9
Random-effects model
Method: REML
Group: **Group**

Study	Effect Size	[95% Conf.	Interval]	% Weight
Group: Bowel/rec~				
Sommers, 2010	0.859	0.833	0.885	11.34
theta	0.859	0.833	0.885	
Group: Erectile ~n				
Sommers, 2009	0.923	0.902	0.944	11.40
Dale, 2008	0.740	0.693	0.787	10.97
Basu, 2009	0.730	0.689	0.771	11.10
theta	0.799	0.675	0.923	
Group: Erectile ~r				
Dale, 2008	0.660	0.608	0.712	10.86
Basu, 2009	0.630	0.582	0.678	10.96
theta	0.644	0.609	0.679	
Group: Urinary i~e				
Sommers, 2008	0.906	0.885	0.927	11.39
Dale, 2008	0.700	0.652	0.748	10.94
Basu, 2009	0.680	0.636	0.724	11.04
theta	0.763	0.620	0.906	
Overall				
theta	0.760	0.688	0.833	



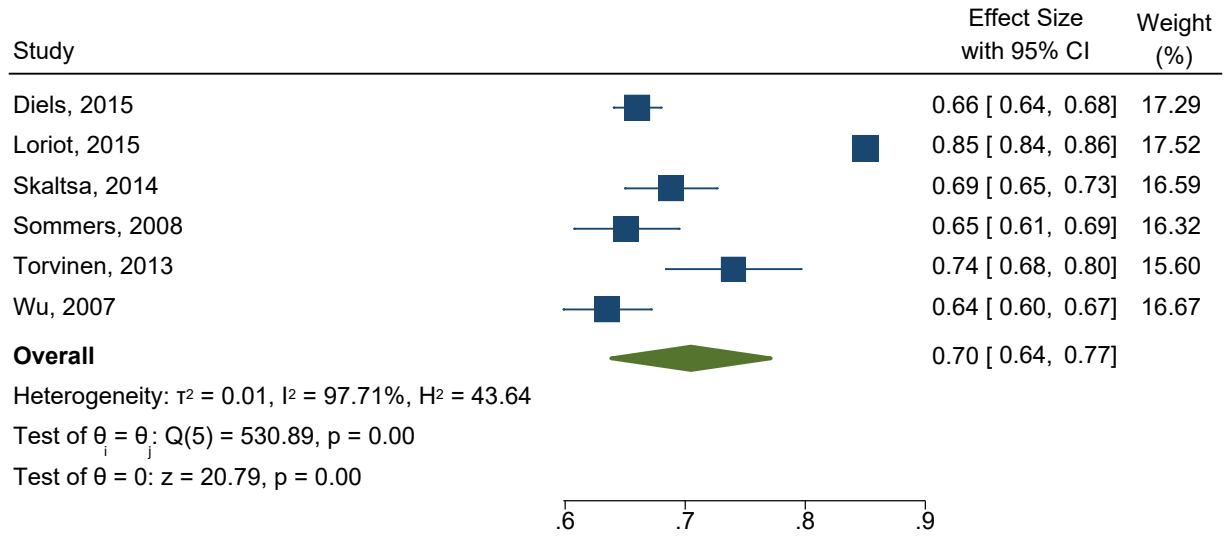
Metanalysis results of all studies reporting utilities in metastatic patients

Meta-analysis summary
Random-effects model
Method: REML

Number of studies = 6
Heterogeneity:
tau2 = 0.0065
I2 (%) = 97.71
H2 = 43.64

Study	Effect Size	[95% Conf.	Interval]	% Weight
Diels, 2015	0.660	0.640	0.680	17.29
Loriot, 2015	0.850	0.843	0.857	17.52
Skaltsa, 2014	0.688	0.650	0.726	16.59
Sommers, 2008	0.651	0.607	0.695	16.32
Torvinen, 2013	0.740	0.684	0.796	15.60
Wu, 2007	0.635	0.599	0.671	16.67
theta	0.705	0.638	0.771	

Test of theta = 0: z = 20.79 Prob > |z| = 0.0000
Test of homogeneity: Q = chi2(5) = 530.89 Prob > Q = 0.0000



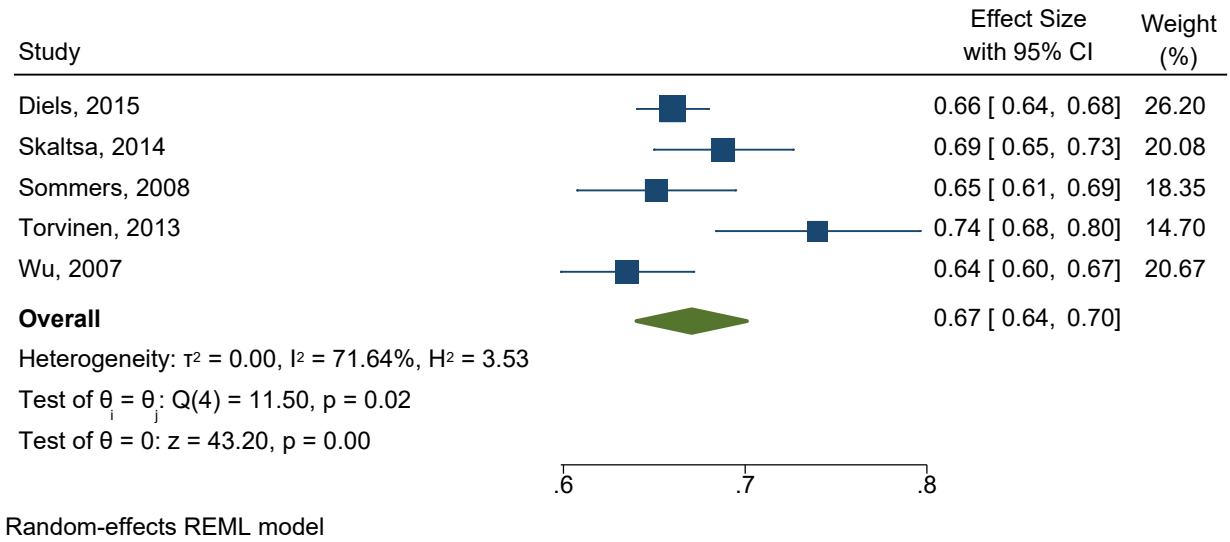
Random-effects REML model

D. Metanalysis results of selected studies reporting utilities in metastatic patients

```
Meta-analysis summary                       Number of studies =      5
Random-effects model                        Heterogeneity:
Method: REML                                            tau2 =  0.0008
                                                      I2 (%) =   71.64
                                                          H2 =    3.53
```

Study	Effect Size	[95% Conf.	Interval]	% Weight
Diels, 2015	0.660	0.640	0.680	26.20
Skaltsa, 2014	0.688	0.650	0.726	20.08
Sommers, 2008	0.651	0.607	0.695	18.35
Torvinen, 2013	0.740	0.684	0.796	14.70
Wu, 2007	0.635	0.599	0.671	20.67
theta	0.671	0.640	0.701	

```
Test of theta = 0: z = 43.20                 Prob > |z| = 0.0000
Test of homogeneity: Q = chi2(4) = 11.50       Prob > Q = 0.0215
```



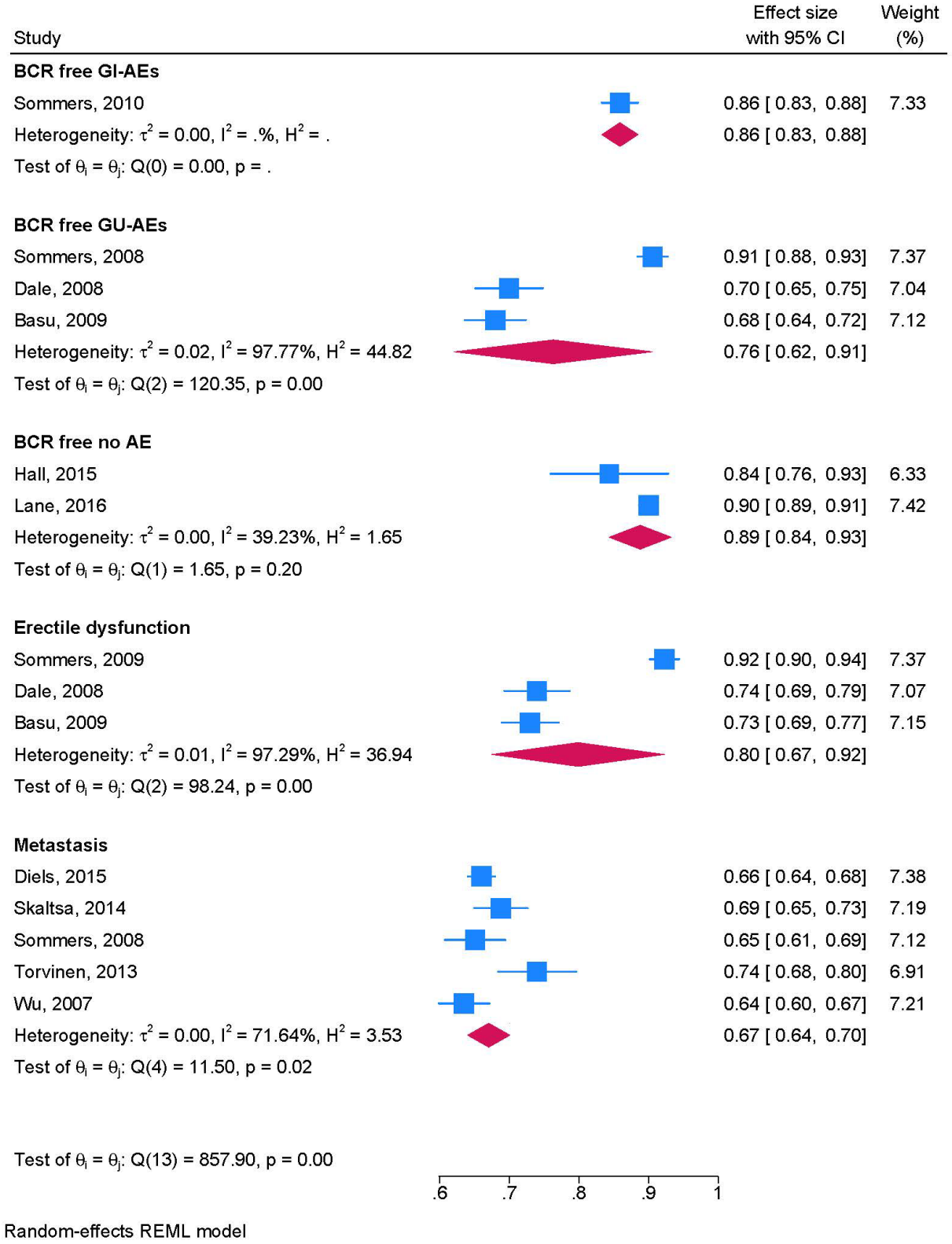
Characteristics of the studies included in the metanalysis.

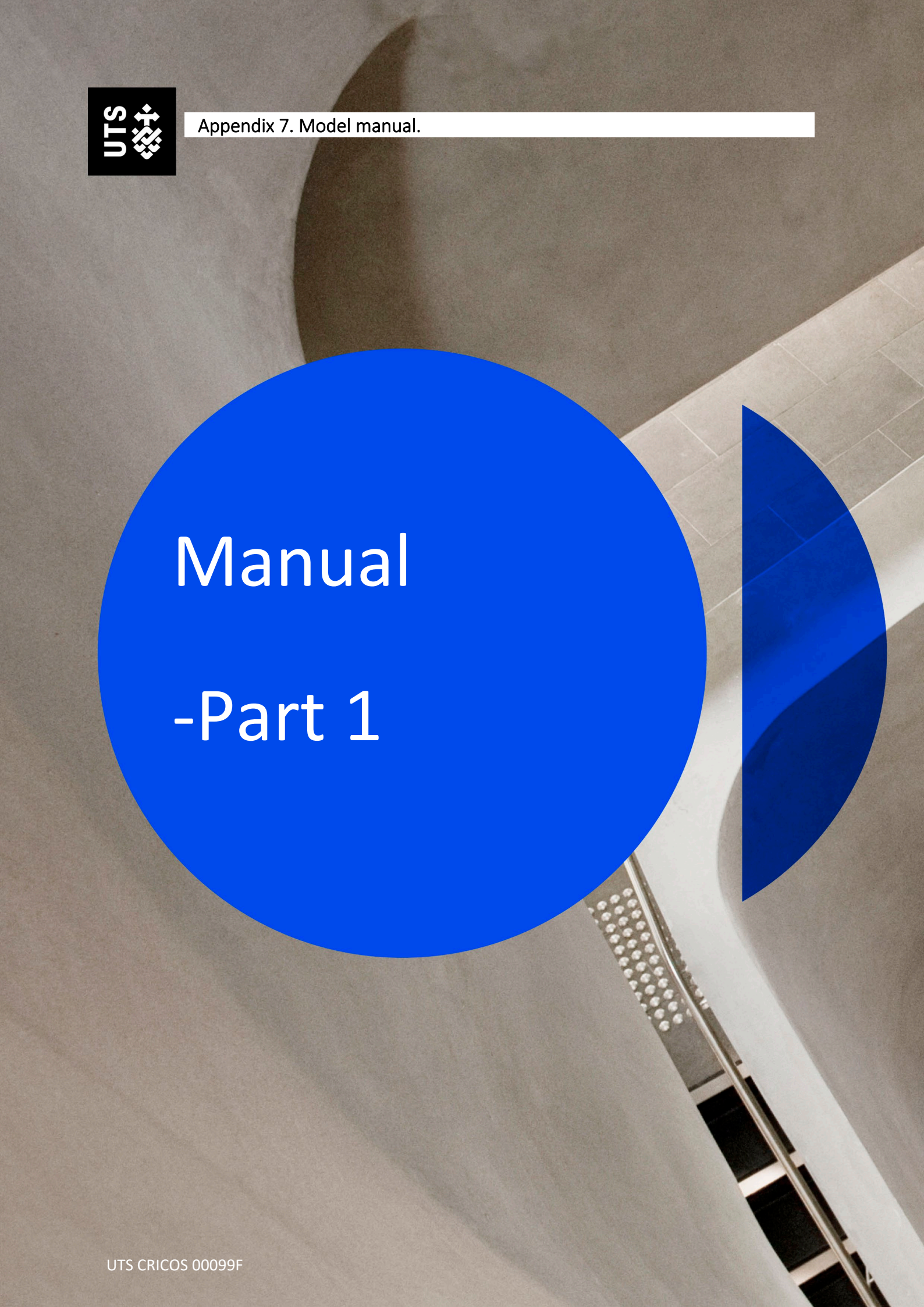
Table 27. Characteristics of the studies included in the metanalysis.

Author	Year	Group	Instrument/Tool	Disease stage	Utility value	Status	Reason for exclusion
Sommers	2008	Urinary incontinence	TTO	localised T1N0M0 or T2N0M0	0.906	I	
Sommers	2009	Erectile dysfunction	TTO	localised T1N0M0 or T2N0M1	0.923	I	
Sommers	2010	Bowel/rectal discomfort	TTO	localised T1N0M0 or T2N0M2	0.859	I	
Dale	2008	Postprostatectomy without complications	TTO	na	0.8	I	
Dale	2008	Erectile dysfunction	TTO	na	0.74	I	
Dale	2008	Urinary incontinence	TTO	na	0.7	I	
Dale	2008	Erectile dysfunction post-prostatectomy	TTO	na	0.72	I	
Dale	2008	Erectile dysfunction and urinary incontinence	TTO	na	0.66	I	
Basu	2009	Erectile dysfunction	TTO	Patient just had abnormal biopsy screening result	0.73	I	
Basu	2009	Urinary incontinence	TTO	Patient just had abnormal biopsy screening result	0.68	I	
Basu	2009	Post-prostatectomy	TTO	Patient just had abnormal biopsy screening result	0.78	I	
Basu	2009	Erectile dysfunction and urinary incontinence	TTO	Patient just had abnormal biopsy screening result	0.63	I	
Basu	2009	Erectile dysfunction post-prostatectomy	TTO	Patient just had abnormal biopsy screening result	0.7	I	
Hall	2015	RT	EQ5D -3L	Gleason scores	0.844	I	
Lane	2016	RT	EQ5D-3L	Localised cancer	0.9	I	
Krahn	2009	RT	SG	T1-T4 . Mostly (75%) T1 and T2	0.92	I	
Glazenor	2011	RP	EQ5D	Post treatment with RP	0.879	I	
Hall	2015	RP	EQ5D -3L	Gleason scores	0.924	I	
Ku	2009	RP	PORPUS	Localised cancer T1 (63%) T2 (37%)	0.89	I	
Lane	2016	RP	EQ5D-3L	Localised cancer	0.88	I	
Krahn	2009	RP	SG	T1-T4 . Mostly (75%) T1 and T2	0.94	I	
Diels	2015	Met	EQ5D-3L	Histologically or cytologically confirmed diagnosis of adenocarcinoma of the prostate	0.66	I	
Loriot	2015	Met	EQ5D	Metastases	0.85	E	Chemotherapy naive patients only
Skaltsa	2014	Met	EQ5D	Stage III or Stage IV	0.688	I	

Sommers	2008	Met	TTO	localised T1N0M0 or T2N0M0	0.651	I
Torvinen	2013	Met	EQ5D-3L	Five mutually exclusive groups : Loc1 <12 mths post diagnosis; Loc2 the following 12 mths from diagnosis; Loc3 Subsequent yrs of remission; Metastatic; Palliative Care.	0.74	I
Wu	2007	Met	EQ5D	Metastatic HRPCa.	0.635	I

Figure 6.11. Forest plot of the pooled utilities values used in the model.





Appendix 7. Model manual.

Manual

-Part 1

Table of Contents

1 Manual overview 312

1.1 Context 312

1.2 The model 312

1.3 Workbook structure 313

2 Population 317

2.1 Time Horizon and discount rate 318

2.2 Perspective 319

3 Economic model 320

3.1 Structure of the economic model: partitioned survival model 320

3.2 Clinical input 323

3.2.1 Overall survival 326

3.2.2 Metastasis free survival 328

3.2.3 Biochemical recurrence free survival 330

3.2.4 General population mortality 331

3.3 The surrogacy relationship 332

3.4 Economic inputs 335

3.4.1 Primary treatment 340

3.4.2 Subsequent Treatment 341

3.4.3 Adverse events costs 343

3.4.4 Follow up 345

3.4.5 Other costs 345

3.5 Quality of life inputs 346

3.6 Model outputs 348

3.6.1 Benefits to society 349

3.7 Sensitivity analysis 350

3.7.1 Deterministic sensitivity analysis 351

3.7.2 Probabilistic sensitivity analysis 352

Attachment 1 353

Attachment 2 356

Appendix 7 Bibliography 365

Index of figures and tables

Figures

Figure 1. Population included in the model. 317
Figure 2. State membership for the overall population in ICECaP. 322
Figure 3. Survival probabilities of the o outcomes of interest for the overall population in ICECaP. ... 323
Figure 4.Overall survival from ICECaP and corresponding extrapolation. 327
Figure 5. Metastasis free survival from ICECaP and corresponding extrapolation. 329
Figure 6.Biochemical failure free survival from ICECaP and corresponding extrapolation. 331

Tables

Table 1. Outline of the workbook. 315
Table 2. Location of all economic inputs in the model. 337
Table 3. Pooled estimates of utility values used in the model. 346
Table 4. Follow-up cost used in the model. 356

1 Manual overview

1.1 Context

In the context of budgetary restrictions, economic evaluations within health technology assessments (HTA) have become a commonly used tool to estimate efficiency. A critical component of these evaluations is the outcome chosen in the clinical trial to measure treatment effect and how this outcome is used in economic evaluations and its impacts the decision-making process timeframe. as they often determine its sample size, duration, cost, and feasibility.

In Australia, prostate cancer is the second most prevalent cancer diagnosed in men with nearly twenty thousand new cases per year, and most of them present localised prostate cancer (LPC). Clinical trials in these patients are usually challenging due to the long natural history that characterise localised disease requiring a median follow-up of 11.5 years to show statistically significant differences when overall survival (OS). The longer we wait to make a product available through public subsidy, potentially, the greater the cost, morbidity and mortality associated with later stages of the disease. In this setting, surrogate outcomes may play an important role, by replacing OS with a surrogate outcome to support an earlier public subsidy decision-making process.

The ICECaP study demonstrated that metastasis free survival is a strong surrogate measure of OS in LPC. The use of metastasis free survival (MFS) in an economic evaluation as the basis for public funding decisions could potentially result in earlier access to new medicines.

The following document provides information about the construction and functioning of an economic model to compare primary treatments in LPC where a surrogate outcome (metastases free survival) is the primary clinical outcome.

The model is designed to allow the user to alter the intervention, comparator or country inputs as required. A four-state (biochemical recurrence free, biochemical recurrence, metastasis and death) partitioned survival model (PSM) is employed, in which disease progression and survival are modelled separately. As a PSM, the model uses projections of overall survival (OS), recurrence free survival (BCR-free) and metastasis free survival (MFS) to estimate the state membership and the time an average patient spends in each health state to estimate treatment costs and outcomes over the time horizon of the model. In the model, patients with LPC remain in BCR-free after initial treatment until they experience progression (based on elevated PSA), develop metastases, or die (Figure 4.6).

1.2 The model

PSM modelling has been widely employed in oncology and has been reviewed and accepted by many health technology assessment agencies (1-7).

The model described in this manual differs from other PSMs in the way OS is estimated. Traditionally, OS curves used in PSM come directly from the observed clinical data and are later “partitioned” into different health states according to the natural history of the disease. Within this model in this manual, OS relies on the treatment effect as defined by the ICECaP[38] initiative which has been determined from the hazard ratios (HR) for MFS. Therefore, the OS curve is explicitly dependant on the MFS curve through a regression equation:

[38] ICECaP goal is the creation of a repository of individual patient data from intermediate and high-risk localized prostate cancer trials to enable the credentialing of earlier clinical endpoints as surrogates for overall survival and clinical benefit to expedite clinical trial conduct and identify which patients benefit from a given therapy. https://icecap.movember.com/publications/

Log(HR)-OS= -0.0056 + 0.7813* Log(HR)-MFS

Patients with LPC have different options for managing their diseases; this model is based on patients undergoing active treatment. The economic and quality of life impact of adverse events (AEs) is also accounted for in the model, based on the relevant survival state and treatment received. The model is structured to allow the user to separate analyses for specific patient groups based on age and risk status.

Health outcomes in the model include time spent in each health state and discounted quality-adjusted life years (QALY). Incremental analysis is presented summarising the differences in costs and health benefits between competing treatments, and these comparisons are summarised as incremental cost-effectiveness ratios.

For exposition purposes, the model has been populated with Australian costs and population inputs and takes the perspective of Australian third-party payers or society. This manual outlines how those values can be updated to allow the model to be executed for any jurisdiction.

Finally, this version of the manual is focused on the construction and operation of the PSM model. A subsequent version will utilise the same data sources to detail the operation of a Markov model. With that forthcoming version, model users will have the capacity to change the modelling approach between PSM or Markov.

1.3 Workbook structure

The model is provided as an Excel workbook, subdivided into tabs that addresses specific topics of the model. For easy navigation, tab names are colour coded according to their content:

- Grey: Administrative information.
- Blue: Input information.
- Orange: Model traces.
- Green: Results of the model and sensitivity analyses.

The topics addressed by each tab of the workbook are described in

Table 1.

Table 1. Outline of the workbook.

Tab name	Characteristics
Population	Epidemiologic characteristics of the cohort being modelled
Outline	Model structure and layout of the workbook
Cost	List of all resources and their corresponding disaggregated cost used in the model.
PSM Comparator	Model traces for the comparator
PSM Intervention	Model traces for the intervention
Dashboard	Visualisation panel for results and adjustments to the model
Deterministic SA	Deterministic sensitivity analysis
Probabilistic SA	Probabilistic sensitivity analysis
Results	Results of the model
Benefits to society	Analysis of the benefits to society
Input	List of all parameters used by the model
OC of death	Probabilities of death by other causes in general population of the modelled cohort.
BCR all pop	Survival analysis of BCR in all population
BCR lowRisk	Survival analysis of BCR in low-risk population
BCR HighRisk	Survival analysis of BCR in high-risk population
Age64 BCR	Survival analysis of BCR in patients aged 64 years or younger
Age65-74 BCR	Survival analysis of BCR in patients between age 65 and 74 years
Age75 BCR	Survival analysis of BCR in patients aged 75 years or older
MFS all pop	Survival analysis of MFS in all population
MFS lowRisk	Survival analysis of MFS in low-risk population
MFS HighRisk	Survival analysis of MFS in high-risk population
Age64 MFS	Survival analysis of MFS in patients aged 64 years or younger
Age65-74 MFS	Survival analysis of MFS in patients between age 65 and 74 years
Age75 MFS	Survival analysis of MFS in patients aged 75 years or older
OS all pop	Survival analysis of OS in all population
OS lowRisk	Survival analysis of OS in low-risk population
OS HighRisk	Survival analysis of OS in high-risk population
Age64 OS	Survival analysis of OS in patients aged 64 years or younger
Age65-74 OS	Survival analysis of OS in patients between age 65 and 74 years

Age75 OS	Survival analysis of OS in patients aged 75 years or older

BCR: biochemical recurrence; MFS: metastasis free survival; OS: overall survival.

2 Population

Cells C4 to C9 in the "Dashboard" tab contain the choice (age and risk characteristics) of the population to be used in the model.

In the "Population" tab:

- Cells B13:D18 contain the epidemiologic characteristics of the population in the model.
- Cells B24:C26 contain the proportion of patients receiving the intervention and comparator according to their risk.

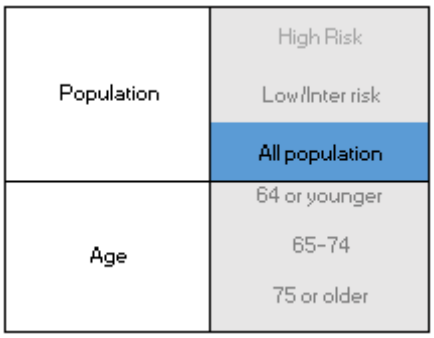
The model includes the analysis populations of patients defined under different categories as per the ICECaP study:

- Risk: Low to intermediate or high risk.
- Age: 64 years or younger, between 65 and 74, and 75 years or older.

Patients are defined as high risk if they had one of the following: high risk by the National Comprehensive Cancer Network (NCCN) or D'Amico criteria, high risk by pathological criteria (pathological Gleason>7, or seminal vesicles involvement, or ≥pT3b stage) or Pathology N1 (ECOG3886 only).

Cells C4 to C9 in the "Dashboard" tab allows the user to analyse all categories included (All population) or choose to run the model for each population separately (Figure 1). When "All population" is selected, the model assumes the cohort of patients being modelled starts at 65 years of age.

Figure 1. Population included in the model.



The methods for estimating clinical outcomes for each subgroup population and their location is described in Section 3.2.

Cells B24:C26 in the "Population" tab can be updated with the proportion of patients receiving the intervention and comparator according to their risk. The latter, along with the epidemiologic characteristics of the population, can be updated in cells B13:D18 in the "Population" tab and is used to estimate the total benefits to society arising from the use of the surrogacy relationship in the decision-making process (see Section 3.6.1 for more information on how benefits to society are estimated in the model).

2.1 Time Horizon and discount rate

In the "Dashboard" tab:

- Cell C44 contains the discount rate (used for health benefits and cost).
- Cell C46 contains the time horizon in years to be used in the analysis.

By default, the model uses a lifetime horizon with a maximum age of 100 years, after which 100% of the starting cohort is assumed dead. As the starting age of the population is a variable that can be modified in the model (depending on the age of the cohort as explained above), the number of cycles varies according to the starting age of the population.

The model evaluates patient's outcomes and accrued cost on a yearly basis. By applying a "half-cycle" correction, the density of patients within each cycle is adjusted to an approximate mid-period value before applying cost and utility payoffs.

A discount rate of 5% is used to discount future costs and health outcomes (QALYs) in the base case. The user can vary the time horizon in years to be used in the analysis using cell C46, while the discount rate can be varied in cell C44.

2.2 Perspective

In the “Dashboard” tab, cells C15:C16 contain the choice of perspective.

The model allows the user to choose between the health care system or the societal perspective. When the health care system perspective is used, only costs incurred by the health care system to make treatments and services available for patients are considered. In practical terms, over-the-counter drugs and drugs not listed in public formularies (the PBS in the case of Australia) are excluded as they have no direct impact on the financial cost to the health care system. The societal perspective includes out-of-pocket costs of treatments and/or medical services (usually in the form of co-payments).

Both perspectives can consider direct and indirect costs. For direct costs, three main cost categories were identified: diagnosis, treatment, and monitoring of the disease (see Section 3.4.4).

By changing cells C15:16 in the “Dashboard” tab, the model allows for the societal or the health care system perspective to be used in the analysis.

3 Economic model

3.1 Structure of the economic model: partitioned survival model

The structure of the model can be found in the “Outline” tab.

In a PSM, individual patient data are commonly used in a survival analysis of the outcomes of interest to determine the proportion of patients at each health state at all times. For example, the proportion of patients alive is calculated as the area under the OS curve. This curve is later disaggregated or “partitioned” into different health states that coexist. As they are non-mutually exclusive, state membership must be derived from the survival analysis rather than through transition probability matrices like in a state transition model. PSMs assign values of cost and benefits to health states, which together with the time spent in each state determines the total cost and benefit of the intervention and comparator.

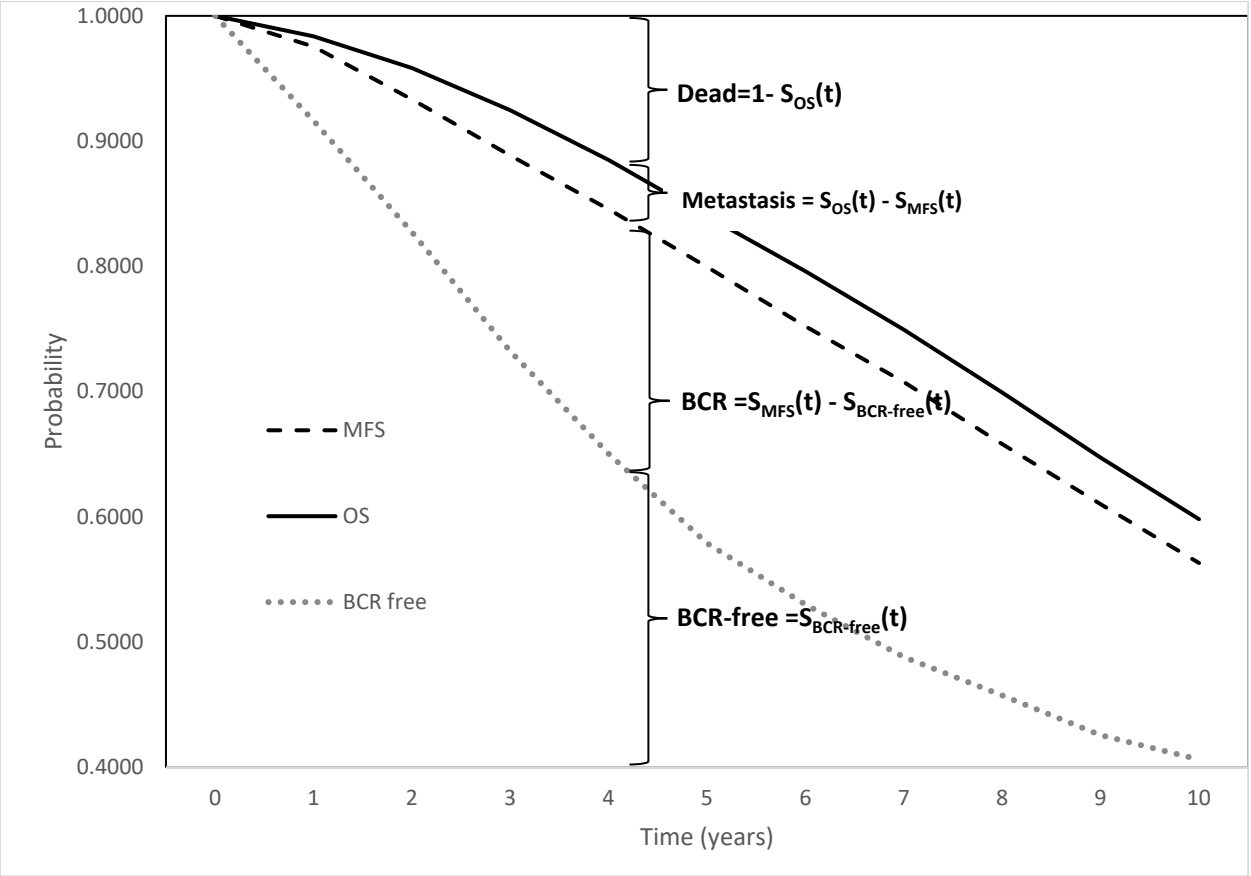
Physicians treating patients experiencing BCR must make challenging decisions, whether to treat with a view to delaying metastases and its potential early disease mortality at a risk of treatment adverse events. impacting quality of life, or not to treat a condition that may never really impact the wellbeing of the patient. Androgen deprivation therapy (ADT) impacts negatively the quality of life of patients who have failed initial treatment. Therefore, a BCR health state was also considered relevant to reflect the natural history of the disease.

To represent the course of a disease in a PSM of n-states, n-1 survival curves are required where the cumulative survival function S(t), represents the probability of survival up to time t. State membership in each alive health state (n=2, …, (n−1)) can be calculated as the subtraction of Sn(t) minus the previous health state S(n−1) (t). The probability of being in the absorbing or final health state is calculated as 1- SOS(t), or the area above the OS survival curve. Therefore, to depict how LPC patient transit through their diseases, the model considers four heath states:

- BCR-free: where patients start their treatment and remain after primary treatment until progression.
- BCR: where patient transit after initial treatment failure after a PSA elevation with no signs of metastasis.
- Metastasis: Where patients can transit after initial treatment or BCR treatment failure and receive treatment for their progressed state.
- Death: All causes of death. Absorbing state.

State membership of a PSM can be represented graphically as the difference between the respective survival curves. Figure 2 shows the state membership of the ICECaP study in a four-health state model (BCR-free, BCR, Metastasis and Death) for the overall population using OS, MFS and BCR-free survival curves. The model was built upon a case study comparing SoC (ADT+ radiotherapy) with abiraterone acetate plus prednisolone (AAP) + SoC. The survival curves from the comparator arm were sourced from the ICECaP high-risk patients whereas the treatment effect for BCR-free and MFS was based on HR and sourced from the STAMPEDE metanalysis [9].

Figure 2. State membership for the overall population in ICECaP.



Abbreviations: BCR, Biochemical recurrence; MFS, Metastasis free survival; OS, Overall survival.

The time unit used in the model to reflect all costs and benefits is called a cycle and this model uses a cycle length of 1 year.

In the "Dashboard" tab, cells

- C48:C49 contain the HRs for MFS and BCR-free.
- C21:C23 contain the data cut-off times for each outcome modelled.

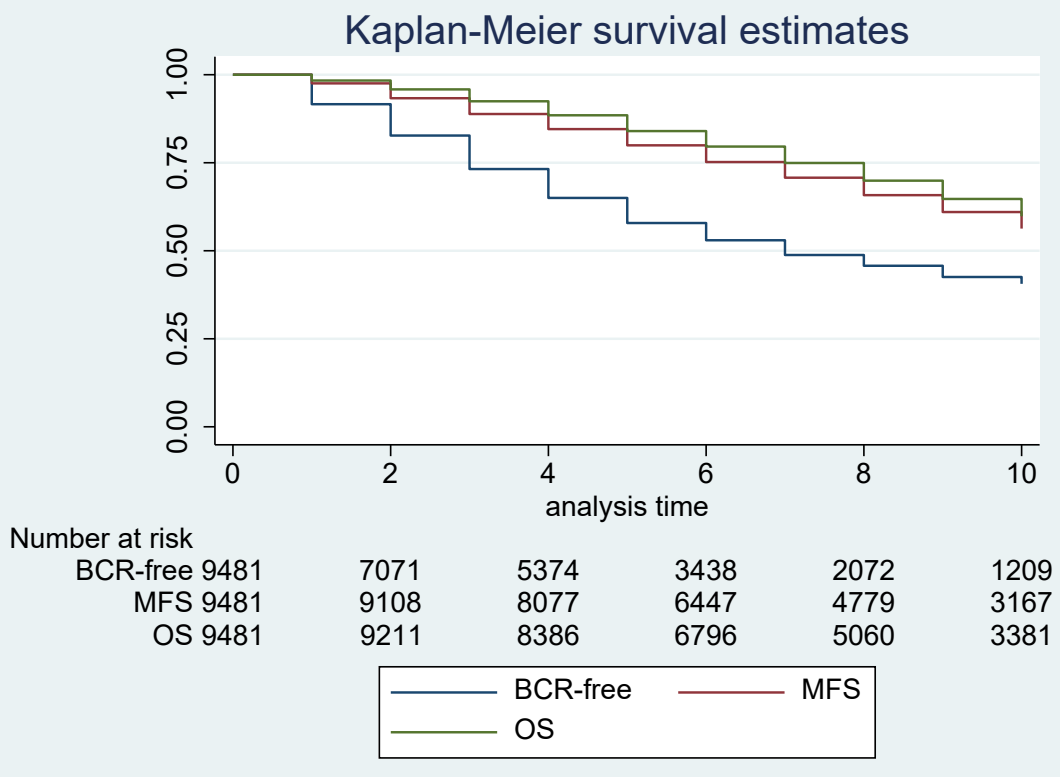
Parametrised curves can be found in the model in the tabs which name starts with an outcome abbreviation followed by the population under analysis (e.g. BCR lowRisk is for patients with low-risk in the BCR category). Cells N2:S4 in each respective tab contain the parameters of all distributions used in the survival analysis.

3.2 Clinical input

The treatment effect used in the model for the comparator was sourced from the ICECaP analysis (ADT plus radiotherapy), with a proportional hazards assumption applied to estimate the treatment effect for the intervention (AAP+ SoC). The comparator as well as intervention to be compared to in the model can be modified by the user as required.

For the purpose of this manual, the survival probabilities of BCR-free, MFS, and OS in the 10 years follow-up period after primary treatment with radiotherapy (RT) from the ICECaP data are used as the basis of the PSM. The outcomes were differentiated according to initial risk assessment (low/intermediate or high risk) and age (64 or younger, 65 to 74 and 75 or older). Figure 3 shows the survival probabilities for MFS, BCR-free and OS for the overall population.

Figure 3. Survival probabilities of the o outcomes of interest for the overall population in ICECaP.



BCR: Biochemical failure; MFS: Metastasis free survival; OS: Overall survival; ADT: Androgen deprivation therapy.

The choice of parametric function applied to all time to event data was hierarchical to OS. That is, the parametric model for OS was first assessed, then parametric models for MFS and BCR-free were guided by the curve which produced values that were below the best fitting OS curve. This was to avoid curves crossing among different populations and between outcomes evaluated for each population. All parametric models were evaluated for all subsets of the population in ICECaP (64 or younger, 65 to 74 years of age, 75 or older, high risk and low/intermediate risk).

Visual inspection of the log cumulative hazard plots for BCR-free, MFS and OS, showed a monotone hazard function as well as a consistent shape and gradient during the follow-up time, therefore, conventional parametric distribution models were evaluated for all outcomes of interest in the ICECaP data.

The survival analysis was conducted in STATA® version 16, using exponential, Weibull, generalised gamma, gompertz, log-logistic and log-normal distributions models to determine the best fit of the ICECaP data. The outputs of the survival analysis in STATA® (in terms of the shape and scale parameters) were used in Excel® to extrapolate all the outcomes beyond the follow-up time (10 years) observed in the ICECaP analysis.

As the extrapolations are conducted in Excel® based on the parameters of the most commonly used probability distributions, it is possible to adapt the survival curves of each outcome to other treatments under comparison by changing the parameters determining the curves. By changing cells N2:S4 in each of the tabs storing the survival analysis, the user can specify different parameters for the distributions used to extrapolate the data. Also, the models can incorporate different data cut-off times depending on the length of follow-up. By default, this time is fixed at 10 years as it corresponds to the analysis time of the ICECaP data.

In addition, as outcomes modelled in a PSMs are structurally independent, curves may cross in the extrapolation period (e.g., where the hazard of progressed stages of the disease is higher than the hazard of death). To overcome the occurrence of such inconsistent outcomes, adjustment of the curves is usually conducted. Within the model, restrictions are applied such

that if the probability of extrapolated MFS or PFS is greater than OS, it is assumed to be the same as OS (i.e., MFS=OS). When the adjustment is needed, survival probabilities in column I, H and J in “PSM comparator” and “PSM intervention” arm will be highlighted orange from the time where the adjustment begins.

The parametrised curves can be found in the model in the tabs with a name which starts with an outcome abbreviation followed by the population (e.g., “BCR all pop”) and contain:

- the ICECaP data for the population analysed used to extrapolate, the six curves,
- the parameters tested together with their regression outputs from STATA, and,
- the corresponding curves calculated for the “Intervention” based on proportional hazards assumption.

In the base case analysis, OS and MFS are extrapolated using a Weibull function whereas BCR-free is extrapolated by means of a log-logistic distribution. The choice of extrapolation model can be selected from the “Dashboard” tab for each outcome independently.

Note

- As the cycle length of the ICECaP analysis was 1 year, the survival analysis and the model use 1-year cycles. If the data to be used in the model are structured in other than one yearly cycles, modify the output of the statistical software and convert their source shape and scale parameters to refer to cycles of 1 year (12 months), so the Kaplan Meier function and parameters of the extrapolation relate to the cycle length used in the model. The cycle length in the model is a fixed variable that is not possible to change.

3.2.1 Overall survival

The survival probabilities used in the model for each subgroup can be find in tabs: OS all pop, OS lowRisk, OS HighRisk, Age64 OS, Age65-74 OS and Age75 OS. In these tabs:

- Cells N1:S4 contain the parameters used to extrapolate the comparator.
- Cells B24:I34 contain the Kaplan Meier survival probabilities output for the comparator.
- Cells K49:K85 contain the survival probability for the intervention based on proportional hazards.

In tab "Dashboard":

- Cells C37:C42 allow the user to change the distribution chosen to extrapolate.
- Cell C21 allow the user to change the time the model use the Kaplan Meier results from the study.
- Cell C48 contain the MFS-HR.

In this case study, the log-logistic distribution was chosen as it produced a survival curve that was consistent with clinical expectations of the natural history of the condition among all populations analysed (see Figure 3.5). To change the choice of distribution function for the extrapolation of OS, the user can choose from among the six most used distributions in cells C37:C42 in tab "Dashboard".

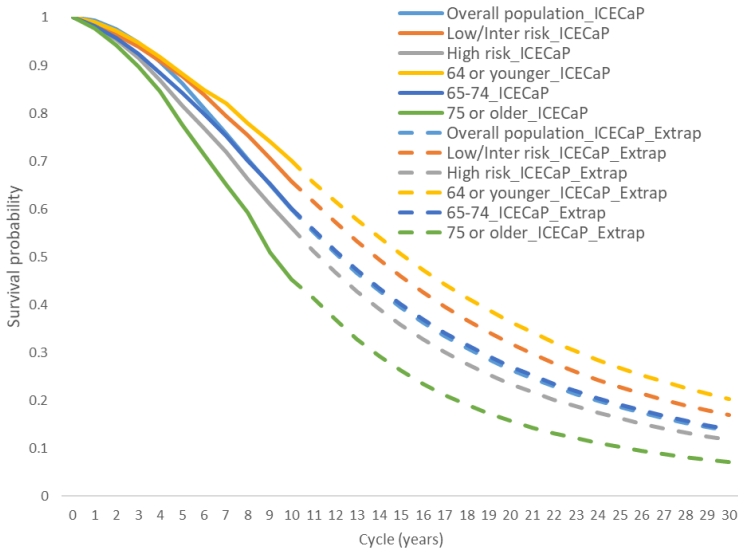
If new data about the comparator or intervention are input into the model for OS, cells F24:F34 of all tabs of OS survival analysis containing the survival probabilities at 1 year cycle needs to be updated as per Section 3.2. The resulting survival probabilities will be used in the model to depict the comparator arm and to estimate the survival probabilities for the intervention based on proportional hazards. Specifically, to estimate the survival probabilities of OS for the intervention, the survival probabilities of OS in the comparator arm are elevated to the potency of the value of the OS-HR estimated through the surrogacy relationship (see Attachment 1 for more information about the use of HRs to estimate survival probabilities).

This model is based on the surrogacy relationship to estimate OS-HR from 5-year MFS; cell C48 in the “Dashboard” tab contains the MFS-HR. OS-HR is estimated by the model and is a non-modifiable cell. Section 3.3 explains the methodology and how to use the surrogacy relationship in the model.

Note:

The model applies the first 10 years of the Kaplan Meier data from ICECaP and then the extrapolated data. If less than 10 years data are available for the model, input the survival probabilities on cells F24:F34 in any tab where the survival analysis of OS is located and adjust cell C21 in the “Dashboard” tab with the corresponding data cut-off time.

Figure 4.Overall survival from ICECaP and corresponding extrapolation.



3.2.2 Metastasis free survival

The survival probabilities used in the model for each subgroup can be found in tabs: MFS all pop, MFS lowRisk, MFS HighRisk, Age64 MFS, Age65-74 MFS and Age75 MFS. In these tabs:

- Cells N1:S4 contain the parameters used to extrapolate the comparator.
- Cells B24:I34 contain the Kaplan Meier survival probabilities output for the comparator.
- Cells K49:K85 contain the survival probability for the intervention based on proportional hazards.

In tab "Dashboard":

- Cells C25:C30 allow the user to change the distribution chosen to extrapolate.
- Cell C22 allows the user to change the timepoint for which the model uses the Kaplan Meier results from the study.
- Cell C48 contains the MFS-HR.

In this case study, a Weibull distribution showed clinically acceptable probabilities and comparable shapes among all populations analysed and therefore was chosen to extrapolate MFS in the model. Nevertheless, log-normal and Weibull were the best fit for 64 or younger and 75 or older, respectively. Figure 3.7 shows the ICECaP data and the corresponding extrapolation for MFS in all populations analysed. To change the choice of distribution function for the extrapolation of MFS, user can choose from among the six most used distributions from cells C25:C30 in the "Dashboard" tab.

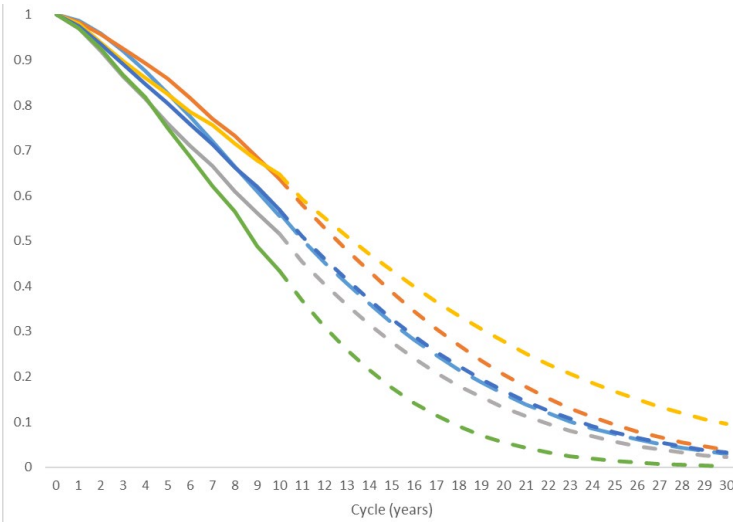
If new data for the comparator and intervention are to be used in the model for MFS, cells F24:F34 of all tabs with MFS survival analysis with the survival probabilities at 1 year cycle will need to be updated as described in Section 3.2. As with OS, the survival probabilities will be used in the model to inform the comparator arm and to estimate the survival probabilities for the intervention based on proportional hazards. Cell C48 in the "Dashboard" tab contains the MFS-HR; to estimate the survival probabilities of MFS for the intervention, the survival

probabilities of MFS in the comparator arm are elevated to the potency of the value of the MFS-HR.

Note:

The model uses the first 10 years of the Kaplan Meier curve and then the extrapolated data. If less than 10 years of data are available for the model, input the survival probabilities in cells F24:F34 in any tab where the survival analysis of MFS is located and adjust cell C22 with the corresponding data cut-off time.

Figure 5. Metastasis free survival from ICECaP and corresponding extrapolation.



3.2.3 Biochemical recurrence free survival

The survival probabilities used in the model for each subgroup can be found in tabs: BCR-free all pop, BCR-free lowRisk, BCR-free HighRisk, Age64 BCR-free, Age65-74 BCR-free and Age75 BCR-free. In these tabs:

- Cells N1:S4 contain the parameters used to extrapolate the comparator.
- Cells B24:I34 contain the Kaplan Meier survival probabilities output for the comparator.
- Cells K49:K85 contain the survival probability for the intervention based on proportional hazards.

In tab "Dashboard":

- Cells C31:C36 allow the user to change the distribution chosen to extrapolate.
- Cell C23 allow the user to change the timepoint for which the model uses the Kaplan Meier results from the study.
- Cell C49 contains the MFS-HR.

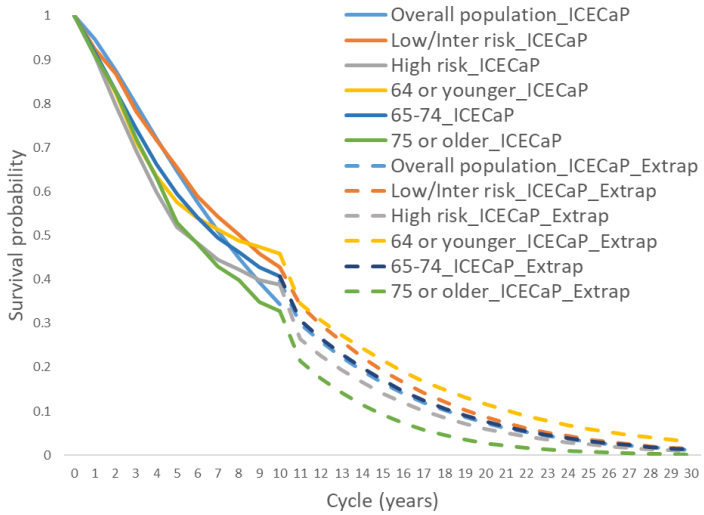
For patients in the BCR-free state in the case study, the Weibull distribution was chosen for the base case analysis as it showed consistent comparative results to other outcomes analysed (MFS and OS) and conservative hazards and a clinically plausible length of the tail compared to other distributions analysed.

If new data for the comparator and intervention are input into the model for BCR-free, cells F24:F34 with the survival probabilities at 1 year cycle needs to be updated as described in Section 3.2. The survival probabilities is used the model to inform the comparator arm and to estimate the survival probabilities for the intervention based on proportional hazards. Cell C49 in the "Dashboard" tab contains the BCR-HR; to estimate the survival probabilities of BCR-free for the intervention, the survival probabilities of BCR-free in the comparator arm are elevated to the potency of the value of the BCR-free-HR.

Note:

The model uses the first 10 years of the Kaplan Meier curve and then the extrapolated data. If less than 10-years of data are available for the model, input the survival probabilities in cells F24:F34 in the tab where the survival analysis of BCR is located and adjust cell C23 with the corresponding data cut-off time.

Figure 6.Biochemical failure free survival from ICECaP and corresponding extrapolation.



3.2.4 General population mortality

Cell C51 in the “Dashboard” tab turns on and off the adjustment of overall survival by the probability of death by other causes observed in the general population.

Cells F7:F62 in the “OC of death” tab contain the proportion of people dying between exact age x and exact age x+1 (qx) from where the background survival probabilities are estimated.

The model includes the option to apply age-adjusted all-cause general population mortality using survival curves with a plateau in the tail. When this option is selected on the “Dashboard” tab, the model compares the estimates of the period specific survival probability to each time window in the model.

In this case study where the Australian setting is used, Australian specific, age-adjusted annual mortality rates are taken from life tables prepared by the Australian Bureau of Statistics[39]. The model uses an input for the mean starting age (65 years old by default) and tracks the present age over the time horizon. From 100% surviving at the start, the model constructs a background survival curve from the piecewise general population mortality probabilities. It then compares the estimated overall survival probabilities with the survival probabilities of the general population at each cycle and converges the prostate cancer OS curve towards the general population survival if the latter is smaller than the OS probabilities predicted by the model.

The model assumes the cohort of patients transiting through the different health states does not survive after 99 years and therefore, there is 100% probability of dying at 100 years.

In the tab "OC of death", cells F7:F62 contain qx (proportion of people dying between exact age x and exact age x+1) that is used to obtain the probability of dying in a given year. The overall survival probability of the population is then calculated in column I as 1-qx.

If a different setting is used in the model, the probability of dying by other causes will need to be updated according to the population being modelled to estimate the background survival probabilities of the general population. Therefore, to adapt the model to other settings, the user should introduce the qx of the population of interest in column F or introduce the new background survival probability in column I of the "OC of death" tab.

3.3 The surrogacy relationship

Cell C48 in the "Dashboard" tab contains the HR of MFS used in the model to estimate the HR of OS based on the ICECaP surrogacy relationship.

[39] Australian Bureau of Statistics. Causes of Death, Australia. https://www.abs.gov.au/statistics/health/causes-death/causes-death-australia/2020

The model relies on the surrogacy relationship between MFS and OS given by ICECaP, and proportional hazard assumption (Attachment 1) to estimate the survival curves of the intervention. The MFS HR between intervention and comparator is used to estimate the OS HR using ICECaP's correlation equation:

Log (HR)-OS= -0.0056 + 0.7813* Log (HR)-MFS.

The resulting OS-HR is then used to estimate the survival probabilities of OS for the intervention as explained in Attachment 1. In this case study, AAP + SoC for patients with high-risk prostate cancer was used in the model. Therefore, the OS probabilities of AAP + SoC are estimated using the OS-HR resulting from the surrogacy relationship based on the OS probabilities of SOC, which in this case is given by the ICECaP patients who received RT after diagnosis.

The parameters in the ICECaP equation come directly from the survival analysis of all studies metanalysed In the study and therefore, they are a fixed fields in the model. The OS HR resulting from the surrogacy relationship is estimated in the model in cell C50 in all tabs where OS is used (OS all pop, OS lowRisk, OS HighRisk, Age64 OS, Age65-74 OS, Age75 OS). The use of the surrogacy relation is limited by the data used to estimate it. In this case, extreme MFS HRs may result in implausible results[40].

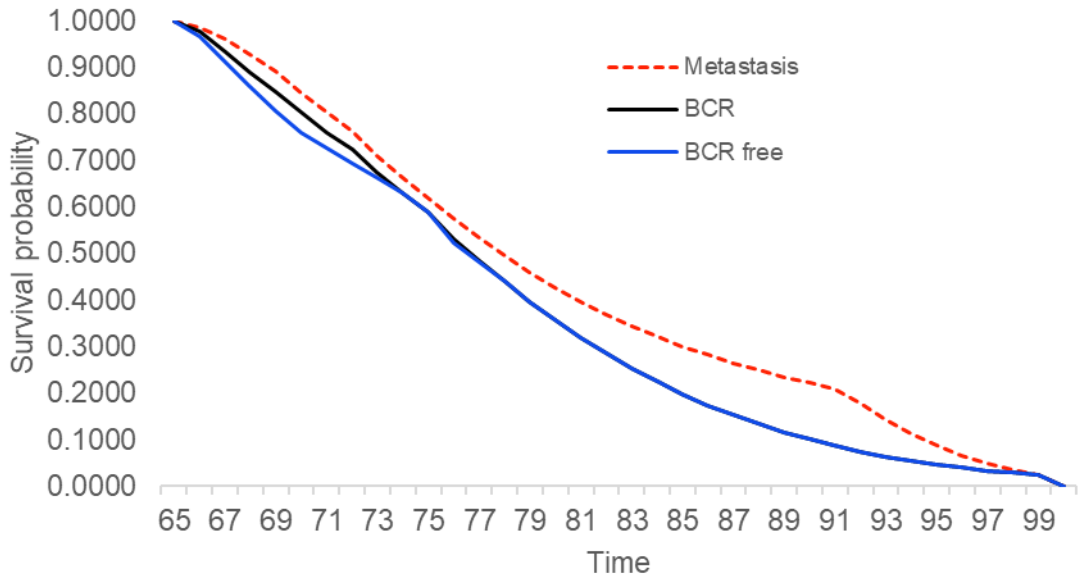
A surrogacy relationship is based on HRs to estimate the treatment effect on OS of the intervention (see section 3.3). HRs are also used in the model to estimate the survival curves of all outcomes in the intervention arm based on the survival curves reported in the ICECaP database for the comparator (Attachment 1 describes the methods). Considering the independence of outcomes associated with the use of PSMs, the use of HRs to estimate treatment effect can impact on how results are interpreted. The interdependence of outcomes and the base relationship of parameters in the ICECaP relationship restrict the model and can

[40] For example, assuming a survival probability of 0.95 for MFS and 0.960 for OS for the comparator, would yield abnormal survival probabilities in the intervention arm resulting in an HR-MFS below 0.3. The surrogacy equation predicts a HR of 0.385 for OS when the HR-MFS is equal to 0.3. Under the proportional hazard assumption, the survival probability of the intervention for MFS was equal to 0.985 ($0.951^{0.3}$) and for OS equal to 0.984 ($0.960^{0.385}$) which are implausible results.

produce outcomes that are not clinically plausible when we have inputs that are at the limit of the surrogacy model parameters. For example, when the ICECaP data are used as comparator for the overall population, a BCR-free HR<0.46 when MFS HR<0.8 would result in similar BCR-free and MFS curves for the intervention. This is translated in almost no patients suffering BCR when receiving the intervention, which can be clinically challenging to accept (Figure 6.7).

The risk of overlapping curves for different outcomes in the model is flagged by a message box that alert the user of the limitations of the model and suggest a visual check of the curves.

Figure 6.7. Survival analysis for patients receiving the intervention when MFS-HR=0.8 and BCR-free-HR=0.4



BCR=biochemical recurrence.

A similar situation occurs when MFS and OS curves are similar, translating to a low number of patients suffering metastasis with the intervention.

As the OS HR resulting from the surrogacy relationship is used to estimate the survival probabilities of the intervention. As mentioned above, the independence of outcomes in a PSM is carried over the extrapolation period, therefore, the chance of curves crossing depends on the shape of the survival curve and the HRs used for BCR-free and MFS.

Note:

- The parameters in the ICECaP equation come directly from the survival analysis of all studies metanalysed in the study and therefore, they are a fixed fields in the model. The OS HR resulting from the surrogacy relationship is estimated in the model in cell C50 in all tabs where OS is used (OS all pop, OS lowRisk, OS HighRisk, Age64 OS, Age65-74 OS, Age75 OS).
- The use of the surrogacy relation is limited by the data used to estimate it. In this case, extreme MFS HRs may result in implausible

In the "Cost" tab, column:

- C: contains the description of resources used in the model for the comparator.
- E: contains the quantity of each resource used in the model for the comparator.
- F: contains the proportion of patients assumed to use the corresponding resource in the model for the comparator.
- G: contains the unit costs to payers of each resource used in the model for the comparator.
- H: contains the unit costs to society of each resource used in the model for the comparator.

In the "Input" tab, cells:

- C10:C47 contain the probabilities of developing adverse events.
- C96 and C97 contain the average weight and height of the population to calculate the surface area for chemotherapy treatment.

3.4 Economic inputs

Economic inputs in this model include treatment (recurrent and subsequent for failure) and administration costs, AE costs, and are largely similar for the intervention and comparator. A discount rate of 5% is used as per PBAC guidelines[41] in Australia and treatment-specific and administration costs per cycle is based on the number of treatment years, dosage per

[41] Department of Health. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (PBAC)
Version 5.0. September 2016. https://pbac.pbs.gov.au/information/about-the-guidelines.html

administration, cost per package and the frequency of use for each item. Cost used in the model are Australian-relevant costs.

ICECaP patients received RT as primary treatment, and this constitutes the comparator arm in the model, therefore, resources used for planning and delivery of RT were considered. Patients who were failed by primary treatment were considered to incur the cost of subsequent treatment. Costs associated with the BCR-free state are given by the cost of primary treatment and the management of AEs associated with it (short and long term). The cost of AE of subsequent treatment are also considered in the model for metastasis and BCR state. Patients transitioning from BCR-free to BCR and from BCR to metastases were assumed to incur the cost of conventional imagining to confirm their disease stage. The model uses a wastage cost approach (i.e., costs are estimated based on the number vials required) for I.V drugs and no administration for oral or I.V drugs. For hormonal implants, the cost of professional consultation (code: MBS105) was used as a proxy for administration cost.

Follow-up was also considered in the model for patients in all states depending on the severity of their disease.

As the resources used are subject to change depending on the intervention being compared, the model allows the user to modify and update the resources used and their costs. In the “Cost” tab, a button located at the top of each table, will help the user to insert as many rows as new resources are needed to consider for each group of resources used. Table 2 describes the location of all economic inputs in the “Cost” tab.

Note:

- The quantity of resources for each treatment, service or medical device described in the “Cost” tab represents the amount estimated to be used in 1 year time.
- The proportion of patients input reflects an estimate of the proportion of patients that would use the resource from the whole population under description.
- If there are resources identified that are not valid to other settings, delete the values in column G and H. Do not delete an entire row.
- No indirect costs are included in this case study; however, the model allows the input of indirect costs in each health state using rows 160 to 178 in the “Cost” tab.

Table 2. Location of all economic inputs in the model.

Description	Resource name	Doses	Quantity	Proportion of patients	Unit cost to payers	Unit cost to society
Primary treatment						
Comparator	C100:C106	D 100:D 106	E100:E106	F100:F106	G100:G106	H100:H106
Intervention	C148:C150	D 148:D 150	E148:E150	F148:F150	G148:G150	H148:H150
ADT	C30:C33	D 30:D 33	E30:E33	F30:F33	G30:G33	H30:H33
Subsequent treatment [α]						
M0CRPC	C123:C124	D 123:D 124	E123:E124	F123:F124	G123:G124	H123:H124
M1CRPC	C129:C135	D 129:D 135	E129:E135	F129:F135	G129:G135	H129:H135
Treatment for metastatic patients	C140:C143	D 140:D 143	E140:E143	F140:F143	G140:G143	H140:H143
Follow-up						
Monitoring Intermediate and high Risk	C7:C8	D 7:D 8	E7:E8	F7:F8	G7:G8	H7:H8
Monitoring ADT	C12:C15	D 12:D 15	E12:E15	F12:F15	G12:G15	H12:H15
Monitoring Metastasis	C19:C25	D 19:D 25	E19:E25	F19:F25	G19:G25	H19:H25

Description	Resource name	Doses	Quantity	Proportion of patients	Unit cost to payers	Unit cost to society
Indirect cost						
Indirect cost BCR-free	C162:C164	D 162:D 164	E162:E164	F162:F164	G162:G164	H162:H164
Indirect cost BCR	C169:C171	D 169:D 171	E169:E171	F169:F171	G169:G171	H169:H171
Indirect cost Metastasis	C176:C178	D 176:D 178	E176:E178	F176:F178	G176:G178	H176:H178
AE$^{\beta}$						
Short term AE Urinary toxicity	C38:C42	D 38:D 42	E38:E42	F38:F42	G38:G42	H38:H42
Short term gastrointestinal toxicity	C46:C52	D 46:D 52	E46:E52	F46:F52	G46:G52	H46:H52
Long term urinary toxicity	C57:C64	D 57:D 64	E57:E64	F57:F64	G57:G64	H57:H64
Long term gastrointestinal toxicity	C69:C74	D 69:D 74	E69:E74	F69:F74	G69:G74	H69:H74
Erectile dysfunction	C65:C82	D 65:D 82	E65:E82	F65:F82	G65:G82	H65:H82
Intervention AE	C155:C157	D 155:D 157	E155:E157	F155:F157	G155:G157	H155:H157
Febrile neutropenia	C117:C118	D 117:D 118	E117:E118	F117:F118	G117:G118	H117:H118
Other AE	C93:C96	D 93:D 96	E93:E96	F93:F96	G93:G96	H93:H96
Imaging						

Description	Resource name	Doses	Quantity	Proportion of patients	Unit cost to payers	Unit cost to society
CI	C111:C112	D 111:D 112	E111:E112	F111:F112	G111:G112	H111:H112

[α] The dosage required for chemotherapy is estimated using the median weight and height of Australian male population between 65 and 74 years of age and surface area, assuming a mean weight of 85.4 kg and body surface area of 1.99 m^2.

[β]Short-term events were those occurring immediately after treatment and lasting no more than 90 days and those starting 90 days after primary treatment with a long-term effect on patients were classified as long-term.

3.4.1 Primary treatment

3.4.1.1 Comparator

In the “Cost” tab, cells:

- C100:C106 contain the description of resources used in the model for the comparator.
- Cell E100:E106 contain the quantity of each resource used in the model for the comparator.
- Cell F100:F106 contain the proportion of patients assumed to use the corresponding resource in the model for the comparator.
- Cell G100:G106 contain the unit cost to payers of each resource used in the model for the comparator.
- Cell H100:H106 contain the unit cost to society of each resource used in the model for the comparator.

ICECaP patients received RT as primary treatment, and this constitutes the comparator arm in the model - therefore, resources used for planning and delivery of RT were considered. If the quantity of resources needs to be changed (e.g. the number of fractions delivered), column E must be updated. As the population modelled corresponds to patients willing to receive treatment, it was assumed all patients receive all resources used in the delivery of RT. If the

In the “Cost” tab, cells:

- C152:C154 contain the description of resources used in the model for the intervention.
- Cell E152:E154 contain the quantity of each resource used in the model for the intervention.
- Cell F152:F154 contain the proportion of patients assumed to use the corresponding resource in the model for the intervention.
- Cell G152:G154 contain the unit cost to payers of each resource used in the model for the intervention.
- Cell H152:H154 contain the unit cost to society of each resource used in the model for the intervention.

proportion of patients using the resources identified in the table is341ifferrent, column F should be updated.

3.4.1.2 Intervention

AAP + SoC was used in the case study, therefore, the cost associated with all resources involved in the acquisition and delivery of this treatment were included for the intervention arm. If the number of resources needs to be changed, column E must be updated, while the proportion of patients being treated (per resource) can be modified in column F.

In the "Cost" tab, cells:

- C30:C33 contain the ADT schedule after initial treatment failure in the comparator arm.
- Cell C183:C186 contain the ADT schedule after initial treatment failure in the intervention arm.
- Cell C123:C124 contain the treatment schedule for patients with hormone insensitive disease in the BCR state.
- Cell C129:C135 contain the treatment schedule for patients with hormone insensitive diseases in the metastasis state.
- Cell C140:C143 contain the treatment schedule for hormone sensitive patients in the metastasis state.

In the "Input" tab, cells:

- C16 contains the rate of hormone insensitive disease.
- C93:C94 contain the average weight and height of the population used in the model.

3.4.2 Subsequent Treatment

Treatment alternatives for patients in BCR after failure by initial treatment depends on their prior treatment and therefore on the intervention to be evaluated by the model. As ICECaP patients received RT as primary treatment it is assumed patients in BCR in the comparator arm received an ADT schedule (as described in cells C30:C33 in the "Cost" tab). Treatment after

intervention failure depends on the intervention evaluated and is described in cells C183:C186 in the “Cost” tab.

Patients in the BCR and metastasis state are assumed to progress to hormone insensitive disease at a yearly rate (see cell C16 in the “Input” tab). This cell can be modified by the user depending on the setting, intervention and population selected.

Chemotherapy treatment for patients in the metastatic state is assumed to be the same regardless of the initial or subsequent treatment received. The dosage required for chemotherapy is estimated using the mean weight and height of the Australian male population between 65 and 74 years of age (resulting in a mean body surface area of 1.99 m^2). Average weight and height of the population used in the model can be modified in cell C93:C94 in the “Input” tab.

In the “Cost” tab, cells:

- C38:C42 contain the resources used to treat patients with urinary toxicity in the short-term.
- C45:C52 contain the resources used to treat patients with gastrointestinal toxicity in the short-term.
- C57:C64 contain the resources used to treat patients with urinary toxicity in the long-term.
- C69:C74 contain the resources used to treat patients with gastrointestinal toxicity in the long-term.
- C79:C82 contain the resources used to treat patients with erectile dysfunction.
- C87:C88 contain the resources used to treat patients with urethral stricture.
- C93:C96 contain the resources used to treat patients with acute myocardial infraction, coronary atherosclerosis, thrombosis, skeletal related event, and fracture during ADT treatment.
- C117:C119 contain the resources used to treat patients with febrile neutropenia.
- C155:C157 contain the resources used to treat patients with other specific AEs related to the intervention.
- C191:C193 contain the resources used to treat patients with other specific AEs related to the comparator.

In the “Input” tab, cells:

- C23:C25 contain the probabilities of developing short term urinary or gastrointestinal toxicities, or erectile dysfunction with the intervention.
- C27:C29 contain the probabilities of developing short term urinary or gastrointestinal toxicities, or erectile dysfunction with the comparator.
- C32:C34 contain the probabilities of developing long term urinary or gastrointestinal toxicities, or erectile dysfunction with the intervention.
- C36:C38 contain the probabilities of developing long term urinary or gastrointestinal toxicities, or erectile dysfunction with the comparator.
- C40 contains the probability of developing febrile neutropenia during metastases treatment.
- C42:C47 contain the probabilities of developing erectile dysfunction, acute myocardial infraction, coronary atherosclerosis, thrombosis, skeletal related events, and fracture during ADT treatment.
- C18:C19 contain the probabilities of developing other specific AE for the intervention and comparator respectively.

The model focuses on AEs that may have a negative impact on sexual, urinary[42], and bowel function after primary treatment as these have been widely reported in the prostate cancer treatment field (10) (11). AEs were classified according to when they occurred and duration: short-term were those occurring immediately after treatment and lasting no more than 90 days, while those starting 90 days after primary treatment with a long-term effect on patients were classified as long-term.

All patients in subsequent treatment (and those considered high-risk) are assumed to receive ADT. The AE are assumed to be the same regardless of the primary treatment received. Patients undergoing hormonal treatment with ADT have a different AE profile from those receiving primary treatment. A recent study that assessed AEs after continuous and intermittent ADT, found that the most common events were endocrine (41%), bone related events (14% having a fracture), ischemic and thrombotic events (10% having ischemic heart disease) (12). As this model focused on serious AEs requiring treatment, the AEs considered were acute myocardial infraction, thrombosis, ischemic heart diseases, and fractures. The aggregated cost of these AEs was extracted from the National hospital cost data collection. (NHCDC). Coronary atherosclerosis was used as a proxy for ischemic heart disease and the cost of distal femoral fractures was used as a proxy for skeletal fractures (unspecified in the study).

As different interventions and comparator could be used in the model, specific AE profile are used for the comparator and intervention. The probability of other specific AEs for the intervention and comparator can be added in cell C18:C19 in the “Input” tab. Cost associated with these AEs can be added in cells C155:C157 and cells C191:C193 in the “Cost” tab for the intervention and comparator groups respectively.

[42] Urinary toxicity includes treatment for irritative voiding symptoms and incontinence. Gastrointestinal toxicity includes treatment for rectal bleeding, pain, and diarrhea.

3.4.4 Follow up

In the "Cost" tab, cells:

- C7:C8 contain the follow-up schedule for patients in intermediate and high-risk disease in the BCR-free state.
- C12:C16 contain the follow-up schedule for patients receiving ADT treatment in the BCR state.
- C19:C15 contain the follow-up schedule for patients in metastatic state.

Follow-up costs in the model are assigned on the basis of treatment and health state membership. Healthcare resources used for management of different stages of prostate cancer were identified based on the recommendation of international clinical guidelines (8, 9) and local clinical practice in Australia as advised by expert opinion. The model assumes that follow-up costs are the same regardless of initial treatment received.

3.4.5 Other costs

In the "Cost" tab, cells:

- C160:C178 contain the description of indirect resources used for each health state in the model.
- E160:E178 contain the quantity of indirect resources used for each health state in the model.
- F160:F178 contain the proportion of patients for which the corresponding indirect resources are applied in each health state in the model.
- G160:G178 contain the unit cost to payers of each indirect resources used for each health state in the model.
- H160:H178 contain the unit cost to society of each indirect resources for each health state in the model.

The model includes the capacity to capture resources lost due to the cessation of work or premature mortality e.g. productivity losses or resources used to provide home care. The

model assumes all indirect costs depend on patients' health state membership. Therefore, indirect costs are described in the "Cost" tab for patients under BCR-free, BCR or metastasis.

Note:

The Australian version of the model does not have the indirect cost parameters populated (functionality has been included for later use as relevant).

3.5 Quality of life inputs

Cells C83:C91 in the "Input" tab contain the utilities (quality of life weights) for all health states used in the model to estimate the quality of life of the cohort of patients.

For the purpose of this case study, a literature review was conducted in Medline to identify studies reporting utility values in patients with prostate cancer through all stages used in the model. Several studies reporting health utilities of patients in relevant health states were found, nevertheless the most updated systematic review and meta-analysis helped to identify and analyse the studies reporting health utilities in patients with prostate cancer (13). There were 29 studies included in this meta-analysis that reported utilities for patients after RT, RP, or ADT, metastatic disease, and who suffered AEs of treatments (14-42). Metastasis was reported in nine papers, four studies reported utilities of patients after RT (27, 28, 32, 34) and one reported the utility for patients on adjuvant ADT (32). Table 4.14 summarise the pooled utility values resulting from the meta-analysis conducted for this case study.

Table 3. Pooled estimates of utility values used in the model.

Health state	Utility value	CI-L	CI-U	I^2
AAP + SoC	0.79	0.594	0.986	NA
SoC	0.79	0.594	0.986	NA
BCR	0.790	0.594	0.986	NA
Metastasis	0.671	0.64	0.701	71.64%

Bowel/Rectal discomfort in BCR-free	0.859	0.833	0.885	NA
Erectile dysfunction in BCR-free	0.799	0.675	0.923	97.29%
Urinary incontinence in BCR-free	0.763	0.62	0.906	97.77%

BCR: Biochemical recurrence; CI-L: Low confidence interval; CI-U: Upper confidence interval; I2: variability of pooled utility value due to heterogeneity; RP: Radical prostatectomy; Radiotherapy.

RT pooled utility values were used to populate the BCR-free state for patients who survived the initial treatment and had no AEs related to the initial treatment whereas ADT utility was used as a proxy for the quality of life of patients after AAP. The utility values of urinary and bowel/rectal toxicities and erectile disfunction were used for patents in the BCR-free state presenting at least one of these events after initial treatment and applied only for the number of cycles over which the AE is assumes to occur. Although there is a long-term probability of developing these AEs for all patients independent of their response to initial treatment, it was assumed that utility values in progressed (BCR and metastatic) health states capture the detrimental effect in quality of life caused by the long-term AEs after primary therapy.

Cells C83:C91 in the “Input” tab of the model can be modified to adapt utility values to the local setting. Changes in quality of life associated with treatments are assumed to be captured by the utility values assigned to each health state. On the other hand, for BCR-free, the utility values of the AE most seen after primary treatment of prostate cancer are used to estimate the quality of life of these patients.

3.6 Model outputs

Results of the cost-effectiveness model are available in cells B4:B42:

- B4 and B5 contain the total cost for intervention and comparator.
- B7 and B8 contain total QALYs for intervention and comparator.
- B6 and B9 contain incremental cost and QALYs.
- B10 contains the ICER.
- B11:B21: contain disaggregated cost and QALYs for the intervention.
- B27:B37 contain disaggregated cost and QALYs for the comparator.
- B22:B26 contain the time in years spent in each health state in the intervention arm.
- B38:B42 contain the time in years spent in each health state in the comparator arm.

Model outputs can be reviewed from the "Results" tab. Model outputs include:

- Total accrued life-years (Lys): Time spent in each health state in the model is estimated as the area under the curve using the trapezoid rule. Time spent in OS corresponds to the total survival time.
- QALYs: Calculated based on the time spent in each health state multiplied by the utility associated with the health state.
- Disaggregated costs: All costs accrued by an average patient in treatment, follow-up, AE, and diagnosis imaging[43] by health state (BCR-free, BCR and metastasis).
- The incremental cost-utility ratio (ICER) estimated as the incremental cost per QALY for the average patient.

Note:

All costs described in the "Results" tabs are dependent on the time horizon chosen in C46 in the "Dashboard" tab.

[43] The model assumes the cohort of patients start the model with localised diseases. Diagnosis imaging refers to the technique used to confirm BCR and metastasis.

3.6.1 Benefits to society

In the “Dashboard” tab, cells:

- C53 contains the time in years to estimate the benefits to society.
- N59:N61 and O59:O61 contain the number of patients BCR-free, BCR and metastasis in the comparator and intervention arms respectively.
- P66 contains differences between the number of patients dead in the intervention and comparator within the time gained from using the surrogacy relationship.
- V59:V67 contain the disaggregated difference in life years, cost and QALY per patient when long-term results of the model are compared based on making a public reimbursement decision now (using the surrogate outcome) compared to making that decision when final outcomes are known.

This aspect of the model estimates the benefits to society when the reimbursement decision-making process is based on the surrogacy relationship rather than waiting on final outcomes data. The ICECaP study was able to correlate the OS rate at 8 years with MFS rate at 5 years (1), showing a potential reduction of 3 years when a reimbursement decision could be made. Hence, the model uses the difference in cost and QALYs between intervention and comparator in the first 3 years, as that corresponds to the potential time saved by an earlier reimbursement decision based on 5-year MFS. The model allows for the time period over which the impact to society is assessed to be varied: the user can choose the duration over which to estimate the benefits to society in Cell C53 in the “Dashboard” tab.

The difference that a 3-year delayed reimbursement decision would produce in terms of absolute values for the whole cohort of patients was calculated to estimate the potential benefit to society. The epidemiologic characteristics of the population starting the model can be modified in cells D13:D18 in the “Population” tab and in cell B24:E26 in the “Population” tab, for the estimates used to calculate the benefits to society. This allows the cohort specific benefits to be applied at the population level.

As for the long-term benefits, the model assumes that only a proportion of the initial cohort will be able to receive the intervention as their primary treatment, as some will have progressed to an advanced stage of the disease due to treatment delay. The model compares

the cost and benefits over the long term for “timely reimbursement decision arm” and the “delayed reimbursement arm”. The difference in number of patients in each health state during the first 3 years for the short-term analysis between intervention and comparator is used to populate the delayed reimbursement arm of the model. For this purpose, it was assumed the proportion of patients who progressed under the comparator arm do not receive primary treatment and start in the BCR or metastasis health state instead. On the other hand, when a timely reimbursement decision is made, all patients of the cohort are assumed to start in BCR-free and receive the intervention as primary treatment. The model is run to evaluate the cost and benefits accounting for the difference in the starting population between both arms in the long term.

3.7 Sensitivity analysis

Sensitivity analysis

- Cell F2 in the “Dashboard” tab contains the option choice of sensitivity analysis.
- Cell D2 in the “One way SA” tab contains the button to run a one-way sensitivity analysis.
- Cell E4 in the “Probabilistic SA” contains the button to run a probabilistic sensitivity analysis.
- Columns C, E and F in the “Input” tab contain the point estimate, minimum and maximum values used for the sensitivity analyses respectively.
- Column G in the “Input” tab contains the type of distribution used for the probabilistic sensitivity analysis.
- Columns H, I, and J in the “Input” tab contain the standard error, alpha and beta values respectively used to estimate the probability (according to their distribution) in column K, which is later used in the probabilistic sensitivity analysis.
- Column K in the “Input” tab contains the distribution assigned to each variable used for the random sampling method of the probabilistic sensitivity analysis.
- Cell G2 in the “Probabilistic SA” tab contains the threshold values used for the net benefit analysis.

Sensitivity analysis is performed to understand the impact on the model results of the assumptions and uncertainty with the inputs used. The workbook contains two types of analysis:

Deterministic sensitivity analyses including a tornado plot summarising the impact of inputs on the ICER estimates (see “Deterministic SA” tab).

Probabilistic sensitivity analysis (PSA): an incremental cost-effectiveness plane and net monetary benefit acceptability curves (NBAC; see the “Probabilistic SA” in the “Prob SA” tab).

Both analyses use the values stored in column E and F in the “Input” tab, which contain the minimum and maximum deviation of the point estimate for each input variable. If this information is not available, the model applies a ± 20% variation.

Note:

Before running any sensitivity analysis, confirm cell F2 in the “Dashboard” tab contains the choice of sensitivity analysis to be run. The results of probabilistic sensitivity analysis when the value of cell F2 is “One way SA” will result in a single point after 5,000 runs. Likewise, results of the one-way sensitivity analysis when the value of cell F2 is “Probabilistic SA” will not be able to capture the correct impact of each variable on the ICER.

3.7.1 Deterministic sensitivity analysis

The deterministic sensitivity analysis explores the ICER when extreme values, within the range defined for each variable, are used in place of its point estimate. In the “One way SA” tab:

- Columns B and C contain the variable name and its description respectively.
- Column D is the base case ICER value.
- Column E and F are the ICER when the minimum and maximum values within the range defined for each variable are used.

Using the values explained above, a rank list is constructed based on the difference of the estimated ICER against the base case, with their values plotted in a tornado figure located at cell N4 in “One way SA” tab.

Note:

A macro estimates the ICERs using the minimum and maximum values of each variable and then sorts the list according to the magnitude of the impact on the ICER given by each variable. If the shape of the figure is not as expected, the user can modify column K in the “One way SA” with the desired ranking and then sort it from smallest to largest to organise the variables accordingly.

3.7.2 Probabilistic sensitivity analysis

In this type of sensitivity analysis, variables used in the model are represented by distributions around the point estimate. From the distribution assigned to each variable in the “Input” tab, a value is drawn by random sampling to explore the ICER when the model is run using that value. This is repeated 5,000 times for all variables at the same time, resulting in a new distribution of ICERs plotted on the cost-effectiveness plane located at cell AL5 of the “Probabilistic SA” tab.

The net benefit analysis and the cost-effectiveness acceptability curve can be found in columns AI and AJ in the “Probabilistic SA” tab. A plot with these values summarises the impact of uncertainty expressed as ICERs in relation to the probability that the intervention will be cost-effective at a given threshold. The threshold value can be found in cell G2 in the “Probabilistic SA” and the user can modify this value according to the setting evaluated.

Attachment 1

Proportional hazard assumption for RP

The construction of MFS and BCR survival curves for radical prostatectomy was based on the HR comparing SoC versus AAP + SoC under the assumptions of Cox's proportional hazard model (43):

$$\ln\left(\frac{h(t,X)}{h_0(t)}\right) = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots$$

In this model, h(t,X) is the hazard (instant rate of failure or the probability of failure of a patient that has survived up to time t, between t and Δt, divided by Δt, as Δt approaches zero) at time t for a patient with covariates X_1, X_2, X_3, … h0(t) represents the baseline hazard at time t or the hazard of a patient when all covariates = 0, and X_1, X_2, X_3, stand for variables that can influence the failure probability like patient's age, presence of co-morbidities or treatment received. By exponentiating this equation, it can be rewritten in terms of the hazard:

$$\mathrm{h(t,X)} = h_0(\mathrm{t})\exp[\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots]$$

As an example, the hazard function for distant metastasis developing of patients undergoing radiotherapy (i.e., $X_1 = 1$), is given by

$$h_{RT}(\mathrm{t,X}) = h_0(\mathrm{t})\exp[\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots],$$

while for patients undergoing radical prostatectomy (i.e., $X_1 = 0$), the hazard function for distant metastasis is given by

$$h_{RP}(\mathrm{t,X}) = h_0(\mathrm{t})\exp[\beta_2 X_2 + \beta_3 X_3 + \cdots],$$

Then the ratio of these hazard functions (HR):

$$HR_{DM} = \frac{h_{RT}(t,X)}{h_{Abi}(t,X)} = \frac{h_0(t)\exp[\beta_1 + \beta_2 X_2 + \beta_3 X_3 \ldots]}{h_0(t)\exp[\beta_2 X_2 + \beta_3 X_3 \ldots]}$$

$$= \frac{h_0(t)\exp[\beta 1]\exp[\beta_2 X_2 + \beta_3 X_3 + \cdots]}{h_0(t)\exp[\beta_2 X_2 + \beta_3 X_3 + \cdots]} = \exp[\beta_1]$$

As the hazard function can be understood in terms of the survival probability (and vice versa), survival at time t (which is denoted by S(t)) can be written in terms of the hazard function:

$$S(t,X) = \exp\left[-\int_0^t h(u,X)du\right]$$

$$= \exp\left(-\exp(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots)\left[-\int_0^t h_0(u)du\right]\right.$$

$$S(t,X) = \left(\exp\left[-\int_0^t h(u,X)du\right]\right)^{\exp(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots)}$$

$$S(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots)}$$

Thus, the survival function for patients that receive radiotherapy is given by

$$S_{RT}(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots)}$$

Then, the survival function for patients that receive radical prostatectomy is given by

$$S_{Abi}(t,X) = S_0(t)^{\exp(\beta_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots)} = S_0(t)^{\exp[\beta 1]\exp[\beta_2 X_2 + \beta_3 X_3 + \ldots)}$$

$$= (S_{RT}(t))^{\exp[\beta 1]}$$

$$= (S_{RT}(t))^{[HR]}$$

Thus, in this model the probability of dying, developing distant metastasis or biochemical recurrence for patients receiving the intervention were calculated from the survival function of patients who received radiotherapy exponentiated to the hazard ratio when both alternatives are compared under the Cox's proportional hazard assumption.

Attachment 2

Cost tables used in the model.

Table 4. Follow-up cost used in the model.

Monitoring intermediate and high risk							
Mbs cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
66660	Psa	1 every 5 months	2	100%	$ 31.75	$ 37.30	B5:c8
105	Professional consultation + dre	1 every 5 months	2	100%	$ 38.25	$ 45.00	
Monitoring adt							
Mbs cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
66660	Psa	1 every 3 months	4	100%	$ 31.75	$ 37.30	B10:c15
105	Professional consultation + dre	1 every 3 months	4	100%	$ 38.25	$ 45.00	
66695	Testosterone	1 every 5 months	2	80%	$ 25.95	$ 30.50	
61441	Bone scan		1	26%	$ 416.25	$ 489.70	
Monitoring metastasis							
Mbs cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
66660	Psa	1 every 3 months	4	100%	$ 31.75	$ 37.30	B17:c25
105	Professional consultation + dre	1 every 3 months	4	100%	$ 38.25	$ 45.00	

66500	Blood test	1 every 5 months	2	80%	$ 41.25	$ 48.50	
66500	Crea	1 every 5 months	2	80%	$ 8.25	$ 9.70	
66500	Alkaline ph	1 every 5 months	2	80%	$ 8.25	$ 9.70	
66695	Testosterone quantitation	1 every 5 months	2	80%	$ 25.95	$ 30.50	
61441	Bone scan	Bone pain or other symptoms of progression	1	50%	$ 416.25	$ 489.70	
Adt after compartor							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
8094b	Bicalutamide 50mg	50mg/day	28	100%	$ 0.60	$ 2.07	B28:c33
8093y	Goserelin 10.8 mg implant	1 implant/12weeks	4.33	70%	$ 873.85	$ 914.85	
8859g	Leuprorelin acetate 45 mg implant	1 implant/6 months	2.00	30%	$ 1,709.50	$ 1,750.50	
105	Professional consultation		2.17	100%	$ 38.25	$ 45.00	
Adt after intervention							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
8094b	Bicalutamide 50mg	50mg/day	28	100%	$ 0.60	$ 2.07	B36:c41
8093y	Goserelin 10.8 mg implant	1 implant/12weeks	4.33	70%	$ 873.85	$ 914.85	
8859g	Leuprorelin acetate 45 mg implant	1 implant/6 months	2.00	30%	$ 1,709.50	$ 1,750.50	
105	Professional consultation		2.17	100%	$ 38.25	$ 45.00	
M0crpc							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual

10174l	Enzalutamide 40 mg	160mg/day	1460	100%	$ 31.72	$ 31.72	B44:c47
Rt							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
15565	Preparation of an imrt dosimetry plan, which uses one or more ct image volume datasets	1	1	100%	$ 3,332.65	$ 3,417.00	B50:c58
15555	Simulation for intensity-modulated radiation therapy (imrt), with or without intravenous contrast medium	1	1	100%	$ 648.05	$ 732.75	
15248	Radiation oncology treatment 1 field - treatment delivered to primary site (prostate)	1	1	100%	$ 52.30	$ 61.50	
15263	Radiation oncology treatment, 2 or more fields up to a maximum of 5 additional fields (rotational therapy being 3 fields) - treatment delivered to primary site (prostate)	1	5	100%	$ 39.15	$ 39.15	
15715	Radiation oncology treatment verification	1	6	100%	$ 67.15	$ 79.00	
104	Radiation oncologist 1st visit	1	1	100%	$ 76.15	$ 89.55	
105	Radiation oncologist review during treatment and follow up	1	6	100%	$ 38.25	$ 45.00	

Gi toxicity short-term							
105	Professional consultation + dre: professional attendance by a specialist in the practice of his or her specialty following referral of the patient to him or her-an attendance after the first in a single course of treatment, if that attendance is at consulting rooms or hospital, other than a service to which item 16404 applies		3	100%	$ 38.25	$ 45.00	
23	Gp visits		2	100%	$ 38.75	$ 38.75	
65070	Blood test: erythrocyte count, haematocrit, haemoglobin, calculation or measurement of red cell index or indices, platelet count, leucocyte count and manual or instrument generated differential count - not being a service where haemoglobin only is requested - one or more instrument generated set of results from a single sample; and (if performed)		2	100%	$ 41.25	$ 48.50	B60:c68
4039n	Oinment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg		3	100%		$ 18.45	
4040p	Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg	2/day/3 months	180	100%		$ 1.46	
1215y	Paracetamol 500 mg + codeine 30 mg tablet	1tbl/12hrs/3m	180	100%		$ 0.64	
1571q	Loperamide hydrochloride 2 mg capsule	1caps/12hrs/3m	180	100%		$ 1.08	
Long term urinary toxicity							

Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
105	Professional consultation + dre		3	100%	$ 38.25	$ 45.00	B71:c80
23	Gp visits		5	100%	$ 38.75	$ 38.75	
65070	Blood test		2	100%	$ 41.25	$ 48.50	
11900	Urine flow study		2	100%	$ 24.15	$ 28.40	
Na	Incontinence pads	2/day	730	11%		$ 0.60	
Na	Incontinence pads	4/day	1460	2%		$ 0.60	
1215y	Paracetamol 500 mg + codeine 30 mg tablet	1tbl/12hr/6months	365	100%	$ -	$ 0.13	
4049d	Ural (3 sachets a day)	1scht/8hrs/year	1095	100%	$ 0.24	$ 0.65	
Long term gastrointestinal toxicity							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
105	Professional consultation + dre		3	100%	$ 38.25	$ 45.00	B83:c90
23	Gp visits		5	100%	$ 38.75	$ 38.75	
4039n	Oinment: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg		2	100%	$ -	$ 18.45	
4040p	Suppository: zinc oxide 300 mg + peru balsam 50 mg + benzyl benzoate 33 mg	4/day/6 weeks	168	100%	$ -	$ 1.46	
10582y	Paracetamol 500 mg	2tbl/8hr/3 months	540	100%	$ -	$ 0.13	
1571q	Loperamide hydrochloride 2 mg capsule	1caps/12hrs/6months	365	100%	$ -	$ 1.08	
Erectile dysfunction							

Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
105	Professional consultation + dre		1	100%	$ 38.25	$ 45.00	B93:c98
23	Gp visits		2	100%	$ 38.75	$ 38.75	
4585h	Sildenafil 50 mg tablet	2/month	24	90%	$ 10.39	$ 16.99	
4579b	Alprostadil 10 microgram injection	1/month	12	10%	$ 99.99	$ 106.59	
Urethral stricture							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
L66z	Urethral stricture						B101:c104
Other ae							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
F76b	Acute mi (arrhy,card arrst&cond dis, minc)			0.9%			B107:c112
F66b	Coronary atherosclerosis, minc			0.8%			
F63b	Thrombosis (venous thrombosis, minc)			2.0%			
I61z	Skeletal related event and fracture (distal femoral fractures)			1.5%			
Rt							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
15565	Preparation of an imrt dosimetry plan, which uses one	1	1	100%	$ 3,332.65	$ 3,417.00	B114:c124

	or more ct image volume datasets						
15555	Simulation for intensity-modulated radiation therapy (imrt), with or without intravenous contrast medium	1	1	100%	$ 648.05	$ 732.75	
15248	Radiation oncology treatment 1 field - treatment delivered to primary site (prostate)	1	1	100%	$ 52.30	$ 61.50	
15263	Radiation oncology treatment, 2 or more fields up to a maximum of 5 additional fields (rotational therapy being 3 fields) - treatment delivered to primary site (prostate)	1	5	100%	$ 39.15	$ 39.15	
15715	Radiation oncology treatment verification	1	6	100%	$ 67.15	$ 79.00	
104	Radiation oncologist 1st visit	1	1	100%	$ 76.15	$ 89.55	
105	Radiation oncologist review during treatment and follow up	1	6	100%	$ 38.25	$ 45.00	
Ci							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
56507	Computed tomography		1	100%	$ 414.20	$ 487.25	B125:c128
61441	Bone marrow study		1	100%	$ 416.25	$ 489.70	
Neutropenia febril							

Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
Na	Febril neutropenia cost in 2010 brought to 2020			100%	$ 4,184.92	$ 4,184.92	B131:c135
M1crpc hormon therapy							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
8094b	Enzalutamide 40 mg	160mg/day	1460	8%	$ 31.35	$ 31.72	B137:c141
11206t	Abiraterone 500mg	1gr/day	730	28%	$ 56.95	$ 57.63	
1917x	Prednisolone 5 mg tablet	10mg/day per 6 weeks	84	14%	$ 0.25	$ 0.35	
M1crpc chemotherapy							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
4376h	Cabazitaxel 60 mg/1.5 ml injection 1,5ml vial	20 mg/m2 i.v. Every 21 days x 10 cycles	10	7%	$ 2,894.78	$ 2,935.78	B144:c148
10148d	Docetaxel 160 mg/16 ml injection, 16 ml vial	Six 3-weekly cycles of 75 mg/m² x 18 cycles	18	14%	$ 141.26	$ 182.26	
1291y	Dexamethasone 8mg	8 mg i.v. Every 21 days	10	12%	$ 21.27	$ 26.95	
10061m	Denosumab 120mg/1.7ml	120 mg s.c. Every 4 weeks	13	11%	$ 470.10	$ 511.10	
Met treatment: stampede (treatment for metastasis patients)							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
10148d	Docetaxel 160 mg/16 ml injection, 16 ml vial	Six 3-weekly cycles of 75 mg/m²	18	100%	$ 141.26	$ 182.26	B151:c156
1917x	Prednisolone 5 mg tablet	10mg/day per 6 weeks	84	100%	$ 0.25	$ 0.35	

1291y	Dexamethasone 8mg	8 mg i.v. Every 21 days	10	100%	$ 21.27	$ 26.95	
Aap + soc cost							
Cod	Description	Dose	Quantity	Frequency	Unitary cost payer	Unitary cost society	Reference in the manual
L05a	Abiraterone 500mg	1000mg/day	730	100%	$ 56.68	57	B159:c163
L05b	Prednisolone 5mg	5mg/day	365	100%	$ 0.25	0.25	
66500	Quantitation in serum of aminotransferase	Every 2 weeks for 3 month then monthly	15	100%	$ 7.30	9.70	

Caps: Capsules; GP: General practitioner; NA: Not applicable; Scht: Sachet; Tbl: Tablets.

Appendix 7 Bibliography

1. Anstrom KJ, Reed SD, Allen AS, Glendenning GA, Schulman KA. Long-term survival estimates for imatinib versus interferon-alpha plus low-dose cytarabine for patients with newly diagnosed chronic-phase chronic myeloid leukemia. Cancer. 2004;101(11):2584-92.

2. Bagust A, Beale S. Survival analysis and extrapolation modeling of time-to-event clinical trial data for economic evaluation: an alternative approach. Medical decision making : an international journal of the Society for Medical Decision Making. 2014;34(3):343-51.

3. Delea TE, Amdahl J, Wang A, Amonkar MM, Thabane M. Cost effectiveness of dabrafenib as a first-line treatment in patients with BRAF V600 mutation-positive unresectable or metastatic melanoma in Canada. PharmacoEconomics. 2015;33(4):367-80.

4. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making. 2013;33(6):743-54.

5. Oriana C, Martin H, Toby P, Chris C, Ruth G, Claudius R, et al. Complete cytogenetic response and major molecular response as surrogate outcomes for overall survival in first-line treatment of chronic myelogenous leukemia: a case study for technology appraisal on the basis of surrogate outcomes evidence. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2013;16(6):1081-90.

6. Reed SD, Anstrom KJ, Li Y, Schulman KA. Updated estimates of survival and cost effectiveness for imatinib versus interferon-alpha plus low-dose cytarabine for newly diagnosed chronic-phase chronic myeloid leukaemia. PharmacoEconomics. 2008;26(5):435-46.

7. Reed SD, Anstrom KJ, Ludmer JA, Glendenning GA, Schulman KA. Cost-effectiveness of imatinib versus interferon-alpha plus low-dose cytarabine for patients with newly diagnosed chronic-phase chronic myeloid leukemia. Cancer. 2004;101(11):2574-83.

8. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. European urology. 2017;71(4):618-29.

9. Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2019;17(5):479-505.

10. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. The New England journal of medicine. 2016;375(15):1415-24.

11. Hayes JH, Ollendorf DA, Pearson SD, Barry MJ, Kantoff PW, Stewart ST, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. Jama. 2010;304(21):2373-80.

12. Hershman DL, Unger JM, Wright JD, Ramsey S, Till C, Tangen CM, et al. Adverse Health Events Following Intermittent and Continuous Androgen Deprivation in Patients With Metastatic Prostate Cancer. JAMA Oncol. 2016;2(4):453-61.

13. Magnus A, Isaranuwatchai W, Mihalopoulos C, Brown V, Carter R. A Systematic Review and Meta-Analysis of Prostate Cancer Utility Values of Patients and Partners Between 2007 and 2016. MDM Policy Pract. 2019;4(1):2381468319852332.

14. Avila M, Becerra V, Guedea F, Suarez JF, Fernandez P, Macias V, et al. Estimating preferences for treatments in patients with localized prostate cancer. Int J Radiat Oncol Biol Phys. 2015;91(2):277-87.

15. Avila M, Pardo Y, Castells M, Ferrer F, Boladeras A, Pera J, et al. Adaptation and validation of the Spanish version of the Patient-Oriented Prostate Utility Scale (PORPUS). Qual Life Res. 2014;23(9):2481-7.

16. Basu A, Dale W, Elstein A, Meltzer D. A linear index for predicting joint health-states utilities from single health-states utilities. Health Econ. 2009;18(4):403-19.

17. Basu A, Dale W, Elstein A, Meltzer D. A time tradeoff method for eliciting partner's quality of life due to patient's health states in prostate cancer. Med Decis Making. 2010;30(3):355-65.

18. Bremner KE, Mitsakakis N, Wilson L, Krahn MD. Predicting utility scores for prostate cancer: mapping the Prostate Cancer Index to the Patient-Oriented Prostate Utility Scale (PORPUS). Prostate Cancer Prostatic Dis. 2014;17(1):47-56.

19. Bremner KE, Tomlinson G, Krahn MD. Marker states and a health state prompt provide modest improvements in the reliability and validity of the standard gamble and rating scale in prostate cancer patients. Qual Life Res. 2007;16(10):1665-75.

20. C Glazener CB, B Buckley, C Cochran,, G Dorey AG, S Hagen, M Kilonzo,, A McDonald GM, K Moore,, J N'Dow JN, C Ramsay and L Vale. Conservative treatment for urinary incontinence in Men After Prostate Surgery (MAPS): two parallel randomised controlled trials. Health Technology Assessment. 2011;26.

21. Cantor SB, Deshmukh AA, Krahn MD, Volk RJ. Use of forecasted assessment of quality of life to validate time-trade-off utilities and a prostate cancer screening decision-analytic model. Health Expect. 2015;18(5):1610-20.

22. Cantor SB, Volk RJ, Krahn MD, Cass AR, Gilani J, Weller SC, et al. Concordance of couples' prostate cancer screening recommendations from a decision analysis. Patient. 2008;1(1):11-9.

23. Dale W, Basu A, Elstein A, Meltzer D. Predicting utility ratings for joint health States from single health States in prostate cancer: empirical testing of 3 alternative theories. Med Decis Making. 2008;28(1):102-12.

24. Diels J, Hamberg P, Ford D, Price PW, Spencer M, Dass RN. Mapping FACT-P to EQ-5D in a large cross-sectional study of metastatic castration-resistant prostate cancer patients. Qual Life Res. 2015;24(3):591-8.

25. Farkkila N, Torvinen S, Roine RP, Sintonen H, Hanninen J, Taari K, et al. Health-related quality of life among breast, prostate, and colorectal cancer patients with end-stage disease. Qual Life Res. 2014;23(4):1387-94.

26. Gries KS, Regier DA, Ramsey SD, Patrick DL. Preferences for Prostate Cancer Outcomes: A Comparison of the Patient Perspective, the General Population Perspective, and a Population at Risk for Prostate Cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(2):218-25.

27. Hall PS, Hamilton P, Hulme CT, Meads DM, Jones H, Newsham A, et al. Costs of cancer care for use in economic evaluation: a UK analysis of patient-level routine health system data. Br J Cancer. 2015;112(5):948-56.

28. Jayadevappa R, Schwartz JS, Chhatre S, Wein AJ, Malkowicz SB. Association between utility and treatment among patients with prostate cancer. Qual Life Res. 2010;19(5):711-20.

29. Korfage IJ, de Koning HJ, Habbema JD, Schroder FH, Essink-Bot ML. Side-effects of treatment for localized prostate cancer: are they valued differently by patients and healthy controls? BJU international. 2007;99(4):801-6.

30. Krahn M, Bremner KE, Tomlinson G, Ritvo P, Irvine J, Naglie G. Responsiveness of disease-specific and generic utility instruments in prostate cancer patients. Qual Life Res. 2007;16(3):509-22.

31. Krahn MD, Bremner KE, Alibhai SM, Ni A, Tomlinson G, Laporte A, et al. A reference set of health utilities for long-term survivors of prostate cancer: population-based data from Ontario, Canada. Qual Life Res. 2013;22(10):2951-62.

32. Krahn MD, Bremner KE, Tomlinson G, Naglie G. Utility and health-related quality of life in prostate cancer patients 12 months after radical prostatectomy or radiation therapy. Prostate Cancer Prostatic Dis. 2009;12(4):361-8.

33. Ku J, Krahn M, Trachtenberg J, Nesbitt M, Kalnin R, Lockwood G, et al. Changes in health utilities and health-related quality of life over 12 months following radical prostatectomy. Can Urol Assoc J. 2009;3(6):445-52.

34. Lane A, Metcalfe C, Young GJ, Peters TJ, Blazeby J, Avery KN, et al. Patient-reported outcomes in the ProtecT randomized trial of clinically localized prostate cancer treatments: study design, and baseline urinary, bowel and sexual function and quality of life. BJU international. 2016;118(6):869-79.

35. Loriot Y, Miller K, Sternberg CN, Fizazi K, De Bono JS, Chowdhury S, et al. Effect of enzalutamide on health-related quality of life, pain, and skeletal-related events in asymptomatic and minimally symptomatic, chemotherapy-naive patients with metastatic castration-resistant prostate cancer (PREVAIL): results from a randomised, phase 3 trial. The Lancet Oncology. 2015;16(5):509-21.

36. Shimizu F, Fujino K, Ito YM, Fukuda T, Kawachi Y, Minowada S, et al. Factors associated with variation in utility scores among patients with prostate cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2008;11(7):1190-3.

37. Skaltsa K, Longworth L, Ivanescu C, Phung D, Holmstrom S. Mapping the FACT-P to the preference-based EQ-5D questionnaire in metastatic castration-resistant prostate cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2014;17(2):238-44.

38. Sommers BD, Beard CJ, D'Amico AV, Kaplan I, Richie JP, Zeckhauser RJ. Predictors of patient preferences and treatment choices for localized prostate cancer. Cancer. 2008;113(8):2058-67.

39. Torvinen S, Farkkila N, Sintonen H, Saarto T, Roine RP, Taari K. Health-related quality of life in prostate cancer. Acta Oncol. 2013;52(6):1094-101.

40. Watson E, Shinkins B, Frith E, Neal D, Hamdy F, Walter F, et al. Symptoms, unmet needs, psychological well-being and health status in survivors of prostate cancer: implications for redesigning follow-up. BJU international. 2016;117(6B):E10-9.

41. Wu EQ, Mulani P, Farrell MH, Sleep D. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2007;10(5):408-14.

42. Zhang AY, Fu AZ. Cost-effectiveness of a behavioral intervention for persistent urinary incontinence in prostate cancer patients. Psychooncology. 2016;25(4):421-7.

43. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972;34(2):187-202.

Appendix 8. Systematic review search strategy and PRISMA diagram.

Systematic review search strategy and PRISMA diagram of the literature search conducted to identify studies of preference for prostate cancer treatment.

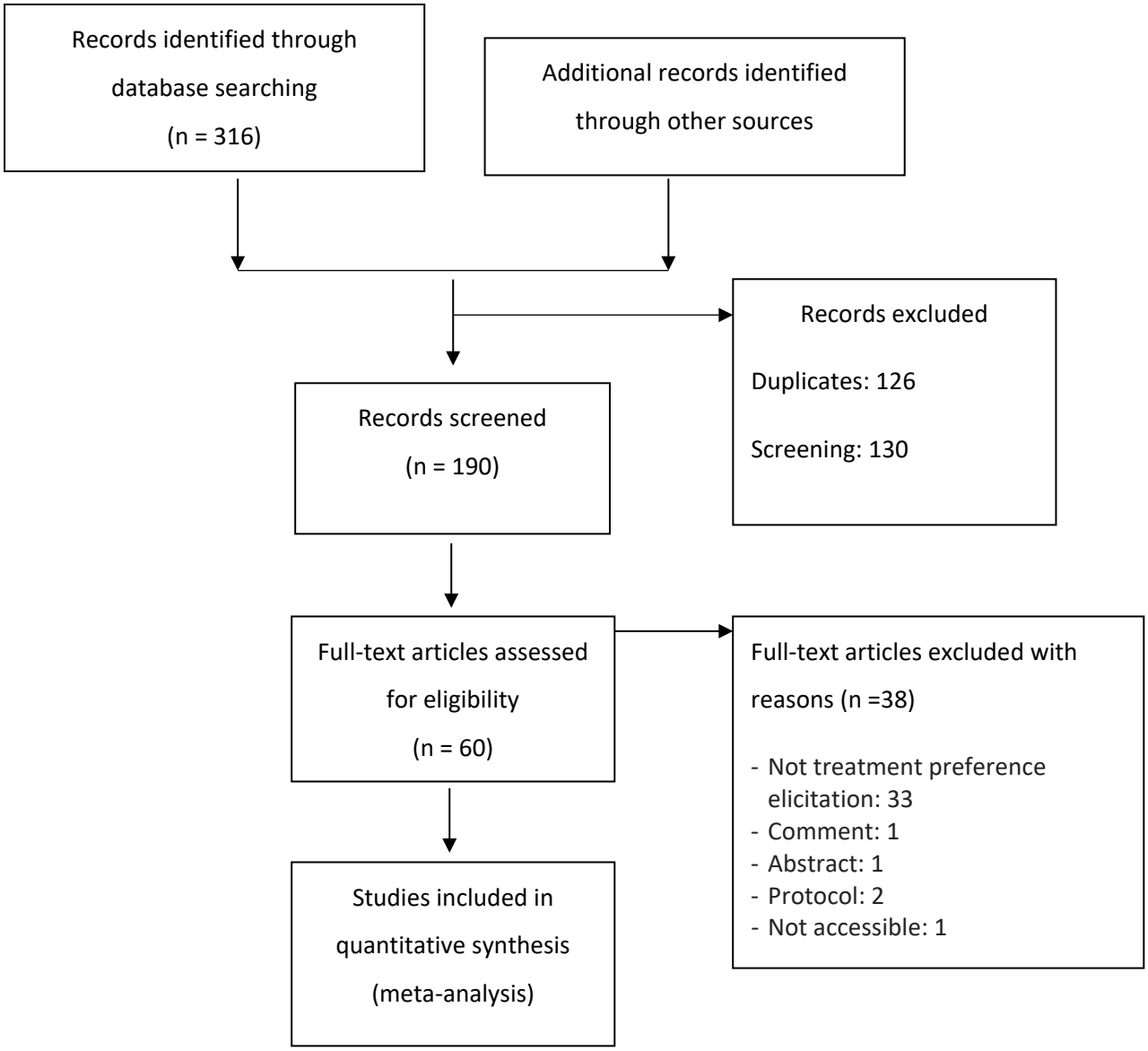
Search strategy:

- Pubmed: (neoplasm prostatic[MeSH Terms]) AND ("Choice Behavior" OR "discrete choice" or "discrete-choice" OR "discrete choice experiment" OR "discrete choice modelling" OR "discrete choice" OR "conjoint experiment" OR "conjoint analysis" or "conjoint measurement" OR "conjoint choice experiment" OR "adaptive conjoint analysis" OR "adaptive choice based conjoint" OR "multi-criteria decision analysis" OR "multicriteria decision analysis" OR "stated preference" OR "choice experiment")

- Cochrane: (Prostatic Neoplasms) explode all trees AND AND ("Choice Behavior" OR "discrete choice" or "discrete-choice" OR "discrete choice experiment" OR "discrete choice modelling" OR "discrete choice" OR "conjoint experiment" OR "conjoint analysis" or "conjoint measurement" OR "conjoint choice experiment" OR "adaptive conjoint analysis" OR "adaptive choice based conjoint" OR "multi-criteria decision analysis" OR "multicriteria decision analysis" OR "stated preference" OR "choice experiment") NOT screening

- EBSCO: CINAHL EconLit MEDLINE with Full Text Psychology and Behavioral Sciences Collection APA PsycInfo APA PsycArticles: ("neoplasm prostatic" OR"prostate cancer" OR prostatae carcinoma" OR "prostate adenocarcinoma") AND ("discrete choice" or "discrete-choice" OR "discrete choice experiment" OR "discrete choice modelling" OR "discrete choice" OR "conjoint experiment" OR "conjoint analysis" or "conjoint measurement" OR "conjoint choice experiment" OR "adaptive conjoint analysis" OR "adaptive choice based conjoint" OR "multi-criteria decision analysis" OR "multicriteria decision analysis" OR "stated preference" OR "choice experiment") NOT screening

Figure 6.8. PRISMA diagram of the systematic review of attributes in PC.



Appendix 9. DCE vignette

Vignette

Imagine that you are a member of a panel providing advice to the Government on which drugs it should subsidise through the Pharmaceutical Benefits Scheme. You have been asked to assess drugs for the treatment of men with localised prostate cancer. Localised prostate cancer is where the cancer is contained within a man's prostate and has not spread elsewhere in his body.

The treatments available for localised prostate cancer differ in a number of areas including:

- value for money or the cost-effectiveness of treatment; this is calculated as the cost required to produce an additional year of full health (which is the absence of any illness, symptoms or negative impacts on quality of life) compared with usual care;
- the extra cost to the Government per year to pay for the drug (given by the cost of the drug times the number of men to be treated);
- the side effects of the treatment;
- the proportion of men who are expected to benefit from treatment;
- the nature of those benefits – including remaining free from prostate cancer being in the blood, prostate cancer progressing to the bone or other parts of the body, remaining free from symptoms of prostate cancer, or the impact on how long men live; and
- the completeness of the data available, including whether it is from an ongoing study and how many years of experience with the treatment are reflected in the available clinical study.

You will now see X hypothetical questions. For each, you will be asked which drug you would support being funded. Each time, please read all of the information on the page before you make your recommendation.

Some of the words are highlighted – hover your mouse or pointer over the highlighted words if you want to see more information about these.

Example Choice Set:

Imagine that you are on a panel providing advice to the Government on which of the following drugs it should subsidise through the Pharmaceutical Benefits Scheme for the treatment of localised prostate cancer. Consider these two options:

Description	Option A	Option B
The benefit to men from treatment is:	The disease does not impact on how long they live.	The disease does not progress to their bones or other parts of the body.
The proportion of men who are likely to benefit from treatment:	30 in 100	20 in 100
The proportion of men who would experience side effects from treatment:	45 in 100	15 in 100
The cost required to produce an additional year of full health compared with usual care is:	$25,000	$75,000
The additional cost per year to the national health budget from funding this drug is:	$200,000,000	$50,000,000
The evidence for the treatment is from a:	Completed study	An ongoing study
The evidence available from the study reflects:	10 years of experience (including 1 year of treatment)	5 years of experience (including 1 year of treatment)

Which option would you recommend funding:

- ☐ Option A
- ☐ Option B

Assuming it was possible, would you support funding the option you recommended?

- ☐ Yes
- ☐ No

Questions to characterise respondents included in the final survey

1. In general, would you say your health is:

- ○ Excellent
- ○ Very good
- ○ Good
- ○ Fair
- ○ Poor

2. How often do you have someone to help you read hospital materials (e.g. letters from your child's doctor or other health related material)?

- ○ Never
- ○ Rarely
- ○ Sometimes
- ○ Often
- ○ Always

3. How confident are you in filing out medical forms by yourself?

- ○ Not at all confident
- ○ A little bit confident
- ○ Somewhat confident
- ○ Quite a bit confident
- ○ Extremely confident

4. What is your sex?

- ○ Male
- ○ Female
- ○ Prefer not to say
- ○ Other _____

5. What is your age? _____

6. Do you identify as Aboriginal and/or Torres Strait Islander origin?

○ Yes

○ No

7. Where do you live?
    - NSW
    - Vic
    - Qld
    - WA
    - NT
    - SA
    - TAS
    - ACT

8. What is the post-code of the area in which you live (specify)?_____

9. Your country of birth:_____ If not Australia, year of arrival to Australia:_____

10. Language(s) other than English spoken at home:
    _____

11. Highest Education/Qualification Completed

○ Finished school before Year 10

○ Year 10 or equivalent

○ Year 12 or equivalent

○ Trade/apprenticeship (e.g. hairdresser, chef) O Certificate/diploma (e.g. childcare, technician)

○ Undergraduate university degree

○ Postgraduate university degree (e.g. Graduate Diploma, Masters, Doctorate, PhD)

12. What is your current relationship status?

- ○ Married
- ○ Not currently, but in a long-term committed relationship
- ○ Separated
- ○ Divorced
- ○ Widowed
- ○ Single

13. What is your current employment status?

- ○ Full time
- ○ Part time
- ○ Casual
- ○ Home duties
- ○ Studying
- ○ Unemployed
- ○ Parental leave

14. In the last financial year, what was your total household income before taxes (include income from wages/salaries, government benefits, pensions, investments and other incomes that might have been received)?

- ○ Negative or zero Income
- ○ $1 - $9,999 per year ($1 - $189 per week)
- ○ $10,000 - $19,999 per year ($190 - $379 per week)
- ○ $20,000 - $29,999 per year ($380 - $579 per week)
- ○ $30,000 - $39,999 per year ($580 - $769 per week)
- ○ $40,000 - $49,999 per year ($770 - $959 per week)
- ○ $50,000 - $59,999 per year ($960 - $1,149 per week)

○ $60,000 - $79,999 per year ($1,150 - $1,529 per week)

○ $80,000 - $99,999 per year ($1,530 - $1,919 per week)

○ $100,000 - $124,999 per year ($1,920 - $2,399 per week)

○ $125,000 - $149,999 per year ($2,400 - $2,879 per week)

○ $150,000 - $199,999 per year ($2,880 - $3,839 per week)

○ $200,000 or more per year ($3,840 or more per week)

○ Declined to answer

○ Don’t know

15. Please indicate which of the following cards you use:

- Health Care Card
- Pensioner Concession Card
- Commonwealth Seniors Card
- Department of Veterans’ Affairs Gold or White Card
- None

16. Have you ever been diagnosed with cancer?
    - No – go to 23
    - Yes – go to 20

17. Which of the following types of cancer have you been diagnosed with (choose all that apply)?
    - Breast cancer
    - Skin cancer
    - Leukaemia or lymphoma
    - Multiple myeloma
    - Prostate cancer
    - Lung cancer
    - Bone cancer
    - Brain cancer
    - Cervix cancer
    - Ovarian cancer
    - Testicular cancer
    - Liver cancer

- Kidney cancer
- Other (specify)

18. Which of the following therapies have you had for your cancer? (Choose all that apply):
    - Radiation therapy
    - Surgery
    - Chemotherapy
    - Hormone therapy
    - Immunotherapy
    - Stem cell transplant
    - Other (specify)

19. Including you, how many people live in your home?
    _____

Follow-up questions in the pilot survey

1. Please fill out the following rating table about the survey.

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
Instructions for the survey were clear.	○	○	○	○	○
The description of prostate cancer was clear.	○	○	○	○	○
The description of prostate cancer and the decision context was relevant to the task of answering the questions.	○	○	○	○	○

2. Please provide any comments you wish to make about those same aspects of the survey.

- ○ The instructions for the survey were clear.
- ○ The description of prostate cancer was clear.
- ○ The description of prostate cancer and the decision context was relevant to the task of answering the questions.

3. Thinking about the 8 questions you just answered, please fill out the following rating table.

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
The language used in the questions was clear.	○	○	○	○	○
The questions were difficult to understand.	○	○	○	○	○
The task was difficult to complete.	○	○	○	○	○

4. Thinking about the 8 questions you just answered, please provide any comments you wish to make.

- ○ The language used in the questions was clear.
- ○ The questions were not difficult to understand.
- ○ The task was not difficult to complete.

5. Thinking about the 8 questions you just answered, please fill out the following rating table on whether the text used was clear.

	Strongly disagree (1)	Disagree (2)	Neither agree nor disagree (3)	Agree (4)	Strongly Agree (5)
Text about the cost required to produce an additional year of full health was clear.	○	○	○	○	○
Text about the proportion of men who would experience side effects was clear.	○	○	○	○	○
Text about the proportion of men likely to benefit from treatment was clear.	○	○	○	○	○
Text about the benefit of treatment was clear.	○	○	○	○	○
Text about the additional cost per year of funding the drug was clear.	○	○	○	○	○
Text about the source of evidence was clear.	○	○	○	○	○
Text about how much evidence is available was clear.	○	○	○	○	○

6. Thinking about the 8 questions you just answered, please provide any comments you wish to make about the text that was used:
    - Text about the cost required to produce an additional year of full health was clear.
    - Text about the proportion of men who would experience side effects was clear.
    - Text about the proportion of men likely to benefit from treatment was clear.
    - Text about the benefit of treatment was clear.
    - Text about the additional cost per year of funding the drug was clear.
    - Text about the source of evidence was clear.
    - Text about how much evidence is available was clear.

7. Thinking about the factors which varied in each of the 8 questions, which one was the most important when choosing between the options you saw?
    - The cost required to produce an additional year of full health.
    - The proportion of men who would experience side effects.
    - The proportion of men likely to benefit from treatment.
    - The benefit of treatment.
    - The additional cost per year of funding the drug.
    - The source of evidence.
    - How much evidence is available.

8. Thinking about the factors which varied in each of the 8 questions, which one was the least important when choosing between the options you saw?
    - The cost required to produce an additional year of full health.
    - The proportion of men who would experience side effects.
    - The proportion of men likely to benefit from treatment.
    - The benefit of treatment.
    - The additional cost per year of funding the drug.
    - The source of evidence.
    - How much evidence is available.

9. Thinking about the questions you just answered, how did you interpret the factor related to the cost required to produce an additional year of full health?

- ○ As the cost to the individual being treated for each additional year of full health.
- ○ As the cost to the individual being treated for each year of treatment.
- ○ As the cost to the government for each additional year of full health.
- ○ As the cost to the government for each additional year of treatment.

10. Thinking about the questions you just answered and the factor related to how much evidence was available, how did you interpret the phrase “years of experience”:

- ○ As how long each man in the study was treated with the drug
- ○ As the total period of time over which information had been collected as part of the study

Appendix 10. Modifications to DCE after steering committee discussion

1- Value for money

a. Descriptions provided in the background information (setting) text:

i. Change the definition of cost-effectiveness:

Current	**Suggested**
Value for money or the cost-effectiveness of treatment; this is calculated as the cost to the Government of the treatment required to produce an additional year of full health (which is the absence of any illness, symptoms or negative impacts on quality of life) compared with usual care.	Value for money of the new treatment compared to usual care: This is the extra cost <u>to the government</u> to produce an extra year of full health in those with prostate cancer. Full health means no illness, symptoms or negative impacts on quality of life. Value for money is also known as the *cost-effectiveness* of the new treatment. A lower cost per year of full health means it is better value for money.

b. Attribute descriptor text

i. To underline the word to the Government in the descriptor:

Current	**Suggested**
The cost to the Government required to produce an additional year of full health compared with usual care is	The extra cost <u>to the Government</u> required to produce an additional year of full health compared with usual care is

ii. To include a hover text in the descriptor on extra cost to address the directionality of the attribute:

Hover text
Extra cost: A lower cost per year of full health means it is better value for money.

c. Level text

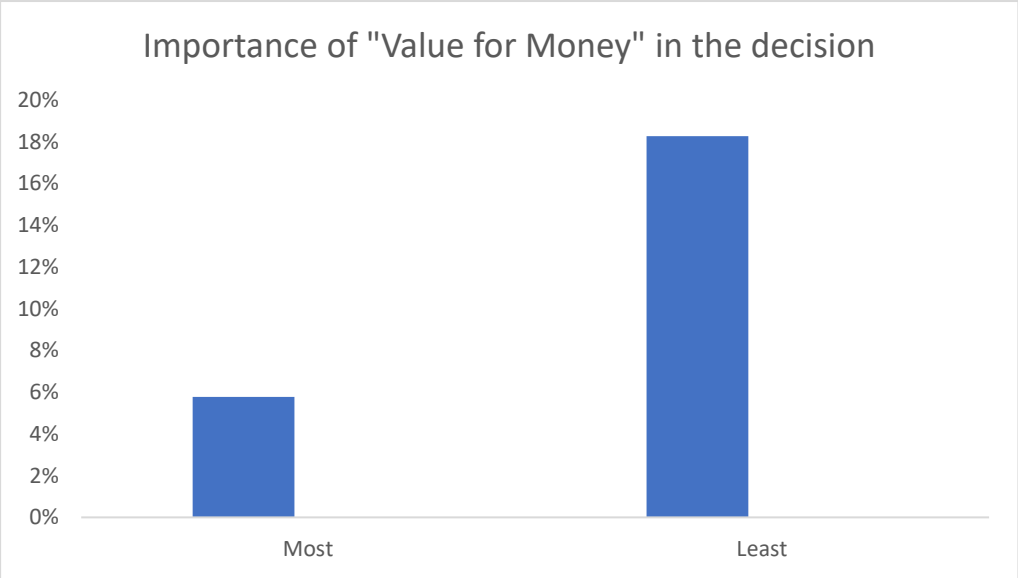
i. Include the units in the level

Current	**Suggested**
$75,000	$75,000 per year of full health

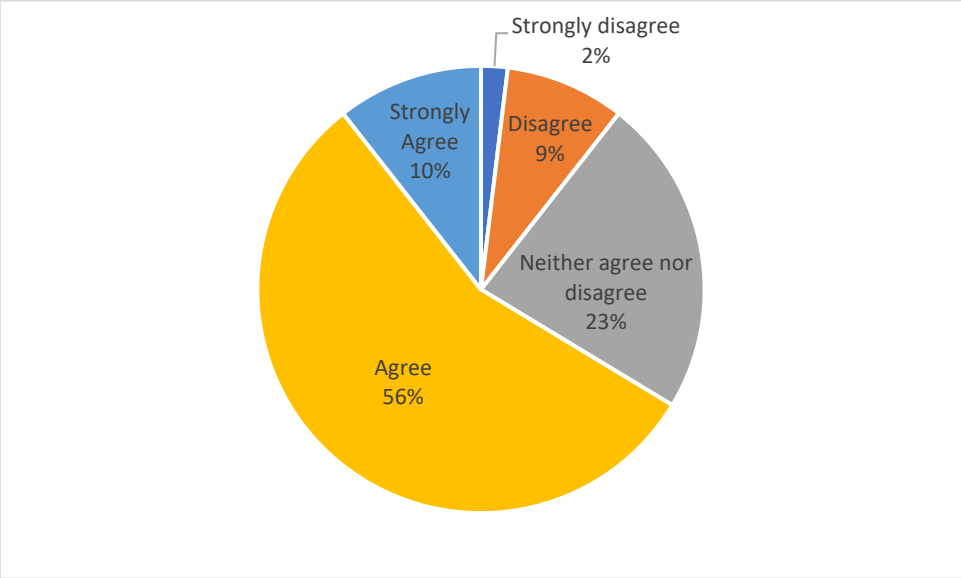
d. Comments and ranking of “Value for Money” attribute:

i. Ranking:

- From the 104 respondents 6 (5,77%) mentioned this attribute as the most important when choosing between the 2 options of treatments.
- From the 104 respondents 19 (18,27%) mentioned this attribute as the least important when choosing between the 2 options of treatments.

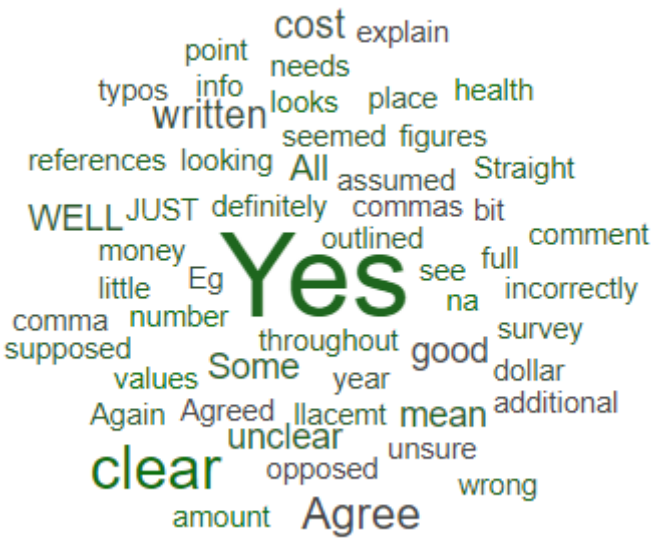


ii. When respondents were asked if the text to describe the cost required to produce an additional year of full health was clear, the majority (66%) at least agree the text was clear:



iii. Comments

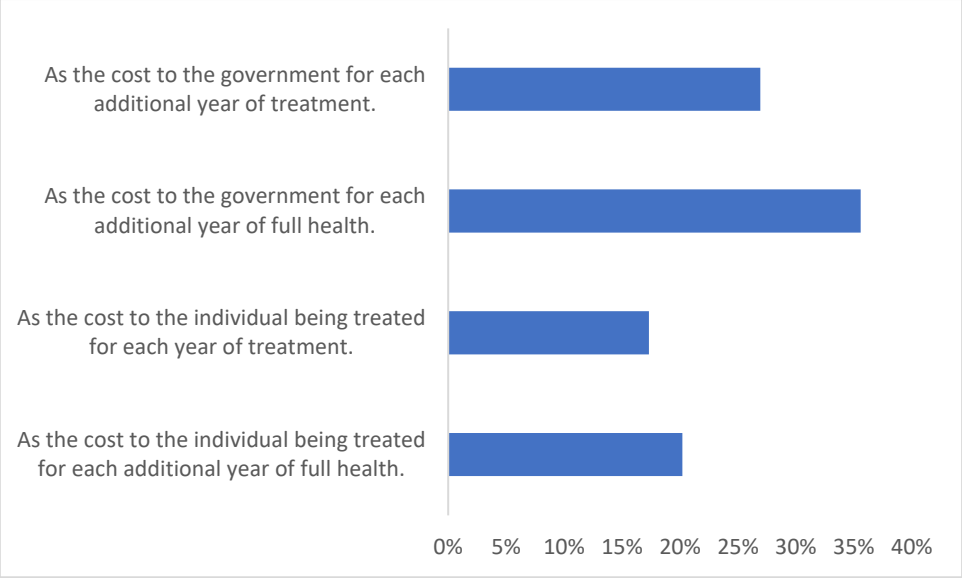
- When respondents were asked to provide feedback about the text that was used (text about the cost required to produce an additional year of full health was clear) majority of the 41 comments received were positive. The word cloud below was developed with the answers received from the question: was the text about the cost required to produce an additional year of full health clear?.



- Negative comments received were more related to how the numbers were displayed on the survey with only few respondents replaying a negative answer. Negative comments:

Comma misplacement	**Not specific**	**Misunderstanding**
there were typos in some figures (e.g. 1,00,000 -- I assumed this was 100,000 but it was a little unclear)	No	Yes i could see the amount of money it would cost.
The references to a cost of $1,00,000 were very unclear. Was that supposed to mean $100,000 or $1,000,000 ?	A bit unsure clear	
JUST WAS NOT WRITTEN WELL	Not	
Some of the dollar values seemed to written incorrectly		
Again- comma llacemt needs looking at. Eg 10,0000. Also- explain what they mean by 'additional year of full health'- as opposed to what?		
The commas were in the wrong place with the number written: $1,00,000		

iv. Probing questions: When respondents were asked how did they interpret the factor related to the cost required to produce an additional year of full health, 63% of respondents understood the cost accruing to the government and 35% of respondents correctly understood the question. 37.5% of respondents interpreted the cost as accruing to the individual:



Conclusion:

- A sense of directionality in the descriptor will help respondents to compare the choice options.
- Adding the unit "year of full health" in the level will help respondents to differentiate between the attributes value for money and cost of treatment.
- Underlining the word "to the government" will help respondents to understand attribute context.

2- Source of evidence is not reflecting the uncertainty in the evidence and is possible given a wrong message about timing.

a. Change the hover text to:

Level	Current	Suggested
Completed study	Completed study: patients are no longer being examined or treated (that is, the last patient's last visit has occurred).	Completed study: All patients have completed the study and the final results have been reported.
An ongoing study	Ongoing study: some patients are still receiving an intervention or being examined.	Ongoing study: Some patients are continuing to receive treatment as part of the study, but some results are available.

Bibliography

1. Lothgren M, Hechmati G, Jacobs I, Qian Y, Arellano J. Modeling lifetime effectiveness of denosumab versus placebo in men with castration resistant prostate cancer (CRPC) at high risk of developing bone metastasis (BM). Value in Health. 2012;15(7):A410.

2. Mathurin K, Beauchemin C, Lachaine J. Development of a global economic model to evaluate the cost-effectiveness of targeted treatments using companion diagnostics in advanced/metastatic cancer treatment. Value in Health. 2014;17(7):A561.

3. Bergman J, Logan S, Fink A, Ganz DA, Peterson MA, Litwin MS. Caring for the Uninsured with Prostate Cancer: A Comparison of Four Policy Alternatives in California. J Community Health. 2010;35(1):18-26.

4. Yun JE, Park DA, Jh A, Oh SH, Lee NR, Jm K, et al. [The economic analysis of robotic surgery for prostate cancer]. Seoul: National Evidence-based Healthcare Collaborating Agency (NECA). 2014.

5. Matsumoto T, Hatakeyama S, Ookubo T, Mitsuzuka K, Narita S, Inoue T, et al. Cost-effectiveness comparison between neoadjuvant chemohormonal therapy and extended pelvic lymph node dissection in high-risk prostate cancer patients treated with radical prostatectomy: a multi-institutional analysis. Medical oncology (Northwood, London, England). 2017;34(12):190.

6. Schapira DV, Jarrett AR. Prostate-cancer surveillance: a cost-effective strategy. International Journal of Oncology. 2006;2(3):100.

7. Veestraeten D. An alternative approach to modelling relapse in cancer with an application to adenocarcinoma of the prostate. Mathematical Biosciences. 2006;199(1):38-54.

8. Fraundorfer MR, Gilling PJ, Kennett KM, Dunton NG. Holmium laser resection of the prostate is more cost effective than transurethral resection of the prostate: results of a randomized prospective study. Urology. 2001;57(3):454-8.

9. Jackson T, Street A, Costello A, Crowe H. Cost-effectiveness of laser ablation of the prostate. Premature evaluation. Int J Technol Assess Health Care. 1995;11(3):595-610.

10. Showalter TN, Foley KA, Jutkowitz E, Lallas CD, Trabulsi EJ, Gomella LG, et al. Costs of early adjuvant radiation therapy after radical prostatectomy: a decision analysis. Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23(3):701-6.

11. Carter HE, Martin A, Schofield D, Duchesne G, Haworth A, Hornby C, et al. A decision model to estimate the cost-effectiveness of intensity modulated radiation therapy (IMRT) compared to three dimensional conformal radiation therapy (3DCRT) in patients receiving radiotherapy to the prostate bed. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2014;112(2):187-93.

12. Hummel S, Paisley S, Morgan A, Currie E, Brewer N. Clinical and cost-effectiveness of new and emerging technologies for early localised prostate cancer: a systematic review. Health Technol Assess. 2003;7(33):iii, ix-x, 1-157.

13. Konski A. Cost effectiveness of prostate cancer radiotherapy. Translational Andrology and Urology. 2018;7(3):371-7.

14. Bayoumi AM, Brown AD, Garber AM. Cost-effectiveness of androgen suppression therapies in advanced prostate cancer. Journal of the National Cancer Institute. 2000;92(21):1731-9.

15. Bloomfield DJ. Should bisphosphonates be part of the standard therapy of patients with multiple myeloma or bone metastases from other cancers: an evidence-based review. Journal of Clinical Oncology. 1998;16(3):1218-25.

16. Buron C, Le Vu B, Cosset JM, Pommier P, Peiffert D, Delannes M, et al. Brachytherapy versus prostatectomy in localized prostate cancer: results of a French multicenter prospective medico-economic study. Int J Radiat Oncol Biol Phys. 2007;67(3):812-22.

17. Close A, Robertson C, Rushton S, Shirley M, Vale L, Ramsay C, et al. Comparative cost-effectiveness of robot-assisted and standard laparoscopic prostatectomy as alternatives to open radical prostatectomy for treatment of men with localised prostate cancer: a health technology assessment from the perspective of the UK National Health Service. European urology. 2013;64(3):361-9.

18. de Carvalho TM, Heijnsdijk EAM, de Koning HJ. When should active surveillance for prostate cancer stop if no progression is detected? The Prostate. 2017;77(9):962-9.

19. Garcia de Paredes Esteban JC, Alegre Del Rey EJ, Asensi Diez R. Docetaxel in hormone-sensitive advanced prostate cancer; GENESIS-SEFH evaluation reporta. Farmacia hospitalaria : organo oficial de expresion cientifica de la Sociedad Espanola de Farmacia Hospitalaria. 2017;41(4):550-8.

20. Hatoum HT, Crawford ED, Nielsen SK, Lin SJ, Marshall DC. Cost-effectiveness analysis comparing degarelix with leuprolide in hormonal therapy for patients with locally advanced prostate cancer. Expert review of pharmacoeconomics & outcomes research. 2013;13(2):261-70.

21. Hayes JH, Ollendorf DA, Pearson SD, Barry MJ, Kantoff PW, Lee PA, et al. Observation versus initial treatment for men with localized, low-risk prostate cancer: a cost-effectiveness analysis. Annals of internal medicine. 2013;158(12):853-60.

22. Helou J, Torres S, Musunuru HB, Raphael J, Cheung P, Vesprini D, et al. Stereotactic Body Radiotherapy versus Low Dose Rate Brachytherapy for Localised Prostate Cancer: a Cost–Utility Analysis. Clinical Oncology. 2017;29(11):718-31.

23. Hodges JC, Lotan Y, Boike TP, Benton R, Barrier A, Timmerman RD. Cost-effectiveness analysis of SBRT versus IMRT: an emerging initial radiation treatment option for organ-confined prostate cancer. Am J Manag Care. 2012;18(5):e186-93.

24. Hohwu L, Borre M, Ehlers L, Venborg Pedersen K. A short-term cost-effectiveness study comparing robot-assisted laparoscopic and open retropubic radical prostatectomy. J Med Econ. 2011;14(4):403-9.

25. Holko P, Kawalec P. Economic evaluation of sipuleucel-T immunotherapy in castration-resistant prostate cancer. Expert Rev Anticancer Ther. 2014;14(1):63-73.

26. Hummel SR, Stevenson MD, Simpson EL, Staffurth J. A model of the cost-effectiveness of intensity-modulated radiotherapy in comparison with three-dimensional conformal radiotherapy for the treatment of localised prostate cancer. Clinical oncology (Royal College of Radiologists (Great Britain)). 2012;24(10):e159-67.

27. Koerber F, Waidelich R, Stollenwerk B, Rogowski W. The cost-utility of open prostatectomy compared with active surveillance in early localised prostate cancer. BMC health services research. 2014;14(1):163.

28. Konski A. Cost-effectiveness of intensity-modulated radiation therapy. Expert review of pharmacoeconomics & outcomes research. 2005;5(2):137-40.

29. Konski A, Sherman E, Krahn M, Bremner K, Beck JR, Watkins-Bruner D, et al. Economic analysis of a phase III clinical trial evaluating the addition of total androgen suppression to radiation versus radiation alone for locally advanced prostate cancer (Radiation Therapy Oncology Group protocol 86-10). International journal of radiation oncology, biology, physics. 2005;63(3):788-94.

30. Konski A, Speier W, Hanlon A, Beck JR, Pollack A. Is proton beam therapy cost effective in the treatment of adenocarcinoma of the prostate? Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25(24):3603-8.

31. Konski A, Watkins-Bruner D, Feigenberg S, Hanlon A, Kulkarni S, Beck JR, et al. Using decision analysis to determine the cost-effectiveness of intensity-modulated radiation therapy in the treatment of intermediate risk prostate cancer. Int J Radiat Oncol Biol Phys. 2006;66(2):408-15.

32. Konski A, Watkins-Bruner D, Brereton H, Feigenberg S, Hanks G. Long-term hormone therapy and radiation is cost-effective for patients with locally advanced prostate carcinoma. Cancer. 2006;106(1):51-7.

33. Lazzaro C, Bartoletti R, Guazzoni G, Orestano F, Pappagallo GL, Prezioso D, et al. Economic evaluation of different hormonal therapies for prostate cancer. Final results from the Quality of Life Antiandrogen Blockade Italian Observational Study (QuABIOS). Archivio italiano di urologia, andrologia : organo ufficiale [di] Societa italiana di ecografia urologica e nefrologica. 2007;79(3):104-7.

34. Lee D, Porter J, Gladwell D, Brereton N, Nielsen SK. A cost-utility analysis of degarelix in the treatment of advanced hormone-dependent prostate cancer in the United Kingdom. J Med Econ. 2014;17(4):233-47.

35. Lyth J, Andersson SO, Andren O, Johansson JE, Carlsson P, Shahsavar N. A decision support model for cost-effectiveness of radical prostatectomy in localized prostate cancer. Scand J Urol Nephrol. 2012;46(1):19-25.

36. Ollendorf DA, Hayes J, McMahon P, Kuba M, Tramontano A, Pearson SD. Brachytherapy/proton beam therapy for clinically localized, low-risk prostate cancer. Boston: Institute for Clinical and Economic Review (ICER). 2008:114.

37. Ollendorf DA, Hayes J, McMahon P, Pearson SD. Active surveillance and radical prostatectomy for clinically localized, low-risk prostate cancer. Boston: Institute for Clinical and Economic Review (ICER). 2009.

38. Organ M, Wood L, Wilke D, Skedgel C, Cheng T, North S, et al. Intermittent LHRH therapy in the management of castrate-resistant prostate cancer (CRPCa): results of a multi-institutional randomized prospective clinical trial. American journal of clinical oncology. 2013;36(6):601-5.

39. Parthan A, Pruttivarasin N, Davies D, Taylor DC, Pawar V, Bijlani A, et al. Comparative cost-effectiveness of stereotactic body radiation therapy versus intensity-modulated and proton radiation therapy for localized prostate cancer. Front Oncol. 2012;2(2 Aug):81.

40. Perez CA, Michalski J, Ballard S, Drzymala R, Kobeissi BJ, Lockett MA, et al. Cost benefit of emerging technology in localized carcinoma of the prostate. International Journal of Radiation Oncology Biology Physics. 1997;39(4):875-83.

41. Peters M, Piena MA, Steuten LMG, Van Der Voort Van Zyp JRN, Moerland MA, Van Vulpen M. Comparative cost-effectiveness of focal and total salvage 125I brachytherapy for recurrent prostate cancer after primary radiotherapy. Journal of Contemporary Brachytherapy. 2016;8(6):484-91.

42. Ramsay C, Pickard R, Robertson C, Close A, Vale L, Armstrong N, et al. Systematic review and economic modelling of the relative clinical benefit and cost-effectiveness of laparoscopic surgery and robotic surgery for removal of the prostate in men with localised prostate cancer. Health Technol Assess. 2012;16(41):1-313.

43. Ramsay CR, Adewuyi TE, Gray J, Hislop J, Shirley MD, Jayakody S, et al. Ablative therapy for people with localised prostate cancer: A systematic review and economic evaluation. Health Technology Assessment. 2015;19(49):1-8.

44. Ramsey S, Veenstra D, Clarke L, Gandhi S, Hirsch M, Penson D. Is combined androgen blockade with bicalutamide cost-effective compared with combined androgen blockade with flutamide. Urology. 2005;66(4):835-9.

45. Samant RS, Dunscombe PB, Roberts GH. A cost-outcome analysis of long-term adjuvant goserelin in addition to radiotherapy for locally advanced prostate cancer. Urol Oncol. 2003;21(3):171-7.

46. Sanyal C, Aprikian AG, Cury FL, Chevalier S, Dragomir A. Management of localized and advanced prostate cancer in Canada: A lifetime cost and quality-adjusted life-year analysis. Cancer. 2016;122(7):1085-96.

47. Seidenfeld J, Samson DJ, Aronson N, Albertson PC, Bayoumi AM, Bennett C, et al. Relative effectiveness and cost-effectiveness of methods of androgen suppression in the treatment of advanced prostate cancer. Evidence report/technology assessment (Summary). 1999(4):i-x, 1-246, I1-36, passim.

48. Shah C, Lanni TB, Jr., Ghilezan MI, Gustafson GS, Marvin KS, Ye H, et al. Brachytherapy provides comparable outcomes and improved cost-effectiveness in the treatment of low/intermediate prostate cancer. Brachytherapy. 2012;11(6):441-5.

49. Sher DJ, Parikh RB, Mays-Jackson S, Punglia RS. Cost-effectiveness analysis of SBRT versus IMRT for low-risk prostate cancer. Am J Clin Oncol. 2014;37(3):215-21.

50. Wex J, Sidhu M, Odeyemi I, Abou-Setta AM, Retsa P, Tombal B. Leuprolide acetate 1-, 3- and 6-monthly depot formulations in androgen deprivation therapy for prostate cancer in nine European countries: evidence review and economic evaluation. Clinicoecon Outcomes Res. 2013;5:257-69.

51. Yong JH, Beca J, McGowan T, Bremner KE, Warde P, Hoch JS. Cost-effectiveness of intensity-modulated radiotherapy in prostate cancer. Clinical Oncology. 2012;24(7):521-31.

52. Evans SR. Fundamentals of clinical trial design. J Exp Stroke Transl Med. 2010;3(1):19-27.

53. Royle K-L, Meads D, Visser-Rogers JK, White IR, Cairns DA. How is overall survival assessed in randomised clinical trials in cancer and are subsequent treatment lines considered? A systematic review. Trials. 2023;24(1):708.

54. Chakraborty R, Efficace F. Importance of quality of life in early phase clinical trials: A case study of selinexor in multiple myeloma. British Journal of Haematology. 2020;189(3):e112-e3.

55. Kemp R, Prasad V. Surrogate endpoints in oncology: when are they acceptable for regulatory and clinical decisions, and are they currently overused? BMC Med. 2017;15(1):134.

56. Cappelleri JC, Zou KH, Bushmakin AG, Alvir JMJ, Alemayehu D, Symonds T. Patient-reported outcomes: measurement, implementation and interpretation: Crc Press; 2013.

57. Johannesson M, Jönsson B, Karlsson G. Outcome measurement in economic evaluation. Health Economics. 1996;5(4):279-96.

58. Burzykowski T, Molenberghs G, Buyse M. The Evaluation of Surrogate Endpoints. New York: Springer-Verlag; 2005.

59. Naci H, Wouters OJ, Gupta R, Ioannidis JPA. Timing and Characteristics of Cumulative Evidence Available on Novel Therapeutic Agents Receiving Food and Drug Administration Accelerated Approval. Milbank Q. 2017;95(2):261-90.

60. Kim C, Prasad V. Cancer Drugs Approved on the Basis of a Surrogate End Point and Subsequent Overall Survival: An Analysis of 5 Years of US Food and Drug Administration Approvals. JAMA internal medicine. 2015;175(12):1992-4.

61. Group IEEW. ICH Harmonised Tripartite Guideline. Statistical principles for clinical trials. International Conference on Harmonisation E9 Expert Working Group. Stat Med. 1999;18(15):1905-42.

62. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89-95.

63. Health AGDo. PBAC Guidelines | Appendix 5 Translating comparative treatment effects of proposed surrogate measures to target clinical outcomes. Australian Government Department of Health; 2016.

64. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics. 2000;1(1):49-67.

65. Grigore B, Ciani O, Dams F, Federici C, de Groot S, Möllenkamp M, et al. Surrogate Endpoints in Health Technology Assessment: An International Review of Methodological Guidelines. PharmacoEconomics. 2020;38(10):1055-70.

66. Ghijben P, Gu Y, Lancsar E, Zavarsek S. Revealed and Stated Preferences of Decision Makers for Priority Setting in Health Technology Assessment: A Systematic Review. Pharmacoeconomics. 2018;36(3):323-40.

67. Solà-Morales O, Volmer T, Mantovani L. Perspectives to mitigate payer uncertainty in health technology assessment of novel oncology drugs. Journal of Market Access & Health Policy. 2019;7(1):1562861.

68. Pan Y, Dieck G, Stemhagen A. Does the use of a surrogate outcome in solid state oncology htas decrease the chances of positive recommendations? Value in Health. 2014;17:A1-A295.

69. Ciani O, Buyse M, Drummond M, Rasi G, Saad ED, Taylor RS. Time to Review the Role of Surrogate End Points in Health Policy: State of the Art and the Way Forward. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2017;20(3):487-95.

70. Cardiac Arrhythmia Suppression Trial I. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. The New England journal of medicine. 1989;321(6):406-12.

71. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med. 1989;8(4):431-40.

72. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. Stat Med. 1992;11(2):167-78.

73. Buyse M, Molenberghs G. Criteria for the validation of surrogate endpoints in randomized experiments. Biometrics. 1998;54(3):1014-29.

74. Surrogate to Final Outcome Working Group. Report of the Surrogate to Final Outcome Working Group to the Pharmaceutical Benefits Advisory Committee: a framework for evaluating proposed surrogate measures and their use in submissions to PBAC.; 2008.

75. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ (Clinical research ed). 1996;312(7023):71-2.

76. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the Economic Evaluation of Health Care Programmes: Oxford University Press; 2005.

77. Incerti D, Thom H, Baio G, Jansen JP. R You Still Using Excel? The Advantages of Modern Software Tools for Health Technology Assessment. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2019;22(5):575-9.

78. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices--overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Med Decis Making. 2012;32(5):667-77.

79. Smare C, Lakhdari K, Doan J, Posnett J, Johal S. Evaluating Partitioned Survival and Markov Decision-Analytic Modeling Approaches for Use in Cost-Effectiveness Analysis: Estimating and Comparing Survival Outcomes. PharmacoEconomics. 2020;38(1):97-108.

80. Briggs AH, Claxton K, Sculpher MJ. Decision Modelling for Health Economic Evaluation: Oxford University Press; 2006.

81. Weinstein MC, Torrance G, McGuire A. QALYs: The Basics. Value in Health. 2009;12:S5-S9.

82. Health Technology Assessment, Consumers and Communities [Internet]. Australian Government Department of Health and Ageing. 2019 [cited 2019-12-11]. Available from: https://www1.health.gov.au/internet/hta/publishing.nsf/Content/hta-1.

83. Pharmaceutical Benefits Scheme. Pharmaceutical Benefits Advisory Committee (PBAC) Membership 2019 [updated 21 June 2019. Available from: https://www.pbs.gov.au/info/industry/listing/participants/pbac.

84. Popiolek M, Rider JR, Andren O, Andersson SO, Holmberg L, Adami HO, et al. Natural history of early, localized prostate cancer: a final report from three decades of follow-up. European urology. 2013;63(3):428-35.

85. Sweeney C, Nakabayashi M, Regan M, Xie W, Hayes J, Keating N, et al. The Development of Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP). Journal of the National Cancer Institute. 2015;107(12):djv261.

86. Australian Institute of Health and Welfare. Cancer in Australia. Canberra: AIHW; 2022.

87. Roehrborn CG, Black LK. The economic burden of prostate cancer. BJU international. 2011;108(6):806-13.

88. Sutcliffe S, Pontari MA. Chapter 2 - Inflammation and Infection in the Etiology of Prostate Cancer. In: Mydlo JH, Godec CJ, editors. Prostate Cancer (Second Edition). San Diego: Academic Press; 2016. p. 13-20.

89. Rawla P. Epidemiology of Prostate Cancer. World J Oncol. 2019;10(2):63-89.

90. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. Int J Cancer. 2010;127(12):2893-917.

91. Australian Institute of Health and Welfare. GRIM (General Record of Incidence of Mortality) books 2016: Prostate cancer. Canberra: AIHW.; 2016.

92. Evans S.M. TG, Brooks A, Currow D., Davis I.D., Delprado W., Frizelle F, Frydenberg M., Heathcote P., James E., Marks S., Mcneil J.J., Miller C, Millar J.L., Moretti K., Pryor D., Roder D., Skala M., Smith D., Villanti P., Walker A., White C. Prostate Cancer Outcomes Registry-Australia and New Zealand Report 2018. Reporting on data 2015–2016. 2018.

93. Pietzak E, Mucksavage P. Chapter 53 - Bone Health in Prostate Cancer. In: Mydlo JH, Godec CJ, editors. Prostate Cancer (Second Edition). San Diego: Academic Press; 2016. p. 491-507.

94. Bill-Axelson A, Holmberg L, Ruutu M, Garmo H, Stark JR, Busch C, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. The New England journal of medicine. 2011;364(18):1708-17.

95. Johansson E, Bill-Axelson A, Holmberg L, Onelov E, Johansson JE, Steineck G, et al. Time, symptom burden, androgen deprivation, and self-assessed quality of life after radical prostatectomy or watchful waiting: the Randomized Scandinavian Prostate Cancer Group Study Number 4 (SPCG-4) clinical trial. European urology. 2009;55(2):422-30.

96. Gupta GP, Massague J. Cancer metastasis: building a framework. Cell. 2006;127(4):679-95.

97. Aus G, Robinson D, Rosell J, Sandblom G, Varenhorst E, South-East Region Prostate Cancer G. Survival in prostate carcinoma--outcomes from a prospective, population-based cohort of 8887 men with up to 15 years of follow-up: results from three countries in the population-based National Prostate Cancer Registry of Sweden. Cancer. 2005;103(5):943-51.

98. Grozescu T, Popa F. Prostate cancer between prognosis and adequate/proper therapy. J Med Life. 2017;10(1):5-12.

99. Johnston TJ, Shaw GL, Lamb AD, Parashar D, Greenberg D, Xiong T, et al. Mortality Among Men with Advanced Prostate Cancer Excluded from the ProtecT Trial. European urology. 2017;71(3):381-8.

100. Australia PCFoAaCC. Prostate Cancer Foundation of Australia and Cancer Council Australia PSA Testing Guidelines Expert Advisory Panel. Draft clinical practice guidelines for PSA testing and early management of test-detected prostate cancer. 2016.

101. Pezaro C, Woo HH, Davis ID. Prostate cancer: measuring PSA. Intern Med J. 2014;44(5):433-40.

102. Punglia RS, D'Amico AV, Catalona WJ, Roehl KA, Kuntz KM. Effect of Verification Bias on Screening for Prostate Cancer by Measurement of Prostate-Specific Antigen. New England Journal of Medicine. 2003;349(4):335-42.

103. Rodrigues G, Warde P, Pickles T, Crook J, Brundage M, Souhami L, et al. Pre-treatment risk stratification of prostate cancer patients: A critical review. Can Urol Assoc J. 2012;6(2):121-7.

104. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. European urology. 2017;71(4):618-29.

105. Rodrigues G, Warde P, Pickles T, Crook J, Brundage M, Souhami L, et al. Pre-treatment risk stratification of prostate cancer patients: A critical review. Canadian Urological Association journal = Journal de l'Association des urologues du Canada. 2012;6(2):121-7.

106. McKay RR, Feng FY, Wang AY, Wallis CJD, Moses KA. Recent Advances in the Management of High-Risk Localized Prostate Cancer: Local Therapy, Systemic Therapy, and Biomarkers to

Guide Treatment Decisions. American Society of Clinical Oncology Educational Book. 2020(40):e241-e52.

107. Venderbos LDF, Aluwini S, Roobol MJ, Bokhorst LP, Oomens E, Bangma CH, et al. Long-term follow-up after active surveillance or curative treatment: quality-of-life outcomes of men with low-risk prostate cancer. Qual Life Res. 2017;26(6):1635-45.

108. Mohler JL, Antonarakis ES, Armstrong AJ, D'Amico AV, Davis BJ, Dorff T, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. Journal of the National Comprehensive Cancer Network : JNCCN. 2019;17(5):479-505.

109. Garibaldi C, Jereczek-Fossa BA, Marvaso G, Dicuonzo S, Rojas DP, Cattani F, et al. Recent advances in radiation oncology. Ecancermedicalscience. 2017;11:785-.

110. Sasse AD, Sasse E, Carvalho AM, Macedo LT. Androgenic suppression combined with radiotherapy for the treatment of prostate adenocarcinoma: a systematic review. BMC Cancer. 2012;12:54.

111. Perera M, Roberts MJ, Klotz L, Higano CS, Papa N, Sengupta S, et al. Intermittent versus continuous androgen deprivation therapy for advanced prostate cancer. Nat Rev Urol. 2020;17(8):469-81.

112. Zhang X, Wu H, Chen C, Zhuang H, Tian J. Optimal duration of androgen deprivation therapy for moderate and high risk prostate cancer after radiotherapy: A Meta-analysis. Tumor. 2019;39(5):388-97.

113. Virgo KS, Basch E, Loblaw DA, Oliver TK, Rumble RB, Carducci MA, et al. Second-Line Hormonal Therapy for Men With Chemotherapy-Naïve, Castration-Resistant Prostate Cancer: American Society of Clinical Oncology Provisional Clinical Opinion. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(17):1952-64.

114. Cornford P, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer. Part II-2020 Update: Treatment of Relapsing and Metastatic Prostate Cancer. European urology. 2021;79(2):263-82.

115. Teo MY, Rathkopf DE, Kantoff P. Treatment of Advanced Prostate Cancer. Annu Rev Med. 2019;70:479-99.

116. Nader R, El Amm J, Aragon-Ching JB. Role of chemotherapy in prostate cancer. Asian journal of andrology. 2018;20(3):221-9.

117. Wallis CJD, Saskin R, Choo R, Herschorn S, Kodama RT, Satkunasivam R, et al. Surgery Versus Radiotherapy for Clinically-localized Prostate Cancer: A Systematic Review and Meta-analysis. European urology. 2016;70(1):21-30.

118. Petrelli F, Vavassori I, Coinu A, Borgonovo K, Sarti E, Barni S. Radical prostatectomy or radiotherapy in high-risk prostate cancer: a systematic review and metaanalysis. Clinical genitourinary cancer. 2014;12(4):215-24.

119. Lei JH, Liu LR, Wei Q, Yan SB, Song TR, Lin FS, et al. Systematic review and meta-analysis of the survival outcomes of first-line treatment options in high-risk prostate cancer. Sci Rep. 2015;5:7713.

120. Jayadevappa R, Chhatre S, Wong YN, Wittink MN, Cook R, Morales KH, et al. Comparative effectiveness of prostate cancer treatments for patient-centered outcomes: A systematic review and meta-analysis (PRISMA Compliant). Medicine (Baltimore). 2017;96(18):e6790.

121. Fahmy O, Khairul-Asri MG, Hadi SHSM, Gakis G, Stenzl A. The Role of Radical Prostatectomy and Radiotherapy in Treatment of Locally Advanced Prostate Cancer: A Systematic Review and Meta-Analysis. Urologia internationalis. 2017;99(3):249-56.

122. Chen L, Li Q, Wang Y, Zhang Y, Ma X. Comparison on efficacy of radical prostatectomy versus external beam radiotherapy for the treatment of localized prostate cancer. Oncotarget. 2017;8(45):79854-63.

123. Alva A, Hussain M. Intermittent androgen deprivation therapy in advanced prostate cancer. Current Treatment Options in Oncology. 2014;15(1):127-36.

124. Mottet N, van den Bergh RCN, Briers E, Van den Broeck T, Cumberbatch MG, De Santis M, et al. EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer-2020 Update. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. European urology. 2021;79(2):243-62.

125. Attard G, Murphy L, Clarke NW, Cross W, Jones RJ, Parker CC, et al. Abiraterone acetate and prednisolone with or without enzalutamide for high-risk non-metastatic prostate cancer: a meta-analysis of primary results from two randomised controlled phase 3 trials of the STAMPEDE platform protocol. The Lancet. 2022;399(10323):447-60.

126. Rush HL, Murphy L, Morgans AK, Clarke NW, Cook AD, Attard G, et al. Quality of Life in Men With Prostate Cancer Randomly Allocated to Receive Docetaxel or Abiraterone in the STAMPEDE Trial. Journal of Clinical Oncology. 2021;40(8):825-36.

127. Denham JW, Steigler A, Wilcox C, Lamb DS, Joseph D, Atkinson C, et al. Time to biochemical failure and prostate-specific antigen doubling time as surrogates for prostate cancer-specific mortality: evidence from the TROG 96.01 randomised controlled trial. The Lancet Oncology. 2008;9(11):1058-68.

128. D'Amico AV, Chen MH, de Castro M, Loffredo M, Lamb DS, Steigler A, et al. Surrogate endpoints for prostate cancer-specific mortality after radiotherapy and androgen suppression therapy in men with localised or locally advanced prostate cancer: an analysis of two randomised trials. The Lancet Oncology. 2012;13(2):189-95.

129. Ray ME, Bae K, Hussain MHA, Hanks GE, Shipley WU, Sandler HM. Potential Surrogate Endpoints for Prostate Cancer Survival: Analysis of a Phase III Randomized Trial. JNCI: Journal of the National Cancer Institute. 2009;101(4):228-36.

130. D'Amico AV, Moul JW, Carroll PR, Sun L, Lubeck D, Chen MH. Surrogate end point for prostate cancer-specific mortality after radical prostatectomy or radiation therapy. Journal of the National Cancer Institute. 2003;95(18):1376-83.

131. Valicenti RK, DeSilvio M, Hanks GE, Porter A, Brereton H, Rosenthal SA, et al. Posttreatment prostatic-specific antigen doubling time as a surrogate endpoint for prostate cancer-specific survival: an analysis of Radiation Therapy Oncology Group Protocol 92-02. Int J Radiat Oncol Biol Phys. 2006;66(4):1064-71.

132. Xie W, Regan MM, Buyse M, Halabi S, Kantoff PW, Sartor O, et al. Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2017;35(27):3097-104.

133. Wang X, Parast L, Han L, Tian L, Cai T. Robust approach to combining multiple markers to improve surrogacy. Biometrics. 2023;79(2):788-98.

134. Xu J, Zeger SL. The evaluation of multiple surrogate endpoints. Biometrics. 2001;57(1):81-7.

135. Turchetti G, Geisler E, editors. Health technology assessment in the context of private and public models of national health systems. PICMET '09 - 2009 Portland International Conference on Management of Engineering & Technology; 2009 2-6 Aug. 2009.

136. World Health Organization. Global Survey on Health Technology Assessment by National Authorities. WHO, editor: WHO; 2015.

137. Maeda H, Takeda K, Urushihara H, Kurokawa T. Searching for potential surrogate endpoints of overall survival in clinical trials for patients with prostate cancer. Cancer Rep (Hoboken). 2021;4(3):e1334.

138. Sweeney C, Nakabayashi M, Regan M, Xie W, Hayes J, Keating N, et al. The Development of Intermediate Clinical Endpoints in Cancer of the Prostate (ICECaP). Journal of the National Cancer Institute. 2015;107(12):djv261.

139. Taylor RS, Elston J. The use of surrogate outcomes in model-based cost-effectiveness analyses: a survey of UK Health Technology Assessment reports. Health Technol Assess. 2009;13(8):iii, ix-xi, 1-50.

140. Mathes T, Walgenbach M, Antoine SL, Pieper D, Eikermann M. Methods for systematic reviews of health economic evaluations: a systematic review, comparison, and synthesis of method literature. Med Decis Making. 2014;34(7):826-40.

141. Ensor H, Lee RJ, Sudlow C, Weir CJ. Statistical approaches for evaluating surrogate outcomes in clinical trials: A systematic review. Journal of biopharmaceutical statistics. 2016;26(5):859-79.

142. Chiou CF, Hay JW, Wallace JF, Bloom BS, Neumann PJ, Sullivan SD, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. Medical care. 2003;41(1):32-44.

143. Ofman JJ, Sullivan SD, Neumann PJ, Chiou CF, Henning JM, Wade SW, et al. Examining the value and quality of health economic analyses: implications of utilizing the QHES. J Manag Care Pharm. 2003;9(1):53-61.

144. Al Mukdad M, Al-Badriyeh D, Elewa HF. Cost-effectiveness Evaluations Among the Direct Oral Anticoagulants for the Prevention and Treatment of Venous Thromboembolism: Systematic Review. Clin Appl Thromb Hemost. 2019;25:1076029619849103.

145. Al-Badriyeh D, Alameri M, Al-Okka R. Cost-effectiveness research in cancer therapy: a systematic review of literature trends, methods and the influence of funding. BMJ Open. 2017;7(1):e012648.

146. Roberts SLE, Healey A, Sevdalis N. Use of health economic evaluation in the implementation and improvement science fields-a systematic literature review. Implement Sci. 2019;14(1):72.

147. Lundkvist J, Ekman M, Ericsson SR, Jonsson B, Glimelius B. Proton therapy of cancer: potential clinical advantages and cost-effectiveness. Acta Oncol. 2005;44(8):850-61.

148. Bloomfield DJ, Krahn MD, Neogi T, Panzarella T, Smith TJ, Warde P, et al. Economic evaluation of chemotherapy with mitoxantrone plus prednisone for symptomatic hormone-resistant prostate cancer: based on a Canadian randomized trial with palliative end points. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1998;16(6):2272-9.

149. Hayes S. Translating evidence into practice: Incorporating exercise into standard cancer care. Asia-Pacific Journal of Clinical Oncology. 2013;9:63-4.

150. Yong JHE, Beca J, McGowan T, Bremner KE, Warde P, Hoch JS. Cost-effectiveness of intensity-modulated radiotherapy in prostate cancer. Clinical oncology (Royal College of Radiologists (Great Britain)). 2012;24(7):521-31.

151. Pollack A, Hanlon AL, Movsas B, Hanks GE, Uzzo R, Horwitz EM. Biochemical failure as a determinant of distant metastasis and death in prostate cancer treated with radiotherapy. International Journal of Radiation Oncology • Biology • Physics. 2003;57(1):19-23.

152. Slater JD, Yonemoto LT, Rossi CJ, Reyes-Molyneux NJ, Bush DA, Antoine JE, et al. Conformal proton therapy for prostate carcinoma. International Journal of Radiation Oncology*Biology*Physics. 1998;42(2):299-304.

153. Hanks GE, Hanlon AL, Pinover WH, Horwitz EM, Price RA, Schultheiss T. Dose selection for prostate cancer patients based on dose comparison and dose response studies. Int J Radiat Oncol Biol Phys. 2000;46(4):823-32.

154. Consensus statement: guidelines for PSA following radiation therapy. American Society for Therapeutic Radiology and Oncology Consensus Panel. Int J Radiat Oncol Biol Phys. 1997;37(5):1035-41.

155. Daniels MJ, Hughes MD. Meta-analysis for the evaluation of potential surrogate markers. Stat Med. 1997;16(17):1965-82.

156. Gail MH, Pfeiffer R, Van Houwelingen HC, Carroll RJ. On meta-analytic assessment of surrogate outcomes. Biostatistics. 2000;1(3):231-46.

157. Baker SG. A simple meta-analytic approach for using a binary surrogate endpoint to predict the effect of intervention on true endpoint. Biostatistics. 2006;7(1):58-70.

158. Collette L, Burzykowski T, Schröder FH. Prostate-specific antigen (PSA) alone is not an appropriate surrogate marker of long-term therapeutic benefit in prostate cancer trials. Eur J Cancer. 2006;42(10):1344-50.

159. Xie W, Regan MM, Buyse M, Halabi S, Kantoff PW, Sartor O, et al. Event-Free Survival, a Prostate-Specific Antigen-Based Composite End Point, Is Not a Surrogate for Overall Survival in Men With Localized Prostate Cancer Treated With Radiation. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2020:Jco1903114.

160. Tan SH, Abrams KR, Bujkiewicz S. Bayesian Multiparameter Evidence Synthesis to Inform Decision Making: A Case Study in Metastatic Hormone-Refractory Prostate Cancer. Med Decis Making. 2018;38(7):834-48.

161. Collins R, Fenwick E, Trowman R, Perard R, Norman G, Light K, et al. A systematic review and economic model of the clinical effectiveness and cost-effectiveness of docetaxel in combination with prednisone or prednisolone for the treatment of hormone-refractory metastatic prostate cancer. Health Technol Assess. 2007;11(2):iii-iv, xv-xviii, 1-179.

162. Cookson MS, Aus G, Burnett AL, Canby-Hagino ED, D'Amico AV, Dmochowski RR, et al. Variation in the definition of biochemical recurrence in patients treated for localized prostate cancer: the American Urological Association Prostate Guidelines for Localized Prostate Cancer Update Panel report and recommendations for a standard in the reporting of surgical outcomes. The Journal of urology. 2007;177(2):540-5.

163. Kestin LL, Vicini FA, Martinez AA. Practical application of biochemical failure definitions: what to do and when to do it. International Journal of Radiation Oncology*Biology*Physics. 2002;53(2):304-15.

164. De Abreu Lourenco R, Williams S. Health Technology Assessment and Health Care Utilisation in the Management of Early Prostate Cancer. European Urology Focus. 2018.

165. Evans C, Dawe R, Carson-Stevens A. Re: Branney et al.: Choosing health, choosing treatment: patient choice after diagnosis of localized prostate cancer (Urology 2009;74:968-971). Urology. 2010;75(3):751; author reply -2.

166. Beauchemin C, Lapierre M, Letarte N, Yelle L, Lachaine J. Use of Intermediate Endpoints in the Economic Evaluation of New Treatments for Advanced Cancer and Methods Adopted When Suitable Overall Survival Data are Not Available. Pharmacoeconomics. 2016;34(9):889-900.

167. Food and Drug Administration. Subpart H - Accelerated Approval of New Drugs for Serious or Life-Threatening Illnesses April 2020, [Jan 2021]. Available from: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=314&showFR=1&subpartNode=21:5.0.1.1.4.8.

168. Schuster Bruce C, Brhlikova P, Heath J, McGettigan P. The use of validated and nonvalidated surrogate endpoints in two European Medicines Agency expedited approval pathways: A cross-sectional study of products authorised 2011-2018. PLoS Med. 2019;16(9):e1002873-e.

169. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Annals of internal medicine. 1996;125(7):605-13.

170. Ciani O, Grigore B, Blommestein H, de Groot S, Möllenkamp M, Rabbe S, et al. Validity of Surrogate Endpoints and Their Impact on Coverage Recommendations: A Retrospective Analysis across International Health Technology Assessment Agencies. Medical Decision Making. 2021;41(4):439-52.

171. Nelsen RB, editor Properties and applications of copulas: A brief survey. Proceedings of the first brazilian conference on statistical modeling in insurance and finance; 2003: Citeseer.

172. Roach M, 3rd, Hanks G, Thames H, Jr., Schellhammer P, Shipley WU, Sokol GH, et al. Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: recommendations of the RTOG-ASTRO Phoenix Consensus Conference. Int J Radiat Oncol Biol Phys. 2006;65(4):965-74.

173. Medical Services Advisory Committee Technical Guidelines for preparing assessment reports for the Medical Services Advisory Committee – Service Type: Investigative. In: Health Do, editor. July 2017.

174. Latimer NR. Survival analysis for economic evaluations alongside clinical trials--extrapolation with patient-level data: inconsistencies, limitations, and a practical guide. Med Decis Making. 2013;33(6):743-54.

175. Gallacher D, Kimani P, Stallard N. Extrapolating Parametric Survival Models in Health Technology Assessment: A Simulation Study. Medical decision making : an international journal of the Society for Medical Decision Making. 2021;41(1):37-50.

176. Woods BS, Sideris E, Palmer S, Latimer N, Soares M. Partitioned Survival and State Transition Models for Healthcare Decision Making in Oncology: Where Are We Now? Value in Health. 2020;23(12):1613-21.

177. Messori A, Becagli P, Trippoli S. Median versus mean lifetime survival in the analysis of survival data. Haematologica. 1997;82(6):730.

178. Damuzzo V, Agnoletto L, Leonardi L, Chiumente M, Mengato D, Messori A. Analysis of Survival Curves: Statistical Methods Accounting for the Presence of Long-Term Survivors. Frontiers in oncology. 2019;9:453-.

179. Cox DR. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological). 1972;34(2):187-202.

180. Latimer N. NICE DSU technical support document 14: survival analysis for economic evaluations alongside clinical trials-extrapolation with patient-level data. Report by the Decision Support Unit. 2011.

181. Jackson C, Stevens J, Ren S, Latimer N, Bojke L, Manca A, et al. Extrapolating Survival from Randomized Trials Using External Data: A Review of Methods. Medical Decision Making. 2017;37(4):377-90.

182. van Oostrum I, Ouwens M, Remiro-Azócar A, Baio G, Postma MJ, Buskens E, et al. Comparison of Parametric Survival Extrapolation Approaches Incorporating General Population Mortality for Adequate Health Technology Assessment of New Oncology Drugs. Value in Health. 2021;24(9):1294-301.

183. Cowen ME, Halasyamani LK, Kattan MW. Predicting life expectancy in men with clinically localized prostate cancer. The Journal of urology. 2006;175(1):99-103.

184. AIHW. General Record of Incidence of Mortality workbooks. Life Tables, 2017-2019. 2019.

185. Hofman MS, Lawrentschuk N, Francis RJ, Tang C, Vela I, Thomas P, et al. Prostate-specific membrane antigen PET-CT in patients with high-risk prostate cancer before curative-intent surgery or radiotherapy (proPSMA): a prospective, randomised, multi-centre study. The Lancet. 2020.

186. Pesapane F, Czarniecki M, Suter MB, Turkbey B, Villeirs G. Imaging of distant metastases of prostate cancer. Medical oncology (Northwood, London, England). 2018;35(11):148.

187. Ganesh K, Massagué J. Targeting metastatic cancer. Nature Medicine. 2021;27(1):34-44.

188. Horwitz EM, Bae K, Hanks GE, Porter A, Grignon DJ, Brereton HD, et al. Ten-year follow-up of radiation therapy oncology group protocol 92-02: a phase III trial of the duration of elective androgen deprivation in locally advanced prostate cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008;26(15):2497-504.

189. Bolla M, Maingon P, Carrie C, Villa S, Kitsios P, Poortmans PM, et al. Short Androgen Suppression and Radiation Dose Escalation for Intermediate- and High-Risk Localized Prostate Cancer: Results of EORTC Trial 22991. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016;34(15):1748-56.

190. Beer TM, Miller K, Tombal B, Cella D, Phung, Holmstrom S, et al. The association between health-related quality-of-life scores and clinical outcomes in metastatic castration-resistant prostate cancer patients: Exploratory analyses of AFFIRM and PREVAIL studies. Eur J Cancer. 2017;87:21-9.

191. Wang G, Zhao D, Spring DJ, DePinho RA. Genetics and biology of prostate cancer. Genes Dev. 2018;32(17-18):1105-40.

192. Chin JL, Al-Zahrani AA, Autran-Gomez AM, Williams AK, Bauman G. Extended followup oncologic outcome of randomized trial between cryoablation and external beam therapy for locally advanced prostate cancer (T2c-T3b). The Journal of urology. 2012;188(4):1170-5.

193. Beckendorf V, Guerif S, Le Prise E, Cosset JM, Bougnoux A, Chauvet B, et al. 70 Gy versus 80 Gy in localized prostate cancer: 5-year results of GETUG 06 randomized trial. Int J Radiat Oncol Biol Phys. 2011;80(4):1056-63.

194. Dearnaley DP, Jovic G, Syndikus I, Khoo V, Cowan RA, Graham JD, et al. Escalated-dose versus control-dose conformal radiotherapy for prostate cancer: long-term results from the MRC RT01 randomised controlled trial. The Lancet Oncology. 2014;15(4):464-73.

195. Bolla M, de Reijke TM, Van Tienhoven G, Van den Bergh AC, Oddens J, Poortmans PM, et al. Duration of androgen suppression in the treatment of prostate cancer. The New England journal of medicine. 2009;360(24):2516-27.

196. Mottet N, Peneau M, Mazeron JJ, Molinie V, Richaud P. Addition of radiotherapy to long-term androgen deprivation in locally advanced prostate cancer: an open randomised phase 3 trial. European urology. 2012;62(2):213-9.

197. Zapatero A, Guerrero A, Maldonado X, Alvarez A, Gonzalez San Segundo C, Cabeza Rodriguez MA, et al. High-dose radiotherapy with short-term or long-term androgen deprivation in localised prostate cancer (DART01/05 GICOR): a randomised, controlled, phase 3 trial. The Lancet Oncology. 2015;16(3):320-7.

198. Armstrong JG, Gillham CM, Dunne MT, Fitzpatrick DA, Finn MA, Cannon ME, et al. A randomized trial (Irish clinical oncology research group 97-01) comparing short versus protracted neoadjuvant hormonal therapy before radiotherapy for localized prostate cancer. Int J Radiat Oncol Biol Phys. 2011;81(1):35-45.

199. Jones CU, Hunt D, McGowan DG, Amin MB, Chetner MP, Bruner DW, et al. Radiotherapy and short-term androgen deprivation for localized prostate cancer. The New England journal of medicine. 2011;365(2):107-18.

200. Lawton CA, DeSilvio M, Roach M, 3rd, Uhl V, Kirsch R, Seider M, et al. An update of the phase III trial comparing whole pelvic to prostate only radiotherapy and neoadjuvant to adjuvant total androgen suppression: updated analysis of RTOG 94-13, with emphasis on unexpected hormone/radiation interactions. Int J Radiat Oncol Biol Phys. 2007;69(3):646-55.

201. Denham JW, Steigler A, Lamb DS, Joseph D, Turner S, Matthews J, et al. Short-term neoadjuvant androgen deprivation and radiotherapy for locally advanced prostate cancer: 10-year data from the TROG 96.01 randomised trial. The Lancet Oncology. 2011;12(5):451-9.

202. Ciani O, Davis S, Tappenden P, Garside R, Stein K, Cantrell A, et al. Validation of surrogate endpoints in advanced solid tumors: systematic review of statistical methods, results, and implications for policy makers. Int J Technol Assess Health Care. 2014;30(3):312-24.

203. Collett D. Modelling survival data in medical research: CRC press; 2015.

204. Food and Drug Administration Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics Guidance for Industry. In: U.S. Department of Health and Human Services, editor.: Clinical/Medical; December 2018.

205. Bulsei J, Darlington M, Durand-Zaleski I, Azizi M. How to perform a cost-effectiveness analysis with surrogate endpoint: renal denervation in patients with resistant hypertension (DENERHTN) trial as an example. Blood Pressure. 2018;27(2):66-72.

206. Philippou Y, Hadjipavlou M, Khan S, Ahmed K, Rane A. Localised prostate cancer: clinical and cost-effectiveness of new and emerging technologies. Journal of Clinical Urology. 2014;7(4):239-51.

207. Ilic D, Evans SM, Allan CA, Jung JH, Murphy D, Frydenberg M. Laparoscopic and robotic-assisted versus open radical prostatectomy for the treatment of localised prostate cancer. Cochrane Database of Systematic Reviews. 2017(9).

208. O'CALLAGHAN M PM, FRYDENBERG M, MARK S, MORETTI K, MAQSOOD S, SMITH D, WALKER T, WHITE C, MILLAR J. Prostate Cancer in Australian and New Zealand Men, Patterns of care within PCOR-ANZ 2015–2017. Melbourne, VIC: Monash University and Movember. 2019.

209. Loblaw DA, Mendelson DS, Talcott JA, Virgo KS, Somerfield MR, Ben-Josef E, et al. American Society of Clinical Oncology recommendations for the initial hormonal management of androgen-sensitive metastatic, recurrent, or progressive prostate cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2004;22(14):2927-41.

210. Cassileth BR, Soloway MS, Vogelzang NJ, Schellhammer PS, Seidmon EJ, Hait HI, et al. Patients' choice of treatment in stage D prostate cancer. Urology. 1989;33(5 Suppl):57-62.

211. Gillessen S, Omlin A, Attard G, de Bono JS, Efstathiou E, Fizazi K, et al. Management of patients with advanced prostate cancer: recommendations of the St Gallen Advanced Prostate Cancer Consensus Conference (APCCC) 2015. Annals of oncology : official journal of the European Society for Medical Oncology. 2015;26(8):1589-604.

212. Perlmutter MA, Lepor H. Androgen deprivation therapy in the treatment of advanced prostate cancer. Reviews in urology. 2007;9 Suppl 1(Suppl 1):S3-S8.

213. Crona DJ, Milowsky MI, Whang YE. Androgen receptor targeting drugs in castration-resistant prostate cancer and mechanisms of resistance. Clinical pharmacology and therapeutics. 2015;98(6):582-9.

214. Foroughi Moghadam MJ, Taheri S, Peiravian F. A Systematic Review of Clinical Practice Guidelines for Castration-Resistant Prostate Cancer. Iranian journal of pharmaceutical research : IJPR. 2018;17(Suppl):17-37.

215. James ND, Sydes MR, Clarke NW, Mason MD, Dearnaley DP, Spears MR, et al. Addition of docetaxel, zoledronic acid, or both to first-line long-term hormone therapy in prostate cancer (STAMPEDE): survival results from an adaptive, multiarm, multistage, platform randomised controlled trial. The Lancet. 2016;387(10024):1163-77.

216. Krahn M, Ritvo P, Irvine J, Tomlinson G, Bezjak A, Trachtenberg J, et al. Construction of the Patient-Oriented Prostate Utility Scale (PORPUS): a multiattribute health state classification system for prostate cancer. Journal of clinical epidemiology. 2000;53(9):920-30.

217. Magnus A, Isaranuwatchai W, Mihalopoulos C, Brown V, Carter R. A Systematic Review and Meta-Analysis of Prostate Cancer Utility Values of Patients and Partners Between 2007 and 2016. MDM Policy Pract. 2019;4(1):2381468319852332.

218. Skaltsa K, Longworth L, Ivanescu C, Phung D, Holmstrom S. Mapping the FACT-P to the preference-based EQ-5D questionnaire in metastatic castration-resistant prostate cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2014;17(2):238-44.

219. Krahn M, Bremner KE, Tomlinson G, Ritvo P, Irvine J, Naglie G. Responsiveness of disease-specific and generic utility instruments in prostate cancer patients. Qual Life Res. 2007;16(3):509-22.

220. Hanna TP, King WD, Thibodeau S, Jalink M, Paulin GA, Harvey-Jones E, et al. Mortality due to cancer treatment delay: systematic review and meta-analysis. BMJ (Clinical research ed). 2020;371:m4087.

221. Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1989;7(1):36-44.

222. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. Stat Med. 1990;9(11):1259-76.

223. Hayden AJ, Martin JM, Kneebone AB, Lehman M, Wiltshire KL, Skala M, et al. Australian & New Zealand Faculty of Radiation Oncology Genito-Urinary Group: 2010 consensus guidelines for definitive external beam radiotherapy for prostate carcinoma. J Med Imaging Radiat Oncol. 2010;54(6):513-25.

224. Freedland SJ, Abrahamsson PA. Androgen deprivation therapy and side effects: are GnRH antagonists safer? Asian J Androl. 2021;23(1):3-10.

225. Duchesne GM, Woo HH, Bassett JK, Bowe SJ, D'Este C, Frydenberg M, et al. Timing of androgen-deprivation therapy in patients with prostate cancer with a rising PSA (TROG 03.06 and VCOG PR 01-03 [TOAD]): a randomised, multicentre, non-blinded, phase 3 trial. The Lancet Oncology. 2016;17(6):727-37.

226. Scher HI, Solo K, Valant J, Todd MB, Mehra M. Prevalence of Prostate Cancer Clinical States and Mortality in the United States: Estimates Using a Dynamic Progression Model. PLoS One. 2015;10(10):e0139440.

227. Parker CC, James ND, Brawley CD, Clarke NW, Hoyle AP, Ali A, et al. Radiotherapy to the primary tumour for newly diagnosed, metastatic prostate cancer (STAMPEDE): a randomised controlled phase 3 trial. Lancet (London, England). 2018;392(10162):2353-66.

228. Hayes JH, Ollendorf DA, Pearson SD, Barry MJ, Kantoff PW, Stewart ST, et al. Active surveillance compared with initial treatment for men with low-risk prostate cancer: a decision analysis. Jama. 2010;304(21):2373-80.

229. Hershman DL, Unger JM, Wright JD, Ramsey S, Till C, Tangen CM, et al. Adverse Health Events Following Intermittent and Continuous Androgen Deprivation in Patients With Metastatic Prostate Cancer. JAMA Oncol. 2016;2(4):453-61.

230. Australian Institute of Health and Welfare (2013). Australia's health 2013. Canberra: AIHW. 2013.

231. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. The New England journal of medicine. 2016;375(15):1415-24.

232. Lingaratnam S, Worth LJ, Slavin MA, Bennett CA, Kirsa SW, Seymour JF, et al. A cost analysis of febrile neutropenia management in Australia: ambulatory v. in-hospital treatment. Aust Health Rev. 2011;35(4):491-500.

233. Higgins JPT TJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021) Available from www.training.cochrane.org/handbook2021 [Available from: www.training.cochrane.org/handbook.

234. Avila M, Becerra V, Guedea F, Suarez JF, Fernandez P, Macias V, et al. Estimating preferences for treatments in patients with localized prostate cancer. Int J Radiat Oncol Biol Phys. 2015;91(2):277-87.

235. Avila M, Pardo Y, Castells M, Ferrer F, Boladeras A, Pera J, et al. Adaptation and validation of the Spanish version of the Patient-Oriented Prostate Utility Scale (PORPUS). Qual Life Res. 2014;23(9):2481-7.

236. Basu A, Dale W, Elstein A, Meltzer D. A linear index for predicting joint health-states utilities from single health-states utilities. Health Econ. 2009;18(4):403-19.

237. Basu A, Dale W, Elstein A, Meltzer D. A time tradeoff method for eliciting partner's quality of life due to patient's health states in prostate cancer. Med Decis Making. 2010;30(3):355-65.

238. Bremner KE, Mitsakakis N, Wilson L, Krahn MD. Predicting utility scores for prostate cancer: mapping the Prostate Cancer Index to the Patient-Oriented Prostate Utility Scale (PORPUS). Prostate Cancer Prostatic Dis. 2014;17(1):47-56.

239. Bremner KE, Tomlinson G, Krahn MD. Marker states and a health state prompt provide modest improvements in the reliability and validity of the standard gamble and rating scale in prostate cancer patients. Qual Life Res. 2007;16(10):1665-75.

240. C Glazener CB, B Buckley, C Cochran,, G Dorey AG, S Hagen, M Kilonzo,, A McDonald GM, K Moore,, J N'Dow JN, C Ramsay and L Vale. Conservative treatment for urinary incontinence in Men After Prostate Surgery (MAPS): two parallel randomised controlled trials. Health Technology Assessment. 2011;26.

241. Cantor SB, Deshmukh AA, Krahn MD, Volk RJ. Use of forecasted assessment of quality of life to validate time-trade-off utilities and a prostate cancer screening decision-analytic model. Health Expect. 2015;18(5):1610-20.

242. Cantor SB, Volk RJ, Krahn MD, Cass AR, Gilani J, Weller SC, et al. Concordance of couples' prostate cancer screening recommendations from a decision analysis. Patient. 2008;1(1):11-9.

243. Dale W, Basu A, Elstein A, Meltzer D. Predicting utility ratings for joint health States from single health States in prostate cancer: empirical testing of 3 alternative theories. Med Decis Making. 2008;28(1):102-12.

244. Diels J, Hamberg P, Ford D, Price PW, Spencer M, Dass RN. Mapping FACT-P to EQ-5D in a large cross-sectional study of metastatic castration-resistant prostate cancer patients. Qual Life Res. 2015;24(3):591-8.

245. Farkkila N, Torvinen S, Roine RP, Sintonen H, Hanninen J, Taari K, et al. Health-related quality of life among breast, prostate, and colorectal cancer patients with end-stage disease. Qual Life Res. 2014;23(4):1387-94.

246. Gries KS, Regier DA, Ramsey SD, Patrick DL. Preferences for Prostate Cancer Outcomes: A Comparison of the Patient Perspective, the General Population Perspective, and a Population at Risk for Prostate Cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(2):218-25.

247. Hall PS, Hamilton P, Hulme CT, Meads DM, Jones H, Newsham A, et al. Costs of cancer care for use in economic evaluation: a UK analysis of patient-level routine health system data. Br J Cancer. 2015;112(5):948-56.

248. Jayadevappa R, Schwartz JS, Chhatre S, Wein AJ, Malkowicz SB. Association between utility and treatment among patients with prostate cancer. Qual Life Res. 2010;19(5):711-20.

249. Korfage IJ, de Koning HJ, Habbema JD, Schroder FH, Essink-Bot ML. Side-effects of treatment for localized prostate cancer: are they valued differently by patients and healthy controls? BJU international. 2007;99(4):801-6.

250. Krahn MD, Bremner KE, Alibhai SM, Ni A, Tomlinson G, Laporte A, et al. A reference set of health utilities for long-term survivors of prostate cancer: population-based data from Ontario, Canada. Qual Life Res. 2013;22(10):2951-62.

251. Krahn MD, Bremner KE, Tomlinson G, Naglie G. Utility and health-related quality of life in prostate cancer patients 12 months after radical prostatectomy or radiation therapy. Prostate Cancer Prostatic Dis. 2009;12(4):361-8.

252. Ku J, Krahn M, Trachtenberg J, Nesbitt M, Kalnin R, Lockwood G, et al. Changes in health utilities and health-related quality of life over 12 months following radical prostatectomy. Can Urol Assoc J. 2009;3(6):445-52.

253. Lane A, Metcalfe C, Young GJ, Peters TJ, Blazeby J, Avery KN, et al. Patient-reported outcomes in the ProtecT randomized trial of clinically localized prostate cancer treatments: study design, and baseline urinary, bowel and sexual function and quality of life. BJU international. 2016;118(6):869-79.

254. Loriot Y, Miller K, Sternberg CN, Fizazi K, De Bono JS, Chowdhury S, et al. Effect of enzalutamide on health-related quality of life, pain, and skeletal-related events in asymptomatic and minimally symptomatic, chemotherapy-naive patients with metastatic castration-resistant prostate cancer (PREVAIL): results from a randomised, phase 3 trial. The Lancet Oncology. 2015;16(5):509-21.

255. Shimizu F, Fujino K, Ito YM, Fukuda T, Kawachi Y, Minowada S, et al. Factors associated with variation in utility scores among patients with prostate cancer. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2008;11(7):1190-3.

256. Sommers BD, Beard CJ, D'Amico AV, Kaplan I, Richie JP, Zeckhauser RJ. Predictors of patient preferences and treatment choices for localized prostate cancer. Cancer. 2008;113(8):2058-67.

257. Torvinen S, Farkkila N, Sintonen H, Saarto T, Roine RP, Taari K. Health-related quality of life in prostate cancer. Acta Oncol. 2013;52(6):1094-101.

258. Watson E, Shinkins B, Frith E, Neal D, Hamdy F, Walter F, et al. Symptoms, unmet needs, psychological well-being and health status in survivors of prostate cancer: implications for redesigning follow-up. BJU international. 2016;117(6B):E10-9.

259. Wu EQ, Mulani P, Farrell MH, Sleep D. Mapping FACT-P and EORTC QLQ-C30 to patient health status measured by EQ-5D in metastatic hormone-refractory prostate cancer patients. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2007;10(5):408-14.

260. Zhang AY, Fu AZ. Cost-effectiveness of a behavioral intervention for persistent urinary incontinence in prostate cancer patients. Psychooncology. 2016;25(4):421-7.

261. Norman R, Church J, van den Berg B, Goodall S. Australian health-related quality of life population norms derived from the SF-6D. Australian and New Zealand journal of public health. 2013;37(1):17-23.

262. Wang S, Gum D, Merlin T. Comparing the ICERs in Medicine Reimbursement Submissions to NICE and PBAC—Does the Presence of an Explicit Threshold Affect the ICER Proposed? Value in Health. 2018;21(8):938-43.

263. The Senate Community Affairs References Committee. Availability of new, innovative and specialist cancer drugs in Australia. . 2015.

264. Meier P, Karrison T, Chappell R, Xie H. The Price of Kaplan-Meier. Journal of the American Statistical Association. 2004;99:890-6.

265. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ. 1999;18(3):341-64.

266. Wilson ECF. A Practical Guide to Value of Information Analysis. PharmacoEconomics. 2015;33(2):105-21.

267. Philips Z, Claxton K, Palmer S. The Half-Life of Truth: What Are Appropriate Time Horizons for Research Decisions? Medical Decision Making. 2008;28(3):287-99.

268. Amin NP, Sher DJ, Konski AA. Systematic review of the cost effectiveness of radiation therapy for prostate cancer from 2003 to 2013. Appl Health Econ Health Policy. 2014;12(4):391-408.

269. Earnshaw SR, Brogan AP, McDade CL. Model-based cost-effectiveness analyses for prostate cancer chemoprevention : a review and summary of challenges. Pharmacoeconomics. 2013;31(4):289-304.

270. Hatoum HT, Crawford ED, Nielsen SK, Lin SJ, Marshall DC. Review of the economic evaluations of hormonal therapy for patients with locally advanced prostate cancer. Expert review of pharmacoeconomics & outcomes research. 2013;13(2):251-9.

271. Muralidhar V, Nguyen PL. Maximizing resources in the local treatment of prostate cancer: A summary of cost-effectiveness studies. Urol Oncol. 2017;35(2):76-85.

272. Becerra V, Ávila M, Jimenez J, Cortes-Sanabria L, Pardo Y, Garin O, et al. Economic evaluation of treatments for patients with localized prostate cancer in Europe: a systematic review. BMC health services research. 2016;16(1):541-.

273. Moore TJ, Zhang H, Anderson G, Alexander GC. Estimated Costs of Pivotal Trials for Novel Therapeutic Agents Approved by the US Food and Drug Administration, 2015-2016. JAMA internal medicine. 2018;178(11):1451-7.

274. Weng X, Zhong L, Xiang P, Li Y, Paciorek A, Dong L, et al. Cost-effectiveness analysis of primary treatments for localised prostate cancer: A population-based Markov analysis using real-world evidence. European Journal of Cancer Care. 2022;31(6):e13740.

275. Dorth JA, Lee WR, Chino J, Abouassaly R, Ellis RJ, Myers ER. Cost-Effectiveness of Primary Radiation Therapy Versus Radical Prostatectomy for Intermediate- to High-Risk Prostate Cancer. International Journal of Radiation Oncology*Biology*Physics. 2018;100(2):383-90.

276. Trueman P, Grainger DL, Downs KE. Coverage with Evidence Development: applications and issues. Int J Technol Assess Health Care. 2010;26(1):79-85.

277. Williams C, Lewsey JD, Briggs AH, Mackay DF. Cost-effectiveness Analysis in R Using a Multi-state Modeling Survival Analysis Framework: A Tutorial. Medical decision making : an international journal of the Society for Medical Decision Making. 2017;37(4):340-52.

278. Wong H-H, Jessup A, Sertkaya A, Birkenbach A, Berlind A, Eyraud J. Examination of clinical trial costs and barriers for drug development final. Office of the Assistant Secretary for Planning and Evaluation, US Department of Health & Human Services. 2014.

279. Lenzer J, Brownlee S. Should regulatory authorities approve drugs based on surrogate endpoints? BMJ (Clinical research ed). 2021;374:n2059.

280. Davis ID, Martin AJ, Stockler MR, Begbie S, Chi KN, Chowdhury S, et al. Enzalutamide with Standard First-Line Therapy in Metastatic Prostate Cancer. New England Journal of Medicine. 2019;381(2):121-31.

281. de Feria Cardet RE, Hofman MS, Segard T, Yim J, Williams S, Francis RJ, et al. Is Prostate-specific Membrane Antigen Positron Emission Tomography/Computed Tomography Imaging Cost-effective in Prostate Cancer: An Analysis Informed by the proPSMA Trial. European urology. 2021;79(3):413-8.

282. Gu Y, Lancsar E, Ghijben P, Butler JR, Donaldson C. Attributes and weights in health care priority setting: A systematic review of what counts and to what extent. Soc Sci Med. 2015;146:41-52.

283. Einav L, Finkelstein A, Ryan SP, Schrimpf P, Cullen MR. Selection on Moral Hazard in Health Insurance. American Economic Review. 2013;103(1):178-219.

284. Ryan M, Gerard K, Amaya-Amaya M. Using discrete choice experiments to value health and health care: Springer Science & Business Media; 2007.

285. Zweifel P. Preference measurement in health using experiments. Central European Journal of Operations Research. 2022;30(1):49-66.

286. Knottnerus JA, Knottnerus BJ. Decision-making given surrogate outcomes. Journal of clinical epidemiology. 2022.

287. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P, Ed, Frontiers in Econometrics, Academic Press,. 1973:105-42.

288. Lancaster KJ. A New Approach to Consumer Theory. Journal of Political Economy. 1966;74(2):132-57.

289. Train KE. Discrete choice methods with simulation: Cambridge university press; 2009.

290. Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, et al. Conjoint Analysis Applications in Health—a Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value in Health. 2011;14(4):403-13.

291. DeShazo JR, Fermo G. Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. Journal of Environmental Economics and Management. 2002;44(1):123-43.

292. Viney R, Lancsar E, Louviere J. Discrete choice experiments to measure consumer preferences for health and healthcare. Expert review of pharmacoeconomics & outcomes research. 2002;2(4):319-26.

293. Hollin IL, Craig BM, Coast J, Beusterien K, Vass C, DiSantostefano R, et al. Reporting Formative Qualitative Research to Support the Development of Quantitative Preference Study Protocols and Corresponding Survey Instruments: Guidelines for Authors and Reviewers. Patient. 2020;13(1):121-36.

294. Vass C, Rigby D, Payne K. The Role of Qualitative Research Methods in Discrete Choice Experiments. Medical decision making : an international journal of the Society for Medical Decision Making. 2017;37(3):298-313.

295. O.Nyumba T, Wilson K, Derrick CJ, Mukherjee N. The use of focus group discussion methodology: Insights from two decades of application in conservation. Methods in Ecology and Evolution. 2018;9(1):20-32.

296. McCarthy M, De Abreu Lourenco R, McMillan LJ, Meshcheriakova O, Cao A, Gillam L. Finding out what matters in decision-making related to genomics and personalized medicine in pediatric oncology: Developing attributes to include in a discrete choice experiment. The Patient: Patient Centered Outcomes Research. 2020;13:347-61.

297. Brown A, Pain T, Tan A, Anable L, Callander E, Watt K, et al. Men's preferences for image-guidance in prostate radiation therapy: A discrete choice experiment. Radiotherapy and oncology : journal of the European Society for Therapeutic Radiology and Oncology. 2022;167:49-56.

298. Antonio ALM, Weiss RE, Saigal CS, Dahan E, Crespi CM. A Bayesian hierarchical model for discrete choice data in health care. Statistical Methods in Medical Research. 2018;27(12):3544-59.

299. de Bekker-Grob EW, Bliemer MC, Donkers B, Essink-Bot ML, Korfage IJ, Roobol MJ, et al. Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. Br J Cancer. 2013;109(3):633-40.

300. de Freitas HM, Ito T, Hadi M, Al-Jassar G, Henry-Szatkowski M, Nafees B, et al. Patient Preferences for Metastatic Hormone-Sensitive Prostate Cancer Treatments: A Discrete Choice Experiment Among Men in Three European Countries. Advances in therapy. 2019;36(2):318-32.

301. Eliasson L, de Freitas HM, Dearden L, Calimlim B, Lloyd AJ. Patients' Preferences for the Treatment of Metastatic Castrate-resistant Prostate Cancer: A Discrete Choice Experiment. Clinical therapeutics. 2017;39(4):723-37.

302. Feldman-Stewart D, Brundage MD, Van Manen L, Svenson O. Patient-focussed decision-making in early-stage prostate cancer: Insights from a cognitively based decision aid. Health Expectations: An International Journal of Public Participation in Health Care & Health Policy. 2004;7(2):126-41.

303. Gorin MA, Soloway CT, Eldefrawy A, Soloway MS. Factors that influence patient enrollment in active surveillance for low-risk prostate cancer. Urology. 2011;77(3):588-91.

304. Jayadevappa R, Chhatre S, Gallo JJ, Malkowicz SB, Schwartz JS, Wittink MN. Patient-Centered Approach to Develop the Patient's Preferences for Prostate Cancer Care (PreProCare) Tool. MDM policy & practice. 2019;4(1):2381468319855375.

305. Kaplan RM, Crespi CM, Dahan E, Saucedo JD, Pagan C, Saigal CS. Comparison of Rating Scale, Time Tradeoff, and Conjoint Analysis Methods for Assessment of Preferences in Prostate Cancer. Medical decision making : an international journal of the Society for Medical Decision Making. 2019;39(7):816-26.

306. King MT, Viney R, Smith DP, Hossain I, Street D, Savage E, et al. Survival gains needed to offset persistent adverse treatment effects in localised prostate cancer. Br J Cancer. 2012;106(4):638-45.

307. Lloyd A, Penson D, Dewilde S, Kleinman L. Eliciting patient preferences for hormonal therapy options in the treatment of metastatic prostate cancer. Prostate cancer and prostatic diseases. 2008;11(2):153-9.

308. Nakayama M, Kobayashi H, Okazaki M, Imanaka K, Yoshizawa K, Mahlich J. Patient Preferences and Urologist Judgments on Prostate Cancer Therapy in Japan. American journal of men's health. 2018;12(4):1094-101.

309. Rentz AM, Mansukhani SG, Liu J, Lloyd AJ, Heidenreich S, Matza LS, et al. Patients' preferences for delaying metastatic castration-resistant prostate cancer: Combining health state and treatment valuation. Urologic oncology. 2021;39(6):367.e7-.e17.

310. Richman MB, Forman EH, Bayazit Y, Einstein DB, Resnick MI, Stovsky MD. A novel computer based expert decision making model for prostate cancer disease management. The Journal of urology. 2005;174(6):2310-8.

311. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. BMJ (Clinical research ed). 2004;328(7436):382.

312. Srinivas S, Mohamed AF, Appukkuttan S, Botteman M, Ng X, Joshi N, et al. Physician preferences for non-metastatic castration-resistant prostate cancer treatment. BMC urology. 2020;20(1):73.

313. Srinivas S, Mohamed AF, Appukkuttan S, Botteman M, Ng X, Joshi N, et al. Patient and caregiver benefit-risk preferences for nonmetastatic castration-resistant prostate cancer treatment. Cancer medicine. 2020;9(18):6586-96.

314. Suzuki K, Grillo V, Chen Y, Singh S, Ledesma DA. Understanding Treatment Strategies and Preferences in Nonmetastatic Castration-Resistant Prostate Cancer From the Japanese Physician Perspective. JCO global oncology. 2021;7:302-10.

315. Uemura H, Matsubara N, Kimura G, Yamaguchi A, Ledesma DA, DiBonaventura M, et al. Patient preferences for treatment of castration-resistant prostate cancer in Japan: a discrete-choice experiment. BMC urology. 2016;16(1):63.

316. Uemura H, Matsushima H, Yokomizo A, Kobayashi K, Arai G, Satoh T, et al. Unmet needs in non-metastatic castration-resistant prostate cancer from the Japanese patient perspective: a discrete choice experiment. BMJ open. 2021;11(8):e052471.

317. Underwood W, 3rd, Orom H, Poch M, West BT, Lantz PM, Chang SS, et al. Multiple physician recommendations for prostate cancer treatment: a Pandora's box for patients? The Canadian journal of urology. 2010;17(5):5346-54.

318. Volk RJ, McFall SL, Cantor SB, Byrd TL, Le Y-CL, Kuban DA, et al. 'It's not like you just had a heart attack': decision-making about active surveillance by men with localized prostate cancer. Psycho-oncology. 2014;23(4):467-72.

319. Watson V, McCartan N, Krucien N, Abu V, Ikenwilo D, Emberton M, et al. Evaluating the Trade-Offs Men with Localized Prostate Cancer Make between the Risks and Benefits of Treatments: The COMPARE Study. The Journal of urology. 2020;204(2):273-80.

320. Trapero-Bertran M, Rodríguez-Martín B, López-Bastida J. What attributes should be included in a discrete choice experiment related to health technologies? A systematic literature review. PloS one. 2019;14(7):e0219905-e.

321. Goto R, Uda A, Hiroi S, Iwasaki K, Takashima K, Oya M. Cost analysis of leuprorelin acetate in Japanese prostate cancer patients: comparison between 6-month and 3-month depot formulations. Journal of medical economics. 2017;20(11):1155-62.

322. Devarakonda S, Cottini F, Bumma N, Khan A, Sharma N, Chaudhry M, et al. Multiple myeloma: Clinical updates from the american society of clinical oncology annual scientific symposium 2020. Journal of Clinical Medicine. 2020;9(11):1-20.

323. Heindel W, Gubitz R, Vieth V, Weckesser M, Schober O, Schafers M. The diagnostic imaging of bone metastases. Deutsches Arzteblatt international. 2014;111(44):741-7.

324. Eton DT, Lepore SJ. Prostate cancer and health-related quality of life: a review of the literature. Psychooncology. 2002;11(4):307-26.

325. Soekhai V, de Bekker-Grob EW, Ellis AR, Vass CM. Discrete Choice Experiments in Health Economics: Past, Present and Future. PharmacoEconomics. 2019;37(2):201-26.

326. Hensher DA, Louviere JJ, Swait JD. Stated Choice Methods: Analysis and Applications. Cambridge: Cambridge University Press; 2000.

327. Street DJ, Burgess L. The construction of optimal stated choice experiments: theory and methods: John Wiley & Sons; 2007.

328. Huber J, Zwerina K. The Importance of Utility Balance in Efficient Choice Designs. Journal of Marketing Research. 1996;33(3):307-17.

329. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2013;16(1):3-13.

330. Campbell D, Erdem S. Including Opt-Out Options in Discrete Choice Experiments: Issues to Consider. Patient. 2019;12(1):1-14.

331. Hauber AB, Gonzalez JM, Groothuis-Oudshoorn CG, Prior T, Marshall DA, Cunningham C, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2016;19(4):300-15.

332. Greene WH, Hensher DA. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological. 2003;37(8):681-98.

333. Hoedemakers M, Karimi M, Jonker M, Tsiachristas A, Rutten-van Mölken M. Heterogeneity in preferences for outcomes of integrated care for persons with multiple chronic diseases: a latent class analysis of a discrete choice experiment. Quality of Life Research. 2022.

334. Weller BE, Bowen NK, Faubert SJ. Latent Class Analysis: A Guide to Best Practice. Journal of Black Psychology. 2020;46(4):287-311.

335. Yoo HI. lclogit2: An enhanced command to fit latent class conditional logit models. The Stata Journal. 2020;20(2):405-25.

336. Hensher DA, Greene WH. The Mixed Logit model: The state of practice. Transportation. 2003;30(2):133-76.

337. Hole A. WTP: Stata module to estimate confidence intervals for willingness to pay measures. 2007.

338. Hole AR. A comparison of approaches to estimating confidence intervals for willingness to pay measures. Health Econ. 2007;16(8):827-40.

339. Dziak JJ, Coffman DL, Lanza ST, Li R, Jermiin LS. Sensitivity and specificity of information criteria. Brief Bioinform. 2020;21(2):553-65.

340. Koopmanschap MA, Stolk EA, Koolman X. Dear policy maker: Have you made up your mind? A discrete choice experiment among policy makers and other health professionals. International Journal of Technology Assessment in Health Care. 2010;26(2):198-204.

341. Watson V, McCartan N, Krucien N, Abu V, Ikenwilo D, Emberton M, et al. Evaluating the Trade-Offs Men with Localized Prostate Cancer Make between the Risks and Benefits of Treatments: The COMPARE Study. The Journal of urology. 2020;204(2):273-80.

342. Ciani O, Manyara AM, Taylor RS. Surrogate endpoints in trials—a call for better reporting. BMJ (Clinical research ed). 2022;378:o1912.

343. Bikdeli B, Punnanithinont N, Akram Y, Lee I, Desai NR, Ross JS, et al. Two Decades of Cardiovascular Trials With Primary Surrogate Endpoints: 1990-2011. J Am Heart Assoc. 2017;6(3).

344. Culyer AJ. Equity - some theory and its policy implications. J Med Ethics. 2001;27(4):275-83.

345. Hampson G MD, Shah K, Devlin N. **Public Preferences for Health Gains and Cures: A Discrete Choice Experiment. OHE Consulting Report**. London: Office of Health Economics. 2018.

346. Whitty JA, Scuffham PA, Rundle-Thiele SR. Public and decision maker stated preferences for pharmaceutical subsidy decisions: a pilot study. Appl Health Econ Health Policy. 2011;9(2):73-9.

347. Ciani O, Grigore B, Taylor RS. Development of a framework and decision tool for the evaluation of health technologies based on surrogate endpoint evidence. Health Economics. 2022;n/a(n/a).

348. Vivot A, Jacot J, Zeitoun JD, Ravaud P, Crequit P, Porcher R. Clinical benefit, price and approval characteristics of FDA-approved new drugs for treating advanced solid cancer, 2000–2015. Annals of Oncology. 2017;28(5):1111-6.

349. Bostwick DG. Prostate-specific antigen. Current role in diagnostic pathology of prostate cancer. American journal of clinical pathology. 1994;102(4 Suppl 1):S31-7.

350. Williams S. Surrogate endpoints in early prostate cancer research. Transl Androl Urol. 2018;7(3):472-82.

351. Adaptive Platform Trials Coalition. Adaptive platform trials: definition, design, conduct and reporting considerations. Nat Rev Drug Discov. 2019;18(10):797-807.

352. Bognar K, Romley JA, Bae JP, Murray J, Chou JW, Lakdawalla DN. The role of imperfect surrogate endpoint information in drug approval and reimbursement decisions. Journal of Health Economics. 2017;51:1-12.

353. Zhang X, Zhang Y, Ye X, Guo X, Zhang T, He J. Overview of phase IV clinical trials for postmarket drug safety surveillance: a status report from the ClinicalTrials.gov registry. BMJ Open. 2016;6(11):e010643.

354. Muthuswamy V. Ethical issues in clinical research. Perspect Clin Res. 2013;4(1):9-13.

355. Djulbegovic B, Lacevic M, Cantor A, Fields KK, Bennett CL, Adams JR, et al. The uncertainty principle and industry-sponsored research. Lancet (London, England). 2000;356(9230):635-8.

356. Harmon A. New drugs stir debate on rules of clinical trials. New York Times. 2010 18/09/2010.

357. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2013;16(2):231-50.

358. Artibani W, Porcaro AB, De Marco V, Cerruto MA, Siracusano S. Management of Biochemical Recurrence after Primary Curative Treatment for Prostate Cancer: A Review. Urologia internationalis. 2018;100(3):251-62.

359. Msaouel P, Pissimissis N, Halapas A, Koutsilieris M. Mechanisms of bone metastasis in prostate cancer: clinical implications. Best Practice & Research Clinical Endocrinology & Metabolism. 2008;22(2):341-55.

360. Shen G, Deng H, Hu S, Jia Z. Comparison of choline-PET/CT, MRI, SPECT, and bone scintigraphy in the diagnosis of bone metastases in patients with prostate cancer: a meta-analysis. Skeletal Radiol. 2014;43(11):1503-13.

361. Turpin A, Girard E, Baillet C, Pasquier D, Olivier J, Villers A, et al. Imaging for Metastasis in Prostate Cancer: A Review of the Literature. Frontiers in oncology. 2020;10:55-.

362. Johansson J-E, Holmberg L, Johansson S, Bergström R, Adami H-O. Fifteen-Year Survival in Prostate Cancer: A Prospective, Population-Based Study in Sweden. Jama. 1997;277(6):467-71.

363. Austin PC, Lee DS, Fine JP. Introduction to the Analysis of Survival Data in the Presence of Competing Risks. Circulation. 2016;133(6):601-9.

364. Erinjeri JP, Clark TWI. Cryoablation: mechanism of action and devices. J Vasc Interv Radiol. 2010;21(8 Suppl):S187-S91.

365. Harris WP, Mostaghel EA, Nelson PS, Montgomery B. Androgen deprivation therapy: progress in understanding mechanisms of resistance and optimizing androgen depletion. Nat Clin Pract Urol. 2009;6(2):76-85.

366. Australia CC. TNM classification of prostate tumours 2020 [Available from: https://wiki.cancer.org.au/australia/Guidelines:Prostate_cancer/Management/Locally_advanced_and_metastatic/TNM_classification_of_prostate_tumours.

367. Hermanek P, Sobin LH. TNM classification of malignant tumours: Springer Science & Business Media; 2012.

368. Divo MJ, Martinez CH, Mannino DM. Ageing and the epidemiology of multimorbidity. Eur Respir J. 2014;44(4):1055-68.

369. Australian Government DoH. PBAC Guidelines. 2.3 Trial design and execution 2016 [Available from: https://pbac.pbs.gov.au/section-2/2-3-trial-design-and-execution.html.