QUANTIFYING THE BENEFITS AND HARMS OF TREATING DUCTAL CARCINOMA IN SITU FOR USE IN THE ECONOMIC EVALUATION OF BREAST CANCER SCREENING PROGRAMMES

By

HANNAH LOUISE BROMLEY

ORCID ID: 0000-0002-6650-1639

A thesis submitted to the University of Birmingham and the University of Melbourne in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2019

Institute of Applied Health Research, University of Birmingham

Melbourne School of Population and Global Health, University of Melbourne

Abstract

Evaluation of breast cancer screening programmes present results in terms of cost per QALY but fail to include any disutility for patients who may have been overdiagnosed and receive unnecessary treatment. Ductal carcinoma in situ (DCIS) is a heterogeneous disease likely to account for a large proportion of overdiagnosis. This thesis presents an overview of the challenges associated with valuing breast cancer states, using quantitative and qualitative evidence to demonstrate that the inclusion of the harm from overdiagnosis is important in the economic evaluation of breast cancer interventions.

A systematic review found the utilities informing breast screening policy were unlikely to fully capture all benefits and risks. An empirical study deriving utilities from 172 women for health states explicitly describing DCIS treatment were lower than those from the literature, where it is unlikely that individuals were informed treatment may be unnecessary. The utilities were applied in an illustrative model and suggested this disutility had the potential to change the cost-effectiveness of treating low-risk, screen detected disease. Qualitative interviews of 26 patients validated overdiagnosis was important in treatment decisions. This thesis contributes important methodological considerations for modelling breast cancer interventions and suggests future research to validate this harm in economic evaluations.

Declaration

This is to certify that:

i. This thesis comprises only my original work towards the Degree of Doctor of Philosophy;

ii. Due acknowledgement has been made in the text to all other material used;

iii. The thesis is fewer than 80,000 words in length, exclusive of tables, figures, references and appendices.

Hannah Louise Bromley

01/09/2019

For Adele

I hope this thesis serves justice to the many you helped,

And the women you inspired into research.

Acknowledgements

Firstly, to my supervisors Professor Tracy Roberts, Doctor Carolyn Nickson, Professor Dennis Petrie, Professor Bruce Mann, Professor Daniel Rea and the late Professor Adele Francis, I really did hit the supervisor jackpot. I could not have wished for a better team and I am eternally grateful for all your hard-work, support and endless patience running late-night supervisions. It might have taken us four years, but you have turned me into a health economist at last!

I would like to thank my funders at the University of Melbourne and the University of Birmingham for awarding me a joint Melbourne International Research Scholarship as part of the Universitás 21 initiative, without which this research would not have been made possible. I am also indebted to Jennie Young and the LORIS trial team for their ongoing support.

I would also like to thank the many women who participated in the study pilot, empirical analysis and qualitative interviews. Without your time and good-will, I would never have completed my analysis for which I am humbled. Special thanks must also go to Lisa Devereux, Grant Lee and Lisa Morstyn from the Breast Cancer Network Australia and Lifepool registries for their assistance with recruitment. Although it took a few months, things really got going in the end!

Much of this research has been presented at completion seminars, local institute lectures and international breast cancer and health economics conferences. I am especially grateful for the opportunity to present my work and for the comments and suggestions

received, both of which have helped enormously to shape the direction of the discussion.

Of course, I could not have written this thesis without the love and support of my wonderful colleagues, friends and family. I would be lost without your quick wit and support. Your ability to offer help, advice and a handful of aptly themed jokes have inspired me to persevere throughout the trials and tribulations of the PhD. To Rebecca, for the many dinners and cups of tea you have provided during the writing stage, I am eternally grateful as, I am sure is my mother! And Anthony, I thank you for your patience and understanding whilst I jetted off to pursue my dreams. I hope you enjoyed your visit down under as much as I did.

Special thanks must go to my companions in Australia. To the folks at International House college thank you for being my home away from home. To all the wonderful girls at Melbourne University Women's Soccer and Cricket Clubs, thank you for helping me to settle into life in Melbourne. Despite some of the trickier times collecting data, your comradery made me feel part of the family, even 12,000 miles away from home, and I truly do miss our Sunday road trips around Victoria (all day!) and goal side celebrations.

Finally, I would like to pay tribute to the late Adele Francis, who sadly passed away in the second year of my candidature. I will fondly remember our time scrubbed in theatre listening to One Direction. I hope this thesis pays homage to your legacy. I will continue to learn from your advice, good humour and humility for many more years to come.

Publications to date

In Press:

Bromley et al. (2019) Valuing the health states associated with breast cancer screening programmes: A systematic review of economic measures. *Social Science and Medicine* 228:142-154.

Bromley et al. (2019) Valuing preferences for treating screen detected ductal carcinoma in situ. *European Journal of Cancer* 123:130-137.

Papers under review:

Bromley et al. (2019) What factors influence treatment decision for low-risk ductal carcinoma in situ? *Submitted*

Select conference presentations:

- 'Valuing Health States Associated with Breast Cancer: a primary study exploring methods, challenges and the issue of overdiagnosis.' Paper presented in part at the Health Economics Study Group meeting, Norwich, UK; July 3-5, 2019.
- 'Active monitoring is an acceptable choice for women with low risk ductal carcinoma in situ: An assessment of quality of Life.' Paper presented in part at the 4th World Congress on Controversies in Breast Cancer and Melbourne International Joint Breast Congress, Melbourne, Australia; October 11-13, 2018.
- 'Valuing Health States Associated with Breast Cancer Screening and Overdiagnosis: A Systematic Review.' Paper presented in part at the Australian Health Economics Association Conference, Tasmania, Australia; September 25-26, 2018.
- 'Identifying the utilities and disutilities associated with the overdiagnosis of early breast cancers in breast screening programmes.' Paper presented in part at the European Health Economics Association Conference, Hamburg, Germany; July 13-16, 2016.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AM; Active monitoring

AQoL; Assessment of Quality of Life

BCS; Breast conserving surgery

BCNA; Breast Cancer Network Australia

BSP; Breast screening programme

CBA; Cost-benefit analysis

CEA; Cost-effectiveness analysis

CEAC; Cost-effectiveness acceptability curve

CI; Confidence interval

CUA; Cost-utility analysis

DES; Discrete event simulation

DCIS; Ductal carcinoma in situ

EQ-5D; EuroQol 5 dimensions

ER/PR/HER2; Oestrogen, progesterone, Herceptin 2 receptor

EUT; Expected utility theory

HRQoL: Health related quality of life

HSUV: Health state utility value

HUI; Health Utilities Index

IBC; Invasive breast cancer

ICER; Incremental cost-effectiveness ratio

ISM; Individual sampling model

LORIS: Low-Risk DCIS trial

LYG; Life years gained

MAUI; Multi-attribute utility instrument

NHS; National Health Service

NICE; National Institute for Health and Care Excellence

NMB; Net monetary benefit

PBM; Preference-based measure

PSA; Probabilistic sensitivity analysis

QALY; Quality adjusted life year

QoL; Quality-of-life

RCT; Randomised controlled trial

RR; Relative risk

RT; Radiotherapy

SEER; Surveillance, epidemiology and end results

SF-6D; Short-Form 6 dimensions

SG; Standard gamble

TTO; Time trade off

VAS; Visual analogue scale

WHO; World Health Organisation

WW: Watchful waiting

UK NSC: United Kingdom National Screening Committee

CHAPTER 1

Introduction and aims of the thesis

1.1 Introduction

Breast cancer is a major cause of morbidity and mortality (1), and resources are spent to prevent, diagnose and treat the disease early. Mammography screening has lowered the cost and morbidity of treatment and improved survival (2). However, the tests used to screen for breast cancer are imperfect and may result in the overdiagnosis and overtreatment of low-risk disease such as ductal carcinoma in situ (DCIS). There are concerns about whether the implications of overdiagnosis are adequately captured in the economic evaluation of breast cancer screening programmes (3). This thesis uses DCIS as a vehicle to explore whether the health states and utilities informing breast screening policy are valued appropriately.

The aim of the thesis is to identify and quantify the benefits and harms of treating ductal carcinoma in situ. This opening chapter introduces the concept of utility and the valuation of health benefits for use in economic evaluation. In doing so, the challenges of quantifying the sequelae of breast cancer screening inspiring this research are briefly discussed. An overview of the disease characteristics and treatment is provided but is expanded further in the following chapters. The aims and objectives of the thesis are then outlined, and lastly the structure of the overall thesis is set out.

1.2 Economic evaluation in healthcare

Economic evaluation plays an important role in the allocation of finite resources (4). Healthcare funds are scarce (5), and decisions must be made as to the most efficient

intervention or programme to fund. There will always exist an "opportunity cost" from the benefits forgone by not selecting an alternative option (6). Economic evaluation facilitates the objective assessment of the relative benefits and costs (7), so that decision makers can choose the strategy with the most efficient net outcome.

It is necessary to have information on the costs and benefits of one or more interventions for economic analysis (8). This usually comprises of one or more new treatments or programmes against standard practice. There are two frameworks for evaluating healthcare benefits, namely welfarism and extra-welfarism, both with different methodological and theoretical basis (9).

Briefly, welfarism appraises healthcare according to the extent to which it contributes to maximise overall social wellbeing or "utility" (10). Utility in this sense is defined in terms of the consumption of goods and the satisfaction derived (4). Within welfarism, health services are some of the many goods that may be consumed, and utility may be gained by improving outcomes outside the domains of health (11). This requires the use of a cost-benefit approach where both benefits and costs are measured in monetary terms (e.g. willingness to pay). Such methods are conventionally adopted in other publicly funded sectors, such as transport and environmental sciences, but have limited application in healthcare due to heterogeneity in the perception of ability to pay for health among other methodological challenges (9).

In contrast, the extra-welfarism approach seeks to maximise utility in terms of health contribution (9). Health is viewed as a physical entity of which individuals have stock, and of which gains may be produced (as health outcomes or "health-related utility") through the allocation of health resources (11). When outcomes are measured in costs and natural units

specific to the nature of the disease (e.g. blood pressure, life years gained etc.), this is termed a cost-effectiveness analysis (CEA). A variant of cost-effectiveness analysis is the cost-utility analysis (CUA) in which outcomes are measured in terms of a single unit measure, namely the cost per Quality-Adjusted-Life-Year (QALY). QALYs combine survival and quality of life, permitting the direct comparison of outcomes across different conditions or interventions (12), and are the measure of cost-effectiveness recommended by decision makers such as the National Institute for Health and Care Excellence (NICE). Extra-welfarism was introduced to expand the evaluative space to include health in policy decisions, and although Coast (11) and Lorgelly (13) argue that it may in fact narrow the valuation of utility to include only health, it is widely adopted in the appraisal of health services.

1.3 Valuing health states and health-related utility

The use of QALYs in economic evaluation requires the numerical adjustment of life expectancy to reflect underlying health (6). Utilities are cardinal measures of quality of life that represent the strength of preference for a health state or intervention (14). They are measured on a scale from 0 to 1, equivalent to being dead and in perfect health respectively. Utility here (and for the remainder of the thesis) is defined using an extra-welfarism framework and focuses on health dimensions, in keeping with national recommendations for the economic evaluation of health interventions in the United Kingdom (UK) (8). Utilities are used to inform QALYs and are therefore integral to the way healthcare is measured.

The issue of how to elicit or whose utilities should be measured is a matter of significant debate (15). Traditionally, utilities were commonly derived by clinical experts using direct approaches such as the visual analogue scale, standard gamble or time-trade-off (16). However, concerns about the ability of providers to accurately determine patient health has

led many to argue that utilities should, in fact be elicited from patients (17). More recently there has been movement toward the valuation of utilities by the general population using generic utility instruments (e.g. EuroQol-5D), given that healthcare is primarily publicly funded (18). Regardless of the approach chosen, the valuation of utility must be undertaken using valid and reliable methods relevant to the condition under study. Furthermore, it is important that utilities capture all the relevant benefits and harms during the valuation process. This concern is central to the research presented in this thesis and will be considered in more detail in subsequent chapters.

1.4 Breast cancer screening programmes

Breast cancer screening programmes were introduced in the late 1980s to reduce the morbidity and mortality associated with the rise in breast cancer among the general population (19). The purpose of screening is to advance the time of diagnosis so that survival may be improved through earlier intervention (20). Digital mammography has largely replaced film screening in recent years, but eligible women are routinely invited and are informed of the result by letter (21). In the United Kingdom, women aged 50-70 years are invited for screening every three years. Elsewhere in continental Europe and Australia, mammography screening is conducted biennially up to the age of 74 years (1).

Since the introduction of breast screening programmes, there has been debate about the magnitude of the benefits and harms to the general population (3). The expected benefits of mammography are the reduction in breast cancer mortality and in morbidity from less intensive treatment (22). The major harm among others (e.g. false positives and irradiation) is overdiagnosis and its sequelae. Overdiagnosis refers to *"the diagnosis of disease that would otherwise not go on to cause symptoms or death"* ((23): p.605). Estimates of overdiagnosis

are diverse and widely contested, with reported rates from meta-analyses varying between 0% and 54% (24). Nonetheless, there is agreement that whilst both invasive and in situ cancers account for those overtreated (25) a significant proportion of overdiagnosis is associated with low-risk disease, specifically ductal carcinoma in situ.

The controversy surrounding national breast screening programmes intensified in 2012 following publication of the UK Independent Panel Report (3) on the net benefits and harms of breast cancer screening. The panel concluded that for every life saved by breast cancer screening, three women potentially received unnecessary treatment as a result of overdiagnosis. Consequently, there has been increasing debate among both clinicians and health economists about national breast screening policy (26). There is a call to reform the current strategy in line with changes in technology and incidences of the disease, so as to maximise the benefits for high-risk women but limit the harms for those with lower baseline risk of invasive progression.

1.5 Why ductal carcinoma in situ?

The pathways associated with the sequelae of mammography screening are complex. To model the entirety of the benefits and harms associated with invasive and non-invasive disease would be beyond the scope of this thesis. Therefore, for the purposes of this research ductal carcinoma in situ is used as a vehicle to explore whether the utilities and disutilities of breast screening are captured appropriately. It is hypothesized that only 10-30% of low-grade DCIS will progress into invasive breast cancer during the lifetime of those diagnosed (27), and many clinicians think that this estimate may be lower with surveillance. The condition thus provides an ideal case study to explore the benefits and harms following breast cancer screening.

Most women undergo breast conserving surgery, with or without radiotherapy, and endocrine treatment (for those with receptor positive disease) following diagnosis of DCIS (28). However, there is evidence to suggest that active monitoring of low-risk DCIS may reduce the harms of unnecessary treatment resulting from overdiagnosis (29). Phase III randomised controlled trials such as the Low Risk DCIS (LORIS) trial (30) in the UK are underway to explore the clinical safety and efficacy of active monitoring, yet little is known about the potential implications of such strategies upon quality of life or breast cancer screening programmes.

1.6 Rationale for the thesis

Mammography screening is deemed cost-effective for women aged 50-70 in the UK general population. However, such policy decisions are only as robust as the evidence informing them. The national breast cancer screening programme, introduced after the Forrest Report (19), concluded screening was a cost-effective method for the UK National Health Service (NHS) to reduce breast cancer morbidity and mortality by approximately a third. Whilst the report did apply a cost per QALY approach, the model did not include any disutilities beyond the initial surgery in the analysis.

In light of increasing concern about the benefits and harms of screening, Raftery and Chorozoglou (31) updated the original lifetable model from the Forrest report to include the harms of false positives and the side-effects from treatment. The explicit inclusion of this additional disutility was found to change the relative cost-effectiveness of the British breast screening programme for the initial eight years of screening and further raised questions as to whether the decision on the current breast screening strategy was adequately informed. However, the authors were explicit in that they did not include the harm of overdiagnosis in

their analysis due to a lack of published data around the disutility of unnecessary treatment. Specifically, they state:

"None of the surveys of quality of life included overtreatment, implicitly assuming all surgery was necessary. To answer this question, surveys would have to ask each woman whether her quality of life would be affected if it could be shown that her surgery had been unnecessary." ((31): p.8)

The difficulty of working within the breast screening remit is that the natural history of the disease is not well understood (32). It is therefore difficult to predict which women will be overdiagnosed at the individual level. Nonetheless, at a societal level there is an obligation for policy makers to ensure that the benefits and harms of screening are adequately captured so that the most efficient strategy is implemented. Misplaced decisions about screening not only pose questions about inappropriate resource use or opportunity costs forgone, but also raise ethical issues associated with inappropriate cancer diagnoses and side-effects from treatment on women who ultimately may not have needed it.

The importance of using utilities that accurately reflect the benefits and harms of the process is therefore of high importance in informing the breast screening debate. Whilst much of the literature to date has focused on clinical outcomes and estimates of overdiagnosis (3, 33), little has been done to address the economic concerns and implications that this might have on screening policy. It is unclear whether screening may or may not be based on missing information. An evaluation of how the benefits and harms are conceptualised in breast cancer screening policy is explored in this thesis, using ductal carcinoma as a vehicle to explore outcomes relevant to overdiagnosis.

1.7 Significance of the research

It is important that decision makers are adequately informed about the risks, as well as the benefits, in the appraisal of the likely implications of breast screening strategies. There is a need to critique the current evidence-base to determine how such health states have been conceptualised and valued, and whether the utilities applied have been valued appropriately.

Furthermore, it is important that individuals valuing screening and the treatment of low-risk disease are aware of the likely benefits and risks in the measurement of quality of life. Without reliable and consistent evidence to inform the economic evaluation of breast screening programmes, women with low-risk disease such as DCIS may be subject to inappropriate health resource use and overtreatment.

1.8 Aim and objectives

The aim of the thesis is to quantify the benefits and harms associated with the treatment of ductal carcinoma in situ for use in the economic evaluation of breast cancer screening programmes.

Four study objectives are explored to address this research question:

1. To ascertain how the benefits and harms associated with breast cancer screening programmes are currently conceptualised and valued.
2. To value the health-related quality of life (utility) associated with different treatments for low-risk DCIS, explicitly including the risk of overdiagnosis.
3. To understand the factors which influence women's preferences in the trade-off between the management of DCIS and the risk of overtreatment.

4. To develop an economic model of low-risk DCIS, predicting costs and outcomes for a hypothetical cohort of patients using the derived utilities, and to consider the potential impact of the findings upon the future evaluation of breast cancer screening programmes.

1.9 Thesis outline

This chapter has introduced the key concepts and purpose for undertaking the research presented throughout the thesis. The subsequent chapters describe the findings from each stage of the research, namely the review of the literature, empirical study, qualitative interviews and economic model, and are structured as follows:

Chapter 2 summarises the debate on breast cancer screening programmes, and highlights the challenges associated with valuing health states related to overdiagnosis and DCIS. The Low-Risk DCIS (LORIS) trial is also described from which this thesis was motivated.

The principles of economic evaluation and health state valuation are described in **Chapter 3.** In doing so, the theoretical foundations and methodology underpinning the economic measures applied in the empirical analysis of utility are explored in detail. Whilst predominantly descriptive in nature, the ideologies discussed are essential for understanding the theory applied later in the thesis.

Chapter 4 presents a systematic review of studies that attempted to use or measure utilities in the economic evaluation of breast cancer screening programmes. The review was conducted to determine how the benefits and harms associated with the sequelae of screening are valued, including the overdiagnosis and overtreatment of low-risk disease.

The methodology for an empirical study valuing health states for DCIS is reported in **Chapter 5**. The objective of this chapter was to address the limitations identified in the systematic review. The chapter concludes with the theoretical recommendations as to what might be an appropriate approach for capturing the disutility of overdiagnosis from first principles.

Chapter 6 presents the main results from the empirical study valuing utilities for low-risk DCIS. This chapter is separated into two parts. The main clinical and health economic findings are presented in part one, and in part two, the methodological limitations are outlined in relation to the data collected.

A review of economic evaluations for the treatment of ductal carcinoma in situ is presented in **Chapter 7.** The aim of this chapter was to identify a suitable vehicle for the development of an economic model of DCIS treatment, specifically active monitoring versus standard treatment.

The construction of an economic model of low-risk DCIS treatment is summarised in **Chapter 8**. The principal aim of the model was to illustrate the potential implications of including the disutility of overdiagnosis (from applying the utilities elicited versus those from the literature) in an economic analysis depicting the LORIS trial.

In **Chapter 9** a qualitative analysis of the preferences of breast cancer patients for managing low-risk DCIS is presented. This chapter supplements the findings from the empirical study and model by providing an in-depth analysis of the factors influencing the utilities derived.

Chapter 10 revisits the objectives specified at the beginning of this chapter. The main findings from each chapter are consolidated and compared with the literature, along with a critical interpretation of the strengths and limitations of the work undertaken. The chapter concludes with an evaluation of the potential research implications of the main findings and recommendations for future research.

CHAPTER SUMMARY

What is known?

- Economic evaluations provide evidence for decision-makers on how best to allocate limited healthcare resources.
- Many decision makers prefer a cost per QALY approach and require utilities to adjust life expectancy by the quality of life of a health state.
- The benefits and harms associated with breast cancer screening programmes are widely contested, yet little research has been undertaken on how these have been conceptualised in economic evaluation.

What this chapter adds

- This thesis aims to quantify the benefits and harms associated with the treatment of DCIS for use in the economic evaluation of breast screening programmes.
- Ductal carcinoma in situ provides a useful vehicle in which to explore how the benefits and harms of breast cancer screening are captured in economic evaluations, including the risk of overdiagnosis and potential for overtreatment.
- The ten chapters in the thesis provide a framework in which to explore how women's preferences of screening for and treating DCIS are captured in the evidence informing breast cancer decision-making.

CHAPTER 2

Breast cancer screening and ductal carcinoma in situ

2.1 Introduction

In this chapter a clinical summary of breast cancer screening programmes and ductal carcinoma in situ are reported, alongside reasons for using this condition as a case study for exploring the disutility associated with overdiagnosis. The chapter is structured as follows; first a summary of the debate surrounding breast cancer screening programmes is presented, followed by a brief discussion of potential benefits and harms. Second, an overview of ductal carcinoma in situ is given, including its incidence, management and relevance to breast cancer screening and treatment. Finally, a summary is provided of the Low-Risk Ductal Carcinoma in Situ (LORIS) trial, specifically its importance and potential application in the debate about overdiagnosis and breast cancer screening programmes.

2.2 Breast cancer and demand on healthcare

2.2.1 Overview

Breast cancer is a major cause of morbidity and mortality. It is the most common cancer among women worldwide and the second highest cause of cancer death (34). Approximately 55,200 women are diagnosed with invasive breast cancer in the United Kingdom every year (35) and 19,535 across Australia (36). Observed incidence rates have increased in all age groups since the 1990s with around 300 new cases diagnosed per 100,000 women each year (37), but numbers are even higher in countries with established population screening programmes including Australia, the USA and UK.

Breast cancer is defined as malignant disease that has developed from the growth of abnormal or cancerous cells which originate from the tissues of the breast (20). Each breast is composed of a structure of mammary glands surrounded by a fatty connective tissue stroma which gives the breast its characteristic shape (1). Mammary glands are modified sweat glands and consist of a series of milk ducts and secretory lobules necessary for lactation (35). Breast cancer diagnosis and management is thus characterised by the origin of the abnormal cells within the breast and their propensity to grow and divide, as well as the protein receptors expressed on the cell surface (1).

Breast cancers are primarily characterised into non-invasive and invasive disease (2). Non-invasive disease (in situ or stage zero breast carcinoma) is considered pre-cancerous in nature as the abnormal cells have not spread beyond the epithelium and into the surrounding normal breast tissue (32). Ductal carcinoma in situ (DCIS) is the most common type of non-invasive breast cancer and derives from the milk ducts of the breast (35). Lobular carcinoma in situ (LCIS) stems from the milk producing glands (lobules) at the ends of the ducts. Both DCIS and LCIS are considered non-invasive but may increase the risk of developing an invasive breast cancer later in life (36).

In situ breast cancers may be further characterised by grade based on the differentiation of the cancer cells under the microscope compared with normal breast cells (1). This is of importance as the propensity for invasive progression is likely to depend on the predicted growth and appearance of the cells identified (32). DCIS is graded as low, moderate or high-grade disease (28). Low grade DCIS has an almost normal appearance and growth rate on histopathological examination, and a lower risk of invasive progression and recurrence (25).

In contrast, moderate and high-grade DCIS appear more abnormal, have higher cell turnover and are more likely to develop into invasive breast cancer (25).

Invasive breast cancer refers to malignant breast disease whereby the cancerous cells have grown into the surrounding breast tissue (1). Invasive breast cancers may originate from in situ disease in the ducts, milk lobules or (less commonly) the surrounding connective tissue, nipple or chest wall of the breast (2). If untreated invasive breast cancer may metastasize to the lymph nodes and potentially to other parts of the body (21). Early breast cancer (stage I-II) encompasses any invasive breast cancer that is contained within the breast or surrounding lymph nodes in the axilla but not elsewhere in the body (Table 2.1). Locally advanced or loco-regional breast cancer (stage III) involves invasive breast that is either larger in size (>50mm) or cancer which has spread beyond the breast to the nearby skin, chest wall musculature or ipsilateral lymph nodes surrounding the breast. Metastatic or advanced breast cancer (stage IV) suggests that the cancer has spread to other parts of the body such as the lungs, liver, brain or bones (1). Breast cancers are also characterised by their hormone receptor status. Two thirds of breast cancers are hormone receptor positive and require female hormones (oestrogen +/- progesterone) to grow and develop (37). Similarly, HER-2 positive breast cancers have excess surface protein human epidermal growth factor receptor-2 which promotes the growth of cancer cells (1).

Mammography screening programmes were introduced to counteract the rise in breast cancer harm and significant demand on healthcare (19). According to US surveillance, epidemiology and end results (SEER) risk data, 12% of women will develop breast cancer during their lifetime and 3% will die from the disease (38). Trends in breast cancer-specific mortality have decreased in most developed countries as a likely result of earlier detection through

screening, better breast awareness and improvements in treatment (39). Prognosis for women diagnosed with in situ breast cancer is particularly favourable. At 20-years, breast cancer-specific survival among 108,196 women with DCIS was 96.7% (40).

Table 2.1: A summary of breast cancer staging

Stage	**Description**
Stage 0 DCIS LCIS	Tumours that have not grown beyond their state of origin into the surrounding breast tissue: (low/moderate/high grade disease) - Ductal carcinoma in situ: non-invasive disease which originates in the milk ducts - Lobular carcinoma in situ: non-invasive disease which originates in the milk glands
Stage 1	Invasive tumour <20mm with no evidence of lymphatic involvement
Stage 2	Invasive tumour >20mm but <50mm in diameter +/- involvement of up to 1-3 axillary lymph nodes
Stage 3	Invasive tumour >50 mm +/- involvement of up to 4-10 axillary lymph nodes or local structures (e.g. chest wall)
Stage 4	Invasive tumour of any size with distant metastases

Whilst overall incidence of breast cancer has remained stable over the last few years, in situ breast cancer rates have almost tripled following the introduction of breast cancer screening programmes (35). Incidence trends, like most cancer types, reflect the changing epidemiology and prevalence of risk factors in the population (41) (e.g. age, family history, lifestyle factors – obesity, alcohol consumption and reduced physical activity, nulliparity, prolonged hormonal therapy etc.), improvements in diagnostic technology and better data recording (35). Nonetheless, the introduction of breast cancer screening programmes in the late 1980s has also contributed to the dramatic rise in early stage disease for those aged 50-74.

Figure 2.1 illustrates the change in breast cancer patterns observed in Australian women over the last 40 years (42). One would expect a stage shift from regional and metastatic breast

cancer to in situ disease, based on the criteria for successful screening programmes set out by the World Health Organisation (WHO) (43). However, despite substantial increases in the number of in situ breast cancers (both lobular and ductal carcinoma in situ encompassing the milk glands or ducts), screening has only marginally decreased the rate at which women present with late-stage or advanced breast cancer (44). Although the burden of breast cancer mortality has decreased, there is also likely to be increased demand on health services from the excess detection of early stage cancer that is not matched by a reduction in late-stage disease (45).

2.2.2 Breast cancer screening programmes

The national breast cancer screening programme was introduced to reduce breast cancer morbidity and mortality in the general population through the identification of cancer at an earlier stage, when treatment may be improved by earlier intervention (3). Currently, women aged 50-70 years in the UK are routinely invited for a mammogram every three years. Breast screening policy varies widely internationally, however, with women aged 50-74 invited biennially in Australia and Western Europe or annually in those aged ≥40 years in some US based clinics (46).

The typical pathway for a woman invited to routine breast cancer screening is shown in Figure 2.2. Eligible women are invited to participate by letter and are asked to attend a specialist mobile unit or clinic to have a mammogram (x-ray of the breasts). Digital mammography has generally replaced film mammography, and those with mammographically dense breasts may be offered digital breast tomosynthesis (47, 48). The screening images are reviewed by specialist radiologists. Women with abnormal mammograms (true or false positives) are recalled for further investigation, usually further

imaging with or without core biopsy. Women are subsequently reassured and are informed of their results (rates of false negative are low) or diagnosed with breast cancer and sent for treatment with the multidisciplinary breast team. The average screening cycle in the National Health Service (NHS) system from invitation to result takes approximately 28 days (49).

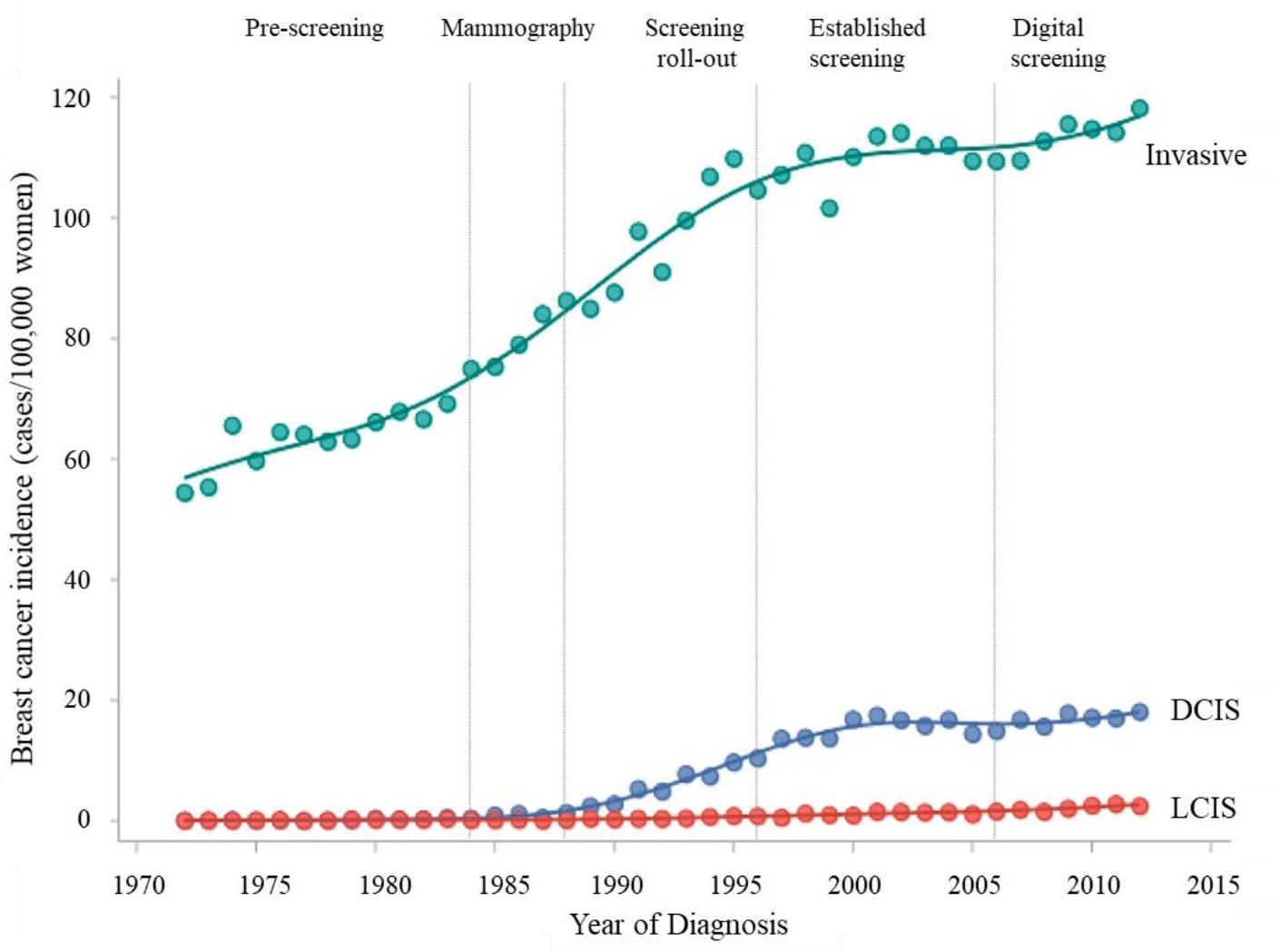


Figure 2.1: Breast cancer incidence in women aged 50-69 years in New South Wales, Australia (sourced from: (42))

*Legend: DCIS : ductal carcinoma in situ; LCIS: lobular carcinoma in situ; Invasive: Invasive breast cancer

In the UK, the national breast screening programme is estimated to have an annual running cost of £96 million (50). Similarly, in Australia, an estimated 36% of the $331 million health expenditure on breast cancer in 2004-5 was spent on mammography screening services through Breast Screen Australia (51). The cost of screening may be offset by the reduced cost of treating less advanced screen detected cancers (52). However, with increasing numbers of women diagnosed with early stage breast cancer, mammography screening places significant demand on limited health resources and thus warrants rigorous health economic review (53).

2.2.3 The debate over the benefits and harms of breast cancer screening

There has been widespread debate over the magnitude of the benefits and harms associated with screening the general population for breast cancer. Mammography screening has been shown to improve patient survival and reduce the morbidity and cost of treatment through the identification of early stage cancer when outcomes are more favourable (3), but it also results in the diagnosis and treatment of some pre-cancerous lesions and cancers that would not otherwise have become symptomatic (54). The benefits and harms of breast cancer screening are subsequently considered.

Breast cancer screening undoubtedly saves lives. A Cochrane meta-analysis of 11 randomised controlled trials of breast screening with 13 years of follow-up estimated a 20% reduction in breast cancer mortality among women invited to screening (RR: 0.8, CI 0.73-0.89) (20). Systematic reviews of US and Australian data have also reported similar breast cancer mortality reductions of 25 to 31% in those invited to and attending breast cancer screening (33, 55). Women with screen detected early breast cancer also receive substantially less intensive treatment than non-screeners (56) and incur fewer costs (57). Furthermore, those

who attend screening and are given a true-negative result may also gain utility from the reassurance of knowing that they are cancer free (58).

However, there are increasing concerns about the potential harms of population breast screening. Mammography may identify benign abnormalities and cause unnecessary anxiety (false positives) and there may be harms associated with the screening process itself (e.g. irradiation) (59). Similarly, whilst screening may save the life of patients with high-grade or aggressive breast cancer (22), it is less likely to benefit those with low-risk disease or reduced life expectancy from comorbid conditions (60). The valuation of breast screening programmes is thus challenging for decision makers seeking to trade-off the uncertainties surrounding the benefits and harms.

In 2012, the UK Independent Panel for Breast Cancer Screening (3) estimated that 11% of the cancers diagnosed by screening programmes were overdiagnosis (i.e. for every life saved by breast cancer screening, three women were the subject of overdiagnosis and potentially unnecessary treatment). Furthermore, a substantial number of screening participants were recalled for further assessment of a positive screening test, of which only 12% had a cancer diagnosis and 88% did not (33). Such assessment involves further imaging, for example mammography and ultrasound, or potentially invasive procedures including core biopsy, all of which may induce significant anxiety and have potential morbidity (61).

Consequently, there has been increasing debate among both clinicians and health economists about national breast screening strategies and the need to maximise the benefits for women at higher risk of developing breast cancer and limit the harms (including overdiagnosis) for those with a lower baseline risk (62).

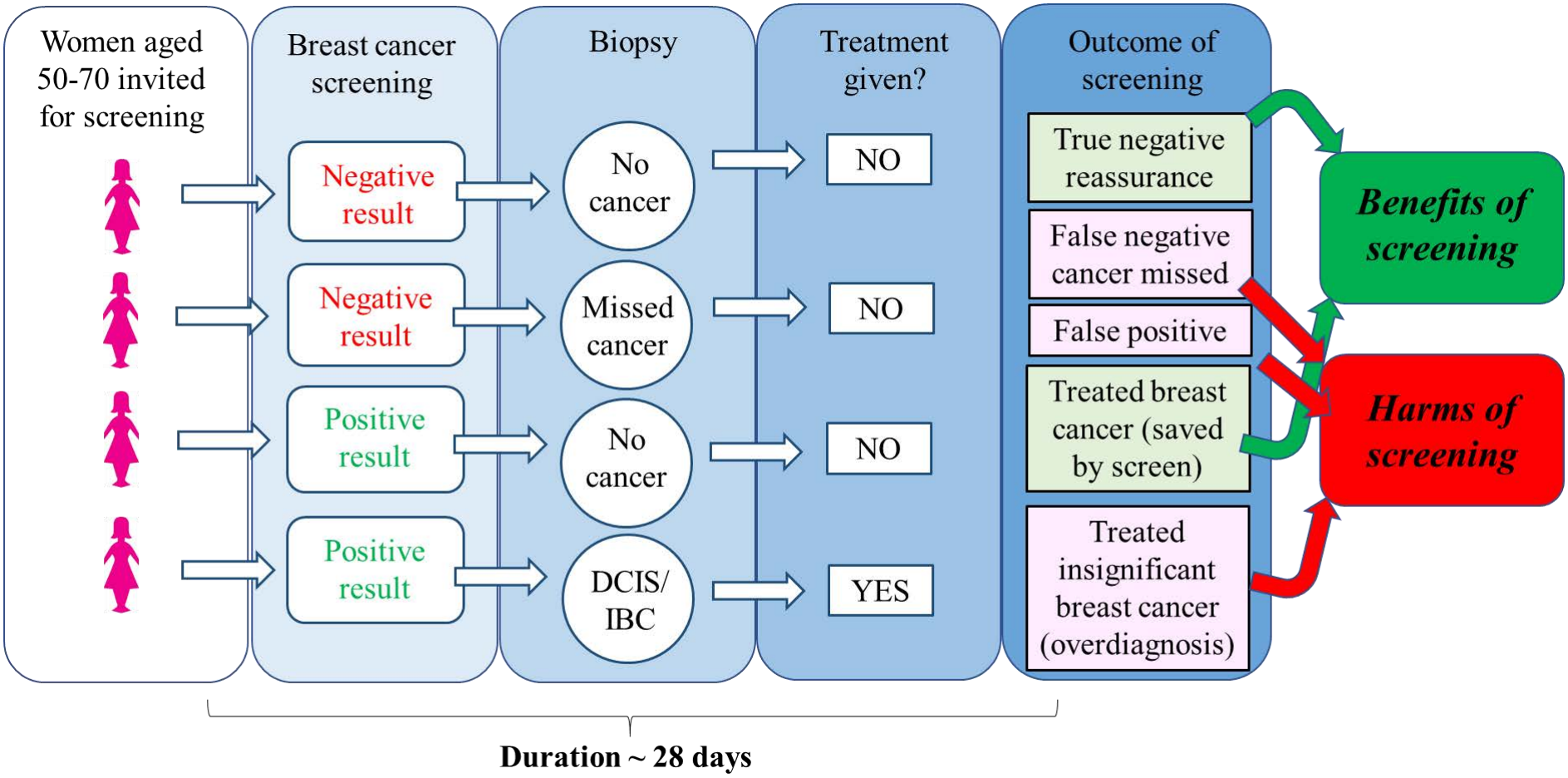


Figure 2.2: An illustration of the typical UK breast cancer screening pathway (adapted from: (63))

*Legend: DCIS: ductal carcinoma in situ; IBC: invasive breast cancer

2.2.4 Overdiagnosis

Overdiagnosis is defined as *"the diagnosis of cancer that would never cause clinical harm,"* ((27): p.6). It is important to distinguish overdiagnosis from false-positive cases arising from screening, in that patients have been histologically confirmed as having cancer and not a benign abnormality. Overdiagnosis refers to all cancers, invasive or in situ, although in situ disease is considered to account for a significantly higher proportion due to the non-invasive nature of the disease and uncertainty in natural history (64). There is no agreement on the extent of overdiagnosis. Rates are highly variable with estimates in one meta-analysis ranging from 0-54% (33), but many think it is substantially higher in those with low-risk DCIS and longer follow-up. This distinction is important because the rate of progression of untreated DCIS is critical to how patients and clinicians may value the benefit of treatment (65, 66).

Underpinning the concept of overdiagnosis is the hypothesis that different breast cancers grow at variable rates. As illustrated in Figure 2.3, breast cancer may be aggressive, slow growing or indolent in nature. Low-risk cancers such as ductal carcinoma in situ (DCIS) may regress, remain dormant or progress so slowly that they might never present symptomatically in the absence of screening (28). As a result, selection of these cancers at mammography turns women into cancer patients and may lead to treatment that may not be beneficial or may adversely impact quality of life and cause harm (67).

In any screening programme, a degree of overdiagnosis is inevitable because of competing causes of death. However, it is also possible that screening identifies some cancers that would never have become clinically detectable during the patient's lifetime (68). These cases, which would not have resulted in any morbidity or required treatment had they not been identified through screening, are therefore considered a harm of breast cancer screening programmes as

women are subjected to the potential morbidity and costs from incurring a cancer label and associated unnecessary treatment (27).

Overdiagnosis also has significant financial implications for society. Breast cancer costs the UK economy an estimated £1.5 billion annually (69). Therefore, if some women are receiving potentially unnecessary treatment as a result of overdiagnosis, resources are being spent inappropriately at the opportunity cost of healthcare expenditure elsewhere (4).

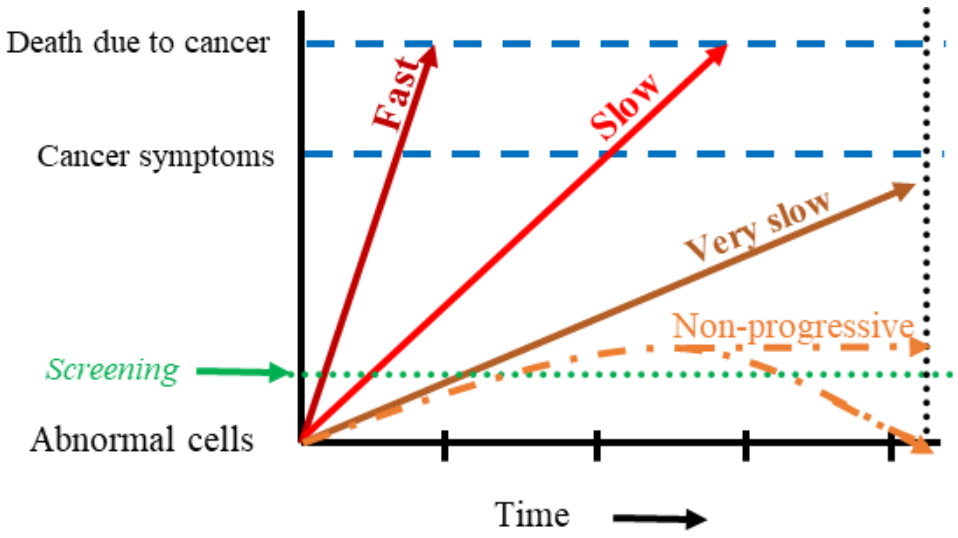


Figure 2.3: Different progression pathways for breast cancers identified at screening (adapted from: (28))

The national breast cancer screening programme was introduced in the UK following the publication of the Forrest report (19) which concluded that screening was a cost-effective method of reducing breast cancer mortality by approximately a third. Whilst the report used a recommended cost per QALY approach in the analysis of costs and benefits (8), the analysis did not include any harm on quality of life beyond the initial surgery for those with positive screens.

Raftery et al. (31) re-evaluated the economic model from the Forrest report after concerns arose surrounding the benefits and harms of mammography and the rise in overdiagnosis. After factoring in the disutility of false positive mammograms and treatment in the reported QALYs, they concluded that breast cancer screening programmes might have caused net harm for up to eight years, changing the relative cost-effectiveness of the current screening strategy. However, they were unable to include the disutility of overdiagnosis in their analysis due to a lack of published data. Whilst much of the literature has focused on the clinical aspects surrounding the rate of overdiagnosis (3, 24, 70), little has been done to address the economic concerns and implications of unnecessary treatment on breast cancer screening policy. This thesis has used ductal carcinoma in situ (DCIS) as a case study in order to appraise this harm.

2.3 Ductal carcinoma in situ

2.3.1 Why DCIS for capturing overdiagnosis?

Ductal carcinoma in situ is a pre-invasive condition encompassing a spectrum of abnormal cells confined within the mammary ducts of the breast (32). It is often considered a non-obligate precursor to invasive breast cancer because although the abnormal cells within the duct display signs suggestive of malignancy (71), they have not invaded into the surrounding tissue and therefore cannot technically be labelled as a cancer *per se*.

DCIS was deemed a suitable vehicle for exploring the challenges associated with valuing breast screening outcomes because:

(i) It is primarily a condition of screening

(ii) Low-risk DCIS has a slow pathway of progression (and is therefore hypothesised to account for a significant proportion of overdiagnosis)

(iii) Different DCIS grades may have different treatment options and clinical outcomes (and can therefore be used to model the different pathways following screening)

2.3.2 Incidence and economic burden

Since the introduction of the NHS Breast Screening Programme in 1988, the incidence of DCIS has increased substantially and now represents approximately 20-25% of screen detected breast cancer (72). In the UK alone, over 8,000 women are diagnosed with DCIS every year (35). Consequently, the condition places considerable strain on limited healthcare resources; the estimated cost to the health system of treating one woman with DCIS is £9,000 (73). With the planned extension of screening to include women aged 47-74 years (74), the relative impact on resource use is likely to increase.

Whether DCIS is a precursor and a risk factor for invasive breast cancer is unknown (33). The major challenge in the current understanding of DCIS is that it is not yet known which DCIS will develop into invasive breast cancer. Consequently, women with low-risk DCIS face being harmed by treatment without any benefit in survival (75, 76). The natural history of DCIS and breast cancer is not well understood. This uncertainty in understanding arises from the difficulty in observing disease progression over time, as the condition is typically treated upon diagnosis following screening and histopathological confirmation (28).

There are two hypotheses about the likely pathophysiology of in situ disease (77). The first theory ("multistep transformation") assumes that breast cancer develops through a multi-step linear transition, in keeping with the ideology behind the principles of screening and the

detection of earlier stage disease. Using this theory (78), increasing genetic alteration along the cell pathway turns normal epithelial cells into atypical aplasia, low then high grade in situ disease and finally invasive breast carcinoma (Figure 2.4). Evidence from older retrospective studies (79-81) support the hypothesis that DCIS is a precursive lesion to invasive ductal carcinoma.

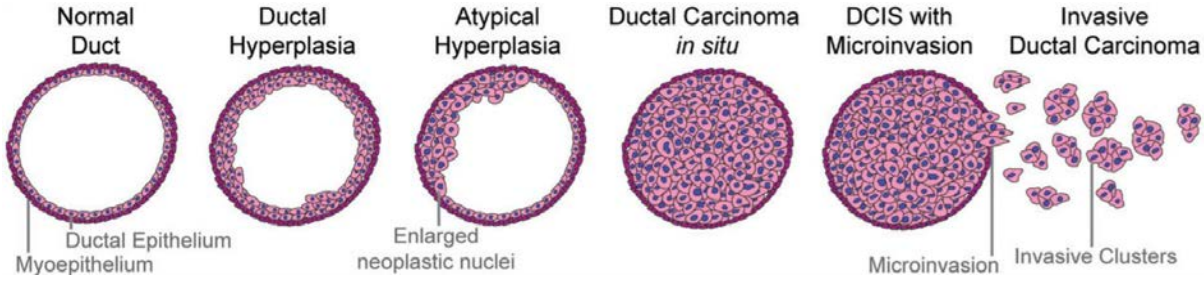


Figure 2.4: A schematic of possible breast carcinoma progression (Sourced from: (78))

More recently, the paradigm has shifted toward a more distinct dual pathway of progression (82). DCIS is divided into three grades of disease dependent on morphological features, hormonal receptors and genetic characteristics: low, intermediate and high-grade (83). Genetic and histopathological advancements suggest that a linear model is too simplistic and that there are in fact two pathways of in situ disease progression (84). According to this model, early genetic changes in the cell pre-determine its propensity for disease progression, i.e. low and high-grade DCIS are different entities with different disease pathways (77). By this theory, low-grade DCIS may remain dormant or progress slowly into low-grade invasive carcinoma. Conversely, high-grade DCIS has a more aggressive pathway of progression. A US based observational study by Narod et al. (82) of 108,196 women demonstrated that high-grade DCIS may have occult invasive cancer unidentified on pathological assessment, and is therefore likely to follow a more aggressive route of invasive progression.

Despite the identification of genetic markers (85), none of the proposed classification systems can reliably predict or rule out invasive progression. As such, it is impossible for clinicians to distinguish indolent DCIS from more aggressive disease at the time of diagnosis, and so all women are treated regardless of disease grade (86). Risk factors for DCIS and invasive breast cancer are similar (87-89), the most common being increasing age, family history and increased breast density. However, it is argued that not all DCIS will progress into an invasive breast cancer without treatment (3). Low-risk DCIS (defined as Grade I-II DCIS, <50mm and ER+/PR+/HER2-), thus provides an ideal proxy health state to explore the valuation of overdiagnosis resulting from breast cancer screening programmes.

2.3.3 Screening, investigation and diagnosis

DCIS is primarily a disease of screening (90). Nearly 90% of cases are diagnosed through microcalcification on imaging studies, primarily mammography screening, rather than through symptomatic presentation or clinical examination. Diagnosis is subsequently confirmed by core biopsy and histopathological review (91).

DCIS itself has an excellent prognosis, with breast cancer survival exceeding 96% after 20 years (25). Nonetheless, there is variation in treatment intensity and a suggestion that women with low-risk DCIS may often undergo more extensive treatment than is required. Overtreatment is a growing concern (92) and has implications for both patients (adverse treatment effects, out of pocket costs) and society (ethical issues, public health expenditure).

2.3.4 Treatment and prognosis

The current treatment for screen-detected, asymptomatic DCIS aims at preventing invasive breast cancer progression (28). Standard local therapy for DCIS is breast conserving surgery

(BCS), with or without radiotherapy, or a total mastectomy. Endocrine treatment to reduce the incidence of subsequent breast cancer may be considered (74). In the UK, 30% of women diagnosed with screen detected DCIS are treated by mastectomy and 70% by breast conserving surgery (27), comparable with mastectomy rates for invasive breast disease. Mastectomy is an excellent treatment in terms of reduced risks of recurrence or progression to invasive cancer, but may be regarded as extensive treatment for a non-life threatening condition (93).

There are many factors that influence treatment decisions including the size and grade of the abnormality, the age and health of the patient, possible side-effects of treatment and patient preference (71). For example, the role of radiotherapy requires a trade-off between the potential side-effects, costs and intensity of treatment versus the risk of developing a (treatable) local recurrence and its subsequent psychological sequelae in the patient (94). Similarly, the consequences of treatment can have a long-term negative impact upon quality of life. A cross-sectional study of 333 Australian women by Tian et al. (95) demonstrated that many breast cancer patients reported chronic morbidity, lower functional status and issues with body image after treatment that was unaffected by time following surgery.

However, the risk of disease recurrence is significantly reduced with more intensive treatment. A meta-analysis (96) of four randomised controlled trials (n = 3,729) demonstrated that radiotherapy after BCS for DCIS reduced the risk of recurrence by up to 50% (HR 0.46) but did not improve breast cancer-specific survival (97). Similarly, rates of ipsilateral invasive recurrence were lower in women treated with mastectomy compared with BCS alone (98), but again there was no significant difference in survival. Endocrine therapy is not routinely recommended for women with DCIS in the UK (74), although both tamoxifen and

anastrozole have been shown to reduce the 10-year risk of recurrence and new breast cancer events in those with hormone (ER/PR) positive disease (99, 100).

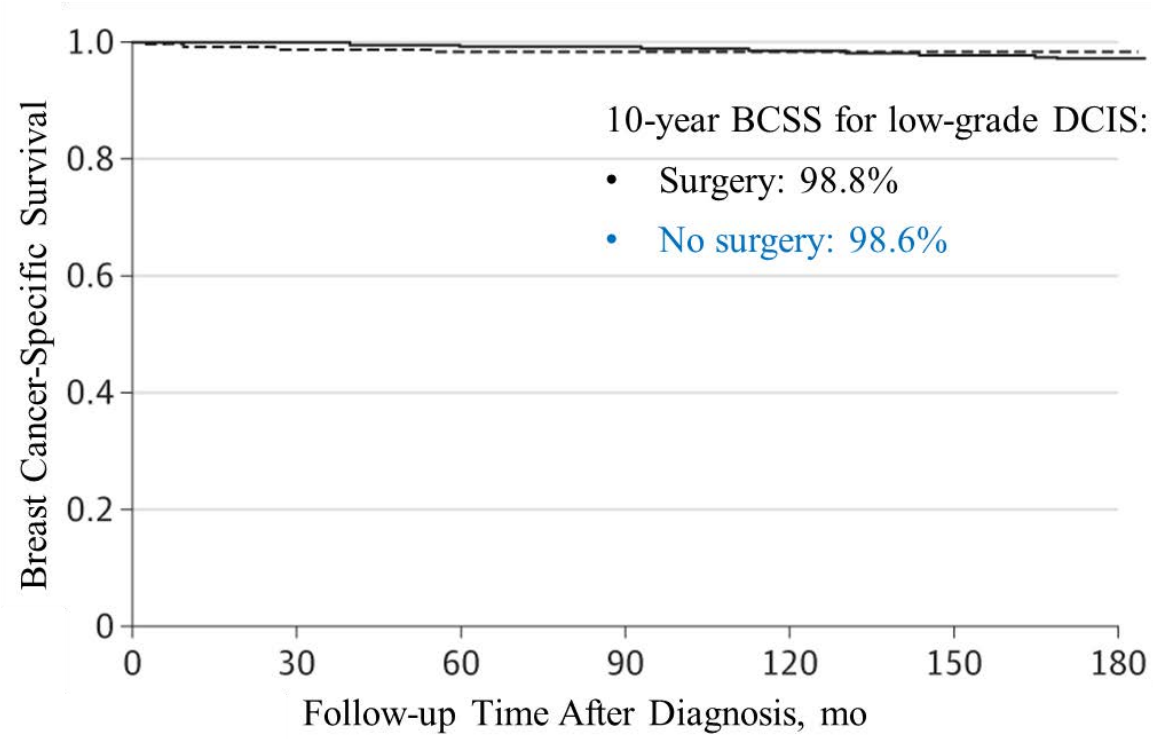
2.4 Controversies in DCIS treatment

The management of DCIS remains controversial. Treatment encompasses a multidisciplinary approach consuming considerable healthcare resources and requires long term follow-up (101). The rationale behind treatment is to prevent the development of invasive disease. Without treatment it is estimated that 10-30% of DCIS will progress to invasive breast cancer (102). Yet once diagnosed, 97% are treated according to current guidelines of surgery, radiotherapy and endocrine therapy, i.e. treatments comparable to those recommended for invasive breast cancer (29). Women are therefore not only exposed to the morbidity of treatment, but the risk of anxiety associated from having a "cancer" diagnosis (27), potentially leading to widespread overtreatment (103).

A retrospective cohort analysis (25) of 57,222 women with DCIS from the US showed no significant difference in survival between those with low-grade disease who underwent surgery and those who did not (Figure 2.5). Among those with low-grade DCIS, the weighted disease specific survival after 10 years was 98.8% and 98.6% for those who had surgery and no surgery respectively. Subsequently, many women undergo the morbidity of invasive treatment to reduce the potential risk of future recurrence without any benefit in survival.

Another study by Solin et al. (104) followed 665 patients with low, intermediate and high-grade DCIS managed via breast conserving surgery alone. At a median follow-up of 12.3 years, ipsilateral breast recurrence was 14.4% (7.5% invasive) for women with low-intermediate grade DCIS, but 24.6% in the high-grade group (13.4% invasive). The results

suggest that not all DCIS will become an invasive breast cancer in patients with favourable DCIS characteristics. Other retrospective reviews prior to screening (80, 81), whereby women with low grade DCIS were treated by biopsy alone, concluded that around half of the local recurrences that occurred were in situ disease with no threat to life.



Legend: BCSS: Breast cancer specific survival; mo: months

Figure 2.5: Kaplan-Meier curves for breast cancer specific survival in low-grade DCIS (adapted from:(25))

Whilst there is ongoing debate about recurrence, there is growing belief that many of the recurrences reported for DCIS may in fact be new primaries or concurrent high-risk disease not previously identified on biopsy rather than true progression (105). Invasive breast cancer has been found in 8-43% of resection specimens from patients primarily diagnosed as having DCIS (28). A US observational study of 108,196 women who received a diagnosis of DCIS

from 1988 to 2011 showed that at 20 years, the breast cancer specific mortality rate was 3.3%, comparable with general population mortality (40). Those who developed metastatic breast cancer were unlikely to have had low-grade disease in isolation and concurrent high-grade DCIS was associated with most of women who went on to die from breast cancer.

2.5 Active monitoring: a possible solution to overtreatment?

For women with low-grade DCIS, trials of observational treatment strategies, known as *active monitoring*, are underway that may serve as an alternative to immediate surgical management (27, 106, 107). Active monitoring is an expectant treatment option that delays and potentially avoids curative treatment unless and until there is evidence of high-grade or invasive disease progression (30). Patients have yearly mammograms to examine for potential cancer progression, and a core biopsy only if there is any suggestion of new abnormalities. Monitoring has the benefit of allowing women to delay and possibly avoid the side-effects of potentially unnecessary treatment. But the risk of invasive progression for DCIS managed non-surgically is uncertain and may induce additional anxiety or more intensive treatment if invasive disease ensues (108). Consequently, such strategies are only available under observation in clinical trials at present, the first of which is the Low-Risk DCIS (LORIS) trial at the University of Birmingham.

2.6 The Low-Risk DCIS (LORIS) trial

The LORIS trial was established to address the issue of overdiagnosis (30). It is a multi-centre, phase III, randomised controlled trial of patients with low-risk DCIS (defined as low or low-intermediate grade DCIS with no comedo-necrosis on histopathology). Patients are randomised to standard surgical treatment and postoperative follow-up versus active monitoring (annual mammography follow-up). The primary outcome of the trial is to assess

whether active monitoring is non-inferior to surgery in terms of ipsilateral breast cancer free survival. A schematic of the trial is provided in Figure 2.6.

Randomised controlled trials are used to gather data and provide reliable information about the effects of an intervention. They are less prone to distorting effects or biases than observational studies and are therefore considered the gold standard practice for investigating new medical interventions. The aim of the LORIS trial is to determine whether women with newly diagnosed low-risk DCIS can safely avoid surgery, enabling women with screen-detected low-risk DCIS to make an informed decision regarding treatment and the potential prevention of unnecessary intervention following overdiagnosis.

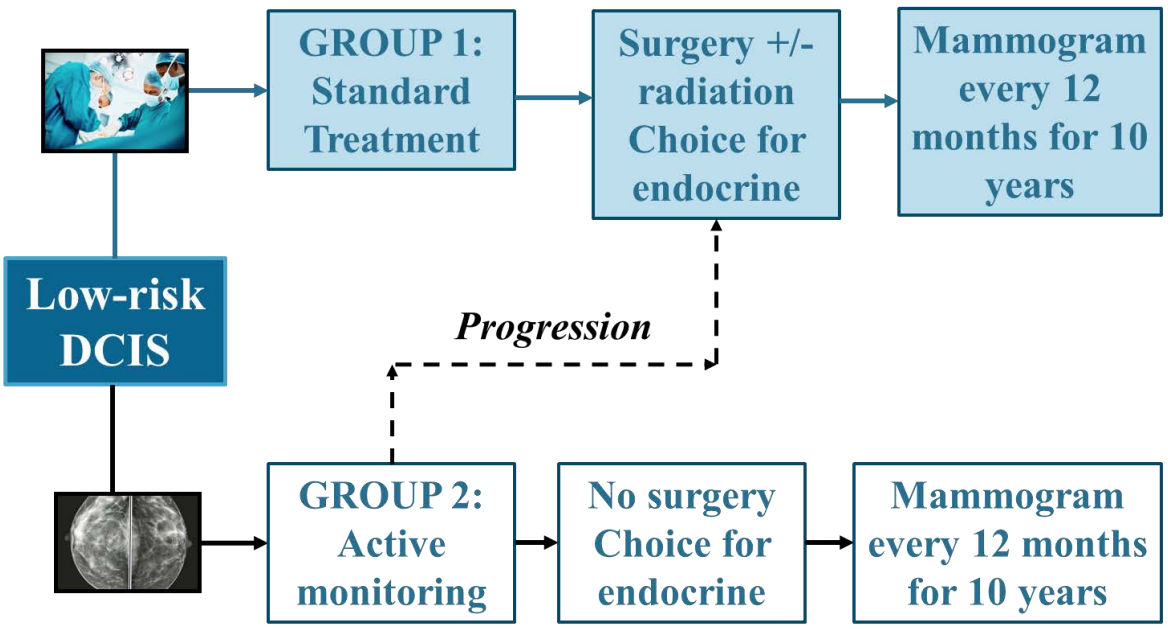


Figure 2.6: A schematic of the LORIS trial

Similar trials have been rolled out internationally, specifically the LORD (107), COMET (106) and LORETTA (109) trials in the Netherlands, USA and Japan. Trials of active

monitoring have been successful in other oncological screening programmes (prostate and cervical) (67, 110), and thus provide a suitable framework for modelling the impact of overdiagnosis (and overtreatment) for the purposes of this thesis.

2.6.1 Issues relating to costs and utility

Although the LORIS trial provides a strong foundation for exploring the disutility of overdiagnosis, the pre-trial research undertaken in this thesis is necessary to ascertain the potential impact of overdiagnosis among women who are not likely to be biased by participation in the trial. It is unlikely that the utilities captured in the surgical arm of the trial will truly capture the disutility of overdiagnosis, as women will have the knowledge of being randomised to a treatment and may therefore perceive it as having been necessary. Patients in clinical trials are more closely monitored than in standard clinical practice and so the data elicited may not reflect the true preferences of the general population. Consequently, the quality of life values reported in the LORIS trial may not be representative and generalisable to women in the community. Gaining an understanding of women's preferences beyond the trial and how they balance the associated benefits and risks will help decision makers to implement effective screening and treatment which maximises the benefits and minimises harms (111).

If active monitoring is found to be a safe and effective approach in the treatment of low-risk DCIS then it is likely that there will be important cost implications for health and social care services. For example, patients will avoid immediate surgery and radiotherapy and will instead incur the smaller costs of monitoring and outpatient follow-up. Resources may thus be saved by avoiding expensive inpatient care. However, active monitoring may incur undue costs from extra appointments for reassurance or counselling between mammograms. Only

one study to date has estimated the burden associated with the overtreatment of DCIS (73), by multiplying the average cost of DCIS treatment by the estimated rates of overdiagnosis in the US. Yet these estimates only offer an approximate calculation of the cost of overtreatment and must be balanced against the trade-offs of monitoring anxiety and invasive progression (112).

The difficulty of working within the breast screening remit is that the natural history of the disease is unknown. It is hard to predict which individuals are overdiagnosed at the point of diagnosis (90). Nonetheless, at the societal level there is an obligation for policy makers to ensure that both the benefits and harms of screening are adequately captured. Misplaced policy decisions about screening and treatment not only pose questions about inappropriate resource use and opportunity costs forgone, but also raise the ethical dilemma around inflicting an unnecessary cancer label and the impact of treatment on women who may not have needed it (31). This thesis aims to address these issues and explore the extent to which such harms are captured in economic evaluations, using DCIS as a proxy measure.

2.7 Conclusion

This chapter has summarised the issues surrounding the benefits and harms associated with ductal carcinoma in situ and breast cancer screening programmes. Although mammography has improved breast cancer survival and reduced treatment intensity, it has also contributed to a dramatic increase in the number of low-risk breast cancers potentially subject to overdiagnosis. Despite the mass of clinical evidence researching the benefits and harms of breast cancer screening, there are important gaps in the economic evidence informing the debate. The next chapter explores the theory of how these benefits and harms are captured in economic evaluations and the methodological challenges faced in valuing health states.

CHAPTER SUMMARY

What is known?

- Breast cancer is a major cause of population morbidity and mortality worldwide
- Breast cancer screening programmes were introduced to reduce the harms and economic burden of breast cancer
- The benefits and harms of breast cancer screening are widely contested; mammography screening reduces breast cancer mortality, but it also contributes to the overdiagnosis and overtreatment of low-risk disease.

What this chapter adds

- Low and low-intermediate risk DCIS is likely to account for a significant proportion of overdiagnosis arising from mammography screening.
- The LORIS trial comparing standard surgical management with active monitoring provides a useful vehicle in which to explore the potential implications of overdiagnosis and overtreatment in terms of quality of life.
- Further research is necessary to explore the values of women unbiased by a clinical trial (generalisability) and to model the potential costs and implications beyond the limited time horizon of the trial.

CHAPTER 3

Theoretical foundations and methods in the economic valuation of health

3.1 Introduction

The aim of this chapter is to provide an overview of the health economic theory and valuation methods applied in the subsequent chapters of the thesis. It begins with an explanation of the trade-offs implied in the allocation of health resources, alongside the theoretical concepts of normative economics and frameworks used in the analysis of healthcare. This is followed by a description of the different types of economic evaluation and their associated methods for measuring and valuing health benefits. The chapter concludes with a brief overview of how these values are applied in economic evaluation, and highlights some of the issues associated with valuing health benefits and harms that are particularly pertinent to the thesis.

3.2 Theoretical foundations in health economics

3.2.1 Economic evaluations in healthcare

Healthcare resources are limited (4). Decisions must be made to determine which health programme or treatment to fund. Yet pursuing one course of action will require forgoing health benefits elsewhere. Should breast screening programmes be extended to older women or high-risk groups in the population? Or should a new oncology drug be funded instead? Such trade-offs will affect the type of healthcare available for the recipients and incur opportunity costs to others. An evidence-based approach is therefore necessary to inform choices in public health systems (8). Economic evaluations provide such a framework to assess the relative benefits and costs of alternative options.

Drummond (4) defines economic evaluation as *"the comparative analysis of alternative courses of action in terms of both their costs and consequences,"* ((4): p.4). The objective is to generate information that will assist decision makers to determine the most efficient allocation of health services. Economic evaluations thus play a key role in the allocation of healthcare, through the systematic comparison of costs and benefits of competing demands.

3.2.2 Positive and normative economics

Economics is broadly divided into two streams: positive and normative economics. Positive economics is concerned with the objective relationship between economic variables and events (5). In contrast, normative economics applies a subjective approach that relies on value judgement to answer questions about how things *should* be done (113). The different approaches therefore distinguish between factors that determine the demand for healthcare, and the desirability of such relationships and whether they are equitable and efficient. Normative judgments provide the basis for economic evaluation by ordering different states of allocation according to benefit, enabling an assessment of the value of providing health resources for one group over another (114). However, the healthcare market does not resemble a normal market and preferences cannot be simply measured by assessing individuals' responses to changes in price (115).

In a normal market, the value of a good is determined directly by observing how individuals respond to changes in quantity, price and their willingness and ability to pay for it (116). Yet in healthcare the conditions that define a normal market are unmet. According to neoclassical economic assumptions individuals make rational decisions with the aim of maximising utility (117), but individuals are unlikely to know when they will become ill, what type of healthcare will be necessary nor what the likely benefit will be at the point of need.

Furthermore, the process of consuming health services may be extremely unpleasant (e.g. chemotherapy) but patients consume healthcare regardless for the expected benefit it may bring in the future. This decisional uncertainty and information asymmetry between clinicians and patients may create conflict in agency (118). Therefore, decision makers act as an advocate on behalf of patients to ensure efficient and objective policy decisions are made in the allocation of healthcare resources (6). Such value judgements require a normative approach, of which either a welfarist or extra-welfarist framework may be applied.

3.2.3 Welfarism

Welfare economics stems from the traditional theoretical framework for economic evaluation (119). It is defined by the assessment of outcomes in terms of the extent to which they contribute to maximise social wellbeing in terms of utility (6, 116). Utility in welfarism represents the strength of an individual's preference in terms of the "satisfaction" or "happiness" they gain from a good (9). The primary purpose of welfarism is therefore to provide a framework in which value judgements can be made (in terms of changes in utility) to assist decision makers on how best to allocate limited health assets (120).

There are four key principles in which welfarism aims to achieve economic efficiency in resource allocation (121):

1. Utility maximisation: individuals act rationally to maximise their own welfare.
2. Sovereignty: individuals are the best judges of their own utility.
3. Consequentialism: utility may only be derived as a consequence of the consumption of resources and not the process itself.

4. Welfarism: the 'goodness' of any given state can only be assessed via the utility obtained from individuals in that state.

As trade-offs in utility cannot be measured directly in welfarist theory, monetary values are used as a proxy for changes in utility in determining economic efficiency. The concept of value judgment is based on the Pareto principle, which states: *"social welfare unambiguously increases only if the welfare of any member of society increases and that of no one falls."* ((9): p.326). That is, if an increase of utility for one individual may be made without making another worse off, then the intervention should be implemented as society is better off (122).

The Pareto principle is not without criticism. Tsuchiya and Williams (123) argue that all policies require a trade-off that will make some individuals better and others worse off in the real world. Healthcare budgets are fixed and so there will always be gainers and losers in healthcare allocation decisions. Consequently, a compensatory adjustment was introduced to operationalise Pareto efficiency in the healthcare setting. Kaldor (124) and Hicks (125) suggested Pareto improvement could be achieved if the gainers (from a change in resource allocation) could adequately compensate the losers so that everyone would be better off. A health programme is thus deemed worthwhile if the sum of the values derived after compensation is greater than zero.

The practical application of Pareto improvement in welfarism is achieved through cost benefit analysis (CBA). CBA is a type of economic evaluation in which all costs and health outcomes (utility) are valued in monetary terms (116). Individuals are conventionally asked to value utility in terms of hypothetical willingness to pay for (or avoid) the intervention under examination (4), and the intervention is justified if there is an overall net gain to

society, i.e. net monetary benefit outweighs the costs of implementation (126). Welfarism has the advantage in that preferences for health may be considered alongside non health-based attributes, and thus may address the question of whether services are worthwhile for society across the wider public sector (6).

A key argument against welfarism is the public opinion that healthcare ought to be distributed fairly based on need and not strength of preference (9). Although welfarism has strong theoretical foundations in consumer choice theory, Sen (120) argues that the assessment of utility alone is unlikely to capture individual wellbeing and preferences for resource allocation in the health setting. For example, individuals differ in terms of their ability to derive utility from resources and may adapt to expectations from ill health or happiness in real life situations. A second criticism of welfarism is that there may be practical and ethical limitations of valuing health in monetary terms. (9) Some individuals may be willing to compensate for those in poor health which violates the central tenets of welfare theory (127). Third, as judgements on different states are based on individuals' strength of preference for a state measured in monetary terms, the perceived ability or willingness to pay of the individual may skew the allocation of resources toward those with higher income (128).

Criticism from decision bodies and the public around valuing health benefits in monetary terms led to development of an alternative method of appraising health interventions, namely extra-welfarism.

3.2.4 Extra-welfarism

Extra-welfarism seeks to maximise health (129). The extra-welfarism framework was developed after Sen (130) argued that a broader perspective was necessary to capture the

wider benefits that may be acquired from ensuing good health. Outcomes in extra-welfarism are thus not solely based on social preferences, as in welfarism, but rather community values are used to assess the impact of resource allocation where changes in health are considered as changes in utility (10).

The notion of extra-welfarism, by definition, implies that something "extra" has been added to the welfarism approach. Indeed, Brouwer et al. (9) identified four methods in which extra-welfarism may be distinguished from welfarism. First, extra-welfarism permits the measurement of non-utility outcomes (e.g. survival) so that health itself is captured in the valuation of health services or treatment. Second, the valuation of benefit may be provided from those not directly affected by the intervention but who may still be relevant to the consideration of public health funding, i.e. the general population (11). Third, different weights may be applied based on the socio-economic characteristics of those receiving the intervention to address concerns regarding equity. Finally, extra-welfarism allows for interpersonal comparison across multiple dimensions of well-being (e.g. pain, mobility, usual activities etc.) as well as utility, which may have an important impact on how people value health.

The extra-welfarism approach is operationalised through cost-effectiveness analysis (CEA). CEA considers the best method of achieving a given objective, where results are presented in terms of cost per natural unit of effect (6). For example, the analysis of a new treatment for breast cancer might be expressed in terms of cost per life year gained or cost per cancer prevented. The results of cost-effectiveness analysis can therefore be used to assess technical efficiency, especially whether a new intervention can minimise the costs associated with achieving a specific outcome or benefit, compared with the current standard of care (7).

However, the CEA is limited in determining allocative efficiency, because findings cannot be compared across different conditions, due to the disease specific nature of the outcome measured. Consequently, this led to the development of the cost-utility analysis (CUA). The CUA is a special sub-type of cost-effectiveness analysis, whereby outcomes are measured in a single metric instead of using natural units, namely the quality-adjusted-life-year (QALY) (6). The main advantage of the CUA is that it is thus able to overcome the issue of comparability through the application of a common unit across disease groups and interventions (14).

The extra-welfarism approach is widely implemented in public health systems and is recommended by several decision-making bodies such as the National Institute for Health and Care Excellence (NICE) (8). However, the valuation of interventions which may have benefits outside of health (e.g. capability) might be negatively impacted as extra-welfarist measures cannot capture non-health dimensions in the decision (10). Furthermore, it is not always clear how to allocate scarce resources using an extra-welfarism perspective (e.g. if a new intervention is more effictive but more costly).

3.3 Quality Adjusted Life Years

The QALY was developed to capture the impact of interventions on both quality and quantity of life, and acts as a common measure which can inform healthcare allocation decisions (12). QALYs are calculated by multiplying the length of time spent in a health state by a numerical weighting, called a utility, corresponding to the quality of life of the health state measured on a 0 to 1 scale, where a weight of 0 is equivalent to being dead and 1 in perfect health. The benefits of a health intervention are then measured as the *"difference over a period of time between expected QALYs with a particular procedure and without it (or with an alternative)"*

((129): p.52). For example, in the decision between hormone treatment for early breast cancer, suppose taking standard treatment (tamoxifen) requires a quality adjustment value of 0.8 and the alternative (anastrozole) is 0.9. A ten-year period of living with treatment is therefore equivalent to 8 QALYs on tamoxifen and 9 QALYs on anastrozole, yielding 1 QALY in favour of the latter (or 1 year in perfect health).

The use of the QALY as a measure of health benefit is recommended by NICE to enable a standardised approach for comparing economic evaluations across different areas of health (131). Indeed, QALYs are especially important in the assessment of oncological health services where quality of life may have a more meaningful impact than quantity of life (132). However, further to concerns about using health alone in the evaluative space, Dolan and Tsuchiya (133) argue that it is unlikely that the value of health states will remain constant over time in the assessment of benefit. The QALY has also faced similar issues of equity as observed in welfarism. Whilst its impartiality may be fair in principle (i.e. a gain in QALYs is held equal regardless of who benefits from them), it impedes moral claims based on need for healthcare and level of deprivation (134).

Nonetheless, QALYs are used widely in the valuation of breast cancer screening and treatment programmes (135, 136). Possible methods for eliciting utilities and QALYs are discussed next.

3.4 Utility

Utilities are cardinal values assigned to each health state on a scale anchored from 0 to 1 equivalent to being dead and in perfect health (4). Values less than 0 indicate a state of health deemed worse than being dead. Utilities are integral to the QALY and therefore also integral

in how benefits and harms are captured in the valuation of health. In this sense, 'utility' refers to the preferences individuals (or society) may have for any set of health outcomes in terms of health-related quality of life (HRQoL) (14).

The measurement of utilities involves first defining a set of health states and then valuing them (137). Valuation may be derived directly or indirectly. The three most common methods of valuing health states directly are the visual analogue scale, time trade-off and standard gamble (16). Other variants such as the person trade-off (PTO) are more relevant to social choice contexts than preference elicitation and were therefore unsuitable in view of the objectives of this thesis.

3.5 Direct Methods

3.5.1 Visual analogue scale

The visual analogue scale (VAS) is the simplest of the direct methods (138). It requires individuals to rate health states on a scale conventionally represented by a single thermometer type line, with clearly defined anchors representing dead and perfect health or worst and best imaginable health respectively (6). Individuals are asked to indicate where on the scale they consider the health states being valued should lie, such that the intervals between placements reflect the perceived differences in how good or bad the health states are perceived. An example of the VAS is shown in Figure 3.1.

Due to its simplicity, the VAS has a high completion rate and represents the most feasible of the health state valuation techniques (139). However, values are elicited in a choiceless context and thus do not require individuals to make trade-offs between preferences (140). Consequently, many have argued that the VAS does not technically elicit utilities per se

(141). The VAS is also subject to scaling biases, whereby individuals may be reluctant to place health states at the extreme ends of the scale (142).

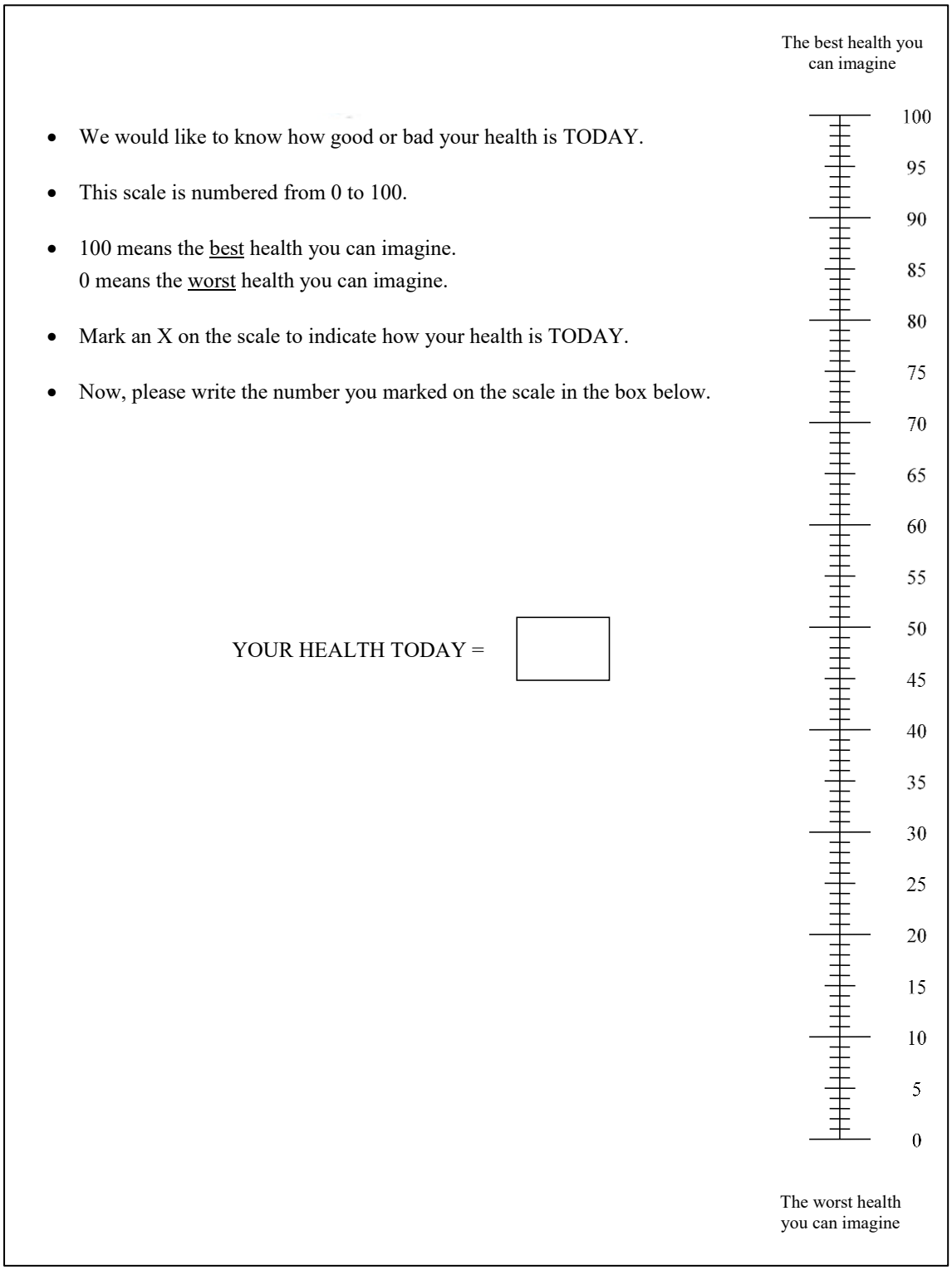


Figure 3.1: An example of a VAS (sourced from: (143))

3.5.2 Time trade-off

The time trade-off (TTO) asks individuals to consider the relative amount of time they are willing to sacrifice to avoid a certain worse health state (16). Individuals are offered two certain alternatives (Figure 3.2): Option 1 to live in health state H for time t followed by death, or Option 2, to live in perfect health for a shorter time (x<t) followed by death. Time x is varied until the individual is indifferent between the two alternatives, at which point the required utility for state H is given by: $U_H = x/t$. For health states deemed worse than dead, the trade-off is rearranged so that the first option is to die immediately, and the alternative is the amount of time spent in the worse health state followed by full health (6).

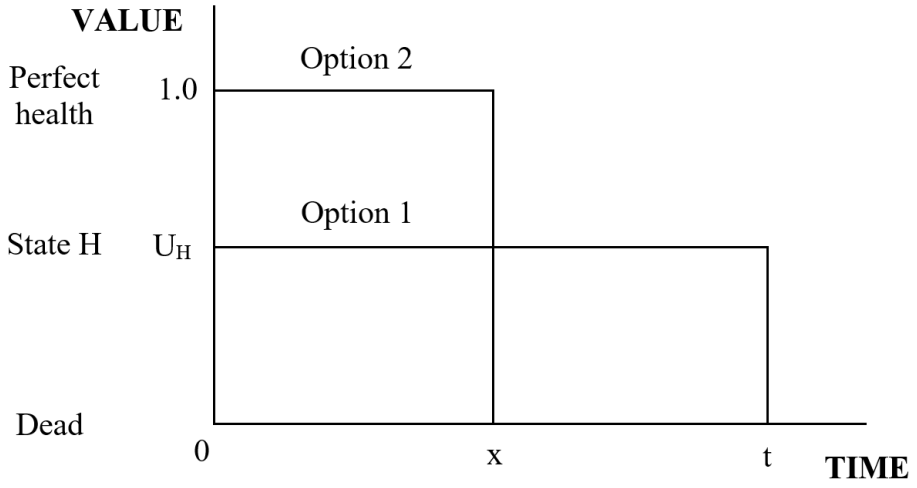


Figure 3.2: An illustration of the TTO for chronic health states better than being dead

The TTO was designed to overcome the issues of explaining probability and is viewed as a reliable and practical alternative to other direct choice-based methods (e.g. standard gamble). It has demonstrated good consistency in the literature (139), including health state valuation among women with invasive breast cancer (144). However, the applicability of the TTO in the context of medical decision-making has been criticised because the technique requires

individuals to choose between two certain outcomes, when health itself is characterised by uncertainty (145). The underlying assumption that individuals are prepared to trade-off life years to improve health may not be met either, if the condition being valued is less severe or amenable to time preferences (146). Indeed, some individuals may prefer to experience ill health immediately to eliminate anxiety, e.g. those with a positive family history of the BRCA gene may opt for prophylactic bilateral mastectomy to reduce the risk of developing breast cancer (147).

3.5.3 Standard gamble

Standard gamble (SG) uses notions of risk and uncertainty (as probability) to elicit utility (6). In contrast to the TTO, the individual is presented with a choice between the certainty of remaining in the health state (H) for t years (Figure 3.3), or the uncertainty of a hypothetical treatment with two possible outcomes, one of which is better and the other worse (e.g. perfect health and immediate death). The probability of perfect health (P) is varied until the individual is indifferent between the certainty and the gamble, corresponding to utility for the health state, where $U_H = P$. The more severe the health state, the greater the risk of being dead the individual is likely to accept (4). For health states worse than dead the gamble is adjusted so that the certainty of being immediately dead is traded against a gamble between full health and the worst health state, i.e. $U_H = -P/(1-P)$.

The standard gamble is advocated on grounds that almost all healthcare decisions involve conditions of uncertainty (140, 145). In developing Expected Utility Theory (EUT), von Neumann and Morgenstern (117) suggested that preferences are defined over a domain of lotteries, whereby if a cardinal utility could be expressed as equivalent to a gamble, the level of risk in the gamble is also linear to utility. Individuals will then choose between prospects

in such a way to maximise their expected utility. Consequently, the standard gamble is regarded by many as the 'gold standard' for health state valuation given its groundings in the axioms of expected utility theory. (16) SG utilities have been adopted in the valuation of breast cancer treatment (148) and in informing indirect health utility instruments (149).

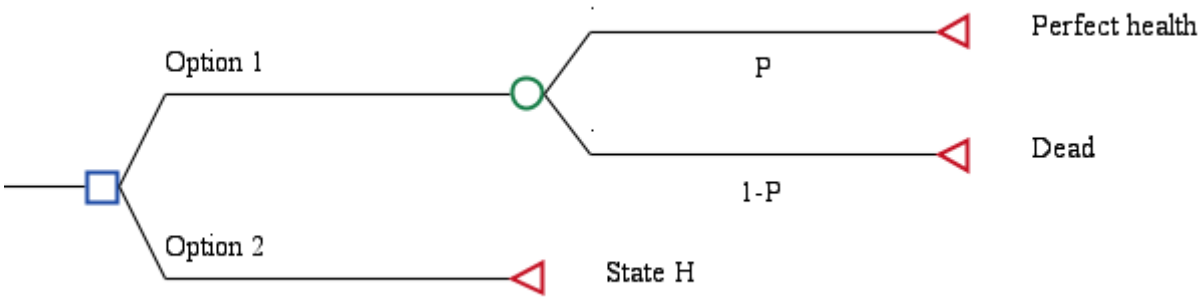


Figure 3.3: An illustration of the SG for a chronic health state better than being dead

However, Richardson (150) argues that the standard gamble only holds if the axioms of EUT are upheld, yet there is evidence that people may not always act rationally in decisions around healthcare. Understanding of probability is cognitively demanding (151), and the utilities elicited by the SG may be contaminated by risk attitude, gambling effects and loss aversion (150, 152).

3.6 Indirect methods

Utilities cannot be derived for every possible health state. Consequently, generic preference-based measures, known as multi-attribute utility instruments (MAUIs), are used to inform economic evaluation indirectly (6). MAUIs have two components; a classification system for describing health states and a scoring algorithm (tariffs or value set) assigned to alternative characteristics of health.

Multi-attribute utility instruments are increasingly being applied in the evidence informing healthcare decisions (153). In the UK, NICE recommends the use of generic descriptive systems, specifically the EQ-5D, to achieve a level of consistency when comparing different healthcare conditions and populations (8). Indirect methods are based on a generic classification system describing a broad range of health dimensions (e.g. physical functioning, pain etc.), where each dimension has a specified number of levels of severity (4). A health state is thus composed of one level in each dimension of health. Value sets are generated for each possible health state using an algorithm that assigns a score to each level in each dimension based on general population values (TTO, SG). Patients then complete the questionnaire and the appropriate tariff from the public scoring algorithm is applied (14).

A range of generic MAUIs are available for application in economic evaluation, each varying in the descriptive system and valuation method informing the design. The most widely reported MAUIs in the literature are the EQ-5D, SF-6D, HUI3 and AQoL (Table 3.1). This is by no means an extensive list (many more descriptive systems exist), but rather these instruments are briefly examined to highlight the strengths and limitations of indirect measures in valuing health benefits and harms.

3.6.1 EQ-5D

The most widely used MAUI is the EQ-5D (154). The questionnaire describes five dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression, with three (3L) or five (5L) level instruments in press, ranging from no problem to extreme problems or unable to (143, 155). An example of the EQ-5D-5L instrument is provided in Figure 3.4. Preferences for the scoring algorithm were developed

using the TTO on a random sample of the UK general population, but the instrument has since been extended worldwide to reflect other population scoring norms (156).

Table 3.1: A comparison of generic multi-attribute utility instruments

Instrument	Dimension	Recall period	Utilities	Scoring
EQ-5D (UK)	1. Mobility 2. Self-care 3. Usual activities 4. Pain/ Discomfort 5. Anxiety/Depression	Health today	TTO	Each item rated on a 3 or 5-point scale: **3L:** 1. No problems 2. Some/Moderate problems 3. Unable to/Extreme problems **5L** 1. No problems 2. Slight problems 3. Moderate problems 4. Severe problems 5. Unable
SF-6D (UK)	1. Physical functioning 2. Role limitations 3. Social functioning 4. Pain 5. Mental health 6. Vitality	Mapped from SF-36 (4 weeks)	SG	Each item rated on a 4-6 point scale
HUI3 (Canada)	1. Vision 2. Hearing 3. Speech 4. Ambulation 5. Dexterity 6. Emotion 7. Cognition 8. Pain	Current or usual health	VAS mapped onto SG	Each item rated on a 4-5 point scale
AQoL (Australia)	1. Independent living 2. Happiness 3. Mental health 4. Coping 5. Relationships 6. Self-worth 7. Pain 8. Senses	Past week	TTO	Each item rated on a 4-6 point scale

Under each heading, please tick the ONE box that best describes your health TODAY.

MOBILITY

I have no problems in walking about ❑
I have slight problems in walking about ❑
I have moderate problems in walking about ❑
I have severe problems in walking about ❑
I am unable to walk about ❑

SELF-CARE

I have no problems washing or dressing myself ❑
I have slight problems washing or dressing myself ❑
I have moderate problems washing or dressing myself ❑
I have severe problems washing or dressing myself ❑
I am unable to wash or dress myself ❑

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*

I have no problems doing my usual activities ❑
I have slight problems doing my usual activities ❑
I have moderate problems doing my usual activities ❑
I have severe problems doing my usual activities ❑
I am unable to do my usual activities ❑

PAIN / DISCOMFORT

I have no pain or discomfort ❑
I have slight pain or discomfort ❑
I have moderate pain or discomfort ❑
I have severe pain or discomfort ❑
I have extreme pain or discomfort ❑

ANXIETY / DEPRESSION

I am not anxious or depressed ❑
I am slightly anxious or depressed ❑
I am moderately anxious or depressed ❑
I am severely anxious or depressed ❑
I am extremely anxious or depressed ❑

Figure 3.4: EQ-5D-5L descriptive system

The EQ-5D is preferred for its *"comparability, transparency and consistency… in informing resource allocation in healthcare."* ((137): p.62). However, the EQ-5D may lack sensitivity in milder conditions or those where only minor changes in quality of life are observed from an intervention (157). Whist the 5L version was introduced to address concerns around this 'ceiling effect' (5L allows for 3125 possible health states versus 243 in 3L), it does not cover all relevant dimensions of health (e.g. hearing, vision) or non-health related benefit (e.g. well-being or capability) (158).

3.6.2 SF-6D

The SF-6D is a preference-based instrument developed from a non-preference based measure (Short Form-36) (159). It was developed in part due to the widespread use of the SF-36 in clinical trials, and subsequent demand to convert quality of life data into utilities for use in economic evaluation (159). The SF-6D consists of a descriptive system with six dimensions of health: physical functioning, role limitations, social functioning, pain, mental health, vitality, each with 4-6 levels of scoring severity (= 18,000 health states). The scoring algorithm was developed by obtaining weights using SG on a random sample of the UK general population (n =836) mapped against SF-36 data (160). The SF-6D has faced similar limitations to the EQ-5D relating to insensitivity to certain health domains (161), and the results are not always congruent with other MAUIs due to the different dimensions of quality of life captured by the descriptive system and valuation method (SG vs TTO).

3.6.3 HUI

Health utilities index (HUI) is a generic descriptive system used to measure general health status and HRQoL. It consists of two systems (HUI2 and HUI3) developed from the VAS and SG in the Canadian general population (149). The HUI3 is used to indirectly value utility in

adults and has eight health dimensions: vision, hearing, speech, ambulation, dexterity, emotion, cognition and pain, each with 5-6 levels of severity (= 972,000 possible health states). The HUI3 is sensitive to a wider range of conditions due to the broader dimensions of health covered but takes significantly longer to complete and has limited application in those with multimorbidity (162).

3.6.4 AQoL

The Assessment of Quality of Life (AQoL) instrument was developed in Melbourne, Australia to address the limitations of other descriptive systems to certain health dimensions (163). Four versions are available to date, ranging from four to eight dimensions of health. The AQoL-8D is the most comprehensive version covering eight health dimensions: independent living, happiness, mental health, coping, relationships, self-worth, pain and senses, each with 4-6 levels of severity (164). Scoring uses a multi-placative model for combining HRQoL dimensions, and tariffs were derived using the TTO from a sample of the Australian general population (165). Whilst the AQoL has demonstrated good convergence with other MAUIs, it is less widely used due to the limited generalisability of the Australian scoring tariff and longer completion time required.

3.7 Condition-specific measures

Condition-specific quality of life instruments (e.g. EORTC-QLQ-C30 in cancer patients) provide additional information about the impact of treatment in relevant dimensions to the condition under study (166). Where MAUIs are unresponsive to changes in utility or where data on utility is not included in clinical trials, condition-specific measures may be used to crosswalk results onto preference-based instruments, such as the EQ-5D, to calculate QALYs (167, 168). Mapping involves estimating the relationship between the condition-specific

measure and a generic multi-attribute utility instrument using statistical regression (4). However, in practice few condition-specific measures are anchored appropriately to be used to generate QALYs or do not use a choice based method which incorporates preferences (i.e. SG, TTO). The relationship between condition-specific and generic utility instruments has also demonstrated mixed results; Gerard et al. (169) found substantial differences between preference-based and condition-specific health state values for breast cancer screening health states using mapping.

3.8 Whose values?

The choice of whose values to elicit is as important as the method used to derive them. Much of the debate concerns whether the preferences of patients or the general public ought to be considered in informing healthcare decisions (6). Yet the distinction between the two groups is not dichotomous; the general population will inevitably contain patients and many healthy people will experience poor health during their lifetime (170). However, the decision is an important one and may change the decision on how healthcare is allocated; different groups report distinctly different utilities and cost-effectiveness results (171).

The main argument in favour of patient preferences rests on the premise that those receiving healthcare are best placed to value the impact of healthcare (172). Welfare economics implies that *"it is the preferences of the gainers and losers from a public programme that should be elicited and not [the] general population who will be unaffected by the change."* ((15): p.204). Thus, a normative approach suggests that resource decisions should reflect the preferences of those likely affected by the decision, and therefore patients are likely the best judges of their own well-being (173). Yet patients typically report higher values than the public for ill health (and hence fewer QALYs from smaller incremental gains in utility)

because of adaptation to the condition, differences in health state perception or response shifts in the expectation of good health (172).

It could be argued however that patients do not represent all those likely to be affected by a resource allocation decision. Decision makers advocate the use of general population preferences in a publicly funded healthcare system (8), as all individuals may be gainers and losers and thus have no vested interest or bias in the decision. However, the general population may have little or no experience of poor health and if inadequately informed may not give truly representative values of the condition being assessed (174). Members of the public may not want to be involved in healthcare decision making and may prefer that the values of those experiencing ill health be used (175).

A limited number of studies have applied a third perspective to address these concerns, whereby expert utilities are used as proxies in the health decision setting (148, 176). Health professionals are likely to be the most knowledgeable about ill health and therefore make good surrogates for public and patient preferences (177). Nonetheless, there is potential for clinician bias toward a certain intervention or a tendency to focus on attributes unimportant to those receiving the intervention (6).

Whilst there is no consensus on whose values should be used to inform resource allocation, there is at least agreement that patient representatives and experts should be involved in assisting social decision-making alongside the general population (15).

3.9 Application of health outcome measures

The focus of this chapter so far has been on the economic approach and the population used to value utility. Yet the primary purpose of collecting utilities is to use them in an economic evaluation to inform resource allocation decisions (6). Such data can be used in randomised controlled trials to determine whether an intervention produces greater gains in quality of life or losses from adverse side-effects, or applied in decision analytic modelling to improve the current standard of clinical practice (178).

For a meaningful comparison of costs and consequences it is necessary to estimate the cost per additional unit effect associated with one intervention over another (e.g. new versus standard treatment) (4). Conventionally, costs and benefits associated with competing interventions are compared as an incremental cost-effectiveness ratio (ICER), i.e. the extra cost of an intervention for a given condition over the next best alternative divided by benefit gained (179).

For example, if a decision is to be made between options A and B, the ICER is calculated by dividing the difference in costs and effects of both options, as:

$$ICER = Cost_A - Cost_B / Effect_A - Effect_B = \Delta Cost / \Delta Effect$$

If one programme costs less and produces more benefits, it is unequivocally considered cost-effective (i.e. it is dominant). More commonly, new interventions generate QALY gains at an additional cost, but may still be considered cost-effective if society is willing to pay for it (8).

In economic evaluation, the ICER is usually represented using the cost-effectiveness plane, a graphical representation of the decision in terms of incremental costs and benefits (180). The plane is divided into four quadrants indicating four possible conclusions following the comparison of the interventions (Figure 3.5). The x-axis represents incremental effects of the new intervention in relation to comparator and the y-axis the incremental costs.

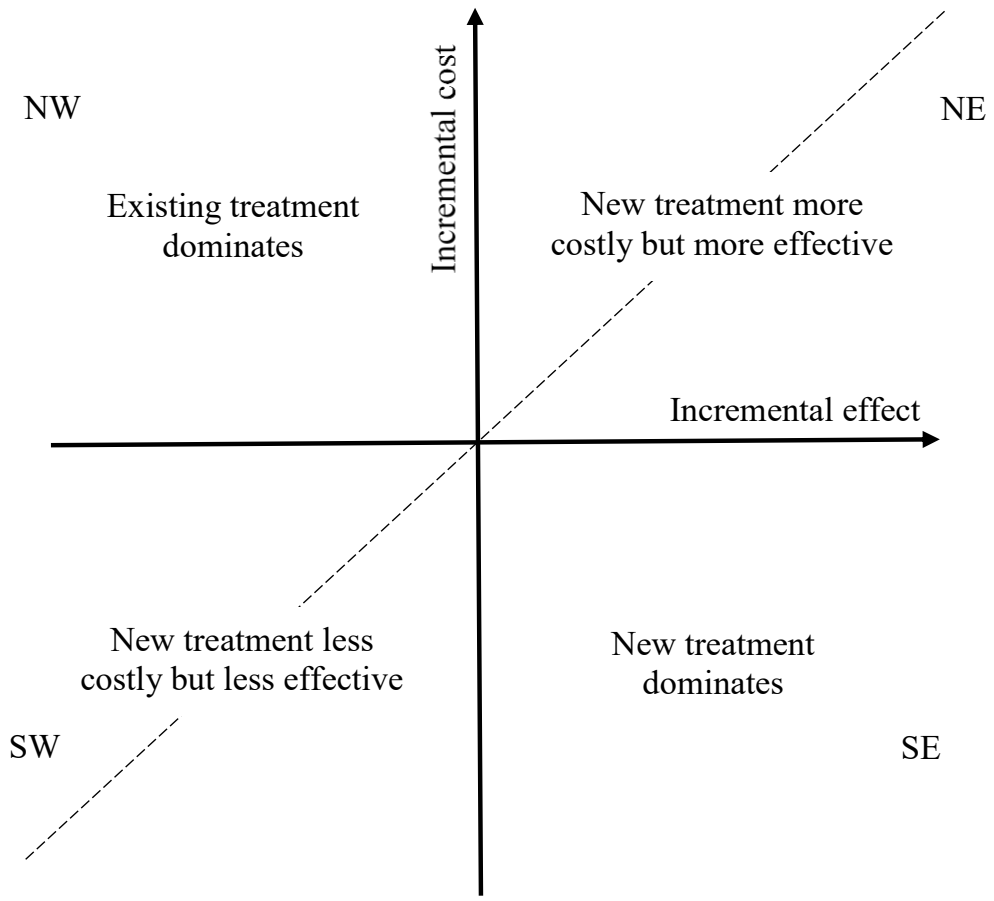


Figure 3.5: Cost-effectiveness plane

There are four possible outcomes from an economic evaluation (180). First, the new intervention may be less costly and more effective than the comparator (south-east quadrant) and is said to be *dominant*, i.e. the intervention should be implemented. Second, the new intervention may be less effective and more costly (north-west quadrant) in which case it is

dominated, i.e. the intervention should be rejected. Third, the intervention may yield more benefits at a higher cost (north-east quadrant) or fourth, the intervention may lower expenditure at the detriment of fewer benefits (south-west quadrant), whereby either decision implies a trade-off between the costs and benefits accrued.

Ultimately, the decision on what policy makers are willing to pay for an additional gain in health is dependent on the societal willingness to pay threshold (dashed line). In the UK, this is stipulated by NICE at £20,000-30,000 per QALY gained after considering the opportunity costs forgone (8). Consequently, the way in which health benefits are valued and appraised in economic evaluation can have a dramatic impact on whether new programmes or services are implemented in practice.

3.10 Conclusion

This chapter has provided an overview of the approaches used to measure health outcomes in economic evaluation. The purpose was to set the scene of health state utility valuation and the theoretical foundations central to the methods applied in the thesis. The main techniques for valuing health states and utility were summarised in detail, and the potential challenges and issues underpinning each method were highlighted. The next chapter explores how these methods have been applied in the economic evaluation of breast cancer screening programmes to value benefit and harm.

CHAPTER SUMMARY

What is known?

- Health economics adopts a normative economics approach due to the failure of health decisions to meet the normal market.
- There exist two frameworks for conducting economic evaluation; welfarism and extra-welfarism.
- Extra-welfarism assesses outcomes in terms of health gains, with the most commonly adopted outcome being the cost per QALY.
- Utilities are cardinal measures of quality of life measured on a scale from 0 to 1, equivalent to being dead and being in perfect health.

What this chapter adds

- Utilities may be derived directly or indirectly.
- The commonest approaches are the VAS, TTO and standard gamble (direct) and EQ-5D, SF-6D, AQoL and HUI (indirect).
- There is no consensus on who is best placed to value health; experts, patients or the general population. All have been used to value breast cancer health states.
- For a meaningful decision on how best to allocate healthcare resources, incremental costs and effects (QALYs) must be compared between two or more interventions.

CHAPTER 4

A systematic review of economic measures used to value health states in breast cancer screening programmes

4.1 Introduction

This chapter presents a systematic review of economic measures used to value the health states associated with breast cancer screening programmes. The aim of the review was to systematically assess if the utilities associated with the sequelae of mammography screening were adequately captured in economic evaluations. In doing so, the methodological challenges associated with health state valuation and overdiagnosis are discussed. The chapter concludes with an appraisal of whether the benefits and risks are adequately reflected in the utilities informing policy, including a summary of appropriate economic methods and recommendations for further research. The main findings from the systematic review are also published in *Social Science and Medicine* (Appendix 1).

4.2 Background

4.2.1 Economic evaluation and breast cancer screening

Evidence regarding the cost-effectiveness of healthcare technology is increasingly required to inform the decision on whether to fund and implement new treatment (4). Many decision-making bodies require interventions to be assessed using cost per quality-adjusted-life-years (QALYs) (8), a single summary measure combining life expectancy with utility (12, 181). As was discussed in Chapter 3, health state utility values (HSUVs) are cardinal measures of preference rated on a utility scale anchored from dead (0) to perfect health (1). Utilities can be valued directly or indirectly using expert, patient or public preferences. It is thus important that the methods used to identify, select and appraise utilities are transparent and systematic to reduce model bias and the potential misallocation of resources (182).

Economic evaluations impact health policy decisions and so the methodological quality of the parameters used to inform such analyses must be robust (183). Whilst a growing wealth of literature has explored the importance of economic analysis, less attention has been given to the methods and quality of the evidence used to inform utilities and QALYs (184). Several criteria are important for the selection of relevant utilities (185). The first relates to the health states, methods, descriptive system and population used to elicit the utilities. Where utilities have been measured directly, the validity, reliability and feasibility of the generated values should also be explicitly considered (169). Second, the duration of impact applied must be measured appropriately for both temporary and chronic health states associated with the intervention (58). The third relates to the generalisability of the condition, severity and population characteristics in the utility study to those in the economic evaluation using them (8).

The quality of the utilities applied is particularly pertinent in the appraisal of oncological interventions, where quality of life may have greater influence on QALYs than the modest gains in life expectancy (136). Many studies have evaluated the cost-effectiveness of breast cancer screening (186, 187), with those including quality of life in their evaluation reporting fewer net benefits, yet few have commented on the quality of the utility estimates used to inform them. When deciding on preferred screening policy it is critical to be able to accurately value the options available to women of being able to attend routine screening (3). This means valuing all associated benefits and risks associated with the alternative screening policy in terms of utility (22). Whilst clinical outcomes relating to breast cancer screening have been reviewed, an appraisal of how the benefits and risks are reflected in the utilities informing breast screening policy has not been well researched.

Screening for breast cancer in women aged 50 to 74 is recommended because of the ability to capture disease earlier and reduce treatment intensity and disease mortality (33, 188). Decision makers must value the risk that screening would lead to a woman having necessary (and perhaps less intense) treatment at an earlier stage than she would have otherwise had, against the risk of the woman having an unnecessary diagnosis and treatment (45). This valuation is made even more challenging because there is limited evidence on the rate of progression for many breast tumour types (27). If policy makers are to interpret cost-effectiveness analyses of mammography screening and balance the benefits and harms of such interventions appropriately, the utilities used in such evaluations must reflect the health states and impact on those experiencing the sequelae, including overdiagnosis and overtreatment (3). It is unclear how such utilities have been captured and conceptualised within the economic literature.

4.2.2 The challenges associated with valuing health states for mammography screening

There are several specific challenges relating to the identification and valuation of utilities for use in the economic evaluation of breast cancer screening programmes. A summary of the major issues is illustrated in Figure 4.1.

First, the natural history of breast cancer is poorly understood (24, 28). Not all valuation methods for deriving utility may account for the uncertainty in disease progression in the valuation process (189).

Second, overdiagnosis and overtreatment from screening create a “paradoxical popularity” because individual women may value unnecessary treatment inappropriately if screening and intervention for benign disease is misconstrued as life-saving (31). Qualitative evidence

suggests that both population and patient understanding of overdiagnosis is poor (65), with most perceiving the prognosis of pre-cancerous disease equal to that of an invasive breast cancer (190).

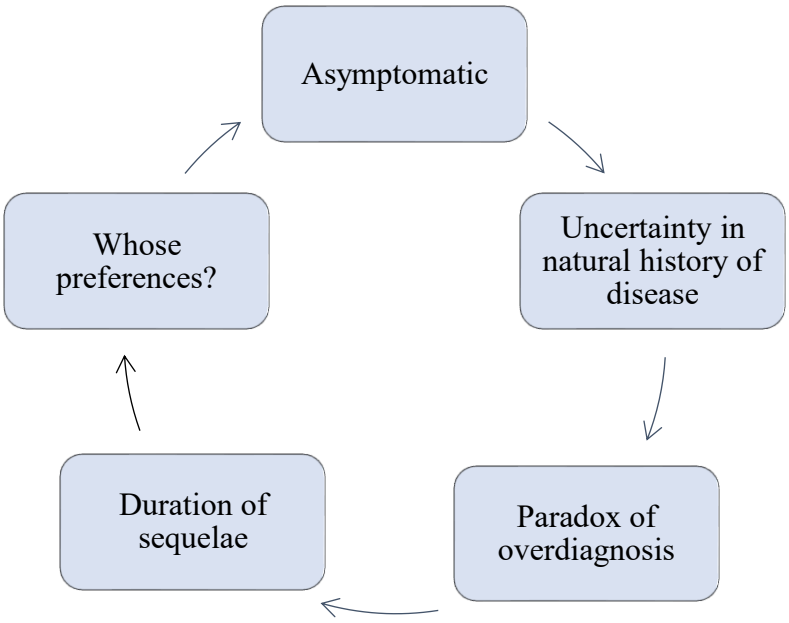


Figure 4.1: Overview of the challenges of health state valuation in breast cancer screening

Third, the sequelae associated with breast screening last for different durations (176). The long-term implications of a mastectomy are permanent (191), but the anxiety or reassurance associated with mammography screening or diagnostic investigation may only be temporary (192). Temporary health states require modification of conventional valuation methodology and economic evaluations must consider how both temporary and chronic health states are valued simultaneously within a single model (193).

Fourth, it is unclear whose preferences would be best placed to assess the benefits and harms of breast screening (194). The National Institute for Health and Care Excellence advocate the use of general population preferences in a publicly funded healthcare system, (8) yet given the complexity involved in valuing screening it may be difficult for the lay individual to

quantify using conventional utility instruments. The preferences and disease characteristics of individual women and breast cancers also vary significantly by demographic (195) and so the generalisability of the population in the primary and economic studies may influence the generated QALYs (196). Such challenges may impact utility instruments, therefore an assessment of the methodology used to overcome these issues is critical in the appraisal of appropriate utilities.

It is important to determine whether the utilities applied in economic analyses have addressed the issues highlighted and are thus able to truly capture the benefits and harms associated with the sequelae of mammography.

4.2.3 Rationale for the review

Systemic reviews are a widely practiced methodology for gaining a thorough understanding of the literature and are generally top of the hierarchical evidence for informing research and policy (8). They apply robust, replicable and extensive search strategies to produce a summary of the evidence relevant to a specified question of research interest. The use of systematic reviews and meta-analyses is well documented in the breast cancer literature. For example, the UK Independent Panel (3), Cochrane (20) and US Task Force (33) reviews applied a synthesis of clinical mortality data to inform the debate on the benefits and harms of mammography, but there have been fewer inquiries in relation to economic outcomes.

Published systematic reviews of economic evaluations relevant to mammography (59, 186, 197) have focused on the modelling approach, costs, population or screening intervention rather than the measures of quality of life. Similarly, whilst the clinical outcomes relating to

breast cancer screening have been reviewed, an appraisal of how these benefits and risks are reflected in the economic measures informing policy has not yet been researched.

Peasgood et al. (135) conducted a meta-analysis of health state utility values in breast cancer but could only identify six studies valuing the utility of mammography screening. Although their review was comprehensive, the authors did not consider the methods used or whether the disutility of overdiagnosis had been included. It was therefore necessary to conduct a new systematic review to explore the quality of breast screening utilities and identify areas of uncertainty or missing data in the evidence informing cost-effectiveness.

4.3 Aim and objectives of the review

The aim of this review was to explore how health states and utilities have been conceptualised and valued in the economic evaluation of breast cancer screening programmes. Primary studies that had measured utilities for relevant health states were also evaluated to examine the quality of the evidence informing economic evaluations.

The objectives of this study were to:

- Critically appraise and assess how economic evaluations have captured the health states and utilities associated with mammography screening.
- Examine the quality of the utilities used in economic evaluation or derived in primary studies to determine whether the sequelae of breast cancer screening have been valued appropriately.

4.4 Methods

The review followed the UK Centre for Review and Dissemination (198) guidelines and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (199).

4.4.1 Scoping search

A scoping search was undertaken to identify the most efficient search strategy of potentially relevant studies whilst minimising the number of irrelevant hits. The initial scoping search was developed by adapting the terms applied in systematic reviews of breast screening programmes (20, 135) to include terms relevant to the valuation of utility. Initially, no limitations were placed on the scoping search criteria. However, this yielded an unmanageable number of hits (>12,000 studies). An iterative process was used to restrict the number of irrelevant articles by removing terms which reproduced the same results (e.g. "cost effectiveness analysis" and "cost utility analysis") and the addition of strict exclusion criteria such as English language and full text limits.

4.4.2 Final search strategy

The full search strategy was developed from the scoping criteria using the terms published in other systematic reviews of breast cancer screening programmes (3, 135) and Cochrane review guidelines (198). Both Medical Subject Heading and keyword searches were used relating to the term 'mammography', 'breast cancer', 'screening', 'overdiagnosis', 'economic evaluation' and 'utility', with truncation used where appropriate (see Appendix 2). There was no restriction placed on publication year to ensure all relevant studies to date were included in the review.

Eleven electronic databases were searched for studies published up to 1 September 2018: MEDLINE, EMBASE, PsycInfo, CINAHL, Econlit, Social Citation Index, Social Sciences Citation Index, Cochrane library, NHS Economic Evaluation Database, Database of Abstracts of Reviews and Effects, and Health Technology Assessment. The reference lists of relevant studies were also hand-searched to identify any further relevant studies for potential inclusion.

4.4.3 Eligibility criteria

The systematic review included studies published in English whereby utilities were either used or elicited for health states relating to mammography screening. Studies were included if they met the following criteria:

- ***Participants:*** Adult women in the general population at risk of breast cancer
- ***Intervention:*** Population mammography screening (for economic evaluations) or breast cancer screening and its sequelae (primary studies)
- ***Comparator:*** No screening
- ***Outcomes:*** Cost per QALY (for economic evaluations) and utility values (primary studies)

Studies were excluded if they were:

- Reviews, editorials or abstracts
- Interventions to improve screening participation
- Evaluations of breast screening programmes using technologies other than mammography (e.g. magnetic resonance imaging) that are not routinely used to screen the general population

4.4.4 Study selection

A two-stage process (200) was used to identify relevant studies for inclusion in the final review. In the first stage, the title and abstract of retrieved studies were checked against the specified eligibility criteria. Relevant studies or those where a decision could not be made based on the title and abstract proceeded to the second stage. In stage two, the full text was assessed for relevance and the reference lists of key articles were hand-searched to identify other potentially relevant studies. Studies citing or reporting utilities which met the required eligibility criteria were included in the final review. A second reviewer screened and checked a sub-sample of studies (10%) to negate any selection bias (201). Information retrieved by the database search was managed via Endnote referencing software version X7 (202).

4.4.5 Data extraction and analysis

An electronic template was used to extract data on the characteristics of the included studies. For each included study, one reviewer extracted data about the study characteristics, the health states and utilities reported and the methodology, population, instruments and duration for which utilities were applied. The data were tabulated and analysed by narrative description as the retrieved utilities were too heterogeneous to usefully combine in a meta-analysis.

4.4.6 Quality assessment

A formal quality appraisal was not performed as there is no agreed quality assessment checklist for assessing studies of utility (183). However, the methods suggested by Brazier (194) and Papaioannou (183) for the systematic identification, selection and assessment of utilities from the literature were used to assess the validity, reliability and robustness of the identified utilities to inform the narrative review (189, 203).

4.5 Results

4.5.1 Main findings from the review

The database search retrieved 9,447 studies, of which 3,562 were removed as duplicates. A further 3 studies were identified through hand-searching the reference lists of relevant studies. A flow diagram of the studies selected or excluded at each stage, with reasons, is provided in Figure 4.2. Data were extracted for the 50 relevant studies included in the narrative review: 40 economic evaluations using cost per QALYs and 10 primary studies that measured utilities for health states associated with mammography screening.

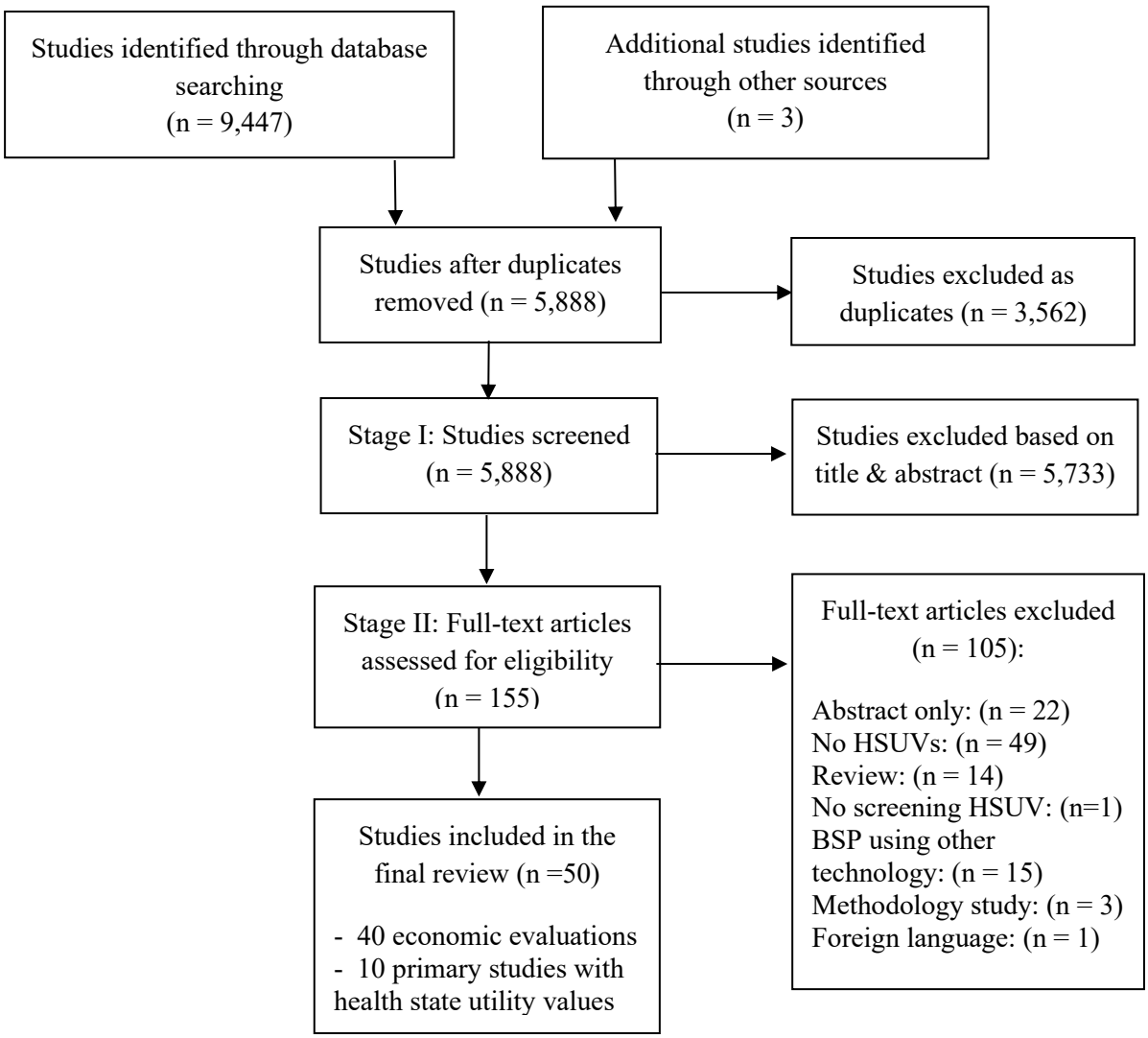


Figure 4.2: A PRISMA flow chart of studies included and excluded at each stage

4.5.2 Economic evaluations using cost per QALY

Table 4.1 summarises the characteristics of the 40 economic evaluations using QALYs in their analysis of breast screening programmes. Most evaluated alternative breast screening strategies (19, 52, 204-223), although five studies (224-228) only presented results using cost per QALY in the sensitivity analysis because of the uncertainty around published utilities for mammography. Seven studies (55, 60, 224, 226, 228-230) explored the cost-effectiveness of screening elderly women, whereas two studies (231, 232) evaluated extending the lower age limit of screening. Four studies (26, 227, 233, 234) assessed the benefits of risk stratified mammography screening and one study (235) appraised opportunistic versus organised mammography screening. The remaining three studies (31, 57, 236) evaluated the benefits and harms of breast cancer screening but reported QALYs without costs in their main analysis.

The utilities associated with breast cancer screening are difficult to compare because each study made different assumptions about the value used, the duration over which they were applied and the sequelae included in each of the health states. The values used for screening varied significantly (0.100-0.994) and were applied for a duration of between 2 hours and 7 days. There were similar issues with heterogeneity between utilities for diagnosis (0.100-0.895) and treatment (0.100-0.990). Utilities were applied for between 5 days and 6 months for a positive mammogram and a duration of one month to the rest of the woman's life for treatment depending on classification by intervention or disease stage. The duration of utilities or disutilities when applied in economic models can be a key driver in influencing results using QALYs (237), yet few

studies justified the duration enforced (31, 233, 234) or considered whether the utilities for temporary health states had used an appropriate chaining adaptation (238).

Figure 4.3 illustrates the range of utilities listed for the health states included in the review in more detail, where values were provided within the text. The spread of points within in each theme is reflective of the heterogeneity in the methods and populations used to derive them and therefore values could not be appropriately pooled into a single point estimate for each health state. Estimates ranged widely between individual processes in the pathway: 0.800-1.000 for screening attendance, 0.100-0.940 for screening results, 0.100-0.895 for diagnostic follow-up, 0.100-0.990 for treatment related health states (dependent on whether this was valued by in situ or invasive staging or intervention).

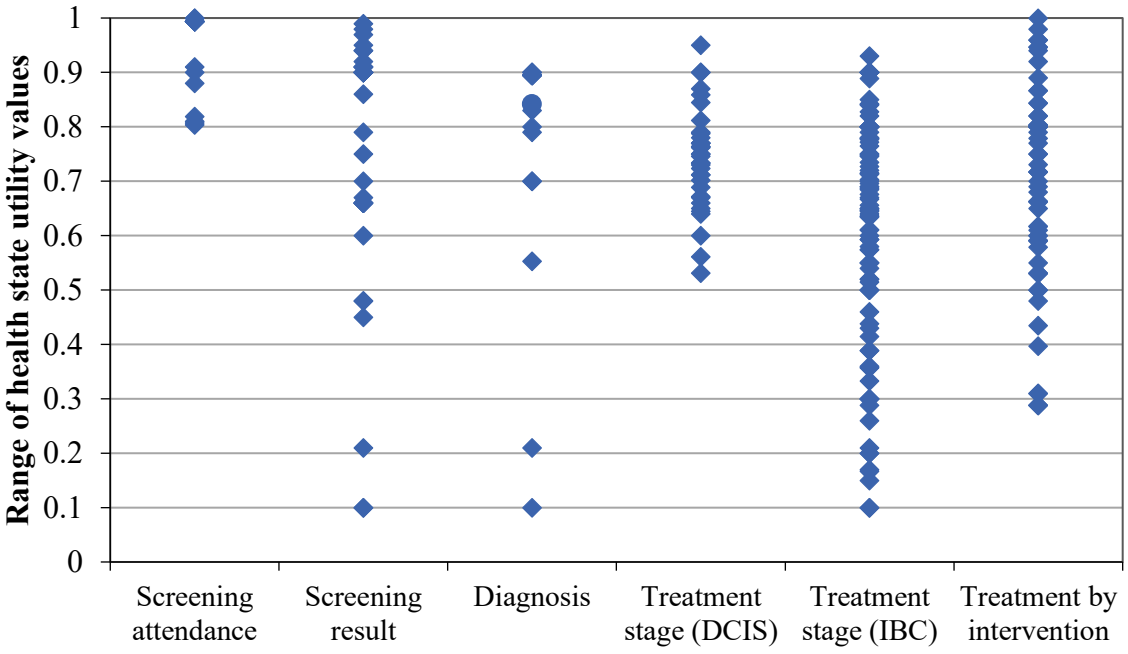


Figure 4.3: A spread of the utilities for each health state listed in the review

Most economic evaluations used the same two sources (176, 218) for their utilities, although there was variation in the actual value used and the generalisability of the population for which they were applied. The first of these sources by Stout et al. (218) applied tariffs based on assumptions for screening, diagnosis and treatment of breast cancer by stage at diagnosis to adjust US healthy population EQ-5D estimates (239). Although the generated utilities were deemed consistent with those reported in other studies (136, 240), it is not clear how the assumed adjustment for each health state was determined, and yet nine economic evaluations (31, 57, 205, 207, 219-221, 232, 236) applied this method. For screening disutility, almost half of the economic evaluations (31, 55, 206, 209, 211, 213, 215, 216, 219, 221, 222, 224, 229, 232, 236) used expert VAS utilities derived from a second study in the Netherlands (176), but only three economic evaluations (55, 209, 211) considered the generalisability of the expert sample to the general population in the model to which this was applied. Other economic evaluations made their own adjustments to local population EQ-5D or SF-6D data (26, 52, 205, 207, 208, 225, 234) or used utilities elicited directly (VAS, TTO, SG) from samples of women with comparable demographics to try and improve the relevance of the applied utilities to the population in their economic model (19, 55, 231). The remaining evaluations cited utilities from another economic model (210, 223), systematic review (204) or made their own assumption of an appropriate value (214, 226-228) but did not provide a detailed critique of how these were derived.

Sensitivity analyses were used to analyse the uncertainty around utilities in the majority of the 40 economic evaluations, with at least half reporting quality of life as having a significant effect on cost-effectiveness results. However, not all economic evaluations

included all relevant phases of the mammography screening pathway in their analysis and therefore implicitly assumed they had no impact on quality of life. 22 studies (19, 26, 31, 52, 60, 204, 205, 210, 212-215, 220, 223-225, 227, 228, 231, 233, 234) did not integrate the potential reassurance or disutility of screening anxiety and diagnostic follow-up in their analyses and a further 27 did not explicitly capture the disutility relating to the risk of overdiagnosis (19, 31, 52, 55, 60, 204, 206, 207, 209-215, 217, 218, 220, 223-227, 229-231). Consequently, no utility loss was applied to reflect this uncertainty in more than half of the economic evaluations, which may bias results (QALYs) toward more frequent screening (19, 31, 52, 207, 211). This limitation was justified in five studies due to the lack of robust utilities for mammography screening. For the 13 studies (26, 57, 205, 208, 216, 219, 221, 222, 232-236) which did attempt to value overdiagnosis in their analysis, an assumption was made that this was captured in the QALYs across screening strategies by including the temporary disutility of diagnosis and treatment without a corresponding gain in life years. However, the utilities applied used sources which had not highlighted that there was a risk the treatment was unnecessary during the valuation process and therefore is unlikely to fully capture the impact of the risk of overdiagnosis on quality of life.

Table 4.1: Characteristics of the economic evaluations (n = 40)

Lead Author	Study objective	Country; Population; Study type	Health states included	Method used by cited HSUVs	Information presented about HSUVs	HSUVs in study	Duration	Cited sources for HSUVs
Ahern (204)	Assess the cost-effectiveness of mammography screening and breast examination	USA; Women aged 40-79; 10 MM strategies (1-2y +/- CBE); MSM	Treatment (intervention)	VAS	Utilities from another model and systematic review (expert VAS transformed to SG using $SG=1-(1-VAS)^{2.29}$	0.590-1.000	6 months, 1 year, lifelong	(136, 241)
Arrospide (205)	Retrospective economic evaluation of Basque BSP	Spain; Women aged 50-69; 2y MM, MSM	Diagnosis Treatment (disease stage)	EQ-5D assumption	Applied assumptions from another model: tariff for the disutility of breast cancer applied to healthy population EQ-5D data (Spain).	0.338-0.824	1 year/ life expectancy	(218)
Barratt (55)	Assess the cost-effectiveness of extending BSP for women over 70	Australia; Women aged over 70, 2y MM, MSM	Screening Diagnosis Treatment (intervention)	VAS	Extrapolated the QALYs from another model which used expert VAS (systematic review).	0.288-0.994	Unclear	(176, 229)
Beemsterboer (206)	Economic evaluation of different screening strategies in Germany	Germany; Women aged 50-69, 2y MM, MSM	Screening Diagnosis Treatment (intervention)	VAS	Expert VAS utilities and durations (transformed to TTO)	0.288-0.994	1 week-lifelong	(176)
Boer (229)	Economic evaluation of extending the upper age limit of BSP	Netherlands; Women aged 50-69 and >70; 2y MM; MSM	Screening Diagnosis Treatment (intervention)	VAS	Expert VAS utilities and durations (transformed to TTO)	0.288-0.994	1 week-lifelong	(176)
Carles (207)	Economic evaluation of breast screening strategies in Catalonia	Spain; Women aged 50-79; 1-2y MM, Probabilistic model	Screening Diagnosis Treatment (disease stage)	EQ-5D assumption	Used the assumptions for duration and loss from healthy population EQ-5D in another model (US)	0.657-0.994	7 days-lifelong	(218)
Christensen (208)	Evaluate the cost-effectiveness of mammography screening in Greenland.	Greenland; Women aged 50-69; 2y MM; CEA	Screening Diagnosis Treatment (intervention)	Systematic review, assumption	Population QoL (Greenland) adjusted using values from a systematic review. Methods not reported.	0.480-0.810	6 months	(135)
De Gelder (235)	Economic evaluation of opportunistic and organised population mammography screening	Switzerland; Women aged 50-69; 2y MM, MSM	Screening Diagnosis Treatment (intervention)	VAS	Expert VAS utilities and durations (transformed to TTO) used in another model	0.288-0.994	1 week- lifelong	(176, 209)
De Koning (209)	Evaluate the cost-effectiveness of different BSP strategies	Netherlands; Women aged 40-75; 5 variants 1.3-3y MM, MSM	Screening Diagnosis Treatment (intervention)	VAS	HSUVs and durations based on 27 experts VAS (transformed to TTO)	0.289-0.994	1 week-lifelong	(176)
Forrest (19)	Cost-effectiveness of implementing a national BSP in the UK.	UK; Women aged 50-65; 3y MM; CUA	Treatment (intervention)	Rosser scale	Rosser ratio rating scale values used for the disutility of surgery	0.920	Lifelong	(242)
Haghighat (210)	Economic evaluation of mammography screening in Iran	Iran; Women aged 40-70, 3y MM, Markov model	Treatment (disease stage)	Assumption	Used the assumptions in another economic model of BSP	0.300-0.950	Unclear	(223)
Hakama (211)	Economic evaluation of Nordic breast screening strategies.	Nordic region; Women aged 50-69, CUA	Screening Diagnosis Treatment (intervention)	VAS	Expert VAS utilities and durations (transformed to TTO)	0.288-0.994	1 week-lifelong	(176)

IMS Health (212)	Economic evaluation of BSP in Australia	Australia; strategies for women aged 40-79, 2y MM, MSM	Treatment (disease stage)	VAS	Expert VAS ratings, authors adjusted weighting and duration using local treatment data	0.774-0.864	Unclear	(209)
Kerlikowske (224)	Economic evaluation of mammography screening in elderly women.	USA; Women aged 65-79; 2y MM, Markov model	Treatment (disease stage)	Assumption	Authors made assumptions of plausible estimates based on published HSUVs (TTO/VAS).	0.300-0.900	Lifelong	(176, 243)
Madan (225)	Cost-effectiveness of extending the lower age limit of BSP	UK; Women 47-49 years; 3y MM; MSM	Diagnosis	EQ-5D Assumption	Baseline UK healthy general population EQ-5D scores adjusted by assumption in sensitivity analysis	Not reported	Unclear	(244)
Mandelblatt (226)	Evaluate the cost-effectiveness of BSP in elderly women with and without comorbid disease	USA; Women aged 65-85 years; 1-2y MM; Decision model	Screening Diagnosis Treatment (disease stage)	Assumption	Assumption of plausible HSUVs based on expert VAS in the literature for similar health states	0.100-0.900	5 days, 30 days, life expectance	(176)
Mandelblatt (227)	Economic evaluation of targeted mammography screening in African American women	USA; African American women aged ≥40,; 1-2y MM; MSM	Treatment (disease stage)	Assumption	Assumption of HSUVs by disease stage, no description of how values were determined is provided.	0.500-1.000	Unclear	No cited source
Mandelblatt (228)	Evaluate the cost-effectiveness of a BSP in older women	USA; Women aged 50+; 2y MM; MSM	Treatment (disease stage)	Assumption	Assumption of HSUVs by disease stage, no description of how values were determined is provided.	0.550-0.950	1 year	No cited source
Mandelblatt (57)	Partial evaluation of mammography strategies considering screening and treatment advances	USA; Women aged 40-74; 1-2y MM; MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D Assumption, VAS	Expert VAS values (screening and diagnosis) and assumptions from another model (treatment) for US population EQ-5D tariffs	0.354-0.856	1 weeks, 5 weeks, 2 years	(176, 218)
Messecar (60)	Economic evaluation of BSP for older women with and without cognitive impairment	USA; Women aged 75-85y; 2y MM; Decision model	Treatment (disease stage)	TTO	Used general population TTO preferences for treatment	0.260-0.800	Lifelong	(245)
Mittmann (213)	Updated cost-effectiveness of BSP in Canada.	Canada; Women aged 40-74; 1-3y MM; MSM	Screening Diagnosis	EQ-5D Assumption, VAS	Expert VAS values (screening and diagnosis) and assumptions from another model (treatment) for US population EQ-5D tariffs	0.895-0.994	1 week, 5 weeks, 2 years	(176)
Morton (214)	Economic analysis of the BSP in the UK	UK; Women aged 50-70; 3y MM; CUA	Screening Treatment (intervention)	Assumption	Used QALYs from another economic model of the UK BSP	Not reported	Unclear	(31, 225)
Pashayan (26)	Cost-effectiveness or risk-stratified screening for breast cancer	UK; Women aged 50-69; 3y MM (risk); Lifetable	Treatment (intervention)	EQ-5D Assumption	Used adjusted population EQ-5D utilities from another economic model (systematic review)	Not reported	1 year, lifelong	(52, 135, 246)
Pataky (215)	Cost-effectiveness of population BSP by age & frequency	Canada; Women aged 40-74; 1-2y MM; MSM	Diagnosis, Treatment (disease stage)	VAS, SG	Systematic review (expert and population VAS to SG)	0.389-1.000	2 weeks-lifelong	(136, 176, 247, 248)
Pharoah (52)	Economic evaluation of the National Health Service BSP	UK; Women aged 50-70; 3y MM; Lifetable	Treatment (intervention)	EQ-5D Assumption	UK general population EQ-5D adjusted by a 0.9 relative reduction	Not reported	Lifelong	(135, 246)
Rafia (230)	Cost-effectiveness of extending the upper age limit of the UK BSP.	UK; Women aged 50-90; 3y MM; MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D Assumption modified by VAS, SG	Expert VAS, population SG and expert opinion used to adjust baseline UK population EQ-5D	0.360-0.910	2 hours, 3 weeks, 1-3 years, lifetime	(244, 247, 249)
Raftery (31)	Assess the benefit and harms of the UK BSP (partial evaluation)	UK; Women aged 50-70; 3y MM, Lifetable	Diagnosis Treatment (intervention)	Assumption	Systematic review, other models (expert VAS, population EQ-5D)	Not reported	0.2 years-Lifelong	(176, 218) (135)
Rojnik (216)	Economic evaluation of alternative breast screening strategies in Slovenia	Slovenia; Women aged 40-80 years; 1-3y MM; MSM	Screening Diagnosis Treatment (intervention)	VAS, SG	Expert VAS and SG utilities (oncology nurses), literature review	0.515-0.994	1 month-lifelong	(176)

Salzmann (231)	Cost-effectiveness of extending mammography screening to women aged 40 to 49 years.	USA; Women aged 40-49 years and 50-69 years; 1.5-2y MM, Markov model	Treatment (disease stage)	TTO	Australian patient TTO utilities in sensitivity analysis	0.300-0.800	Unclear	(243)
Sankatsing (232)	Cost-effectiveness of mammography screening before the age of 50.	Netherlands; Women aged 40-74 years; 2y MM; MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D, Assumption VAS	Expert VAS utilities (screening, diagnosis) Decrements in US healthy general population EQ-5D (treatment) from another model.	Unclear	1 week, lifelong	(176, 218)
Schousboe (233)	Cost-effectiveness of mammography screening by risk factors.	USA; Women aged 40-79; 1-2y MM, MSM	Treatment (disease stage)	EQ-5D	Swedish breast cancer patient EQ-5D applied to Swedish general female population EQ-5D	0.620-1.000	5 days- lifelong	(250)
Souza (217)	Economic evaluation of implementing a national BSP in Brazil	Brazil; Women aged 40-69; 1-2y MM; MSM	Diagnosis Treatment (disease stage and intervention)	SF-6D assumption	Author assumption of plausible estimate for false positive MM. HSUVs were estimated based on patient SF-6D scores (Brazil)	0.686-0.800	2 months-lifelong	(251, 252)
Stout (218)	Economic evaluation comparing alternative screening strategies.	USA; Women aged 40-80; 1-5y MM; DESM	Screening Diagnosis Treatment (disease stage)	EQ-5D assumption	Age-sex specific EQ-5D for healthy women (US) adjusted for negative effects of breast cancer diagnosis and treatment.	0.354-0.856	1 week- lifelong	(218)
Stout (219)	Assess the benefit, harms and costs of digital mammography screening	USA; Women aged 40-74; 1-2y MM; MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D assumption, VAS	Population EQ-5D (US) adjusted using assumptions from another model. Expert VAS utilities included in sensitivity analysis.	0.354-0.586	1 week-lifelong	(176, 218)
Tosteson (220)	Evaluate the cost-effectiveness of digital mammography screening	USA; Women aged ≥ 40;1y MM; MSM	Treatment (disease stage)	EQ-5D assumptions	Applied the duration and weighting assumptions from another economic model of BSP to healthy population EQ-5D data (USA)	0.430-0.860	Unclear	(218)
Trentham-Dietz (221)	Economic evaluation of tailored mammography screening for women over 50 years	USA; Women aged 50-74; 1-3y MM (risk); MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D assumption, VAS	Expert VAS (screening and diagnosis). US population EQ-5D (treatment) adjusted using assumptions from another model	Unclear	1 week-lifelong	(176, 218)
Van Luijit (2017) (80)	Economic evaluation of the Norwegian BSP	Norway; Women aged 50-69; 2y MM; MSM	Screening Diagnosis Treatment (disease stage)	VAS	Expert VAS utilities (transformed to TTO) from the literature	0.288-0.994	1 week-lifelong	(176)
Van Ravesteyn (236)	Assess the benefits and harms of mammography after age 74 years (partial evaluation)	USA; Women aged 50-94; 2y MM; MSM	Screening Diagnosis Treatment (disease stage)	EQ-5D assumption, VAS	Expert VAS for (screening and diagnosis). US population EQ-5D assumptions adopted from another economic model.	0.600-0.994	1 week, 5 weeks, 2 years, life expectancy	(176, 218)
Vilaprinyo (234)	Cost-effectiveness of risk-based breast screening strategies for breast cancer	Spain; Women aged 40-74; 1-5y MM, Probabilistic model	Diagnosis Treatment (disease stage)	EQ-5D	Patient EQ-5D (Sweden) for treatment extrapolated using the methods from another model.	0.655-0.859	2 months-5 years	(233, 250)
Wong (223)	Economic evaluation of biennial mammography screening in Hong Kong.	China; Women aged 40-79; 2y MM, Markov model	Treatment (disease stage)	Assumption	HSUVs from another economic model (but values do not match those cited)	0.300-0.950	Lifelong	(227)

Legend: BSP: breast screening programme, CBE: Clinical breast exam; EQ-5D: Euroqol-5D, HSUV: health state utility value, LYG: life years gained, MM: mammogram, MSM: microsimulation model, QALY: quality adjusted life year, QoL: quality of life, , TTO: Time trade off, SG: standard gamble, VAS: visual analogue scale

4.5.3 Primary empirical studies

Ten primary studies (58, 169, 176, 192, 243, 245, 247, 249, 253, 254) valued utilities for health states relevant to mammography screening. A summary of the study characteristics and methodology is shown in Table 4.2. The studies' aims were diverse and measured utilities for a range of relevant health states including: screening attendance and anxiety, mammography result (true positive, false positive, true negative and false negative), diagnostic investigation of a positive mammogram, treatment of a screen detected breast cancer, breast cancer recurrence and terminal care. The risk of overdiagnosis was not valued independently or explicitly captured within the descriptions of treatment health states in any of the primary studies of breast screening.

4.5.3.1 Methodology

The main method for valuing health states include direct and indirect empirical measurement or expert opinion (255). Multiple approaches were taken to elicit utilities for breast screening health states identified by this review, with more than four different valuation techniques reported within the primary studies. These four primary approaches included the VAS (256), which was anchored from worst to best imaginable health, standard gamble which compared the health state against a gamble of dead and perfect health (257), time trade-off which trades years lived in full health against living longer in the health state being valued (6) and EQ-5D which asked trial participants to report their own or hypothetical health on a generic scale and applied general population tariffs to estimate final utility scores (258). More than one technique was used to value screening health states in six studies (58, 169, 192, 249, 253, 254), whilst the remainder used a single technique (176, 243, 245, 247).

Table 4.2: Characteristics of the 10 primary studies which had elicited health state utility values (HSUVs)

Author; Country	Study aim	Participants	Health states valued	Utility range	Duration range	Technique	Further information methods
Bonomi USA (247)	Obtain QoL values for mammography screening and breast cancer treatment	131 women sampled from a population breast screening programme (aged 50-79)	Screening attendance Screening result (FP, TN) Diagnostic mammogram Treatment (intervention) Disease free at 1 year Recurrence at 1 year Terminal care	0.804 0.457-0.891 0.553 0.397-0.530 0.768 0.330 0.358	2 hours 2 weeks 2 weeks 4 months-5 years Lifelong 4 months 3 months	VAS	14 vignettes via in-person or telephone interview. VAS anchored dead-perfect health
Chie Taiwan (253)	Utility in different clinical phases of breast cancer.	21 clinical and public health experts	Screening attendance Diagnosis Initial treatment (intervention) Post-treatment (intervention) Recurrence at 1 year Terminal care	0.900-1.000 0.700-0.900 0.500-0.800 0.600-0.800 0.250-0.300 0.100-0.150	20 years for all	VAS, TTO SG	17 vignettes via face-to-face interview (visual aids). VAS anchored dead-perfect health.
De Haes Netherlands (176)	Elicit utilities for use in an economic model of BSP	27 clinical and public health experts	Screening attendance Diagnosis Initial treatment (intervention) Post-treatment (intervention) Disease free >1 year Terminal care	0.994 0.895 0.717-0.820 0.844-0.914 0.947-0.960 0.288	1 week 5 weeks 2 months-2 years 10 months Lifelong 1 month	VAS	15 vignettes via face-to-face interview. VAS anchored worst-best imaginable health. VAS scores transformed to TTO using the formula: TTO=1-(1-VAS) ^1.82
Gerard Australia (245)	Explore framing and labelling effects on breast cancer values.	180 women from the local general population (aged 45-69)	Treatment (intervention) of screen detected breast cancer with and without breast cancer death	0.150-0.750	10-30 years (age dependent)	TTO	9 different presentations of two breast cancer vignettes (varied cancer terminology and pronoun).
Gerard UK (169)	Determine the feasibility of mapping EQ-5D to TTO for validating breast cancer descriptions.	440 women from the general population eligible for breast screening (aged 40-64)	True negative False positive True negative False positive	0.910-0.940 0.210-0.790 0.480-0.660 0.450-0.660	12 months Lifelong	TTO (chain) EQ-5D	Two-stage chaining used to adjust temporary health states onto dead-full health scale. EQ-5D mapped onto TTO using 3/5 dimensions.
Hall Australia (243)	Derive utilities for use in an economic evaluation of BSP in Australia.	44 women from the general population and 60 breast cancer patients (aged 45-69)	Treatment (intervention) of a screen detected breast cancer	0.270-0.800	10-30 years (age dependent)	TTO	6 vignettes via face-to-face interview
Johnston; UK (58)	Derive QoL values for key breast screening outcomes	440 women from the general population eligible for breast screening (aged 40-64)	True negative False positive True negative False negative	0.91 0.66 0.66 0.66	12 months Lifelong	VAS TTO (chain)	Two-stage chaining method used to adjust temporary health states onto dead-full health scale
Kim Korea (254)	Determine the utility of breast cancer health states in Korean population.	509 general population men and women (aged >19)	Treatment (intervention) of screen detected non-invasive, invasive or advanced breast cancer, recurrence, terminal care	VAS: 0.170-0.681 SG: 0.352-0.804	Lifelong	VAS, SG	8 vignettes via face-to-face interview. VAS anchored worst-best health (readjusted to dead).
Rijnsburger Netherlands (249)	Assess the QoL of screening high-risk women for breast cancer.	334 women in a high-risk breast screening trial (mean age 40.9).	Screening attendance	VAS: 0.807-0.819 EQ-5D: 0.880	Unclear	VAS, EQ-5D, SF-36	Direct measurement at time points of 2 months prior, during and 1-4 weeks after attending screening.
Tosteson USA (192)	Measure the QoL impact of false-positive mammograms	1028 women in digital breast screening trial: 534 = negative, 494 = false positive	Screening (negative mammogram) Diagnosis (positive mammogram)	VAS: 0.830-0.860 EQ-5D: 0.900-0.910	Unclear	EQ-5D VAS	Direct measurement at baseline and up to 1 year after screening. VAS anchored worst-best imaginable health

The standard gamble, initially presented by Neumann and Morgenstern (259) and reported in Tolley (260), is the gold standard method for valuing conditions of uncertainty (261), yet only two studies (253, 254) used this technique to capture the potential benefit and risks associated with screening. An alternative choice-based method (TTO) was justified by five studies to reduce cognitive burden associated with the standard gamble (58, 169, 243, 245, 253). De Haes et al. (176) did not use a choice-based method but mapped visual analogue scale scores into utilities using a power function (VAS): TTO = 1-(1-VAS)^1.*82.* (262), although there are reported issues with the reliability of conversion formulas (167). A combination of both direct and indirect methods was used by the remaining studies (192, 249) using tariffs from the US (239) and Dutch general population (263) for the EQ-5D descriptive instrument before and after screening. The Short Form-36 questionnaire was also used by Rijnsburger (249), but the values were never mapped into SF-6D utilities (159). Only half of the studies considered whether the chosen method was appropriate for overcoming the methodological challenges associated with screening health states (58, 169, 176, 243, 245).

4.5.3.2 Duration

Traditional methods such as the standard gamble, TTO and VAS are targeted towards chronic health states (14, 16, 261). For valuing temporary health states, a two-stage technique known as ‘cascading’ or ‘chaining’ is recommended and can be applied to modify the traditional TTO or SG approach (264). For chaining, the worst temporary health state is known as the anchor health-state because it is used as the lower anchor instead of dead (193). The anchor state is subsequently valued against full health and dead to realign values onto the traditional utility scale (16). Only two studies (58, 169) used a chaining adaptation of the conventional TTO to appropriately value temporary health states for screening attendance and diagnostic investigation.

A combination of direct and indirect assumptions (265) were used to specify duration in the remaining studies. Four studies (243, 245, 253, 254) specified a single duration of impact for both temporary and chronic health states and applied the same method (TTO or SG) to ensure consistency. The same technique (VAS) was used in two studies (176, 247) to specify the time within the vignettes, although the durations applied varied depending on the timeframe assumed. Other studies (192, 249) did not specify the health state duration per se but indirectly measured utility at discrete time points during the screening process. However, due to variation in follow-up time some women were aware of their results a priori which may have inadvertently biased results.

4.5.3.3 Descriptive system

The validity of the health state and utility elicited is dependent on the accuracy of the vignette and should be informed by a thorough review of the literature or input from those well acquainted with the condition (167, 266). Utilities were generated using health state descriptions in eight (80%) of the primary studies (58, 169, 176, 243, 245, 247, 253, 254). Although the vignettes in all eight studies were informed by clinical guidelines and expert input, only five studies (58, 169, 176, 243, 247) validated the clinical scenarios through patient piloting or focus group discussion. Similarly, the framing and labelling of health descriptions can systematically bias choices and perceived quality of life due to the negative connotations associated with cancer and dying (267, 268), yet only two studies (243, 245) explicitly considered the impact of this on their results. The remaining two studies (192, 249) did not use vignettes but indirectly measured the disutility associated with screening by asking women enrolled in a clinical trial of tailored mammography to value their own health ex-ante and ex-post screening using validated health instruments (EQ-5D). Interestingly, both

studies commented on the limitations of the sensitivity of the EQ-5D domains in capturing changes in utility for the short-term duration of screening.

No primary study explicitly considered the impact of the risk of overdiagnosis or unnecessary treatment in any of the health states described. Only Gerard (245) and Hall (243) introduced the notion of dying of causes other than breast cancer in their vignettes, although they did not explicitly include the risk of unnecessary follow-up and treatment. Kim et al. (254) explicitly included risk in their health state descriptions of surgery and radiotherapy but only provided estimates for recurrence and survival, assuming all treatment was necessary for non-invasive disease.

4.5.3.4 Population

Health states relevant to breast screening can be valued by three populations groups; the general population, patients and clinical experts (194). Seven (58, 169, 192, 245, 247, 249, 254) of the ten primary studies used general population values, which are preferred by most publicly funded healthcare departments (8, 258), although there was some selection bias toward women of breast screening age. One study (243) collected a mixed sample of public and patient preferences and reported significant differences between the utilities measured by those with and without experience of breast cancer. Patient preferences are typically higher than those elicited from the public due to adaptation to the condition or a feeling of necessary intervention (194, 269), but Hall (243) justified their approach as they felt patients were best placed to value the complex side-effects associated with breast surgery. The remaining two studies (176, 253) used an expert sample to overcome the cognitive difficulties experienced in their feasibility piloting of TTO health states.

4.5.3.5 Quality assessment

Most studies did not explicitly comment on the quality of the reported utilities in terms of the validity, reliability and feasibility of the methods used. Among the four studies (58, 169, 176, 245) that reported on reliability, four assessed ranking order and only one (169) examined test-retest consistency. None of the primary studies commented on the time taken to complete the task, although this is routinely recommended for appraising participant comprehensibility (189, 257). At least half of the authors commented on comprehensibility issues relating to the SG and TTO techniques, although only one study (169) provided quantitative evidence to measure the reported cognitive burden using a Likert scale. Whilst most studies justified the VAS based on task acceptability, only three studies (58, 169, 176) considered the theoretical validity of this approach in capturing the temporary or uncertain benefits and risks associated with breast screening specifically.

4.6 Discussion

4.6.1 Main findings from the review

Population based mammography screening for breast cancer is a major public health investment and significant time investment for women and therefore warrants rigorous scrutiny (3). This systematic review provides the first synthesis of economic measures and health states used to value mammography screening, explicitly including overdiagnosis, and summarises the evidence base informing the population screening debate. The identified evaluations found that quality of life had a significant effect on cost-effectiveness results in sensitivity analyses. Determining whether the associated benefits and harms have been captured appropriately is therefore not only of clinical importance but may impact how screening policy is determined or overdiagnosis is conceptualised (20).

Deciding how breast screening utilities should be captured is fraught with challenges (45, 135). There is no consensus on the most appropriate economic measure and population to use when valuing outcomes in cancer screening programmes. Half of the identified studies in this review used the same two sources to value quality of life (176, 218), but the remainder used values that were based on assumption, used out of context or were not methodologically sound. Unlike prostate and cervical cancer, the natural history of in situ breast disease is not well understood (28), yet the way in which the utilities were assigned to represent the associated health states for screening and its sequelae were not described in detail in any of the studies. Balancing the availability and quality of published utilities to inform economic evaluations can be problematic where primary evidence is limited (140), but it is imperative that such limitations are made explicit so that decision makers may consider the implications upon cost-effectiveness results (270).

The heterogeneity in utility values raises the question of what economic measure should be used, or whether health related quality of life is suitable for measuring outcomes associated with screening and overdiagnosis. The commonest approach used in the empirical studies was the VAS, despite this technique being considered methodologically inferior to other choice-based techniques (138, 142). Ideally, the measure chosen should reflect the underlying decision within the valuation process, in line with traditional axioms of utility theory (117). When trading length of life against quality, TTO is more appropriate (238), whereas in a situation in which there is also risk (such as screening and treatment uncertainty), the standard gamble may be more suitable (257). A systematic review of metastatic breast cancer utilities (135) found that the SG was the most frequently used technique for capturing uncertainty in survival and preferred for valuing risk-based utility (269), although there are concerns it may inappropriately conflate health with risk aversion (124). Conversely,

preference-based instruments (EQ-5D) are considered the method of choice by NICE (271), but it is unclear whether an indirect approach would be sufficiently sensitive to detect minor changes in utility (192) or reflect the true risks involved, unless respondents are adequately informed about the benefits and harms during the valuation process or vignette. With most identified studies using utilities based on author assumption, new empirical evidence to reliably inform such analyses is clearly necessary.

The clinical outcomes associated with breast cancer screening programmes are widely contested, yet the benefits and harms of mammography are inadequately appraised in the economic literature informing the debate. Few studies identified by this review integrated all relevant phases of care associated with breast cancer screening into the assessment of quality of life, and the values used were limited in their ability to truly capture the disutility. Thirteen studies included overdiagnosis in their evaluation but applied the same utilities for diagnosis and treatment as a non-over diagnosed cancer (110, 272, 273), even if the costs and quality of life losses were ultimately not necessary or entirely representative. Estimates of screen detected overdiagnosis vary significantly from 0 to 54% (24, 33). Whilst several economic evaluations cite the lack of published utilities as a justification for not including screening or overdiagnosis in their analysis (19, 31, 207, 211), ignoring this harm may inadvertently lead to inappropriate advice to women, decisions on the value of screening programmes and potential misallocation of resources. Similarly, none of the primary studies explicitly considered the impact of unnecessary treatment in their vignettes.

The inclusion of overdiagnosis in qualitative descriptions has been shown to change general population preferences toward more conservative management or surveillance strategies (274, 275). The limitations of the economic measures and health states outlined in this review

(276) raises concerns about information asymmetry, and whether women can make an informed decision about screening without information on the full benefits and risks. Any potential advantages and risks should be explicitly listed within the descriptions of relevant health states. These findings are not limited to breast cancer; appraising the impact of unnecessary treatment may be relevant in other public interventions such as prostate cancer (273), cervical screening (277) or the management of cardiovascular disease, where treatments reduce the risk of future morbidity and mortality but have side-effects (92). Indeed, a number of cancer screening initiatives have reported varying outcomes when different sets of utilities are assumed (272, 278).

The literature is similarly heterogeneous in the duration and methods used to apply utilities. A difficulty in valuing screening interventions is that the process encompasses both temporary and chronic health states (58). The intensity and duration of the utilities associated with screening (169) and diagnostic anxiety (61, 279) vary significantly to the long-term sequelae associated with treatment (280), depending on whether this is classified by intervention or disease stage. There is ongoing debate about how best to overcome such issues (281), including the adaptation of conventional direct approaches (248, 269) or clinical guidelines on the duration of impact for each of the health states (3). Whether such adjustments are practical for screening interventions is debated and there are limitations of QALYs in interventions such as breast screening which may only have a transient impact on utility yet may be highly valued (132). Thorough sensitivity analysis of the durations applied to QALYs should be undertaken in any economic evaluation of population mammography screening to ascertain the effect of key drivers on cost-effectiveness (193, 196).

Two systematic reviews (187, 282) have previously explored the outcomes of economic evaluations relating to breast screening programmes. Schiller-Fruhwirth et al. (187) reported on the lack of breast screening specific utilities and insufficient reporting of validation in their review of economic models. A second review (282) reported similar findings relating to a paucity of methodologically appropriate utilities relevant to mammography screening. Other systematic reviews (135, 136) of economic outcomes in the broader breast cancer literature have been equally unable to combine screening values in meta-analyses due to insufficient numbers and inconsistencies in the approach and population used to derive them.

4.6.2 Strengths and limitations

The value of this review is that it provides a critical appraisal of the utilities used in economic evaluations of breast screening programmes, alongside a wider appreciation of the methodological issues and challenges associated with the empirical valuation of mammography and its sequelae. It offers new insight into the methodological issues informing the screening and overdiagnosis debate, and recommendations on where to direct future research to improve the appraisal of population screening services. Nonetheless, this review also has limitations. Some studies were not explicit in stating that the condition under study was relevant to mammography screening. Therefore, a subjective judgment had to be made by the reviewers about the health states measured and their relevance for inclusion. Second, the review only included studies published in English and may have excluded relevant utilities in other publications. Finally, a summary statistic for the health states associated with mammography screening could not be determined due to the heterogeneity between studies and the methods used to derive reported utilities.

4.6.3 Implications for breast cancer screening and treatment policy

Economic evaluations of breast screening programmes are hindered by poor quality utilities which are largely inappropriately valued or rely on assumption. There is a clear lack of appropriate data on overdiagnosis and overtreatment, most likely relating to the methodological challenges associated in valuing the uncertainty. However, to not capture the benefits and harms associated with the sequelae of screening is negligent and likely to lead to misplaced policy decisions for national breast screening strategy. This review has identified a clear gap in the inappropriate valuation of utility in breast cancer screening, and the collection of robust empirical data which captures the disutility of overdiagnosis is necessary before a re-evaluation of cost-effectiveness can be undertaken.

As screening evolves in line with technological advancements and improvements in genetic understanding (future risk), quality of life values should also be adjusted. Similarly, as the identification of low-risk disease from screening becomes more prevalent (27), it is likely that more personalised, risk-stratified utilities for active monitoring strategies will be required in breast cancer screening models.

4.6.4 Recommendations for future research

The economic evaluation of mammography screening remains problematic due to uncertainties in the natural history of the disease, duration of sequelae and risk of potential unnecessary treatment. The following methodological recommendations are highlighted for researchers planning future economic evaluations of population breast cancer screening:

- Economic evaluations should explicitly include all relevant utilities and disutilities associated with mammography screening and its sequelae. Overdiagnosis should be

explicitly captured in the evaluation of population screening policy, alongside extensive uncertainty analysis where there is debate on the extent of unnecessary treatment.

- New empirical evidence based on adequately informed utility data is needed to inform breast cancer screening decisions. The findings suggest the standard gamble and EQ-5D as the most appropriate economic measures to value screening health states, but vignettes should explicitly describe the advantages and risks of screening during the valuation process.

- Groups at high or low-risk for breast cancer should be considered in sub-group analysis, and quality of life values risk-stratified accordingly. It is likely that the management and prognosis for ductal carcinoma in situ will have markedly different implications than the disutilities associated with high-risk, invasive disease.

- Consistency in the duration for which the penalties are applied to screening, diagnosis and treatment related health states should be standardised by a panel of experts, clinicians and patients to prevent study heterogeneity driving cost-effectiveness results.

- Breast cancer screening evaluations assume perfect compliance with treatment which may not be reflective of clinical practice. The utility of active surveillance or non-invasive management, included in other population cancer screening evaluations, may be adopted by some women with low-risk disease and should be considered in the breast cancer setting.

4.7 Conclusion

Breast cancer screening programmes are deemed cost-effective for women aged 50-74 in the general population. Nonetheless, the evidence informing breast cancer screening policy have several limitations that must be addressed to determine what would be the most cost-effective approach. This review highlights the methodological challenges associated with valuing the utilities and disutilities associated with breast cancer screening, and suggests economic measures are unlikely to adequately capture the outcomes of screening in terms of quality of life.

There is no single recommended approach for valuing the health states associated with breast cancer screening and its sequelae, but women should be properly informed about the benefits and risks during the valuation process and vignettes. Overdiagnosis is not appropriately accounted for in the appraisal of mammography screening and undervaluation may lead to inappropriate decisions on the value of screening programmes. The measurement of health state utility values derived from adequately informed individuals, as well as sub-group analysis by risk group, is necessary if the debate on population screening programmes is to be adequately addressed.

The next chapter will seek to address the issues highlighted in this review, by exploring what would be an appropriate economic approach to value the disutility of overdiagnosis and the risk of unnecessary treatment.

CHAPTER SUMMARY

What is known?

- The methodological quality of utilities informing economic evaluation may be assessed according to the health states, methods, descriptive system and population used to derive them, and the generalisability of the condition, severity and population against those in which they are being applied to.
- Direct utilities should also be assessed for validity, reliability and feasibility.
- Valuing health states for mammography screening poses additional challenges related to the unknown natural history of the disease, paradox of overdiagnosis, variable duration of impact, and mixed preferences in decision-making.
- None of the systematic reviews to date have explored how the benefits and risks of breast screening have been captured in terms of utility, specifically considering the methods used and whether overdiagnosis has been captured appropriately.

What this chapter adds

- Economic evaluations of breast cancer screening programmes are unlikely to fully capture the benefits and risks associated with the sequelae of screening.
- None of the studies explicitly valued the risk of overdiagnosis and unnecessary treatment during the measurement of utility.
- The utilities applied for DCIS are the same as those for invasive breast cancer, despite markedly different prognoses and treatment pathways for low-risk disease.

CHAPTER 5

Methodology for an empirical study valuing health states for low-risk ductal carcinoma in situ

5.1 Introduction

The methodology applied in an empirical study valuing health states for low-risk ductal carcinoma in situ is outlined in this chapter. The study was motivated by the findings from the systematic review, which demonstrated a paucity of utilities for low-risk breast cancers in the economic evaluation of breast cancer screening programmes. Although women may never know whether they are the subject of overdiagnosis, the harm from the treatment that ensues may potentially be captured in the debate. The chapter begins with a summary of the main objectives of the empirical work, and the justification behind the techniques applied. This is followed by a detailed description of the study design, economic instruments and analytic techniques applied. Finally, the feasibility, reliability and likely ability of the methods to capture the true harm of overtreatment is considered against the current literature.

5.2 Background

5.2.1 Controversies in breast cancer screening

Choosing between treatment alternatives requires a balance of costs and risks against the possible benefits (283). In earlier chapters, the benefits and risks associated with breast cancer screening were contested. It was observed that whilst breast screening programmes improve breast cancer survival and quality of life through earlier diagnosis, (20, 33, 56), they also contribute to the overdiagnosis and overtreatment of low-risk disease that may never have caused symptomatic harm (3). It is imperative that the evidence informing policy

decisions appropriately capture these benefits and harms in the valuation process, so that the risks are balanced, and the most efficient strategy is implemented.

The major challenge in valuing the sequelae of breast cancer screening relates to the uncertain natural history of the disease (75). Patients must make treatment decisions based on how they feel about the trade-offs that exist in often unknown benefits and risks. For most women this is related to the anxiety about recurrence versus the side-effects and morbidity of treatment on quality of life (276). Ductal carcinoma in situ is a pre-invasive condition primarily identified through mammography screening (90). It is not known at the individual level whether all DCIS will develop into an invasive breast cancer, and so treatment is given to reduce concerns about the risk of progression (28). At the same time, treatment can also induce serious side-effects, costs and detriments to quality of life (176), which may have been unnecessary for those with low-risk disease. However, the utilities informing breast screening and treatment policy are limited in their ability to adequately capture these benefits and risks (284).

Choosing an economic measure capable of overcoming such issues is difficult. There is no single ‘gold standard’ method for capturing utility. The systematic review presented in Chapter 4 raised a number of methodological challenges associated with eliciting utilities for breast cancer (284). Uncertainties in the natural history of the disease, duration of effect, and heterogeneity in methodology and population applied, render the valuation of overdiagnosis problematic. Determining the impact of breast cancer treatment on quality of life is complex but the ideal economic measure should address some if not all these limitations.

Whilst there is agreement that quality of life should be captured in the evaluation of breast cancer interventions (285), no consensus exists on how this should be measured. The current assumption held in economic evaluations is that overdiagnosis and overtreatment are captured in outcomes, through the side-effects and the additional cases overtreated or treated earlier (286). Such assumptions are inappropriate as they imply that all treatment is held equal, regardless of outcome, and so do not truly capture the potential risk of unnecessary treatment in their valuation. There is evidence to suggest that women who are treated may report higher utility or satisfaction if they have been told or believe that the treatment was necessary (274). As Raftery and Chorozoglou (31) point out, to quantify the true harm of overdiagnosis would require surveys to ask women if their quality of life would be different if it could be shown that there was a risk that their treatment had been unnecessary.

The objective of this chapter was to explore potential economic methods to derive utilities associated with breast cancer, which capture the risk of overdiagnosis and unnecessary treatment. In doing so, the methodological challenges in health state valuation are discussed, including the uncertainties associated with valuing population screening outcomes, overdiagnosis and the impact of risk on utility.

5.2.2 Economic methods and the issues of valuing overdiagnosis

There are two main approaches to deriving health state values (6). The first is to directly value discrete descriptions of health or its treatment from first principles. Alternatively, another approach is to measure health indirectly using a standardised descriptive system. It is not the purpose of this chapter to summarise the theoretical foundations behind these methods, which were already established in Chapter 3, but to deliberate if and how such

methods are appropriate and might be applied to quantify the disutility of overdiagnosis. A brief reminder of the advantages and disadvantages of utility measures is shown in Table 5.1.

More traditionally, empirical studies have derived utilities through ***direct*** measurement of health descriptions using the visual analogue scale (VAS), time trade-off (TTO) or standard gamble (SG). The visual analogue scale is the simplest approach but is often criticised for not producing 'true' utilities, due to the choice-less nature of the rating task (138). There are limited data on the use of the VAS in quantifying health states associated with risk, but the lack of uncertainty in the decision process is likely to limit its suitability in capturing the harms of overtreatment. Dyer and Sarin (287) argue the VAS demonstrates a stable relationship with other preference-based measures, but the wider evidence (141, 288) suggests this relationship is highly variable and the VAS is unlikely to truly capture the trade-off of decisions in isolation.

The standard gamble and TTO are regarded as superior to the VAS because they directly capture preferences for discrete descriptions of health (16). Whilst the TTO is often preferred to the standard gamble due to its lesser cognitive demands (valuing time is easier than valuing probability), it requires conditions of certainty (264) which are violated by the concept of overdiagnosis. Specifically, it may not be methodologically appropriate or valid to ask participants to trade-off years in perfect health against years spent in a health state with an unknown outcome, i.e. a health state which may change within the traded time from DCIS to invasive cancer. Conversely, the standard gamble *can* overcome this limitation because it deals with risk and probability (140). This point of view is disputed by Richardson (150) who argues that the uncertainties embodied by the SG may bear little resemblance to the actual clinical decision made. However, the standard gamble has demonstrated reasonable validity

in two analogous studies eliciting utilities for cancer surveillance with variable disease progression (289, 290).

The most widely adopted method for eliciting utility is through generic multi-attribute measures (270) that capture utility ***indirectly***. Breast cancer health states are disaggregated into standardised health dimensions, and pre-specified weights are applied to adjust patient scores to reflect those from the general population. The EQ-5D (3L/5L) has been used extensively in the breast cancer literature for its validity and feasibility of use alongside other trial-based reported methods (291). Whilst decision bodies recommend the use of standardised descriptive systems to measure benefit in economic evaluations (8), generic instruments have been criticised for being insensitive to important oncologic domains of quality of life (292). Furthermore, there have been concerns that the response options may not capture changes in utility in low-risk conditions (293), such as DCIS, partly due to the ceiling effects of the limited levels of severity or patient biases in self-validating treatment.

Ultimately, the valuation of overdiagnosis requires economic measures which can capture the risk of uncertainty associated with the disease. A review of the literature (135) suggested both direct and indirect methods were appropriate for deriving utilities relevant to breast cancer. However, the values derived are unlikely to have truly captured the disutility of overdiagnosis if the women were inadequately informed about the risks during the decision process.

Simply applying a standardised descriptive system to patients in a trial may not value the true impact of unnecessary treatment because women may report higher utility or satisfaction if they have been told, or believe, that the treatment was necessary. Tosteton et al. (192) observed that although false positive mammograms had a significant impact on anxiety, this

was not reflected by prospective EQ-5D utilities. Furthermore, knowledge of uncertainty and risk is likely to change over time and therefore the 'value' of certain states may also change. The use of vignettes, which explicitly stipulate both benefits and risks, may offer a tangible option to negate this bias and improve the validity of the utilities derived. Descriptions of health including prognosis and recurrence risk have been applied previously to elicit utilities for other low-risk cancers using both public and patient populations (254, 289). Whilst this approach has been less commonly used alongside indirect methods, it has demonstrated reasonable validity provided the vignettes are designed to reflect the clinical condition and are not simply mapped to the instrument dimensions (294).

This rest of this chapter investigates the validity, reliability and feasibility of direct and indirect methods to value the disutility of the sequelae of overdiagnosis, using low-risk DCIS as a case-study for women potentially subject to unnecessary treatment.

Table 5.1: A summary of the most commonly adopted methods used to elicit utilities

Category	Method	Overview	Strengths	Limitations
Direct	Standard gamble	Choice between (i) varying probabilities of a gamble with two possible outcomes, and (ii) certain outcome	• 'Gold standard' (rooted in the axioms of EUT) • Can account for risk or uncertainty • Widely used in oncology	• Cognitively & time demanding • Potential bias or contamination from individual risk-attitude, gambling effect or loss aversion • Ceiling effect
	TTO	Choice between two alternatives of certainty of different durations	• Easier to complete than the SG • Good test-retest reliability and consistency in ranking	• Unable to accommodate the risk of uncertainty in the trade-off • Prone to duration bias, time preference and unwillingness to trade
	VAS	Health states are valued on a rating or interval scale (usually anchored best-worst imaginable health or dead-perfect health)	• Most feasible technique • High completion rate • Can be used alongside indirect methods (EQ-5D)	• Does not technically elicit utilities as no uncertainty in the decision process • Mapping formulae vary widely • Susceptible to response spreading and end-point bias
Indirect	AQoL	Generic questionnaire with multiple versions including 4-8 dimensions (independent living, social relationships, physical senses, psychological well-being); 4+ levels of severity	• Relevant population set • Covers a wide range of health issues and domains	• Generalisability limited • Complex number of health states • Limited practicality/completion rate
	EQ-5D	Generic questionnaire with 5 dimensions of health (mobility, self-care, usual activities, pain, anxiety/depression); 3 or 5 level version of severity available	• Widely used internationally • Preferred by NICE • Practical and easily administered in RCTs	• Insensitive to small QoL changes • Limited dimensions of impairment • Prone to ceiling effect
	HUI-3	Adult version has 8 health dimensions (vision, hearing, speech, ambulation, dexterity, emotion, cognition, pain); 4-6 levels of severity	• Covers a wider range of health impairments • Widely used internationally	• Does not cover role/social function • Takes longer to complete • Less widely used in oncology
	SF-6D	Generic questionnaire with 6 health dimensions (physical functioning, role limitation, social functioning, pain, energy, mental health); 4-6 levels of severity	• Widely used internationally • Cover a wide range of health dimensions	• Limited dimensions of impairment • Prone to a 'floor-effect' • Limited practicality/completion rate

5.2.3 Rationale for the study

There is no consensus on the most appropriate measure to value utility. Both direct and indirect methods have been used to elicit utilities for health states associated with breast cancer screening, but the EQ-5D or an alternative a direct risk-based approach (standard gamble, TTO) were the commonest methods applied in the economic evaluation of breast screening programmes (Chapter 4). The literature review undertaken earlier in this thesis demonstrated that DCIS was inadequately valued in current appraisals of breast cancer screening programmes. Although the utility of treating DCIS has been valued previously (148, 254), none of the studies valued active monitoring, overdiagnosis or the potential for unnecessary treatment in the vignettes.

The systematic review (284) also raised several methodological challenges associated with eliciting utilities for breast screening related health states. Uncertainties in the natural history of the disease, risk perception, duration of the sequelae and heterogeneity in methodology render the valuation of overdiagnosis problematic. Current utility estimates assume all treatment is held equal, regardless of outcome, and so do not truly capture the potential risk of overdiagnosis in their valuation. Similarly, the review of active monitoring models did not identify any utilities for active monitoring in the breast cancer setting.

Primary research to derive a set of appropriate valuations was therefore deemed necessary to address the paucity of utilities for breast screening health states potentially subject to overdiagnosis. Whilst addressing all the gaps in the literature would be beyond the scope of this thesis, this study aimed to apply economic methods to address the major limitations around overdiagnosis, so that both the benefits and harms are captured in the outcomes applied in the economic evaluation of breast screening programmes.

5.3 Aim and objectives of the empirical study

The aim of the empirical study was to quantify the utilities associated with the treatment of low-risk DCIS.

There were three main sub-objectives of the quantitative work:

1. To quantify women's preferences for managing low-risk disease potentially subject to overdiagnosis, including the acceptability of active monitoring as an alternative treatment.
2. To address the limitations identified by the review by explicitly including the risk of unnecessary treatment in the valuation process.
3. To elicit utilities on the sequelae of mammography to inform QALYs in the economic evaluation of breast cancer screening programmes.

5.4 Methods

Utilities were elicited for hypothetical health states describing the different treatments for low-risk DCIS. The empirical research was divided into three stages:

(i) Pilot study and focus groups to refine the methodology

(ii) Collection of utilities in the main study

(iii) Analysis and application of the derived utilities

This chapter focuses on parts one and two and describes the techniques and population used in the pilot and main study analyses.

5.4.1 Health state definition

5.4.1.1 Development of the vignettes

Bespoke descriptions of health were designed to represent the treatments and potential outcomes associated with low-risk ductal carcinoma in situ. A literature review (176, 284) was conducted to identify the relevant dimensions of quality of life affected by breast cancer screening and treatment. Seven hypothetical health states describing low-risk DCIS were defined:

A. Breast conserving surgery alone
B. Breast conserving surgery with radiotherapy
C. Mastectomy +/- immediate reconstruction
D. Active monitoring with 40% risk of progression in 10 years
E. Active monitoring with 20% risk of progression in 10 years
F. Active monitoring with 10% risk of progression in 10 years
G. Treatment for progressed DCIS

The sequelae and uncertainty attached to outcomes were guided by consultation with clinicians and population health experts. Three hypothetical levels (10%, 20%, 40%) of risk were specified to ascertain the impact of unnecessary treatment on utility, and were consistent with those used in studies valuing low-risk cancer surveillance (289) and estimated breast cancer recurrence after treatment (295, 296). Risks were displayed as both percentages and in absolute numbers (event per 100 women), because combined communication of probability has been shown to increase lay understanding (111). Progressed DCIS was also included to determine the magnitude of effect on utility of an adverse outcome following surveillance.

Endocrine treatment was explicitly described as optional for progressed disease due to the variable compliance and side-effects experienced by women, as reported in the literature.

5.4.1.2 Framing and labelling

The way in which health state descriptions are developed and framed can dramatically influence the resulting valuation (297). The term 'cancer' is known to introduce negative connotations, lower utility and change preferences toward more invasive treatment. Research conducted by Hersch et al. (298, 299) exploring treatments for early breast cancer found that preferences favoured more invasive surgery after changing the terminology for in situ disease from "abnormal cells" to "cancer". Similarly, O'Connor (268) and McNeil (300) reported significantly lower utilities when negative framing was used to describe the side-effects of cancer treatment compared to positive clinical outcomes.

Vignettes were unlabelled and used standardised mixed (neutral) descriptions to minimise any bias in the framing of potential benefits and risks (245). This permitted both advantages (reduced risk of recurrence) and risks (side-effects, overdiagnosis) to be considered concurrently during the valuation process. The same lead in was used to introduce each health state to ensure consistency in the presentation of the vignettes (Appendix 3). DCIS was explicitly clarified as a *"pre-invasive disease of the breast,"* to limit the impact of cancer terminology on preferences. However, women were informed that there was a risk that DCIS might become invasive breast cancer so that the descriptions accurately reflected the real-life decisions made. An example of the pilot health state description and presentation of information for breast conserving surgery is presented in Figure 5.1.

Gamble 1:

Imagine you attended a routine breast screening mammogram. The x-ray identified some abnormal cells, but they were confined to the milk ducts of the breast.

The abnormal cells were described by the doctors as LOW risk DCIS as they had a very low chance of developing into a cancer.

You received the treatment in the health state described on the card:

Health state A

- The **affected part** of the breast has been removed by **day** surgery.
- You have some scarring and shrinkage in the affected breast.
- You feel more conscious of your body image and your sexual drive is reduced.
- The **chance of the cells coming back** in the next 10 years is **20%.** That is, **20 out of every 100** women will experience a recurrence within 10 years.
    - You would require further surgery, radiation and possibly anti-cancer medication but there is no increased risk of dying
- You are told by your doctor that there is a **60% chance** that **you never needed the surgery:**
    - This means you may have unnecessarily experienced side-effects from treatment as the cells would likely not have become harmful.

Figure 5.1: An example vignette for health state A as presented in the pilot

5.4.1.3 Piloting and focus group feedback

The vignettes were extensively piloted on patient focus groups to check the content validity of the descriptions and then modified to reflect any additional attributes deemed important by patients and clinicians. This consisted of two focus groups of between six and ten women from the Patient Advisory Group at Breast Cancer Network Australia, all of whom had undergone treatment for an in situ or invasive breast cancer.

Although the phrase ‘abnormal cells’ was initially used in lieu of cancer as recommended in the literature, feedback from patient focus groups suggested that this term was too ambiguous and created some confusion as to what participants were being asked to value. Therefore, the phrase ‘abnormal cells’ was changed to ‘ductal carcinoma in situ’ in the questionnaire lead in and a brief explanation was suggested at interview to facilitate participant understanding of the condition.

A compromise had to be made between the volume of information presented and relevant side-effects selected for inclusion. Only minor iterations to the wording of side-effects and the presentation of monitoring risk were deemed necessary, as illustrated in the changes presented for ‘Scenario D’ in Figure 5.2. The bold highlighting of benefit and risks within the main text was also removed to avoid undue attention or bias in favour of attributes. The final vignettes were reviewed by a group of patients, clinicians and health economists, and may be viewed in full in Appendix 3.

5.4.2 Choice of elicitation method

5.4.2.1 Choosing the instrument

The systematic review suggested that both direct and indirect methods were appropriate for deriving utilities for breast cancer. Therefore, both approaches were used to elicit and contrast utilities for the purposes of this research. Three techniques were applied in the empirical study; the visual analogue scale (VAS), standard gamble and EQ-5D-5L questionnaire. The methods were selected following a systematic review and were based on their ability to incorporate risk into the valuation process, and consistency with the NICE recommendation (271) for a choice-based technique as a reference case alongside an indirect multi-attribute utility instrument (EQ-5D).

Figure 5.2: Health state F as displayed in (i) the pilot and (ii) with changes for the main study

(i) Gamble 4:

Imagine you attended a routine breast screening mammogram. The x-ray identified some abnormal cells, but they were confined to the milk ducts of the breast.

The abnormal cells were described by the doctors as LOW risk DCIS as they had a very low chance of developing into a cancer.

You received the treatment in the health state described on the card:

Health state D

- The abnormality identified on the mammogram is closely **monitored** by regular imaging.
- You have a **mammogram once a year** and are told the result by letter.
- You do not need to see a doctor or have a breast examination unless any changes are picked up on your mammogram.
- You may be anxious at the time of the mammogram but otherwise do **not** experience any side effects from the treatment.
- You **do not** experience **any** unnecessary treatment (**0%** chance).
    - This means you **avoided** any unnecessary side-effects from invasive surgery if the cells never developed into an invasive breast cancer.
- The **chance of the cells progressing** in the next 10 years is **40%.** That is, **40 out of every 100** women will experience a progression within 10 years.
    - If this happens you would need to undergo surgery to remove either part of or the entire breast, possibly with radiation.
    - You may need additional treatment with hormone tablets or possibly anti-cancer medication but there is **no** increased risk of dying
    - The risk of death from breast cancer is no greater with this strategy

(ii) Scenario D

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario D:

Scenario D

- The abnormality identified on the mammogram is closely monitored by regular imaging.
- You have a mammogram once a year and are told the result by letter.
- You do not need to see a doctor unless any changes are picked up on your mammogram.
- You are advised about breast awareness and do regular breast checks at home.
- You may feel anxious or worried from time to time.
- The chance of the cells progressing in the next 10 years is 40%. That is, 40 out of every 100 women with low risk DCIS will experience some progression within 10 years.
    - If this happens you would require surgery and radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
    - You may need hormone tablets or a biopsy of the glands in the arm
- You are told there is a 60% chance that the cells will never progress into an invasive cancer during your lifetime. This means that:
    - 60 out of 100 women who choose to have monitoring will avoid overtreatment and unnecessary side-effects or pain from invasive treatment.
    - 40 out of 100 women would need to have surgery at some point during their lifetime.

The visual analogue scale was chosen to familiarise participants with the concept of valuing health. Whilst the VAS has been criticised for not technically producing utilities (no uncertainty during the decision process), it has been widely used to supplement other direct and indirect methods in the literature (6), including the EQ-5D. In particular, the VAS has strengths in facilitating participants to conceptualise the concept of valuing or ranking health (142). The VAS was therefore selected as the first method to help acquaint participants with the health states, prior to completing the more cognitively demanding tasks.

The standard gamble was the second method chosen from the systematic review. The main justification for using this method was its ability to incorporate risk and value preferences under conditions of uncertainty (e.g. disease progression). Whilst the TTO is often preferred to the standard gamble due to its lesser cognitive demands (valuing time is easier than valuing probability) (264), it requires conditions of certainty which are violated by the concept of overdiagnosis. Specifically, individuals cannot be certain whether DCIS will progress or not due to the unknown natural history of the disease. As such, it would not be methodologically appropriate or valid to ask participants to trade-off years in perfect health against years in a health state with an unknown outcome, i.e. a health state which may change within the traded time from DCIS to cancer. In contrast, the standard gamble *can* overcome this limitation because it deals with risk and probability and has been successfully applied in other studies eliciting utilities for cancer screening (254), in situ disease (248) and active monitoring previously (289).

The third and final method applied was the EQ-5D-5L questionnaire. An indirect technique was deemed necessary to supplement the standard gamble, in line with the recommendations of decision bodies (NICE) which stipulate the use of MAUIs to value health (8). Selecting

which method to use is important given that the instruments differ in structure, dimensions of health, methods and the populations used to derive them. The EQ-5D was chosen for this study because of its accessibility, validity and ease of completion alongside other direct methods (291). Furthermore, the EQ-5D was deemed the most commonly adopted indirect approach in the breast cancer literature and is recommended for use in trials of active monitoring (e.g. LORIS). Whilst other questionnaires such as the SF-6D and AQoL have better specificity and wider health dimension coverage (301), it was decided that the longer completion time required for these instruments would not be practical or feasible without incurring significant fatigue in a study with seven health states. Although Australian population tariffs are more limited for the EQ-5D, several studies have demonstrated good consistency with the utilities derived using Australian and UK population tariffs (302-304).

The final decision to be made for the indirect approach was whether to adopt the 3 or 5 level version of the instrument. The EQ-5D-5L was chosen specifically to avoid any bias in favour of favour of oncological treatment (305). Whilst the 5L questionnaire has developed some criticism within the oncological field for producing less favourable results than the 3L (306), it has not been employed in the mammography screening setting where the margins between health states may be more subtle and therefore warrant a wider range of severities. Similarly, the 3L has been criticised for its limited responsiveness for low-risk conditions (291), partly due to ceiling effects and therefore would be less likely to detect the smaller changes in utility observed in DCIS.

5.4.2.2 Duration

The health states associated with breast cancer screening and treatment have variable effects during the cycles of treatment. However, there is a wealth of evidence suggesting that the

side-effects following invasive treatment and ongoing surveillance may persist for years beyond the initial cancer diagnosis (95). Similarly, overdiagnosis is unlikely to have only short-term implications due to the ongoing physical and psychological effects of a potentially unnecessary intervention. Therefore, all health states associated with the treatment of DCIS were considered chronic in nature.

For comparability, participants were asked to consider that all health states in the standard gamble lasted 10-years in line with clinical trials of active monitoring (27). The EQ-5D and VAS are more limited to capturing a "snapshot" of the health state under valuation, but participants were encouraged to remember that the health states were chronic in nature and would similarly be followed up over a period of 10 years.

5.4.2.3 Face-to-face versus online

Women were interviewed face-to-face by a single interviewer (HB) to ensure consistency in the approach to data collection. Although face-to-face interviews have been shown to be more time and resource demanding (6), the completion rate and reliability of the derived utilities are improved over those collected via online surveys. Feedback from the pilot testing also suggested that the tasks would have been too cognitively demanding to complete online without assistance or clarification of what was being asked.

5.4.2.4 Interview format

The interview used an adaptation of the format described by Hayman (148) and Furlong (257) and consisted of four parts: a rating scale exercise, EQ-5D-5L, standard gamble and socio-demographic questionnaire. Before beginning the valuation tasks, participants were provided with a brief description of DCIS as some participants were less familiar with the

concept of the disease. However, this was limited to a few sentences explaining the nature of the disease, rather than the treatment, so as not to bias the generalisability of potential values.

A questionnaire booklet was designed for completion during the interview to facilitate the collection of utilities and socio-demographic data (Appendix 4). The booklet was divided into five sections (A-E) so that each exercise could be separated to reduce cognitive fatigue. An overview of each of the questionnaire tasks is illustrated in Figure 5.3 is summarised in the following section.

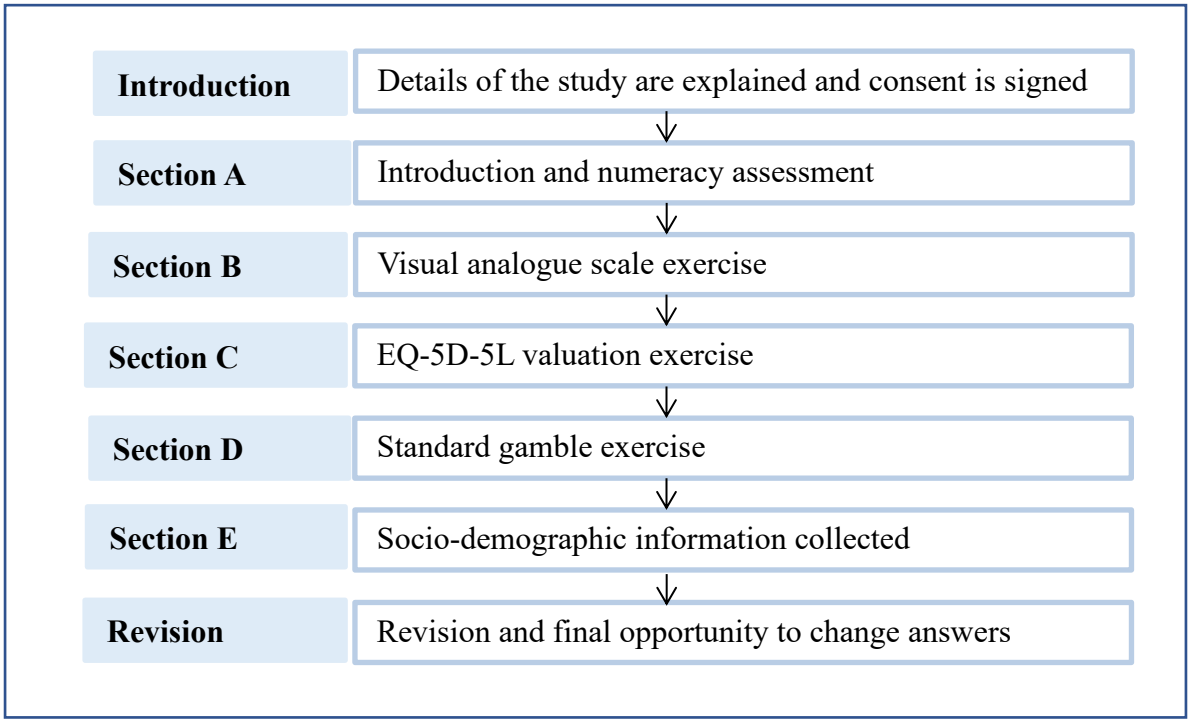


Figure 5.3: A flow diagram summarising the questionnaire tasks

5.4.2.5 Design of the questionnaire

Section A

Following a brief introduction to the study, participants were asked to complete a few short questions to test basic numeracy, risk and time preference. The objective behind these

questions was to familiarise subjects with the concept of risk-benefit scenarios (307, 308) and to aid the reviewer's judgement on the reliability of the utilities derived. There is substantial evidence demonstrating a general lack of risk literacy among the general population (309, 310), although Torrance (16) suggests this is reduced among those with knowledge of the condition. Proxy numeracy and risk comprehension was assessed using four standardised questions from the literature (311-313) describing varying probabilities of tossing a coin or winning a prize. Although none of the participants were excluded based on their answers to these questions (309), an initial assessment of participant numeracy and understanding of benefit and risk information was deemed helpful in ascertaining the level of support likely required by the participant.

Section B

Section B initiated the process of eliciting utilities using the visual analogue scale. Participants were asked to read, in random shuffled order, seven hypothetical health states and rank them from least to most desirable on a 20cm visual analogue scale, anchored from worst (0) to best (100) imaginable health. This involved quantifying how good or bad they felt the health state was by drawing a cross point on the scale and writing the corresponding figure on the dotted line next to letter for each health state. They were also asked to place being dead on the same scale so that the health states could be adjusted to the dead-perfect health scale, accordingly, using the formula:

$$VAS = (x-d)/(100-d)$$

where x corresponds to the scale placement of the health state and d corresponded to the value of dead (4).

Section C

In the third section, participants were asked to rate each of the health states in terms of the five EQ-5D-5L dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension has five levels of severity to describe the health state being valued. For example, in the usual activities domain the impact of treatment may be scored from having no problems, slight, moderate or severe problems, or being unable to do them. Each dimension was valued using one of the five hierarchical levels of severity, and the final reported scores were converted into utilities using general population tariffs (18).

Section D

For the standard gamble exercise, participants were asked to choose between two alternatives: (i) a gamble between varying probabilities of perfect health and being dead, or (ii) the health state being evaluated. The probabilities of the gamble were varied until the participant was indifferent between the two alternatives, corresponding to the utility for the health state. For comparability, it was assumed that all states were chronic in nature with a 10-year duration specified as per the LORIS trial protocol (27).

The methods were based on the format described by the Measurement and Valuation of Health Group (314) and have been used in previous studies to elicit utilities using the conventional standard gamble. Probabilities were varied using the recommended ping-pong technique (315) (i.e. values are alternated from high to low and back again), to reduce the risk of procedural bias observed with continuous ascending or descending probability changes associated with titration. Visual aids consisting of an A3-sized chance board and probability wheel were also used to aid comprehension (Appendix 5).

To ensure consistency in technique and participant understanding of the exercise, the initial gamble was set to 100/0 and the question phrased as follows (6):

> *"I want you to imagine you have two options; one option involving a risk and a second involving a certain outcome.*
>
> *I want you to decide between the two options, or if you think they are both equal please tell me. The amount of risk in Option 1 will be changed until we find out how much risk you will take to avoid the certain outcome in Option 2.*
>
> *The first choice I would like you to decide on is the health state described on the green card. You have already seen this card in the earlier exercises.*
>
> *As you can see on the board, Option 1 is a 100% chance of ten years in perfect health, with zero chance of immediate death. Option 2 is a 100% chance of being in the health state described on the card.*
>
> *Would you prefer Option 1 or 2?"*

According the participant's response, a tick, cross or equals sign was marked on the corresponding scoring sheet (Appendix 6). Rationally, participants were expected to choose Option 1 because guaranteed perfect health ought to be preferable to an inferior health state. If they did not choose Option 1, efforts were made to ensure that they had correctly understood the task. (If after further explanation they were indifferent, the health state was scored a utility of 1.0). If Option 1 was chosen, the gamble was changed to 10/90:

"Option 1 is now a 10% chance of ten years in perfect health, with a 90% chance of immediate death. Option 2 is still a 100% chance of being in the health state described.

Would you choose Option 1 or Option 2 now?"

The probabilities of the gamble in Option 1 were changed iteratively in an alternating manner using 10 per cent intervals until the point where the participant was indifferent between the two options. In situations where the participant could not decide on a point of indifference, the probability in the gamble was reduced to the midpoint between the two preferences (e.g. 95% for preferences between 90% and 100%) to determine the final utility value.

Section E

Demographic data were collected in the final section of the interview to enable sub-group and regression analyses. This section comprised of 20 questions covering a range of socio-demographic characteristics and information on breast cancer screening and treatment history. The questions were adapted from those used in the BCNA Review & Survey group (316) questionnaire (a survey gathering detailed scientific and demographic data relevant to breast cancer patients) and the Australian national census (317). A Likert scale and free text box was added at the end of the booklet to appraise the feasibility of each of the methods, and to assist the interviewer's appraisal of whether the participant had understood the tasks (179, 245).

5.4.3 Sampling

5.4.3.1 Population

A decision had to be made about whose preferences to elicit (17). Gaining an understanding of the preferences of a relevant group may help policy makers quantify overdiagnosis given that the implications are cognitively difficult to understand and so poorly researched. In this exploratory study, utilities were elicited from women with and without breast cancer, both of whom are likely best placed to appraise the impact of overdiagnosis, as potential subjects of unnecessary treatment from screening attendance. It could be argued that men's values are also relevant as taxpaying members of the general population (270), but given that this was an exploratory study the sample was limited to a relevant population for pragmatic reasons. Experts were used during the development of the vignettes but not during the main analysis as their familiarity with the health states may have biased the results.

5.4.3.2 Eligibility Criteria

Utilities were obtained from women in Melbourne, Australia between April and September 2018. Eligible women had to be over 30 years (breast cancer incidence is uncommon below this age) and have basic English language skills to participate in the tasks. Both patient (DCIS or early invasive breast cancer) and non-patient preferences were elicited so that differences among women with varying breast cancer experience could be explored. It was hypothesized that patients would have a higher preference toward more invasive treatment (248). Women undergoing active treatment (except for endocrine therapy) or advanced breast cancer were excluded because they were deemed high-risk cases.

5.4.3.3 Recruitment

Potential participants were recruited by email invitation from the Lifepool cohort (318) and Breast Cancer Network Australia (BCNA) Review & Survey group (316). Both groups have access to a register of members interested in participating in breast cancer research. The use of these groups for recruitment was justified for two reasons. First, these women had expressed an interest in participating in breast cancer research and were therefore more likely to ensure a higher response rate than anonymous email invitation. Second, recruiting patients directly from the hospital clinic would have required additional health services ethical approval and would been more time-consuming to establish eligibility for the study within the limited time constraints of the PhD.

Women were asked to contact the researcher directly if they were interested to arrange the interview. The project coordinators at the two registries stratified the sample prior to invitation so that women with a diverse group of socio-demographic variables were included where possible. An additional group of women outside of the two registries contacted the researcher directly through snowball sampling of participants' family, friends or colleagues.

5.4.3.4 Ethical approval

Individual interviews were undertaken at the Melbourne School of Population and Global Health and local community facilities in Victoria. Two interviews were conducted using telecommunication software to accommodate participants' disabilities for those unable to travel. No financial incentive was provided. The study was approved by the Health Sciences Ethics Sub-Committee at the University of Melbourne (ID: 1750252) (Appendix 7).

5.5 Data preparation and analysis

5.5.1 Sample size

Sample size calculations are routinely undertaken prior to conducting an empirical study to determine the number of observations necessary to make inferences about a population from the sample (319). However, due to the expected heterogeneity in individual risk aversion, sample size calculations are not routinely performed for studies deriving utility.

The methods described by Hayman (248) and Hall (243) were used to estimate a proxy sample size for the sample population, as stipulated by the recruitment registries, to accommodate the expected heterogeneity relating to individual risk aversion. The standard deviation of SG utilities for patients reported in the literature is 0.2 (257). A sample size of ≥100 women was therefore estimated to detect mean utilities at a 95% confidence level of 0.04 and 90% power.

5.5.2 Data preparation

The values derived in the empirical study had to be transformed onto the conventional 0-1 utility scale (dead-perfect health) to enable statistical analysis and comparison of the computed values.

1. Visual analogue scale

The VAS values collected in the questionnaire were converted from a 0-100 to a 0-1 scale by dividing each score by 100 if being dead was valued as being the worst imaginable health. For example, a VAS score of 70 was recorded as 0.70 (70÷100). Where death was not placed at zero by participants, utilities were adjusted accordingly using the formula (4):

Equation 1: $VAS = (x-d)/(100-d)$

Where x corresponds to the scale placement of the health state and d is the value of being dead. For example, if health state x was valued at 70 and being dead at 10, then the corresponding final utility would be: VAS = (70-10)/(100-10) = 0.67.

2. EQ-5D-5L

The reported scores from the EQ-5D-5L dimensions were converted into utilities using general population conversion tariffs. In the absence of Australian general population tariffs for the 5L instrument, the methods adopted by McCaffrey et al. (302) were applied to convert the scores into utilities using a comparable general population dataset (18). Given the similarities between the British and Australian breast cancer screening programmes and populations, a UK mapping matrix was deemed appropriate to convert the aggregated scores. Further, several studies have demonstrated good consistency between the utilities derived using UK and Australian tariffs (303, 304).

The EQ-5D-5L health states were converted into a single index value using the 'EQ-5D-5L Crosswalk Index Value Calculator,' available on the European-Qol website (320). The formula uses a crosswalk function based on the existing value sets for the 3L which are mapped against the new values elicited using the 5L descriptive system (291). Each dimension of health has a summary index (or weight) attached to each level of severity. The calculator then deducts the appropriate weights from 1, the value for full health (i.e. state 11111), to generate the final utility score.

For example, a health state scored as 11134 generates a final utility score equivalent of 0.512 using the UK EQ-5D-5L crosswalk table value set on Microsoft Excel (320).

3. Standard gamble

The utilities for the standard gamble were equivalent to the probability at which the participant was indifferent between the health state being valued and the gamble between perfect health and death (16), i.e.

Equation 2: $U_x = (p)$

For health states considered worse than death, the equation is modified so that death is compared to a gamble between perfect health (p) and the worst health state (1-p). The probability is varied until the participant is indifferent, at which point the utility for the health state is given by:

Equation 3: $U_x = -p/(1-p)$

5.5.3 Statistical analysis

Data analysis was carried out using STATA Version 15 for Windows 7 (321). A priori significance levels were set at 0.05 for comparing the differences between utilities, consistent with clinical studies of this nature (264). Mean and median values were summarised with associated degrees of dispersion. Differences in characteristics between patients and non-patients were analysed using t tests and chi-squared tests, where appropriate for continuous and discrete or categorical variables (322).

Further statistical analyses were conducted to determine whether any differences in utility within or between the two population groups were of meaningful significance. The decision of which statistical test was most appropriate was determined through evaluation of the distribution of the data, sample size and type of variables included in the analysis. The recommendation is to use parametric tests for data with an approximate normal distribution and non-parametric tests for those that violate this assumption (323).

Visual inspection of the histograms and formal skewness testing confirmed that while some data were approximately normally distributed, others were left skewed (Shapiro-Wilkes: negatively skewed). Therefore, non-parametric tests were used to analyse differences in the utilities due to the non-normal distribution of the data (324). A Mann-Whitney U test was used to determine whether utilities for the seven health states differed between patient and non-patient groups. Correlation between the different methods was assessed using the Spearman's correlation coefficient for pairwise comparisons (323), and linear regression was used to determine if demographic variables were significant predictors of utility.

5.5.4 Quality appraisal

The criteria set out by Brazier et al. (6) were applied to assess the performance of the economic measures and quality of the utilities derived in the study. Although there are no formal quality checklists for evaluating utilities, appraising the quality of the data in terms of reliability, validity and feasibility is important to assess whether the utilities are suitable for use in economic analyses (182).

i) Validity

A three-part approach was used to assess the validity of the utilities. The first part concerns the validity of the economic measures applied to elicit the utilities. There is extensive literature (4, 178) to support the validity and reliability of choice-based techniques (standard gamble) and the EQ-5D instrument (325) in many populations and conditions, including breast cancer. The derived utilities therefore meet with this first criterion.

Second, convergent or empirical validity between the three methods was assessed by comparing the rank order of the health states, with standard gamble utilities hypothesized to be higher than those elicited using the VAS *a priori* (178). A logic test was also included to test internal consistency. Whilst the rank order between the seven health states was expected to vary between individual women, the utility for active monitoring (10%) should, in theory, be at least as or more preferable than an active monitoring state with a worse prognosis (20% or 40%).

Third, the health descriptive systems were assessed in terms of content, face and construct validity. Descriptive validity was evaluated directly by patient and expert feedback and refinement during the piloting of the vignettes and study questionnaire, as described in the pilot section earlier in the chapter.

ii) Reliability

The reliability of the derived utilities was ascertained by demonstrating that the values were consistent and based on adequately informed choices. Focus group feedback and interviewer judgement was used to appraise participant understanding of the vignettes and tasks. Reliability was also objectively measured by asking participants to rate the difficulty of the

tasks on a Likert scale in the final section of the questionnaire. Limited test-retest reliability was performed for the six women who agreed to participate in both the pilot and main study analyses and results were deemed within the expected range for patients (0.2 SD) (245, 248).

iii) Feasibility

The vignettes and questionnaire were extensively piloted and reviewed by patients, clinicians and advisory committees, to ensure that the tasks were reasonable and simple enough to administer. Additional feasibility checks were included through both qualitative feedback from the focus groups and quantitative measurement of perceived difficulty at the end of the interview. Study participation rates, completion time and degree of missing data were used as a proxy to determine the practicality of the methods (238).

5.6 Conclusion

In this chapter the methodology for the development of an empirical study valuing the treatment of low-risk DCIS was reported. The first half of the chapter focused on the methodological issues in health state valuation, particularly with regards to the application and uncertainty surrounding the risk of overdiagnosis. The second half of the chapter was dedicated to the development of the vignettes and study questionnaire. The VAS, SG and EQ-5D were chosen as the most appropriate methods to derive the utilities of treatment using risk in the vignettes, with patient and non-patient preferences being used to analyse the data. The findings of the empirical study are presented in the next chapter.

CHAPTER SUMMARY

What is known

- Current breast cancer estimates of utility do not capture the risk of overdiagnosis and overtreatment in the valuation process.
- An empirical study seeking to capture the benefits and risks is necessary to inform the economic evaluation of breast cancer screening and treatment programmes.
- Low-risk DCIS managed via standard surgery and active monitoring (LORIS) provides an ideal case-study in which to explore preferences related to the benefits and harms of treatment.

What this chapter adds

- The methods for an empirical study aiming to address these limitations were described in detail.
- Standard gamble and EQ-5D were determined as the best approaches to facilitate the incorporation of risk into the valuation process alongside a choice-based technique (NICE recommendations).
- The validity, reliability and feasibility of the instruments chosen were justified and extensively piloted by patients, academics and clinical experts to ensure content validity.

CHAPTER 6

Results of an empirical study valuing health states for low-risk ductal carcinoma in situ

6.1 Introduction

This chapter presents the results of the empirical study, the methodology of which was described in the previous chapter. It covers an in-depth analysis of the utilities in relation to the population and methods used to derive them, and a comparison with values for in situ cancer treatment and active monitoring reported previously in the literature. The chapter concludes with a critique of the study validity, strengths and limitations, and consideration of the potential implications of the main findings. The results were presented in part, prior to the publication of the thesis, at the 4th World Congress on Controversies in Breast Cancer Congress in Melbourne, Australia (October 2018), and Health Economists' Study Group meeting in Norwich, United Kingdom (July 2019).

6.2 Overview of the empirical research

The objective of the empirical study was to quantify women's preferences for managing low-risk ductal carcinoma in situ. As outlined in Chapter 5, utilities were elicited for seven hypothetical health states describing treatments for low-risk ductal carcinoma in situ, with the risk of overdiagnosis or cancer progression explicitly stated within the vignettes. Women were recruited via two breast cancer registries (BCNA and Lifepool), over a six-month period between April and September 2018. Both patient (DCIS/early invasive breast cancer) and

non-patient preferences were obtained from women so that any potential differences between those with varying levels of disease education or experience could be explored.

Women were invited via email and were encouraged to contact the primary researcher to arrange an interview if interested in participating. All interviews were conducted by the primary researcher (HB) in Melbourne, Australia; either at the university, a local community facility or via online teleconferencing software for those unable to travel. Participants were asked to value each of the health states in turn, as part of a questionnaire, using three measures of utility: visual analogue scales, standard gambles and the EQ-5D-5L. Socio-demographic data and a history of breast cancer were examined as predictors of utility.

The following section provides a breakdown of the questionnaire response rates, followed by a summary of the socio-demographic characteristics of each of the sample groups and associated mean and median utilities elicited using the three different instruments.

6.3 Results

6.3.1 Sample characteristics

A total of 929 women were invited to participate in the main study analysis, with a further 36 women identified through snowball sampling. One duplicate invitation sent out by both the BCNA and Lifepool registries was excluded. Out of the 254 (26%) responses, 172 (68%) eligible women attended and completed the interview in full and were included in the final analysis. A summary of participant response rates and completion at each stage is provided in Figure 6.1.

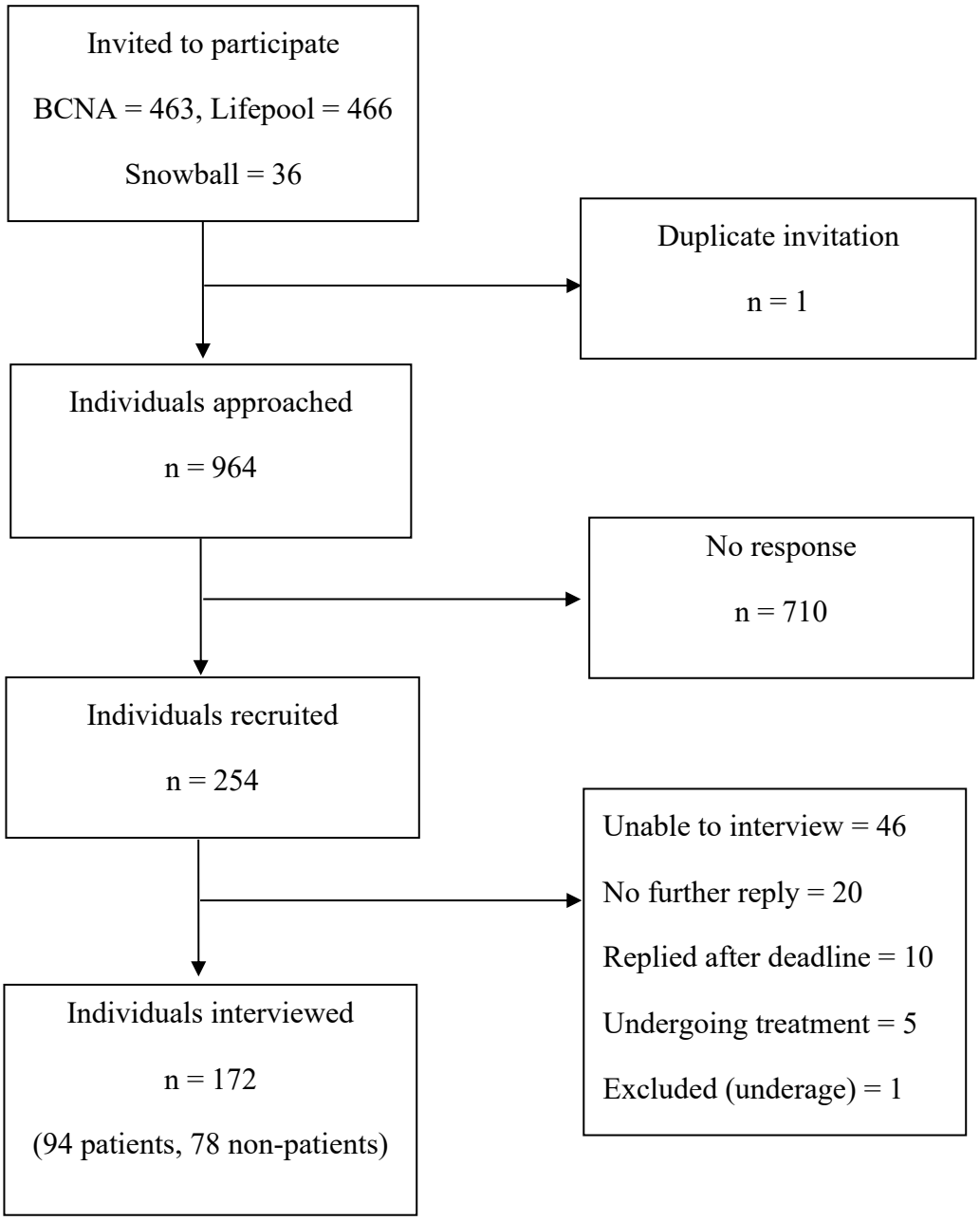


Figure 6.1: A flow chart of patient recruitment and participation in the main study

Utilities were obtained from 172 women (94 patients and 78 non-patients) who all successfully completed the interview in full. Patients took slightly longer to complete the interview compared to non-patients (56 vs 45 minutes) but reported fewer cognitive difficulties in understanding the tasks (33% versus 40% rated difficult or very difficult).

Table 6.1 shows the key characteristics of the sample. Although there were no significant differences between the two groups with respect to socio-demographic variables, women without a history of breast cancer reported lower levels of worry, were more likely to consider active monitoring and were slightly younger than patients on average. The demographics in the patient group were also generally representative of the female breast cancer population in Australia (37), where the average age of a first diagnosis is 61 years. However, screening participation in the sample was above the national average in both patient and non-patient sample groups (2), reflective of the recruitment of women from registries engaged in breast cancer research.

6.3.2 Utilities

Utilities are scaled from 0-1 (on a spectrum equivalent from being dead to in perfect health), so that values closer to 1 are indicative of a stronger preference for the reported health state. For all utility measures of health states under various scenarios, higher utility values indicate a more positive view (i.e. a preference).

6.3.2.1 VAS

Patient and non-patient values for the VAS are presented in Table 6.2. For both groups, the median utilities were higher than the means reflecting the skewed distribution of response values. The exception to this was for breast conserving surgery alone, which was rated more favourably by patients using the VAS (0.691, SD 0.19). There was considerable variability in the utilities elicited between individuals, with the widest dispersion around mean and median values observed for mastectomy and progressed DCIS among patients. Patients also had a stronger preference for invasive treatment compared to non-patients but valued active monitoring slightly lower, the difference of which was significant on statistical testing.

Table 6.1: A comparison of participant demographics with associated p value

Demographic	Patients n = 94 (%)	Non-patients n = 78 (%)	P value
Age in years: Mean (SD)	66.2 (6.7)	58.7 (11.9)	0.007
Education level: n (%)			
Secondary school	14 (15)	7 (9)	0.094
Certificate or diploma	23 (25)	16 (21)	
Undergraduate degree	19 (20)	29 (37)	
Postgraduate degree	38 (40)	26 (33)	
Socio-economic income:			
Low	12 (13)	6 (8)	0.356
Middle	25 (27)	16 (21)	
High	44 (46)	47 (60)	
Prefer not to say	13 (14)	9 (11)	
Relationship status:			
Single	9 (10)	13 (17)	0.432
Married	67 (70)	52 (67)	
Separated or divorced	10 (11)	6 (7)	
Widowed	8 (9)	5 (6)	
Prefer not to say	0 (0)	2 (3)	
Current employment status:			
Full or part time	40 (43)	42 (54)	0.032
Unemployed	2 (2)	0 (0)	
Retired	48 (51)	28 (36)	
Carer, volunteer or other	4 (4)	8 (10)	
Ethnicity:			
Australian	69 (73)	61 (78)	0.286
New Zealand/Torres Strait Islander	4 (4)	4 (5)	
European	11 (12)	6 (8)	
Asian	3 (3)	5 (6)	
American	7 (8)	0 (0)	
Middle Eastern	0 (0)	2 (3)	
Attended breast screening	76 (81)	59 (76)	0.408
Undergone a breast biopsy	94 (100)	15 (19)	0.000
Breast cancer diagnosis:			
DCIS	33 (35)	-	N/A
Invasive breast cancer	61 (65)	-	
***Treatment received:**			
Breast conserving surgery	65 (69)	-	N/A
Mastectomy	39 (41)	-	
Adjuvant radiotherapy	54 (57)	-	
Endocrine therapy	56 (60)	-	
Chemotherapy	36 (38)	-	
Sentinel lymph node	56 (60)	-	
Worried about breast cancer:			
Not at all	23 (24)	32 (41)	0.015
A little	49 (52)	35 (45)	
Some of the time	15 (16)	10 (13)	
A lot or all the time	7 (8)	1 (1)	
Difficulty of questionnaire			
Very easy or easy	44 (37)	35 (45)	0.296
Neither easy nor difficult	19 (20)	12 (15)	
Difficult or very difficult	31 (33)	31 (40)	
Would consider active monitoring?	57 (61%)	76 (97%)	0.000
Mean 10-year risk of recurrence free survival for active monitoring women were happy to accept (SD)	84.0% (0.22)	72.0% (0.16)	0.000

*Totals add up to greater than 100% as some patients received multiple treatments

Table 6.2: Visual analogue scale utilities

Health state (VAS)	Patient Mean (SD)	Patient Median (IQR)	Non-patient Mean (SD)	Non-patient Median (IQR)
A: Breast conserving surgery alone	0.691 (0.19)	0.703 (0.600-0.800)	0.623 (0.20)	0.649 (0.500-0.750)
B: Breast conserving surgery and radiotherapy	0.615 (0.23)	0.638 (0.500-0.800)	0.575 (0.19)	0.589 (0.474-0.706)
C: Mastectomy (+/- reconstruction)	0.567 (0.23)	0.553 (0.400-0.750)	0.454 (0.22)	0.456 (0.300-0.600)
D: Active Monitoring (40% risk of progression)	0.547 (0.25)	0.550 (0.400-0.710)	0.696 (0.16)	0.700 (0.611-0.800)
E: Active Monitoring (20% risk of progression)	0588 (0.25)	0.618 (0.450-0.800)	0.733 (0.15)	0.750 (0.680-0.850)
F: Active Monitoring (10% risk of progression)	0.645 (0.25)	0.693 (0.450-0.875)	0.775 (0.14)	0.800 (0.722-0.875)
G: Disease progression	0.498 (0.26)	0.500 (0.300-0.700)	0.519 (0.22)	0.500 (0.313-0.700)

6.3.2.2 Standard gamble

The SG values are presented in Table 6.3. Utilities were highest for active monitoring health states followed by surgery, radiotherapy and progressed disease in both populations. The utilities for most health states describing low-risk DCIS were close to full health, as indicted by the negative skew of the data, but more invasive treatments were associated with worse quality of life. This ceiling effect was more pronounced among patients, where a difference of only 0.090 was observed between the best and worst ranked health states (B/C and F).

Table 6.3: Standard gamble utilities

Health state (SG)	Patient Mean (SD)	Patient Median (IQR)	Non-patient Mean (SD)	Non-patient Median (IQR)
A: Breast conserving surgery alone	0.820 (0.18)	0.900 (0.800-0.900)	0.740 (0.21)	0.800 (0.600-0.900)
B: Breast conserving surgery and radiotherapy	0.790 (0.19)	0.875 (0.700-0.900)	0.700 (0.21)	0.700 (0.600-0.900)
C: Mastectomy (+/- reconstruction)	0.790 (0.19)	0.900 (0.700-0.900)	0.650 (0.23)	0.700 (0.500-0.800)
D: Active Monitoring (40% risk of progression)	0.850 (0.14)	0.900 (0.900-0.950)	0.870 (0.13)	0.900 (0.900-0.950)
E: Active Monitoring (20% risk of progression)	0.870 (0.13)	0.900 (0.900-0.950)	0.890 (0.11)	0.950 (0.900-0.950)
F: Active Monitoring (10% risk of progression)	0.880 (0.13)	0.900 (0.900-0.950)	0.890 (0.11)	0.950 (0.900-0.950)
G: Disease progression	0.810 (0.17)	0.900 (0.700-0.900)	0.740 (0.22)	0.800 (0.600-0.900)

6.3.2.3 EQ-5D

The EQ-5D-5L demonstrated the most consistent ranking of health states in both populations. As expected, utilities increased in line with a decrease in the risk of disease progression or reduced intensity of treatment (Table 6.4). The values of patient and non-patients using the EQ-5D were numerically similar, and no statistically significant differences were found within or between the two groups for any of the health states using this method.

Table 6.4: EQ-5D-5L utilities

Health state (EQ-5D)	Patient Mean (SD)	Patient Median (IQR)	Non-patient Mean (SD)	Non-patient Median (IQR)
A: Breast conserving surgery alone	0.767 (0.17)	0.768 (0.696-0.848)	0.800 (0.14)	0.817 (0.681-0.879)
B: Breast conserving surgery and radiotherapy	0.699 (0.19)	0.729 (0.592-0.837)	0.700 (0.19)	0.724 (0.592-0.837)
C: Mastectomy (+/- reconstruction)	0.606 (0.23)	0.649 (0.531-0.748)	0.578 (0.23)	0.614 (0.510-0.728)
D: Active Monitoring (40% risk of progression)	0.819 (0.15)	0.848 (0.791-0.879)	0.840 (0.15)	0.879 (0.848-0.879)
E: Active Monitoring (20% risk of progression)	0.826 (0.16)	0.879 (0.848-1.000)	0.863 (0.12)	0.879 (0.848-1.000)
F: Active Monitoring (10% risk of progression)	0.863 (0.15)	0.879 (0.848-1.000)	0.884 (0.13)	0.879 (0.879-1.000)
G: Disease progression	0.558 (0.28)	0.622 (0.388-0.745)	0.553 (0.27)	0.624 (0.422-0.736)

The utilities for active monitoring health states (whereby the risk of potential unnecessary treatment is reduced) were primarily driven by the lack of impairment in the four dimensions beyond the anxiety/depression dimension, as shown in Figure 6.2, whereas the conventional treatment health states were perceived to impact on all five dimensions of health. This resulted in lower values for the health states involving intensive treatment (A-C), where the perceived harm of overtreatment from surgery and radiotherapy or progressed disease (G) may be more extreme, because this was perceived to have impacted upon all five dimensions of the EQ-5D-5L as opposed to the health effects of anxiety/depression in isolation.

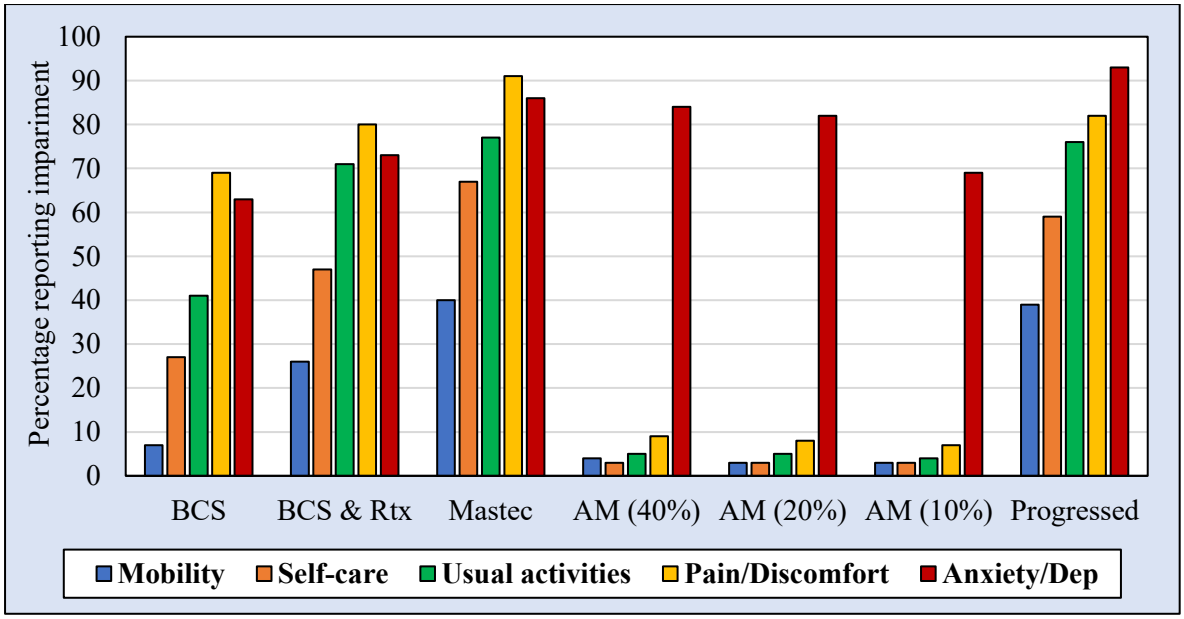


Figure 6.2: EQ-5D-5L Profile of participants estimating any level of impairment (%)

(Legend: BCS: Breast conserving surgery, Rtx: Radiotherapy, Mastec: Mastectomy, AM: Active monitoring, Progressed: Progressed DCIS, Dep: Depression)

6.3.3 Comparison of populations

Overall, both patients and non-patients had the strongest preference for active monitoring followed by breast conserving surgery alone, breast conserving surgery with radiotherapy, mastectomy and lastly treatment for progression. The only exception to this was for the VAS rating where breast conserving surgery alone was rated more favourably by patients (Mean VAS: 0.691, SD 0.19). As expected, the mean utilities of active monitoring increased in both groups as the risk of disease progression decreased from 40% to 10%. Utilities for most health states describing the management of low-risk DCIS were close to perfect health, as indicted by the upward skewed data, but more invasive treatments were associated with worse average quality of life (mastectomy: 0.454, breast conserving surgery with adjuvant radiotherapy: 0.575), particularly in the non-patient group.

6.3.4 Variation in preferences

Whilst there was no evidence to suggest that health states were ranked differently by individuals by the instrument used, there were differences in the grouped utilities derived between the two populations. Patient utilities were significantly higher than non-patient utilities for conventional treatments and lower for monitoring states using the VAS and standard gamble, but there were no group differences for any health state using the EQ-5D-5L (Table 6.5).

Table 6.5: Statistical comparison between patient and non-patient utilities (P values)

Health state	**P (VAS)**	**P (SG)**	**P (EQ-5D)**
A: Breast conserving surgery alone	0.017	0.005	0.260
B: Breast conserving surgery and radiotherapy	0.205	<0.001	0.925
C: Mastectomy (+/- reconstruction)	0.002	<0.001	0.221
D: Active Monitoring (40% risk of progression)	<0.001	0.121	0.138
E: Active Monitoring (20% risk of progression)	<0.001	0.123	0.149
F: Active Monitoring (10% risk of progression)	0.002	0.253	0.302
G: Disease progression	0.607	0.022	0.732

Paired testing of mean and median patient values found statistically significant differences between the overall utilities reported for monitoring and surgery (i.e. patients with a preference toward surgery reported lower values for active monitoring). Regression suggested that socio-demographic factors accounted for less than 5% of the variability, but a lower risk threshold for progression and history of invasive breast cancer did seem to significantly impact upon utility in favour of more aggressive treatment, regardless of the risk of overdiagnosis included. Reported task difficulty did not have a significant impact upon utility in either sample, but responses were highly heterogeneous, and the sample may not have sufficient power to detect differences in those who may not have understood the task.

6.3.5 Validity, reliability and feasibility of the approaches

The validity of the utilities derived was assessed in three parts. First, the economic measures applied used validated methods from the literature (178). Although the use of indirect methods alongside vignettes is less common, caution was applied in the construction of health states to ensure they reflected the treatment and not the dimensions of the EQ-5D (empirical validity). Second, descriptive validity was ensured through extensive piloting and refinement of the vignettes with clinicians and patient focus groups. Third, despite small differences in the utilities derived between different populations, the three methods demonstrated reasonable convergent validity (Appendix 8) and logical consistency in ranking across the seven health states among individuals (40% progression was ranked equal to or worse than the 10% risk by all women as expected).

Focus group feedback from the pilot study and interviewer judgement was used to appraise participant understanding of the vignettes and task reliability. Test-retest reliability among the six women participating in both the pilot and main study analyses were within expected level of agreement (0.2 SD) (148) and only 11 (6%) participants rated the tasks as “very difficult” on a Likert scale.

Finally, the study demonstrated reasonable feasibility due to the face-to-face nature of the interviews, despite some participants finding the task cognitively challenging. This was reflected in the lack of missing data, high completion rate and general consistency in the ranking between the three methods.

6.3.6 Comparison to utilities from the literature

Utilities for health states describing surgical or radio-therapeutic management of DCIS or recurrence in our study were lower than those previously elicited from the literature, where women may not have been informed that there was a chance the treatment may have been unnecessary. Figure 6.3 compares the mean utilities reported in three studies from the literature which had derived utilities for DCIS (148, 248, 254) versus those elicited in this study. For comparability, the utilities derived using standard gambles, vignettes and non-patients are shown. Although the methods used by Hayman (248) varied in the gamble anchor, the utilities in this empirical study (labelled as "Bromley 2019") were consistently lower than in others where the risk of overdiagnosis was not made explicit. Although other sampling or methodological differences might also have contributed to the variation in utilities between the studies, it is possible that the description of overdiagnosis in the vignettes may have contributed toward this at least in part.

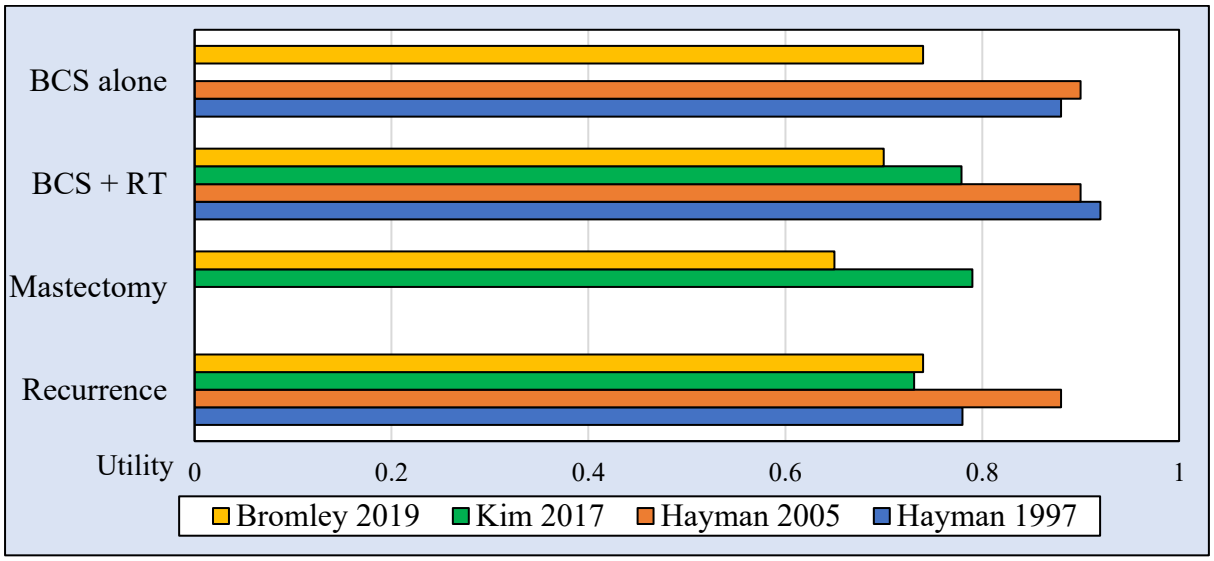


Figure 6.3: Comparison of standard gamble utilities (non-patient) for treatment

6.4 Can the disutility associated with overdiagnosis be estimated?

The disutility associated with the overdiagnosis of breast cancer has not been quantified in the economic literature, despite widespread acknowledgement that this harm exists (326). In this next section, this thesis will explore whether this harm may be isolated and estimated using the empirical data collected. It is hypothesised that if the harm associated with overdiagnosis arises from the potentially unnecessary treatment that ensues, then this harm might be captured by the difference in the disutilities for treatment with and without this risk.

Thus far this chapter has presented the utilities for the different health states associated with the treatment of low-risk DCIS. The findings suggest:

1. Most women report higher utilities for less intensive treatment (active monitoring) compared with standard treatment.
2. The probability of overdiagnosis or progression described correlate with the utility reported (i.e. vignettes with a higher risk are more likely to have lower utility).
3. Utilities in this study are lower than in other studies of DCIS using comparable populations and methods, and may be due in part to the explicit description of potential overdiagnosis in the vignettes.

The results suggest that the risk of overdiagnosis may be important in the assessment of benefits and harms (in terms of QALYs), although the utilities in this study may also have been affected by other methodological variables (e.g. country of origin, recruitment, interview technique etc.). Nonetheless, the consistency in the values elicited in the empirical study in both patients and non-patients using three different methods demonstrate that the harm from unnecessary treatment is likely to be important in the decision-making process.

The findings thus far have only detailed the utilities associated with the different treatments for low-risk DCIS, explicitly including the benefits and risks in the vignettes. Yet the harm from overdiagnosis itself corresponds the utility detriment, *or disutility,* incurred from undergoing a treatment which may have been unnecessary. This may be calculated as 1-utility for the corresponding health states (327). It would be naive to suggest that the utilities collected have completely captured the harm of overdiagnosis in isolation. Women can never truly know whether they are the subject of overdiagnosis and therefore can never really know whether they have been overtreated. DCIS is different to invasive cancer in that women must make decisions about treatment with imperfect information. Therefore, the real harm of overdiagnosis arises from the disutility associated with the additional loss in quality of life from the risk of unnecessary treatment.

Theoretically, this may be represented by the difference between how women value utility for treatment that might not have been needed (immediate surgery) and treatment that does not carry this risk (active monitoring). The interval between these two values thus embodies the disutility of unnecessary treatment (decision regret) for each individual (184).

The mean disutility of unnecessary treatment for a population may be calculated as:

Equation 1: $\sum(1-U_{ST}) - (1-U_{AM})/n$

Where n is the number of observations, U_{AM} is the utility of active monitoring (10%) and U_{ST} is the utility of standard treatment for each individual. Both the best valued surgical state (A-C) and usual treatment of BCS with radiotherapy (B) were substituted for U_{ST} in exploring

the methodological feasibility of this approach. The equation, $(1-U_{ST}) - (1-U_{AM})$, may also be rearranged to calculate the median and uncertainty around the estimates.

Consider one patient who reported VAS values of: 0.750 for BCS alone, 0.600 for usual treatment (BCS+RT) and 0.900 for AM (10%). The disutility of unnecessary treatment may be calculated as: $(1-U_{ST}) - (1-U_{AM})$ which is equal to: (1-0.750) – (1-0.900) = 0.150 if the best treatment state is considered, or (1-0.600) – (1-0.900) = 0.300 if conventional treatment with BCS with radiotherapy is used. This is then done for each woman in the population sample, summed and then divided by the total number of women in each group so that the results can be aggregated to a mean estimated disutility.

Table 6.6 summarises the overall disutilities calculated for each population and method using the best surgical and usual treatment utility values. Interestingly, patients using the VAS reported a gain in utility from undergoing immediate treatment compared to active monitoring if their preferred treatment option was chosen (e.g. BCS alone/mastectomy), or else they incurred no disutility if standard treatment (BCS with radiotherapy) was used as the comparator. In contrast, non-patients reported a mean VAS disutility of 0.130-0.200 from undergoing treatment that may not have been necessary compared with active monitoring.

Table 6.6: Disutilities associated with unnecessary treatment from overdiagnosis

Method	Patient Mean (SD)	Patient Median (IQR)	Non-patient Mean (SD)	Non-patient Median (IQR)
VAS (best surgical)	-0.123 (0.28)	-0.054 (0.35)	0.130 (0.21)	0.118 (0.20)
VAS (BCS + RT)	0.030 (0.30)	0.000 (0.33)	0.200 (0.23)	0.118 (0.27)
SG (best surgical)	0.037 (0.14)	0.000 (0.05)	0.148 (0.20)	0.199 (0.25)
SG (BCS + RT)	0.090 (0.18)	0.000 (0.15)	0.203 (0.22)	0.125 (0.35)
EQ-5D (best surgical)	0.042 (0.17)	0.071 (0.29)	0.040 (0.15)	0.042 (0.14)
EQ-5D (BCS + RT)	0.164 (0.19)	0.163 (0.22)	0.184 (0.21)	0.165 (0.25)

Most women reported losses in quality of life from having more invasive and potentially unnecessary treatment. This trend is supported in both patients and non-patients using the disutilities derived from the SG and EQ-5D-5L. There was a quantifiable harm from undergoing management for DCIS which carried the risk of overtreatment, whereby the estimated disutility might be as little as 0.037 or as high as 0.203. In general, the disutility was higher if active monitoring was compared against conventional treatment (BCS with radiotherapy), as some women reported smaller losses from less intensive treatment (BCS alone) or definitive management (mastectomy). This was also reflected at the individual level, where a few women who had undergone mastectomy or those who had reported higher breast cancer worry reported a gain in utility from having immediate surgery, regardless of whether it was ultimately necessary or not.

6.5 Discussion

6.5.1 Summary of main findings

This chapter has explored different methods to derive the utilities associated with DCIS, explicitly including the risk of overdiagnosis and unnecessary treatment in the valuation process. In general, mean and median utilities were greatest for health states minimising the risk of potential overtreatment (active monitoring). The exception to this was for breast conserving surgery, which was rated more favourably by patients using the VAS. Direct and indirect utilities varied in magnitude (standard gamble > EQ-5D > VAS), with the standard gamble yielding the highest values, consistent with the health economics literature (12). The EQ-5D demonstrated the most consistent ranking of the vignettes across populations. The disutility associated with unnecessary treatment derived from these values varied from -0.123 to 0.203 depending on the methods and sub-group.

Overall, the utilities in this study were lower than those produced in other studies of DCIS (148, 248, 254) or cancer surveillance (289). Whilst there is no single set of 'true' values for health states or profiles, it is important to consider the instrument, variant and perspective used to collect them (178). In this study, vignettes explicitly describing the benefits and harms for breast cancer treatment were used to elicit utility and to clearly incorporate the risk of overdiagnosis in the valuation process, and the results suggest this did impact upon women's preferences.

The standard gamble has been used in three other studies eliciting utilities for DCIS (148, 248, 254) and two studies that valued active monitoring (289, 290). Two (148, 248) used a modification of the traditional standard gamble, known as partial cascading, to avoid repeatedly having to confront participants with dead as the worst outcome, due to the low-risk nature of the disease. Instead, participants were asked to rank the vignettes on a visual analogue scale, and the lowest ranked health state was used as an intermediate anchor state instead of dead. The worst health state was then valued relative to dead and perfect health, as in a traditional standard gamble, to realign derived utilities onto a perfect health-dead (0-1) scale (16).

Whilst this adaptation does negate the potential ceiling effects associated with the conventional standard gamble (140), there are methodological issues observed in modifying the technique. Llewellyn-Thomas et al. (328) argue that replacing the anchors of the gamble with an intermediate health state can systematically bias utilities upward due to the re-framing effect of the gamble. A change in utility has been observed in other chained gambles relating to the anchoring and salience of the method-order of the intermediate health states substituted (329). A conventional standard gamble was used in this study as piloting

suggested a chained approach would add unnecessary cognitive complexity. The rank order of the worst health state varied in our patient sample, and would have led to the double counting of risks if monitoring health states had been substituted in place of 'dead' (e.g. 40% risk in AM scenario plus the added risk from the gamble) and utilities outside the 0 to 1 range, had a chained standard gamble approach been applied.

Interestingly, use of the EQ-5D instrument was found to be the most consistent of the three methods and able to differentiate the domains of health affected when treatment was de-escalated in the vignettes. However, there are concerns regarding the sensitivity of the 5L in oncological settings and the generalisability of the mapping conversion from the 3L using another population index (330). The EQ-5D is also best placed to value actual rather than hypothetical health, but it would have been unethical to ask women to value their own health in the context of overdiagnosis until the risks associated with disease progression are confirmed in clinical trials (27). It is unlikely that such a 'snapshot' of health would be able to capture the uncertainty in the decision surrounding overtreatment. Further research comparing vignettes with and without overdiagnosis or consecutive instrument valuation over time is necessary to confirm this.

The disutility of overdiagnosis and overtreatment is unlikely to be appropriately accounted for if unnecessary treatment is perceived to have the same impact as treatment that is required. How then should overdiagnosis be valued in terms of utility if not through vignettes? Monetary penalties have been applied in some economic models of mammography screening (52), but it is unclear as to what compensation would adequately offset the harm of unnecessary treatment. Furthermore, this approach does not negate the issue of valuing treatment conceptualised as being necessary in the QALY weights applied.

For many women, the harm of overdiagnosis is not perceived from having a diagnosis, as most women want to know that they have a condition so that they can take the appropriate action. Instead, the disutilities related to overdiagnosis are likely to arise from how the condition is managed after the condition is detected and whether the treatment received was appropriate (298). Ensuring information on benefits and risks are adequately communicated, ether in vignettes or clinical practice, is key to valuing this concept and ethically moral. This study addressed this issue by valuing health states which explicitly described both benefits and uncertainties involved in the decision, and quantified the disutility as the difference between utilities observed for treatment that may potentially be unnecessary and treatment which does not carry this risk.

6.5.2 Perspective of the population

It is assumed that patients will give a higher value for health states than either clinicians or the general population (331). The differences between the two samples in our study were only significant for the VAS surgery and SG monitoring health states. Values for active treatment (and overdiagnosis) were higher in patients than in women who had not undergone breast cancer treatment. Although participants were asked to value the health states described in the vignettes, it possible that the patient utilities were influenced by their own health. This could occur through adaptation or the desire to validate their treatment, which they may have been told was necessary by their doctors. This phenomenon was observed in another study valuing breast cancer health states (248), where patients rated hypothetical scenarios significantly higher than the general public, despite the same outcome of recurrence.

There was also considerable variation in individual responses within each group that could not be explained by socio-demographic characteristics. The acceptability of active monitoring

was highly dependent on risk trait or perceived anxiety about progression. Although on average a 10% risk of progression in 10 years was an acceptable trade-off to most, some women were willing to forgo the risk of possible side-effects associated with treatments for mastectomy or radiotherapy to reduce local recurrence. Consequently, individuals may place a higher value on a health state which they perceive as lifesaving or necessary, regardless of the risk of overdiagnosis presented, if they are risk averse. This raises an interesting question about whose utilities to use in economic evaluations, when individuals are given a choice of different treatment options or only certain people may choose each option. The information provided may change the preferred choice of women, as indicated by qualitative interviews of early breast cancer management which found that providing information on overdiagnosis to women changed stated preferences about treatment (299). Evaluations would ideally incorporate variation around estimates to account for different values among women.

The utility associated with DCIS management was high regardless of the treatment decision, and thus perhaps the focus ought to move from a single approach toward supporting informed patient decisions. The acceptability of de-escalation or active monitoring was highly dependent on the risk of progression and perceived anxiety. It is hard to predict the risk in an individual patient, and therefore treatment is given in case it is beneficial (101). This study shows that if the risk of progression is low, most women will choose monitoring over surgical treatment, however, if the risk is high that number is substantially reduced. Ultimately, the acceptability of active monitoring is dependent on defining these risks in clinical trials and communicating oncologic outcomes to patients and those valuing health states.

There are no published utilities for active monitoring in the breast cancer setting but the results are comparable to those elicited in prostate (289, 332) and cervical cancer screening

(290). Active monitoring strategies have been successfully implemented in other low risk cancer settings (333). The utilities derived in this study show theoretical acceptability of DCIS surveillance, and are encouraging information for clinicians and policy makers aiming to reduce the harms of overdiagnosis in the wider screening population.

6.5.3 Strengths and limitations

This main strength of this study is its methodological roots in exploring the feasibility of economic measures to quantify the harm of overdiagnosis, which may have resulted in lower utilities than those previously derived. Women were provided with health states which were explicit in detailing the benefits and harms within the vignettes of treatment, and the results may better reflect the value of treatment than those published to date. The methods were justified through a rigorous and systematic review of the literature and applied validated instruments or direct measures grounded in the axioms of utility theory. The methods adopted have been previously applied to derive utilities for uncertain health states in other relevant health conditions with risks, including other cancers identified through screening (281, 289).

Descriptive validity of the vignettes was carefully conducted to permit both direct and indirect approaches to be used concurrently. Vignettes were piloted on clinical experts, applied health researchers and patients to ensure they reflected the condition and not the dimensions of the EQ-5D. Overdiagnosis was explicitly listed in the description to ensure balanced information of potential benefits and harms was provided during the valuation process. This balanced description was of importance in addressing the limitations identified by a recent systematic review of breast screening valuation methods (284).

This is the first empirical study to value the utility of active monitoring in breast cancer. This is especially significant because it provides real-world data to suggest the acceptability of more conservative management of DCIS. Many women may find active monitoring a sensible alternative to surgery in terms of quality of life, which is likely to have significant implications for future health policy. However, there was considerable variability in individual preferences, and it is unclear what characteristics may be driving the different preferences between women.

It should be noted that there are also limitations to this exploratory research. The standard gamble can conflate the concept of health with risk aversion, and this ceiling effect was observed in part in the patient population for some health states, although the impact of risk was important in valuing uncertain health states. The time trade-off might have been cognitively easier, but its assumptions are violated by this uncertainty in risk (16). The standard gamble also assumed rationality but not all individuals may behave in a rational way when presented with a cancer diagnosis (334). Cancer treatment decisions may be emotively led where the perceived risk of recurrence supersedes other harms, whether the perceived risk reflects the truth or not. This phenomenon was displayed by some patients who preferred mastectomy over conservative management regardless of the risk of overdiagnosis or side-effects presented.

It is difficult to justify the exclusive use of one perspective for the allocation of healthcare resources. This study adopted a mixed approach to enable the methodological issues in valuing overdiagnosis to be explored from two alternative perspectives. There are limitations in the recruitment of women through breast cancer or screening groups who were likely to be more motivated and educated about breast cancer and overdiagnosis than the general

population. Similarly, the generalisability of the findings may not convey the true preferences of the entire population given the eligibility criteria specified. However, health state preferences are not fully informed if the participants do not have accurate expectations about the consequences of ill health (335), and one could argue in favour of the use of subjects with potential experience of overdiagnosis for exploratory methodological research.

6.5.4 Further research and policy implications

The utilities derived in this exploratory study suggest that the disutility of overdiagnosis does seem to be quantifiable and important to women diagnosed with breast cancer. The results imply that unnecessary treatment may have a negative impact on quality of life, but further research is needed to validate this finding, and determine what impact this might have for economic evaluations of breast cancer screening programmes.

When society knows more about the risks and benefits of de-escalated DCIS management from clinical trials (27), the true 'value' of the health states may change. Further empirical research to ascertain the effect of overdiagnosis on utility is necessary, by randomising the population to value vignettes both with and without the description of overdiagnosis, or by assessing how changes in the level of uncertainty reported in the vignettes corresponds to the unknown risks associated with breast cancer.

The results may also have an impact on guiding future clinical practice or treatment policy. Whilst clinical trials (e.g. LORIS) may provide much needed efficacy data on the value of de-escalating treatment (336), the potential policy implications of adopting such approaches is thus far unknown. The utilities from this empirical study can be used to inform the economic evaluation of screening or treatment programmes and estimated policy implications of

implementing a risk-stratified approach to managing DCIS. More research is necessary to investigate the potential oncologic outcomes of balancing treatment with monitoring and to ascertain what a surveillance protocol might look like prior to offering it to patients.

6.6 Conclusion

The utilities derived in this study show theoretical acceptability of DCIS monitoring, under certain conditions. The findings suggest that, when trading off benefits and risks in deciding about treatment, some women consider improved quality of life through conservative management and the avoidance of potentially unnecessary treatment to be as important as reducing local recurrence. Active monitoring of low-risk DCIS also appears to be an acceptable choice for reducing the harm of overdiagnosis in terms of utility. The magnitude of the disutilities for less invasive therapy suggest that active monitoring may be as cost-effective as standard treatment but further validation in an economic evaluation or clinical trial is required.

Overall, this chapter raises important methodological issues associated with valuing overdiagnosis and overtreatment, including conditions where the truth about benefits and risks may be commonly withheld from patients, or a patient's perception of their health is influenced by their doctors. Whilst some economic methods are limited in their ability to capture this risk, vignettes appear to be a feasible method of overcoming such limitations to capture the harm of overdiagnosis using both direct and indirect approaches. Further empirical testing is necessary to quantify and validate this. The following chapters will explore the application of the derived utilities in an economic model to ascertain the impact of overdiagnosis on treatment decisions. A qualitative analysis of the factors influencing patient utilities (and disutilities) from the empirical study is then presented in Chapter 9.

CHAPTER SUMMARY

What is known

- Utilities for seven health states describing the treatment of low-risk DCIS were elicited from patients and non-patients using the VAS, SG and EQ-5D-5L.
- Both benefits and harms (e.g. risk of overdiagnosis) were made explicit in the health state descriptions.
- Socio-demographics and a history of breast cancer were examined as predictors of utility.

What this chapter adds

- In general, both patient and non-patient utilities were more favourable toward less intensive treatments for low-risk DCIS (active monitoring).
- Values were highest for SG > EQ-5D-5L > VAS in general (consistent with the health economics literature), although the EQ-5D-5L produced the most consistent ranking of health states between individuals in both groups.
- Utilities for health states in which the risk of potential overtreatment was made explicit were lower than those from the literature, where women may not have been made aware that the treatment may have been unnecessary.
- The estimated disutility associated with overdiagnosis is quantifiable and varied from -0.123 to 0.203 dependent on the method and population used.

CHAPTER 7

A targeted systematic review of economic evaluations for the treatment of ductal carcinoma in situ

7.1 Introduction

This chapter presents a targeted systematic review of economic evaluations for the treatment of DCIS. The objective of the review was to identify the available modelling approaches in the literature to inform the design of an economic model in which the utilities from Chapters 5 and 6 could be applied. The chapter begins with an introduction of the types of models available for use in the economic evaluation of healthcare. The methods and results of the systematic review are presented next. Finally, the chapter concludes with a discussion of the most suitable approach identified by the review to apply in the design of an economic model representative of the LORIS trial.

7.2 Overview of modelling approaches

Economic evaluation is increasingly used to inform health policy decisions (4). As observed in Chapter 3, economic analyses may be conducted alongside a clinical trial or applied in a decision analytic model to predict the long-term outcomes and costs. Decision analytic models are used in health economics to consider the cost-effectiveness of alternative healthcare interventions beyond the direct application of observational data (337). Economic modelling is useful in the prediction of costs and benefits beyond the data observed in clinical trials, in extrapolating the results for an intervention from one clinical setting to another, and for informing resource allocation decisions in the absence of hard data (e.g. a pre-trial analysis of active monitoring for low-risk DCIS) (338).

The primary purpose of this chapter is to assess the evidence informing the cost-effectiveness of different interventions for DCIS treatment, with a specific focus on the model structures used and data inputs required. However, before presenting the methods and results of the systematic review, a brief overview of the main modelling approaches applied in economic evaluation are presented to ensure that the methods identified are appropriately defined and characterised.

7.2.1 A taxonomy of economic models

There exist several different models for undertaking economic evaluation in healthcare. According to Brennan et al. (339), economic models may be broadly classified into four main categories dependent on the relationship between the analytical approach and capacity for population interaction. These are: decision trees and Markov models, individual state-transition models, system dynamics models and discrete event simulation (Table 7.1).

Table 7.1: A taxonomy of economic models

	Cohort analysis	**Individual analysis**
No interaction allowed	Decision tree, Markov cohort	Individual sampling model
Interaction allowed	System dynamics model	Discrete event simulation

7.2.1.1 Decision trees

Decision trees are among the most widely used aggregated model due to their simplicity and familiarity (337). As illustrated in the example tree in Figure 7.1, all possible patient pathways and outcomes are shown with appropriate probabilities and output measures. The

chance of following any path is determined by the preceding node. Decision nodes (square) illustrate the choice of intervention (e.g. screening versus no screening) and chance nodes (round) represent the uncertainty of outcomes following the decision. Terminal nodes (triangles) represent the final outcomes. For example, in Figure 7.1 for an individual screened with cancer the upper branch for a true positive test is compared with the lower branch for a true negative test. The final cost-effectiveness for each strategy is then calculated by rolling back the expected pay outs from each tree, working from right to left. In general, decision trees are useful if the time frame is short and the intervention is simple, for example the decision to have a simple diagnostic test or not (7).

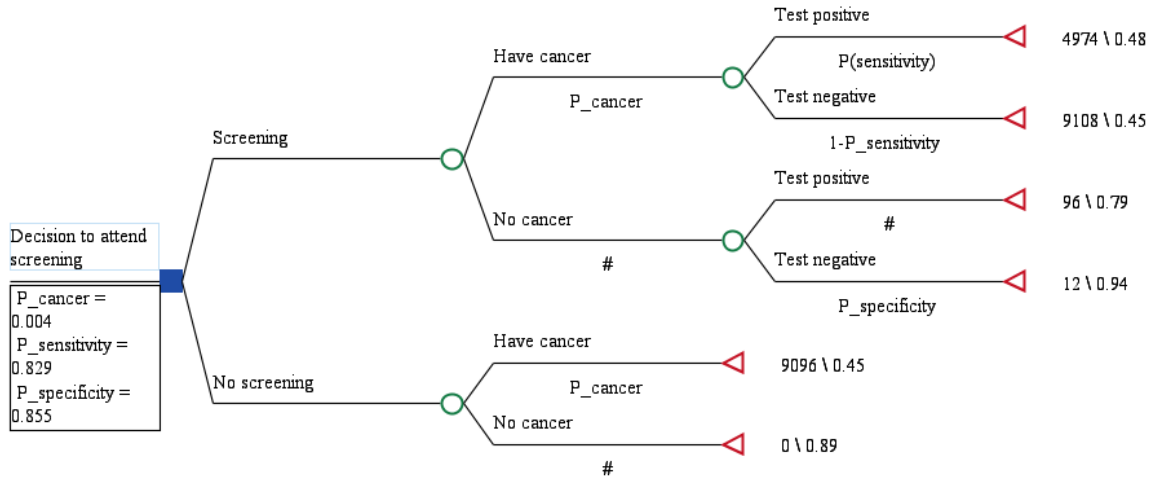


Figure 7.1: An example decision tree

7.2.1.2 Markov models

Markov models analyse uncertain decisions over time (7). Unlike decision trees, they allow for more complex health transitions and are suitable for decisions where the timing of events is important or a set of outcomes reoccur over time (340). Markov models are based on a series of 'health states' that an individual can occupy at a given point in time, with the probability of occupying the given state assessed over a series of discrete time periods or

cycles (4). The model is run for a specified number of cycles, and patients move between health states (defined by transition probabilities) or remain in their current health state, accruing costs and rewards as they progress through the model. An example of a Markov model for cancer treatment is shown in Figure 7.2.

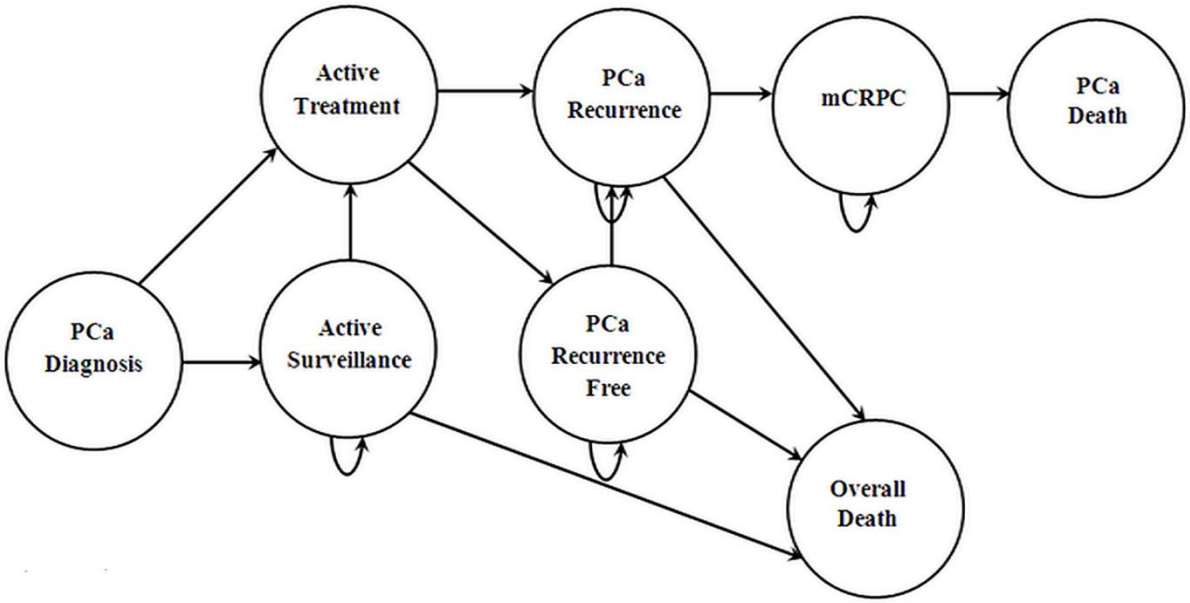


Figure 7.2: A Markov model for prostate cancer treatment (sourced from: (341))

The major benefit of the Markov model is that it allows for recurrent events and the analysis of costs and effects over a longer period of time (179). Such assumptions are useful in modelling chronic conditions (e.g. COPD) (342) and long-term cyclical events (e.g. menstrual bleeding) (343). However, they do not permit interaction between individuals, and future health transitions are limited by the current health state rather than prior events (6). Whilst such restrictions may be overcome through the addition of temporary or 'tunnel' states, they add significant complexity to the model and are unable to store patient history (4).

7.2.1.3 Individual sampling models

An alternative to Markov cohort analysis is to model individual patients as they progress through the model. Individual sampling models (ISMs) can incorporate memory into the model through microsimulation, permitting specific individuals and thus future probabilities to combine past events (337). For example, cancer microsimulation may use trackers so that transition probabilities (e.g. probability of developing metastases) may depend on previous interventions (adjuvant radiotherapy) or events (number of recurrences). Similarly, individual sampling allows individuals to progress independently of each other, whereby patients within the population may have potentially heterogenous characteristics that may affect their progression through the model (7). Such assumptions are helpful in conditions where past medical history or individual risk factors might influence disease progression or recurrence. However, ISMs do not allow for any interaction between individuals and thus may not be helpful in certain conditions such as infectious diseases. They are also more computationally and time demanding to construct and less amenable to sensitivity analysis (337).

7.2.1.4 System dynamic models

Dynamic transition models allow for some sort of interaction between individuals within the model, unlike the models discussed thus far. Consequently, they are particularly useful in the context of infectious disease modelling (344), where the incidence of new infections may depend on the existing number of individuals who are infected (stocks) and rates of disease transmission and recovery following interaction with others in the population (flow). Whilst system dynamic models allow for some aggregated individual interaction, they treat individuals as continuous variables so can only include a limited account of their past medical histories (337).

7.2.1.5 Discrete event simulation

In contrast to system dynamic models, discrete event simulation (DES) works at the individual patient level, permitting full representation of every individual's history and complex interaction between specific individuals in the model. DES can accommodate a larger number of individuals in each health state and variability over time by modelling outcomes as time to event (4). The advantage of DES is that it can overcome the limitations of individual sampling models in that more heterogeneous patient characteristics and histories may be modelled, but this requires specialist computational skills and processing power to run and may not always be worthwhile if the simulation makes no difference to cost-effectiveness results (337).

7.2.2 Economic models and DCIS

Decision modelling is an important vehicle for economic evaluation and identifying the appropriate type of model is an important aspect of the decision modelling process. Generally speaking simplicity in models is an advantage (4). However, the model structure is also dependent on the nature of the condition and impact of the interventions. Barton et al. (337) suggest that the key initial consideration is whether individuals are regarded as independent from one another. Where interaction between individuals is not important then the decision may be restricted to a decision tree, Markov model or ISM, and is dependent on the pathways and number of health states required. If individual-level modelling is necessary, then system dynamics or DES modelling may be more appropriate.

How then is DCIS modelled in economic evaluations? In Chapter 4 it was observed that overdiagnosis was not truly captured in economic evaluations of breast cancer screening programmes. The benefits and harms associated with treating low-risk breast disease,

specifically DCIS, were not adequately differentiated from that of invasive breast cancer. In order to apply the utilities derived in the empirical study which sought to address these limitations, an economic vehicle is required to demonstrate their use. A range of modelling approaches have been applied previously to evaluate breast cancer screening and treatment strategies including Markov models, individual sampling models and DES (187). However, there has been little research undertaken to explore the appropriateness of the economic models, data requirements and uncertainty in the economic evaluation of treatment for DCIS specifically.

7.3 Aim of the review

The aim of this targeted review was to examine economic evaluations of treatments for DCIS. In doing so, the objective was to identify the available methods and modelling approaches employed specifically for DCIS to inform an economic model of the LORIS trial and to act as a vehicle in which to insert the utilities elicited in the empirical study.

7.4 Methods

7.4.1 Search strategy

The targeted review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (199). Five electronic databases were searched from inception to June 2019; MEDLINE, EMBASE, CINAHL, PsycInfo and Econlit. Reference lists from the eligible studies were also checked for further potentially relevant studies.

The literature search strategy included index and free text terms relating to DCIS, economic evaluation and economic modelling. The full search strategy is provided in Appendix 9. The search utilised the terms applied in a Cochrane review of post-operative radiotherapy for

DCIS (345) and NICE appraisal guidelines on the identification of economic evaluations in the literature (131). Screening was conducted by the primary researcher (HB) and a second reviewer examined a sample of titles and abstracts. The results were managed using Endnote Reference Manager software X7 (202).

7.4.2 Eligibility criteria

Full economic evaluations were analysed if they met the following inclusion criteria:

- **Participants:** Adult women (aged 18 years and above) with DCIS.
- **Intervention:** Primary treatment for DCIS (including surgery, adjuvant radiotherapy, endocrine therapy and active monitoring).
- **Comparator:** No adjuvant therapy or standard treatment.
- **Outcome:** Cost per unit outcome, Cost per QALY.
- **Studies:** Cost-effectiveness, cost-utility, cost-benefit and cost-consequence analyses.

Only full text and English language studies were eligible for inclusion. Studies evaluating invasive breast cancer treatment, secondary in situ treatment or diagnostic interventions (e.g. biopsy techniques) were excluded.

7.4.3 Selection of studies for review

A two-stage process was applied to select studies for inclusion using the methods described by Mugford (346) and Roberts (347). First, each study was categorised based on the title and abstract using the eligibility criteria. The full paper was then obtained if the title and abstract were deemed potentially relevant or if no abstract was available. Studies were further categorised for selection by study type and relevance. The following criteria were adapted to assess the relevance of the studies for inclusion in the review.

Stage 1: Initial categorisation

A) The study reports primary research and includes a formal economic evaluation.
B) The study discusses economic aspects of healthcare and contains useful primary or secondary cost and utilisation data.
C) The study may have useful information but does not obviously fall into (A) or (B).
D) The study does not have any relevance to the economic evaluation of DCIS treatment.

Studies in categories (A) and (B) were considered potentially relevant to the review. A sub-sample (25%) of those in (C) were reviewed for relevance. Studies allocated to (D) were not reviewed any further.

Stage 2: Final categorisation

Studies classified as (A), (B) or (C) (if relevant) were further categorised after full text review into the following secondary categories:

1. Economic model or evaluation
2. Effectiveness study with some assessment of costs
3. Review article or abstract only
4. Not relevant

Studies classified as (A)1, (A)2 or (B)1 or (B)2 were included in the final review.

7.4.4 Data extraction and quality assessment

Data on the methods, characteristics and main findings were extracted and analysed narratively against the criteria set out in Table 7.2. The data extracted included: population, perspective, setting, intervention and comparator, model characteristics, analytical methods and the main results reported.

The quality of the included studies was assessed according to the criteria recommended by Philips et al. (348) because it is more comprehensive than other tools and is recommended for the assessment of model-based economic evaluations. Studies were not excluded from the analysis based on their quality, but the quality was considered in the synthesis of the results.

Table 7.2: Data extraction criteria

Criterion	Assessment
Setting (country)	*Are the results generalisable to the wider population? Are data sources appropriate for the country of study?*
Population	*Who are the population being evaluated in the study? Is this relevant to the demographic observed in DCIS?*
Perspective	*Healthcare or societal? Justification and appropriate inputs included for the perspective adopted?*
Interventions	*What treatments for DCIS treatment are being compared? What is the primary objective of the analysis?*
Type of economic evaluation and method	*What economic evaluation has been undertaken? Is this appropriate for the disease?*
Data sources	*Where is the data taken from for costs, utilities, transition probabilities and outcomes? Is the source suitable for the population? Is the extrapolation of data appropriate?*
Costs	*Was the most current year for the study used? Were all data inflated to the same price year? Discounting?*
Utilities	*Population and methods appropriate? Duration? Discounted? What health states are captured?*
Measure of benefit (effectiveness)	*Cost per natural unit or QALY? Is the outcome measure appropriate for the perspective of the evaluation? Are any additional clinical outcomes assessed in the evaluation?*
Model type	*Is the model type suitable for DCIS health states?*
Model structure	*How was this determined? Validated? Rationale provided?*
Time horizon	*Is the time horizon suitable for the intervention and disease?*
Cycle length	*Does the model cycle allow for costs and outcomes to be captured accurately?*
Model assumptions	*Are key assumptions explicitly stated and justified? Are they appropriate for the treatment pathway of DCIS?*
Sensitivity analysis	*Has a comprehensive uncertainty analysis been undertaken? One-way and probabilistic? Appropriate distributions?*
Results	*Which intervention was most cost-effective? How are the main findings presented against model strengths and limitations or those previously reported from the literature?*

7.5 Results

7.5.1 Study selection

The literature search identified 801 studies. Two additional studies were retrieved from searches of the reference lists. Of these, 50 studies were considered potentially relevant after the removal of duplicates and initial categorisation and were reviewed in full. 47 studies were excluded at full text review (34 were irrelevant, four were reviews and nine reported only abstracts). Three studies fulfilled all criteria and were included in the final review. The studies included and excluded at each stage of the selection process is detailed in Figure 7.3.

7.5.2 Characteristics of the included studies

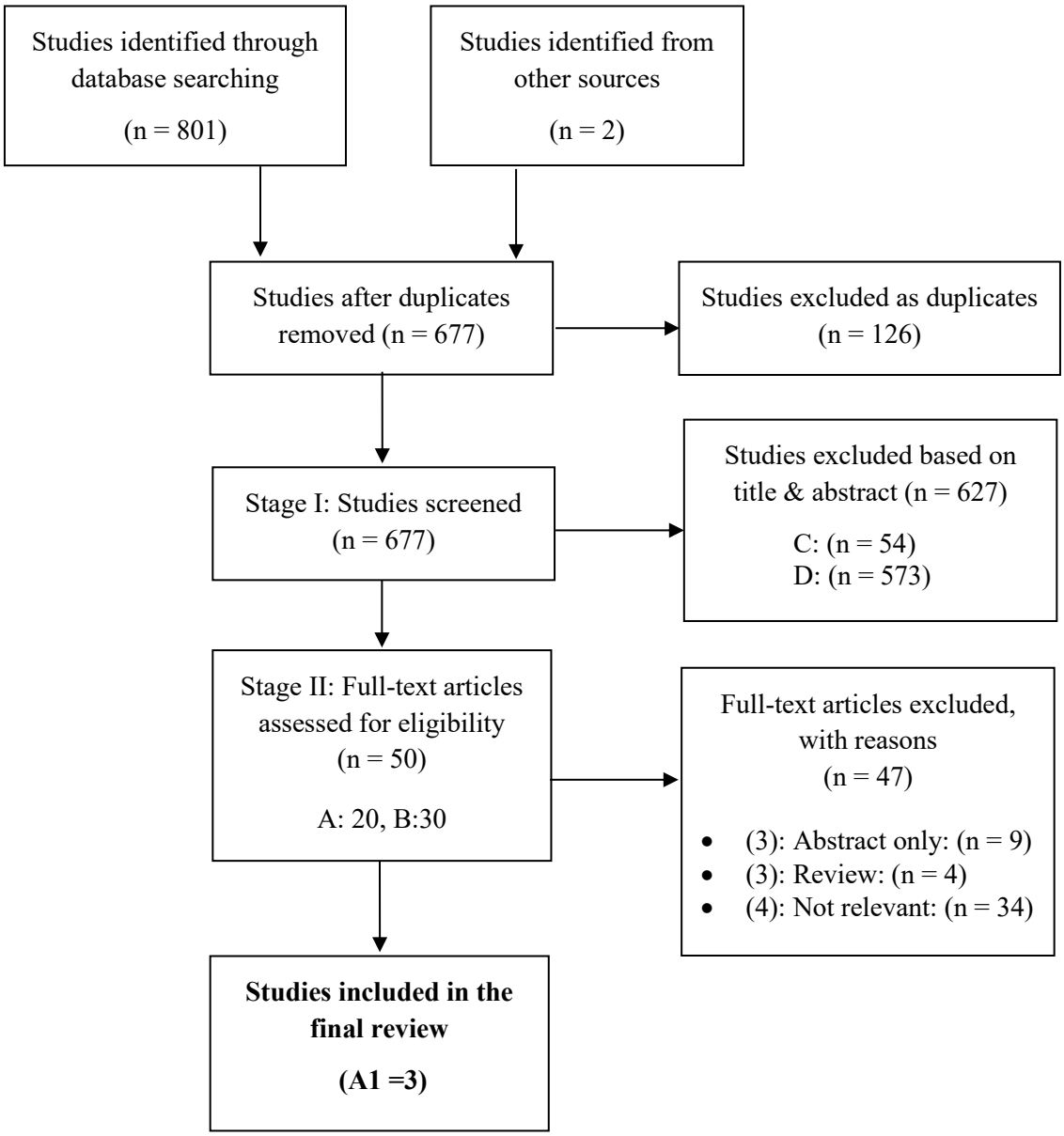
Table 7.3 summarises the main characteristics of the three studies (349-351) identified by the review. All three studies evaluated treatments for women of screening age diagnosed with DCIS in the USA. Two studies (350, 351) explicitly adopted a healthcare perspective in keeping with the health insurance system evaluated, whereas Suh et al. (349) explicitly applied a societal approach to capture the extra costs incurred by breast cancer patients (for transport expenses and missed work) related to radiotherapy.

Most studies evaluated breast conserving surgery with and without radiotherapy, as is standard treatment for DCIS. Trentham-Dietz et al. (351) modelled a hypothetical prognostic marker to compare standard treatment against active surveillance, but the study was not reflective of clinical practice or ongoing trials of active monitoring (27) due to an assumed 100% test accuracy in identifying non-progressive disease. Overdiagnosis (and overtreatment) was not explicitly mentioned in any of the evaluations. However, there was some consideration of de-escalating adjuvant therapy for patients with low-intermediate grade DCIS compared to high-grade disease in two (349, 351) of the three studies.

Table 7.3: Characteristics of the economic studies evaluating treatments for DCIS

Study	Raldow (2016) (350)	Suh (2005) (349)	Trentham-Dietz (2018) (351)
Population	Women aged 60 years with low-intermediate or high-grade DCIS	Women aged 55 years with DCIS	Women aged 40-74 years with non-progressive DCIS
Setting	USA	USA	USA
Perspective	Healthcare service	Societal	Healthcare service
Interventions	BCS +/- prognostic test +/- radiotherapy	BCS +/- adjuvant radiotherapy	Hypothetical prognostic test and either standard treatment or AM
Type of EE	Cost utility analysis	Cost utility analysis	Cost utility analysis
Data sources	RCT, empirical utility study, Medicare database	National database, RCT, empirical utility study	National database, RCT, national age-specific utility data
Costs	2013 USD (discounted 3%)	2002 USD (discounted 3%)	2015 USD (discounted 3%)
Utilities	SG utilities (patient) (248)	SG utilities (patient and general population) (248)	EQ-5D utilities (general population adjusted using assumptions (218))
Measure of benefit	Cost per QALY, ICER, cost per IBC prevented (discounted 3%)	Cost per QALY, ICER (discounted 3%)	Cost per QALY, Cost per LYG, BC deaths (discounted 3%)
Model type	Markov model	Markov model	Discrete event simulation
Model structure	Microsimulation (appropriate health states)	Cohort (appropriate health states)	Microsimulation (structure suitable for disease pathway)
Time horizon	10 years	Lifetime	Lifetime
Cycle length	1 year	1 year	N/A
Model assumptions	For recurrence after BCS: 50% treated with BCS+RT and 50% mastectomy. For recurrence after radiotherapy, 100% undergo mastectomy.	Recurrence salvaged by mastectomy (if BCS + RT) or BCS + RT/ mastectomy (if BCS alone). Metastases treated with chemotherapy	Test identifies DCIS as progressive or non-progressive. Patients have 100% adherence with treatment.
Sensitivity analysis	One-way, two-way	One-way, two-way	One-way
Results	No strategy was cost-effective, but test did reduce overtreatment	BCS + RT cost-effective (ICER in 2002 USD $36,700)	Prognostic test and AM dominant (0.04 extra QALYs, $5,316 fewer costs)

Figure 7.3: PRISMA flow diagram of the study selection process



7.5.3 Modelling approach

All the studies identified in the review used an analytical model to assess cost-effectiveness, either to extrapolate outcomes to a lifelong time horizon or to evaluate an approach not yet validated in clinical trials for DCIS. Both Raldow et al. (350) and Suh et al. (349) used Markov models (analysed by individual patient simulation and cohort analysis respectively)

to capture health states relevant to DCIS treatment, recurrence and death. Markov models were deemed appropriate for modelling DCIS treatment because they allowed for recurrent health states relating to events after treatment such as ipsilateral breast recurrence, which is the primary outcome in most trials of DCIS, and the long-term assessment of outcomes. In contrast, Trentham-Dietz et al. (351) used discrete event simulation to approximate the complex disease pathways associated with the condition. However, whilst this did allow for additional patient variables to be captured in the model (e.g. ER/HER2 status), it was more computationally demanding and reliant on progression time assumptions because of the unknown natural history of the condition and hypothetical treatment explored.

There was notable variety in the time horizons applied across the studies. Whilst two studies (349, 351) adopted a lifelong extrapolation of data to account for the continuous risk of disease progression, Raldow et al. (350) explicitly used a 10-year time horizon because of the limited clinical trial data and follow-up for the different grades of the disease. There are advantages and disadvantages associated with each of the time horizons used. A 10-year horizon is in keeping with the time frame used in clinical trials of DCIS, allowing more accurate modelling of the likely progression of women as they move through each treatment pathway. However, given that the risk of recurrence continues for the remainder of each woman's lifetime, this time horizon might not capture all relevant costs and effects incurred by women with DCIS. A longer time frame may be more suitable, particularly for modelling women with increased risk of recurrence if allocated to less intensive treatment or surveillance. In contrast, there was at least consistency in the cycle length applied in the three studies (1 year), as per the recommended follow-up of women after treatment and annual mammography modelled.

7.5.4 Type of economic evaluation

All three studies performed cost-utility analyses, with a focus on quality of life and health over clinical outcomes. The impact of treatment intensity on costs and quality of life, where breast-cancer-specific survival for DCIS is favourable regardless of treatment, was felt to be more meaningful in the analysis than an alternative method of economic evaluation such as cost-effectiveness where life years gained (LYG) might be minimal. Two studies (350, 351) did report another measure of effectiveness alongside QALYs in their analysis (cost per IBC prevented, cost per LYG) but it did not significantly change the results.

7.5.5 Utilities

Utilities informing the reported QALYs were obtained from two sources (218, 248). Trentham-Dietz et al. (351) applied adjustments to US female population EQ-5D utilities for stage-specific breast cancer treatment as previously described in an economic evaluation (218) of breast cancer screening programmes. However, it is unclear whether all relevant benefits and harms were captured in terms of quality of life; the adjustments were based on assumption and no further disutility was applied to women in the prognostic test or active surveillance health states in their model. The remaining two studies (349, 350) applied the general population standard gamble utilities from Hayman et al. (248) which were explicitly derived for health states describing DCIS treatment and recurrence. Suh et al. (349) also used a second set of SG utilities (taken from the same study but elicited from patients with DCIS) to ascertain the impact of the population preference in a sensitivity analysis. None of the studies explicitly considered the disutility of overtreatment, applying the same utilities for treatment regardless of outcome.

7.5.6 Effectiveness measures

Model inputs for the transition probabilities were largely sourced from US randomised controlled trials of DCIS therapy (352, 353) or national cancer registries (354). Whilst the data reported were of high quality, follow-up in the trials was restricted to 20 years and longer-term outcomes applicable to the model had to be assumed. Similarly, data for the accuracy of prognostic markers and disease progression under surveillance were based on assumption due to a lack of published estimates.

Costs and health resource use were generally more homogeneous and primarily sourced from the Medicare literature or other economic models (218, 355). Suh et al. (349) included indirect costs (time and transport) in their model related to the societal perspective of the analysis. All costs and benefits were discounted at 3%, as is standard convention in the US system on which the models were based (356).

7.5.7 Sensitivity analysis and quality assessment

In general, all three studies adequately met most of the quality appraisal criteria outlined in Appendix 10. The type of economic evaluation, model structure, intervention and comparator were consistently reported, and reasonable conclusions were drawn from the main findings. Trentham-Dietz et al. (351) did not actually report the ICER for standard treatment and active monitoring in their results, but this could be calculated from the data presented elsewhere in the paper.

All studies conducted at least a one-way sensitivity analysis to assess the uncertainty of the results. In a one-way sensitivity analysis, model parameters are varied individually to determine the impact of each variable on overall cost-effectiveness (7). Probability of

recurrence and the utilities, costs and effectiveness of treatment were identified as being key drivers of cost-effectiveness in managing DCIS. However, for a robust analysis of uncertainty it is necessary to perform a probabilistic sensitivity analysis where several model parameters are changed simultaneously using distributions (348). None of the studies conducted a probabilistic sensitivity analysis in their analysis. Given the widespread uncertainty (28) and debate surrounding the natural history of disease, recurrence risk, and implications of surveillance on costs and quality of life, it was unclear as to why an extensive sensitivity analysis had not been undertaken in any study.

7.6 Discussion

7.6.1 Summary of the main findings from the review

This targeted systematic review has highlighted the paucity of economic evaluations for the treatment of DCIS in the literature. The findings corroborate previous concerns (3, 31) about the lack of economic evidence in informing the debate on screening and the management of low-risk breast disease. Most studies looked at the cost-effectiveness of standard treatment (BCS +/- adjuvant radiotherapy), with only one study (351) considering the potential implications of active monitoring in DCIS. However, the hypothetical marker used in the analysis to stratify women to surveillance assumed 100% test sensitivity which does not reflect the ongoing clinical trials of active monitoring reported in the literature (27, 106, 357). There were mixed results on whether the de-escalation of adjuvant treatment was cost-effective. However, the analyses in the studies identified by this review suggest that the distinction between low-intermediate and high-grade DCIS is important in terms of results (i.e. less intensive treatment was generally cost-effective and clinically feasible in economic analysis of low-intermediate grade DCIS but not for high-grade DCIS).

The limited number of economic evaluations for DCIS treatment is likely related in part to how DCIS has been conceptualised in the literature. Most studies (26, 57, 219) have analysed the cost-effectiveness of interventions for DCIS within economic models of breast cancer screening programmes because it is primarily a condition of screening. Yet it may not be appropriate to exclusively consider DCIS alongside invasive breast cancer in screening if the potential for overtreatment may significantly differ (358). Whilst the inclusion of DCIS in the economic evaluation of breast cancer screening is evidently important given the debate about the benefits and harms of overdiagnosis (3), it is difficult to differentiate the outcomes related to DCIS in isolation from those of early breast cancer when the results of the two conditions are analysed concurrently. For example, the CISNET group of US breast cancer screening models evaluated DCIS extensively in their analyses (359), yet only reported on outcomes related to screening and did not present a sub-group analysis for women diagnosed with low-risk disease.

There are limited data on the cost-effectiveness of interventions for DCIS, thereby making it challenging for policy makers to inform the decision on the most appropriate treatment. Given the clinical prognosis of DCIS is favourable regardless of treatment (25), it is feasible to assume that cost-effectiveness is likely to be particularly important in determining the optimal treatment policy. There is little data on active monitoring or consideration of women who may be overtreated in the current analyses of DCIS management, despite initiatives toward the safe de-escalation of breast cancer treatment in recent years (336). A standalone appraisal of the cost-effectiveness of alternative DCIS management outside of screening is necessary to address this gap and guide policy decisions for those diagnosed with low-risk disease at mammography.

7.6.2 Strengths and limitations

The major strength of this systematic review is the comprehensive review of the literature undertaken to identify economic evaluations of DCIS interventions. The quality of the studies was appraised using the validated Philips checklist (348) and the review identified a clear gap in the literature associated with DCIS management. The results provide suggestions for future cost-effectiveness analyses of DCIS, which to current knowledge has not been done specifically in economic evaluations of DCIS treatment before. However, the search strategy was limited to English language and full text studies only and may have missed other potentially relevant evaluations not identified by the review. Second, it was difficult to identify economic evaluations exclusively assessing DCIS treatment outside of other interventions such as screening and diagnostic tests and thus some studies may not have been identified appropriately by the search. Third, this was a targeted systematic review to identify and replicate a suitable economic model in which to insert the utilities derived in earlier chapters and thus focused on strategies for in situ breast cancer rather than other cancers.

7.6.3 Comparison with other studies

This review highlights an important literature gap relating to the lack of evidence on the cost-effectiveness of interventions for DCIS. Previous studies of DCIS treatment (101, 360, 361) have consolidated only the clinical efficacy of treatment and have not included the cost-effectiveness of strategies in the presentation of results. Similarly, other systematic reviews (55, 186, 187, 362) of economic evaluations related to DCIS have assessed the cost-effectiveness of breast cancer screening programmes or invasive breast cancer treatment but not DCIS specifically.

There are very few studies evaluating active monitoring in the breast cancer setting, which may be expected given that clinical trials, such as LORIS (27), are still ongoing and the findings are yet to be published. One epidemiological study (363) has modelled active monitoring versus standard treatment but did not include any analysis of costs and quality of life. Economic evaluations of active monitoring using a cost per QALY approach have been conducted in the wider oncological setting (e.g. prostate cancer surveillance) and could potentially be applied in the design of an economic model the for the evaluation of active monitoring for breast cancer (333, 364, 365), but was beyond the scope of this targeted review, the purpose of which was to identify models relevant to DCIS specifically.

7.6.4 Implications for how the review informs economic models

There are five important implications raised by this systematic review to consider in the planning of future economic evaluations of DCIS:

1) A Markov model or DES *may* be a suitable approach for modelling the cost-effectiveness of treatments for women with DCIS, based on the structures adopted in the three studies identified in this limited scoping review and the taxonomy of models suggested by Barton et al. (337) earlier in this chapter for modelling chronic conditions with recurrent health states, but further validation is necessary through exploration of the models used in the economic evaluation of other low-risk screen detected cancers.
2) Interventions for DCIS should be compared against a clear comparator (e.g. standard treatment of BCS with radiotherapy) to enable the consistent interpretation of results.
3) Analyses should be presented by DCIS grade (low, intermediate and high-grade) as this is likely to have a considerable impact upon cost-effectiveness results.

4) Results should be presented as a cost-utility analysis. Quality of life is important given that 10-year breast cancer specific survival between surgically and non-surgically managed low-risk DCIS is the same (25). Utilities informing QALYs and cost-utility analyses should capture the benefits and risks appropriately, including the potential for overtreatment.
5) An extensive sensitivity analysis (including PSA) should be conducted given the uncertainty in the unknown natural history of the disease and variable probability of progression.

7.6.5 Further research

The paucity of studies identified by the review demonstrates the need for further research on the economic evaluation of DCIS treatment. The aim was to conduct a targeted review and critique the evidence on the cost-effectiveness of in situ treatment to inform the design of an economic model in which to apply the empirical study utilities from Chapters 5 and 6. However, due to the lack of economic models identified it is difficult to ascertain the most appropriate health states for the evaluation of unnecessary treatment. A rapid review to look at how active monitoring has been modelled in other low-risk cancers detected by screening is recommended to confirm the most suitable approach for modelling the LORIS trial.

7.7 Conclusion

This chapter has described in detail the different methodological modelling approaches that can be used in the economic evaluation of healthcare interventions. A systematic review of economic evaluations in DCIS treatment was then reported to identify a suitable approach in which to apply the utilities from Chapters 5 and 6. There were very little data on the cost-effectiveness of DCIS treatment outside of the economic evaluation of breast cancer

screening programmes. Only three studies describing an economic evaluation of DCIS treatment were identified. The findings from this chapter suggest that a Markov (individual sampling model) or DES may be most appropriate for undertaking a cost-utility analysis of DCIS treatment, but further adjustment is likely required to accommodate the uncertainties surrounding active monitoring and low-risk DCIS specifically. In the next chapter, a targeted review of active monitoring models is presented, alongside the development and results of an economic model of treatments for low-risk DCIS.

CHAPTER SUMMARY

What is known

- Decision analytic modelling is useful in the prediction of costs and outcomes beyond the time horizon of clinical trials, for extrapolating the results of an intervention from one clinical setting to another and for informing resource decisions in the absence of hard data.
- Economic models may be classified into a taxonomy of decision trees or Markov models, individual sampling models, system dynamics models or DES.
- A targeted systematic review was necessary to ascertain what model was best suited for the economic evaluation of DCIS treatment.

What this chapter adds

- The paucity of studies identified by the review demonstrates the need for further research on the economic evaluation of DCIS treatment.
- From the limited evidence available, a Markov model, ISM or DES may be suitable for modelling the long-term progression of women with DCIS.
- Sub-group analysis analysis should be conducted for low-intermediate and high-grade DCIS as it is likely to have a considerable impact upon cost-effectiveness.
- An extensive PSA is required given the widespread uncertainty in the natural history of the disease and unknown probability of invasive progression.

CHAPTER 8

An economic model comparing active monitoring with standard treatment for low-risk ductal carcinoma in situ

8.1 Introduction

This chapter describes the rationale and methodology for developing an economic model to simulate the costs and benefits associated with the treatment of low-risk ductal carcinoma in situ. Two treatment strategies are compared: active monitoring and standard treatment with surgery and radiotherapy. The aim of the model is solely to illustrate how the utilities derived in the empirical study might be used in an economic evaluation, using an example model to simulate outcomes for a hypothetical cohort of women with localised, low-risk DCIS, as described in the LORIS trial.

The chapter begins with an overview of active monitoring models in the cancer literature. This is followed by a summary of the DCIS model structure, health states and parameters. The results are then presented in terms of cost per quality-adjusted-life-year (QALY), comparing findings using the utilities from Chapter 6 with those from the literature. Finally, both deterministic and probabilistic sensitivity analyses are run to compare the significance of the two results, alongside a discussion of the strengths, limitations and potential implications for breast cancer screening programmes.

8.2 Background

8.2.1 Modelling DCIS and active monitoring

The incidence of DCIS has risen substantially since the introduction of breast cancer screening programmes (366). Despite the increasing number of women affected the risk of

invasive progression and optimal management remain uncertain, leading to widespread variation in patient treatment and decision making (360). Standard treatment consists of breast conserving surgery, radiotherapy and endocrine therapy – treatment comparable to that recommended for invasive breast cancer. There is concern that women with low-risk DCIS may therefore be overtreated (3), incurring potentially unnecessary extra costs and morbidity.

Active monitoring with annual mammography is proposed as an alternative to surgery to reduce the harm of overtreatment (27). Clinical trials such as LORIS are underway to explore the clinical feasibility of active monitoring. Whilst costs and utilities have been examined in the models of active monitoring of some cancers, they have not yet been researched in the breast cancer setting. Decision analytic models are economic tools which can estimate the potential impact of new interventions until such data are available (7).

In Chapter 7, a targeted literature review was undertaken of economic evaluations for the treatment of DCIS. Whilst several economic models were identified by the review, there were few evaluations of active monitoring in the breast cancer setting. Consequently, a rapid literature review of economic models of active monitoring in other screen-detected cancers was undertaken to inform the design of the in-situ model structure.

8.2.2 Literature search

Briefly, a rapid literature review was undertaken of five electronic bibliographic databases (MEDLINE, EMBASE, Web of Science, PsycInfo, EconLit) to identify any economic evaluations or economic models of active monitoring in low-risk cancers identified by screening (e.g. prostate, colorectal and cervical cancer). Searches were run from the earliest date available to 1 July 2019, using a search strategy of subject heading and key words

including: (prostate cancer OR cervical cancer OR colorectal cancer OR breast cancer) AND (economic evaluation OR cost-effectiveness analysis OR cost utility analysis OR (cost and QALY)) AND (active monitoring OR active surveillance OR watchful waiting).

The search strategy identified a total of 147 potentially relevant studies, of which 49 were removed as duplicates. Two additional studies were identified from searches of the reference lists. 69 studies were excluded after reading the title and abstract (no economic evaluation or model) and a further 21 were excluded on full text review (conference abstract, review article or irrelevant), leaving 10 studies in the final analysis.

All ten studies had applied an economic model of active monitoring for an analogous low-risk cancer. The characteristics of the economic models are summarised in Table 8.1. All 10 studies used a state-transition microsimulation model and modelled active surveillance versus immediate treatment for low-risk prostate cancer.

The economic models simulated men of screening age (50-69 years) with localised, low-risk prostate cancer managed via active surveillance (defined as 6-12 monthly blood tests, clinical examination, imaging +/- biopsy) or immediate radical treatment. Follow-up varied widely and there was noticeable heterogeneity in the time horizon applied across the studies. Six studies (60%) used a lifetime horizon to reflect the indolent nature of the condition, but others used shorter simulations of 10 or 20 years due to the paucity in long term data published in clinical trials. A similar decision must be made in modelling in low-risk DCIS between the paucity of long-term data and lifelong propensity for invasive progression.

Table 8.1: A summary of economic models for active monitoring identified by the literature review (n = 10)

Study	Year	Country	Population	Intervention	Model type	Analytic approach	Model cycle	Time horizon	Uncertainty analysis	Outcome
De Carvalho (367)	2018	USA	Prostate cancer; Men aged 55-69	AM vs. immediate radical treatment	State-transition model	Micro-simulation	1 year	Lifetime	1-way, multi-way & deterministic	Cost/QALY, LYG, ICER
Hayes (333)	2013	USA	Prostate cancer; Men aged 65 & 75	AM vs. initial treatment (WW, BT, IMRT, RP)	State-transition model	Micro-simulation	1 year	Lifetime	1-way, Multi-way and PSA	Cost/QALY, ICER
Hayes (273)	2010	USA	Prostate cancer; Men aged 65	AM vs. initial treatment (BT, IMRT, RP)	State-transition model	Micro-simulation	1 year	Lifetime	1-way, Multi-way and PSA	QALY
Keegan (368)	2012	USA	Prostate cancer; Men (age unspecified)	AM vs. immediate treatment (ADT, BT, IMRT, RP)	State-transition model	Micro-simulation	1 year	10 years	1-way	Costs
Keller (278)	2017	Australia	Prostate cancer; Men aged 50-69	Screening with AM vs. screening with immediate treatment	State-transition model	Micro-simulation	1 year	20 years	1-way & PSA	Cost/QALY, LYG, ICER
Lao (365)	2017	New Zealand	Prostate cancer; Men aged 45-70	AM vs. WW vs. RP	State-transition model	Micro-simulation	1 year	Lifetime	1-way & Deterministic	Cost/QALY, ICER
Loeb (364)	2017	USA	Prostate cancer; Men aged 50	AM vs. WW	State-transition model	Micro-simulation	1 month	Lifetime	1-way & 2-way	Cost/QALY, LYG, ICER
Sanyal (341)	2016	Canada	Prostate cancer; Men aged 65	AM vs. initial treatment (ADT, BT, IMRT, RP)	State-transition model	Micro-simulation	1 year	15 years	1-way	Cost/QALY, ICER
Sathianathen (369)	2019	USA	Prostate cancer; Men aged 50	AM vs. WW vs. initial treatment	State-transition model	Micro-simulation	6 months	Lifetime	Deterministic & PSA	Cost/QALY, ICER
Sharma (370)	2019	USA	Prostate cancer; Men aged 50-69	AM vs. RT vs. RP	State-transition model	Micro-simulation	1 year	10 years	1-way & multi-way	Cost/QALY, ICER

ADT: Androgen deprivation therapy, AM: Active monitoring, BT: Brachytherapy, ICER: Incremental cost-effectiveness ratio, IMRT: Intensity-modulated radiation therapy, LY: Life years, NA: QALY: Quality-adjusted-life-year, RP: Radical prostatectomy, RT: Radiotherapy (unspecified), USA: United States America, WW: Watchful waiting

It is clear from the summary of the models that a state-transition microsimulation is the best approach for modelling active monitoring. Individual simulation modelling (ISM) facilitates the movement of patients through different strategies of conservative management and permits recurrent health states and uncertainty to be captured in invasive progression. All ten studies adopted a Markov model (analysed via microsimulation) to model a hypothetical cohort of patients with low-risk disease in predefined health states from the point of diagnosis (at screening) to treatment and post-treatment events. During each cycle patients could either remain in a low-risk cancer state under monitoring (no progression), undergo treatment for progression, remain disease-free after treatment, develop a recurrence or metastasis or die. The fundamentals of the surveillance pathway are comparable to those outlined in the LORIS trial protocol. This structure thus offers a suitable framework for application in DCIS.

All but one study used cost per QALY as the main outcome measure. QALYs were informed by utilities derived from first principles (standard gamble) for active surveillance reported in the literature. This suggests that the utilities derived in the empirical study can be applied in a similar economic vehicle for active monitoring in low-risk DCIS.

From the models of active monitoring identified in the rapid review, the model deemed best placed to represent the evaluation of women with low-risk DCIS is a state-transition model analysed via microsimulation. This is consistent with the findings from the review of DCIS treatment recommended in Chapter 7. Although breast and prostate cancer vary in terms of pathophysiology, the outcomes and health states related to low-

risk and progressed disease under monitoring are likely to be comparable to those anticipated in the LORIS trial. The model structures identified in the review thus illustrate a suitable framework on which to build the active monitoring simulation and health states specific to low-risk DCIS.

8.3 Aim and scope of the model

As discussed in Chapter 7, the primary purpose of this chapter is to use an economic model to illustrate how the elicited utilities for low-risk DCIS may be applied in practice. The aim of the model is not to ascertain the true cost-effectiveness of active monitoring (such estimates are reliant on the availability of clinical efficacy and resource data from clinical trials), but to illustrate the potential impact of including the disutility of overdiagnosis in economic evaluations of breast cancer.

The objectives of the model are to:

1) Illustrate how the utilities derived in the empirical study might be applied in economic evaluation
2) Estimate the benefits and costs of active monitoring compared with standard treatment for women with low-risk DCIS
3) Compare approximate cost-effectiveness outputs using the utilities from the empirical study (explicitly valuing overdiagnosis) versus those from the literature (without overdiagnosis) to ascertain whether this may potentially impact cost-effectiveness
4) Use the findings to ascertain the potential implications within the wider context of breast cancer screening and treatment programmes

8.4 Methods

All methods and results are reported against the CHEERS (371) and Philips (348) checklists, the details of which are provided in Appendix 11, as is good practice in economic evaluation.

8.4.1 Type of evaluation

A cost-utility analysis was used to analyse standard treatment versus active monitoring so that the impact of each strategy could be determined using a conventional cost per QALY approach. This allowed for the application of the utilities derived in earlier in the thesis (Chapter 6) to be included in the analysis of outcomes.

8.4.2 Model structure

A Markov state transition model was designed to simulate the outcomes of a hypothetical cohort of women with clinically localised, low risk DCIS in the United Kingdom. This design was chosen to allow for the multiple and recurrent health states that occur during the management of DCIS and was informed by the literature review. A decision tree is inappropriate because it would require a significant number of branches to allow for recurrent health states (337). Conversely, a discrete time simulation would be unnecessarily complex given the objectives of the model as an illustrative vehicle (339).

The model structure was informed by the LORIS trial protocol (27) and compares the potential outcomes for women undergoing standard treatment and active monitoring. The health states and clinical pathways were developed with specialist input from

surgeons, health economists and epidemiologists with experience of modelling breast cancer screening and treatment.

8.4.3 Model development

A Markov state transition model was constructed using decision analytic software (TreeAge Pro Version 2018) (372) and used Monte Carlo simulation to estimate the benefits and harms associated with the management of women with newly diagnosed low-risk DCIS. Individual microsimulation was chosen over cohort analysis to allow for the different probabilities of recurrence to be linked against treatment and recurrence history.

The model structure was built from the pathways outlined in the LORIS trial protocol (373). However, pathways were extended beyond the 10-year duration of the clinical trial to extrapolate outcomes to a lifelong time horizon more reflective of the chronic nature of the disease and its low propensity for invasive progression. The structure of the usual treatment arm was informed by the economic models of surgery and radiotherapy identified by the review in Chapter 7. The initial structure was then refined with advice from breast surgeons and oncologists to ensure health states reflected the likely progression of low-risk DCIS.

Several iterations were made after running an initial prototype of the model. Given that the natural history of DCIS is unknown, the simplification of potential outcomes was necessary so that only relevant health states were included (e.g. a single recurrence). Finally, the model was presented at both local (Victorian Comprehensive Cancer Centre

Breast Tumour Grand Round, Melbourne, Australia; December 5, 2018) and international (4th World Congress on Controversies in Breast Cancer, Melbourne, Australia; October 11-13, 2018) conferences for external review and face validity prior to running the analysis presented in this thesis. An overview of the model development is illustrated in Figure 8.1.

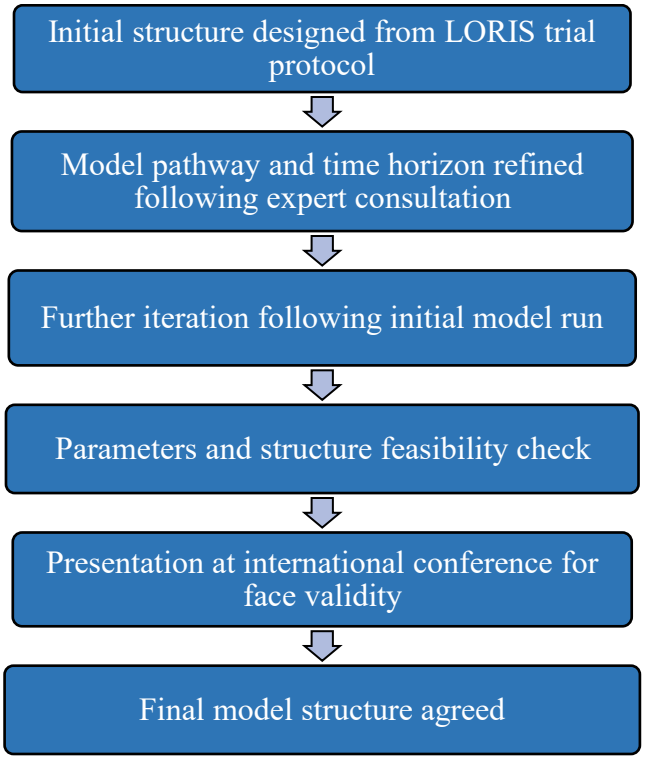


Figure 8.1: An overview of the stages of the model development and validation

8.4.4 Population

The simulation starts with women at age 50 years for the main analysis, coinciding with the starting age at which women are eligible for screening in the UK (374). However, the model was rerun for starting ages of 40, 60 and 70 years in the sensitivity analysis as it was hypothesized that active monitoring would be more cost-effective in the elderly

(where the risk of death from competing illness may be higher than the risk of invasive cancer progression).

8.4.5 Model Pathway

The model was based around the expected pathway of patients participating in the LORIS trial. Briefly, the model simulates a cohort of 1,000,000 hypothetical women aged 50 years with a new diagnosis of low-risk DCIS. Women enter the model at the decision node at time 0 corresponding to the point at which they are diagnosed with low-risk DCIS. For each run of the model, women are simulated under each strategy and are:

(i) treated via standard surgical excision +/- radiotherapy

(ii) managed via active monitoring

and then continue the post-treatment pathway until death (from breast cancer or competing sources) or until the maximum number of cycles (n = 50) are completed. Health states are mutually exclusive (i.e. women can only be in one health state at any one point in time) and represent the natural history of women as they are treated and progress through the model.

The diagram presented in Figure 8.2 shows a simplified structure of the model health states. The clinical pathways for each of the treatments are almost identical, except for that women in the active monitoring arm may stay in the low-risk DCIS health state beyond the first cycle. The health states are represented by a round circle and the arrows

represent the possible transitions between health states within each cycle. Women may move to a different health state (straight arrows), continue in same health state (round arrows) or remain in an absorbing health state (e.g. dead) for the remainder of the model simulation (square box). Women may also transition through intermediate health states within each cycle to accrue incremental costs and rewards from additional procedures (e.g. surgery for a recurrence), as highlighted in the dashed circles. The sum of the transition probabilities moving to and from each health state is equal to 1. All health states may lead to dead from all-cause mortality.

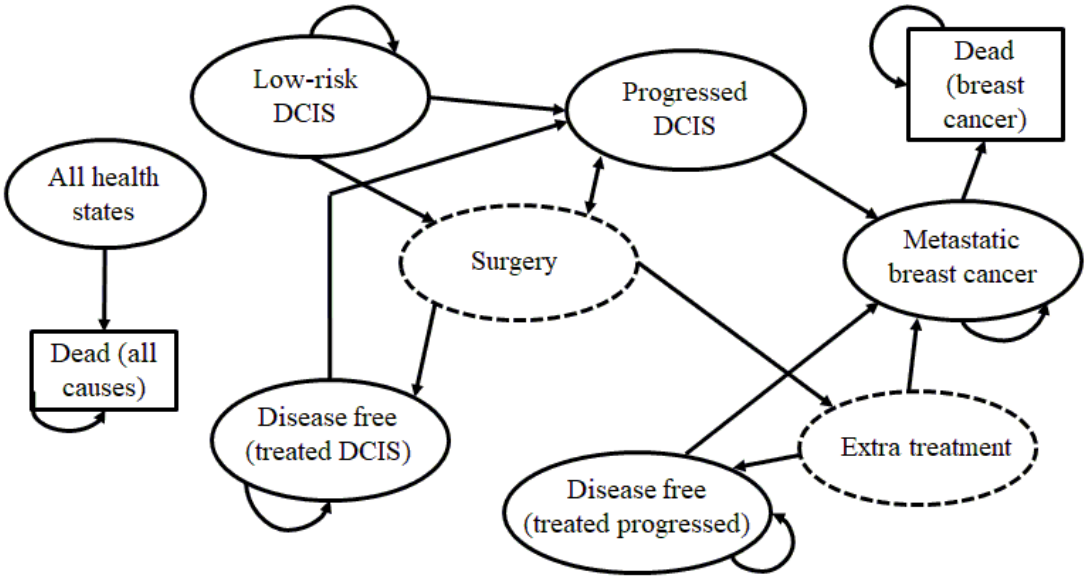


Figure 8.2: Simplification of the model decision following diagnosis of low-risk DCIS

Definition of health states for the standard treatment pathway:

- **Low-risk DCIS:** Women start the simulation in this health state.
- **Surgery:** Women in the treatment arm undergo surgical excision following diagnosis in the first cycle of the trial. They are either cured (disease free and treated) or not cured (progressed disease).

- **Disease free (treated DCIS):** Following successful treatment of low-risk DCIS women may remain cancer free or they may develop a progression (recurrence).
- **Extra treatment:** Women who have progressed disease (DCIS or invasive recurrence) may require additional treatment during the next cycle. They are either cured (disease free and treated progressed) or develop metastatic disease.
- **Disease free (treated progressed):** Women with progressed disease who are successfully treated with further treatment may remain disease free or may relapse and develop metastatic breast cancer.
- **Metastatic breast cancer:** Women may die from metastatic breast cancer.
- **Dead (breast cancer):** Women with metastases may die from breast cancer.
- **Dead (all-causes):** Women in all health states may die from co-morbid conditions.

Definition of health states for the active monitoring pathway:

- **Low-risk DCIS:** Every woman starts the simulation in this health state. Women may stay in this health state if they continue with monitoring and do not develop disease progression, choose to opt out of active monitoring and undergo surgery for low-risk DCIS (disease free treated DCIS) or have a progression (progressed DCIS).
- **Surgery:** Women in the active monitoring arm may stop and choose to have surgery. They are either cured (disease free treated DCIS) or not cured (progressed disease).

- **Disease free (treated DCIS):** Women who stop monitoring and have surgery for low-risk DCIS may remain cancer free or they may develop a progression (recurrence).
- **Extra treatment:** Women who progressed after monitoring, either with high-grade DCIS or invasive cancer, require more extensive treatment. They are subsequently cured (disease free treated progressed) or not cured (metastatic disease).
- **Disease free (treated progressed):** Women who are successfully treated with further treatment following progression may remain disease free or may relapse and develop metastatic breast cancer.
- **Metastatic breast cancer:** Women may die from metastatic breast cancer.
- **Dead (breast cancer):** Women with metastases may die from breast cancer

  Dead (all-causes): Women in all health states may die from co-morbid conditions.

Definition of intermediate health states:

In order to include the relevant outcomes of the treatment disutilities in the model, decision trees are embedded within each of the Markov health states. For example, women in the health state for *"Low-risk DCIS"* in the active monitoring arm may:

(i) stay in the *'Low-Risk DCIS'* health state

(ii) choose to have surgery and transition to the *"Disease-free treated DCIS"*

(iii) develop invasive disease and transition to "*Progressed disease"*

Additional events may occur within each of these transitions (depicted by intermediate health states or decision branches) to facilitate the accumulation of costs and rewards between each cycle. Figure 8.3 illustrates an example of the intermediate health states and trees for women in the *"Low-risk DCIS"* state who choose to continue monitoring, and may be explained as follows:

- If there is no disease progression, then the patient simply returns to the *"Low-Risk DCIS"* health state, provided she does not die from background mortality.
- If there is disease progression, it may either be missed (in which case the woman transitions to *"Progressed DCIS"*) or detected (they proceed to the next decision branch).
- For women with a detected progression, the next branch decides whether they will undergo surgery with or without radiotherapy (A and B).
- Branches for surgery (BCS +/- radiotherapy or mastectomy) incur different disutilities, costs and probabilities specific to the efficacy of the treatment chosen.
- If cured, women transition to *"Disease free progression"* and if not in remission, they transition to *"Metastatic disease"*.

Consequently, women may accumulate costs and rewards when undergoing extra testing or treatment within the low-risk DCIS state. The full tree structure of the Markov model for both active monitoring and standard treatment arms may be referenced in Appendix 12.

8.4.6 Cycle length and time horizon

A cycle length of one year was used to reflect the one-year interval between active monitoring follow-up in the LORIS trial (27). This is consistent with the chronic nature of the disease and the yearly cycle applied in the breast cancer models from the systematic review. A half-cycle correction was applied to both costs and effects in keeping with best practice because it was not known when the transitions might occur during each cycle (375).

A lifelong time horizon was used to assess costs and outcomes. Although the LORIS trial has a 10-year duration, this was not deemed long enough to evaluate the outcomes associated with low-risk DCIS due to the low propensity for disease progression. Therefore, outcomes were extrapolated to a lifelong time horizon to enable follow-up beyond the limitations of the trial. A total of 1,000,000 Monte Carlo simulations were run over 50 cycles and was deemed equivalent to a lifetime for a cohort of women entering the model at age 50 years.

Figure 8.3: A view of the decision tree within the Markov health state for low-risk DCIS in the active monitoring arm

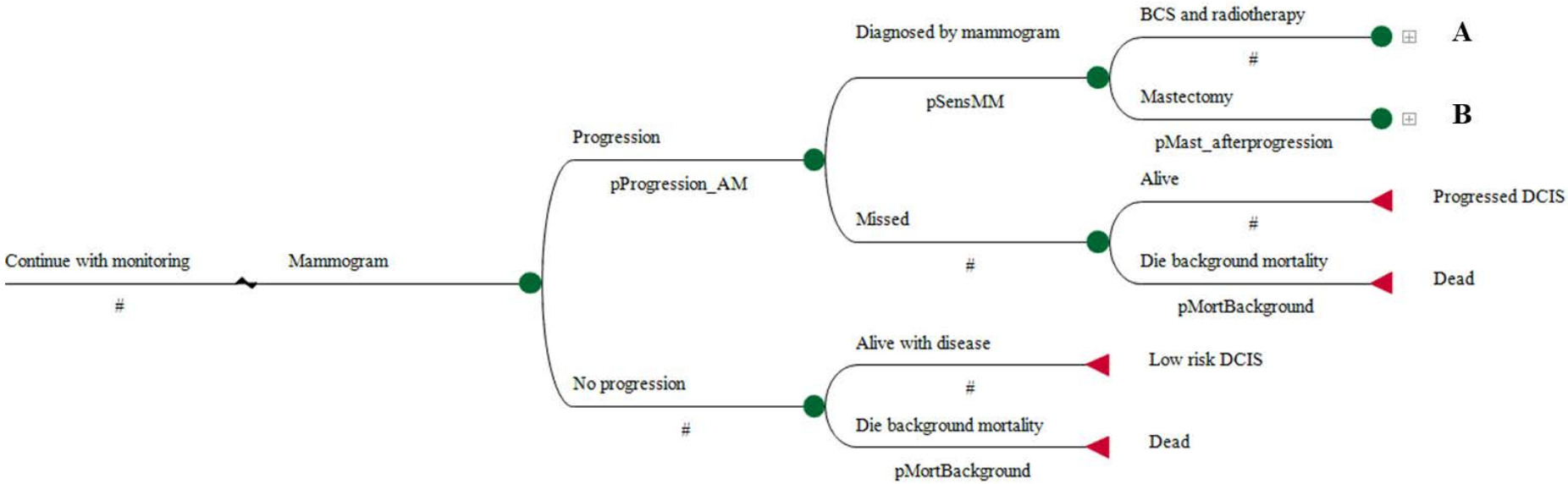
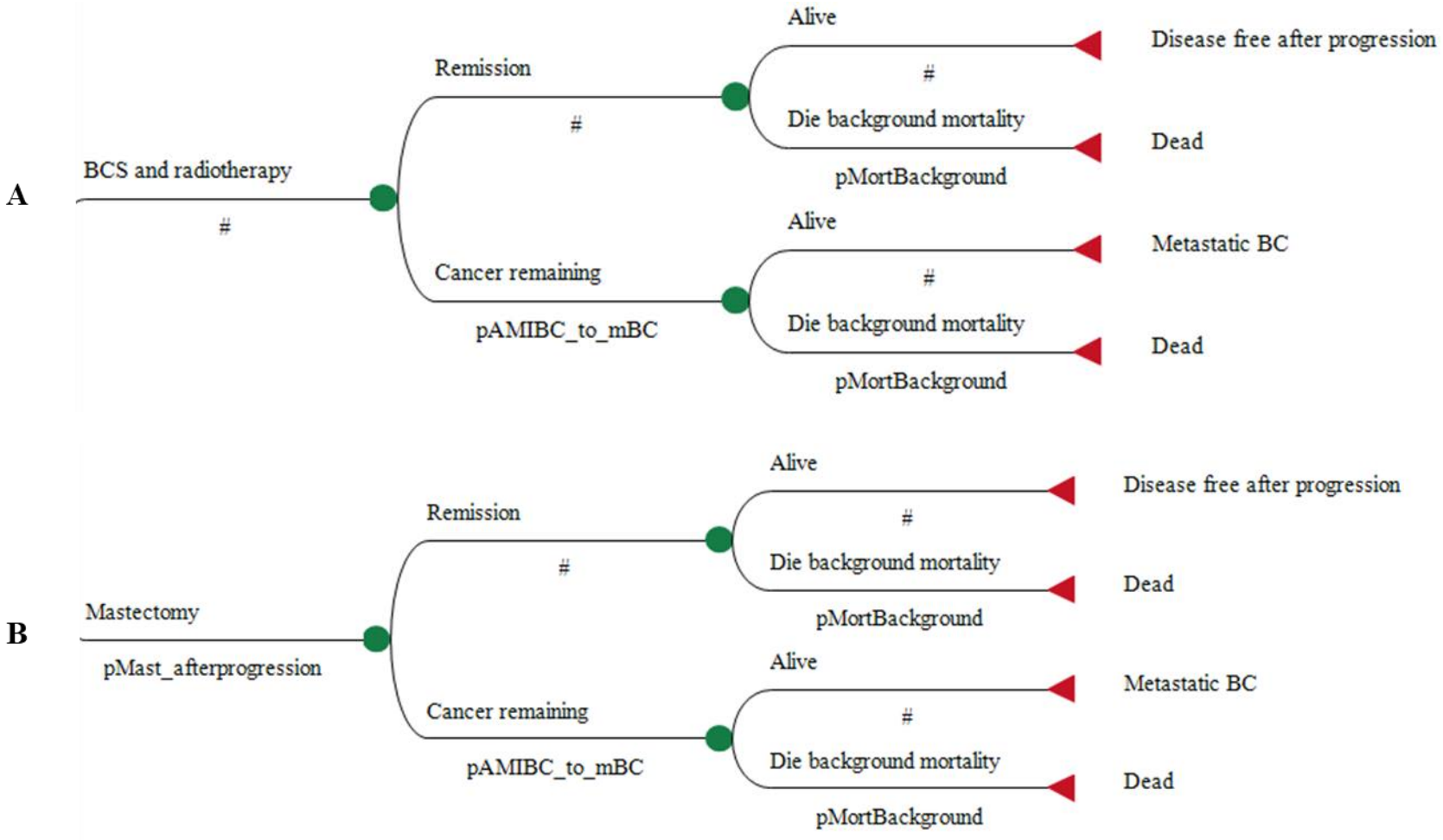


Figure 8.3 continued: Sub-trees for A and B:



8.4.7 Perspective

A UK national healthcare perspective was adopted in the analysis to reflect the taxpayer healthcare system on which the population was modelled (8). Although it was anticipated that active monitoring may incur indirect or unexpected costs related to additional appointments for reassurance or transport needs, information on the indirect impact of active monitoring is unknown (112) and there were insufficient data on the wider impact of active monitoring to run a societal analysis.

8.4.8 Model Assumptions

The results of the LORIS trial are yet to be published. Therefore, several assumptions had to be made in the model due to the complex and unknown natural history of the condition.

- **Compliance:** 100% compliance with treatment was assumed, reflective of the high rates (>96%) of surgical management reported in national breast cancer databases (376). However, patients could opt out of the active monitoring arm in the model, based on the predicted estimates from active surveillance trials in prostate cancer (273).
- **First mammogram:** The first mammogram (and costs) for active monitoring occur in cycle 1 (i.e. 12 months after starting the trial).
- **Active monitoring:** The sensitivity of mammography was assumed to be 86% in keeping with the rates from other UK models of mammography (235).
- **Progressed DCIS:** If a progression was missed on the first mammogram, it was diagnosed in the subsequent active monitoring cycle.

- **Decision to stop monitoring:** Women were able to stop active monitoring. Rates were highest in the first year (when anxiety peaks) and lower in the subsequent years (333).
- **Risk of progression:** The risk of progression whilst monitoring varied with time.
    - Year 1: progression was highest in the first year to account for the upgrade of under staged invasive disease at diagnosis (377)
    - Year 2-5: progression was then lower for five years of monitoring (378)
    - Years ≥6: and then followed the trends reported in the limited retrospective studies of low-risk DCIS (379)
- **Treatment intensity:** Treatment of progressed DCIS required more intensive treatment than low-risk DCIS treated upon diagnosis (103).
- **Adjuvant treatment**: Based on local hospital data, it was assumed that 80% of DCIS was treated via BCS (of which 60% have radiotherapy) and 20% with mastectomy (30). Endocrine treatment was only assumed for progressed disease.
- **Recurrence:** Only one recurrence was permitted (as in other breast screening models)
- **Breast cancer mortality:** Patients could only die from breast cancer if they had metastatic breast disease (82). Therefore, there is no difference in survival for DCIS managed with and without immediate surgery (25).

8.4.9 Model inputs

Model inputs were categorised into transition probabilities, utilities and costs. Estimates were calculated from several different sources including a systematic review of the literature, national cost and mortality tables and expert opinion, as follows.

i) Transition probabilities

The estimation of transition probabilities for the distributions followed the methods described by Briggs (179). Probability parameters are constrained by limits of 0 and 1 and therefore the probabilities of mutually exclusive events within each cycle must also sum up to 1. Briggs et al. (179) suggest that such data are best represented by a binomial distribution, specifically the beta distribution, which is constrained between the values of 0 and 1 and characterised by two parameters alpha (α) and beta (β). If the data are represented by number of events of interest, r, observed in a sample of size n, then the proportion of events to the total sample gives the point estimate of the probability. Uncertainty related to this probability can then be calculated by a (α,β) distribution by setting:

Equation (1) $\alpha = r$

and

Equation (2) $\beta = n\text{-}r$

In TreeAge (372), this is calculated by selecting *'Distributions'* from the *'Values'* ribbon and the beta distribution for probabilities. For integer parameters the software only requires the values of n and r from *equations (1)* and *(2)*. Otherwise, α and β are

generated by inputting the values for the mean and standard deviation from the literature.

This method was used to estimate the transition probabilities for DCIS progression from the number of events reported in an analysis of retrospective studies of non-surgically managed DCIS (378, 379). The exception is for the progression during the first cycle, which was assumed to be 21% based on the findings from a meta-analysis of 1,142 women with under staged invasive cancer at excision for low-intermediate grade DCIS (377).

Where rates were reported in the literature for breast cancer recurrence after treatment, the methods described by Gray (7) were used to convert the values into probabilities so that the values could be applied in a yearly Markov simulation. For the probabilities of progression associated with treatment, the risks from RCTs had to be first adjusted to calculate a yearly transition rate relevant for the model and then the probability (7). Reported estimates (e.g. probability over 10 years follow-up) were first converted to a rate for 1 year, assuming the event occurs over a constant rate r, over a time t:

Equation (3) $r = -ln(1/p)/t$

This equation is then rearranged to convert the yearly rates back into transition probabilities, where ‘e’ is the base of natural logarithm (7):

Equation (4) $P = 1- e^{-rt}$

In this way a yearly transition probability could be estimated from the 10-year recurrence from meta-analyses and RCTs. Transition probabilities for metastatic disease were taken from the estimates reported in other economic models of DCIS treatment.

Table 8.2: Probability parameters used in the analysis

Probability parameters	Base case	PSA range*	Source
Active monitoring			
Sensitivity of mammogram	0.860	0.800-0.900	(235)
Low-risk DCIS to Progressed DCIS			
Year 1	0.210	(0.150-0.280)	(377-379)
Year 2-5	0.100	(0.050-0.180)	
Year ≥6	0.110	(0.090-0.250)	
Choose to stop monitoring (surgery)			(273, 333)
Year 1	0.018	(0.005-0.036)	
Subsequent years	0.005	(0.000-0.018)	
Treatment received (Progressed DCIS)			(376)
BCS with radiotherapy	0.850	Not varied	
Mastectomy	0.150		
Endocrine therapy	0.800		
Local recurrence after treatment (Progressed)			(380-383)
BCS with radiotherapy	0.008	(0.002-0.018)	
Mastectomy	0.004	(0.002-0.022)	
Standard treatment			
Treatment received (low-risk DCIS)			(30, 376, 384)
BCS alone	0.230	(0.200-0.400)	
BCS with radiotherapy	0.570	(0.300-0.800)	
Mastectomy	0.200	(0.100-0.400)	
Progression after treatment (low-risk DCIS)			(380, 381, 385)
BCS alone	0.011	(0.005-0.025)	
BCS with radiotherapy	0.006	(0.004-0.020)	
Mastectomy	0.001	(0.001-0.010)	
Treatment received (Progressed DCIS)			(376)
BCS with radiotherapy	0.850	Not varied	
Mastectomy	0.150		
SNL	0.006		
Endocrine therapy	0.016		
Local recurrence after treatment (Progressed)			(380-383)
BCS with radiotherapy	0.008	(0.002-0.018)	
Mastectomy	0.004	(0.002-0.022)	
Metastatic breast cancer after treated progression	0.030	(0.001-0.005)	(82)
Baseline probabilities			
Death from breast cancer	0.022	Not varied	(386, 387)
Death from all-cause mortality	0.002-0.332		
Discount rate	0.035	0.015-0.060	(8)

**PSA ranges are the lower and upper 95% limits of a beta distribution calculated in TreeAge using the reported error from the studies referenced.*

For events which do not occur at a constant rate with time (e.g. survival), probabilities were adjusted per cycle using life tables. All-cause mortality data were derived from UK life tables, with adjustment for gender, age and breast cancer mortality (386). A 1970 birth cohort was applied, starting at age 50 years. Breast cancer mortality was based on published estimates (Kaplan Meier curves) but only applied to women with metastatic disease (25, 40). The transition probabilities used in the model are given in Table 8.2.

ii) Costs and resource use

Resource use and costs incurred to healthcare services were inputted as model parameters in accordance with the assumed national health perspective. All costs were reported in 2018 GBP Sterling (£) using the UK hospital and community health services index (388) and were discounted at 3.5% (8).

The costs incorporated into the model are outlined in Table 8.3. Inpatient and follow-up costs are included. Direct medical costs for active monitoring, initial treatment, ongoing treatment and follow-up were derived from local hospital data, NHS cost unit tables (389), the British National Formulary and estimates reported in economic evaluations of breast screening programmes (26, 390). Staff costs were calculated using nationally recognised reference costs (388). A weighted average was taken where a combination of treatment was applied simultaneously for the same intervention (e.g. diagnostic investigation includes the cost of imaging and biopsy). Resource use for active monitoring was estimated by clinical experts due to the uncertainty in health service use relating to the blinded nature of the LORIS trial.

For the probabilistic sensitivity analysis, cost data were constrained to be non-negative and were made up of counts of resource use weighted by unit costs. In standard statistics this type of data is represented by a Poisson distribution, or a gamma distribution (179). A gamma distribution is constrained by an interval of 0 to positive infinity and is thus appropriate to represent cost parameter uncertainty. To fit a gamma distribution parameterised as gamma, $\gamma(\alpha, \beta)$ to the data in Tree Age (372), the mean and variance in the observed sample from the literature is set equal to the corresponding expression for mean (μ) and variance (s^2) of the distribution (179), where:

Equation (5) $\mu = \alpha.\beta$

and

Equation (6) $s^2 = \alpha.\beta^2$

The above equations are subsequently rearranged to solve for the two unknowns:

Equation (7) $\alpha = \mu^2/s^2$

and

Equation (8) $\beta^2 = s^2/\mu$

Where the standard error was not reported in cost tables it was assigned the same value as the mean. For example, for mammography the parameters of the gamma distribution are estimated from the equations above, where: $\alpha = 56/56 = 1$ *and* $\beta = 56^2/56 = 56$, resulting in a gamma (1, 56).

When distributions were entered in Tree Age Pro (372), the software automatically parameterised the gamma distribution with the reciprocal of the beta parameter (λ).

Table 8.3: Cost data used in the analysis

Item	Unit cost (£)	Source
Active monitoring		
Mammography	56	(390)
Result by letter	0.61	(391)
Specialist Breast Nurse follow-up (non-face to face)	24	(389)
Diagnostic investigation (imaging and biopsy)	289	(389)
Breast multi-disciplinary team review	96	(389)
Standard treatment		
Breast conserving surgery	1450	(26, 389)
Adjuvant radiotherapy	1800	(26, 389)
Mastectomy (no reconstruction)	2810	(26, 389)
Mastectomy (with reconstruction)	2920	(389)
Endocrine therapy (Tamoxifen)	287	(391)
Sentinel lymph node procedure	3074	(389)
Extra treatment for progressed disease	11,918	(390)
Metastatic breast cancer	15,787	(390)
Palliative care	20,685	(26)
Mammogram	56	(390)
Breast cancer specialist follow-up	130	(389)

**All costs are presented in 2018 (£) Sterling*

iv) Utilities

Utilities representing the impact of treatment on quality of life on a scale of 0 (dead) to 1 (perfect health) were applied to women transitioning through each health state in the model. Baseline utility was assigned to each woman on entry into the model using age-specific female EQ-5D tariffs for the UK general population (244).

Utilities for individual health states collected in the empirical study (Chapter 6) and systematic review (Chapter 4) were then applied for each of the health states to adjust

baseline utility as women moved through the model and accrued benefit and harm. This included intermediate payoffs, such as the disutility of undergoing further investigation for a suspicious monitoring mammogram, so that all relevant effects on quality of life were represented in the model. The non-patient population EQ-5D values from the empirical study were used in the base-case analysis to ensure consistency between the utilities applied in line with NICE recommendations (8), and to avoid the ceiling effect observed in the standard gamble.

For the base-case analysis, utilities from both the empirical study (capturing overdiagnosis) and systematic review (284) (unlikely to capture overdiagnosis) were applied in two separate runs of the model to observe the potential effect of including this disutility on QALYs. Where the duration of utilities lasted less than the yearly time cycle, QALYs were adjusted so that the total utility for each of the health states added up to one year. For example, if the utility for diagnostic investigation lasted for 5 weeks, followed by a utility of 0.880 for the remaining weeks of the year for a negative result (i.e. they remain in low-risk DCIS monitored) then the annual utility in a year is calculated as: (0.895 x 5/52) + (0.880 x 47/52) = 0.881 and the QALY disutility is 1-0.881 = 0.119. Other disutilities were calculated using the same technique. A full list of the utilities and durations applied in the model is shown in Table 8.4.

iv) Discounting

Both costs and benefits were discounted at 3.5% per annum in the base-case analysis, as recommended by NICE (8). Discount rates were varied in a one-way sensitivity analysis

to reflect the variation in tariffs stipulated by the British Treasury green book (i.e. 3.5% costs, 1.5% benefits) (392) and other decision bodies (356).

Table 8.4: Health state utility data used in the model

Health state	Empirical utility (SD)	Sys. review utility (SD)	Source	Duration
Base-line population age-specific utility				
45-54 years 55-64 years 65-74 years ≥ 75 years	-	0.846 0.815 0.777 0.712	(244)	1 year
Active monitoring				
Low-risk DCIS (monitoring)	0.880 (0.13)	0.830 (0.24)	(332)	Lifelong
Patient elects to stop AM	-	0.990 (0.01)	(393)	1 week
Progressed DCIS	0.550 (0.27)	0.890 (0.16)	(248)	10 months
Diagnostic investigation	-	0.895 (0.2)	(176)	5 weeks
Missed progression	-	0.660 (0.29)	(58)	1 year
Standard treatment				
Breast conserving surgery alone	0.800 (0.14)	0.900 (0.15)	(248)	1 year
Breast conserving surgery + radiotherapy	0.700 (0.19)	0.900 (0.14)	(248)	1 year
Mastectomy	0.580 (0.23)	0.790 (0.27)	(254)	1 year
Disease free (treated low-risk DCIS)	-	0.844 (0.20)	(250)	1 year
Recurrence (progression)	0.550 (0.27)	0.779 (0.20)	(250)	10 months
BCS and radiotherapy for progression	-	0.810 (0.19)	(248)	1 year
Mastectomy for progressed disease	-	0.840 (0.18)	(248)	1 year
Endocrine therapy	-	0.820 (0.17)	(176)	10 years
Sentinel lymph node procedure	-	0.900 (0.14)	(235)	2 months
Disease free (treated progression)	-	0.779 (0.22)	(250)	1 year
Invasive progression (recurrence)				
Metastatic breast cancer	-	0.685 (0.19)	(250)	1 year
Palliative care	-	0.288 (0.15)	(176)	Until death

** Utility values are rounded to 3 decimal places*

8.4.10 Analysis

8.4.10.1 Model outputs

Outcomes were primarily measured in terms of cost per QALY, although clinical outcomes such as the number of surgeries, progressions and breast cancer deaths were also captured in the model. A summary of the model parameters and outputs is provided in Table 8.5.

For the base-case analysis outcomes were compared for active monitoring and standard treatment using the utilities derived in the empirical study and the utilities from the literature. Results were reported as an incremental cost effectiveness ratio (ICER), in terms of the cost per QALY gained between standard treatment and active monitoring. If an intervention was less costly and generated a greater number of QALYs, it was reported as dominant.

Table 8.5: A summary of the model parameters and outcomes

Model	***Parameter or analysis***
Population	*Women aged 50 years with low-risk DCIS* *(1970 UK female birth cohort starting at age 50 years)*
Intervention	*Active monitoring (yearly mammogram)*
Comparator	*Standard treatment (surgery +/- radiotherapy)*
Outcomes	*Cost per QALY, ICER, breast cancer deaths*
Perspective	*UK National Health Service (2018 GBP)*
Time horizon	*Lifelong*
Cycle length	*1 year*
Sensitivity analysis	*One-way, deterministic and PSA.* *Cost-effectiveness threshold* £20,000 per QALY gained

8.4.10.2 Sensitivity analyses

The model parameters described thus far were used in the base-case analysis. To evaluate how the results were influenced by the model assumptions and parameter uncertainty, one-way sensitivity analyses were conducted around key variables (probability of invasive progression, costs and discounting etc.). Specifically, this required changing one variable at a time to determine how the cost-effectiveness varied according to each parameter. Where one parameter was changed to create a new

possible scenario (e.g. changing the start age of women in the model from 50 to 60 years) this was termed a deterministic sensitivity analysis.

Probabilistic sensitivity analysis was also performed using distributions instead of expected values to test parameter and overall decision uncertainty. This was calculated by running 1,000 individual trials via second order Monte Carlo simulation, with sets of draws randomly sampled from probability distributions around the parameters. The advantage of this type of analysis is that all parameter uncertainties can be varied simultaneously. Uncertainty around event probabilities and utilities were represented using β distributions and costs using γ distributions as outlined above (179).

8.4.10.3 Model validation and quality assessment

The model was validated using three separate processes. First, the number of breast cancer deaths simulated in the model were compared with breast cancer mortality in UK national cancer tables for women (394). Breast cancer specific mortality in the model was comparable to that observed in national data and other treatment models of breast cancer surveillance (103). Second, a series of logic tests in which mortality and transition probabilities were modified to extreme values (+/- 25% of the baseline value) were undertaken to test whether the model worked as expected, i.e. increasing/decreasing the probability of progression actually resulted in an increase/decrease in the number of women developing cancer in the model. Third, the model pathway was reviewed by both clinicians and health economists to determine whether the outputs were sensible and indicative of the likely natural history of the disease (face validity).

The cost-effectiveness of active monitoring is dependent on the outcomes of the LORIS trial, the results of which are not expected until the trial is completed in 2025. Clinical implementation of active monitoring is ultimately dependent on the efficacy and safety in women with low-risk DCIS and not the sensitivity of costs and utility estimates at this stage of the investigation. Therefore, a value of information analysis was not conducted as it was beyond the scope of the model objectives and the thesis.

8.5 Results

8.5.1 Base-case analysis

Cost utility analysis

The results for the base-case analysis are presented in Table 8.6. Two model simulations are shown; the first corresponds to the results using the utilities from the empirical study (Simulation 1) and the second using those from the systematic review (Simulation 2). The results from the empirical utilities analysis (Simulation 1) suggest that active monitoring might be cost-effective when compared with standard surgical treatment. Active monitoring is more expensive but also yields more QALYs, with a reported ICER of £2,125 per QALY.

In comparison, when the utilities from the literature were used in place of those from the empirical study (Simulation 2), the magnitude and direction of cost-effectiveness is reversed. Active monitoring is now dominated by standard treatment, i.e. active monitoring is more expensive *and* yields fewer QALYs than surgical excision. In this case, where the risk of overtreatment is not included in the valuation of QALYs, active

monitoring would be unlikely to be considered by policy makers as it is inferior to standard treatment in terms of both costs and outcomes over a lifelong time horizon.

Table 8.6: Results from the cost-effectiveness analysis using empirical study and systematic review utilities

Treatment	Mean Cost (£)	Incremental Cost (£)	Effectiveness (QALYs)	Incremental effectiveness (QALYs)	ICER (Cost/QALY)
Simulation 1: Results of the model using the utilities from the empirical study					
Standard treatment	20,503		12.84		
Active monitoring	21,098	595	13.12	0.28	*2,125*
Simulation 2: Results of the model using the utilities from the literature					
Standard treatment	20,483		12.99		*Dominant*
Active monitoring	21,108	625	12.58	-0.41	*Dominated*

The results were plotted onto the cost-effectiveness plane, presented in Figure 8.4, to further illustrate how the disutility of overdiagnosis might impact upon the cost-effectiveness decision. For women aged 50 years, active monitoring is dominated in comparison to standard treatment (i.e. northwest quadrant) if the utilities from the literature are applied (no valuation of unnecessary treatment). However, active monitoring may be considered cost-effective (north-east quadrant), provided society is willing to pay for it, if utilities which explicitly capture overdiagnosis are included in the analysis. This suggests the impact of unnecessary treatment from overdiagnosis is important in the decision around the treatment of low-risk DCIS.

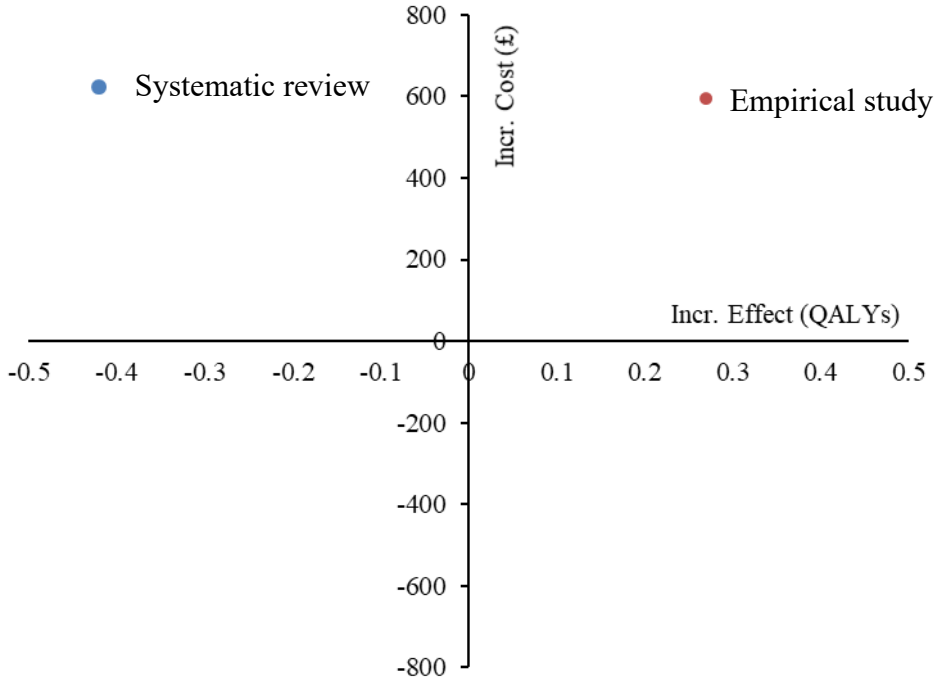


Figure 8.4: Cost-effectiveness plane comparing active monitoring against standard treatment (surgery) using the empirical study and systematic review results

Clinical outcomes

In the base case analysis (Simulation 1: empirical study utilities in women aged 50 years), 65% of women initially managed with active monitoring in the model undergo definitive treatment because of progressed disease or patient choice over a lifelong time horizon. The 20-year risk of ipsilateral breast cancer recurrence was higher in the active monitoring group (24% versus 11%), but this did not adversely impact upon breast cancer survival.

Importantly, there was no significant difference in breast cancer-specific mortality between the two groups (20-year BCSM: 0.98, P<0.001). However, those in the active

monitoring group required more intensive treatment, with 6% more women requiring a mastectomy for progression compared with the standard treatment group. The risks were reduced in women over 70 years in the monitoring group as fewer women lived long enough for the condition to progress (18% recurrence and 4% rate of mastectomy).

8.5.2 Sensitivity analyses

8.5.2.1 One-way

One-way sensitivity analyses were performed on each of the input parameters using the variability around the mean to determine the potential impact and key drivers of the cost-effectiveness results. The parameters associated with active monitoring, in particular, were critical to the model results due to the high level of uncertainty surrounding the point estimates. The results suggest that age, duration of follow-up, discount rate, probability of invasive progression and utility of active monitoring are key drivers in the decision. For example, active monitoring may be cost-effective until the associated utility falls below a threshold of 0.78.

8.5.2.2 Threshold sensitivity analysis

Further threshold sensitivity analysis was undertaken for the parameters identified as driving cost-effectiveness results in the one-way analysis. A series of threshold sensitivity analyses were conducted to test the main assumptions relating to; (1) start age of women in the population, (2) time horizon, (3) discount rate, (4) probability of progression and (5) utility of active monitoring.

(1) Start age

If the start age of the population with low-risk DCIS is ≥60 years, then active monitoring becomes dominant in comparison to standard surgical treatment. Conversely, starting the simulation with women aged 40 years resulted in an ICER of £8,456 per QALY compared to standard treatment. As one would expect monitoring becomes more expensive in line with the longer follow-up time required with younger women.

Table 8.7: Results of the cost-utility analysis per start age of women in the model

Population	Strategy	Mean Cost (£)	Inc. Cost (£)	Effectiveness (QALYs)	Inc effect (QALYs)	ICER (Cost/ QALY)
Start age 40	Standard	24,488	-	14.88	-	-
	Tx AM	26,771	2,283	15.15	0.27	8,456
Start age 50	Standard	20,503	-	12.84	-	-
	Tx AM	21,098	595	13.12	0.28	2,125
Start age 60	Standard	15,515	-	10.44	-	*Dominated*
	Tx AM	14,355	-1160	10.71	0.27	*Dominant*
Start age 70	Standard	10,818	-	7.77	-	*Dominated*
	Tx AM	8,090	-2728	8.01	0.24	*Dominant*

(2) Time horizon

Similar to the start age specified in the model, the cost-effectiveness of active monitoring and standard treatment is also dependent on the duration of follow-up accrued after diagnosis. At 31.78 years active monitoring is no longer dominant in comparison to immediate surgery, as shown in the lower end of the ten-year categories in Table 8.8.

Table 8.8: Results of the cost-utility analysis for time horizon

Time horizon	Strategy	Mean Cost (£)	Inc. Cost (£)	Effectiveness (QALYs)	Inc effect (QALYs)	ICER (Cost/ QALY)
10 years	Standard Tx AM	7,012 3,005	- -4,007	5.82 6.00	 0.18	*Dominated* *Dominant*
20 years	Standard Tx AM	11,816 9,322	- -2,494	9.50 9.76	- 0.26	*Dominated* *Dominant*
30 years	Standard Tx AM	16,661 15,869	- -792	11.67 11.97	- 0.30	*Dominated* *Dominant*
40 years	Standard Tx AM	19,768 20,415	- 647	12.65 12.93	- 0.28	- 2,311
50 years	Standard Tx AM	20,503 21,098	- 595	12.84 13.12	- 0.28	- 2,125

(3) Discount rate

If the UK Treasury discount rates are applied (392) (i.e. 3.5% for costs, 1.5% for QALYs) then the ICER falls from the base-case value of £2,125 per QALY to £2,052 per QALY. The costs are unchanged from the base-case at 3.5% but active monitoring may delay or avoid unnecessary treatment, and thus future outcomes are favoured by the 1.5% discount rate. If no discount rate is applied (0% in both) the delayed intervention accrued in the active monitoring arms becomes dominant.

Table 8.9: Results of the cost-utility analysis for discount rate

Discount rate	Strategy	Mean Cost (£)	Inc. Cost (£)	Effectiveness (QALYs)	Inc effect (QALYs)	ICER (Cost/ QALY)
0%	Standard Tx AM	40,330 38,071	- -2,259	22.05 22.34	- 0.29	*Dominated* *Dominant*
3.5% / 1.5%	Standard Tx AM	20,503 21,098	- 595	17.13 17.42	- 0.29	- 2,052
3.5%	Standard Tx AM	20,503 21,098	- 595	12.84 13.12	- 0.28	- 2,125

(4) Probability of progression

There was considerable uncertainty surrounding the transition probability of invasive progression in the model due to a lack of data for untreated DCIS beyond a 20 year time horizon. As such this was identified as a key driver of cost-effectiveness. A 20% risk of progression in 20 years was modelled in the base case analysis using the limited estimates published in a retrospective cohort analysis of DCIS managed via biopsy alone (379), resulting in an ICER of £2,125 per QALY. If the probability of invasive progression is reduced to 10% then active monitoring is *dominant*, however if the risk is increased to 40% then the direction is reversed (i.e. active monitoring is *dominated*).

Table 8.10: Results of the cost-utility analysis for probability of progression

Prob. of progression	Strategy	Mean Cost (£)	Inc. Cost (£)	Effectiveness (QALYs)	Inc effect (QALYs)	ICER (Cost/ QALY)
10%	Standard	20,438	-	12.83	-	*Dominated*
	Tx AM	13,917	-6,521	13.35	0.52	*Dominant*
20%	Standard	20,503	-	12.84	-	-
	Tx AM	21,098	595	13.12	0.28	2,125
40%	Standard	20,440	-	12.84	-	*Dominant*
	Tx AM	33,811	13,371	12.71	-0.13	*Dominated*

(5) Utility of AM

The utility of active monitoring was a key driver of cost-effectiveness. In the base case analysis, the utility from the empirical study (0.880) yielded an ICER of £2,125 per QALY. However, if the utility of active monitoring is reduced (0.780), so that the quality of life associated with surveillance is less accepted, then active monitoring is *dominated* by standard treatment (Table 8.11). Conversely, if the utility of active monitoring is increased (or the utilities associated with standard treatment are decreased) then the analysis favours the direction of active monitoring.

Table 8.11: Results of the cost-utility analysis for utility of active monitoring

Utility of AM	Strategy	Mean Cost (£)	Inc. Cost (£)	Effectiveness (QALYs)	Inc effect (QALYs)	ICER (Cost/ QALY)
0.780	Standard Tx AM	20,363 21,042	- 679	12.85 11.89	- -0.96	*Dominant* *Dominated*
0.880	Standard Tx AM	20,503 21,098	- 595	12.84 13.12	- 0.28	- 2,125
0.980	Standard Tx AM	20,566 20,848	- 282	12.83 14.35	- 1.52	- 186

8.5.2.3 Probabilistic sensitivity analysis

Figure 8.5 presents the results of a probabilistic sensitivity analysis which demonstrates the uncertainty surrounding the optimal management of low-risk DCIS. Although there is a 70% chance that active monitoring is cost-effective with no willingness-to-pay threshold, this probability is reduced to 47% at the standardised NICE threshold of £20,000 per QALY. In contrast, the probability that standard treatment is cost-effective at the £20,000 threshold is 53%. As such, there is no clear decision in favour of one treatment over the other in terms of costs and effects. This finding is reinforced by the incremental cost-effectiveness plane (Figure 8.6), which displays the widespread heterogeneity in the distribution of potential costs and effects for active monitoring versus standard surgical treatment, correlating with individual heterogeneity in preferences (utilities) toward treatment.

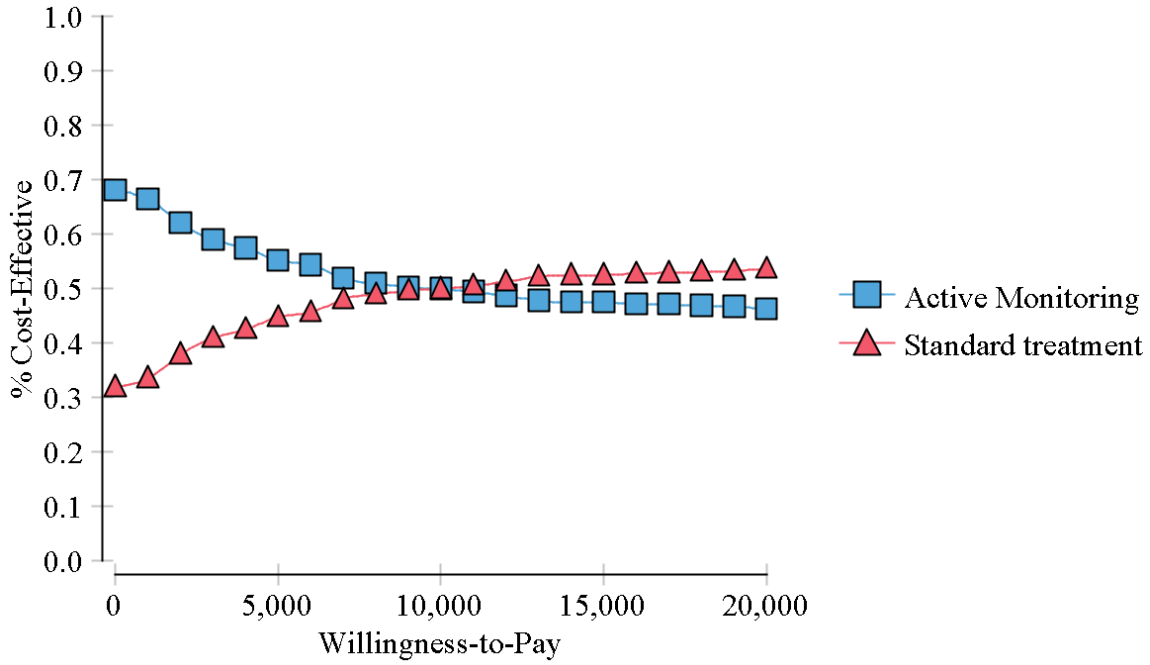


Figure 8.5: Cost-effectiveness acceptability curve for standard treatment and active monitoring

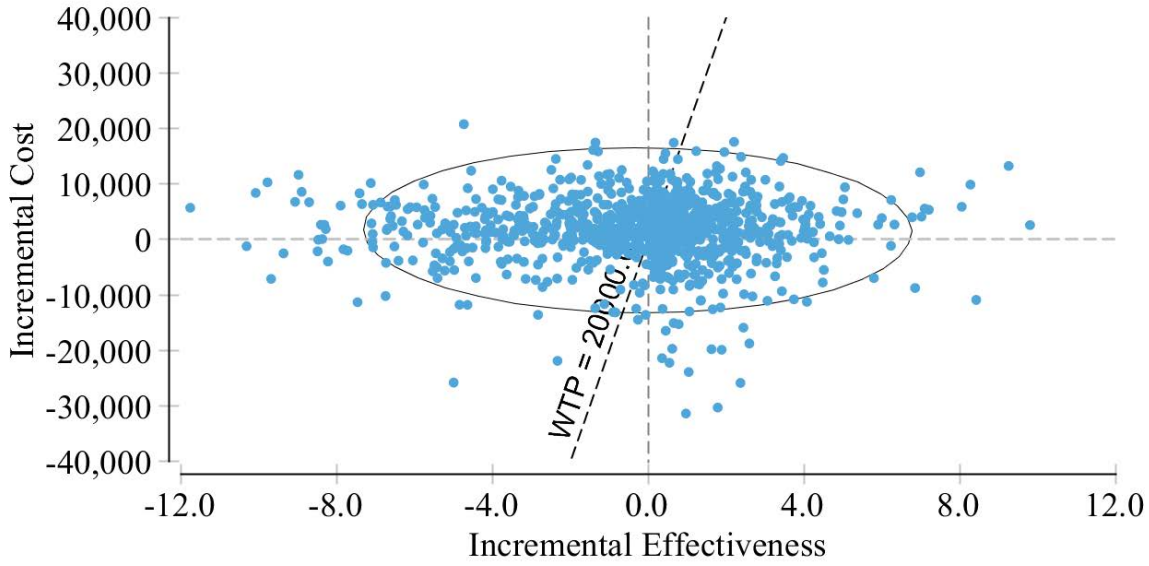


Figure 8.6: Incremental cost-effectiveness plane for standard treatment (origin) versus active monitoring

8.6 Discussion

8.6.1 Main findings

In this chapter, an economic model comparing active monitoring with standard treatment was presented to illustrate the application of the utilities from the empirical study. The economic evaluation was first conducted using the values from the empirical study to generate QALYs, and then re-estimated using the utilities identified in the literature review to assess the impact of including the disutility of overdiagnosis on the cost-effectiveness results. It was assumed from Chapter 6 that the description of overdiagnosis explained (in part) the differences between the utilities in the empirical study and other studies, so that the findings could be applied in the economic model presented in this chapter.

Keeping everything else constant, if cost per QALY results are compared using the utilities from the empirical study and the literature then the inclusion of the disutility of overdiagnosis changed both the magnitude and direction of cost-effectiveness in the treatment of low-risk DCIS. The relative cost-effectiveness of active monitoring compared to standard treatment for women aged 50 years is £2,125 per QALY using the utilities derived in this thesis. This means that every additional QALY costs an extra £2,125. As the NICE guidelines recommend new interventions into practice if the ICER is below a threshold of £20,000 per QALY (8), active monitoring might be considered cost-effective and recommended for low-risk DCIS provided the clinical efficacy data from LORIS is favourable.

However, the decision would be reversed if the results using the utilities from the literature (which do not factor overdiagnosis into the valuation) are applied. When this harm is not explicitly valued in terms of quality of life, active monitoring was *dominated* by standard treatment over a lifelong time horizon. This finding is important because it suggests that the disutility of overdiagnosis and unnecessary treatment is likely to impact the decision on breast cancer treatment. Misplaced breast cancer screening and treatment decisions may be made unless the benefits and harms are valued appropriately. This suggests that it is important that the harms of overdiagnosis are factored into future breast cancer policy through the selection of appropriate utilities.

The impact of overdiagnosis on quality of life when quantified appropriately does matter. However, there is considerable uncertainty surrounding the magnitude of this harm. Deterministic (one-way) analysis showed that the cost-effectiveness of treatment for low-risk DCIS was sensitive to the risk of invasive progression, perceived utility of treatment, duration of follow-up and age. This indecision was highlighted further in probabilistic sensitivity analysis whereby both surgery and monitoring strategies yielded similar results. Given the uncertainty around outcomes related to overdiagnosis, decision making for low-risk DCIS will remain problematic until the results of clinical trials for active monitoring are known. The results of this preliminary economic model do not suggest that active monitoring is likely to be a viable and cost-effective option given the uncertainty and PSA results. However, the threshold analysis suggests that future analyses directed at older women (>60 years) with localised disease who may

wish to reduce the risk of unnecessary treatment may benefit from further research once the results from LORIS are known.

8.6.2 Strengths and limitations

A strength of the evaluation presented in this chapter is that it has applied novel utilities collected prospectively to ascertain the potential impact of including net harms on treatment decisions for low-risk DCIS. It builds on the findings from Raftery (31) by explicitly capturing the disutility of the risk of overtreatment in the economic evaluation of breast cancer treatment. By comparing QALYs using utilities which capture the risk of unnecessary treatment with those from the literature, the direction of cost-effectiveness is changed suggesting it is likely key to the decision on cost-effectiveness.

The results are important in informing the debate about the management of low-risk breast cancers identified through screening as they demonstrate how the explicit inclusion of overtreatment (from overdiagnosis) may impact the evidence informing policy. Furthermore, the model was based on a thorough review of published models for active monitoring in the wider oncological setting, and the illustrative pathway was validated through discussion with clinical experts.

A limitation of the model is the significant uncertainty surrounding invasive progression estimates and resource use in the active monitoring arm. Sensitivity analysis showed that this did have a significant impact on cost-effectiveness estimates, but further refinement of these parameters is dependent on the future publication of the results from clinical trials such as LORIS (27). Assumptions regarding transition probabilities had to

be made given the paucity in published progression data for low-risk DCIS. Consequently, the probability of progression is assumed to continue proportionally despite the chance that low-risk DCIS may recede or remain dormant over time. However, this limitation is unlikely to have had a significant impact on the results given the main objectives of the model as an illustrative vehicle for the application of utilities.

A further potential limitation is the uncertainty surrounding the magnitude of the disutility of overdiagnosis. This analysis applied the values derived in the empirical study, but in reality women are likely to self-select treatment in the direction of their preferences (i.e. women who place a greater disutility on overdiagnosis are more likely to choose and value active monitoring, and those who do not will opt for surgery). Given the diversity observed in treatment preferences and clinical practice, it is unclear as to which utilities should be applied in the evaluation of low-risk DCIS nor whether the observed differences may be due to other factors. The model also assumed that the differences in the results in the base-case analysis from applying the empirical study and systematic review utilities were mostly related to overdiagnosis, but other factors are likely to have contributed to the differences in part. Regardless the results do suggest that the risk of potential overtreatment is likely to impact upon the decision, but further research is necessary to clarify the magnitude of this harm on decisions.

8.6.3 Comparison to the literature

To current knowledge, this is the first model-based evaluation to explore the cost-effectiveness of active monitoring for ductal carcinoma in situ. Whilst active monitoring has been shown to be cost-effective in other low-risk cancers including prostate (364)

and cervical (395), there has been no exploration of the potential policy implications of surveillance for DCIS in the breast cancer setting. One computational analytic model (103) has been published relating to active surveillance for DCIS, but the focus of the model was aimed at estimating the 10-year disease-specific cumulative mortality for AM and the authors did not conduct an analysis of costs or quality of life in the results.

8.6.4 Implications for health policy

Based on the data from the model, the economic evaluation does not suggest that either active monitoring or standard treatment is superior in the management of low-risk DCIS. Rather, the results provide clear evidence in support of the inclusion of utilities which appropriately capture the benefits and harms associated with each treatment, particularly overdiagnosis. The implications of the main findings may have profound impact for both breast cancer screening and treatment policy, as when the net harm of overtreatment is factored into the decision, the direction of cost-effectiveness is potentially changed. The findings support the suggestion by Francis (30) that the overtreatment of low-risk DCIS may be doing more harm than good after 10 years if the disutility of overdiagnosis is not factored into the decision-making process. Policy makers must ensure that the evidence informing decisions appropriately captures the benefits and harms so that treatment is targeted appropriately. Refining the model with more accurate input values is necessary on completion of the LORIS trial so that policy makers can make informed decisions.

8.6.5 Further research and recommendations

The results from this chapter make an important contribution in understanding the association between breast cancer treatment and the harm of overdiagnosis. The explicit inclusion of overdiagnosis in the utilities from the empirical study was found to be a key driver (among other variables) of cost-effectiveness in the evaluation of treatment of low-risk DCIS, re-iterating the need for more conclusive evidence on the true harms of overdiagnosis on quality of life. It is necessary to explore the wider implications of this harm on breast cancer screening policy by including these effects in a decision analytic model of mammography screening. Only then may the debate about breast cancer screening and the potential cost-effectiveness of active monitoring be truly addressed.

8.7 Conclusion

In this chapter, the utilities from the empirical study were applied in an economic model representing the management of low-risk DCIS based on the LORIS trial. The results suggest that the disutilities associated with active monitoring and conventional management of DCIS are important in the assessment of the cost-effectiveness of the benefits and harms of DCIS treatment. When the model was re-run using the utilities from the literature (whereby women may not have been aware that the treatment may have been unnecessary), the results imply that active monitoring is no longer cost-effective. The disutility of overdiagnosis is thus likely to have a significant impact on cost-effectiveness and the evidence informing policy decisions for patients with low-risk disease. The findings illustrate the importance of including the net benefits and harms in the decision. However, the model is exemplar and therefore no conclusions can be made on the likely cost-effectiveness of active monitoring versus standard

treatment, as reiterated in the probabilistic sensitivity analysis and the significant uncertainty around the risk of invasive progression. In the next chapter, the ideas driving the differences in utilities observed in the empirical study and economic model are considered in a qualitative analysis of women's preferences toward DCIS treatment.

CHAPTER SUMMARY

What is known

- The literature on economic models of active monitoring for screen-detected cancers suggest a state-transition model (ISM) with yearly transitions is optimal for comparing active monitoring with standard treatment.
- A state-transition model was designed to model women with low-risk DCIS indicative of the LORIS trial.
- A hypothetical cohort of 50-year old women with low-risk DCIS randomised to active monitoring or standard treatment (surgery +/- radiotherapy) were modelled using a yearly cycle and lifelong time horizon.

What this chapter adds

- The inclusion of the benefits and harms of treatment in the utilities applied in the economic evaluation of DCIS treatment is important. Active monitoring might be cost-effective compared with standard treatment when using the utilities from the empirical study (overdiagnosis captured) but is dominated when the values from the literature are applied (no overdiagnosis).
- Threshold analysis was sensitive to the probability of invasive progression, utility of active monitoring, time horizon, age and discount rate.
- The results suggest that overdiagnosis is important in the decision on breast cancer treatment (further research is required to validate the precise impact upon breast cancer screening programmes).

CHAPTER 9

A qualitative study of women's preferences for treating DCIS

9.1 Introduction

So far in this thesis the primary consideration has been the measurement and valuation of the benefits and harms of treating ductal carcinoma in situ in terms of utility. In Chapter 6, it was observed that women valued active monitoring more favourably than invasive treatment on average, and utilities for health states explicitly describing overdiagnosis were lower than those previously elicited from the literature. These values are important in quantifying the overall preferences of the population and informing economic evaluation, but they do not explain the individual variation relating to how these benefits and harms are perceived. Breast conserving surgery and radiotherapy are common practice for most women diagnosed with DCIS at mammography. However, treatment preferences are diverse with some women opting for a mastectomy regardless of prognosis and others feeling equally as strongly about active monitoring. In order to understand the reasoning behind this heterogeneity, a qualitative study was undertaken to explore the factors influencing women's decisions and preferences for managing low-risk DCIS.

This chapter reports the background, methods and results of the qualitative interviews undertaken alongside the main empirical study in Chapter 6. The objective was to gain an insight into the factors underpinning how utilities are understood and interpreted and a fuller understanding of patient preferences for managing low-risk DCIS. The chapter

begins with a brief introduction into qualitative research methods and their associated theoretical assumptions. The chapter concludes with a thematic analysis of the data collected and a summary on how utilities and disutilities might be conceptualised by women with experience of the treatments and potentially overtreatment. The principal role of the research student as the interviewer and relevance of clinical background is considered in the interpretation of the findings.

9.2 Qualitative research in health economics

9.2.1 Qualitative research methods

Qualitative research is concerned with interpretation of the meanings people attach to their experiences of the social world (396). It goes beyond quantitative hypothesis testing and statistical analysis to gain an understanding of the beliefs, assumptions, values and practices shared by a community (397) through exploration of emotional, social and experiential phenomena. Its origin is grounded in beliefs about the nature of the social world and what can be known (ontology), what knowledge exists and how such knowledge can be acquired (epistemology) and the techniques that can be used to obtain this knowledge (methodology) (398). Although conventionally adopted by social scientists and psychologists (399), there has been increased demand for the use of qualitative or mixed methods research in recent years (400), reflecting a need to understand public preferences for priority setting and healthcare allocation beyond simple numerical representation (utilities).

Understanding how resources are perceived or utilised is imperative to improving the cross-organisation of health and social care services (8). For example, breast cancer

management has adopted a more multidisciplinary approach as a result of the evolution and refinement of screening, diagnosis, staging and treatment (401). Consequently, qualitative research can improve health service delivery by supplementing clinical and cost-effectiveness data with an explanation of the beliefs, behaviours and experiences of those likely subject to its application (402). Such methods have been applied in the oncological setting to explore patient and carer experiences of breast cancer treatment, end of life care and the information needs of breast screening attendees, to improve healthcare delivery (403). Furthermore, decision makers now advocate the involvement of patient preferences in the planning of health services and clinical trials (8). It is therefore important that the evidence informing policy makers includes consideration of patient health behaviour and preferences toward treatment in the values presented (399).

Qualitative research in the field of health economics has conventionally been used to confirm the reliability and descriptive validity of vignettes, discrete choice experiment attributes or economic utility measures (271, 404). Coast (405) argues that robust qualitative methods should move beyond such limited application, to provide scientifically meaningful contribution to economic theory in priority setting and the development of empirical data methodology. Few studies have explored the qualitative reasoning associated with utilities, but those adopting a combined approach were able to explain how subjects viewed and valued health relative to the axioms of expected utility theory (406). Baker and Robinson (407) conducted semi-structured interviews with participants completing a standard gamble. They concluded that the added discussion after eliciting utilities helped to clarify the rationality of the decisions made in valuing health and variation in preferences observed. Similarly, Bronsard et al. (408)

highlighted the value of using in-depth interviews to understand the motivations and behaviours driving willingness to pay preferences as a result of interactions in individual and socio-cultural factors in their empirical study. Preferences toward healthcare are complex and personal, and the application of structured qualitative data collection alongside empirical research is necessary if the differences in the trade-off of benefit and risk are to be fully understood.

9.2.2 Comparison with quantitative research

Qualitative research encompasses a range of techniques. This may include individual or group interviews and the analysis of media, text and observed behaviour, to facilitate the detailed exploration of issues about which informants may hold deep beliefs or attitudes (405). The aim of this section is not to provide a comprehensive list of qualitative methods, but rather an overview of the major challenges and systematic differences in comparison with the empirical methods applied in the preceding chapters, to inform comparison and synthesis of findings from the two approaches. There are methodological and theoretical differences in the recruitment, collection and interpretation of quantitative and qualitative research, all of which are discussed briefly below.

9.2.2.1 Recruitment and sampling

Qualitative research involves the selection of a data sample with relevant characteristics to explore elements of particular interest related to the topic (398). Data can comprise of multiple formats; in-depth individual interviews, focus group discussions facilitated by the researcher, ethnographic observations (observing subjects in their natural

environment) or multi-media data relevant to the subject of interest (409). Unlike empirical analyses, qualitative studies have a stronger emphasis on the quality of the data collected, rather than a specified number required to determine a representative sample of the general population from which statistical generalisations can be inferred (405, 410). As such, a smaller number of cases, if carefully selected to answer the research question, can be equally as insightful for an in-depth analysis of meaning and experience. These cases may be sampled in many different ways (411):

- purposive (cases possess characteristics relevant to the themes of interest)
- opportunistic (recruitment is flexible to opportunities that arise)
- theoretical (samples are driven by theory and iteratively recruited)
- snowball (new recruits are identified through previous participants)
- convenience (cases are selected for their ease of access)

Purposive sampling, whereby women with a diagnosis of DCIS or early breast cancer were actively recruited, was chosen for this research study to provide information-rich data to answer the research question on how benefits and harms of breast cancer treatment are valued by women with actual practical experience.

9.2.2.2 Empirical methods

The purpose of the health economic research will influence the design of the qualitative methods. In choosing which method to use, health economists must consider both economic and qualitative theory (405), dependent on whether the objective of the research is to:

1. Develop or enhance quantitative economic research
2. Help the interpretation of quantitative findings
3. Generate economic understanding

Interviews are the most common approach adopted in qualitative healthcare research. Focus groups or individual interviews are most effective for clarifying the clinical attributes or themes deemed important by those with experience in the area under study (111, 412, 413). This descriptive validation in the development of vignettes or attributes is important for ensuring elicited preferences are meaningful and reflect the genuine utility of the condition described (414). Interviews elicit individual understandings and focus groups elicit shared social understandings (398).

Conversely, semi-structured or “think-aloud" interviewing may be more useful in understanding how people interpret economic instruments (e.g. EQ-5D) (415) or direct approaches such as the time trade-off (144) and standard gamble (416). Economic understanding may be generated by asking participants to explain their thought process whilst completing the interview (413). Such techniques are useful for determining whether the informant has truly understood the task, violated theoretical assumptions or misinterpreted the risks presented in the interpretation of findings (417). Examples of deliberative questioning alongside the application of economic measures have been applied to generate priorities in the provision of breast cancer screening (412), end of life care (418), and policy makers understanding of economic modelling (419).

Participant observation is appropriate for exploring patient behaviour in the context of guiding policy but has been less frequently applied in the framework of health economics (406).

9.2.2.3 Data analysis

The analysis of qualitative research is dependent on the study objectives, epistemological stance of the researcher and nature of the data collected (399). For example, narrative analysis requires interpretation of the language used to describe and narrate experiences of health; discourse analysis focuses on the patterns construed with cultural perspectives and identity (420); thematic analysis is a method applied to identify and explore the themes in relation to the research question (398), of which framework analysis adopts a more rigid and structured analytical approach. In contrast, grounded theory focuses on the development of new theory throughout the data collection process (421). It is therefore important that both the methods and analysis are carefully designed before collaborative qualitative and health economic research in relation to the research question being addressed.

9.2.3 Previous qualitative research on DCIS

Qualitative research investigating the impact of DCIS has increased since the publication of the Marmot publication (3) but has received much less focus than invasive breast disease (422, 423). A recent systematic review by Kim et al. (424) identified 19 qualitative studies relevant to DCIS. Most studies had focused on the psychosocial impact of treatment, the use of decision aids to communicate information about treatment, demographic patterns and prior knowledge of the condition and risk

perception of in situ versus invasive breast disease. Whilst the identified studies demonstrated that most women had little knowledge of DCIS or its associated risks and prognosis, the review highlighted a clear research gap in addressing how the benefits and risks were perceived for low-risk disease, especially in the context of overdiagnosis and overtreatment.

Only three (298, 299, 425) qualitative studies were identified which had explored overdiagnosis in the context of breast cancer screening or DCIS. Hersch et al. (298) used focus groups of 50 Australian women to explore views on overdiagnosis in relation to breast cancer screening. They found that prior awareness of overdiagnosis was poor, attitudes diverse, and women raised concerns about the impact overdiagnosis might have on screening behaviour or management. Similar findings were reported by Waller et al. (425) where British women expressed surprise at the extent of overdiagnosis arising from breast screening programmes. However, women in their study considered the issue of overtreatment to be more important than screening participation or overdiagnosis, in keeping with the higher national screening acceptance and attendance rates observed in the UK. Finally, Nickel et al. (299) explored how treatment preferences for DCIS change with the terminology used to describe the condition. Their qualitative study of 26 Australian women reported higher anxiety and a stronger preference toward more invasive treatment if the phrase *"pre-invasive breast cancer"* was used instead of *"abnormal cells"* to describe in situ disease. Women expressed an interest in active monitoring when educated about its potential to reduce unnecessary treatment but felt that more frequent monitoring (e.g. six-monthly) and better education were needed if the strategy were to be implemented.

9.2.4 Rationale for conducting the qualitative interviews

Little research has been done examining how treatment decisions are made following diagnosis of low-risk DCIS. Whilst the findings from the studies identified by the review (424) suggest prior knowledge of overdiagnosis is poor, less is known about how women weigh up the benefits and harms in the decision. Similarly, there is little evidence regarding the acceptability of active monitoring in women with DCIS. Attitudes toward active monitoring in breast cancer have only been explored in the context of the terminology used or hypothetical anxiety (275). Robust evidence is required to explore how the trade-off might impact upon utility among women who are adequately informed about the decision.

Qualitative interviews were conducted to explore the factors which influence treatment decisions for low-risk DCIS specifically. This included investigation of how attitudes towards overdiagnosis, risk and active monitoring might impact the utility of the treatment chosen. It was considered important to collect data from women with experience of the treatments as they were likely to have insight of the trade-off associated with treating the sequelae of screening and overdiagnosis.

9.3 Aim

The aim of the qualitative study was to examine how women describe their preferences following diagnosis and treatment. In particular, the study aimed to explore in depth the factors influencing how women value the benefits and harms associated with treatment, including the risk of overdiagnosis, and the acceptability of active monitoring as an alternative management option.

The three main objectives of the qualitative research were to:

1) Elicit the range of views relating to the overdiagnosis and treatment of low-risk DCIS
2) Describe how the benefits and harms associated with the sequelae are perceived
3) Explain why differences in utility may exist between patients for low-risk DCIS

9.4 Methods

9.4.1 Study design

A qualitative study was conducted to investigate the factors influencing utilities in the treatment of low-risk DCIS. The theoretical framework applied was one of constructivism-interpretivism (426), which assumes that there are multiple realities, and that the reality or meaning captured is one created through combined understandings of the interviewer and interviewees (427). Specifically, a constructivism-interpretivism approach does not generally begin with a theory but rather aims to generate or inductively develop a theory or pattern of meanings or themes through gaining an understanding of human experiences from the participants' views and the researcher's interpretation of such experiences (426). This perspective was adopted over other qualitative frameworks (e.g. grounded theory) to facilitate systematic comparison to define common themes which might explain, in part, the individual heterogeneity in utility rather than generating new qualitative theory. As seen in Chapter 6, the variation in utilities from the empirical analysis could not be explained by socio-demographic characteristics, suggesting that women's preferences were likely linked to their own understanding or experiences of health risk unanticipated by the researcher. Similarly,

all women's opinions are deemed meaningful and valid whether their own perception of risk reflects the truth or not (428), and constructivism-interpretivism adopts a process of imposing purpose in order to make the best possible assimilation to which such themes are taken to belong to (427).

A scoping review (424) of qualitative studies in DCIS was used to inform the methods applied in this research. The review suggested that focus groups or semi-structured interviews were the most appropriate vehicles in which to gain perspective on women's thoughts, opinions and behaviour toward breast cancer screening and treatment. Whilst focus groups facilitate participant interaction (399), they are more effective for exploring general attitudes toward treatment. In-depth interviews offer opportunities for women to voice their understanding of DCIS within the context of their own personal history, family circumstances or decision making (429).

This thesis used individual interviews to explore how the benefits and harms of DCIS management were conceptualised. Interviews produced a rich and detailed account of each woman's understanding and perception of DCIS treatment rather than shared social meanings. Collecting data individually rather than in focus groups enabled a more in-depth understanding of how women traded the benefits and harms of treatment, without potential bias or confrontation (398). Social interaction in focus groups can reveal shared ideas and social norms but this was not the focus of this study.

9.4.2 Interviews and topic guide

Interviews were semi-structured and informed by an interview guide designed to facilitate the discussion. The guide was designed to cover a range of issues deemed likely to explain the variation in preferences related to management of DCIS, whilst also leaving room for unanticipated factors to be raised. This included attitudes toward disease progression and risk of recurrence, overdiagnosis, active monitoring and potential side-effects of treatment. Opening questions were kept purposefully broad to enable women to expand on the advantages and disadvantages considered most important to them, with specific prompts and structured questions suggested if relevant issues to the research question (e.g. overdiagnosis) had not already been discussed by the interviewee (398). The guide was refined with the input of clinical experts, qualitative health researchers and BCNA Review & Survey group members. A copy of the interview guide is provided in Appendix 13.

All interviews were conducted face-to-face to encourage the establishment of rapport early in the interview so that women felt comfortable discussing the sensitive issues related to breast cancer treatment. This was deemed especially important in the context of overdiagnosis because as Reinharz (430) suggests, preparation and trust are critical to establishing free discussion of the issues important to both participants and the researcher.

9.4.3 Sampling and recruitment

9.4.3.1 Population

Cresswell (431) found that the reliability and depth of their qualitative research was significantly improved by selecting participants with first-hand experience of the subject under investigation. Women with low-risk breast cancer (defined as either DCIS or early invasive breast cancer) were thus selected for interview based on their unique experience and understanding of the benefits and harms associated with the sequelae of breast screening. Firestone (432) argued that the validity of qualitative findings is also dependent on ensuring a range of socio-economic perspectives relevant to the topic are reflected. The sample frame was therefore stratified by the recruitment agencies (Breast Cancer Network Australia and Lifepool) prior to the invitation of participants, so that women from a range of socio-demographic backgrounds and ages were represented in those with experience of breast cancer treatment.

9.4.3.2 Selection

Purposive sampling was used to select individuals with experience of in situ or early breast cancer. This technique is widely adopted in economic qualitative research to identify suitable participants likely to have knowledge or experience of the phenomenon of interest (433). Women with DCIS or early invasive breast cancer were specified for inclusion to enable the real-life issues surrounding treatment decisions to be explored by those with experience of localised disease (434). Women with no breast cancer, advanced or metastatic breast cancer, and those unable to provide informed consent or understand and converse in English were excluded.

9.4.3.3 Recruitment

Women were recruited from the patient cohort in the main empirical analysis (via the Breast Cancer Network Australia Review & Survey group and Lifepool registries). After completing the utility questionnaire, women with a history of localised or in situ breast cancer were invited for a second interview to discuss their preferences in greater detail. The details of interested participants were tabulated in a spreadsheet (Microsoft Excel 2018), and a random number matrix was used to generate a sample list and order to invite for interview. A convenient date and time for interview was arranged by email or telephone for those selected. Interviews were conducted in both affluent and deprived suburbs of Melbourne to ensure diversity of socio-demographic characteristics.

9.4.3.4 Data collection

Interviews were carried out face-to-face in a private room at the university or a local community facility in Melbourne. The venues were selected for comfort and convenience as it was felt that a quiet or relaxed environment would improve interpersonal communication (399). Two interviews were conducted via telecommunication to accommodate those unable to travel due to extremes of distance or disability. Although it is argued that telephone interviews may hinder rapport (435), the interviewer had already established an initial relationship with the women in the quantitative phase of the research which included face-to-face engagement.

The interviews were conducted by a single researcher (HB) who had clinical experience of working in breast cancer surgery. This was made explicit to participants for consistency but may have changed the dynamic of the interview toward a clinician-

patient focus. Participants were required to give written consent prior to commencing the interview and were asked to complete a short socio-demographic questionnaire after finishing the discussion. Interviews followed a semi-structured format to cover the topics perceived to be important to DCIS but remained flexible to allow women to highlight any additional issues deemed relevant to the discussion. The questions were structured using a funnelled approach which encourages logical progression from broad to specific inquiry (436). Initial questions were phrased openly to encourage women to talk freely about their preferences toward treatment, with follow-up questions and probes used to gain clarification or a deeper understanding of the pertinent issues.

9.4.3.5 Sample size

Sample size was not determined *a priori* because qualitative estimates vary enormously depending on the scope and purpose of the research. All opinions in qualitative research are considered valid and thus there are no tests of statistical significance to determine when meaningful conclusions are likely to be met (437). Instead, data were collected until the basis of theoretical saturation, whereby no new theoretical insights or themes were identified through subsequent interviewing (438). This was determined as the point at which participants consistently reported similar accounts, such as repetition of knowledge relating to DCIS, overdiagnosis or the side-effects of treatment. This technique aligns with the principles of sample adequacy described by Bowen who recommends the collection of data until *"the same comments are heard from different participants in different places,"* ((439): p.148), to ensure complete and truthful theoretical representation of the condition is met.

9.4.4 Data transcription and analysis

All data were managed using NVivo V11 software. Interviews were audio-recorded with participant consent, anonymised and transcribed verbatim by the research student (HB) to ensure a reliable account of participants' views.

A framework analysis was adopted to identify potential themes, code and analyse the data. This approach uses a systematic process to identify, analyse and report trends within the data, and was chosen so that the issues most relevant to women could be stratified into collective domains using a transparent approach (398). Furthermore, the framework method is useful for evaluating data where participants may have contrasting opinions or experiences (440), and provides a holistic overview of themes that may not have been anticipated by the researcher but are deemed important to the analysis (441). This caveat was essential given the aim was to understand women's preferences for managing DCIS. Certain codes were induced by the study objectives (e.g. risk perception following treatment, side-effects of treatment), whereas others (e.g. knowledge of DCIS, personal circumstances) were deduced directly from the data.
The five-stage process described by Ritchie and Spencer (441) was used to analyse the data, and is summarised in Figure 9.1.

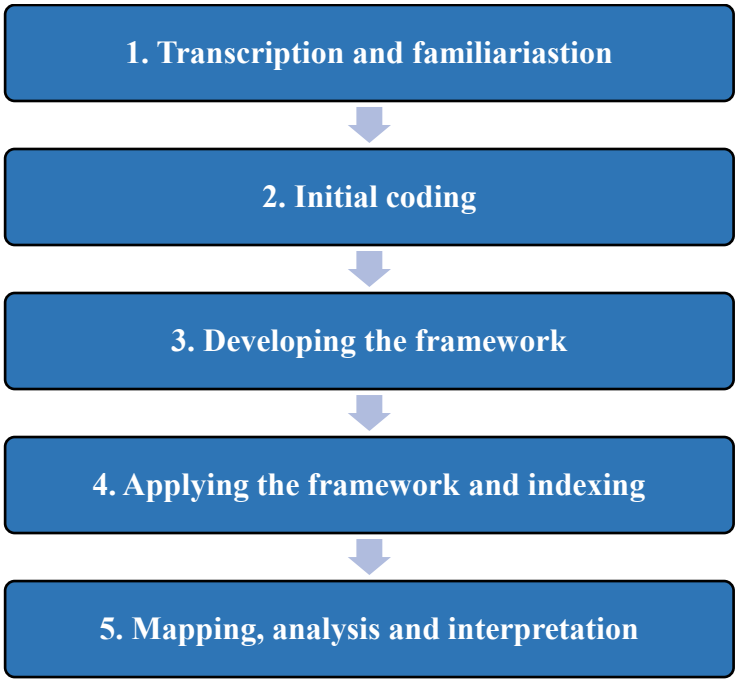


Figure 9.1: A flow diagram of the five-stage process for framework analysis

To ensure consistency in the presentation of data, a single researcher (HB) was responsible for carrying out transcription. Transcripts were subsequently checked against audio-recordings to resolve any inconsistencies and re-read multiple times to improve familiarisation with the data. This required the researcher to become 'fully immersed' in the data and initial impressions, strong views or contrasting opinions toward treatment were noted in margins (441).

Open coding of the transcripts was undertaken in the second step of the analysis using a line-by-line approach. Initial ideas were noted, coded and subsequently grouped into potentially relevant themes. This included single words or phrases and provided a starting point on how the themes could be grouped in the next stage of the analysis (440). As key concepts related to DCIS were identified, the process became more

refined to encompass larger sentences and paragraphs, facilitating a deeper understanding of the underlying issues and themes (421). This was initially performed both manually and electronically for thoroughness, whereby interesting excerpts were highlighted and allocated an initial label. An example of the first codes from one of the interview transcripts is provided in Figure 9.2. Here the patient talks about some of the personal factors and concerns related to treatment. The underlined sections emphasise the researcher's initial thoughts on important aspects worth noting, which were later refined and labelled under 'practical considerations' as part of the 'psychosocial impact of treatment' code.

Following on from the initial coding, recurrent themes were used to develop a thematic framework to sort the data for analysis (Appendix 14). Codes were iteratively refined and re-grouped by three researchers (the research student and two independent academics with experience in qualitative research) to ensure consistency in the labels derived. Where differences in interpretation existed, further clarification was sought to revise the framework. For example, where one code was labelled as 'convenience', another researcher categorised it as 'practical considerations.' There was also overlap between some of the initial themes relating to overdiagnosis, worry about progression and active monitoring which were re-grouped into a single overarching code of 'uncertainties about the benefits and harms of treatment' which better encompassed all three labels. The final framework encompassed four codes split into 14 nodes or themes, as is shown in Figure 9.3.

The final framework matrix was subsequently applied to each transcript as part of the fourth stage of analysis using NVivo version 11. Indexing involved systematically reviewing each interview a second time and applying the index codes to an appropriate passage of text (442). Data were charted against existing labels and regrouped under the thematic framework. Specifically, this required 'lifting' the data from its original context and rearranging it within the themes identified during indexing. This facilitated early interpretation of the main arguments generated within each theme, particularly relating to participants' opinions around uncertainty of overdiagnosis.

Once all the data were charted using the analytical framework, excerpts from transcripts were applied to each code so that the preferences of the participants in the study were reflected in the analysis of the data. Final interpretation was generated by reviewing the matrix and mapping individual interviews to exemplify the important underlying issues (441). This stage of the analysis was critical to answering the objectives of the research, with the support of evidence quoted directly from the interviews.

Figure 9.2: Initial coding notes and labels

Coding labels		**Qualitative interview 18 (Q18)**	*Notes and ideas*
	48	***I: Would you think about the idea of monitoring? Do you think people would***	
	49	***think about it?***	
	50		Urban versus rural differences
Rural divide	51	R: Like having a mammogram every year to keep an eye on it rather than having	
	52	the surgery? Well yes, I mean. Maybe not if you live in the country, trying to get	
	53	there. But for me to get monitored by my GP, it's fifteen minutes from my house.	Treatment preferences dependent on accessibility, location.
Convenience	54	Going once a year is no big deal.	
	55		
	56	***I: You don't think you'd be overly anxious or anything?***	
	57		
	58	R: I don't think so because it's pre-cancerous. I think if you were monitoring an	Differentiation (risk) between DCIS and invasive breast cancer
Perception of risk	59	invasive lump then that's different. I'd quite like to know how long my DCIS had	
	60	been there. I know I asked at the time and they said quite a while. Certainly, within	
	61	two years. I mean, I had a mammogram two years before and there was nothing	
	62	there then.	
	63		
	64	***I: Do you think people would rather monitor than have the surgery? Not in your***	Side-effects of surgery debilitating
	65	***case but if it had been low risk.***	
Side-effects	66		Frequency/ follow-up of monitoring likely important to its acceptability
	67	R: I would have thought people would rather not have surgery and would be	
Frequency of monitoring	68	monitored. I mean, surgery is severe and um…look, all I can say is if I was given	
	69	the information that it's pre-cancerous, we'll keep a close eye on it…it might mean	
	70	you have to go here in a year…I mean, how long would they monitor it for?	

Table 9.1: Coding framework matrix

CODE	DESCRIPTION
Understanding of DCIS	
Awareness	*Prior knowledge of the condition, likelihood of having any experience or understanding of the condition*
Terminology	*Effect of terminology used on level of concern or treatment preferences, discrepancy between diagnosis and treatment*
Relationship to invasive breast cancer	*Expectations of DCIS as a breast cancer, discrepancy between expectation and reality, ability to differentiate from invasive breast cancer*
Perception of risk	*Perception of the risk associated, urgency of treatment, uncertainty in risks with spectrum of disease, connotations with recurrence and death*
Information and communication about the benefits and risks	
Sources of information	*Friends, family, patients, BCNA support networks, decision tools, leaflets, social media, "Dr Google" , framing, research and technological advancements*
Feeling involved with care	*Choice offered, personalised treatment plans, time to process, information balance (benefits and harms discussed)*
Role of the doctor	*Education, explanations, trust, decisional conflict, heterogeneity in practice, power/bias, clinical recommendation*
Uncertainties about the benefits and harms of treatment	
Attitude toward overdiagnosis and overtreatment	*Attitudes, prior knowledge and worry about lack of awareness, scepticism (non-believers) vs cynicism (ulterior motive, media influence), paradox of overtreatment, ethical concerns (women have a 'right to know' and be treated), regret they were not informed/unnecessary treatment*
Implications and worry about progression	*Challenges in decision making due to unknown, fear and anxiety, regret,), impact on breast screening, risk averse*
Active monitoring and de-escalation	*Acceptability, attitude, heterogeneity in views on uptake (positive vs negative), risk threshold, frequency of monitoring, costs (insurance), follow-up (safety-netting), comparison to other screening programmes (e.g. cervical & prostate cancer), flexibility to change mind*
Psychosocial impact of treatment	
Quality of life	*Anticipated impact of side-effects on life and relationships, emotional response, sense of finality*
Social meanings	*Purpose of breasts, body image, identity as a woman, sexuality, "normality", stigma, stereotypes (positive and negative)*
Practical considerations	*Age, health, family (children or breastfeeding infants), support from partner or friends, financial, cover (access to private vs. public healthcare), personality and attitude toward risk, accessibility, rural-urban divide*
Fitting in with role and responsibilities	*Factors more than just convenience, how decisions made to fit in around life and role/sensibility as a woman or mother, impact to working or family practice*

9.4.5 Ethical considerations and validation

Ethics approval was sought to conduct additional qualitative interviews from the University of Melbourne Health Sciences Ethics Sub-Committee (ID: 1750252) and may be found in Appendix 15.

Anticipated ethical challenges were considered prior to commencing the research. All interviews were anonymised to safeguard patient identity and confidentiality. Data were securely secured on a password protected university network that could only be accessed by the research student, and consent forms were stored separately to the transcripts. Patients were signposted to sources of further information and support at the end of the interview. No financial incentive was provided.

A written summary of the results was provided to the registries for circulation to participants. This enabled interviewees to raise any issues identified in the subjective interpretation of the transcripts and ensure accurate interpretation of meaning by the researcher (443). Patterns of convergence between the interviews and original utilities were also undertaken to corroborate the overall interpretation of results (444).

The results are reported against *'Consolidated criteria for reporting qualitative research'* (COREQ) (445) and *'Standards for reporting qualitative research'* (SRQR) (446) guidelines, in keeping with good qualitative research practice (Appendix 16).

9.5 Results

9.5.1 Sample characteristics

Twenty-six women were interviewed between May and September 2018, at which point saturation was achieved and recruitment was stopped. Interviews lasted an average of 39 minutes (range 17-70 minutes) and all women completed the interview in full. Most women had a diagnosis of DCIS (n=13) or early invasive breast cancer (n=9), but few had mixed (n=3) or uncertain (n=1) in situ diagnoses, all within the last 10 years.

The characteristics of the women sampled are given in Table 9.2. Patients ranged in age (43-73 years) and experience of breast cancer treatment. Whilst most women had undergone surgical management (n=24), additional treatment with adjuvant radiotherapy, endocrine therapy or surveillance was more variable. Three-quarters (n=20) of the sample reported little or no worry about breast cancer and two-thirds (n=17) of women would have considered active monitoring for low-risk DCIS had it been a viable option.

9.5.2 Framework analysis

The framework analysis identified four main themes arising from the data relating to the challenges and treatment decisions associated with DCIS:

(1) Understanding of DCIS

(2) Information and communication about the benefits and risks

(3) Uncertainties about the benefits and harms of treatment

(4) Psychosocial impact of treatment

Direct quotations are presented verbatim (with the use of ellipses to represent omitted text) to illustrate the heterogeneity in patient responses highlighted within each theme.

Table 9.2: A summary of participant socio-demographic characteristics (n=26)

ID	Age (years)	Diagnosis	Treatment	Worry about BC	Consider AM?
1	66	DCIS	Mastectomy, endocrine	No worry	Yes
2	63	DCIS	BCS/RT	No worry	Unsure
3	60	Early IBC	BCS/RT, Mastectomy, endocrine	Some	Unsure
4	62	DCIS	BCS, Mastectomy	Little	No
5	65	Early IBC	BCS/RT, endocrine	Little	Yes
6	67	DCIS	Mastectomy	No worry	No
7	57	DCIS	Mastectomy	Some	No
8	53	Early IBC	Mastectomy	Little	Yes
9	70	DCIS/IBC	BCS, Mastectomy	Some	Yes
10	59	Unsure	Monitoring	Little	Yes
11	62	DCIS	Mastectomy, endocrine	Some	Yes
12	56	Early IBC	Mastectomy, endocrine	Little	Yes
13	66	Early IBC	BCS/RT, Mastectomy	No worry	No
14	72	Early IBC	BCS	Little	Yes
15	73	Early IBC	BCS	Little	Yes
16	71	Early IBC	Mastectomy, endocrine	No worry	Yes
17	72	DCIS	BCS	No worry	Yes
18	66	DCIS	BCS/RT, endocrine	Little	Yes
19	60	DCIS/IBC	BCS/RT, Mastectomy, endocrine	Little	Yes
20	51	Early IBC	Mastectomy, endocrine	Some	No
21	43	DCIS	Mastectomy	No worry	Unsure
22	69	DCIS	BCS/RT, Mastectomy	No worry	No
23	71	DCIS	BCS/RT	Little	Yes
24	61	DCIS/IBC	BCS, Mastectomy, endocrine	Little	Yes
25	65	DCIS	BCS/RT, Mastectomy	Some	Yes
26	59	DCIS	Monitoring	Little	Yes

Legend: AM: active monitoring; BC: breast cancer; BCS: Breast conserving surgery; DCIS: ductal carcinoma in situ; IBC: invasive breast cancer; RT: radiotherapy

Theme 1: Understanding of DCIS

Awareness of DCIS and relationship to invasive breast cancer

There was wide variation in understanding of DCIS among those interviewed and asked to value utility. Whilst most women had heard of invasive breast cancer, prior awareness of in situ disease was limited and many women did not know that it was a different condition.

"I don't think I would have been abnormal in that. I think people talk breast cancer or lumps rather than thinking DCIS." [P11]

"I suppose you just think of the big picture. You have breast cancer and you have your breast removed." [P23]

Similarly, most women had not heard of DCIS before attending breast cancer screening. Although patients diagnosed with DCIS felt their knowledge was vastly improved, several remained unsure about the nature of the condition despite having undergone invasive surgical treatment.

"The information I had was that I had breast cancer. It wasn't until years later from my own reading that I really understood that it was a pre-cancer." [P9]

Understanding of DCIS was subsequently linked with how women interpreted and valued the benefits and risks in the vignettes. Those who felt that they had more insight into the pre-invasive nature of the condition tended to value less intensive treatment regimens more favourably.

"It's a risk factor. Pre-cancer sounds like you will get it anyway. That's a leading thing. I just call [DCIS] *a risk factor. So, if I tell people about it, I tell people I have a cancer risk factor and I have it monitored."* [P26]

However, other women felt that there was discrepancy between their expectations of the condition and reality. Women expressed concern that their limited understanding was likely related to a preference toward more intensive treatment. Consequently, attitudes toward treatment were largely representative of those exhibited among women with invasive breast cancer, and a higher quality of life reported for surgery with radiotherapy or mastectomy.

*"It's certainly pre-cancerous but it's treated like…mine was treated so seriously, it felt like it was a cancer…the mastectomy was a big deal." [*P18]

"I know it's changes in the cells. They haven't spread out of the breast ducts and that there are different ideas about whether it will go on to cause cancer. I don't think it does in everyone, but my sister certainly had it and she had her breasts off because of it." [P20]

Attitude, risk threshold and perception of risk

Several women with DCIS were confused about whether they had a cancer that could result in death. Risk perception was diverse, and interviewees found it difficult to accept the perceived paradox between having a pre-cancerous condition and the extensive surgery that is sometimes required.

"It's cancer but it doesn't need to be treated… or it's diagnosed as cancer but actually it isn't? Is that sort of it?" [P5]

"I think that's scary because if you are a lay person and you know nothing, you go to a breast specialist and they say it's nothing to worry about. You want to believe it." [P20]

"Look some women have DCIS. They sort of elect for a mastectomy because they think they don't want to take any chances. I don't want to get full on breast cancer. I don't want to lose my life." [P8]

Subsequently, patients largely perceived the risk of progression and mortality associated with DCIS the same as that observed in invasive breast cancer. Nearly a third of women felt that DCIS was likely life-threatening and required immediate surgical intervention to avoid an increased risk of death or morbidity.

"I think that it's more [about] communicating with DCIS. How it's different to invasive, explaining the risks. I think the whole thing's there about being rushed to make decision on DCIS… the anxiety of thinking you've got to do something straight away." [P4]

"*I do think the majority of people who haven't had much to do with cancer do fear it. It's a sentence. They think you're just going to die and there's no alternative. They're just… get it out of my body type [of] thing!"* [P9]

Risk perception was also deemed important to women and played a key part in treatment decisions. Two thirds of the women interviewed suggested that they might have adopted a more conservative approach had they had an explicit diagnosis of low-risk disease.

"I think that being high grade made it easier for me to decide on the treatment. I think if I had been low, it would have been much more difficult to decide. [P4]

"I certainly think if my grade had been low and I think from those scenarios that I would take [less] – especially having had a mastectomy – I wouldn't be the type to say get it out, get it off." [P18]

In contrast, those with risk averse personalities were reluctant to adopt less invasive treatment regardless of risk categorisation, and this subsequently affected the quality of life reported.

"I don't think there's any such thing as overtreatment if it's going to make you 100% sure. I would hate to be overtreated with chemo[therapy] or something, if that was unnecessary. But just having the options of getting my boobs off, I don't think that's overtreatment. I just think that's safe." [P21]

"A lot of it will make you feel anxious if you're not sure. But then coming from somewhere with the anxiety issues anyway I don't know whether general people in the community would feel so anxious [about de-escalating treatment]." [P6]

Terminology used to describe DCIS

All women reported high concern about DCIS regardless of the terminology used.

"I mean people with this minor DCIS may still say no, get rid of it! I don't want it. You know? That's their prerogative because you know if you've got DCIS the chances are it could get worse." [P24]

However, stronger emotional reactions and levels of concern were reported by women in response to doctors using cancer terminology, even though this was not included in the vignettes.

"As soon as you hear the big 'c' huge emotion comes in. In fact, you might not even be listening or you're not hearing anything because you're just in this kind of panic." [P10]

"It was just unfortunate that they had used the word cancer, I think if they had said DCIS and explained what it was, I think I would have been less frightened." [P23]

Intensity of treatment was positively correlated with prior use of cancer terminology. Patients felt they were more likely to accept toxic interventions in the trade-off against the risk of unnecessary treatment because of the fear induced by the term 'cancer'.

"All it takes is a doctor saying that it could develop into an infiltrating cancer and as soon as they say that, they want it out." [P14]

"I do think that until you change the name, it's not going to be very easy to change the mentality. People just jump into that word…carcinoma... ah, you've got cancer." [P26]

Conversely, women expressed shock at being told that they needed a mastectomy for a condition which was described as non-threatening and felt the terminology should reflect the non-invasive nature of the disease and likely less intensive treatment required, as described in the vignettes.

"He kept talking about this calcification and I thought I've got calcification…so what? Then he said we recommend a mastectomy and I said what! Nothing he had said lead me to think it was that serious. I was in shock." [P26]

Theme 2: Information and communication about benefits and risks

Role of the clinician

The role of the clinician had a major influence on how women valued the utility of treatment. Women trusted the advice given by their surgeon and were more likely to report higher values for adjuvant radiotherapy or mastectomy if they had been recommended it, even if it meant undergoing more invasive treatment than anticipated.

"I just went with what the doctor said. I didn't want to have radiotherapy, but it was only when I saw the radiotherapist who told me, well you really need to…. I was just sort of really in denial that I needed it. In some way, I kind of feel it's quite good the doctor sort of tells you what you need to do because then you do what's best for you." [P2]

"I'd be saying if you think that's the better option or that's what's in my best interests, then this is the way we should go, and I'd go with it." [P17]

Other women reported decisional conflict or differences in clinical practice between the healthcare personnel involved in managing DCIS, which made the process of valuing treatment more challenging.

"You might have two women who go to two different doctors but are treated differently when they've got DCIS too. I think knowing why you have the treatment is

important. Yeah, because some people – some specialists – are like, yeah, you need to do it but then others are like no you don't!" [P7]

There was some concern that clinicians might have a vested interest in a particular outcome or treatment. A few women were worried that this may have biased their values for more intensive treatment, particularly if they had not been provided with the full information and likely outcomes of all available interventions. This resulted in them opting for more invasive surgery, even if they had reservations about undergoing a mastectomy or radiotherapy.

"Well there was no choice. It was just like this is the treatment. I suppose I could have said no treatment, but the percentages of recurring were too high to really contemplate that you wouldn't have treatment." [P2]

"I suppose I come back to that surgeon of mine who said if you don't do this then you'll die. So, I don't know. Do I regret it? It would have been nice to do less but then again I'm here, aren't I?" [P8]

Patient involvement in decisions

Feeling informed and involved in decision making gave women additional confidence in the treatment utilities they reported. Women who recounted having had open and realistic discussions with their treatment team about the true risks and side-effects of treatment tended to value breast conserving surgery more favourably compared to those without any clinical experience.

"Having someone who is a professional who you trust, who you can have that big talk to. Being able to make that decision. They can say that's not what we would recommend at your stage, but it is ultimately your decision. Not just saying you have option A and that's the end of the discussion." [P9]

Women also highly valued personalised treatment plans and information. When patients felt listened to or involved in the decision process, they were generally more accepting of the treatment planned, even if there was a chance that it had ultimately been unnecessary.

"I really appreciated that understanding, it wasn't just my body, it was mind as well that she was treating, if you like?" [P5]

"You're cutting part of your body off but it's going to save your life. The sensible part wins out and I'm happy with what I've done." [P21]

However, almost all participants discussed issues related to communication. It was suggested that medical professionals needed to provide clearer explanation about the non-invasive nature of in situ disease and a balanced view of each of the treatment options with patients and their families.

"He just told me you've got breast cancer. We'll send you to a surgeon. There was no DCIS at that stage. People are luckier – or perhaps unluckier – if they use 'Dr Google' wrongly but at that stage you had to go to the library to find out, so my information was limited. I definitely think information, education, it's so important." [P9]

"*I mean, you don't want to scare people, you just want to be aware. Information and education are so important, so you can make an informed decision."* [P7]

Sources of information

Most interviewees reported difficulties in accessing information specifically about DCIS. Consequently, women were surprised to learn about the risks of overdiagnosis or less invasive treatments for localised low-risk DCIS when valuing the health states.

"I remember reading that article in Time and thinking hey! Why didn't someone tell me that I didn't necessarily have to have surgery and radiation? Even though maybe some aspects of that article were wrong it's still made me think I wasn't told any of this." [P23]

Advice about treatment was largely sought from family, friends or discussion with other patients on breast cancer network forums but were found to be more relevant to invasive cancer rather than DCIS. Resources for DCIS were deemed inadequate or disappointing and thus women felt obliged to be led by their clinician due to their limited information about the condition.

"I guess you could go online to those online networks – like the BCNA – and ask women who've had it. You know, find out from women who have gone through those things like the chemo[therapy]. Like three years down the track what are the effects? That can be reassuring. Some people might trust that information more because that's more real to them, but you can take things the wrong way and make a decision based on that." [P4]

Only two-thirds of breast cancer patients admitted applying clinical data in the vignettes to inform their evaluation of treatment, indicating that the decision had been more of an emotional one than rational. Although these women were largely guided by clinical experience, they suggested that the decision to have more invasive surgery or radiotherapy was dominated by the initial shock of the diagnosis and that they may have valued less invasive therapy more favourably had they been given the time to reflect and process the description of low-risk DCIS.

"There's sort of this hard fact information and this overlay of emotion and perspective. It's not just straight facts and making these sorts of decision." [P3]

"Even when you've got really good information that this treatment is not necessary, you're still up against that human emotion that something needs to be done." [P5]

Theme 3: Uncertainties about the benefits and harms

Attitudes toward conventional treatment

Most women were familiar with conventional breast cancer treatment. Women generally felt that they had a reasonable understanding of the benefits and harms associated in terms of the trade-off between the side-effects versus a reduced chance of recurrence or spread.

"I think if I'd had to have a mastectomy, I would have felt very self-conscious about that and that would have affected me more psychologically." [P2]

Breast conserving surgery with radiotherapy was valued worst by those with traumatic experiences of the side-effects from treatment. Women with extensive radiotherapy burns

reported the lowest individual utilities as they felt that the reduced risk of recurrence was not worth the morbidity personally experienced.

"After about four weeks [of radiotherapy] it was awful. I used to take the dressings off under the shower. I would stand in the shower crying. And then it takes time for the effect to stop. It does go away, but I used to have to drive to my appointment and it was so painful." [P15]

Reaction to overdiagnosis and overtreatment

Interestingly, unlike the conventional risks of treatment most women had not heard of overdiagnosis or overtreatment prior to taking part in the research. Whilst most women indicated that it was unlikely to influence their decision retrospectively, several women expressed concern that they had not been informed about the harm and indicated that they would likely have opted for a more conservative approach had they known about it.

"It was over and done with. I couldn't change it… but I just thought this is wrong. I should have gotten more information." [P23]

"If it was unnecessary treatment, I would regret it but on the other hand I don't think there's enough research." [P16]

Some women had heard about overdiagnosis in the context of other screening programmes (prostate and cervical). Women expressed an interest in whether this information was being discussed with women attending breast cancer screening or patients diagnosed with low-risk DCIS.

"What about how they manage early stage prostate cancer? What do they do? What are they calling that? It seems that a lot of men are happy doing…is it called watchful waiting?" [P15]

In general, women felt that overdiagnosis itself was not an issue because women had a right to know and be treated. However, there was some empathy with decision makers in that unnecessary treatment was likely to instil additional disutility (as reflected in the utilities reported) and costs to both patients and the healthcare system. A few women also expressed anger at the idea of having undergone unnecessary treatment and regret from the detriment to their quality of life.

"I was reading in the paper about it. That's sometimes it's not necessary to have an operation at all. That sometimes you can watch and wait. I stewed on that for a few years actually! Because that really did concern me, the fact that I might have had this operation for no reason." [P6]

About a third of women were more cynical about overdiagnosis and questioned whether there were ulterior motives behind the concept to cut costs or defer blame. They felt that it lessened the importance of their own diagnosis and expressed concerns about the potential negative implications that the term might have on future screening behaviour or health insurance.

"I think that's absolutely abhorrent. It really bothers me that these people are saying you don't need it because what we find is not worth worrying out. Well, even if it's only 1%

of the time, for that particular person it means a lot. I feel very strongly about that I must admit!" [P14]

"I saw this story and I thought, what's going on? For some reason her health insurance was a problem. They wouldn't pay because they said it [DCIS] wasn't cancer." [P19]

These women tended to validate their own treatment in the utilities elicited and did not feel that the inclusion of overdiagnosis in the descriptions was likely to sway their valuation. Paradoxically, women who felt lucky to have been diagnosed and treated with DCIS also reported confusion as to how treating an abnormality could be considered a negative outcome.

"I think I'm lucky that I guess we caught all those cells, but you never really know. The fact that we have DCIS here, I would lie awake at night thinking tonight's the nights it's going to pop through that duct." [P14]

"Part of me does think, well, maybe we are [over diagnosing] but I would rather be over-diagnosed and lose my breasts and live, than be playing that game of wait and see." [P20]

Implications and worry about progression

Worry about progression was a key influence in how women perceived the utility of the various treatments for DCIS. Valuing treatment was found to be challenging due to the fear and anxiety surrounding the unknown risk of invasive cancer progression. An estimated rate

of 10% invasive progression in ten years was viewed as an acceptable trade-off by most women for interventions aiming to reduce the harms of overdiagnosis.

"Different people have a very different attitude toward risk. These are the sorts of things that would worry me. If it's a 10% risk profile then I'm not going to worry about it happening. If it's a 90% chance, then fair enough [have surgery]." [P10]

"Some people would roll the dice with monitoring [at that risk level] because they don't want to be disfigured or go through all that experience." [P11]

However, most women expressed significant concerns where the risk of progression given was over 20% and this anxiety had measurable negative impact on quality of life and utility.

"Well, the fact that I was having the lumpectomy, I wasn't just totally watching and waiting. I was doing something to minimise the risk of return." [P13]

Perception of active monitoring and de-escalation

Despite initial concerns about the uncertainty associated with DCIS progression, many women expressed an interest in active monitoring. Two thirds suggested that they would be willing to consider active monitoring over conventional management for low-risk DCIS, if it was shown to be clinically safe and effective in trials as it had been for prostate or cervical cancer.

"In actual fact that's interesting because when you think of the fact that we have smears, it's the same principal. You can get called back for that, but nobody gets in a knot

about that. You can go on and have a core biopsy and you still don't get tied up. It's really strange that." [P9]

"My brother has got the equivalent prostate cancer and he actually has a blood test every three months and a biopsy then so he's actually having actually much closer monitoring than I am. But he's taking it all in his stride…it's better than the alternative. [P5]

Similarly, several women were happy to trade-off the risk of progression associated with de-escalation to minimise the risks of emotional and physical debilitation from unnecessary intensive treatment in the valuation of health states with overdiagnosis.

"I mean if it's low-risk then monitoring is a good way to go because the surgeries are pretty drastic. I mean, they are a pretty drastic way to go if you don't have invasive breast cancer." [P25]

Nonetheless, a higher frequency of monitoring or follow-up was preferred to provide further reassurance, and adequate education and safety netting was deemed necessary by patients before routinely offering such strategies to other women.

"Well if you're closely monitored… twelve months seems a long time to me but that might be ignorance on my part about the time it takes most cancers to develop. That would be something I would have to find out to feel comfortable." [P2]

"I think being monitored is important. Feeling you've got the security of the next appointment. Maybe doing breast examination every six months and then the mammogram at twelve months. Six monthly check ups might be better for some women if they worry." [P15]

A few sub-groups of women indicated that they would not opt for active monitoring regardless of the risk of overdiagnosis. These women predominantly had a history of invasive breast cancer or exhibited a higher threshold level of worry about cancer. The low utilities reported by these individuals for active monitoring were influenced by a perceived lack of finality, fear of invasive progression and the potential implications of surveillance on finances and screening anxiety.

"My mentality is I wanted it out. I wanted it done. I wanted to be able to live my life knowing that I haven't got a problem in my body. Waiting for a ticking time bomb." [P17]

"Well, the fact that I was having the lumpectomy, I wasn't just totally watching and waiting. I was doing something to minimise the risk of return or becoming, you know?" [P13]

Theme 4: Psychosocial impact of treatment

Quality of life

Anticipated impact of treatment upon quality of life was a key factor in the valuation of treatment for DCIS. Women who had experienced severe side-effects during their own treatment were more likely to report lower utilities for the associated health states.

"I didn't know I'd get burns that meant I wouldn't be able to drive normally. It's just six weeks but I just got horrendous burns." [P13]

In contrast, those who had experienced very few problems during treatment or adapted to the side-effects reported higher values for more invasive treatment compared to other women.

"Initially, mentally my quality of life was absolutely better because I wasn't stressing. Physically it was tricky, but it just took some time to do the things I could do before again." [P21]

Social meanings

Utility was also heavily influenced by a loss in sexuality and social identity as a woman. Many women expressed concern about DCIS requiring a mastectomy. Body image and maintaining a sense of 'normality' was important to younger women and the implications of surgery were perceived as having both physical and psychological effects on relationships and quality of life.

"I think anybody would think a mastectomy was a really significant decision, you know. Certainly, my breasts – women's breasts – are not only a part of your identity, they're part of your sexuality. They're really important. To kind of go, there's a 10% chance of something happening, let's lop off your breasts. No!" [P10]

"I think age has got a heck of a lot to do with it. You know? But yeah for me, all I want to be is as close to normal as I can. In some ways it sounds stupid…but it's like you've got a disability. If you lose your breast, you've lost part of your body." [P8]

Despite this stigma surrounding mastectomy, women with a family history of breast cancer were happy to trade-off the potential morbidity and side-effects of intensive treatment to

offset concerns about recurrence. This sub-group of women were responsible for much of the heterogeneity in utilities in the empirical study where mastectomy was rated higher than breast conserving surgery or active monitoring.

"Definitely family history with me. I wanted [reconstruction] to feel normal because that whole process of all those tests and that sort of stuff…you feel far from normal. If I was going to make a tough decision like this, I wanted to feel like me at the end of it, as much as I possibly could. I was stressed all the time. It probably just gets back to that whole peace of mind thing for me." [P21]

"I would expect that people who've got a family history of breast cancer, who've perhaps seen mums or aunties die early of whatever, would probably be one to take much more drastic action than people who don't have that family history as well." [P5]

Practical considerations

Women were worried not just about the implications of the treatments on themselves, but the negative impact it might have on their family, friends and colleagues.

"It's a long time in which their everyday [life] is altered for them, and not only for them but their immediate family and their workplace as well." [P1]

Those with supportive partners or family nearby were more likely to give higher values to scenarios with radiotherapy or monitoring, where transport or financial constraints might impact upon decisions.

"I also think people will take into consideration their family setup because some treatment can be inconvenient. Like radiotherapy was very inconvenient to me because I had just gone back to work." [P2]

"I was really fortunate that my daughter-in-law was working in the city at the time and they lived within easy commutable distance of the city. So, for me the transport and the accommodation weren't an issue. And I wasn't working so I wasn't having to fit trips in around work." [P5]

Some women from rural areas had experience of staying away from home to facilitate the demands of radiotherapy regimes, but they reported worse quality of life associated with the period of social isolation away from their loved ones.

"People live 800km from the nearest centre. I had to go down for six weeks. They put you up on a one-bedroom apartment and I had ten minutes radiation [every day]. So, when it comes to DCIS particularly, I would think finances and other obligations could make a big difference." [P14]

Fitting in with role and responsibilities

Personal factors were deemed equally as important to some women as the clinical and economic evidence provided in the vignettes. Although some women would have preferred a more conservative approach based on perceived quality of life, ultimately their decision was made to fit in around their role and responsibilities as a mother or carer. For example, several women living in rural communities opted for a mastectomy to accommodate their duties to the working farm and family practice.

"They can't afford to leave the farm…let's just have a mastectomy based entirely on the fact that they did not want to leave the farm for six weeks, which they would have to do if they were going to have radiotherapy for [breast] conserving surgery." [P14]

"I think a lot of people from the country go for a mastectomy to avoid radiotherapy. I think that's massive overtreatment." [P15]

Similarly, younger women reported lower quality of life for breast conserving surgery with radiotherapy because it was likely to have a bigger impact on their work-life commitments or anticipated role as a mother breastfeeding or caring for the children.

"There is a difference for younger women who are thinking about families or breastfeeding, or they want to have kids. Then there's the next group. They've had their kids, they're working, they're at the busiest time of their lives. They don't have the time [for radiotherapy]." [P7]

9.6 Discussion

9.6.1 Principal findings

The aim of this chapter was to explore factors influencing patient preferences for managing low-risk DCIS. Patients identified several important treatment attributes not identified in the clinical and economic research, namely the potential regret from unnecessary intervention and logical concerns around responsibilities and access to treatment. In spite of the relatively favourable prognosis, many patients were unable to differentiate the risk of recurrence or death to that of invasive breast cancer, consistent with previous findings in DCIS (190, 422). Consequently, women expressed high levels of confusion when asked about the different

treatment choices for low-risk DCIS relative to overdiagnosis, and were likely to value invasive treatment more favourably if their physician had recommended it or used terminology suggestive of invasive breast cancer (447). However, most women were generally interested in active monitoring, with de-escalation strategies having a greater impact upon quality of life if the risk of invasive progression was ≤20% in ten years.

What mattered most to patients when considering low-risk DCIS treatment varied but risk played a major role in women's valuation. The benefits and harms of treating low-risk DCIS were perceived differently depending on risk personality, understanding of the condition and roles and responsibilities. This phenomenon was more pronounced among women who had undergone a mastectomy, whereby quality of life was higher than expected from adaptation after having had the abnormality removed (and risk of recurrence reduced). In contrast, attitudes toward overdiagnosis were highly variable and reflected the heterogeneity in patient utilities from the previous chapters. Women did not think that overdiagnosis would impact utility per se as they preferred to know that they had the condition, but overtreatment *was* perceived to have a likely adverse impact on quality of life, particularly if this risk had not been made explicit at the point of making the decision. This is consistent with qualitative research in other low-risk cancers in which participants raised concerns about the need to mitigate overtreatment (424).

Understanding women's preferences for managing DCIS is far from straightforward. Valuing treatment is challenging because of the fear and uncertainties surrounding the risk of disease progression and overdiagnosis. Whilst patients with low-risk DCIS have an excellent prognosis (25), its management remains controversial because patients continue to perceive it as a life-threatening condition. Women are paradoxically reassured that they do not have

cancer yet are recommended extensive treatment which they may subsequently perceive as having saved their life (190). Thus, patients who felt "lucky" to have been diagnosed and treated tended to self-validate their own treatment, even if there was a chance that it might have been unnecessary. Yet other women expressed regret at the idea of having undergone unnecessary intervention. Clarifying the true risks in clinical trials may improve the acceptability of strategies aiming to safely de-escalate treatment and provide reassurances reflected in terms of quality of life (validating the utilities from Chapter 6).

It is unclear where the true benefits and harms lie in managing DCIS. The risks associated with conventional breast cancer treatment were generally well described, with most interviewees able to list potential side-effects from surgery and radiotherapy versus an anticipated gain in recurrence free survival. Overdiagnosis has societal implications on costs and resource use (23) but some individual women attach greater anxiety and disutility to the unknown fear of progression versus the regret of unnecessary treatment. Treatment preferences expressed by these women depended heavily on the magnitude of overdiagnosis and risk of progression presented, with the highest estimate of overdiagnosis having a greater impact on intended preferences. Other women generally felt that they didn't have enough information to make an informed decision about overdiagnosis but suggested that if it could be shown that their treatment had been unnecessary then it might impact upon utility.

Whilst previous assumptions about the impact of overdiagnosis on quality of life appear to hold true (31, 236) (i.e. the disutility is captured from the diagnosis of the disease), unnecessary treatment does seem to be more meaningful than previously assumed and valued in economic evaluations. Whilst some of these harms are likely captured by the morbidity of

treatment, it appears that the added harm or regret of unnecessary treatment is more meaningful to some women and should be valued conceptually.

This poses an interesting question about how quality of life is conceptualised in health states with an unknown outcome. Economic measures of utility assume rationality in decision making (16), but many women reported that their preferences were influenced by an emotional response. Whilst measures such as the standard gamble facilitate the trade-off in benefits and risks, this assumption may be violated if there are other individual factors (e.g. family circumstances, fears of "cancer" etc.) driving the final decision on utility (140, 150). Gaining an understanding of these factors through qualitative research is crucial as it highlights areas of importance for consideration in the planning of services and valuation of health. This emotional response was not only observed in scenarios of active monitoring but in all hypothetical treatment scenarios describing DCIS.

9.6.2 Strengths and limitations of the study

The screening and management of low-risk ductal carcinoma in situ is under review (3). By examining patients' perspectives of the benefits and harms of treatment, this study makes an important and timely contribution to the debate on the management of low-risk DCIS identified at screening. Furthermore, it explores whether active monitoring would be an acceptable alternative to conventional management in the context of overdiagnosis. Our results are consistent with other qualitative research (275), which suggest up to two thirds of women would consider active monitoring to manage in situ disease despite it not being standard practice.

Another strength of this study is the novel insight it offers into women's preferences for treating low-risk DCIS. The use of qualitative methods to explore how patients perceive the risk of overdiagnosis offers valuable explanation into the variation in treatment choices observed in the empirical research. The in-depth analysis of women with experience in the trade-off of risk adds to the literature by highlighting the areas that patients deem most important in policy decisions involving DCIS.

The findings were limited by the recruitment of women with some knowledge about the risk of overdiagnosis from the empirical stage of the research. As participants had chosen to complete both the questionnaire and interview, there may be motivation toward invasive management among those interested in breast cancer research. Similarly, interviewing women who had already undergone surgical management of breast cancer may be potentially biased by their experience of treatment or communication with doctors (448). Indeed, several women with high-risk disease made statements that validated their own treatment. These women were in the minority and it is plausible that their views may not reflect those for low-risk disease in the wider population or in other screening countries such as the UK.

The sample size is in keeping with those observed in published qualitative research (299, 425), and the depth of information obtained from the 26 women had reached theoretical saturation, so that further interviewing would have been unlikely to yield further novel insight. It could be argued that thematic analysis may be prone to the partiality of the interviewer (398), but this approach offered the advantage of identifying relevant issues that may not have been apparent before commencing the interviews and the rigour of the analysis was increased through the independent review of transcripts by two non-supervisory researchers.

9.6.3 Strengths and limitations in the context of existing literature

This study provides important insight into patient understanding of the benefits and risks associated with treating DCIS. The thematic analysis was able to explain, in part, the heterogeneity in patient utilities for health states describing low-risk DCIS observed in the empirical study. As per the literature (274, 298), the results highlighted that many patients do not understand the uncertainty in the risks and benefits of treatment. There was strong desire for improved patient and public education around DCIS to complement decision making. It is likely that some women may value the de-escalation of treatment, a view shared in previous qualitative research (299), but further research is needed to validate this.

There is very little work exploring the management of DCIS in the context of active monitoring. Nickel et al. (299) conducted semi-structured interviews with a community sample of 26 Australian women aged 25-80 years to explore how treatment preferences for DCIS changed with the terminology used. Interestingly, the rates of women hypothetically interested in managing DCIS via active monitoring were comparable to those reported in our study using patients (64% vs 65%). Other qualitative studies (449) have reported similar findings relating to the variation in treatment practices and paradoxical validation of potential overtreatment, but have not evaluated how this would impact screening and treatment decisions for low-risk DCIS.

Issues around overdiagnosis are more widely researched in other analogous cancers identified by screening (278, 290). Active surveillance has become standard practice for managing prostate cancer or cervical intraepithelial neoplasia, although trials are currently underway to establish the safety of active monitoring in DCIS. Understanding of the condition remains a major issue in the valuation of quality of life; some patients felt they were not adequately

informed about low-risk DCIS to value the benefits and harms. Similarly, the discrepancy between the medical language and advice provided to patients by their doctors was highly variable. Given doctors' recommendations are likely to influence patient response (32), qualitative examination from the perspective of healthcare providers might address the heterogeneity in individual treatment observed between women.

9.6.4 Implications for health policy

Understanding patient preferences may have profound impact for both breast cancer screening and treatment policy. Clinicians and policy makers must recognise that patients have different personal needs and thresholds for risk. This research could be extended to determine whether patient reported outcomes in clinical or economic evaluation capture the treatment attributes that are of highest priority to patients and clinicians when making healthcare decisions. This includes deciding which preferences to apply in economic modelling. Women expressed a high level of interest in active monitoring and reducing the harm of overtreatment. Differentiation of DCIS treatment by risk and exploration of the impact upon screening policy may change the decision on how breast cancer management is personalised to those with low-risk disease.

9.6.5 Further research recommendations

Breast cancer treatment is moving toward a personalised, risk-based approach (26). There is undoubtedly a need to screen and treat women with high-risk of invasive disease more aggressively (25). However, women at the lower end of the spectrum should also be managed accordingly and treatment should be de-escalated where safe to do so to prevent the unnecessary over-servicing of women and the associated utilisation of limited healthcare resources. Further exploration of how women would want to be treated based on their own

risk-profile and threshold for risk may help to identify a group of patients where treatment may be safely de-escalated. The views and opinions held in this study were limited to those with previous experience of DCIS or early invasive breast cancer treatment. Interviewing a wider range of women eligible for breast cancer screening (potentially subject to overdiagnosis in the future) could strengthen the findings of this research by exploring the preferences of women prior to the experience of treatment.

9.7 Reflection

As with any qualitative research, the epistemological view of the researcher conducting the interviews may have affected the responses from participants. This is important to acknowledge given that the researcher had clinical experience of working in breast cancer surgery and had already established a relationship with the participants during the empirical phase of the research prior to the qualitative interviews. For example, women may have been more aware about the topic of overdiagnosis having examined the hypothetical descriptions of DCIS treatment when valuing utility. Similarly, it was difficult for the interviewer not to express empathy or surprise when women disclosed some of the difficulties that they had encountered during their own experiences of treatment. This empathetic body language may have changed the narrative of the answers or direction of the interview but is likely to have improved rapport and facilitated the flow of the interview.

Whilst efforts were made to minimise any leading questions during data collection and analysis, through the application of topic guides and a rigid framework index, it is important to acknowledge the potential influencers so that the results may be interpreted with these factors in mind. The clinical background of the interviewer is an important consideration as some women may have felt less comfortable talking about their own experiences or

interactions with medical professionals if they had been negative. However, a sub-sample of transcripts was checked by two independent researchers to ensure the quotes presented in the analysis gave a true reflection of the participant's meaning and intent.

9.8 Conclusion

Understanding of DCIS, communication about the benefits and risks of treatment, uncertainties around overdiagnosis and the psychosocial impact of treatment were identified as important factors in how women value the management of DCIS. This chapter highlights key areas for future research necessary to address the harms perceived by women and where overdiagnosis is likely to impact most significantly in terms of quality of life. Gaining an understanding of why women's views of the benefits and harms vary, alongside the statistical analysis of utilities, offers a rich understanding of how women ultimately value healthcare and the sequelae of screening for low-risk DCIS.

Economic measures assume that patients make rational decisions about treatment, but preferences toward treatment and utility were heavily influenced by emotional response. This will be useful to researchers, clinical advisers and policy makers alike in finding a solution to the problems surrounding overtreatment. Risk perception, accessing treatment, prior knowledge of DCIS and fear of recurrence were reported as key drivers in women's valuation of the relevant health states. The results pose an interesting question as to which utilities should be applied in economic modelling, given the variation in utility related to the psychosocial impact of treatment and self-validation of treatment observed in some patients. Chapter 10 considers the implications of the qualitative findings alongside the other research outputs of the thesis.

CHAPTER SUMMARY

What is known?

- Women with screen detected DCIS may be subject to overdiagnosis and potentially overly intensive or unnecessary treatment.
- De-escalation may reduce the morbidity of overdiagnosis from breast screening programmes, yet the management of low-grade DCIS is highly variable.
- Little is known about the factors driving individual preferences toward DCIS treatment.

What this chapter adds

- Understanding of DCIS, convenience of treatment, communication and attitudes toward risk are important variables in patient treatment decisions.
- Women's responses to overdiagnosis and active monitoring are diverse. Some women value active monitoring as an acceptable alternative to surgery and radiotherapy but require more information and clinical evidence from clinical trials.
- Whilst active monitoring should be discussed as an option in appropriate situations, there continues to be heterogeneity of preferences requiring the need for objective patient choice and discussion.
- Patients may not always make rational decisions in the trade-off in benefits and harms, requiring careful consideration of the utilities applied in modelling.

CHAPTER 10

Discussion and conclusion

10.1 Introduction

The primary aim of this thesis was to quantify the benefits and harms associated with the treatment of ductal carcinoma in situ. The research was motivated in light of rising concerns about whether the benefits and harms were adequately captured in the economic evaluation of breast cancer screening programmes. Ductal carcinoma in situ was used as a vehicle to explore the trade-off between immediate treatment and overdiagnosis in terms of utility.

The empirical evidence in this thesis showed that the harm of overdiagnosis is likely to be important in the economic evaluation of the sequelae of breast cancer screening programmes. Utilities derived for health states explicitly describing the benefits and harms of treatment for low-risk DCIS were lower than those from the literature, where it is unlikely that individuals were informed that their treatment may have been unnecessary. An economic model of the LORIS trial demonstrated that the explicit inclusion of both benefits and harms in QALYs had the potential to change the decision on the likely cost-effectiveness of treating low-risk, screen detected disease. Finally, qualitative interviews with patients further validated that the inclusion of the effects of harms was important in the valuation of quality of life.

This chapter revisits the original aims of the thesis and demonstrates how each chapter has contributed to answering each objective in part. The main findings are then consolidated and discussed within the context of the wider literature, highlighting the strengths and limitations of the approaches used and potential implications for breast cancer treatment and screening policies. The final section outlines recommendations for future research to address the

challenges associated with valuing overdiagnosis and its sequelae, as well as a reflection on the methods and overall conclusions of the thesis.

10.2 Overview and objectives of the thesis

Breast cancer screening programmes have been the focus of significant debate following the concerns raised by the UK Independent Panel about the potential harms of overdiagnosis (3). Briefly, whilst mammography has undoubtedly contributed to a reduction in breast cancer morbidity and mortality in recent years, there is growing evidence to suggest that it also identifies low-risk disease that may never cause symptomatic harm (450). Consequently, some women may be screened and potentially treated (for DCIS) without any benefit in breast cancer survival.

Most of the research to date has focused on the clinical aspects of the screening strategy, with less attention afforded to the economic evidence informing the debate (22). Economic evaluations conducted to appraise breast screening programmes conventionally adopt the 'cost per QALY' as the outcome measurement. Yet there is evidence to suggest that the utilities informing QALYs may not consider all relevant benefits and harms, particularly regarding the sequelae of overdiagnosis or overtreatment (284).

The principal purpose of the thesis was to identify and quantify the benefits and harms associated with the treatment of DCIS for use in the economic evaluation of breast cancer screening programmes. Data were generated from a combination of systematic literature reviews, empirical utility valuation, the development of an economic model and qualitative interviews in order to address this question.

Four study objectives were outlined at the beginning of the thesis:

1. To ascertain how the benefits and harms associated with breast cancer screening programmes are currently conceptualised and valued.
2. To value the utility associated with different treatments for low-risk DCIS, explicitly including the risk of overdiagnosis.
3. To understand the factors which influence women's preferences in the trade-off between the treatment of DCIS and the risk of overtreatment.
4. To develop an economic model of low-risk DCIS, predicting costs and outcomes for a hypothetical cohort of patients using the derived utilities, and to consider the potential impact of the findings upon the future evaluation of breast screening programmes.

The next section summarises the principal findings from the thesis in meeting each of these objectives.

10.3 Summary of principal findings

10.3.1 Findings from the systematic reviews

A necessary objective of the thesis was to first systematically review the economic evidence for breast cancer screening programmes in order to ascertain how the benefits and harms of screening were captured in the literature. Specifically, the review sought to determine what health states had been valued, how they had been quantified and whether the utilities and disutilities had been valued appropriately in terms of cost per QALYs.

Most economic evaluations reviewed and presented in Chapter 4 concluded that breast cancer screening was cost-effective using a cost per QALY approach. Quality of life was found to be

a key driver of cost-effectiveness in sensitivity analyses, yet the evidence informing QALYs did not necessarily capture all benefits and harms associated with the sequelae of mammography. The systematic review of economic evaluations and primary studies suggested that the utilities derived for or applied to breast screening health states were unlikely to fully account for screening outcomes in terms of utility, particularly with regard to the potential risk of overdiagnosis, and were consistent with the concerns raised by the UK Independent Panel on Breast Cancer Screening (3). None of the primary studies had explicitly included overdiagnosis in the valuation of utility and the same values were used for in situ and invasive breast cancer, despite potential differences in prognosis and management.

The review highlighted a gap in the literature related to the limitations of the utilities being applied in economic evaluations of breast screening programmes. It concluded that better empirical data were needed to inform decisions using QALYs, particularly for low-risk disease most likely to be subject to overdiagnosis. Importantly, the results illustrated the methodological challenges associated with valuing screening related health states and DCIS; namely the uncertainties associated with the natural history of the disease and risk of overdiagnosis, and inconsistency in the economic methods, durations and populations applied to measure utility. Consequently, recommendations were made on how best to proceed in potentially capturing these harms in future economic evaluations.

A second targeted systematic review in Chapter 7 was subsequently undertaken of economic evaluations for treatments specific to DCIS. The aim was to gather information on the design of the models, relevant parameters and necessary data for an economic model representing the LORIS trial in which to apply the utilities derived in the empirical study. Only three studies describing an economic evaluation of DCIS treatment were identified beyond the

scope of screening. The review suggested that utilities for DCIS could be applied in an individual sampling or Markov model to compare the benefits and harms associated with different interventions, but that better utility data were required to populate and assess the cost-effectiveness of treatments for low-risk disease.

The review of published breast cancer screening studies presented in Chapter 4, and of economic evaluations for DCIS treatment in Chapter 7, thus provided support for the empirical valuation and application of utilities associated with the potential overdiagnosis and treatment of ductal carcinoma in situ. An empirical study valuing health states describing treatments for low-risk DCIS allowed the exploration of economic methods aiming to capture the utilities and disutilities in an economic model of the LORIS trial.

In summary, despite evidence from published studies that breast cancer screening is cost-effective using a cost per QALY approach, the utilities informing the decision are limited in their ability to appropriately account for all relevant benefits and risks, specifically related to the potential overdiagnosis and treatment of ductal carcinoma in situ. The systematic reviews highlighted an important gap related to the lack of robust utility data informing the decision on population breast cancer screening and treatment programmes and supported the primary data collection of utilities which aimed to capture the potential risk of unnecessary treatment.

10.3.2 Findings from the empirical study

The objective of the empirical study was to quantify women's preferences for managing low-risk DCIS in order to address the paucity in utilities identified by the review. DCIS is a non-obligate precursor to invasive breast cancer (71) and therefore provided a suitable proxy in

which to evaluate the potential disutility from the risk of overtreatment from overdiagnosis following mammography.

Women were recruited via two breast cancer agencies over a six-month period, and both patient and non-patients were asked to value seven hypothetical health states describing treatments for low-risk DCIS, where the benefits and risks of treatment were made explicit in the vignettes. Three different economic measures (VAS, SG, EQ-5D-5L) were applied using the methods identified in the systematic review.

A total of 172 women (94 patients, 78 non-patients) provided utilities for the seven health states. Utilities were greatest for health states minimising the risk of potential overtreatment (active monitoring) and were lower than those previously elicited for surgery and radiotherapy in the literature, whereby it is unlikely that women were made aware that their treatment may have been unnecessary. The exception to this was for breast conserving surgery, which was rated more favourably by patients using the VAS. The disutility associated with unnecessary treatment was quantified (as the difference in utilities for monitoring and potentially unnecessary intervention) and varied from -0.123 to 0.203 depending on the methods and perspective applied.

The utilities elicited in the thesis were generally lower than those from the literature for health states using comparable populations and methods (135). Whilst this decrement in utility could be related to other methodological factors (e.g. sampling, anchoring, demographics etc.), the differences may also be correlated in part with the risk of unnecessary treatment described, particularly given the convergence in the utilities derived in Chapters 5 and 6 across three economic methods and two populations.

10.3.3 Findings from the economic model

The next step following the collection of utilities was to determine whether capturing the benefits and harms in QALYs was likely to have a meaningful impact on future breast cancer policy decisions. It was beyond the scope of this thesis to model the entire breast cancer screening pathway. Consequently, an economic model of the LORIS trial was designed to compare the cost-effectiveness of treatments using the utilities derived in the empirical study. An individual state-transition model was developed to account for the recurrent health states associated with the treatment of DCIS. The model simulated a cohort of 50 year-old women with low-risk DCIS undergoing active monitoring or standard treatment over a lifelong time horizon. The results using the utilities from the empirical study (where the disutility of unnecessary treatment was valued) were compared with those from the literature (where the risk is unlikely to have been adequately captured).

Keeping all other model parameters equal, the cost per QALYs using the empirical utilities from the thesis suggested active monitoring was cost-effective when compared with standard surgical management, but when the utilities from the literature were applied the magnitude and direction of cost-effectiveness was reversed (AM was dominated). This finding is critical to the main objectives of the thesis; not only was the disutility of overtreatment important in economic evaluation, but it also had the potential to change the relative cost-effectiveness of the current strategy and decision on how low-risk DCIS is managed.

The impact of the disutility of overtreatment from overdiagnosis had a more profound effect on older women (>60 years), or where model inputs were set to a lower risk of invasive progression or shorter time horizon, consistent with analyses of active monitoring strategies in other screen detected cancers (333, 451), where patients are more likely to die from

competing causes of mortality. There was however significant uncertainty associated with the model related to the limitations of the data for active monitoring (PSA demonstrated mixed results). Nonetheless, given the illustrative nature of the model the results suggest that capturing the harms of the sequelae of screening in economic evaluations does matter, as per Raftery and Chorozoglou's (31) initial hypothesis about the possible net harms of breast cancer screening.

10.3.4 Findings from the qualitative analysis

The empirical data and economic model thus far indicated that the sequelae of overdiagnosis were important in the decision on how to manage low-risk DCIS. But there was also variability in the individual utilities reported, implying that how women conceptualised the benefits and harms was also likely to be crucial to understanding the impact that this could have on practice and policy. A qualitative analysis was therefore undertaken to explore the reasoning behind the heterogeneity in utility and to gain an understanding of the factors which influenced women's preferences in the trade-off between treatments for DCIS.

Twenty-six in-depth interviews were conducted with breast cancer patients in Melbourne, Australia, who had experience of the different treatments for early stage disease. Interviews were semi-structured and inductive in that women were encouraged to talk at length about their experiences and interpretation of the benefits and harms, but they could also raise issues that they felt were important in the valuation of treatment not suggested in the topic guide. A framework analysis was adopted to allow the preferences of women to be compared using a systematic approach and supplemented the empirical data by exploring the factors influencing utility.

Four common themes were identified from the qualitative interviews relating to how women valued the utilities associated with DCIS; understanding of the condition, information and communication about the benefits and risks of treatment, uncertainty around the outcomes and the psychosocial impact of treatment. Evidence from the interviews, and in line with previous qualitative research, suggested that many patients continued to perceive DCIS as a life-threatening condition despite its favourable prognosis (190, 452). This phenomenon was more pronounced among women who had undergone a mastectomy, where the reported utilities for invasive management were higher than for BCS or active monitoring. Similarly, most women were unaware about the risk of unnecessary treatment but indicated that this would likely have been important in their decision toward less invasive treatment, had they been made aware of it by their clinician.

Women found valuing utilities for DCIS to be particularly challenging because of the uncertain pathway of invasive disease progression. Patients expressed conflicting opinions about the benefits and harms of treatment according to risk threshold (fear of recurrence), family commitments and responsibilities (carer spill-over), access to treatment (rural-urban divide) and perceived tolerance for the side-effects. Nonetheless, patient involvement suggested that the harms were considered just as important as the benefits in the valuation of breast cancer screening and treatment programmes.

10.4 Synthesis and reflection on the main findings

A number of similar overarching themes were raised across the different phases of the research. The findings collectively demonstrated that:

1. Economic evaluations of breast cancer screening programmes are unlikely to adequately capture the benefits and risks associated with its sequelae in QALYs.
2. Quantifying the disutility of overdiagnosis in breast cancer is challenging because of the unknown natural history of the condition and risk of invasive progression.
3. Current utilities are limited by the *paradox of overdiagnosis,* i.e. the more people are (over) treated, the more people think screening saved their lives, especially if each individual justifies their treatment by believing that they have had a dramatic benefit.
4. There are a limited data on the cost-effectiveness of interventions for low-risk disease (DCIS) identified through mammography screening.
5. The utilities associated with the treatment (and risk of overtreatment) of DCIS may feasibly be quantified using standard gambles or EQ-5D and vignettes which explicitly detail the benefits and risks in the health states described.
6. The harm of overdiagnosis may subsequently be captured indirectly by calculating the disutility between treatments with and without the risk of overtreatment.
7. Under a wide range of assumptions, for a 50 year-old woman, active monitoring is a reasonable approach to treating low-risk DCIS based on QALYs compared with standard treatment. However, individual preferences play a key role in the decision.
8. The valuation of the benefits and harms associated with DCIS is important and may change the relative cost-effectiveness of the current treatment (and screening) policy.

In terms of the overall research, the combination of quantitative and the qualitative patient interviews has enhanced the strength and validity of the findings. The utilities provided empirical data that could be applied in the economic evaluation of DCIS interventions, and the interviews were able to capture data on the issues that women felt were important in the valuation of health states. Furthermore, the main themes and issues that emerged from

quantifying the benefits and harms associated with the sequelae of DCIS were corroborated in both phases of the research.

The next section discusses the main strengths and limitations of the overall findings and work undertaken in this thesis.

10.5 Strengths and limitations of the thesis

10.5.1 Strengths

The main strength of this thesis is that it has presented both empirical and qualitative evidence to demonstrate that the explicit inclusion of the benefits and harms associated with the treatment of DCIS does matter in the decision following breast cancer screening.

First, a rigorous and comprehensive review of the literature was undertaken to critique the current evidence informing the breast cancer screening debate. The systematic review illustrated that overdiagnosis had not been adequately captured in the economic evidence informing breast cancer screening policy and suggested potential methods for exploring how this harm might be captured in QALYs. An exploratory study was subsequently undertaken to provide novel utility data in which the risk of overtreatment was quantified, using DCIS as a case study in which to address the limitations identified.

Second, the application of the derived utilities in an economic model of the LORIS trial demonstrated that the explicit inclusion of the effects of this disutility would likely provide results different from those produced using the utilities from the literature. This finding is of vital significance to future economic evaluations in breast cancer because it supports concerns that the net harms may change the decision on how women are treated or screened

(31). Patients also voiced issues related to how the benefits and risks associated with DCIS were communicated in the decision to treat screen-detected disease. This thesis was therefore unique in that it not only explored economic methods to try and quantify the utilities and disutilities associated with DCIS, but it also used in-depth interviews to explore the factors driving the differences in the values themselves.

Thirdly, this is the first study to value the disutility associated with the potential harm of unnecessary treatment in terms of quality of life. It builds on the hypothesis generated by Raftery and Chorozoglou (31) to demonstrate (through the application of the utilities in an economic model) that the disutility associated with the risk of unnecessary treatment may change the relative cost-effectiveness of the current treatment strategy for low-risk DCIS. Further, the utilities derived in this thesis may be applied in future economic evaluations of breast cancer screening programmes to capture the impact of treating (and potentially overtreating) in situ disease following mammography (19).

The results provide key evidence to inform the ongoing debate about breast cancer screening programmes. Breast cancer screening was declared a priority policy area by the UK Department of Health (453), and yet the economic evidence on the harms of overdiagnosis are insufficient to advise future policies on risk-stratified mammography screening. The utilities reported in the literature are unlikely to capture the additional impact upon quality of life from having an intervention which may not be necessary, as shown in Chapter 6. Although the findings from this thesis require further validation in the context of the wider breast screening pathway, the results do demonstrate that vignettes may provide a suitable vehicle in which to quantify this hypothetical disutility as a first step toward capturing the harms of overtreatment in economic evaluation.

10.5.2 Limitations

It is important to acknowledge that the methods applied in this thesis were exploratory in nature. Although the methods and vignettes were informed by clinical experts and an extensive review of the literature (284), it is unlikely that the disutility of overdiagnosis was captured in its entirety. Theoretically, women can never truly know whether they are the subject of overdiagnosis. Therefore, an assumption was made that this harm was best captured indirectly by valuing the impact of the risk of unnecessary treatment among those with low-risk disease. Whilst this risk was important in the valuation of utility, further validation is required by comparing vignettes with and without the risk of overtreatment and confirmation of the likely probability when LORIS is published.

A second limitation of the thesis was that the vignettes were framed to avoid using cancer terminology to describe in situ disease, as recommended in other studies valuing health states for DCIS (275, 299), but this may conflict with the goal of overdiagnosis valuation. If women are willing to tolerate more risk for the sake of avoiding cancer, one could argue that this should be considered in the description of DCIS, even if it may bias the results toward invasive treatment. However, it is unlikely that this would have had a significant impact on the utilities elicited in this thesis given the recruitment of women from breast cancer research registries who may have been more knowledgeable than the lay individual. Ascertaining the values of non-screeners, younger women and men is thus also necessary to ensure the findings are representative of the views of the general population who may be less well informed about DCIS and the concept of overdiagnosis.

Third, the calculation of the EQ-5D-5L utilities was reliant on the generalisability of UK to Australian tariffs (18). Although previous studies have demonstrated reasonable convergence

between the two population demographics and baseline EQ-5D-3L utility data (302, 303), less is known about the applicability of EQ-5D-5L tariffs across the two breast cancer populations.

Finally, the application of utilities was reliant on an economic model with considerable uncertainty. The model parameters were limited by the lack of published data on the probability of in situ disease progression and resource use for active monitoring, as confirmed by the mixed results in the probabilistic sensitivity analysis. However, the purpose of the model was to illustrate the potential impact of the inclusion of the net harms and not to generate the actual cost-effectiveness of active monitoring versus standard treatment. Therefore, by keeping all other variables constant the model was able to demonstrate the potential impact associated with the management and potential overtreatment of low-risk DCIS identified by mammography.

10.6 Comparison with other studies

10.6.1 Challenges associated with overdiagnosis

This study has provided empirical evidence to demonstrate that the benefits and harms captured in the economic evaluation of DCIS treatment matters. The findings build on the foundations set out by Raftery and Chorozoglou (31) who found that the inclusion of the effects of the harms from false positives and treatment changed the initial cost-effectiveness of the current breast cancer screening strategy. Although the economic model in this thesis focused on part of the screening pathway, specifically the management of DCIS based on LORIS trial, the explicit inclusion of utilities capturing the risk of unnecessary treatment did affect the direction and magnitude of the policy recommended for low-risk disease. The

utilities derived may be used to further update the Forrest model (19) to ascertain the potential impact of overtreatment on breast cancer screening programmes.

This is the first study to elicit utilities for active monitoring and the risk of unnecessary invasive treatment in DCIS, but prior studies have attempted to capture this harm in the wider oncological setting (289, 332, 454). Several studies (289, 455, 456) have valued health states for active surveillance in prostate cancer describing the potential for invasive progression, and reported higher utilities compared with standard surgical treatment. Utilities were subsequently applied in economic models of prostate cancer screening but demonstrated mixed cost-effectiveness results on whether men should be routinely screened (333, 364). Similarly, Venkatesh et al. (457) estimated the harm of unnecessary treatment for low-risk thyroid cancer using the differences in disutility (1-utility) between active monitoring of stable disease and disease-free monitoring after surgery. Cost-effectiveness was highly dependent on patient disutility associated with active monitoring, but individuals did not have to make decisions about screening and treatment based on incomplete or missing information.

DCIS on the other hand presents a challenging health state to value and conceptualise because of the unknown risk of invasive progression. A full understanding of the natural history of the disease is required to overcome traditional screening biases and uncertainty. Women are paradoxically reassured that they do not have cancer yet are recommended extensive treatment which they may perceive as having saved their life. Whilst some authors (458, 459) argue that if women don't know there is a disutility then the harm cannot be captured, the results from both the empirical and qualitative interviews presented in this thesis provide firm evidence that knowing the risk of unnecessary treatment does seem to be important and potentially quantifiable in economic evaluation.

Whilst the disutility of overdiagnosis cannot be quantified directly because women cannot ever be sure whether they have or have not been overdiagnosed (31), the disutilities from treatment may be used to capture this harm indirectly by comparing values which include the risk of overtreatment (immediate surgery) with those that do not (active monitoring). Indeed, the qualitative research in Chapter 9 suggested that being informed about both the benefits and risks (including the potential for unnecessary treatment) was important to women when valuing the health states for DCIS. Similar findings about the sequelae of mammography screening have been reported in other qualitative studies exploring the impact of overdiagnosis (298, 299).

10.6.2 Methodological approach

A major contribution of this thesis was the exploratory methods applied to measure the utilities and disutilities associated with DCIS. The standard gamble and EQ-5D approaches adopted were built on the methods applied previously to value health states with uncertainty in prostate cancer (289) and were informed by a comprehensive and systematic review of the literature (284). However, this thesis also explored whether the utilities derived could be applied to capture the disutility associated with overdiagnosis indirectly.

The utilities reported were lower than those previously elicited in the literature for health states describing treatments for DCIS, where the risk of unnecessary treatment was not made explicit in the vignettes (148, 248, 254), but were comparable to those derived for active monitoring states in prostate cancer where the benefits and risks were similarly described (289, 332). Nonetheless, the SG values exhibited a slight ceiling effect between health states (a known complication of the SG) (152), and further validation of the EQ-5D-5L is required by comparing vignettes with and without the risk of overtreatment.

The responses to standard gambles were comparable to those reported by Baker and Robinson (460) where the impact of the choices on others was a natural part of the decision process (e.g. radiotherapy on family life). The standard gamble is often advocated on the grounds that almost all decisions about health are made under conditions of uncertainty (140). The gamble holds if the axioms of EUT are upheld but individuals may not always act rationally in decisions around healthcare (461). Yet preferences are defined over a domain of lotteries, some of which may be constant, whereas others may have multiple outcomes (117). This is reflective of real life where health states and outcomes are never certain. Thus, by including the benefits and risks in the vignettes but keeping the gamble anchors constant (i.e. dead and perfect health), the axioms of EUT were upheld and may be argued as valid.

Llewellyn-Thomas et al. (328) point out that it would be *"naïve to think of any state of health as possessing a single utility or value,"* (1984: p50). Rather, utilities for health states are influenced by the method used to elicit judgements and other circumstances of the rating task. The findings from this thesis raise two different questions; the problem of valuation and of communication. First, the results suggest that there is an intrinsic harm from overdiagnosis related to overtreatment. Women valued surgery lower than active monitoring if it reduced the potential disutility from undergoing unnecessary treatment without a significant trade-off in the risk of invasive progression using all three economic measures. However, the magnitude of this difference was largely dependent on the method and population applied to derive the utility (178). This phenomenon was also observed in another study of breast cancer health states, where patients rated BCS higher than the public, despite the same outcome of recurrence being described in the vignettes (144).

The qualitative study also raised the question of the harm of misinformation and informed choices. Some women were willing to forgo the risk of possible side-effects and overtreatment to reduce the risk of local recurrence, but others indicated that they would have placed a lower value on health states had they known that there was a chance that it might not have been needed. For example, patient utilities were more likely to favour invasive adjuvant therapy or mastectomy if they were told that it had been necessary by their doctors. Issues in the understanding of risk and misinformed decisions about breast cancer screening and treatment practices were previously highlighted as major drivers of responses to treatment among high-risk women in Australia (462).

10.6.3 Economic model

The economic evaluation of active monitoring and standard treatment in Chapter 8 is the first reported for DCIS, but the results were comparable to economic models of prostate cancer surveillance (463). Both models resulted in a higher rate of invasive progression for active monitoring but no difference in disease-specific or all-cause mortality (333). Active monitoring was favoured over initial treatment for both low-risk breast and prostate disease using cost per QALYs, but the results were highly dependent on the utilities for living under active monitoring compared with having been treated (451). Individual preferences thus appear to play a central role in the decision on whether to treat or to pursue active monitoring in the management of low-risk cancers identified through cancer screening programmes.

10.7 Contribution to the literature

This thesis has demonstrated potential methods toward capturing the harm associated with overdiagnosis in the economic appraisal of breast screening programmes and has therefore made an original and relevant contribution to the literature. However, it would be naïve to

suggest that this thesis has captured the disutility of overdiagnosis in its entirety. Even if the risk of invasive progression is validated in the LORIS trial (30), women will always face a degree of uncertainty in deciding whether to have treatment, as is the case for any intervention. Why then is the harm of overdiagnosis different to the risks observed in other health decisions? Indeed, some individuals may prefer to experience ill health immediately to eliminate anxiety, for example, those with a positive family history of BRCA may opt for prophylactic bilateral mastectomy to reduce the risk of developing breast cancer (144).

Ultimately, women with DCIS are treated aggressively for a condition which may never progress as an indirect consequence of screening. The natural history of the condition following diagnosis is unknown (28), yet the current valuations are based on the premise that surgery is necessary or lifesaving. The approach adopted in economic evaluations to date assume that overdiagnosis is captured through the side-effects and costs of cases treated or treated earlier. However, this argument is flawed if the individuals valuing health states (or their own health in the surgical arm of active monitoring trials) are under the impression that the treatment is necessary because this may inappropriately bias utilities and the decision towards screening and invasive treatment. This thesis has demonstrated that if women are informed that there is a risk of overtreatment then the utilities reported are likely to be lower.

Overdiagnosis has obvious harms to society in terms of inappropriate resource use (and opportunities forgone). However, what this thesis has aimed to ascertain is whether this harm also has a tangible impact upon quality of life (and QALYs) among those informed that there is a risk that their treatment may be unnecessary as an indirect result of overdiagnosis. The results of the empirical study suggest that there is likely to be an additional disutility when overdiagnosis is explicitly captured into the decision-making process. Further, when this

harm was included in an economic evaluation it had the potential to change cost-effectiveness results.

10.8 Implications for DCIS and breast cancer screening

The objective of this thesis was to quantify the benefits and harms associated with the overdiagnosis and treatment of DCIS and to highlight where future research may be required in order to minimise the potential harms from screening. The aim was not to provide an all-encompassing model of the breast screening programme, but rather to consider how such harms may be quantified in future economic evaluations by exploring the potential methods for capturing them in a sample of women likely to be the subject of overtreatment.

Evidence from the interviews and empirical study conducted, and in line with previous studies, suggest that women do value the early diagnosis and management of in situ disease. Whilst overdiagnosis from screening is a risk that many women felt worth accepting, the potential consequences of overtreatment, however, were not. The results showed that capturing the harms from unnecessary treatment were important in the economic evaluation of DCIS management, and the values may be further applied in the economic evaluation of breast cancer screening programmes. Indeed, the utilities are more likely to be representative of women's preferences than those elicited in the LORIS trial as women were informed in the vignettes that there was a risk of that the treatment may have been unnecessary, whereas this is not ethically plausible between the two randomised trial arms in practice.

Critically, the results presented are relevant to the national breast cancer screening programme which is seeking to move toward a more risk-stratified approach (26, 390). The results from this thesis do not suggest by any means that we should not be screening women

(mammography undoubtedly saves lives), but rather that the harms of screening the wider population should be considered in the decision process, so that screening and treatment can be tailored appropriately. With plans to extend routine mammography screening to women aged 47-74 years in the UK (230), it is likely that the incidence of low-risk and in situ breast cancer will further increase and therefore a re-evaluation of the most cost-effective breast screening and treatment strategy is necessary.

10.9 Implications for other health conditions

The inclusion of benefits and harms in QALYs is important in the economic evaluation of healthcare. Whilst the results presented in this thesis may impact upon decisions related to breast cancer treatment (and potentially mammography), there are also potentially important implications for the economic evaluation of screening strategies in the wider health economic literature.

The UK National Screening Committee (UKNSC) is responsible for the evidential review of over ten population screening strategies for those who may be at increased risk of a disease or condition (464). When considering who to screen and which conditions to screen for, the benefits of offering a screening programme are weighed up against the potential harms. The UK NSC reviews its recommendations on screening for different conditions every three years, but it is unclear as to whether the economic evidence informing other screening decisions have appraised the relevant benefits and harms in QALYs or come under the same scrutiny as breast cancer screening programmes in recent years.

The inappropriate valuation of harms, such as overdiagnosis, may have contributed to the potentially misleading results about the cost-effectiveness of other screening initiatives. For

example, one such case is the use of pulse oximetry as an adjunct screening test for congenital heart defects in newborn infants to supplement the current newborn screening programme (465). Pulse oximetry uses infrared light to identify congenital heart problems that might otherwise go undetected. It was deemed a safe, non-invasive and cost-effective test to support the current newborn screening examination in terms of 'cost per additional case of timely diagnosis' due the predicted reduction in morbidity and costs from early treatment (466). Yet the analysis failed to include any harm from false positives and additional diagnostic investigation. If the findings from this thesis are applied to newborn screening, then the explicit inclusion of the potential harms from false positive screens (i.e. extra diagnostic tests and lumbar punctures performed) may render the screening programme no longer cost-effective and recommended for those at lower risk.

Similarly, the conclusions from this thesis may also have implications in other screening programmes where the cost-effectiveness of screening may potentially be reversed by including the disutilities associated with the harms alongside the benefits in QALYs. Examples where the harms are likely to be important in other screening decisions include abdominal aortic aneurysm screening (467), prostate cancer screening (286) and population cancer genetic screening (468). The results do not imply that the general population should or should not be screened, but rather that further research and critical appraisal is necessary to ensure that the sequelae of screening are valued appropriately in economic evaluation.

10.10 Future research recommendations

The research developed within this thesis has made a valuable contribution to the existing body of literature, but also serves to outline recommendations for future research.

Firstly, a comparison of utilities using vignettes for health states with and without the description of overdiagnosis would provide more information on the potential magnitude of this harm. This thesis has demonstrated the methodological feasibility of quantifying the benefits and harms associated with the sequelae of screening in terms of utility, including the risk of unnecessary treatment from overdiagnosis, but it is recognised that further analysis is necessary to validate the findings. The measurement of utility using a sample of the UK general population would also provide utilities which are more generalisable for use in economic evaluation as they may have a lower lay understanding of DCIS and overdiagnosis.

Second, the findings support a re-evaluation of the cost-effectiveness of the NHS breast screening programme using the utilities derived in this thesis to account for the disutility of overdiagnosis. A primary objective of this thesis was to provide data for low-risk breast cancers that could be used to inform cost per QALYs in the breast cancer screening debate.

The optimal management of DCIS is uncertain because there are limited data on the natural history of the disease. As has been highlighted in the literature, future research should focus on clarifying the risk of invasive progression and more reliable data about the long-term sequelae associated with DCIS, to reduce the uncertainty in economic models and health valuation studies. Collaboration of clinical and health economic findings from trials seeking to ascertain the appropriate treatment of screen detected low-risk DCIS (e.g. LORIS) is necessary to reduce adverse outcomes incurred through population screening. It is likely that utilities will also be better informed once the evidence from LORIS is published as only then can women make truly informed decisions about the screening and treatment of low-risk disease.

Finally, the qualitative work presented in this thesis raised important concerns about patient understanding of DCIS and the benefits and harms of treatment. Further engagement with the general public and clinicians is necessary to determine how best to communicate decisions related to the risk of overtreatment, to improve lay understanding of overdiagnosis prior to mammography attendance and to improve the standard of information regarding treatment choices for DCIS provided by health professionals.

10.11 Conclusion

This thesis has provided both quantitative and qualitative evidence to demonstrate that the benefits and harms captured in the economic evaluation of DCIS treatment is likely to matter. The contribution of this research is important in the context of the debate on breast cancer screening programmes. Much of the focus has been on the magnitude of overdiagnosis yet little has been done to address the issues in the economic evidence informing the debate.

The inclusion of the benefits and harms associated with DCIS is likely to be important in the economic evaluation of mammography although further research is required to ascertain the exact magnitude of the associated disutility. Utilities derived in this thesis were lower than those previously elicited in the literature, where the risk of overtreatment had not been made explicit, and had the potential to change the decision on the cost-effectiveness of treatment. Capturing the risks of unnecessary treatment was also deemed important to women in the discussion about the sequelae of screening, particularly for those with DCIS which accounts for a considerable proportion of the breast cancers identified through screening.

This thesis has made several original and important contributions to the health economic and breast cancer literature, both methodological and in terms of recommendations for the future

economic evaluation of breast cancer screening programmes. Methodological approaches are suggested on how to quantify the harm of overtreatment and may have implications for breast cancer screening or other screening initiatives. The involvement of both patient and non-patient perspectives was also useful and provided novel data on how treatment is perceived and valued for low-risk cancers potentially subject to overdiagnosis.

To conclude, this thesis has argued that the harms of screening should be considered in QALYs. The explicit inclusion of the utilities and disutilities associated with DCIS is likely to matter in the decision on how women are screened and treated for breast cancer in the future. Recommendations are made for the economic evaluation of breast screening programmes using the data from this thesis, aimed at enhancing the quality and validity of breast cancer screening and treatment strategies.

APPENDICES

APPENDIX 1: Systematic review publication

Social Science & Medicine 228 (2019) 142–154

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Review article

Valuing the health states associated with breast cancer screening programmes: A systematic review of economic measures

Hannah L. Bromley[a,f], Dennis Petrie[b], G.Bruce Mann[c], Carolyn Nickson[a,d], Daniel Rea[e], Tracy E. Roberts[f,*]

[a] Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia
[b] Centre for Health Economics, Monash University, Melbourne, Victoria, Australia
[c] Department of Surgery, University of Melbourne, Melbourne, Victoria, Australia
[d] Cancer Research Division, Cancer Council NSW, Australia
[e] Cancer Research UK Clinical Trials Unit, University Hospital of Birmingham, Birmingham, West Midlands, UK
[f] Health Economics Unit, University of Birmingham, Birmingham, West Midlands, UK

ARTICLE INFO

Keywords:
Health state utility value
Quality of life
QALY
Breast cancer
Mammography
Screening

ABSTRACT

Policy decisions regarding breast cancer screening and treatment programmes may be misplaced unless the decision process includes the appropriate utilities and disutilities of mammography screening and its sequelae. The objectives of this study were to critically review how economic evaluations have valued the health states associated with breast cancer screening, and appraise the primary evidence informing health state utility values (cardinal measures of quality of life). A systematic review was conducted up to September 2018 of studies that elicited or used utilities relevant to mammography screening. The methods used to elicit utilities and the quality of the reported values were tabulated and analysed narratively.

40 economic evaluations of breast cancer screening programmes and 10 primary studies measuring utilities for health states associated with mammography were reviewed in full. The economic evaluations made different assumptions about the measures used, duration applied and the sequalae included in each health state. 22 evaluations referenced utilities based on assumptions or used measures that were not methodologically appropriate. There was significant heterogeneity in the utilities generated by the 10 primary studies, including the methods and population used to derive them. No study asked women to explicitly consider the risk of overdiagnosis when valuing the health states described.

Utilities informing breast screening policy are restricted in their ability to reflect the full benefits and harms. Evaluating the true cost-effectiveness of breast cancer screening will remain problematic, unless the methodological challenges associated with valuing the disutilities of screening are adequately addressed.

1. Introduction

Evidence regarding the cost-effectiveness of healthcare technology is increasingly required to inform the decision on whether to fund and implement new treatment (Drummond et al., 2015). Many decision-making bodies require interventions to be assessed using cost per quality-adjusted-life-years (QALYs) (NICE, 2013), a single summary measure combining life expectancy with individuals' relative preferences for health states in terms of quality of life (Neumann et al., 2000; Whitehead and Ali, 2010). Health state utility values (HSUVs) are cardinal measures of preference rated on a utility scale anchored from dead (0) to perfect health (Drummond et al., 2015). Utilities can be valued directly or indirectly (Tolley, 2009). Direct methods ask individuals to value hypothetical health states, and preferences are directly measured onto the utility scale using the standard gamble (SG), time-trade off (TTO) or visual analogue scale (VAS) (Torrance, 1986). The TTO and SG elicit individual choices under uncertainty in life expectancy or risk of death and good health (Dolan et al., 1996), whereas the VAS provides an intermediate valuation of health on a graduated rating scale (Bass et al., 1994). Indirect methods use a generic multi-attribute utility instrument (MAUI), such as the EuroQol-5 Dimensions (EQ-5D) (Dolan, 1997). Current or hypothetical health is mapped onto a generic health instrument and indirectly valued using tariffs for the generic health states that have previously been estimated using direct valuation methods from the general population (Kind et al., 1999).

Economic evaluations impact health policy decisions and so the

* Corresponding author. Health Economics Unit, University of Birmingham, Birmingham, West Midlands, B15 2TT, UK.
E-mail address: T.E.ROBERTS@bham.ac.uk (T.E. Roberts).

https://doi.org/10.1016/j.socscimed.2019.03.028
Received 7 November 2018; Received in revised form 21 January 2019; Accepted 15 March 2019
Available online 21 March 2019
0277-9536/

APPENDIX 2: Search strategy for the systematic review of economic measures in breast cancer screening programmes

Ovid MEDLINE(R)

1. Breast Neoplasms/
2. (ductal carcinoma in situ or DCIS).ti,ab. or ductal carcinoma in situ/ or DCIS/
3. (breast cancer or breast carcinoma or breast neoplasm or breast tumo*r or breast adenocarcinoma).ti,ab.
4. (overdiagnosis and breast).ti,ab.
5. Mammography/
6. mammogra*.ti,ab.
7. breast screen*.ti,ab.
8. or/1-7
9. Economics/
10. economic evaluation.ti,ab.
11. "Costs and Cost Analysis"/
12. Cost-Benefit Analysis/
13. cost.ti,ab.
14. exp "Quality of Life"/
15. (quality of life or QoL).ti,ab.
16. Quality-Adjusted Life Years/
17. (quality adjusted life year* or QALY*).ti,ab.
18. (cost benefit analysis or cost utility analysis or cost effectiveness analysis).ti,ab.
19. Models, Economic/
20. economic model.ti,ab.
21. ("time trade off" or tto or "standard gamble" or "visual analogue scale" or (EQ5D or "euroqol 5d")).tw.
22. (preference$ or valuation$).tw.
23. (health utilit* or health disutilit*).ti,ab.
24. health state$.tw.
25. or/9-24
26. Mass Screening/
27. screening.ti,ab.

28. 26 or 27
29. 8 and 25 and 28
30. Limit 29 to English language

EMBASE

1. breast cancer/
2. (ductal carcinoma in situ or DCIS).ti,ab. or ductal carcinoma in situ/ or DCIS/
3. (breast cancer or breast carcinoma or breast neoplasm or breast tumo*r or breast adenocarcinoma).ti,ab.
4. (overdiagnosis and breast).ti,ab.
5. mammography/
6. mammogra*.ti,ab.
7. breast screen*.ti,ab.
8. or/1-7
9. economics/
10. economic evaluation.ti,ab.
11. "cost benefit analysis"/
12. economic evaluation/
13. cost.ti,ab.
14. "quality of life"/
15. (quality of life or QoL).ti,ab.
16. quality adjusted life year/
17. (quality adjusted life year* or QALY*).ti,ab.
18. (cost benefit analysis or cost utility analysis or cost effectiveness analysis).ti,ab.
19. economic model.ti,ab.
20. (health utilit* or health disutilit*).ti,ab.
21. ("time trade off" or tto or "standard gamble" or "visual analogue scale" or (EQ5D or "euroqol 5d")).tw
22. (preference$ or valuation$).tw.
23. health state$.tw.
24. or/9-23
25. mass screening/
26. screening.ti,ab.

27. 25 or 26
28. 8 and 24 and 27
29. Limit 28 to English language

PsycINFO

1. Breast Neoplasms/
2. (ductal carcinoma in situ or DCIS).ti,ab. or ductal carcinoma in situ/ or DCIS/
3. (breast cancer or breast carcinoma or breast neoplasm or breast tumo*r or breast adenocarcinoma).ti,ab.
4. (overdiagnosis and breast).ti,ab.
5. MAMMOGRAPHY/
6. mammogra*.ti,ab.
7. breast screen*.ti,ab.
8. or/1-7
9. ECONOMICS/
10. economic evaluation.ti,ab.
11. "Costs and Cost Analysis"/
12. cost.ti,ab.
13. "Quality of Life"/
14. (quality of life or QoL).ti,ab.
15. (quality adjusted life year* or QALY*).ti,ab.
16. (cost benefit analysis or cost utility analysis or cost effectiveness analysis).ti,ab.
17. economic model.ti,ab.
18. (health utilit* or health disutilit*).ti,ab.
19. ("time trade off" or tto or "standard gamble" or "visual analogue scale" or (EQ5D or "euroqol 5d")).tw.
20. (preference$ or valuation$).tw.
21. health state$.tw.
22. or/9-21
23. Cancer Screening/
24. screening.ti,ab.
25. 23 or 24
26. 8 and 22 and 25

27. Limit 26 to English language

Econlit (EBSCO)

1. Breast cancer or breast neoplasm or breast carcinoma or breast adenocarcinoma or breast tumo*r
2. Ductal carcinoma in situ or DCIS
3. Overdiagnosis
4. Mammogra*
5. Breast screen*
6. Or/1-5
7. Economics
8. Economic evaluation
9. Cost
10. Cost benefit analysis or cost utility analysis or cost effectiveness analysis
11. Economic model
12. health utilit* or health disutilit* or "health state*" or preference* or "time trade off" or "standard gamble" or "visual analogue scale" or "EQ5D" or valuation*
13. Quality of life or QoL
14. Quality adjusted life year* or QALY*
15. Or/7-14
16. Cancer screening
17. Screening
18. 16 or 17
19. 6 or 15 or 18
20. Limit 19 to English language

CINAHL Plus (EBSCO)

1. (MH "Breast Neoplasms")
2. (ti dcis or ab dcis) OR (ti ductal carcinoma in situ or ab ductal carcinoma in situ)
3. (MH “carcinoma, ductal, breast”)
4. (ti overdiagnosis or ab overdiagnosis) AND (ti breast or ab breast)

5. (ti "breast cancer" or ab "breast cancer") OR (ti "breast carcinoma" or ab "breast carcinoma") OR (ti "breast neoplasm" or ab "breast neoplasm") OR (ti “breast tumo*r” or ab “breast tumo*r”) OR (ti "breast adenocarcinoma*" or ab "breast adenocarcinoma*")
6. (MH “Mammography”)
7. Ti mammogra* or ab mammogra*
8. Ti "breast screen*" or ab "breast screen*"
9. Or/1-8
10. (MH "Economics")
11. (MH "Costs and Cost Analysis")
12. (MH “cost benefit analysis”)
13. Ti “economic evaluation” or ab “economic evaluation”
14. ti cost or ab cost
15. (ti "cost benefit analysis" or ab "cost benefit analysis") OR (ti "cost utility analysis" or ab "cost utility analysis") OR (ti "cost effectiveness analysis" or ab "cost effectiveness analysis")
16. (MH "Quality of Life")
17. (ti "quality of life" or ab "quality of life") OR (ti "QOL" or ab "QOL")
18. (MH “Quality adjusted life years”)
19. (ti "quality adjusted life year*" or ab "quality adjusted life year*") OR (ti "QALY*" or ab "QALY*")
20. (ti health utilit* or ab health utilit*) OR (ti health disutilit* or ab health disutilit*) OR (ti preference or ab preference) OR (ti "health state" or ab "health state") OR (ti "time trade off" or ab "time trade off") OR (ti "standard gamble" or ab "standard gamble") OR (ti "visual analogue scale" or ab "visual analogue scale") OR (ti "EQ5D" or ab "EQ5D")
21. ti "economic model" or ab "economic model"
22. or/10-21
23. (MH “cancer screening”)
24. Ti screening or ab screening
25. 23 or 24
26. 9 or 22 or 25
27. Limit 26 to English language, female

Cochrane library, NHS EED, DARE and HTA strategy

#1 MeSH descriptor: [Breast Neoplasms] explode all trees

#2 MeSH descriptor: [Mammography] explode all trees

#3 breast screen*:ti,ab,kw

#4 mammogra*:ti,ab,kw

#5 "breast cancer" or "breast neoplasm" or "breast carcinoma" or “breast tumo*r” OR "breast adenocarcinoma":ti,ab,kw

#6 "ductal carcinoma in situ" or "DCIS":ti,ab,kw

#7 overdiagnosis and breast:ti,ab,kw

#8 MeSH descriptor: [Mass Screening] explode all trees

#9 screening:ti,ab,kw

#10 MeSH descriptor: [Economics] explode all trees

#11 MeSH descriptor: [Costs and Cost Analysis] explode all trees

#12 MeSH descriptor: [Cost-Benefit Analysis] explode all trees

#13 "economic evaluation":ti,ab,kw

#14 "cost benefit analysis" or "cost effectiveness analysis" or "cost utility analysis":ti,ab,kw

#15 health utilit* or health disutilit* or "health state" or preference or "time trade off" or "standard gamble" or "visual analogue scale" or "EQ5D":ti,ab,kw

#16 MeSH descriptor: [Quality of Life] explode all trees

#17 "quality of life" or QoL:ti,ab,kw (Word variations have been searched)

#18 MeSH descriptor: [Quality-Adjusted Life Years] explode all trees

#19 "quality adjusted life year*" or QALY*:ti,ab,kw

#20 MeSH descriptor: [Models, Economic] explode all trees

#21 "economic model":ti,ab,kw (Word variations have been searched)

#22 cost:ti,ab,kw (Word variations have been searched)

#23 #1 or #2 or #3 or #4 or #5 or #6 or #7

#24 #8 or #9

#25 #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or # 22

#26 #23 and #24 and #25

Social Sciences Citation Index (SSCI) and Science Citation Index Expanded (SCI-EXPANDED)

1. TS=("breast cancer" OR "breast neoplasm" OR "breast carcinoma" OR "breast adenocarcinoma" OR “breast tumo*r” OR "ductal carcinoma in situ" OR "DCIS" OR mammogra* or "breast screen*") or ts=(overdiagnosis AND breast)
2. TS=(economics OR "economic evaluation" OR cost OR "cost analysis" or "cost benefit analysis" OR "cost utility analysis" OR "cost effectiveness analysis" OR "quality of life" or "QoL" OR "quality adjusted life year*" or "QALY*" or health utilit* or disutilit* OR economic model)
3. TS=("cancer screen*" OR "screening")
4. TI=(breast or mammogra*)
5. #4 AND #3 AND #2 AND #1
6. #4 AND #3 AND #2 AND #1 AND LANGUAGE: (English)
7. #4 AND #3 AND #2 AND #1 AND LANGUAGE: (English) AND Document types: (Article)

APPENDIX 3: Final vignettes used in the empirical study

Example Scenario

Imagine the following example:

Example Scenario

YOUR general health

TODAY

Scenario A

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario A:

Scenario A

- The affected part of the breast is removed by day surgery.
- There is a small chance of some bleeding after surgery, or of developing an infection in the week following the operation. Some post-operative pain may occur, which may be permanent.
- You will have some scarring and shrinkage in the affected breast.
- You may feel more conscious of your body image and your sexual drive reduced. You feel anxious or depressed at times.
- The chance of the cells coming back in the next 10 years is 20%. That is, 20 out of every 100 women will experience a recurrence within 10 years.
    - If this happens you would require further surgery and radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
- You are told there is a 60% chance that you do not need the surgery as the cells may never progress into an invasive cancer during your lifetime. This means that:
    - 60 out of 100 women who choose to have surgery are overtreated and unnecessarily experience side-effects from treatment for an abnormality that would never have caused any harmful symptoms.
    - 40 out of 100 women would need the surgery at some point during their lifetime.

Scenario B

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario B:

Scenario B

- The affected part of the breast is removed by day surgery.
- You attend hospital to have radiotherapy for 3 weeks
    - You may experience redness, itchiness, pain or fatigue during this period
- There is a small chance of some bleeding after surgery, or of developing an infection in the week following the operation. Some post-operative pain may occur, which may be permanent.
- You will have some scarring and shrinkage in the affected breast.
- You may feel more conscious of your body image and your sexual drive reduced. You feel anxious or depressed at times.
- The chance of the cells coming back in the next 10 years is 10%. That is, 10 out of every 100 women will experience a recurrence within 10 years.
    - If this happens you would require further surgery and radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
- You are told there is a 60% chance that you do not need the surgery as the cells may never progress into an invasive cancer during your lifetime. This means that:
    - 60 out of 100 women who choose to have surgery are overtreated and unnecessarily experience side-effects from treatment for an abnormality that would never have caused any harmful symptoms.
    - 40 out of 100 women would need the surgery at some point during their lifetime.

Scenario C

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario C:

Scenario C

- You would undergo surgery to remove the entire breast, and are offered breast reconstruction. You may need to spend a few nights in hospital to recover.
- There is a small chance of some bleeding after surgery, or of developing an infection in the week following the operation. Some post-operative pain may occur, which may be permanent.
- You may experience several side-effects, including:
    - Scarring and breast disfigurement
    - A loss in body confidence and sexual drive
    - Depression and anxiety
- The chance of the cells coming back in the next 10 years is 1%. That is, 1 out of every 100 women will experience a recurrence within 10 years.
    - If this happens you would require more invasive surgery and radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
- You are told there is a 60% chance that you do not need the surgery as the cells may never progress into an invasive cancer during your lifetime. This means that:
    - 60 out of 100 women who choose to have surgery are overtreated and unnecessarily experience side-effects from treatment for an abnormality that would never have caused any harmful symptoms.
    - 40 out of 100 women would need the surgery at some point during their lifetime.

Scenario D

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario D:

Scenario D

- The abnormality identified on the mammogram is closely monitored by regular imaging.
- You have a mammogram once a year and are told the result by letter.
- You do not need to see a doctor unless any changes are picked up on your mammogram.
- You are advised about breast awareness and do regular breast checks at home.
- You may feel anxious or worried from time to time.
- The chance of the cells progressing in the next 10 years is 40%. That is, 40 out of every 100 women with low risk DCIS will experience some progression within 10 years.
    - If this happens you would require surgery and radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
    - You may need hormone tablets or a biopsy of the glands in the arm
- You are told there is a 60% chance that the cells will never progress into an invasive cancer during your lifetime. This means that:
    - 60 out of 100 women who choose to have monitoring will avoid overtreatment and unnecessary side-effects or pain from invasive treatment.
    - 40 out of 100 women would need to have surgery at some point during their lifetime.

Scenario E

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario E:

Scenario E

- The abnormality identified on the mammogram is closely monitored by regular imaging.
- You have a mammogram once a year and are told the result by letter.
- You do not need to see a doctor unless any changes are picked up on your mammogram.
- You are advised about breast awareness and do regular breast checks at home.
- You may feel anxious or worried from time to time.
- The chance of the cells progressing in the next 10 years is 20%. That is, 20 out of every 100 women with low risk DCIS will experience some progression within 10 years.
    - If this happens you would require surgery and possibly radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
    - You may need hormone tablets or a biopsy of the glands in the arm
- You are told there is an 80% chance that the cells will never progress into an invasive cancer during your lifetime. This means that:
    - 80 out of 100 women who choose to have monitoring will avoid overtreatment and unnecessary side-effects or pain from invasive treatment.
    - 20 out of 100 women would need to have surgery at some point during their lifetime.

Scenario F

Imagine you attend a routine breast screening mammogram. The x-ray identifies some abnormal cells, but they are confined to the milk ducts of the breast.

The abnormal cells are described by the doctors as "low risk DCIS" as the risk of them developing into a cancer is very **low**.

You are offered the treatment in Scenario F:

Scenario F

- The abnormality identified on the mammogram is closely monitored by regular imaging.
- You have a mammogram once a year and are told the result by letter.
- You do not need to see a doctor unless any changes are picked up on your mammogram.
- You are advised about breast awareness and do regular breast checks at home.
- You may feel anxious or worried from time to time.
- The chance of the cells progressing in the next 10 years is 10%. That is, 10 out of every 100 women with low risk DCIS will experience some progression within 10 years.
    - If this happens you would require surgery and possibly radiotherapy but there is no increased risk of dying as a result of this diagnosis and treatment.
    - You may need hormone tablets or a biopsy of the glands in the arm
- You are told there is a 90% chance that the cells will never progress into an invasive cancer during your lifetime. This means that:
    - 90 out of 100 women who choose to have monitoring will avoid overtreatment and unnecessary side-effects or pain from invasive treatment.
    - 10 out of 100 women would need to have surgery at some point during their lifetime

Scenario G

Imagine you attend a routine breast screening mammogram. The x-ray identified some abnormal cells, but they were confined to the milk ducts of the breast.

The abnormal cells were described by the doctors as “low risk DCIS” and were closely monitored by yearly mammograms.

There were some changes noted at your repeat mammogram. Further investigation and biopsy showed that the abnormal cells had progressed.

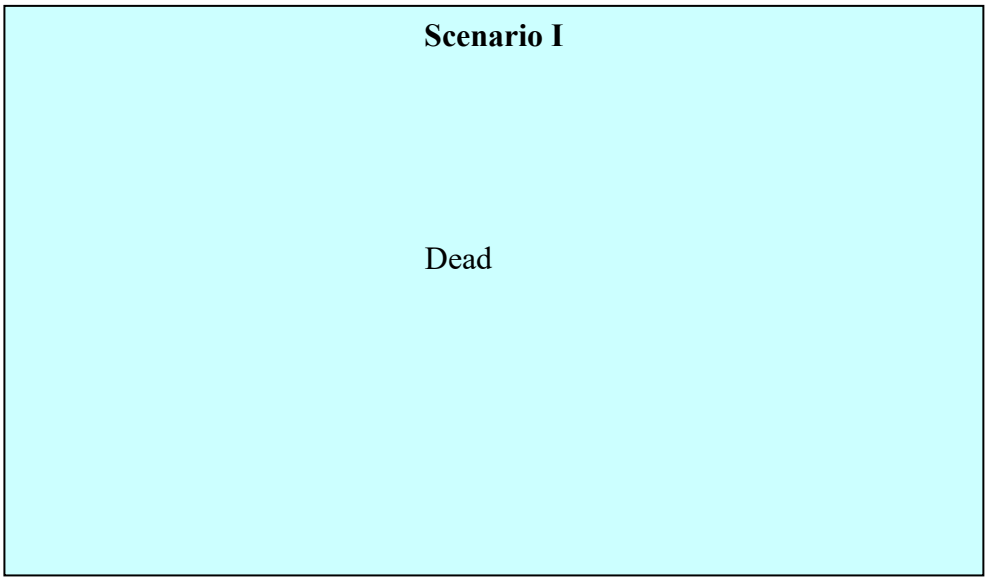
Imagine the treatment in Scenario G:

Scenario G

- The abnormality identified on the mammogram was closely monitored.
- Your repeat mammogram showed some suspicious changes. Further investigation and biopsy confirmed that the abnormal cells have progressed to an early breast cancer.
- You are worried about the delay in treatment.
- You require surgery to remove part of breast and some of the glands in the armpit. You may need to spend overnight in hospital to recover.
- There is a small chance of some bleeding after surgery, or of developing an infection in the week following the operation. Some post-operative pain may occur, which may be permanent.
- You may also need attend hospital to have radiotherapy for 3 weeks and take hormone tablets every day.
- You may experience several side-effects, including:
    - Scarring and breast disfigurement
    - Potentially significant swelling and discomfort of the arm on the affected side
    - The radiotherapy may cause additional redness, swelling, pain or fatigue
    - Side effects of the tablets may include hot flushes and joint pains
    - A loss in body confidence and sexual drive
    - Anxiety regarding the possibility the cancer will return
- The chance of the cancer coming back in the next 10 years is 10%. That is, 10 out of every 100 women will experience a recurrence within 10 years.
    - If this happens you would require further treatment
    - The risk of dying from breast cancer would be <10%

Scenario H:

Scenario I:



APPENDIX 3: Empirical study questionnaire

UNIVERSITY OF BIRMINGHAM

THE UNIVERSITY OF MELBOURNE

Valuing the benefits and risks associated with treating low risk breast conditions

Questionnaire

Date:
Participant ID:

Introduction

Welcome and thank you for choosing to participate in this research.

Before you begin, the following pages will tell you a bit more about what will happen during the study and provide you with some basic information about low risk breast cancer and the types of treatment we will ask you to value.

Background

Overdiagnosis refers to the diagnosis of breast cancers which would not cause symptoms or harm in a woman's lifetime. At present, it is impossible for doctors to know which cancers will spread and which will not, so they are all treated straight away.

Up to 20% of breast cancers identified by breast screening programs are pre-cancerous conditions, such as ductal carcinoma in situ (DCIS). Ductal carcinoma in situ is called a 'pre-cancer' because there are abnormal cells in the milk ducts of the breast but not in any other breast tissue. Doctors can classify DCIS as high, intermediate or low grade. Because high grade DCIS is more likely to become breast cancer, it is treated as if it is an invasive cancer. But many doctors are not sure that low or intermediate DCIS would ever become an invasive cancer because it may grow so slowly or not at all during most women's lifetimes.

Most women have surgery to remove DCIS. This may be surgery to remove the breast (a mastectomy). Or it may be surgery to remove just the area of the breast containing the DCIS (breast conserving surgery), with or without radiotherapy or hormone therapy. There are also trials looking at actively monitoring low risk DCIS. These women will have a mammogram once a year to monitor the disease instead of undergoing invasive surgeries which they may not need.

There is a small chance that women in either group will develop higher grade DCIS or invasive breast cancer that would need more invasive treatment. The risk of disease progression may be slightly higher in the monitoring group. However, there are risks associated with having surgery compared to monitoring, such as bleeding, wound infections and physical changes to the breast. There are also side effects of radiotherapy and hormone therapy.

It is not known what women would feel about monitoring low risk pre-cancers compared to having invasive surgery. We expect different women will think differently about the type of treatment they would prefer. We are keen to know *your* individual view so that we can better plan breast cancer treatment services and reduce the harm of unnecessary invasive treatments for low risk disease.

Outline of the interview

This study is looking to better understand women's preferences and see how women value different types of treatment for low risk pre-cancers identified through breast cancer screening.

You will be given a set of imaginary scenarios to read. Each will describe a different treatment for low risk DCIS. The scenarios might describe health states following surgery, radiotherapy or monitoring of a low risk DCIS.

The interview will be split into three sections. The first two sections will involve different ways of measuring how you value each of the treatments described. The last section will ask some basic some questions about you, to help us analyse the results.

Before you begin, we would like you to answer some questions about risks as a warm-up exercise.

1. Imagine a fair coin is tossed 1,000 times. What is your best guess about how many times the coin would land heads up in 1,000 flips?

   Write your answer here: ……

2. Imagine the chance of winning a prize in a competition is 10%. How many people out of 100 entering the competition would win a prize?

   Write your answer here: ……

3. Suppose you had a choice between a cash bonus of either i) $100 now OR ii) $150 a year from now. Please tick the box to indicate which alternative you would choose:
   - □ $100 now
   - □ $150 a year from now

4. Suppose you had a choice between i) winning a cash bonus of $100 OR ii) a 50-50% chance of winning $50 or $200. Which option would you choose?
   - □ $100 for certain
   - □ 50-50 chance of winning either $50 or $200

We will now move on to think about the benefits and risks relating to the treatment of low risk DCIS. You will be given a set of scenarios describing different treatments.

Please go through each of the scenarios carefully, one at a time. You will notice that the risks involved change slightly between each of the treatments described.

Section One

In this section we would like you to say how good or bad you think each scenario is. To help you measure this, there is a line (like a thermometer) which runs from the best health state you can imagine (marked at 100) to the worst health state you can imagine (marked at 0).

For each scenario, please mark the letter on whichever point on the scale indicates how good or bad the scenario would be if you had to be in that health state for the next 10 years.

We will run through an example using your own general health to familiarise you with the scale first.

You will then be asked to decide how each scenario impacts upon five different areas of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. For each scenario (A-G), tick one box under each of the headings that describes how severely each scenario impacts upon each of the areas of health.

Example

We would like to know how good or bad your general health is TODAY.

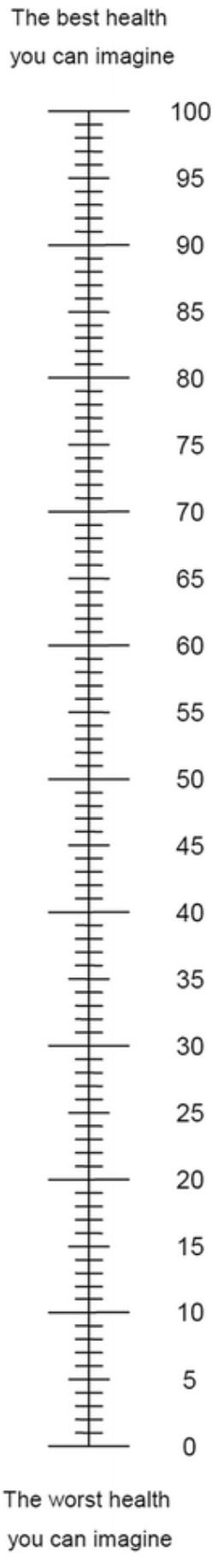
The scale is numbered from 0 to 100.

100 means the **best** health you can imagine.
0 means the **worst** health you can imagine.

Please mark an X on the scale to indicate how your general health is today.

Please write down the number you marked on the scale on the dotted line below.

YOUR HEALTH TODAY = ……



Question 1

Please read each of the scenarios provided. We would like to know how good or bad you think each scenario is.

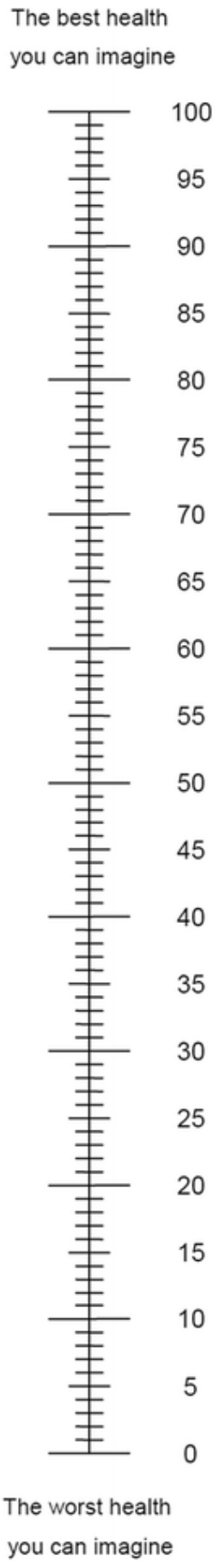
The scale is numbered from 0 to 100.

100 means the **best** health you can imagine.
0 means the **worst** health you can imagine.

Please mark a letter on the scale for each scenario to show how good or bad you think it is compared to the best (100) or worst (0) imaginable health possible.

Write the corresponding number for each scenario on the dotted line below.

A = ……
B = ……
C = ……
D = ……
E = ……
F = ……
G = ……
H = ……
I = ……



Under each heading, tick ONE box that best describes the severity for each scenario (A-G).

	A	B	C	D	E	F	G
MOBILITY							
I have no problems in walking about	❑	❑	❑	❑	❑	❑	❑
I have slight problems in walking about	❑	❑	❑	❑	❑	❑	❑
I have moderate problems in walking about	❑	❑	❑	❑	❑	❑	❑
I have severe problems in walking about	❑	❑	❑	❑	❑	❑	❑
I am unable to walk about	❑	❑	❑	❑	❑	❑	❑
SELF-CARE							
I have no problems washing or dressing myself	❑	❑	❑	❑	❑	❑	❑
I have slight problems washing or dressing myself	❑	❑	❑	❑	❑	❑	❑
I have moderate problems washing or dressing myself	❑	❑	❑	❑	❑	❑	❑
I have severe problems washing or dressing myself	❑	❑	❑	❑	❑	❑	❑
I am unable to wash or dress myself	❑	❑	❑	❑	❑	❑	❑
USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*							
I have no problems doing my usual activities	❑	❑	❑	❑	❑	❑	❑
I have slight problems doing my usual activities	❑	❑	❑	❑	❑	❑	❑
I have moderate problems doing my usual activities	❑	❑	❑	❑	❑	❑	❑
I have severe problems doing my usual activities	❑	❑	❑	❑	❑	❑	❑
I am unable to do my usual activities	❑	❑	❑	❑	❑	❑	❑
PAIN / DISCOMFORT							
I have no pain or discomfort	❑	❑	❑	❑	❑	❑	❑
I have slight pain or discomfort	❑	❑	❑	❑	❑	❑	❑
I have moderate pain or discomfort	❑	❑	❑	❑	❑	❑	❑
I have severe pain or discomfort	❑	❑	❑	❑	❑	❑	❑
I have extreme pain or discomfort	❑	❑	❑	❑	❑	❑	❑
ANXIETY / DEPRESSION							
I am not anxious or depressed	❑	❑	❑	❑	❑	❑	❑
I am slightly anxious or depressed	❑	❑	❑	❑	❑	❑	❑
I am moderately anxious or depressed	❑	❑	❑	❑	❑	❑	❑
I am severely anxious or depressed	❑	❑	❑	❑	❑	❑	❑
I am extremely anxious or depressed	❑	❑	❑	❑	❑	❑	❑

Section Two

In this section, you will be asked to choose between two imaginary options using the scenarios from section one. The choices are labelled Option 1 and Option 2. Note that the scenarios in Option 2 are different in every question.

You will be asked to choose between two imaginary options. For Option 1, we would like you to imagine there is an imaginary treatment that could restore you to perfect health for the next 10 years but at a risk of death. For Option 2, we would like you to imagine that you will live in the health state described on the card for the next 10 years and then you will die.

We would like you to consider which of the options you would prefer: the gamble in Option 1 or certainty of Option 2?

Your interviewer will explain how the next section of the interview will run. They will work through an example so that you are familiar with the task and chance board before going through each of the scenarios with you.

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

SHOWCARD ? Answer 1 or 2?

Section Three

In this section, we would like you to answer a few questions about yourself. This is so that we can try to understand the different choices women might make when we analyse the results.

1. **What is your current age?** ……. years old

2. **What is the main language you speak at home?** *(Please tick one box only)*
   - □ English □ Arabic □ Cantonese
   - □ Greek □ Italian □ Mandarin
   - □ Vietnamese □ Prefer not to say □ Other (please specify) ……

3. **What is your ethnic background?** *(That is, the group or groups with whom you share a common heritage.)*
   - □ Australian □ South East Asian □ North American
   - □ Aboriginal □ British □ Central American
   - □ Torres Strait Islander □ Irish □ Southern American
   - □ New Zealander □ Northern European □ Hispanic
   - □ Maori □ Western European □ Other: ……
   - □ Pacific Islander □ Southern European
   - □ North East Asian □ Middle Eastern/North African
   - □ Southern Asian □ Sub-Saharan African

4. **What gender do you identify with?** *(Please tick one box only)*
   - □ Female □ Male □ Other □ Prefer not to say

5. **What is your highest level of education?** (*Please tick one box only)*
   - □ No formal education □ Primary school □ Secondary school
   - □ Certificate or diploma □ Undergraduate degree □ Postgraduate degree
   - □ Prefer not to say

6. **Which of the following best describes your main activity?** *(Please tick one box only)*
   - □ In full or part-time employment or self-employment
   - □ Seeking work
   - □ Retired
   - □ Student
   - □ Carer/volunteer
   - □ Other (please specify) ……

7. **How would you define your relationship status?** *(Please tick one box only)*
   - □ Not in a relationship □ Married or in a steady relationship (de facto)
   - □ Separated □ Divorced
   - □ Widowed □ Prefer not to say

8. **What is your total yearly household income?** *(Please tick one box only)*
   - □ $1- $7,799 per year □ $41,600 - $51,999 per year
   - □ $7,800 - $15,599 per year □ $52,000 - $64,999 per year
   - □ $15,600 - $20,799 per year □ $65,000 - $77,999 per year
   - □ $20,800 - $25,999 per year □ $78,000- $90,999 per year
   - □ $26,000 - $33,799 per year □ $91,000 per year or higher
   - □ $33,800 - $41,599 per year □ Prefer not to say

9. **Have you ever been invited to attend the breast cancer <u>screening</u> programme?**
   - □ Yes
   - □ No

10. If you have been invited to breast screening, did you attend your <u>screening</u> mammogram?

□ Yes
□ No
□ N/A

11. Have you ever had a breast biopsy? *(A biopsy is when breast tissue is taken to look under the microscope)*

□ Yes
□ No

12. Which of the following statements describes you best? *(Tick all that apply)*

□ I have been diagnosed with ductal cancer in-situ (DCIS) or lobular cancer in-situ (LCIS) (both non-invasive)
□ I have been diagnosed with early stage breast cancer (where the breast cancer is contained within your breast and/or lymph nodes)
□ I have been diagnosed with metastatic breast cancer (breast cancer that has spread beyond the breast tissue and lymph nodes to distant parts of the body)
□ I am a family member/friend of someone diagnosed with breast cancer
□ I am a health professional who works with patients who have breast cancer
□ Other ……..
□ Prefer not to say

13. **If you have had breast cancer, how long has it been since you were first diagnosed?**
(Please tick one box only)

□ 0 to 1 year □ > 1 year to 2 years □ > 2 years to 3 years
□ > 3 years to 4 years □ >4 years to 9 years □ 10 years +
□ Prefer not to say □ N/A

14. If you have had breast cancer, which of the following treatments did you receive? *(tick all that apply)*

□ Breast conserving surgery (lumpectomy) □ Mastectomy (whole breast removed)
□ Sentinel lymph node surgery □ Axillary clearance
□ Radiotherapy □ Chemotherapy
□ Hormone tablets (e.g. Tamoxifen) □ Other: ……
□ None of the above

15. To what extent are you concerned about your own risk of breast cancer? *(Please tick one box only)*

□ I don't worry about breast cancer at all
□ I worry about breast cancer a little
□ I worry about breast cancer some of the time
□ I worry about breast cancer a lot
□ I always worry about breast cancer
□ Unsure

16. How many babies have you given birth to, if any?

…….. number of babies

17. Which statement best describes your smoking status?

□ Current smoker □ Ex-smoker □ Never smoked

18. In the last 12 months, how often have you had an alcoholic drink of any kind?

□ Every day □ 5-6 days per week □ 3-4 days per week
□ 1-2 days per week □ 2-3 days per month □ About 1 day per month
□ Less often □ I do not drink

19. How easy or difficult did you find making choices in section two today? *(Please tick one box only)*

□ Very easy □ Easy □ Neither easy nor difficult
□ Difficult □ Very difficult

20. Would you consider active monitoring (yearly mammograms) if you had a <u>new</u> diagnosis of low risk breast disease?

□ Yes
□ No
□ Unsure

21. Is there anything else you would like to add regarding the questions we have asked you?

……..
……..
……..
……..

The questionnaire has been completed!

Thank you for your time.

For any further information or queries, please contact:

Hannah Bromley (Doctoral research student)
Melbourne School of Population and Global Health, 207 Bouverie Street, University of Melbourne, Parkville, VIC 3053**.**
Email: [redacted]

Please tick the box if you would like to take part in more research relating to this project □

If you wish to receive a copy of the results, please provide your email address in the space below:

……..

Plain Language Statement

Main study

Faculty of Medical, Dental and Health Sciences

Valuing the benefits and risks associated with treating low risk breast conditions

Primary Researcher: Dr Hannah Bromley (Doctoral research student)
Email: [redacted]

Supervisory researchers:
Dr C Nickson, (University of Melbourne), Prof B Mann (Victorian Comprehensive Cancer Centre), Associate/Prof D Petrie (Monash University) Prof D Rea (University of Birmingham), Prof T Roberts, (University of Birmingham)
Email: cnickson@unimelb.edu.au

Introduction

You are invited to participate in this PhD student research project which aims to investigate how women value the different treatments for low risk breast conditions. This is because you have either been referred through the Breast Cancer Network Australia or because you have responded to an advertisement. This plain language summary tells you about the research project. It explains what it is involved to help you to decide if you want to take part.

Please take the time to read this information carefully. You may ask questions about anything you don't understand or want to know more about. Before deciding whether to take part, you might want to talk it through with a relative, friend or your local health worker. You will be given a copy of this plain language summary to keep.

Your participation is this research voluntary. If you don't wish to take part, you don't have to. If you begin participating, you can also stop at any time without needing to give a reason.

What is the purpose of this research project?

The reason this project is being undertaken is that we want to know more about how women value the benefits and risks associated with breast screening programs. In particular, we are interested in seeing how women make choices between the different treatments for low risk breast conditions that have been picked up by routine screening mammograms (x-rays of the breast).

This is because many women do not know about the range of treatments available for low risk breast conditions. We want to help women make better informed choices by exploring how they think the imaginary treatment of a breast cancer might impact upon their quality of life.

We will measure your preferences by asking you to choose between imaginary scenarios relating to treatments for low risk pre-cancers (sometimes referred to as low risk '**ductal carcinoma in situ' or 'DCIS'**). These scenarios might describe having surgery (where either the affected part or the entire breast is removed), radiotherapy or monitoring of the abnormal area with a yearly mammogram.

Why is this research important?

It is not known how women would feel about monitoring low risk breast cancers compared to undergoing invasive surgery. We expect different women will think differently about the type of treatment they would prefer.

We are keen to know *your* individual view so that we can better plan breast treatment services and reduce the harm of unnecessary harmful treatments for DCIS. By finding out how women prefer to manage screen detected breast cancers, the research findings may be used to justify changes to the way low risk cancers are screened and treated.

What will I be asked to do?

You will be invited to attend a face-to-face interview with the research student at the University of Melbourne or a convenient local venue. For those unable to travel the interview may arranged via video-communication software such as Skype. During the interview, you will be asked to complete a short questionnaire asking about your preferences for a number of different treatments for DCIS. The questionnaire will take around 60 minutes to complete and you will have the opportunity to ask the research student any questions at any time during the interview. The questionnaire will have been tested by doctors, patients and women at the University of Melbourne beforehand to make sure the questions are easy enough to understand and answer.

At the interview we will remind you about the details of the project and ask you to sign consent form stating that you are happy to take part. You will then be asked to read through a set of imaginary scenarios to help value the different treatments for DCIS and to answer some basic questions about yourself. We may also ask a small group of women if they would be happy to audio-record their interview, to help us with the analysis. These women will be encouraged to "think out aloud" and describe their feelings about the treatments as they complete the questionnaire, so that we can try to understand how different women might feel about choosing between the different choices available.

The questionnaire is split into three sections:

1) A rating scale and domains of health exercise

You will be asked to rate each of the treatments on a scale from 0 to 100. You will then be asked to think about how each of the treatments might affect different aspects of your health (mobility, self-care, usual activities, pain and mood).

2) A gambling exercise

We will ask you to choose between different treatments by varying the benefits and risks in each of the descriptions, one at a time.

3) A few questions about yourself

Finally, we will ask you a few questions about yourself, simply to help with the analysis.

What are the possible benefits?

This is your opportunity to share your views with breast cancer researchers. We think it is important that women's preferences for treatments are included in the decision making process when it comes to planning breast cancer services. Although there may be no clear immediate benefits to you from your participation, we hope that the results of this project can be used to improve local breast cancer services in Victoria in the future by contributing to the literature. The values obtained from this study will be used to inform an economic evaluation of breast cancer screening and treatment and may be used to justify changes to the way breast cancer resources are allocated in Australia.

Your responses will be used to inform a PhD thesis at the University of Melbourne. The results may be presented at scientific conferences, to policy makers at local breast cancer meetings or the public through the media (e.g. in newspaper articles). The main findings may also be published in scientific journals.

What are the possible risks?

Sometimes talking about cancer can be a sensitive or upsetting topic. We will not ask you to talk about your own experiences and you will be free to withdraw from the study at any time, without needing to give a reason. We can direct you to breast cancer support services if you have any further concerns or questions about breast cancer screening or treatment services in the area.

Some of the interviews may be audio-recorded, with your permission, and then typed up to ensure that we record your comments accurately. We will remove any personal details from the transcript and use a pseudonym to refer to you in any published material to keep everything anonymous. Due to the small size of the sample, there is a small chance that participants may be identifiable in published material, despite using pseudonyms in the write up of the results. The information collected in the questionnaire is also anonymised, grouped and presented as averages so that no personal information is published.

Do I have to take part?

No. Participation is completely voluntary. If you don't wish to take part, you don't have to do so. If you do decide to take part and later change your mind, you are free to withdraw (quit) at any time. If you decide to leave the project, the researchers would like to keep the information about you that has been collected. This is to help them make sure that the results of the research can be measured properly. If you don't want them to do this, you must tell them before you withdraw from the study. Your decision whether to take part or not, or to take part and then withdraw, will not affect your relationship with the researchers.

Will I hear about the results of this project?

A plain language summary of the general findings of this research project will be made available to all participants via email or post, if they have provided consent to receive any further communication from the research team. Please leave your contact details if you are interested in hearing about the results of the project. It is expected that the results will be made available following the completion of the PhD thesis and publication of the results in scientific journals in late 2018.

What will happen to information about me?

A summary of the Review and Survey group demographics will be provided by the BCNA. All information will be kept entirely confidential. We will not be publishing any personal details and data will be stored securely on password protected servers or under lock-and-key at the University of Melbourne.

All information collected in the questionnaire is anonymised. The results from the questionnaire will be grouped and presented as averages so that no personal information is published. Any data collected in the interview or questionnaire will be destroyed after five years following publication as per University of Melbourne data management and ethics guidelines.

In accordance with the relevant Australian and/or Victorian privacy and other relevant laws, you have the right to access the information collected and stored by the researchers about you. Please contact one of the researchers if you would like to access your information.

Who is funding this project?

The study is funded as part of a joint PhD programme between the Universities of Melbourne (Australia) and Birmingham (UK) to try and improve breast cancer screening and treatment services worldwide. The PhD student is supervised by a team of experienced doctors and University of Melbourne researchers to make sure the research is of the highest quality.

Where can I get further information?

If you would like more information about the project, please contact the researchers;
Hannah Bromley (doctoral research student),
Email: [redacted] Melbourne School of Population & Global Health, University of Melbourne, 207 Bouverie Street, Parkville, Melbourne, VIC 3053.

Who can I contact if I have any concerns about the project?

This research project has been approved by the Human Research Ethics Committee of The University of Melbourne. If you have any concerns or complaints about the conduct of this research project, which you do not wish to discuss with the research team, you should contact the Manager, Human Research Ethics, Research Ethics and Integrity, University of Melbourne, VIC 3010. Tel: +61 3 8344 2073 or Email: HumanEthics-complaints@unimelb.edu.au. All complaints will be treated confidentially. In any correspondence please provide the name of the research team or the name or ethics ID number of the research project.

Where can I get more information if I have questions about breast cancer?

There are some excellent resources available if you have any further questions or concerns relating to breast cancer or screening services as a result of participating in this project:

Breast Cancer Network Australia:
An excellent resource talking more about breast screening, breast cancer and women's health and well-being. Website: https://www.bcna.org.au. Telephone: 1800 789 209

Breast Screen Victoria
A resource specifically listing the benefits and risks associated with breast screening programs. The team are happy to provide advice about local breast screening services. Website: https://www.breastscreen.org.au Telephone: 13 20 50.

Lifepool
A source of information about your participation in the Lifepool cohort. Website: http://www.lifepool.org/ Telephone: 1800 198 082.

Cancer Council Victoria
A resource dedicated to local cancer facilities and service in Victoria. Website: http://www.cancervic.org.au/ Telephone: 13 11 20

The Royal Women's Hospital
A summary of the local cancer resources available in the Melbourne Metropolitan area and guidelines on local breast cancer screening and treatment policies. Website: https://www.thewomens.org.au/patients-visitors/clinics-and-services/womens-cancers/breast-cancer/ Telephone: 03 8345 3561.

Consent Form for the Main Study

A joint research project between the University of Melbourne (Australia) and University of Birmingham (UK) School of Medical, Dental and Health Sciences

THE UNIVERSITY OF MELBOURNE

UNIVERSITY OF BIRMINGHAM

Title: *Valuing the benefits and risks associated with treating low risk breast conditions*

Primary Researcher: Hannah Bromley (Doctoral research student)

[redacted]

Supervisory researchers:
Dr C Nickson, (University of Melbourne), Prof B Mann (Victorian Comprehensive Cancer Centre), Associate/Prof D Petrie (Monash University) Prof D Rea (University of Birmingham), Prof T Roberts, (University of Birmingham)
Email: cnickson@unimelb.edu.au

This consent form relates to the participation in a PhD student project valuing different types of treatment for low risk breast conditions. It involves a short interview and a questionnaire to determine your preferences for the different treatments described.

Please read the form and if you are happy to proceed, fill in your details at the bottom of the form.

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I acknowledge that the possible effects of participating in this research project have been explained to my satisfaction.

3. I understand that my participation is voluntary and that I am free to withdraw from this project anytime without explanation or prejudice and to withdraw any unprocessed data that I have provided.

4. I understand that the information provided will be used for research only and that I will not be identified in any way in the analysis and reporting of results.

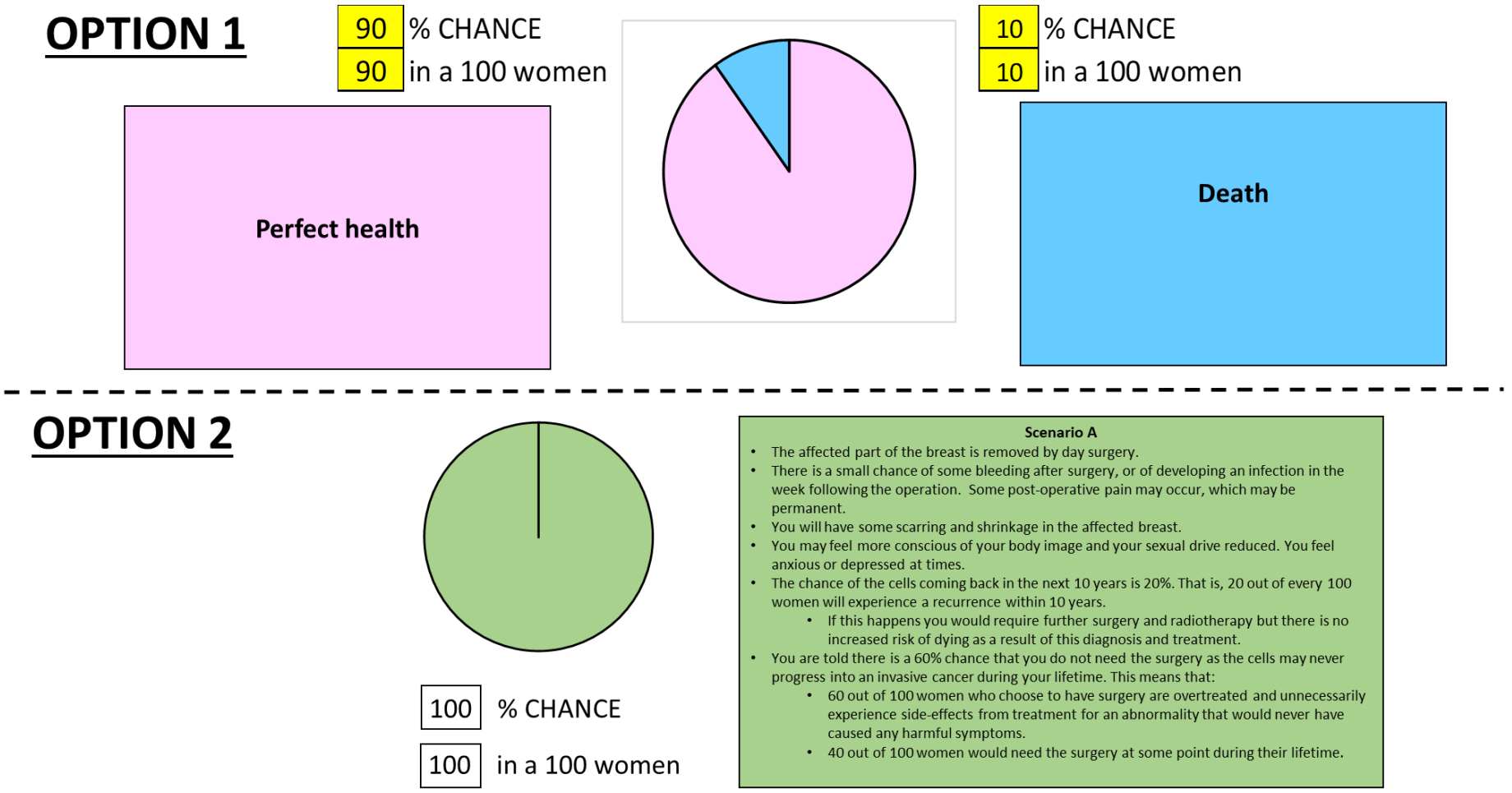
5. I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements and any data collected will be stored as per University of Melbourne guidelines.

6. I understand what is involved and agree to take part in the study.

Name of Participant: _____

Participant Signature: _____ **Date:** _____

APPENDIX 5: Probability wheel and chance board from the empirical study

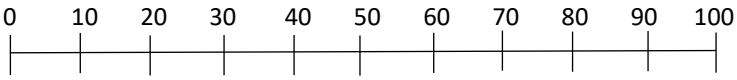


APPENDIX 6: Standard gamble scoring sheet

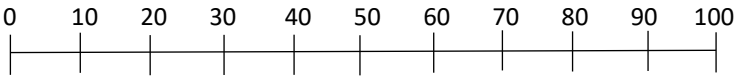
Record sheet for the Standard Gamble:
Date: ……
Participant ID: ……
1. Time at start of SG: ……h
2. Time at end of SG: ……h
3. Time taken to complete the SG: ……minutes
Did the respondent understand the exercise the first time? Y/N

Health state 1

Better than Death:



Worse than Death

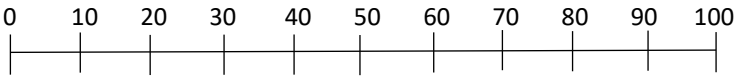


VALUE: +/- ☐☐ . ☐☐

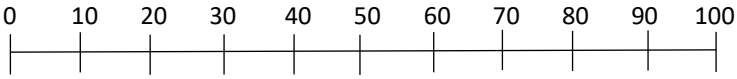
Record verbatim response:
……
……
……

Health state 2

Better than Death:



Worse than Death

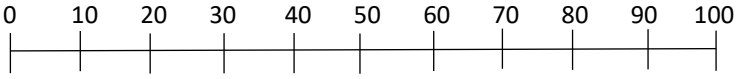


VALUE: +/- ☐☐ . ☐☐

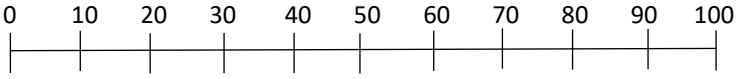
Record verbatim response:
……
……
……

Health state 3

Better than Death:



Worse than Death



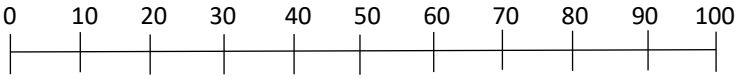
VALUE: +/- ☐☐ . ☐☐

Record verbatim response:

……
……
……

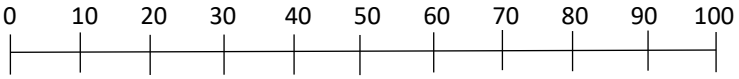
Health state 5

Better than Death:



Worse than Death

VALUE: +/- ☐☐ . ☐☐

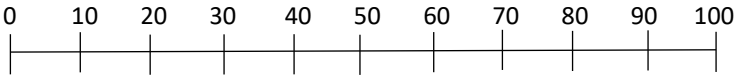


Record verbatim response:

……
……
……

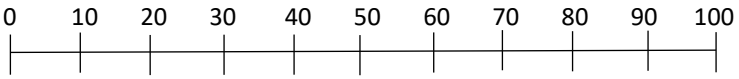
Health state 6

Better than Death:



Worse than Death

VALUE: +/- ☐☐ . ☐☐

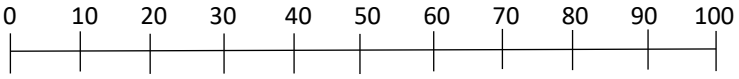


Record verbatim response:

……
……
……

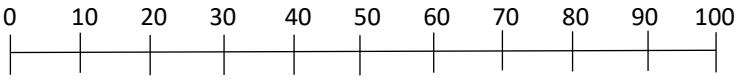
Health state 7

Better than Death:



Worse than Death

VALUE: +/- ☐☐ . ☐☐

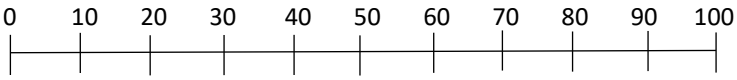


Record verbatim response:

……
……
……

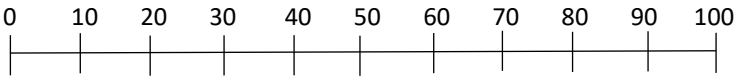
Health state 8

Better than Death:



Worse than Death

VALUE: +/- ☐☐ . ☐☐



Record verbatim response:

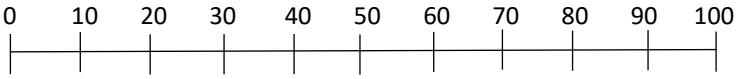
……

……

……

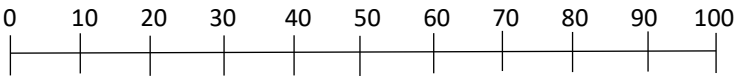
Health state (SPARE)

Better than Death:



Worse than Death

VALUE: +/- ☐☐ . ☐☐



Record verbatim response:

……

……

……

APPENDIX 7: Ethical approval for the empirical study

04 December 2017

Dr Carolyn Nickson
Centre for Molecular, Environmental, Genetic and Analytic Epidemiology
Melbourne School of Population and Global Health
The University of Melbourne

Dear Dr Nickson

I am pleased to advise that the Health Sciences Human Ethics Sub-Committee approved the following Project:

Project title:	**Identifying and quantifying the benefits and harms associated with treating ductal carcinoma in situ (DCIS) for use in the economic evaluation of breast screening programmes**
Researchers:	**Prof D Rea, Prof T Roberts, Dr C A Nickson, Professor B Mann, A/Prof D Petrie, H Bromley**
Ethics ID:	**1750252**

The Project has been approved for the period: **04-Dec-2017 to 31-Dec-2018**

It is your responsibility to ensure that all people associated with the Project are made aware of what has actually been approved.

Research projects are normally approved to 31 December of the year of approval. Projects may be renewed yearly for up to a total of five years upon receipt of a satisfactory annual report. If a project is to continue beyond five years a new application will normally need to be submitted.

Please note that the following conditions apply to your approval. Failure to abide by these conditions may result in suspension or discontinuation of approval and/or disciplinary action.

(a) **Limit of Approval:** Approval is limited strictly to the research as submitted in your Project application.

(b) **Variation to Project:** Any subsequent variations or modifications you might wish to make to the Project must be notified formally to the Human Ethics Sub-Committee for further consideration and approval. If the Sub-Committee considers that the proposed changes are significant, you may be required to submit a new application for approval of the revised Project.

(c) **Incidents or adverse effects:** Researchers must report immediately to the Sub-Committee anything which might affect the ethical acceptance of the protocol including adverse effects on participants or unforeseen events that might affect continued ethical acceptability of the Project. Failure to do so may result in suspension or cancellation of approval.

(d) **Monitoring:** All projects are subject to monitoring at any time by the Human Research Ethics Committee.

(e) **Annual Report:** Please be aware that the Human Research Ethics Committee requires that researchers submit an annual report on each of their projects at the end of the year, or at the conclusion of a project if it continues for less than this time. Failure to submit an annual report will mean that ethics approval will lapse.

(f) **Auditing:** All projects may be subject to audit by members of the Sub-Committee.

If you have any queries on these matters, or require additional information, please contact me using the details below.

Please quote the ethics registration number and the title of the Project in any future correspondence.

On behalf of the Sub-Committee I wish you well in your research.

Yours sincerely

[redacted]

Ms Hilary Young - Secretary
Health Sciences HESC
Phone: 03 8344 8595, Email: hilary.young@unimelb.edu.au

RESEARCH, INNOVATION & COMMERCIALISATION
Incorporating UoM Commercial Ltd and the School of Melbourne Custom Programs
The University of Melbourne, Victoria 3010, Australia
T: +61 3 8344 1539 (external) **T:** 40777 (internal) **W:** unimelb.edu.au/research

THE UNIVERSITY OF MELBOURNE

unimelb.edu.au

Date…… 2018

Dear ……..,

Thank you for being part of the **Life**pool study. You might remember that when you joined **Life**pool we let you know that you may be asked to participate in **Life**pool research again in the future. We also let you know that it is completely up to you to decide if you would like to take part in any new studies. Today I am writing to invite you to take part in a study looking at the different treatment choices for low risk breast cancers picked up by screening. The aim of the research is to help women make better informed choices in the future by finding out how different treatments might impact upon quality of life.

This research is being conducted by Dr Hannah Bromley who is a medical doctor from Birmingham UK and currently undertaking her PhD studies at the University of Melbourne. You can read more about this research by clicking here: [insert online link to PLS]

Being involved will mean either coming to the University of Melbourne in Parkville for an interview lasting approximately 1 hour, or another venue more convenient to you. During the interview you will be asked to fill in a questionnaire.

This study forms part of Dr Bromley's work towards learning more about how women value and balance the benefits and risks associated with breast cancer screening. In particular, the research team are interested in how women choose between the different treatments for low risk breast disease identified by breast screening.
This session should take about 1 hour of your time. Refreshments will be provided.

Please respond to myself or Dr Hannah Bromley at [redacted] if you are interested in participating.

If you are not interested, you can ignore this invitation. You do not have to participate in this study to remain in **Life**pool.

Please feel free to contact myself or Dr Bromley if you have any queries about this.

With very best wishes and thanks again for your support of **Life**pool,

Lisa Devereux
Lifepool Manager

Lisa.Devereux@petermac.org
[redacted]

Want to improve breast cancer services? Have your say!

Dear Review & Survey Group Member

Researchers at the University of Melbourne and Birmingham are inviting you to participate in a study looking different treatment choices for low risk ductal carcinoma in situ (DCIS).

Please note: You do NOT have to have had a diagnosis of DCIS to participate in the research study.

Who can take part?

You can participate in this study if you:

- Are over the age of 18 years
- Are comfortable reading and writing in English
- Have **OR** have not been diagnosed with breast cancer (the research team is interested in seeing what people diagnosed with breast cancer and people who have never had a breast cancer diagnosis think).

What does the study involve?

The researchers are inviting you to attend a short interview at the University of Melbourne or a convenient local venue. The session should take no more than 60 minutes and refreshments will be provided.

During the interview, you will be asked to fill in a questionnaire describing a number of imaginary scenarios relating to possible treatments for DCIS, such as surgery, radiotherapy or active monitoring (e.g. having regular scans, such as yearly mammograms). For each if the imaginary scenarios, the researchers will ask how you might decide what treatment(s) to have.

The aim of the research is to help women make better informed choices about treating DCIS in the future by exploring how different treatments might impact upon quality of life.

How do I find out more about the project?

If you would like to participate or would like more information about the project, please contact Hannah Bromley (Doctoral Research student) at the University of Melbourne by emailing [redacted]

For more information, please also read the **Participant Information Statement**.

Thanks very much for taking the time to consider this opportunity.

Warm regards,

Lisa Morstyn
Policy Officer

APPENDIX 8: Pairwise correlations for the VAS, SG and EQ-5D-5L

Visual analogue scale:

	A BCS alone		B BCS & radio		C Mastectomy		D AM (40%)		E AM (20%)		F AM (10%)		G Progressed	
Health state	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**
A: BCS alone														
Patients	1.000													
Nonpatients	1.000													
B: BCS & radiotherapy														
Patients	0.537	.000	1.000											
Nonpatients	0.883	.000	1.000											
C: Mastectomy +/- recon.														
Patients	0.201	.052	0.371	.000	1.000									
Nonpatients	0.754	.000	0.795	.000	1.000									
D: AM (40%)														
Patients	0.219	.035	0.180	.082	0.089	.392	1.000							
Nonpatients	0.358	.001	0.245	.031	0.289	.010	1.000							
E: AM (20%)														
Patients	0.222	.032	0.145	.162	0.062	.556	0.970	.000	1.000					
Nonpatients	0.285	.011	0.155	.177	0.173	.131	0.927	.000	1.000					
F: AM (10%)														
Patients	0.259	.016	0.200	.053	-0.036	.719	0.886	.000	0.932	.000	1.000			
Nonpatients	0.243	.032	0.085	.458	0.133	.243	0.856	.000	0.945	.000	1.000			
G: Progressed disease														
Patients	0.310	.002	0.570	.000	0.260	.011	0.364	.000	0.389	.000	0.335	.001	1.000	
Nonpatients	0.479	.000	0.536	.000	0.494	.000	0.329	.003	0.219	.055	0.202	.076	1.000	

Abbreviations: AM: active monitoring, BCS: breast conserving surgery, P: P value, Radio: radiotherapy, Recon: reconstruction |

Standard gamble:

	A BCS alone		B BCS & radio		C Mastectomy		D AM (40%)		E AM (20%)		F AM (10%)		G Progressed	
Health state	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**
A: BCS alone														
Patients	1.000													
Nonpatients	1.000													
B: BCS & radiotherapy														
Patients	0.928	.000	1.000											
Nonpatients	0.874	.000	1.000											
C: Mastectomy +/- recon.														
Patients	0.777	.000	0.817	.000	1.000									
Nonpatients	0.721	.000	0.794	.000	1.000									
D: AM (40%)														
Patients	0.475	.000	0.477	.000	0.551	.000	1.000							
Nonpatients	0.341	.002	0.289	.010	0.225	.048	1.000							
E: AM (20%)														
Patients	0.472	.000	0.470	.000	0.554	.000	0.957	.000	1.000					
Nonpatients	0.307	.006	0.194	.088	0.175	.126	0.911	.000	1.000					
F: AM (10%)														
Patients	0.429	.000	0.433	.000	0.499	.000	0.927	.000	0.981	.000	1.000			
Nonpatients	0.273	.016	0.164	.153	0.178	.121	0.811	.000	0.923	.000	1.000			
G: Progressed disease														
Patients	0.670	.000	0.680	.000	0.722	.000	0.688	.000	0.693	.000	0.656	.000	1.000	
Nonpatients	0.506	.000	0.492	.000	0.467	.000	0.354	.002	0.334	.003	0.299	.008	1.000	

Abbreviations: AM: active monitoring, BCS: breast conserving surgery, P: P value, Radio: radiotherapy, Recon: reconstruction

EQ-5D-5L:

	A BCS alone		B BCS & radio		C Mastectomy		D AM (40%)		E AM (20%)		F AM (10%)		G Progressed	
Health state	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**	**Correlation**	**P**
A: BCS alone														
Patients	1.000													
Nonpatients	1.000													
B: BCS & radiotherapy														
Patients	0.361	.000	1.000											
Nonpatients	0.063	.584	1.000											
C: Mastectomy +/- recon.														
Patients	0.246	.017	0.564	.000	1.000									
Nonpatients	-0.020	.861	0.712	.000	1.000									
D: AM (40%)														
Patients	0.140	.179	0.468	.000	0.338	.001	1.000							
Nonpatients	0.318	.005	0.312	.005	0.213	.061	1.000							
E: AM (20%)														
Patients	0.196	.058	0.435	.000	0.347	.001	0.885	.000	1.000					
Nonpatients	0.199	.081	0.280	.013	0.318	.005	0.893	.000	1.000					
F: AM (10%)														
Patients	0.232	.025	0.433	.000	0.299	.003	0.772	.000	0.860	.000	1.000			
Nonpatients	0.209	.067	0.217	.056	0.312	.005	0.763	.000	0.885	.000	1.000			
G: Progressed disease														
Patients	0.305	.003	0.570	.000	0.703	.000	0.449	.000	0.487	.000	0.437	.000	1.000	
Nonpatients	0.146	.204	0.594	.000	0.449	.000	0.322	.004	0.346	.001	0.424	.000	1.000	

Abbreviations: AM: active monitoring, BCS: breast conserving surgery, P: P value, Radio: radiotherapy, Recon: reconstruction

APPENDIX 9: Search strategy for the systematic review of economic evaluations for DCIS treatment

MEDLINE

1. Carcinoma, Intraductal, Noninfiltrating/
2. DCIS.ti,ab.
3. ductal carcinoma in situ.ti,ab.
4. (breast cancer adj3 situ).tw.
5. carcinoma in situ.tw. and breast neoplasm/
6. (intraductal carcinoma$ and breast).mp.
7. 1 or 2 or 3 or 4 or 5 or 6
8. economics/
9. value of life/
10. exp "costs and cost analysis"/
11. exp economics, hospital/
12. exp economics, medical/
13. economics, nursing/
14. economics, pharmaceutical/
15. exp "fees and charges"/
16. exp budgets/
17. budget*.ti,ab.
18. cost*.ti.ab
19. (economic* or pharmaco?economic*).ti,ab.
20. (price* or pricing*).ti,ab.
21. (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ti,ab.
22. (financ* or fee or fees).ti,ab.
23. (value adj2 (money or monetary)).ti,ab.
24. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23
25. 7 and 24
26. limit 25 to (english language and humans)

EMBASE

1. DCIS.ti,ab.
2. ductal carcinoma in situ.ti,ab.
3. (breast cancer adj3 situ).tw.
4. carcinoma in situ.tw. and breast neoplasm/

5. (intraductal carcinoma$ and breast).mp.

6. intraductal carcinoma/

7. 1 or 2 or 3 or 4 or 5 or 6

8. health economics/

9. exp economic evaluation/

10. exp health care cost/

11. exp fee/

12. budget/

13. funding/

14. budget*.ti,ab.

15. cost*.ti.

16. (economic* or pharmaco?economic*).ti.

17. (price* or pricing*).ti,ab.

18. (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ti,ab.

19. (financ* or fee or fees).ti,ab.

20. (value adj2 (money or monetary)).ti,ab.

21. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20

22. 7 and 21

23. limit 22 to (abstracts and english language)

Psycinfo

1. DCIS.ti,ab.

2. ductal carcinoma in situ.ti,ab.

3. (breast cancer adj3 situ).tw.

4. carcinoma in situ.tw. and breast neoplasm/

5. (intraductal carcinoma$ and breast).mp.

6. breast neoplasms/ and in situ.tw.

7. 1 or 2 or 3 or 4 or 5 or 6

8. health economics/

9. exp economic evaluation/

10. exp health care cost/

11. exp fee/

12. budget/

13. funding/

14. budget*.ti,ab.

15. cost*.ti.

16. (economic* or pharmaco?economic*).ti.

17. (price* or pricing*).ti,ab.

18. (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ti,ab.

19. (financ* or fee or fees).ti,ab.

20. (value adj2 (money or monetary)).ti,ab.

21. 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20

22. 7 and 21

23. limit 22 to (abstracts and english language)

EconLit

S13 S11 AND S12

S12 S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10

S11 S1 OR S2 OR S3

S10 AB cost per QALY

S9 AB cost benefit analysis or cost consequence analysis or cost consequence analysis

S8 AB economic model

S7 cost effectiveness analysis

S6 cost utility analysis

S5 cost effectiveness

S4 economic evaluation

S3 breast cancer AND in situ

S2 ductal carcinoma in situ

S1 DCIS

CINAHL

S13 S11 AND S12

S12 S5 OR S6 OR S7 OR S8 OR S9 OR S10

S11 S1 OR S2 OR S3 OR S4

S10 AB economic model

S9 TI economic model

S8 AB cost per QALY

S7 AB cost utility analysis or cost effectiveness analysis or cost minimisation analysis or cost consequence analysis or cost benefit analysis

S6 AB economic evaluation

S5 (MH "Economics+")

S4 "breast cancer AND in situ"

S3 AB ductal carcinoma in situ

S2 "DCIS"

S1 (MH "Carcinoma, Ductal, Breast")

APPENDIX 10: Quality appraisal checklists for the economic evaluations in the systematic review

	Raldow (2016)	Suh (2005)	Trentham-Dietz (2018)
Study design	✓	✓	✓
Research question is stated	✓	✓	✓
Economic relevance of research question is stated	✓	✓	✓
The viewpoint(s) of the analysis is clearly stated and justified	✓	✓	X
Rationale for the choice of the alternatives compared is reported	✓	✓	✓
Alternatives being compared are clearly described	✓	✓	✓
Type of economic evaluation is stated	NC	✓	X
Choice of type of economic evaluation is justified	X	✓	X
Data collection	✓	X	✓
Source(s) of effectiveness estimates is stated	✓	✓	✓
Details of the design and results of the effectiveness study are given	✓	✓	✓
Details of the methods of synthesis or meta-analysis of estimates given	✓	✓	✓
Primary outcome measure(s) for the economic evaluation is clearly stated	✓	✓	✓
Methods used to value health states and other benefits are stated	✓	✓	✓
Details of the subjects from whom valuations were obtained are given	✓	✓	✓
Productivity changes (if included) are reported separately	NA	✓	NA
Relevance of productivity changes to the study question is discussed	NA	✓	NA
Quantities of resources are reported separately from their unit cost	X	X	X
Methods for the estimation of quantities and unit costs are described	✓	✓	✓
Currency and price data are recorded	✓	✓	✓
Were details of price adjustments for inflation or currency conversion given	X	✓	X
Details of any model used are given	✓	✓	✓
Justification for the choice of model used and the key parameters are stated	✓	✓	✓
Analysis and interpretation of results	✓	✓	✓
Time horizon of cost and benefits is stated	✓	✓	✓
The discount rate is stated	✓	✓	X
The choice of discount rate is justified	NA	X	✓
An explanation is given if cost or benefits were not discounted	NA	NA	NA
The details of statistical test(s) and confidence intervals are given for stochastic data	X	X	X
The approach to sensitivity analysis is described	✓	✓	✓
Choice of variables for sensitivity analysis is justified	✓	✓	✓
Ranges over which the parameters were varied are stated	✓	✓	✓
Relevant alternatives are compared	✓	✓	✓
Incremental analysis is reported	✓	✓	X
Major outcomes are presented disaggregated as well as aggregated	✓	✓	✓
The answer to the study question is given	✓	✓	✓
Conclusions follow from the data reported	✓	✓	✓
Conclusions are accompanied by the appropriate caveats	✓	✓	✓

NA - Not applicable; NC – Not clear

APPENDIX 11: CHEERS and Philips checklists for the reporting of economic evaluations

CHEERS checklist:

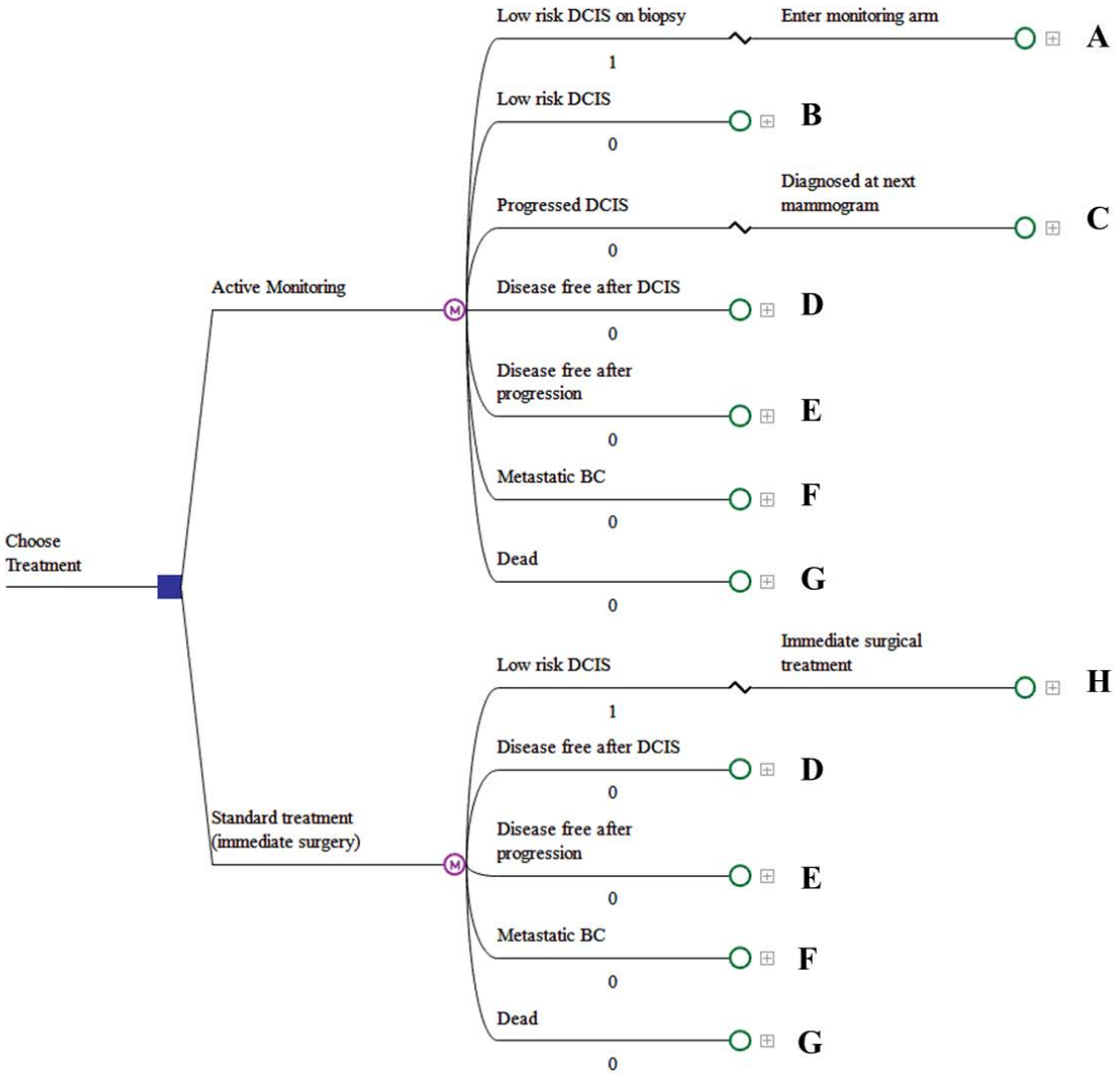
Section/item	Item	Recommendation	Reported
Title and abstract			
Title	1	Identify the study as an economic evaluation or use more specific terms such as "cost-effectiveness analysis", and describe the interventions compared.	✓ (chapter title)
Abstract	2	Provide a structured summary of objectives, perspective, setting, methods (including study design and inputs), results (including base case and uncertainty analyses), and conclusions.	NA (thesis chapter)
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.	✓
Target population and subgroups	4	Describe characteristics of the base case population and subgroups analysed, including why they were chosen.	✓
Setting and location	5	State relevant aspects of the system(s) in which the decision(s) need(s) to be made.	✓
Study perspective	6	Describe the perspective of the study and relate this to the costs being evaluated.	✓
Comparators	7	Describe the interventions or strategies being compared and state why they were chosen.	✓
Time horizon	8	State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate.	✓
Discount rate	9	Report the choice of discount rate(s) used for costs and outcomes and say why appropriate.	✓
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.	✓
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.	✓
Measurement and valuation of preference-based outcomes	12	If applicable, describe the population and methods used to elicit preferences for outcomes.	✓ (previous chapters)
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.	✓
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	✓

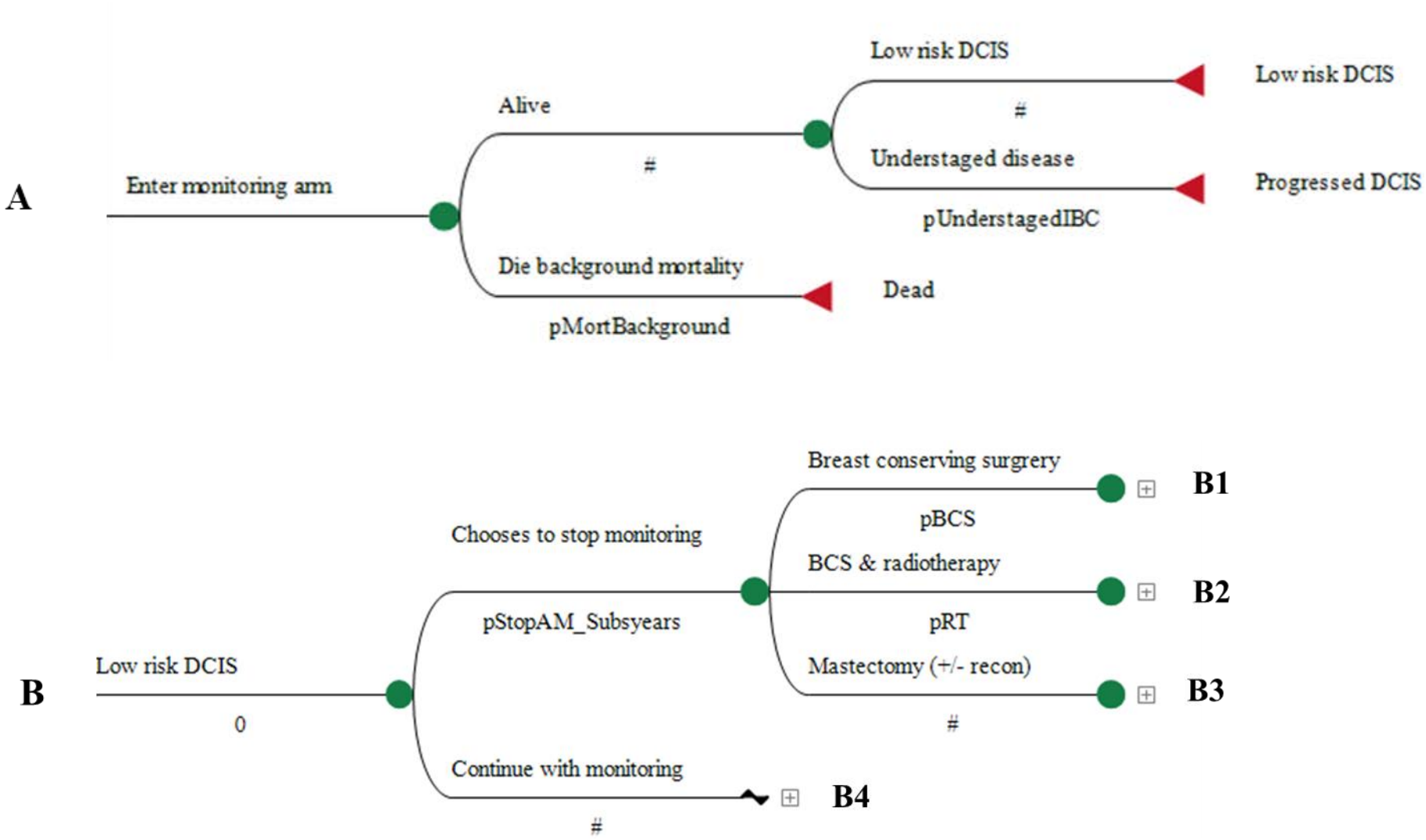
Section/item	Item	Recommendation	Reported
		rate.	
Choice of model	15	Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended.	✓
Assumptions	16	Describe all structural or other assumptions underpinning the decision-analytical model.	✓
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 adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty.	✓
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.	✓
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.	✓
Characterising 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.	✓
Characterising 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.	✓
Study findings, limitations, generalisability, 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.	✓
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.	NA
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.	NA

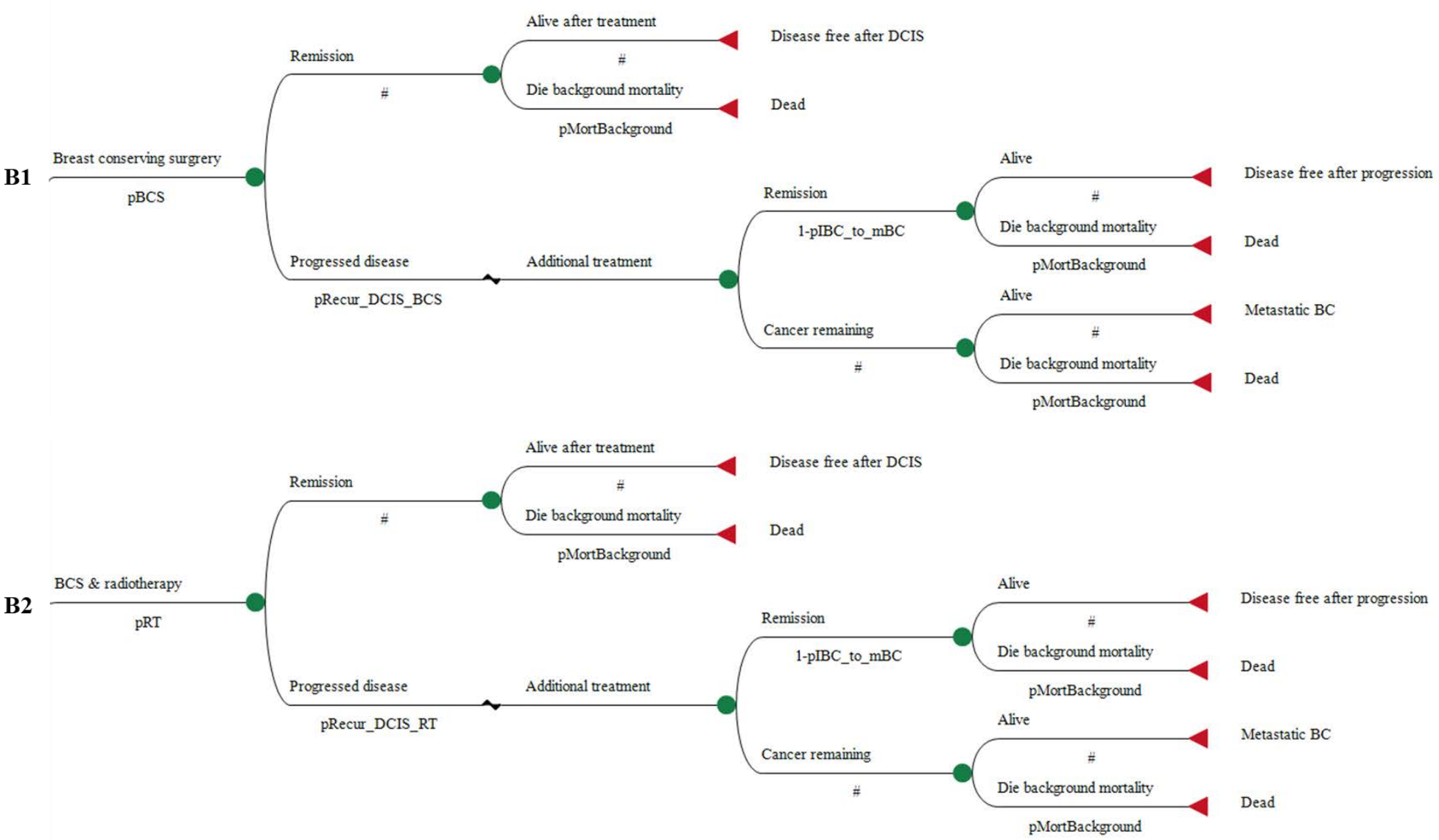
Philips checklist:

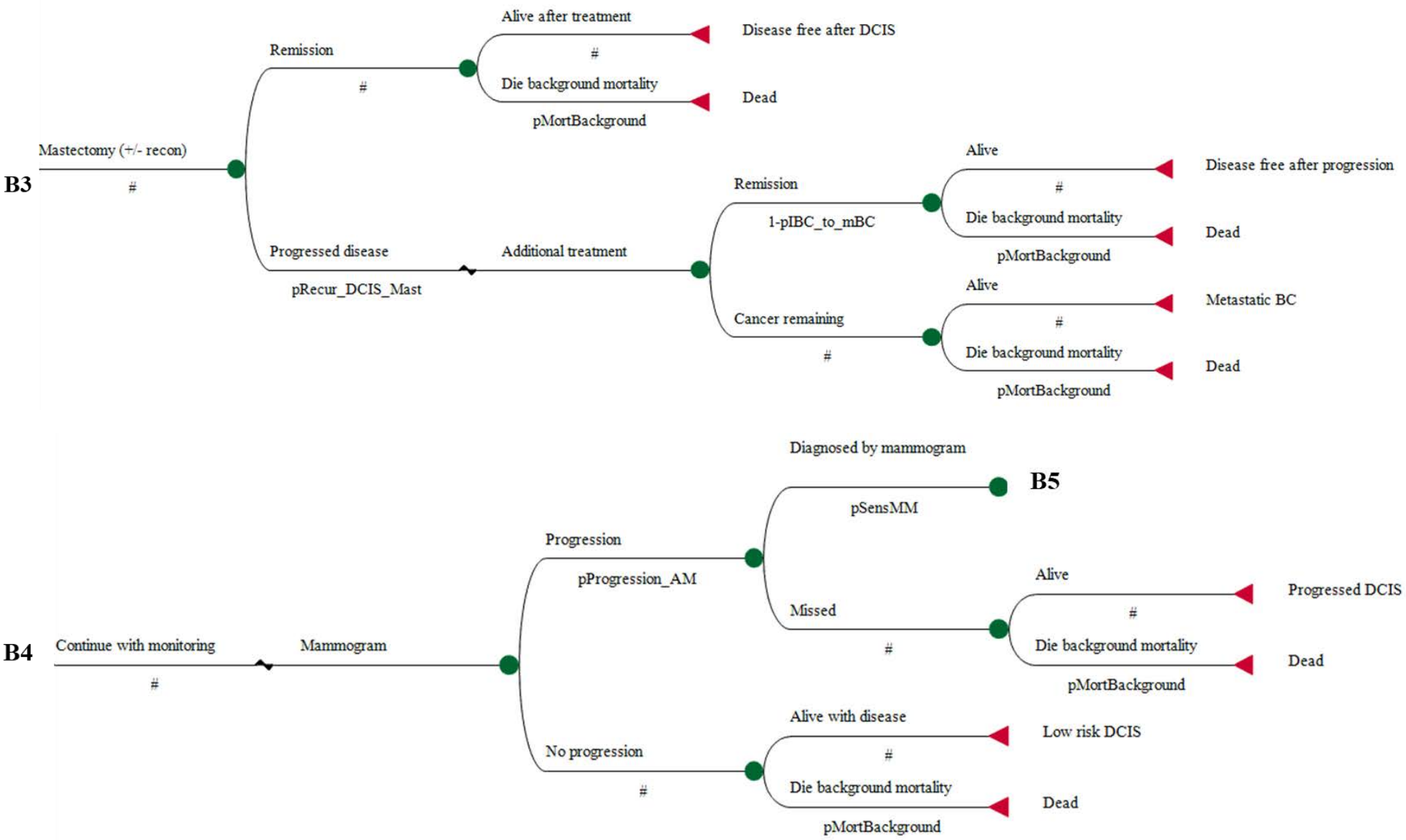
Criteria	Question	Response	Comments
Structure			
S1	Statement of decision problem/objective	✓	Aims and purpose clearly stated
S2	Statement of scope/perspective	✓	Background provided and perspective justified (NHS system in UK)
S3	Rational for structure	✓	Based on literature review
S4	Structural assumptions	✓	Informed by review and LORIS protocol
S5	Strategies/comparators	✓	AM vs. standard treatment
S6	Model type	✓	Markov (individual microsimulation)
S7	Time horizon	✓	Lifelong justified. Limitations discussed
S8	Disease states/pathways	✓	Explain in detail in the methods section
S9	Cycle length	✓	1 year (justified)
Data			
D1	Data identification	✓	References given
D2	Pre-model data analysis	NA	No trial data available yet from LORIS
D2a	Baseline data	✓	Discussed in detail – costs, probabilities and utilities given
D2b	Treatment effects	✓	Table 1
D2c	Costs	✓	Table 2
D2d	Quality of life weights (utilities)	✓	Table 3
D3	Data incorporation	✓	All data used in the model is outlined I the methods, including how this was incorporated at each stage of the model
D4	Assessment of uncertainty	✓	One-way, deterministic and PSA
D4a	Methodological	✓	Input parameters changed to check logic. Trackers used to check progress of women moving in the model.
D4b	Structural	✓	Sensitivity analysis run. Structure checked by clinicians.
D4c	Heterogeneity	✓	Model run for sub-group (by age)
D4d	Parameter	✓	Extensive SA performed. Limitations (lack of DCIS transition probability data is explicitly declared)
Consistency			
C1	Internal consistency	✓	Extreme values inputted to check logic of simulation pathways
C2	External consistency	X	No trial data to compare against. Assumptions comparable to other DCIS models but limited application

APPENDIX 12: Full TreeAge model of the economic model of the LORIS trial

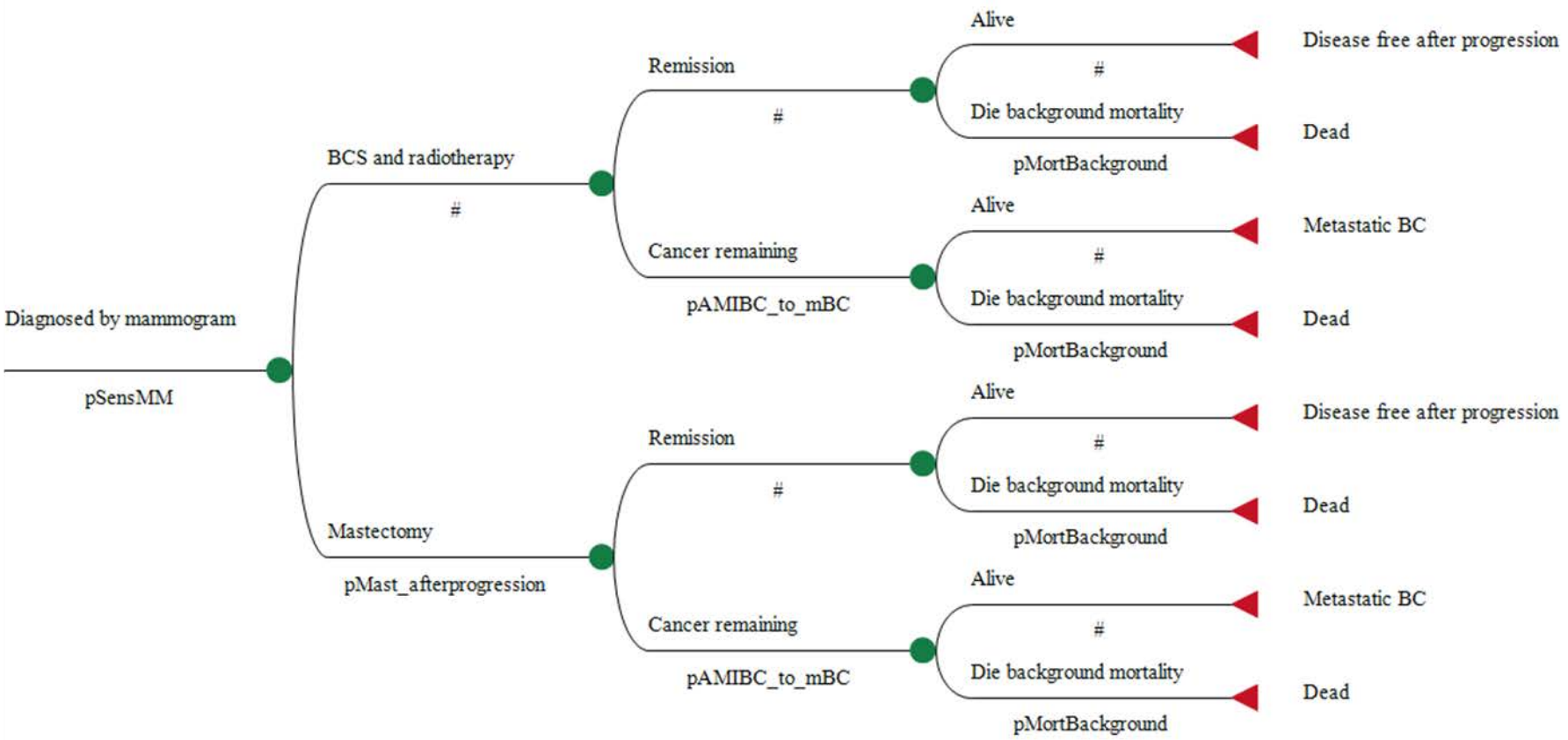


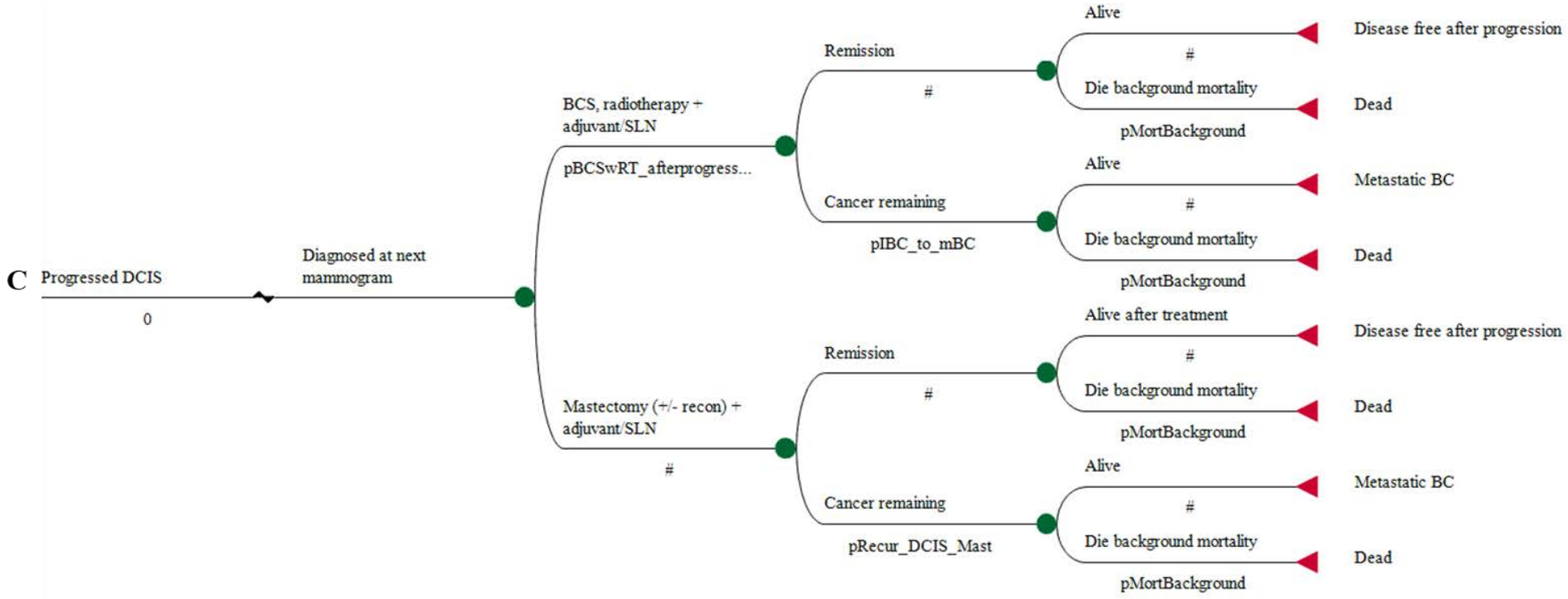

A
Enter monitoring arm
Alive
#
Die background mortality
pMortBackground
Dead
Low risk DCIS
#
Understaged disease
pUnderstagedIBC
Low risk DCIS
Progressed DCIS
B
Low risk DCIS
0
Chooses to stop monitoring
pStopAM_Subsyears
Continue with monitoring
#
B4
Breast conserving surgrery
pBCS
B1
BCS & radiotherapy
pRT
B2
Mastectomy (+/- recon)
#
B3


B1
Breast conserving surgery
pBCS
Remission
#
Alive after treatment
#
Disease free after DCIS
Die background mortality
pMortBackground
Dead
Progressed disease
pRecur_DCIS_BCS
Additional treatment
Remission
1-pIBC_to_mBC
Alive
#
Disease free after progression
Die background mortality
pMortBackground
Dead
Cancer remaining
#
Alive
#
Metastatic BC
Die background mortality
pMortBackground
Dead
B2
BCS & radiotherapy
pRT
Remission
#
Alive after treatment
#
Disease free after DCIS
Die background mortality
pMortBackground
Dead
Progressed disease
pRecur_DCIS_RT
Additional treatment
Remission
1-pIBC_to_mBC
Alive
#
Disease free after progression
Die background mortality
pMortBackground
Dead
Cancer remaining
#
Alive
#
Metastatic BC
Die background mortality
pMortBackground
Dead

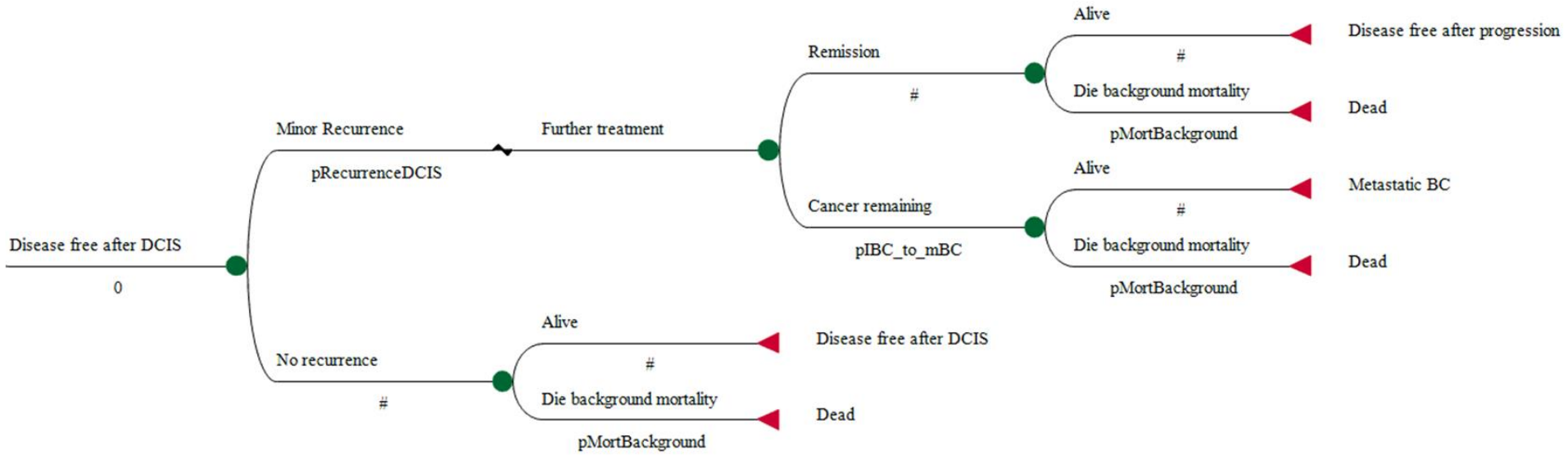

B3
Mastectomy (+/- recon)
#
Remission
#
Alive after treatment
#
Disease free after DCIS
Die background mortality
pMortBackground
Dead
Progressed disease
pRecur_DCIS_Mast
Additional treatment
Remission
1-pIBC_to_mBC
Alive
#
Disease free after progression
Die background mortality
pMortBackground
Dead
Cancer remaining
#
Alive
#
Metastatic BC
Die background mortality
pMortBackground
Dead
B4
Continue with monitoring
#
Mammogram
Progression
pProgression_AM
Diagnosed by mammogram
pSensMM
B5
Missed
#
Alive
#
Progressed DCIS
Die background mortality
pMortBackground
Dead
No progression
#
Alive with disease
#
Low risk DCIS
Die background mortality
pMortBackground
Dead

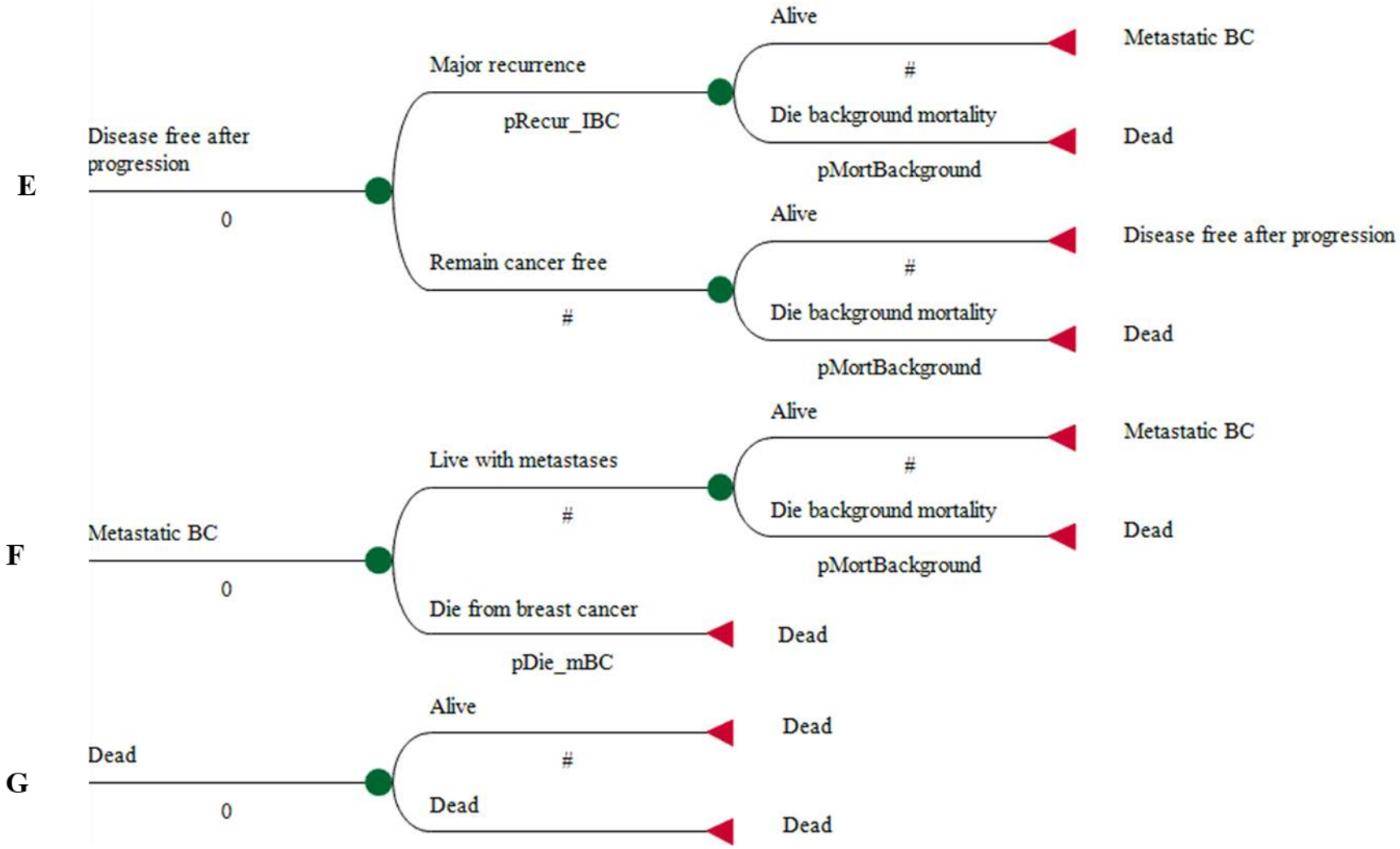
B5

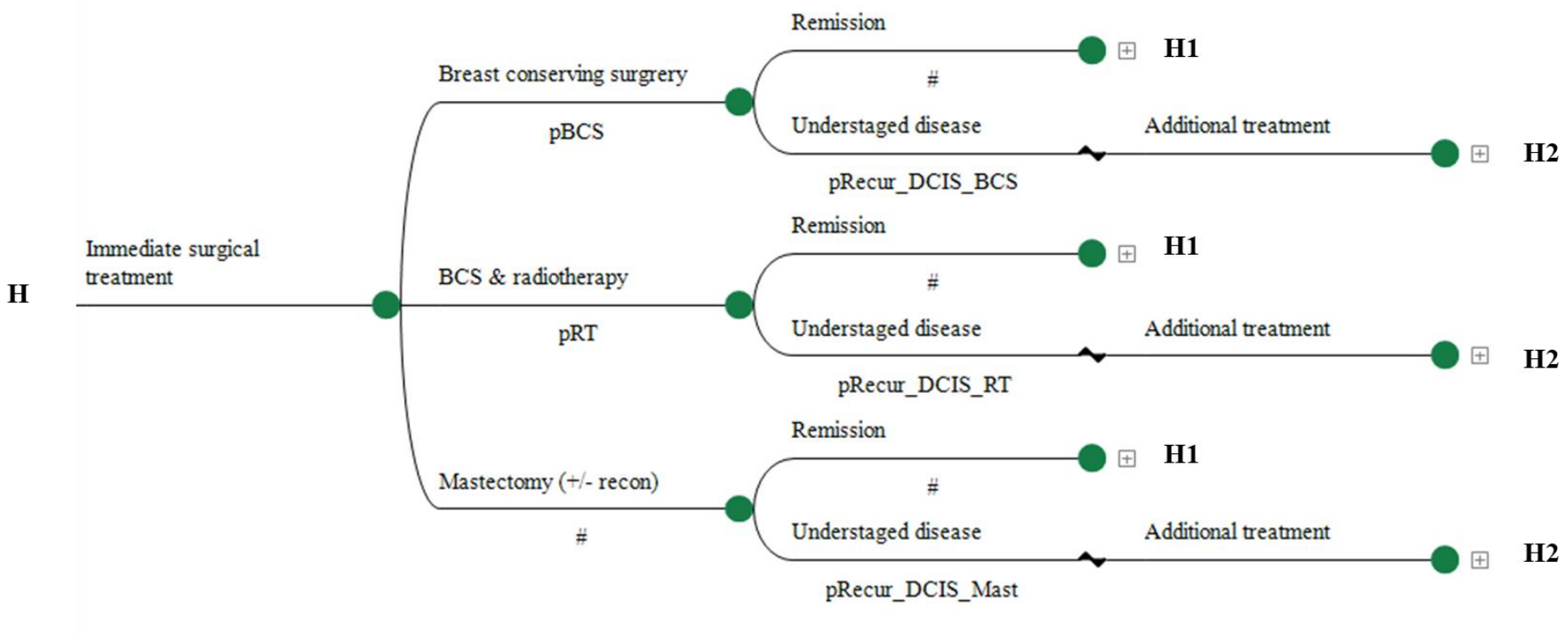

Diagnosed by mammogram
pSensMM
BCS and radiotherapy
#
Mastectomy
pMast_afterprogression
Remission
#
Cancer remaining
pAMIBC_to_mBC
Remission
#
Cancer remaining
pAMIBC_to_mBC
Alive
#
Die background mortality
pMortBackground
Alive
#
Die background mortality
pMortBackground
Alive
#
Die background mortality
pMortBackground
Alive
#
Die background mortality
pMortBackground
Disease free after progression
Dead
Metastatic BC
Dead
Disease free after progression
Dead
Metastatic BC
Dead

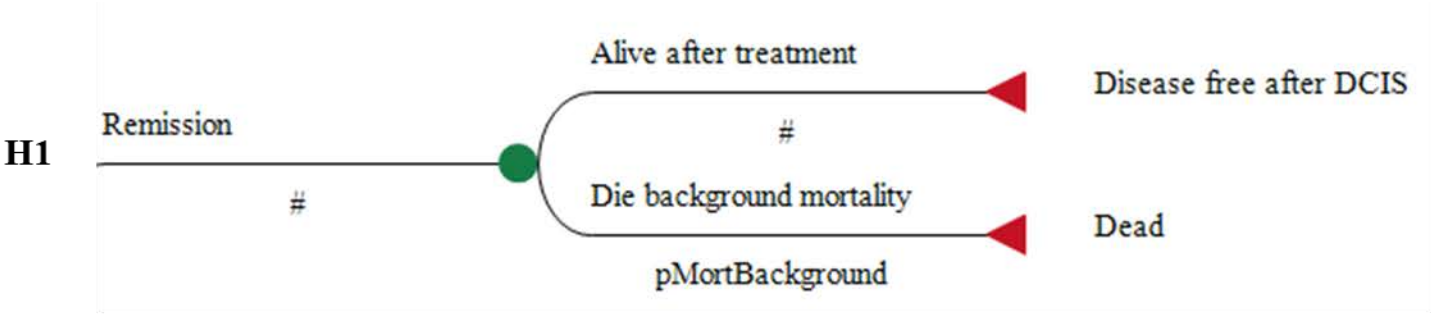

C
Progressed DCIS
0
Diagnosed at next mammogram
BCS, radiotherapy + adjuvant/SLN
pBCSwRT_afterprogress...
Mastectomy (+/- recon) + adjuvant/SLN
#
Remission
#
Cancer remaining
pIBC_to_mBC
Remission
#
Cancer remaining
pRecur_DCIS_Mast
Alive
#
Die background mortality
pMortBackground
Disease free after progression
Dead
Alive
#
Die background mortality
pMortBackground
Metastatic BC
Dead
Alive after treatment
#
Die background mortality
pMortBackground
Disease free after progression
Dead
Alive
#
Die background mortality
pMortBackground
Metastatic BC
Dead

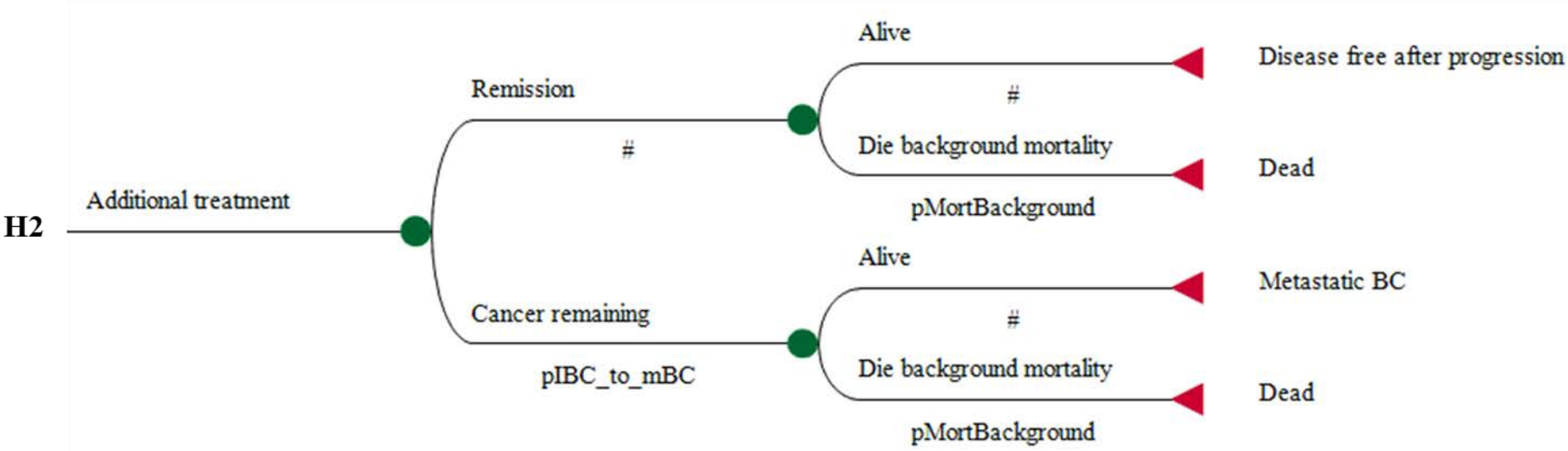
D


Disease free after DCIS
0
Minor Recurrence
pRecurrenceDCIS
Further treatment
Remission
#
Alive
#
Disease free after progression
Die background mortality
pMortBackground
Dead
Cancer remaining
pIBC_to_mBC
Alive
#
Metastatic BC
Die background mortality
pMortBackground
Dead
No recurrence
#
Alive
#
Disease free after DCIS
Die background mortality
pMortBackground
Dead


E
Disease free after progression
0
Major recurrence
pRecur_IBC
Remain cancer free
#
Alive
#
Die background mortality
pMortBackground
Metastatic BC
Dead
Alive
#
Die background mortality
pMortBackground
Disease free after progression
Dead
F
Metastatic BC
0
Live with metastases
#
Die from breast cancer
pDie_mBC
Dead
Alive
#
Die background mortality
pMortBackground
Metastatic BC
Dead
G
Dead
0
Alive
#
Dead
Dead
Dead


H
Immediate surgical treatment
Breast conserving surgrery
pBCS
BCS & radiotherapy
pRT
Mastectomy (+/- recon)
#
Remission
#
Understaged disease
pRecur_DCIS_BCS
Additional treatment
H1
H2
Remission
#
Understaged disease
pRecur_DCIS_RT
Additional treatment
H1
H2
Remission
#
Understaged disease
pRecur_DCIS_Mast
Additional treatment
H1
H2


H1
Remission
#
Alive after treatment
#
Disease free after DCIS
Die background mortality
pMortBackground
Dead


H2
Additional treatment
Remission
#
Alive
#
Disease free after progression
Die background mortality
pMortBackground
Dead
Cancer remaining
pIBC_to_mBC
Alive
#
Metastatic BC
Die background mortality
pMortBackground
Dead

APPENDIX 13: Topic guide from the qualitative interviews

1. **Introductions, overview & rules**

- Welcome, introduction and thank you for giving up their time.
- Information about me, purpose of the study, confidentiality, ground rules etc.
    - Project is looking value the benefits and harms associated with treating low risk pre-cancer (DCIS) so women can make better informed decisions/choices about treatment
    - Reiterate that breast screening saves lives/involves less invasive treatment (benefits)
    - But also identifies low risk conditions that would never cause harm (overdiagnosis)
    - Looking to measure women's preferences for managing low risk conditions identified at screening (DCIS) to try and reduce the harms of overdiagnosis
- Emphasize confidentiality and right to withdraw. (not to share personal experiences)

2. **Topics**

What do women know about the benefits and harms of breast screening? Open ended question (What do you know about…?)

i. Ductal carcinoma in situ (DCIS)
ii. Treatment or choices for low risk disease (surgery, radiotherapy, 'active monitoring)
iii. Tell me about BCS vs MT vs RT vs. AM?
iv. What factors do you think are important when women with low risk disease are choosing treatment? How do you think they choose?
v. Overdiagnosis or overtreatment
vi. What are your thoughts about active monitoring?
    i. How would you feel about it if it was shown to be safe and effective? (ref prostate cancer)
vii. How do you think women would feel if they are told they might not need the surgery? How do you think that would affect their decision?
    i. Would there be any extra impact on their QoL do you think?

Prompts:
Ductal carcinoma in situ (DCIS) What is it? How is it picked up? How is it treated? What choices do women have? "Pre-cancer" terminology. Spectrum of disease. Good knowledge given your experience, but what do you think about the lay person in the general population?

Treatment choices: What choices do you think women with low risk pre-cancer have? (or are they given a choice?) How do you think they might make that choice?

What factors do you think are important for women when choosing between the different types of treatment? Who would choose BCS/MT? Who would choose AM? Why would you choose x?

Active monitoring: Have you heard of active monitoring? Is this something you have been informed about or you think other women are aware of? If research shows AM is safe a safe and effective option, how do you think you would manage the abnormal cells? Do you think women would find this acceptable? Why? What about the other treatments? If this is something that is shown to be safe and effective, how do you think you would prefer to manage DCIS. Why would you choose x?

Overdiagnosis: have they ever heard of over-detection or over-treatment? What are your thoughts/feelings? How does it make you feel about attending screening or choosing treatment? Any thoughts after hearing the information about overdiagnosis/treatments for low risk breast disease? How does that make you feel? Do you think they would feel worse about having surgery or regret if they have?

How would you feel about active monitoring now you have heard more information? Why yes/no over surgery? What factors are important when women are deciding whether to adopt surgery or active monitoring?

3. Go through the different vignettes

Go through each question. Prompts:

- What are you thinking?
- Why did you choose that option?
- What factor swayed your decision?
- What do you think matters to women (did you find yourself concentrating on one characteristic more than others or did you weight things up evenly?)
- Are you happy with your choice? Would you change your choice?

THEME	**KEY QUESTIONS (for each method)**	**ADDITIONAL PROMPTS**
Information in health state description	What do you think about the amount of information provided?	Risk information
Attributes and health states described	What factors were more important to you when deciding on a choice?	Why? Anything missing? Risk adversity
Acceptability of active monitoring (AM)	Feasible choice? Why (not)?	What would make women choose surgery over AM? Why? What would make AM acceptable?
Risks	Is overdiagnosis and overtreatment something you were informed about when attending? Do you think women are aware of the benefits and harms of screening and treating low risk DCIS	What do you think they know? How can this be conveyed better? Other comments?

4. Opportunity for questions

Any further feedback or thoughts?

5. Thanking and closing

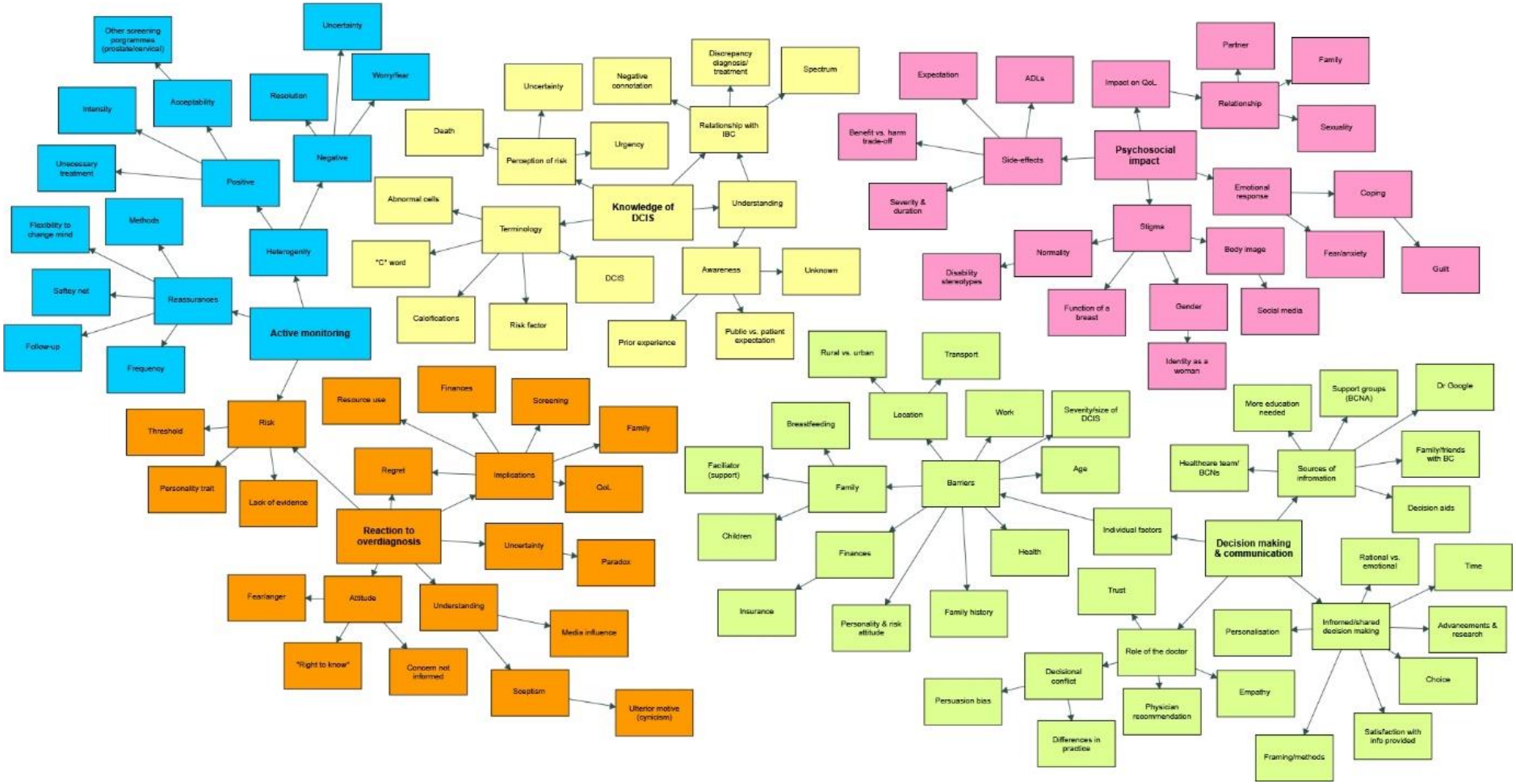
"We would like to thank you for taking part in this discussion. Your answers will be used to develop the descriptions and questions so they can be sent out to women all across Victoria to get an idea of how women value screening and treatment choices for DCIS."

6. Safety netting:

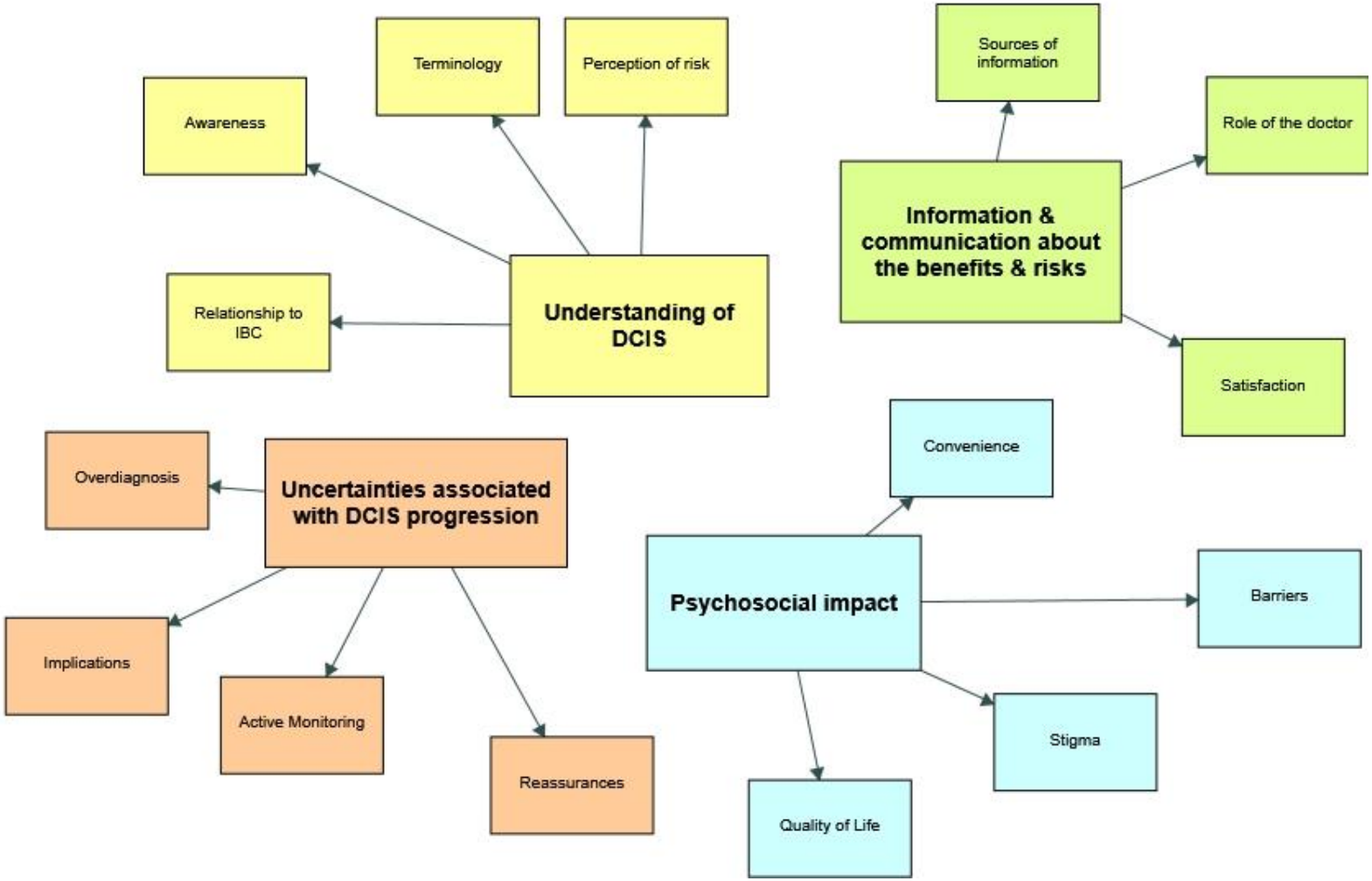
Signpost for where people can access help and further information on breast cancer screening and treatment – **this is included in the plain language summary**

APPENDIX 14: Framework analysis development

Initial coding structure (phase one)



Revised coding structure (phase two)



Word frequency diagram (phase three)

APPENDIX 15: Ethical approval for the qualitative interviews

THE UNIVERSITY OF MELBOURNE

12 March 2018

Dr Carolyn Nickson
Melbourne School of Population and Global Health
The University of Melbourne

Dear Dr Nickson

Project title:	**Identifying and quantifying the benefits and harms associated with treating ductal carcinoma in situ (DCIS) for use in the economic evaluation of breast screening programmes**
Researchers:	**Prof D Rea, Prof T Roberts, Dr C A Nickson, Professor B Mann, A/Prof D Petrie, H Bromley**
Ethics ID:	**1750252**

I am pleased to advise that the amendment 1750252.2 dated 26 February 2018 to this Project has been approved by the Medicine and Dentistry Human Ethics Sub-Committee.

Please note it is your responsibility to ensure that all people associated with the Project are made aware of the amendment.

Yours sincerely

[redacted]

Ms Hilary Young
Secretary, Medicine and Dentistry HESC
Phone: 03 8344 8595, Email: hilary.young@unimelb.edu.au

RESEARCH, INNOVATION & COMMERCIALISATION
Incorporating UoM Commercial Ltd and the School of Melbourne Custom Programs
The University of Melbourne, Victoria 3010, Australia
T: +61 3 8344 1539 (external) **T:** 40777 (internal) **W:** unimelb.edu.au/research

unimelb.edu.au

Consent Form for the Main Study (Audio-recording)

A joint research project between the University of Melbourne (Australia) and University of Birmingham (UK) School of Medical, Dental and Health Sciences

Title: Valuing the benefits and risks associated with treating low risk breast conditions

Primary Researcher: Hannah Bromley (Doctoral research student)

[redacted]

Supervisory researchers:
Dr C Nickson, (University of Melbourne), Prof B Mann (Victorian Comprehensive Cancer Centre), Associate/Prof D Petrie (Monash University) Prof D Rea (University of Birmingham), Prof T Roberts, (University of Birmingham)
Email: cnickson@unimelb.edu.au

This consent form relates to the participation in a PhD student project valuing different types of treatment for low risk breast conditions. It involves a short interview (which may be audio-recorded) and a questionnaire to determine your preferences for the different treatments described.

Please read the form and if you are happy to proceed, fill in your details at the bottom of the form.

1. I consent to participate in this project, the details of which have been explained to me, and I have been provided with a written plain language statement to keep.

2. I acknowledge that the possible effects of participating in this research project have been explained to my satisfaction.

3. I understand that my participation is voluntary and that I am free to withdraw from this project anytime without explanation or prejudice and to withdraw any unprocessed data that I have provided.

4. I acknowledge that participation will involve an interview, which may be audio-recorded.

5. I understand that the information provided will be used for research only. My name will be referred to by a pseudonym in any publications arising from the research. I understand that due to the small sample size, the potential may exist for me to be identified from my comments.

6. I have been informed that the confidentiality of the information I provide will be safeguarded subject to any legal requirements and any data collected will be stored as per University of Melbourne guidelines.

7. I understand what is involved and agree to take part in the study.

Name of Participant: _____

Participant signature _____ **Date:** _____

APPENDIX 16: COREQ checklist for the reporting of qualitative research

i) Consolidated criteria for reporting qualitative studies (COREQ) checklist

No. Item	Guide questions/description	Reported
Domain 1: Research team and reflexivity		
Personal Characteristics		
1. Interviewer/facilitator	Which author/s conducted the interview or focus group?	PhD researcher
2. Credentials	What were the researcher's credentials? E.g. PhD, MD	Not stated
3. Occupation	What was their occupation at the time of the study?	Stated (PhD student/clinician)
4. Gender	Was the researcher male or female?	Female
5. Experience and training	What experience or training did the researcher have?	Qualitative module
Relationship with participants		
6. Relationship established	Was a relationship established prior to study commencement?	Prior relationship established (utilities)
7. Participant knowledge of the interviewer	What did the participants know about the researcher? e.g. personal goals, reasons for doing the research	Role and purpose
8. Interviewer characteristics	What characteristics were reported about the interviewer/facilitator? e.g. Bias, assumptions, reasons and interests in the research topic	Stated in reflection
Domain 2: study design		
Theoretical framework		
9. Methodological orientation and theory	What methodological orientation was stated to underpin the study? e.g. grounded theory, discourse analysis, ethnography, phenomenology, content analysis	Framework analysis
Participant selection		
10. Sampling	How were participants selected? e.g. purposive, convenience, consecutive, snowball	Purposive
11. Method of approach	How were participants approached? e.g. face-to-face, telephone, mail, email	Email
12. Sample size	How many participants were in the study?	26
13. Non-participation	How many people refused to participate or dropped out? Reasons?	Nil
Setting		
14. Setting of data collection	Where was the data collected? e.g. home, clinic, workplace	Community venues
15. Presence of non-participants	Was anyone else present besides the participants and researchers?	No
16. Description of sample	What are the important characteristics of the sample? e.g. demographic data, date	Demographics provided in Table 1
Data collection		
17. Interview guide	Were questions, prompts, guides provided by the authors? Was it pilot tested?	Topic guide provided in Appendices

18. Repeat interviews	Were repeat interviews carried out? If yes, how many?	No
19. Audio/visual recording	Did the research use audio or visual recording to collect the data?	Audio-recording
20. Field notes	Were field notes made during and/or after the interview or focus group?	No
21. Duration	What was the duration of the interviews or focus group?	Stated in results
22. Data saturation	Was data saturation discussed?	Yes
23. Transcripts returned	Were transcripts returned to participants for comment and/or correction?	No
Domain 3: analysis and findings		
Data analysis		
24. Number of data coders	How many data coders coded the data?	3
25. Description of the coding tree	Did authors provide a description of the coding tree?	Yes
26. Derivation of themes	Were themes identified in advance or derived from the data?	Mixed
27. Software	What software, if applicable, was used to manage the data?	NVivo
28. Participant checking	Did participants provide feedback on the findings?	Yes
Reporting		
29. Quotations presented	Were participant quotations presented to illustrate the themes/findings? Was each quotation identified? e.g. participant number	Yes
30. Data and findings consistent	Was there consistency between the data presented and the findings?	Relationship to existing empirical study given
31. Clarity of major themes	Were major themes clearly presented in the findings?	Yes
32. Clarity of minor themes	Is there a description of diverse cases or discussion of minor themes?	Yes

ii) Standards for Reporting Qualitative Research (SRQR) checklist

	Standards for Reporting Qualitative Research (SRQR)*	**Reported?**
Title and abstract		
	Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	✓
	Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	NA
Introduction		
	Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	✓
	Purpose or research question - Purpose of the study and specific objectives or questions	✓

Methods		
	Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., post positivist, constructivist/ interpretivist) is also recommended; rationale**	✓
	Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	✓
	Context - Setting/site and salient contextual factors; rationale**	✓
	Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	✓
	Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	✓
	Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	✓
	Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	✓
	Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	✓
	Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	✓
	Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	✓
	Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	✓
Results/findings		
	Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	✓
	Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	✓
Discussion		
	Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	✓
	Limitations - Trustworthiness and limitations of findings	✓
Other		
	Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	✓
	Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	NA

References

1. IARC Working Group on the Evaluation of Cancer-Preventive Interventions. Handbooks of Cancer Prevention: Breast cancer screening. Lyon, France: International Agency for Research on Cancer; 2014.
2. Australian Institute of Health and Welfare. Breast cancer in Australia: an overview. Canberra, Australia: AIHW; 2012.
3. Marmot MG, Altman D, Cameron D, Dewar J, Thompson S, Wilcox M. The benefits and harms of breast cancer screening: an independent review. British Journal of Cancer. 2013;108(11):2205.
4. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes: Oxford University Press; 2015.
5. Morris S, Devlin N, Parkin D. Economic analysis in health care: John Wiley & Sons; 2007.
6. Brazier J, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Oxford: Oxford University Press; 2017.
7. Gray AM, Clarke PM, Wolstenholme JL, Wordsworth S. Applied methods of cost-effectiveness analysis in healthcare. Oxford: Oxford University Press; 2011.
8. National Institute for Health and Clinical Excellence. Guide to the methods of technology appraisal. London, UK: NICE; 2013.
9. Brouwer WB, Culyer AJ, van Exel NJA, Rutten FF. Welfarism vs. extra-welfarism. Journal of health economics. 2008;27(2):325-38.
10. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. Social science & medicine. 2008;67(7):1190-8.
11. Coast J. Maximisation in extra-welfarism: A critique of the current position in health economics. Social Science & Medicine. 2009;69(5):786-92.
12. Whitehead SJ, Ali S. Health outcomes in economic evaluation: the QALY and utilities. British medical bulletin. 2010;96(1):5-21.
13. Lorgelly PK, Lawson KD, Fenwick EAL, Briggs AH. Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? Int J Environ Res Public Health. 2010;7(5):2274-89.
14. Tolley K. What are health utilities? London, UK: Hayward Medical Communications; 2009.

15. Brazier J, Akehurst R, Brennan A, Dolan P, Claxton K, McCabe C, et al. Should patients have a greater role in valuing health states? Applied health economics and health policy. 2005;4(4):201-8.
16. Torrance GW. Measurement of health state utilities for economic appraisal: a review. Journal of health economics. 1986;5(1):1-30.
17. Brazier J, Akehurst R, Brennan A, Dolan P, Claxton K, McCabe C, et al. Should patients have a greater role in valuing health states? Appl Health Econ Health Policy. 2005;4(4):201-8.
18. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: An EQ-5D-5L value set for England. Health economics. 2018;27(1):7-22.
19. Forrest AP. Breast cancer screening: Report to the Health Ministers of England, Wales, Scotland & Northern Ireland. London, UK: HMSO; 1986.
20. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. Cochrane database of systematic reviews. 2011;19(1):CD001877.
21. Cancer Research UK. Breast screening [internet]2017 [01 June 2019]. Available from: https://www.cancerresearchuk.org/about-cancer/breast-cancer/screening/breast-screening.
22. Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghate S, et al. Benefits and harms of breast cancer screening: a systematic review. Jama. 2015;314(15):1615-34.
23. Welch HG, Black WC. Overdiagnosis in cancer. J Natl Cancer Inst. 2010;102.
24. Puliti D, Duffy SW, Miccinesi G, De Koning H, Lynge E, Zappa M, et al. Overdiagnosis in mammographic screening for breast cancer in Europe: a literature review. Journal of medical screening. 2012;19(1_suppl):42-56.
25. Sagara Y, Mallory MA, Wong S, Aydogan F, DeSantis S, Barry WT, et al. Survival Benefit of Breast Surgery for Low-Grade Ductal Carcinoma In Situ: A Population-Based Cohort Study. JAMA surgery. 2015;150(8):739-45.
26. Pashayan N, Morris S, Gilbert FJ, Pharoah PP. Cost-effectiveness and benefit-to-harm ratio of risk-stratified screening for breast cancer: A life-table model. JAMA oncology. 2018.

27. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JM, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. European journal of cancer. 2015;51(16):2296-303.
28. Groen EJ, Elshof LE, Visser LL, Emiel JT, Winter-Warnars HA, Lips EH, et al. Finding the balance between over-and under-treatment of ductal carcinoma in situ (DCIS). The Breast. 2017;31:274-83.
29. Hwang ES, Hyslop T, Lynch T, Frank E, Pinto D, Basila D, et al. The COMET (Comparison of Operative versus Monitoring and Endocrine Therapy) trial: a phase III randomised controlled clinical trial for low-risk ductal carcinoma in situ (DCIS). BMJ Open. 2019;9(3):e026797.
30. Francis A, Fallowfield L, Rea D. The LORIS trial: addressing overtreatment of ductal carcinoma in situ. Clinical Oncology. 2015;27(1):6-8.
31. Raftery J, Chorozoglou M. Possible net harms of breast cancer screening: updated modelling of Forrest report. Bmj. 2011;343:d7627.
32. Fallowfield L ML, Francis A, Jenkins V, Rea D.,. Low grade Ductal Carcinoma in situ (DCIS): how best to describe it? The Breast,. 2014;23(5):693-6.
33. Nelson HD, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: systematic review and meta-analysis to update the 2009 US Preventive Services Task Force Recommendation. Annals of internal medicine. 2016;164(4):244-55.
34. Stewart B, Wild CP. World Cancer Report 2014. Lyon, France: World Health Organisation; 2014.
35. Cancer Research UK. Breast cancer statistics [internet]2019 [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer.
36. Breast Cancer Network Australia. Breast cancer statistics in Australia. Canberra, Australia: BCNA; 2018.
37. Australian Institute of Health and Welfare. Cancer Data in Australia. Canberra, Australia: AIHW; 2018.
38. Howlader N, Noone AM, Krapcho M, Miller D BK, Kosary CL, Yu M, et al. Cancer Statistics Review 1975-2008. Bethesda, USA: National Cancer Institute; 2014.

39. Autier P, Boniol M, LaVecchia C, Vatten L, Gavin A, Héry C, et al. Disparities in breast cancer mortality trends between 30 European countries: retrospective trend analysis of WHO mortality database. BMJ. 2010;341:c3620.
40. Narod SA, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. JAMA oncology. 2015;1(7):888-96.
41. Torre LA, Islami F, Siegel RL, Ward EM, Jemal A. Global Cancer in Women: Burden and Trends. Cancer Epidemiology Biomarkers & Prevention. 2017;26(4):444.
42. Jacklyn G, Morrell S, McGeechan K, Houssami N, Irwig L, Pathmanathan N, et al. Carcinoma in situ of the breast in New South Wales, Australia: Current status and trends over the last 40 year. Breast (Edinburgh, Scotland). 2018;37:170-8.
43. Wilson JMG, Jungner G. Principles and practice of screening for disease. World Health Organization; 1968.
44. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. New England Journal of Medicine. 2012;367(21):1998-2005.
45. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA internal medicine. 2014;174(3):448-54.
46. IARC Working Group on the Evaluation of Cancer-Preventive Interventions. Handbooks of Cancer Prevention: Breast cancer screening. . Lyon, France: International Agency for Research on Cancer; 2002.
47. Gilbert FJ. Stratification of screening. Breast cancer research and treatment. 2018;167 (1):315.
48. Phi X-A, Tagliafico A, Houssami N, Greuter MJW, de Bock GH. Digital breast tomosynthesis for breast cancer screening and diagnosis in women with dense breasts - a systematic review and meta-analysis. BMC cancer. 2018;18(1):380-.
49. NHS England. Diagnosing cancer earlier and faster [internet]2019 [cited 2019 26 June]. Available from: https://www.england.nhs.uk/cancer/early-diagnosis/.
50. NHS Breast Cancer Screening Programme. How much does the programme cost? [internet]2017 [31 March 2019]. Available from: http://webarchive.nationalarchives.gov.uk/20150506150512/http://www.cancerscreening.nhs.uk/breastscreen/cost.html
51. Institute of Health and Welfare & Cancer Australia. Breast cancer in Australia: an overview,. Canberra: AIHW; 2012.

52. Pharoah PD, Sewell B, Fitzsimmons D, Bennett HS, Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. Bmj. 2013;346:f2618.
53. Wolstenholme JL, Smith SJ, Whynes DK. The costs of treating breast cancer in the United Kingdom: implications for screening. International journal of technology assessment in health care. 1998;14(2):277-89.
54. Jorgensen KJ, Gotzsche PC. Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. BMJ. 2009;339:b2587.
55. BreastScreen Australia. Benefits, harms and costs of screening mammography in women 70 years and over: a systematic review. The Medical Journal of Australia. 2002;176(6):266-71.
56. Elder K, Nickson C, Pattanasri M, Cooke S, Machalek D, Rose A, et al. Treatment Intensity Differences After Early-Stage Breast Cancer (ESBC) Diagnosis Depending on Participation in a Screening Program. Annals of surgical oncology. 2018;25(9):2563-72.
57. Mandelblatt JS, Stout NK, Schechter CB, Van Den Broek JJ, Miglioretti DL, Krapcho M, et al. Collaborative modeling of the benefits and harms associated with different US breast cancer screening strategies. Annals of internal medicine. 2016;164(4):215-25.
58. Johnston K, Brown J, Gerard K, O'Hanlon M, Morton A. Valuing temporary and chronic health states associated with breast screening. Social Science & Medicine. 1998;47(2):213-22.
59. Barratt AL, Irwig LM, Salkeld GP, Glasziou PP, Houssami N. Benefits, harms and costs of screening mammography in women 70 years and over: a systematic review. Medical Journal of Australia. 2002;176(6):266-71.
60. Messecar DC. Mammography screening for older women with and without cognitive impairment. Journal of Gerontological Nursing. 2000;26(4):14-24.
61. Gram I, Lund E, Slenker S. Quality of life following a false positive mammogram. British Journal of Cancer. 1990;62(6):1018.
62. Autier P, Boniol M. Mammography screening: A major issue in medicine. European Journal of Cancer. 2018;90:34-62.

63. Vass C. Using discrete choice experiments to value benefits and risks in primary care, . Manchester, UK: University of Manchester; 2017.
64. Biesheuvel C, Barratt A, Howard K, Houssami N, Irwig L. Effects of study methods and biases on estimates of invasive breast cancer overdetection with mammography screening: a systematic review. Lancet Oncology. 2007;8(12):1129-38.
65. Hersch J, Jansen J, Barratt A, Irwig L, Houssami N, Howard K, et al. Women's views on overdiagnosis in breast cancer screening: a qualitative study. BMJ. 2013;346:f158.
66. Parker LM, Rychetnik L, Carter S. Framing overdiagnosis in breast screening: a qualitative study with Australian experts. BMC cancer. 2015;15:606.
67. Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. JAMA. 2009;302(15):1685-92.
68. Carter SM, Rogers W, Heath I, Degeling C, Doust J, Barratt A. The challenge of overdiagnosis begins with its definition. BMJ. 2015;350.
69. Dolan P, Torgerson DJ, Wolstenholme J. Costs of breast cancer treatment in the United Kingdom. The Breast. 1999;8(4):205-7.
70. Duffy SW, Tabar L, Olsen AH, Vitak B, Allgood PC, Chen THH. Absolute numbers of lives saved and overdiagnosis in breast cancer screening, from a randomized trial and from the Breast Screening Programme in England. J Med Screen. 2010;17.
71. Barnes NL, Ooi JL, Yarnold JR, Bundred NJ. Ductal carcinoma in situ of the breast. BMJ. 2012;344:e797.
72. Wallis M, Clements K, Kearins O, Ball G, Macartney J, Lawrence G. The effect of DCIS grade on rate, type and time to recurrence after 15 years of follow-up of screen-detected DCIS. British Journal of Cancer. 2012;106(10):1611.
73. Ong M-S, Mandl KD. National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at $4 billion a year. Health affairs. 2015;34(4):576-83.
74. National Institute for Health and Care Excellence. Early and locally advanced breast cancer: diagnosis and management (NICE guideline 101). London, UK. 2018.
75. Barrio AV, Van Zee KJ. Controversies in the Treatment of Ductal Carcinoma in Situ. Annual review of medicine. 2017;68:197-211.

76. Ernster VL, Ballard-Barbash R, Barlow WE, Zheng Y, Weaver DL, Cutter G, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. J Natl Cancer Inst. 2002;94(20):1546-54.
77. Cowell CF, Weigelt B, Sakr RA, Ng CKY, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. Molecular Oncology. 2013;7(5):859-69.
78. Casasent AK, Edgerton M, Navin NE. Genome evolution in ductal carcinoma in situ: invasion of the clones. The Journal of Pathology. 2017;241(2):208-18.
79. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. Cancer. 2005;103(12):2481-4.
80. Eusebi V, Feudale E, Foschini MP, Micheli A, Conti A, Riva C, et al., editors. Long-term follow-up of in situ carcinoma of the breast. In Seminars in diagnostic pathology; 1994 (Vol. 11, No 3, pp. 223-235).
81. Silverstein MJ, Waisman JR, Gamagami P, Gierson ED, Colburn WJ, Rosser RJ, et al. Intraductal carcinoma of the breast (208 cases): clinical factors influencing treatment choice. Cancer. 1990;66(1):102-8.
82. Narod SA, Sopik V. Is invasion a necessary step for metastases in breast cancer? Breast cancer research and treatment. 2018;169(1):9-23.
83. Bobrow LG, Happerfield LC, Gregory WM, Springall RD, Millis RR, editors. The classification of ductal carcinoma in situ and its association with biological markers. In Seminars in diagnostic pathology; 1994 (Vol. 11, No. 3, pp. 199-207).
84. Maxwell A, Evans A, Carpenter R, Dobson H, Kearins O, Clements K, et al. Follow-up for screen-detected ductal carcinoma in situ: Results of a survey of UK centres participating in the Sloane project. European Journal of Surgical Oncology 2009;35(10):1055-9.
85. Rakovitch E, Nofech-Mozes S, Hanna W, Baehner FL, Saskin R, Butler SM, et al. A population-based validation study of the DCIS Score predicting recurrence risk in individuals treated by breast-conserving surgery alone. Breast cancer research and treatment. 2015;152(2):389-98.
86. Lee RJ, Vallow LA, McLaughlin SA, Tzou KS, Hines SL, Peterson JL. Ductal carcinoma in situ of the breast. International journal of surgical oncology. 2012;2012.

87. Claus EB, Petruzella S, Matloff E, Carter D. Prevalence of BRCA1 and BRCA2 mutations in women diagnosed with ductal carcinoma in situ. JAMA. 2005;293(8):964-9.
88. Claus EB, Stowe M, Carter D. Breast carcinoma in situ: risk factors and screening patterns. Journal of the National Cancer Institute. 2001;93(23):1811-7.
89. Barclay J, Ernster V, Kerlikowske K, Grady D, Sickles EA. Comparison of risk factors for ductal carcinoma in situ and invasive breast cancer. Journal of the National Cancer Institute. 1997;89(1):76-82.
90. Gorringe KL, Fox SB. Ductal carcinoma in situ Biology, biomarkers, and diagnosis. Frontiers in Oncology. 2017;23(7):248.
91. Doke K, Butler S, Mitchell MP. Current Therapeutic Approaches to DCIS. Journal of mammary gland biology and neoplasia. 2018;23(4):279-91.
92. Bulliard J-L, Chiolero A. Screening and overdiagnosis: public health implications. Public Health Reviews. 2015;36(1):8.
93. Rea D, Francis A. Breast screening saves lives: now is the time to address overtreatment. Journal of comparative effectiveness research. 2014;3(2):111-3.
94. NHS Cancer Screening Programmes. Uncertainties in the management of screen-detected ductal carcinoma in situ. Sheffield, UK: NHSBSP; 2008.
95. Tian Y, Schofield PE, Gough K, Mann GB. Profile and predictors of long-term morbidity in breast cancer survivors. Annals of surgical oncology. 2013;20(11):3453-60.
96. Correa C, McGale P, Taylor C, Wang Y, Clarke M, Davies C, et al. Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. Journal of the National Cancer Institute Monographs. 2010;2010(41):162-77.
97. Goodwin A, Parker S, Ghersi D, Wilcken N. Post-operative radiotherapy for ductal carcinoma in situ of the breast–a systematic review of the randomised trials. The Breast. 2009;18(3):143-9.
98. Morris AD, Morris RD, Wilson JF, White J, Steinberg S, Okunieff P, et al. Breast-conserving therapy vs mastectomy in early-stage breast cancer: a meta-analysis of 10-year survival. Cancer Journal Sci Am. 1997;3(1):6-12.

99. Collins LC, Laronga C, Wong JS. Ductal carcinoma in situ: treatment and prognosis [internet]2018 [cited 2019 01 June]. Available from: https://www.uptodate.com/contents/ductal-carcinoma-in-situ-treatment-and-prognosis.
100. Staley H, McCallum I, Bruce J. Postoperative tamoxifen for ductal carcinoma in situ. The Cochrane database of systematic reviews. 2012;10:Cd007847.
101. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal carcinoma in situ of the breast: a systematic review of incidence, treatment, and outcomes. Journal of the National Cancer Institute. 2010;102(3):170-8.
102. Ozanne EM, Shieh Y, Barnes J, Bouzan C, Hwang ES, Esserman LJ. Characterizing the impact of 25 years of DCIS treatment. Breast cancer research and treatment. 2011;129(1):165-73.
103. Ryser MD, Worni M, Turner EL, Marks JR, Durrett R, Hwang ES. Outcomes of active surveillance for ductal carcinoma in situ: a computational risk analysis. Journal of the National Cancer Institute. 2015;108(5):djv372.
104. Solin LJ, Gray R, Hughes LL, Wood WC, Lowen MA, Badve SS, et al. Surgical excision without radiation for ductal carcinoma in situ of the breast: 12-year results from the ECOG-ACRIN E5194 study. Journal of Clinical Oncology. 2015;33(33):3938.
105. Hanna WM, Parra-Herran C, Lu F-I, Slodkowska E, Rakovitch E, Nofech-Mozes S. Ductal carcinoma in situ of the breast: an update for the pathologist in the era of individualized risk assessment and tailored therapies. Modern Pathology. 2019;32(7):896-915.
106. Lynch T, Frank ES, Collyar DE, Basila D, Pinto D, Partridge A, et al. Comparison of operative to monitoring and endocrine therapy for low-risk DCIS (COMET). Journal of Clinical Oncology. 2018;36(15):S599.
107. Wesseling J, Elshof LE, Tryfonidis K, Poncet C, Aalders K, Van Leeuwen-Stok E, et al. Update of the randomized, non-inferiority LORD trial testing safety of active surveillance for women with screen-detected low risk ductal carcinoma in situ. . Cancer Research Conference: San Antonio Breast Cancer Symposium. 2017;78(4 Supplement 1).
108. Grimm LJ, Shelley Hwang E. Active Surveillance for DCIS: The Importance of Selection Criteria and Monitoring. Annals of surgical oncology. 2016;23(13):4134-6.

109. Kanbayashi C, Thompson AM, Hwang E-SS, Partridge AH, Rea DW, Wesseling J, et al. The international collaboration of active surveillance trials for low-risk DCIS. Journal of Clinical Oncology. 2019;37(15):S603.
110. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. Lancet. 2007;370:890-907.
111. Vass CM, Rigby D, Payne K. Investigating the Heterogeneity in Women's Preferences for Breast Screening: Does the Communication of Risk Matter? Value in Health. 2018;21(2):219-28.
112. Park TS, Hwang ES. Current trends in the management of ductal carcinoma in situ. Oncology. 2016;30(9):823-31.
113. Black J, Hashimzade N, Myles G. A dictionary of economics. Oxford, UK: Oxford University Press 2012.
114. Broome J. Why Economics Needs Ethical Theory. In: Basu K, Kanbu R, editors. Arguments for a better world: essays in honor of Amartya Sen. Oxford, UK: Oxford University Press; 2009. p. 7-14.
115. Harberger AC. Three basic postulates for applied welfare economics: an interpretive essay. Journal of Economic literature. 1971;9(3):785-97.
116. McIntosh E, Clarke, P. M., Frew, E. J., & Louviere, J. J. Applied Methods of Cost Benefit Analysis in Health Care. . Oxford: Oxford University Press.; 2010.
117. Von Neumann J, Morgenstern O. Theory of games and economic behavior. 2nd ed. Princeton, NJ, USA: Princeton University Press; 1947.
118. Johnson EM. Physician-induced demand. In: Culyer AJ, editor. Encyclopedia of health economics. Amsterdam, Netherlands: Elselvier; 2014. p. 77-82.
119. McGuire A. Theoretical concepts in the economic evaluation of health care. In: Drummond M, McGuire, A, editor. Economic Evaluation in Health Care: Merging Theory with Practice. New York: Oxford University Press; 2001. p. 1-21.
120. Sen A. Social choice theory: a re-examination. Econometrica. 1977;1:53-89.
121. Hurley J. An overview of the normative economics of the health care sector. . In: Culyer A, Newhouse J, editors. Handbook of Health Economics. Amsterdam, Netherlands: Elsevier Science; 2000.
122. Mitchell RC, Carson RT. Using surveys to value public goods: the contingent valuation method. Washington, US: RFF Press; 2013.

123. Tsuchiya A, Williams A. Welfare economics and economic evaluation. Economic evaluation in health care: merging theory with practice. 2001:27-8.
124. Kaldor N. Welfare propositions of economics and interpersonal comparisons of utility. The Economic Journal. 1939:549-52.
125. Hicks J. The foundations of welfare economics. The economic journal. 1939;49:696-712.
126. Robinson R. Cost-benefit analysis. BMJ. 1993;307(9):924-6.
127. Hurley J. Welfarism, Extra-Welfarism and Evaluative Economic Analysis in the Health Sector, . In: Barer ML, Getzen, T.E., Stoddart, G.L., editor. Health, Health Care and Health Economics. Chichester: John Wiley & Sons; 1998. p. 373-95.
128. Coast J. Is economic evaluation in touch with society's health values? Bmj. 2004;329(7476):1233-6.
129. Culyer AJ. The normative economics of health care finance and provision. In: A., McGuire PF, and K.Mayhew editors. Providing health care: The economics of alternative systems of finance and delivery. Oxford: Oxford University Press; 1991.
130. Sen AK. Equality of what? In: McMurrin S, editor. The Tanner Lectures on Human Values. Cambridge University Press: Cambridge; 1980.
131. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual [PMG20]. London, UK; 2018.
132. Devlin NJ, Lorgelly PK. QALYs as a measure of value in cancer. Journal of Cancer Policy. 2017;11:19-25.
133. Dolan P, Shaw R, Tsuchiya A, Williams A. QALY maximisation and people's preferences: a methodological review of the literature. Health economics. 2005;14(2):197-208.
134. Weinstein MC. A QALY is a QALY is a QALY—or is it? : North-Holland; 1988.
135. Peasgood T, Ward SE, Brazier J. Health-state utility values in breast cancer. Expert Rev Pharmacoecon Outcomes Res. 2010;10(5):553-66.
136. Earle C, Chapman R, Baker C, Bell C, Stone P, Sandberg E, et al. Systematic overview of cost-utility assessments in oncology. Journal of Clinical Oncology. 2000;18(18):3302-17.

137. Brazier JE, Rowen D, Lloyd A, Karimi M. Future directions in valuing benefits for estimating QALYs: is time up for the EQ-5D? Value in Health. 2019;22(1):62-8.
138. Brazier J, Green C, McCabe C, Stevens K. Use of visual analog scales in economic evaluation. Expert review of pharmacoeconomics & outcomes research. 2003;3(3):293-302.
139. Green C, Brazier J, Deverill M. Valuing health-related quality of life. Pharmacoeconomics. 2000;17(2):151-65.
140. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. Journal of health economics. 1996;15(2):209-31.
141. Bleichrodt H, Johannesson M. An experimental test of a theoretical foundation for rating-scale valuations. Medical Decision Making. 1997;17(2):208-16.
142. Torrance GW, Feeny D, Furlong W. Visual analog scales: Do they have a role in the measurement of preferences for health states? Thousand Oaks, CA: Sage Publications; 2001.
143. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2011;20(10):1727-36.
144. Ashby J, O'hanlon M, Buxton MJ. The time trade-off technique: how do the valuations of breast cancer patients compare to those of other groups? Quality of life research. 1994;3(4):257-65.
145. Mehrez A, Gafni A. The healthy-years equivalents: how to measure them using the standard gamble approach. Medical Decision Making. 1991;11(2):140-6.
146. Van der Pol MM, Cairns JA. Negative and zero time preference for health. Health economics. 2000;9(2):171-5.
147. Grann VR, Patel P, Bharthuar A, Jacobson JS, Warner E, Anderson K, et al. Breast cancer-related preferences among women with and without BRCA mutations. Breast cancer research and treatment. 2009;119(1):177.
148. Hayman JA, Fairclough DL, Harris JR, Weeks JC. Patient preferences concerning the trade-off between the risks and benefits of routine radiation therapy after conservative surgery for early-stage breast cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 1997;15(3):1252-60.

149. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw S, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Medical care. 2002;40(2):113-28.
150. Richardson J. Cost utility analysis: what should be measured? Social science & medicine. 1994;39(1):7-21.
151. Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. Health economics. 2002;11(5):447-56.
152. Gafni A. The standard gamble method: what is being measured and how it is interpreted. Health Services Research. 1994;29(2):207.
153. Rowen D, Zouraq IA, Chevrou-Severac H, van Hout B. International regulations and recommendations for utility data for health technology assessment. Pharmacoeconomics. 2017;35(1):11-9.
154. Richardson JRJ, Mckie JR, Bariola EJ. Multiattribute utility instruments and their use. Encylopedia of Health Economics, Volume 2: Elsevier; 2014. p. 341-57.
155. Brooks R, Group E. EuroQol: the current state of play. Health policy. 1996;37(1):53-72.
156. Dolan P. Modeling valuations for EuroQol health states. Medical care. 1997:1095-108.
157. Brazier J, Roberts J, Tsuchiya A, Busschbach J. A comparison of the EQ-5D and SF-6D across seven patient groups. Health economics. 2004;13(9):873-84.
158. Netten A, Burge P, Malley J, Potoglou D, Towers A-M, Brazier J, et al. Outcomes of social care for adults: developing a preference-weighted measure. Health Technology Assessment. 2012;16(16):1-166.
159. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. Journal of health economics. 2002;21(2):271-92.
160. Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 Health Survey. Journal of clinical epidemiology. 1998;51(11):1115-28.
161. Richardson J, McKie J, Bariola E. Review and critique of health related multi attribute utility instruments. Melbourne, Australia: Centre for Health Economics, Monash University; 2011.

162. Lung T, Howard K, Etherton-Beer C, Sim M, Lewin G, Arendts G. Comparison of the HUI3 and the EQ-5D-3L in a nursing home setting. PloS one. 2017;12(2):e0172796.
163. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Quality of Life Research. 1999;8(3):209-24.
164. Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the Assessment of Quality of Life (AQoL)-8D multi-attribute utility instrument. The patient. 2014;7(1):85-96.
165. Hawthorne G, Osborne R. Population norms and meaningful differences for the Assessment of Quality of Life (AQoL) measure. Australian and New Zealand journal of public health. 2005;29(2):136-42.
166. Fayers P, Bottomley A. Quality of life research within the EORTC-the EORTC QLQ-C30. European Organisation for Research and Treatment of Cancer. Eur J Cancer. 2002;38 (4):S125-33.
167. Brazier J, Rowen D, Mavranezouli I, Tsuchiya A, Young T, Yang Y, et al. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). 2012.
168. Marriott ER, van Hazel G, Gibbs P, Hatswell AJ. Mapping EORTC-QLQ-C30 to EQ-5D-3L in patients with colorectal cancer. Journal of medical economics. 2017;20(2):193-9.
169. Gerard K, Johnston K, Brown J. The role of a pre scored multi attribute health classification measure in validating condition specific health state descriptions. Health economics. 1999;8(8):685-99.
170. Dolan P. Whose preferences count? Medical decision making: an international journal of the Society for Medical Decision Making. 1999;19(4):482-6.
171. Conner-Spady B, Suarez-Almazor ME. Variation in the estimation of quality-adjusted life-years by different preference-based instruments. Medical care. 2003;41(7):791-801.
172. Ubel PA, Loewenstein G, Jepson C. Whose quality of life? A commentary exploring discrepancies between health state evaluations of patients and the general public. Qual Life Res. 2003;12(6):599-607.

173. Buckingham K. A note on HYE (healthy years equivalent). J Health Econ. 1993;12(3):301-9.
174. Litva A, Coast J, Donovan J, Eyles J, Shepherd M, Tacchi J, et al. 'The public is too subjective': public involvement at different levels of health-care decision making. Social science & medicine (1982). 2002;54(12):1825-37.
175. Gold MR, Siegel, J.E., Russell, L.B.,. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.
176. De Haes J, de Koning HJ, van Oortmarssen GJ, van Agt HM, de Bruyn AE, van der Maas PJ. The impact of a breast cancer screening programme on quality adjusted life years. International journal of cancer. 1991;49(4):538-44.
177. Wolfson AD, AJ. Sinclair, C. Bombardier and A. McGeer. Preference measurement for functional status in stroke patients: Inter-rater and inter-technique comparisons,. In: Robert, Kane KaR, editors. Values and long-term care. Lexington, MA: Lexington Books; 1982. p. 191-214.
178. Brazier J, Deverill M, Green C. A review of the use of health status measures in economic evaluation. Journal of health services research & policy. 1999;4(3):174-84.
179. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.
180. Black WC. The CE plane: a graphic representation of cost-effectiveness. Medical decision making : an international journal of the Society for Medical Decision Making. 1990;10(3):212-4.
181. Neumann PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. Annu Rev Public Health. 2000;21:587-611.
182. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Medical decision making. 2012;32(5):722-32.
183. Papaioannou D, Brazier J, Paisley S. Systematic Searching and Selection of Health State Utility Values from the Literature. Value in Health. 2013;16(4):686-95.
184. Ara R, Brazier JE. Populating an economic model with health state utility values: moving toward better practice. Value in Health. 2010;13(5):509-18.

185. Jackson LJ, Auguste P, Low N, Roberts TE. Valuing the health states associated with Chlamydia trachomatis infections and their sequelae: a systematic review of economic evaluations and primary studies. Value in Health. 2014;17(1):116-30.
186. Brown ML, Fintor L. Cost-effectiveness of breast cancer screening: preliminary results of a systematic review of the literature. Breast cancer research and treatment. 1993;25(2):113-8.
187. Schiller-Fruhwirth IC, Jahn B, Arvandi M, Siebert U. Cost-Effectiveness Models in Breast Cancer Screening in the General Population: A Systematic Review. Applied Health Economics & Health Policy. 2017;15(3):333-51.
188. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. Breast Cancer Research. 2015;17(1):63.
189. Badia X, Herdman M, Dipstat MR, Ohinmaa A. Feasibility and validity of the VAS and TTO for eliciting general population values for temporary health states: a comparative study. Health Services and Outcomes Research Methodology. 2001;2(1):51-65.
190. Partridge A, Adloff K, Blood E, Dees EC, Kaelin C, Golshan M, et al. Risk perceptions and psychosocial outcomes of women with ductal carcinoma in situ: longitudinal results from a cohort study. Journal of the National Cancer Institute. 2008;100(4):243-51.
191. Weitzner MA, Meyers CA, Stuebing KK, Saleeba AK. Relationship between quality of life and mood in long-term survivors of breast cancer treated with mastectomy. Support Care Cancer. 1997;5(3):241-8.
192. Tosteson AN, Fryback DG, Hammond CS, Hanna LG, Grove MR, Brown M, et al. Consequences of false-positive screening mammograms. JAMA internal medicine. 2014;174(6):954-61.
193. Wright DR, Wittenberg E, Swan JS, Miksad RA, Prosser LA. Methods for measuring temporary health states for cost-utility analyses. Pharmacoeconomics. 2009;27(9):713-23.
194. Brazier JE, Dixon S, Ratcliffe J. The role of patient preferences in cost-effectiveness analysis. Pharmacoeconomics. 2009;27(9):705-12.

195. Smith RA. Expert Group: IARC Handbooks of Cancer Prevention. Vol. 7: Breast Cancer Screening. Lyon, France: IARC; 2002. 248pp. Breast Cancer Research. 2003;5(4):216.
196. Brazier J, Deverill M. A checklist for judging preference-based measures of health related quality of life: learning from psychometrics. Health economics. 1999;8(1):41-51.
197. Zelle SG, Baltussen RM. Economic analyses of breast cancer control in low-and middle-income countries: a systematic review. Systematic reviews. 2013;2(1):20.
198. Centre for Reviews and Dissemination. CRD Guidance for Undertaking Reviews in Health Care. Layerthorpe: University of York; 2009.
199. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS medicine. 2009;6(7):e1000100.
200. Higgins JP, Altman DG. Assessing risk of bias in included studies. Cochrane handbook for systematic reviews of interventions: Cochrane book series. 2008:187-241.
201. McDonagh M, Peterson K, Raina P, Chang S, Shekelle P. Avoiding bias in selecting studies. In Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Washington, US: Agency for Healthcare Research and Quality; 2013.
202. Endnote Version X7.0. Referencing software. Boston, MA Clarivate Analytics; 2018.
203. Ogwulu CB, Jackson LJ, Kinghorn P, Roberts TE. A systematic review of the techniques used to value temporary health states. Value in Health. 2017;20(8):1180-97.
204. Ahern CH, Shen Y. Cost-effectiveness analysis of mammography and clinical breast examination strategies: a comparison with current guidelines. Cancer Epidemiology and Prevention Biomarkers. 2009;18(3):718-25.
205. Arrospide A, Rue M, van Ravesteyn NT, Comas M, Soto-Gordoa M, Sarriugarte G, et al. Economic evaluation of the breast cancer screening programme in the Basque Country: retrospective cost-effectiveness and budget impact analysis. BMC cancer. 2016;16(1):344.

206. Beemsterboer PM, De Koning HJ, Warmerdam PG, Boer R, Swart E, Dierks ML, et al. Prediction of the effects and costs of breast-cancer screening in Germany. International journal of cancer. 1994;58(5):623-8.
207. Carles M, Vilaprinyo E, Cots F, Gregori A, Pla R, Román R, et al. Cost-effectiveness of early detection of breast cancer in Catalonia. BMC cancer. 2011;11(1):192.
208. Christensen MK, Niclasen BV, Iburg KM. Societal costs and effects of implementing population-based mammography screening in Greenland. International Journal of Circumpolar Health. 2017;76(1):1373580.
209. de Koning HJ, Martin van Ineveld B, van Oortmarssen GJ, De Haes J, Collette HJ, Hendriks JH, et al. Breast cancer screening and cost effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. International Journal of Cancer. 1991;49(4):531-7.
210. Haghighat S, Akbari ME, Yavari P, Javanbakht M, Ghaffari S. Cost-effectiveness of three rounds of mammography breast cancer screening in Iranian women. Iranian journal of cancer prevention. 2016;9(1):e5443.
211. Hristova L, Hakama M. Effect of screening for cancer in the Nordic countries on deaths, cost and quality of life up to the year 2017. Acta Oncologica. 1997;36:1-60.
212. BreastScreen Australia. Economic evaluation and modelling study. Canberra, Australia: Australian Government Department of Health and Ageing; 2009.
213. Mittmann N, Stout NK, Tosteson ANA, Trentham-Dietz A, Alagoz O, Yaffe MJ. Cost-effectiveness of mammography from a publicly funded health care system perspective. CMAJ Open. 2018;6(1):E77-E86.
214. Morton R, Sayma M, Sura MS. Economic analysis of the breast cancer screening program used by the UK NHS: should the program be maintained? Breast cancer (Tokyo, Japan). 2017;9:217-25.
215. Pataky R, Ismail Z, Coldman AJ, Elwood M, Gelmon K, Hedden L, et al. Cost-effectiveness of annual versus biennial screening mammography for women with high mammographic breast density. Journal of medical screening. 2014;21(4):180-8.
216. Rojnik K, Naveršnik K, Mateović-Rojnik T, PrimicŽakelj M. Probabilistic cost effectiveness modeling of different breast cancer screening policies in Slovenia. Value in Health. 2008;11(2):139-48.

217. Souza FH, Polanczyk CA. Is Age-targeted full-field digital mammography screening cost-effective in emerging countries? A micro simulation model. Springerplus. 2013;2(1):366.
218. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. Journal of the National Cancer Institute. 2006;98(11):774-82.
219. Stout NK, Lee SJ, Schechter CB, Kerlikowske K, Alagoz O, Berry D, et al. Benefits, harms, and costs for breast cancer screening after US implementation of digital mammography. JNCI: Journal of the National Cancer Institute. 2014;106(6):dju092.
220. Tosteson AN, Stout NK, Fryback DG, Acharyya S, Herman BA, Hannah LG, et al. Cost-effectiveness of digital mammography breast cancer screening. Annals of internal medicine. 2008;148(1):1-10.
221. Trentham-Dietz A, Kerlikowske K, Stout NK, Miglioretti DL, Schechter CB, Ergun MA, et al. Tailoring breast cancer screening intervals by breast density and risk for women aged 50 years or older: collaborative modeling of screening outcomes. Annals of internal medicine. 2016;165(10):700-12.
222. van Luijt PA, Heijnsdijk EAM, de Koning HJ. Cost-effectiveness of the Norwegian breast cancer screening program. International Journal of Cancer. 2017;140(4):833-40.
223. Wong IO, Kuntz KM, Cowling BJ, Lam CL, Leung GM. Cost effectiveness of mammography screening for Chinese women. Cancer: Interdisciplinary International Journal of the American Cancer Society. 2007;110(4):885-95.
224. Kerlikowske K, Salzmann P, Phillips KA, Cauley JA, Cummings SR. Continuing screening mammography in women aged 70 to 79 years: impact on life expectancy and cost-effectiveness. Jama. 1999;282(22):2156-63.
225. Madan J, Rawdin A, Stevenson M, Tappenden P. A rapid-response economic evaluation of the UK NHS Cancer Reform Strategy breast cancer screening program extension via a plausible bounds approach. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2010;13(2):215-21.

226. Mandelblatt JS, Wheat ME, Monane M, Moshief RD, Hollenberg JP, Tang J. Breast cancer screening for elderly women with and without comorbid conditions: a decision analysis model. Annals of internal medicine. 1992;116(9):722-30.
227. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Muennig P, et al. Benefits and costs of interventions to improve breast cancer outcomes in African American women. Journal of Clinical Oncology. 2004;22(13):2554-66.
228. Mandelblatt JS, Schechter CB, Yabroff KR, Lawrence W, Dignam J, Extermann M, et al. Toward optimal screening strategies for older women. Journal of general internal medicine. 2005;20(6):487-96.
229. Boer R, De Koning H, Van Oortmarssen G, Van der Maas P. In search of the best upper age limit for breast cancer screening. European Journal of Cancer. 1995;31(12):2040-3.
230. Rafia R, Brennan A, Madan J, Collins K, Reed MW, Lawrence G, et al. Modeling the cost-effectiveness of alternative upper age limits for breast cancer screening in England and Wales. Value in Health. 2016;19(4):404-12.
231. Salzmann P, Kerlikowske K, Phillips K. Cost-effectiveness of extending screening mammography guidelines to include women 40 to 49 years of age. Annals of Internal Medicine. 1997;127(11):955-65.
232. Sankatsing VD, Heijnsdijk EA, van Luijt PA, van Ravesteyn NT, Fracheboud J, de Koning HJ. Cost-effectiveness of digital mammography screening before the age of 50 in T he N etherlands. International journal of cancer. 2015;137(8):1990-9.
233. Schousboe JT, Kerlikowske K, Loh A, Cummings SR. Personalizing mammography by breast density and other risk factors for breast cancer: analysis of health benefits and cost-effectiveness. Annals of internal medicine. 2011;155(1):10-20.
234. Vilaprinyo E, Forne C, Carles M, Sala M, Pla R, Castells X, et al. Cost-effectiveness and harm-benefit analyses of risk-based screening strategies for breast cancer. PloS one. 2014;9(2):e86858.
235. de Gelder R, Bulliard J-L, de Wolf C, Fracheboud J, Draisma G, Schopper D, et al. Cost-effectiveness of opportunistic versus organised mammography screening in Switzerland. European Journal of cancer. 2009;45(1):127-38.
236. Van Ravesteyn NT, Stout NK, Schechter CB, Heijnsdijk EA, Alagoz O, Trentham-Dietz A, et al. Benefits and harms of mammography screening after age 74

years: model estimates of overdiagnosis. Journal of the National Cancer Institute. 2015;107(7).
237. Oliver A. Should we maximise QALYs? Applied health economics and health policy. 2004;3(2):61-6.
238. McNamee P, Glendinning S, Shenfine J, Steen N, Griffin SM, Bond J. Chained time trade-off and standard gamble methods. The European Journal of Health Economics. 2004;5(1):81-6.
239. Hoover DR, Crystal S, Kumar R, Sambamoorthi U, Cantor JC. Medical expenditures during the last year of life: findings from the 1992–1996 Medicare current beneficiary survey. Health services research. 2002;37(6):1625-42.
240. Tengs TO, Wallace A. One thousand health-related quality-of-life estimates. Medical care. 2000:583-637.
241. Parmigiani G. Modeling in medical decision making. West Sussex, England: John Wiley and Sons; 2002.
242. Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? International journal of epidemiology. 1978;7(4):347-58.
243. Hall J, Gerard K, Salkeld G, Richardson J. A cost utility analysis of mammography screening in Australia. Social Science & Medicine. 1992;34(9):993-1004.
244. Kind P, Hardman G, Macran S. UK population norms for EQ-5D. York: Centre for Health Economics, University of York; 1999.
245. Gerard K, Dobson M, Hall J. Framing and labelling effects in health descriptions: quality adjusted life years for treatment of breast cancer. Journal of Clinical Epidemiology. 1993;46(1):77-84.
246. Ara R, Brazier JE. Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. Value in Health. 2011;14(4):539-45.
247. Bonomi AE, Boudreau DM, Fishman PA, Ludman E, Mohelnitzky A, Cannon EA, et al. Quality of life valuations of mammography screening. Quality of Life Research. 2008;17(5):801-14.
248. Hayman JA, Kabeto MU, Schipper MJ, Bennett JE, Vicini FA, Pierce LJ. Assessing the benefit of radiation therapy after breast-conserving surgery for ductal

carcinoma-in-situ. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2005;23(22):5171-7.
249. Rijnsburger A, Essink-Bot M-L, van Dooren S, Borsboom G, Seynaeve C, Bartels C, et al. Impact of screening for breast cancer in high-risk women on health-related quality of life. British Journal of Cancer. 2004;91(1):69.
250. Lidgren M, Wilking N, Jönsson B, Rehnberg C. Health related quality of life in different states of breast cancer. Quality of life research. 2007;16(6):1073-81.
251. Souza FH. Mamografia digital em comparação com mamografia convencional no rastreamento de câncer de mama no Brasil: revisão sistemática, custo da doença e analise de custo-efetividade no Sistema Único de Saúde. 2012.
252. Cruz LN. Medidas de qualidade de vida e utilidade em uma amostra da população de Porto Alegre. 2010.
253. Chie W-C, Huang C-S, Chen J-H, Chang K-J. Utility Assessment for Different Clinical Phases of Breast Cancer in Taiwan. Journal of the Formosan Medical Association. 2000;99(9):677-83.
254. Kim SH, Jo MW, Ock M, Lee HJ, Lee JW. Estimation of health state utilities in breast cancer. Patient Preference and Adherence. 2017;11:531-6.
255. Dolan P. Developing methods that really do value the 'Q'in the QALY. Health economics, policy and law. 2008;3(1):69-77.
256. Lamers LM, Stalmeier PF, Krabbe PF, Busschbach JJ. Inconsistencies in TTO and VAS values for EQ-5D health states. Medical Decision Making. 2006;26(2):173-81.
257. Furlong W, Feeny D, Torrance G, Barr R, Horsman J. Guide to design and development of health-state utility instrumentation. Centre for Health Economics and Policy Analysis (CHEPA), McMaster University, Hamilton, Canada; 1992.
258. Longworth L, Rowen D. Mapping to obtain EQ-5D utility values for use in NICE health technology assessments. Value in Health. 2013;16(1):202-10.
259. Von Neuman J, Morgenstern O. Theory of games and economic behavior. 1953.
260. Tolley KH. What are health utilities? What is series, 2009.
261. Shepard DS. Cost effectiveness in Health and Medicine. By MR Gold, JE Siegel, LB Russell, and MC Weinstein (eds). New York: Oxford University Press, 1996. The journal of mental health policy and economics. 1999;2(2):91-2.

262. Bombardier C, Wolfson A, Sinclair A, McGreer A. Comparison of three preference measurement methodologies in the evaluation of a functional status index. Choices in Health Care: University of Toronto, Toronto; 1982.
263. Aaronson NK, Muller M, Cohen PD, Essink-Bot M-L, Fekkes M, Sanderman R, et al. Translation, validation, and norming of the Dutch language version of the SF-36 Health Survey in community and chronic disease populations. Journal of clinical epidemiology. 1998;51(11):1055-68.
264. Torrance GW, Furlong W, Feeny D. Health utility estimation. Expert Review of Pharmacoeconomics & Outcomes Research. 2002;2(2):99-108.
265. Attema AE, Brouwer WB. The value of correcting values: influence and importance of correcting TTO scores for time preference. Value in Health. 2010;13(8):879-84.
266. Hughes R, Huby M. The construction and interpretation of vignettes in social research. Social Work and Social Sciences Review. 2012;11(1):36-51.
267. Hershey JC, Kunreuther HC, Schoemaker PJ. Sources of bias in assessment procedures for utility functions. Management science. 1982;28(8):936-54.
268. O'Connor AM. Effects of framing and level of probability on patients' preferences for cancer chemotherapy. Journal of Clinical Epidemiology. 1989;42(2):119-26.
269. Jansen SJ, Stiggelbout AM, Wakker PP, Vliet Vlieland TP, Leer J-WH, Nooy MA, et al. Patients' utilities for cancer treatments: a study of the chained procedure for the standard gamble and time tradeoff. Medical Decision Making. 1998;18(4):391-9.
270. Brazier J. Valuing health states for use in cost-effectiveness analysis. Pharmacoeconomics. 2008;26(9):769-79.
271. Tosh JC, Longworth LJ, George E. Utility values in National Institute for Health and Clinical Excellence (NICE) technology appraisals. Value in Health. 2011;14(1):102-9.
272. Lew J-B, Simms K, Smith M, Lewis H, Neal H, Canfell K. Effectiveness Modelling and Economic Evaluation of Primary HPV Screening for Cervical Cancer Prevention in New Zealand. PLOS ONE. 2016;11(5):e0151619.

273. 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.
274. McCaffery K, Nickel B, Moynihan R, Hersch J, Teixeira-Pinto A, Irwig L, et al. How different terminology for ductal carcinoma in situ impacts women's concern and treatment preferences: a randomised comparison within a national community survey. BMJ open. 2015;5(11):e008094.
275. Nickel B, Barratt A, Copp T, Moynihan R, McCaffery K. Words do matter: a systematic review on how different terminology for the same condition influences management preferences. BMJ open. 2017;7(7):e014129.
276. Campbell B, Yang JC, Gonzalez J, Reed S, Havrilesky L, Johnson FR, et al. Patient preferences for DCIS treatment strategies: A discrete-choice experiment. Annals of surgical oncology. 2018;25 (2 Supplement 1):38-9.
277. Esserman LJ, Thompson IM, Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. JAMA. 2013;310(8):797-8.
278. Keller A, Gericke C, Whitty JA, Yaxley J, Kua B, Coughlin G. A Cost-Utility Analysis of Prostate Cancer Screening in Australia. Applied Health Economics and Health Policy. 2017;15(1):95-111.
279. Lindfors KK, O'Connor J, Parker RA. False-positive screening mammograms: effect of immediate versus later work-up on patient stress. Radiology. 2001;218(1):247-53.
280. Parker PA, Youssef A, Walker S, Basen-Engquist K, Cohen L, Gritz ER, et al. Short-term and long-term psychosocial adjustment and quality of life in women undergoing different surgical procedures for breast cancer. Annals of surgical oncology. 2007;14(11):3078-89.
281. Stiggelbout AM, Kiebert GM, Kievit J, Leer J-WH, Stoter G, De Haes J. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. Medical Decision Making. 1994;14(1):82-90.
282. Arnold M. Simulation modeling for stratified breast cancer screening - a systematic review of cost and quality of life assumptions. BMC Health Services Research. 2017;17(1):802.

283. Verhoef C, Maas A, Stalpers L, Verbeek A, Wobbes T, van Daal W. The feasibility of additive conjoint measurement in measuring utilities in breast cancer patients. Health Policy. 1991;17(1):39-50.
284. Bromley HL, Petrie D, Mann GB, Nickson C, Rea D, Roberts TE. Valuing the health states associated with breast cancer screening programmes: A systematic review of economic measures. Social science & medicine. 2019;228:142-54.
285. Department of Health and Social Care. The Independent Breast Screening Review 2018. London, England: House of Commons; 2018 13 December 2018.
286. Sanghera S, Coast J, Martin RM, Donovan JL, Mohiuddin S. Cost-effectiveness of prostate cancer screening: a systematic review of decision-analytical models. BMC cancer. 2018;18(1):84.
287. Dyer JS, Sarin RK. Relative Risk Aversion. Management Science. 1982;28(8):875-86.
288. Jansen SJT, Stiggelbout AM, Nooij MA, Kievit J. The effect of individually assessed preference weights on the relationship between holistic utilities and nonpreference-based assessment. Quality of Life Research. 2000;9(5):541-57.
289. Stewart ST, Lenert L, Bhatnagar V, Kaplan RM. Utilities for prostate cancer health states in men aged 60 and older. Medical care. 2005;43(4):347-55.
290. Simonella L, Howard K, Canfell K. A survey of population-based utility scores for cervical cancer prevention. BMC research notes. 2014;7:899-.
291. van Hout B, Janssen M, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value in Health. 2012;15(5):708-15.
292. Payakachat N, Ali MM, Tilford JM. Can The EQ-5D Detect Meaningful Change? A Systematic Review. PharmacoEconomics. 2015;33(11):1137-54.
293. Pennington B, Hernandez-Alava M, Pudney S, Wailoo A. The Impact of Moving from EQ-5D-3L to -5L in NICE Technology Appraisals. PharmacoEconomics. 2019;37(1):75-84.
294. Galante J, Augustovski F, Colantonio L, Bardach A, Caporale J, Marti SG, et al. Estimation and comparison of EQ-5D health states' utility weights for pneumococcal and human papillomavirus diseases in Argentina, Chile, and the United Kingdom.

Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2011;14(5 Suppl 1):S60-4.
295. Jacklyn G, Glasziou P, Macaskill P, Barratt A. Meta-analysis of breast cancer mortality benefit and overdiagnosis adjusted for adherence: improving information on the effects of attending screening mammography. British Journal of Cancer. 2016;114(11):1269.
296. Garg PK, Jakhetiya A, Pandey R, Chishi N, Pandey D. Adjuvant radiotherapy versus observation following lumpectomy in ductal carcinoma in-situ: A meta-analysis of randomized controlled trials. The breast journal. 2018;24(3):233-9.
297. Edwards A, Elwyn G, Covey J, Matthews E, Pill R. Presenting risk information a review of the effects of framing and other manipulations on patient outcomes. Journal of health communication. 2001;6(1):61-82.
298. Hersch J, Jansen J, Barratt A, Irwig L, Houssami N, Howard K. Women's views on overdiagnosis in breast cancer screening: a qualitative study. BMJ. 2013;346.
299. Nickel B, Barratt A, Hersch J, Moynihan R, Irwig L, McCaffery K. How different terminology for ductal carcinoma in situ (DCIS) impacts women's concern and management preferences: A qualitative study. The Breast. 2015;24(5):673-9.
300. McNeil BJ, Pauker SG, Sox Jr HC, Tversky A. On the Elicitation of Preferences for Alternative Therapies. Preference, Belief, and Similarity. 1982;583.
301. Hawthorne G, Richardson J, Day NA. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Annals of medicine. 2001;33(5):358-70.
302. McCaffrey N, Kaambwa B, Currow DC, Ratcliffe J. Health-related quality of life measured using the EQ-5D-5L: South Australian population norms. Health and quality of life outcomes. 2016;14(1):133.
303. Clemens S, Begum N, Harper C, Whitty J, Scuffham P. A comparison of EQ-5D-3L population norms in Queensland, Australia, estimated using utility value sets from Australia, the UK and USA. Quality of Life Research. 2014;23(8):2375-81.
304. Viney R. An Australian Discrete Choice Experiment to Value EQ-5D Health States. Health economics. 2014;23(6):729-42.

305. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. Pharmacoeconomics. 2018;36(6):675-97.
306. Alava MH, Wailoo A, Grimm S, Pudney S, Gomes M, Sadique Z, et al. EQ-5D-5L versus EQ-5D-3L: the impact on cost effectiveness in the United Kingdom. Value in Health. 2018;21(1):49-56.
307. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Affairs. 2007;26(3):741-8.
308. Reyna VF, Nelson WL, Han PK, Dieckmann NF. How numeracy influences risk comprehension and medical decision making. Psychological bulletin. 2009;135(6):943.
309. Lloyd AJ. The extent of patients' understanding of the risk of treatments. Quality in Health Care. 2001;10(suppl 1):i14-i8.
310. Kahneman D, Slovic SP, Slovic P, Tversky A. Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press; 1982.
311. Black WC, Nease Jr RF, Tosteson AN. Perceptions of breast cancer risk and screening effectiveness in women younger than 50 years of age. JNCI: Journal of the National Cancer Institute. 1995;87(10):720-31.
312. Schwartz LM, Woloshin S, Black WC, Welch HG. The role of numeracy in understanding the benefit of screening mammography. Annals of internal medicine. 1997;127(11):966-72.
313. Woloshin S, Schwartz LM, Black WC, Welch HG. Women's perceptions of breast cancer risk: how you ask matters. Medical Decision Making. 1999;19(3):221-9.
314. Gudex C. Standard Gamble user manual: props and self-completion methods. York: Centre for Health Economics, University of York; 1994.
315. Fryback DG, Dasbach EJ, Klein R, Klein BE, Dorn N, Peterson K, et al. The Beaver Dam Health Outcomes Study: initial catalog of health-state quality factors. Medical Decision Making. 1993;13(2):89-102.
316. Breast Cancer Network Australia. Breast Cancer Network Australia [internet]: BCNA; 2019 [cited 2019 12 March]. Available from: https://www.bcna.org.au/.
317. Australian Bureau of Statistics. Census Household Form. Canberra, Australia; 2016

318. Lifepool. Lifepool Cohort Peter MacCallum Cancer Centre, Victoria, Australia: National Breast Cancer Foundation; 2019 [cited 2019 12 March]. Available from: http://www.lifepool.org/.
319. Altman DG. Statistics and ethics in medical research III: How large a sample? BMJ. 1980;281(6251):1336.
320. EuroQol group. EQ-5D-5L Valuation Crosswalk Index Value Calculator London, UK2018 [Available from: https://euroqol.org/eq-5d-instruments/eq-5d-5l-about/valuation-standard-value-sets/crosswalk-index-value-calculator/.
321. StataCorp. Stata Statistical Software. TX, USA: StataCorp LLC; 2017.
322. Walker JaA, P. Interpreting statistical findings : a guide for health professionals and students. Berkshire, England Open University Press; 2010.
323. Altman DG. Practical statistics for medical researchers. London: Chapman and Hall; 1991.
324. Hollander M, Wolfe DA, Chicken E. Nonparametric Statistical Methods. New Jersey, USA: John Wiley & Sons; 2013.
325. Buchholz I, Janssen MF, Kohlmann T, Feng Y-S. A systematic review of studies comparing the measurement properties of the three-level and five-level versions of the EQ-5D. Pharmacoeconomics. 2018;36(6):645-61.
326. Elmore JG, Fletcher SW. Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. Ann Intern Med. 2012;156(7):536-7.
327. York Health Economics Consortium. Disutility [online] York2016 [Available from: https://www.yhec.co.uk/glossary/disutility/.
328. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The Measurement of Patients' Values in Medicine. Medical Decision Making. 1982;2(4):449-62.
329. Rutten-van Molken M, Custers F, Van Doorslaer E, Jansen C, Heurman L, Maesen F, et al. Comparison of performance of four instruments in evaluating the effects of salmeterol on asthma quality of life. European Respiratory Journal. 1995;8(6):888-98.
330. Feng Y, Devlin N, Herdman M. Assessing the health of the general population in England: how do the three- and five-level versions of EQ-5D compare? Health and quality of life outcomes. 2015;13:171.

331. Dolan P, Kahneman D. Interpretations of utility and their implications for the valuation of health. The economic journal. 2008;118(525):215-34.
332. 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. Medical decision making : an international journal of the Society for Medical Decision Making. 2008;28(1):102-12.
333. 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.
334. Perez DJ, Williams SM, Christensen EA, McGee RO, Campbell AV. A longitudinal study of health related quality of life and utility measures in patients with advanced breast cancer. Quality of Life Research. 2001;10(7):587-93.
335. Brazier J, Rowen D, Karimi M, Peasgood T, Tsuchiya A, Ratcliffe J. Experience-based utility and own health state valuation for a health state classification system: why and how to do it. European journal of health economics. 2018;19(6):881-91.
336. Curigliano G, Burstein HJ, E PW, Gnant M, Dubsky P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: the St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. Annals of oncology : official journal of the European Society for Medical Oncology. 2017;28(8):1700-12.
337. Barton P, Bryan S, Robinson S. Modelling in the economic evaluation of health care: selecting the appropriate approach. Journal of health services research & policy. 2004;9(2):110-8.
338. Buxton MJ, Drummond MF, Van Hout BA, Prince RL, Sheldon TA, Szucs T, et al. Modelling in economic evaluation: an unavoidable fact of life. Health economics. 1997;6(3):217-27.
339. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health economics. 2006;15(12):1295-310.
340. Briggs A, Sculpher M. An Introduction to Markov Modelling for Economic Evaluation. PharmacoEconomics. 1998;13(4):397-409.

341. 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.
342. Lambe T, Adab P, Jordan RE, Sitch A, Enocson A, Jolly K, et al. Model-based evaluation of the long-term cost-effectiveness of systematic case-finding for COPD in primary care. Thorax. 2019;74(8):730-9.
343. Sanghera S, Barton P, Bhattacharya S, Horne AW, Roberts TE. Pharmaceutical treatments to prevent recurrence of endometriosis following surgery: a model-based economic evaluation. BMJ Open. 2016;6(4):e010580.
344. Roberts TE, Robinson S, Barton P, Bryan S, Low N, Chlamydia Screening Studies G. Screening for Chlamydia trachomatis: a systematic review of the economic evaluations and modelling. Sex Transm Infect. 2006;82(3):193-201.
345. Goodwin A, Parker S, Ghersi D, Wilcken N. Post-operative radiotherapy for ductal carcinoma in situ of the breast. The Cochrane database of systematic reviews. 2013(11):Cd000563.
346. Mugford M. How does the method of cost estimation affect the assessment of cost-effectiveness in health care? : University of Oxford; 1996.
347. Roberts T, Henderson J, Mugford M, Bricker L, Neilson J, Garcia J. Antenatal ultrasound screening for fetal abnormalities: a systematic review of studies of cost and cost effectiveness. BJOG: An International Journal of Obstetrics & Gynaecology. 2002;109(1):44-56.
348. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health technology assessment (Winchester, England). 2004;8(36):iii-iv, ix-xi, 1-158.
349. Suh WW, Hillner BE, Pierce LJ, Hayman JA. Cost-effectiveness of radiation therapy following conservative surgery for ductal carcinoma in situ of the breast. International journal of radiation oncology, biology, physics. 2005;61(4):1054-61.
350. Raldow AC, Sher D, Chen AB, Recht A, Punglia RS. Cost Effectiveness of the Oncotype DX DCIS Score for Guiding Treatment of Patients With Ductal Carcinoma In Situ. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2016;34(33):3963-8.

351. Trentham-Dietz A, Ergun MA, Alagoz O, Stout NK, Gangnon RE, Hampton JM, et al. Comparative effectiveness of incorporating a hypothetical DCIS prognostic marker into breast cancer screening. Breast cancer research and treatment. 2018;168(1):229-39.
352. Solin LJ, Gray R, Baehner FL, Butler SM, Hughes LL, Yoshizawa C, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. Journal of the National Cancer Institute. 2013;105(10):701-10.
353. Fisher B, Land S, Mamounas E, Dignam J, Fisher ER, Wolmark N. Prevention of invasive breast cancer in women with ductal carcinoma in situ: an update of the National Surgical Adjuvant Breast and Bowel Project experience. Seminars in oncology. 2001;28(4):400-18.
354. Shiyanbola OO, Sprague BL, Hampton JM, Dittus K, James TA, Herschorn S, et al. Emerging trends in surgical and adjuvant radiation therapies among women diagnosed with ductal carcinoma in situ. Cancer. 2016;122(18):2810-8.
355. Yabroff KR, Lamont EB, Mariotto A, Warren JL, Topor M, Meekins A, et al. Cost of care for elderly cancer patients in the United States. Journal of the National Cancer Institute. 2008;100(9):630-41.
356. Attema AE, Brouwer WBF, Claxton K. Discounting in Economic Evaluations. PharmacoEconomics. 2018;36(7):745-58.
357. Elshof LE, Tryfonidis K, Slaets L, Van Leeuwen-Stok AE, Dif N, Skinner VP, et al. The LORD trial: A randomized, non-inferiority trial, between active surveillance versus standard treatment in patients with low risk ductal carcinoma in situ. Cancer Research. 2015;75(9 SUPPL. 1).
358. Esserman L, Yau C. Rethinking the Standard for Ductal Carcinoma In Situ Treatment. JAMA oncology. 2015;1(7):881-3.
359. van Ravesteyn NT, Van Den Broek JJ, Li X, Weedon-Fekjær H, Schechter CB, Alagoz O, et al. Modeling ductal carcinoma in situ (DCIS): An overview of CISNET model approaches. Medical Decision Making. 2018;38(1):S126-39.
360. Soeteman DI, Stout NK, Ozanne EM, Greenberg C, Hassett MJ, Schrag D, et al. Modeling the effectiveness of initial management strategies for ductal carcinoma in situ. Journal of the National Cancer Institute. 2013;105(11):774-81.

361. Hillner BE, Desch CE, Carlson RW, Smith TJ, Esserman L, Bear HD. Trade-offs between survival and breast preservation for three initial treatments of ductal carcinoma-in-situ of the breast. Journal of clinical oncology. 1996;14(1):70-7.
362. Muhlberger N, Sroczynski G, Gogollari A, Beate Jahn BJ, Pashayan N, Widschwendter M, et al. Cost-effectiveness of breast cancer screening and prevention-a systetmatic review of decision-analytic models for European settings. International Journal of Gynecological Cancer. 2017;27 (Supplement 4):365.
363. Ryser MD, Worni M, Turner EL, Marks JR, Durrett R, Shelley Hwang E, et al. Outcomes of Active Surveillance for Ductal Carcinoma in Situ: A Computational Risk Analysis. JNCI: Journal of the National Cancer Institute. 2016;108(5):1-8.
364. Loeb S, Zhou Q, Siebert U, Rochau U, Jahn B, Mühlberger N, et al. Active surveillance versus watchful waiting for localized prostate cancer: a model to inform decisions. European urology. 2017;72(6):899-907.
365. Lao C, Edlin R, Rouse P, Brown C, Holmes M, Gilling P, et al. The cost-effectiveness of active surveillance compared to watchful waiting and radical prostatectomy for low risk localised prostate cancer. BMC cancer. 2017;17(1):529.
366. Sumner WE, 3rd, Koniaris LG, Snell SE, Spector S, Powell J, Avisar E, et al. Results of 23,810 cases of ductal carcinoma-in-situ. Annals of surgical oncology. 2007;14(5):1638-43.
367. de Carvalho TM, Heijnsdijk EAM, de Koning HJ. Comparative effectiveness of prostate cancer screening between the ages of 55 and 69 years followed by active surveillance. Cancer. 2018;124(3):507-13.
368. Keegan KA, Dall'Era MA, Durbin-Johnson B, Evans CP. Active surveillance for prostate cancer compared with immediate treatment: an economic analysis. Cancer. 2012;118(14):3512-8.
369. Sathianathen NJ, Konety BR, Alarid-Escudero F, Lawrentschuk N, Bolton DM, Kuntz KM. Cost-Effectiveness Analysis of Active Surveillance Strategies for Men with Low-risk Prostate Cancer. European Urology. 2019;75(6):910-7.
370. Sharma V, Wymer KM, Borah BJ, Barocas DA, Thompson RH, Karnes RJ, et al. Cost-Effectiveness of Active Surveillance, Radical Prostatectomy, and External Beam Radiotherapy for Localized Prostate Cancer: An Analysis of the ProtecT Trial. The Journal of Urology. 2019;21:10-97.

371. 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. 2013;16(2):231-50.
372. TreeAge Pro. TreeAge Software, . R2. Williamstown, MA: TreeAge; 2018.
373. Francis A. LORIS: A Phase III Trial of Surgery versus Active Monitoring for Low Risk Ductal Carcinoma in Situ (DCIS) Protocol v3.0. Birmingham: University of Birmingham; 2015.
374. Public Health England. NHS public health functions agreement 2018-19: Service specification no.24 Breast Screening Programme,. London, England,: Public Health England with NHS England Public Health Commissioning; 2018.
375. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value in Health. 2012;15(6):812-20.
376. National Clinical Audit and Patient Outcomes Programme. National Audit of Breast Cancer in Older Patients. London, United Kingdom; 2018.
377. Brennan ME, Turner RM, Ciatto S, Marinovich ML, French JR, Macaskill P, et al. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. Radiology. 2011;260(1):119-28.
378. Maxwell AJ, Clements K, Hilton B, Dodwell DJ, Evans A, Kearins O, et al. Risk factors for the development of invasive cancer in unresected ductal carcinoma in situ. European Journal of Surgical Oncology. 2018;44(4):429-35.
379. Ryser M, Weaver DL, Marks JR, Hyslop T, Hwang ES. Quantifying the Natural History and Overtreatment Rate of Ductal Carcinoma in Situ. San Antonio Breast Cancer Symposium; December 5-9, 2017; San Antonio, Texas, USA 2017.
380. Hamilton SN, Nichol A, Wai E, Gondara L, Velasquez Garcia HA, Speers C, et al. Local Relapse After Breast-Conserving Therapy Versus Mastectomy for Extensive Pure Ductal Carcinoma In Situ >/=4 cm. International journal of radiation oncology, biology, physics. 2019;103(2):381-8.
381. Donker M, Litière S, Werutsky G, Julien J-P, Fentiman IS, Agresti R, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma in situ:

15-year recurrence rates and outcome after a recurrence, from the EORTC 10853 randomized phase III trial. Journal of Clinical Oncology. 2013;31(32):4054-9.
382. National Cancer Intelligence Network. All Breast Cancer report: A UK analysis of non-invasive breast cancers diagnosed in England from 1 January 2006 to 31 December 2007 London: National Cancer Intelligence Network; 2011.
383. Worni M, Akushevich I, Greenup R, Sarma D, Ryser MD, Myers ER, et al. Trends in Treatment Patterns and Outcomes for Ductal Carcinoma in Situ. JNCI: Journal of the National Cancer Institute. 2015;107(12).
384. Melvin JC, Purushotham AD, Garmo H, Pinder SE, Fentiman IS, Gillett C, et al. Progression of breast cancer following locoregional ipsilateral recurrence: importance of interval time. British Journal of Cancer. 2016;114(1):88-95.
385. Solin LJ, Fourquet A, Vicini FA, Taylor M, Olivotto IA, Haffty B, et al. Long-term outcome after breast-conservation treatment with radiation for mammographically detected ductal carcinoma in situ of the breast. Cancer: Interdisciplinary International Journal of the American Cancer Society. 2005;103(6):1137-46.
386. Office of National Statistics. National Life Tables, United Kingdom 1980-82 to 2015-2017. London, UK: Office of National Statistics; 2018.
387. Office for National Statistics. Cancer Survival in England: adults diagnosed between 2012 and 2016 and followed up to 2017. London, UK: ONS; 2019.
388. Curtis L, Burns A. Unit Costs of Health and Social Care 2018. University of Kent, Canterbury: Personal Social Services Research Unit; 2018.
389. National Health Service. National schedule of reference costs. London: National Health Service Improvement; 2018.
390. Gray E, Donten A, Karssemeijer N, van Gils C, Evans D, Astley S, et al. Evaluation of a stratified national breast screening program in the United Kingdom: An early model-based cost-effectiveness analysis. Value in Health. 2017;20(8):1100-9.
391. Joint Formulary Committee. British National Formulary (online) London: UK: BMJ Group and Pharmaceutical Press; 2018. Available from: http://www.medicinescomplete.com.
392. Hurst M. The Green Book: Central Government Guidance on Appraisal and Evaluation: HM Treasury, London, UK: OGL Press, 2018 (also available online on www. gov. uk/government/publications). Taylor & Francis; 2019.

393. Heijnsdijk EA, Wever EM, Auvinen A, Hugosson J, Ciatto S, Nelen V, et al. Quality-of-life effects of prostate-specific antigen screening. New England Journal of Medicine. 2012;367(7):595-605.
394. Office for National Statistics. Twenty-year survival data for 2001-2003. Cancer survival rates, Long-term Breast Cancer Survival, England and Wales. London: Office for National Statistics; 2005.
395. Obradovic M, Mrhar A, Kos M. Cost-effectiveness analysis of HPV vaccination alongside cervical cancer screening programme in Slovenia. European Journal of Public Health. 2010;20(4):415-21.
396. Pope C, Mays N. Qualitative methods in health research. Qualitative research in health care. 2006;3:1-11.
397. Harper D, Thompson AR. Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners: John Wiley & Sons; 2011.
398. Braun V, Clarke V. Successful qualitative research: A practical guide for beginners: sage; 2013.
399. Green J, Thorogood N. Qualitative methods for health research: Sage; 2018.
400. Gooberman-Hill R. Qualitative approaches to understanding patient preferences. The patient. 2012;5(4):215-23.
401. Saini KS, Taylor C, Ramirez AJ, Palmieri C, Gunnarsson U, Schmoll HJ, et al. Role of the multidisciplinary team in breast cancer management: results from a large international survey involving 39 countries. Annals of oncology : official journal of the European Society for Medical Oncology. 2012;23(4):853-9.
402. Gagliardi AR, Dobrow MJ. Paucity of qualitative research in general medical and health services and policy research journals: analysis of publication rates. BMC Health Services Research. 2011;11(1):268.
403. Leydon GM, Boulton M, Moynihan C, Jones A, Mossman J, Boudioni M, et al. Cancer patients' information needs and information seeking behaviour: in depth interview study. Bmj. 2000;320(7239):909-13.
404. Coast J, Al-Janabi H, Sutton EJ, Horrocks SA, Vosper AJ, Swancutt DR, et al. Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. Health economics. 2012;21(6):730-41.

405. Joanna C. Qualitative methods for health economics / edited by Joanna Coast. London: Rowman & Littlefield Inc; 2017.
406. Coast J. The appropriate uses of qualitative methods in health economics. Health economics. 1999;8(4):345-53.
407. Baker R, Robinson A. Responses to standard gambles: are preferences 'well constructed'? Health economics. 2004;13(1):37-48.
408. Bronsard A, Geneau R, Shirima S, Courtright P, Mwende J. Why are children brought late for cataract surgery? Qualitative findings from Tanzania. Ophthalmic epidemiology. 2008;15(6):383-8.
409. Ormston R, Spencer L, Barnard M, Snape D. The foundations of qualitative research. Qualitative research practice: A guide for social science students and researchers. 2014;2:52-5.
410. Cleary M, Horsfall J, Hayter M. Data collection and sampling in qualitative research: does size matter? Journal of advanced nursing. 2014;70(3):473-5.
411. Marshall MN. Sampling for qualitative research. Family practice. 1996;13(6):522-6.
412. Vass C, Rigby D, Payne K. The role of qualitative research methods in discrete choice experiments: a systematic review and survey of authors. Medical Decision Making. 2017;37(3):298-313.
413. Al-Janabi H, Keeley T, Mitchell P, Coast J. Can capabilities be self-reported? A think aloud study. Social Science & Medicine. 2013;87:116-22.
414. Canaway A, Al-Janabi H, Kinghorn P, Bailey C, Coast J. Development of a measure (ICECAP-Close Person Measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. Palliative medicine. 2017;31(1):53-62.
415. Keeley T, Al-Janabi H, Lorgelly P, Coast J. A qualitative assessment of the content validity of the ICECAP-A and EQ-5D-5L and their appropriateness for use in health research. PloS one. 2013;8(12):e85287.
416. van Osch SM, Stiggelbout AM. The construction of standard gamble utilities. Health economics. 2008;17(1):31-40.

417. Robinson S, Bryan S. Does the process of deliberation change individuals' health state valuations? An exploratory study using the person trade-off technique. Value in Health. 2013;16(5):806-13.
418. Kinghorn P, Robinson A, Smith RD. Developing a capability-based questionnaire for assessing well-being in patients with chronic pain. Social indicators research. 2015;120(3):897-916.
419. Husbands S, Jowett S, Barton P, Coast J. How qualitative methods can be used to inform model development. PharmacoEconomics. 2017;35(6):607-12.
420. Coyle A. Introduction to qualitative psychological research. In: Lyons E, Coyle A, editors. In Analysing qualitative data in psychology. CA, USA: Sage Publications; 2007. p. 9-29.
421. Corbin JM, Strauss A. Grounded theory research: Procedures, canons, and evaluative criteria. Qualitative sociology. 1990;13(1):3-21.
422. Kennedy F, Harcourt D, Rumsey N. The shifting nature of women's experiences and perceptions of ductal carcinoma in situ. Journal of advanced nursing. 2012;68(4):856-67.
423. Ganz PA. Quality-of-life issues in patients with ductal carcinoma in situ. Journal of the National Cancer Institute Monographs. 2010;2010(41):218-22.
424. Kim C, Liang L, Wright FC, Hong NJL, Groot G, Helyer L, et al. Interventions are needed to support patient-provider decision-making for DCIS: a scoping review. Breast cancer research and treatment. 2018;168(3):579-92.
425. Waller J, Douglas E, Whitaker KL, Wardle J. Women's responses to information about overdiagnosis in the UK breast cancer screening programme: a qualitative study. BMJ Open. 2013;3(4).
426. Crotty M. The foundations of social research: Meaning and perspective in the research process. London: Sage Publications; 1998.
427. Lincoln YS, Denzin NK. The handbook of qualitative research. London: Sage Publications; 2000.
428. Golafshani N. Understanding reliability and validity in qualitative research. The qualitative report. 2003;8(4):597-606.
429. Lewis J, Ritchie J. Generalising from qualitative research. Qualitative research practice: A guide for social science students and researchers. 2003;2:347-62.

430. Reinharz S. Neglected voices and excessive demands in feminist research. Qualitative sociology. 1993;16(1):69-76.
431. Creswell JW, Hanson WE, Clark Plano VL, Morales A. Qualitative research designs: Selection and implementation. The counseling psychologist. 2007;35(2):236-64.
432. Firestone WA. Alternative arguments for generalizing from data as applied to qualitative research. Educational researcher. 1993;22(4):16-23.
433. Palinkas LA, Garcia AR, Aarons GA, Finno-Velasquez M, Holloway IW, Mackie TI, et al. Measuring use of research evidence: the structured interview for evidence use. Research on social work practice. 2016;26(5):550-64.
434. Rubin HJ, Rubin IS. Qualitative interviewing: The art of hearing data. London: Sage Publications; 2011.
435. Novick G. Is there a bias against telephone interviews in qualitative research? Res Nurs Health. 2008;31(4):391-8.
436. Kvale S, Brinkmann S. Interviews: Learning the craft of qualitative research interviewing. London: Sage Publications; 2009.
437. Mays N, Pope C. Rigour and qualitative research. BMJ. 1995;311(6997):109-12.
438. Glaser BG, Anselm L. Strauss, and Elizabeth Strutzel. The discovery of grounded theory; strategies for qualitative research. Nursing research. 1968;17(4):264.
439. Bowen D, Robbins R, Bush N, Meischke H, Ludwig A, Wooldridge J. Effects of a web-based intervention on women's breast health behaviors. Translational behavioral medicine [Internet]. 2017; 7(2):[309-19 pp.]. Available from: http://cochranelibrary-wiley.com/o/cochrane/clcentral/articles/718/CN-01403718/frame.html.
440. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC medical research methodology. 2013;13(1):117.
441. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. The qualitative researcher's companion. 2002;573(2002):305-29.
442. Srivastava A, Thomson SB. Framework analysis: a qualitative methodology for applied policy research. Journal of Administration and Governance. 2009;4(2):72-9.

443. Ziebland S MA. Making sense of qualitative data analysis: an introduction with illustrations from DIPEx (personal experiences of health and illness). Medical education. 2006;40(5):405-14.
444. Pope C, Ziebland S, Mays N. Qualitative research in health care: analysing qualitative data. BMJ. 2000;320(7227):114.
445. Tong A SP, Craig J.,. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. . International journal for quality in health care. 2007;19(6):349-57.
446. O'Brien BC HI, Beckman TJ, Reed DA, Cook DA.,. Standards for reporting qualitative research: a synthesis of recommendations. Academic Medicine. 2014;89(9):1245-51.
447. Prinjha S, Evans J, Ziebland S, McPherson A. 'A mastectomy for something that wasn't even truly invasive cancer.'Women's understandings of having a mastectomy for screen-detected DCIS: a qualitative study. Journal of medical screening. 2011;18(1):34-40.
448. De Morgan S RS, D'Este C, Rogers K.,. Knowledge, satisfaction with information, decisional conflict and psychological morbidity amongst women diagnosed with ductal carcinoma in situ (DCIS). . Patient education and counseling. 2011;84(1):62-8.
449. De Morgan S, Redman S, White KJ, Cakir B, Boyages J. 'Well, have I got cancer or haven't I?'The psycho-social issues for women diagnosed with ductal carcinoma in situ. Health Expectations. 2002;5(4):310-8.
450. Loberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. Breast cancer research. 2015;17:63.
451. Hayes JH, Ollendorf DA, Barry MJ, Pearson SD, McMahon PM. Therapeutic options for low-risk prostate cancer: A cost-effectiveness analysis. Journal of Clinical Oncology. 2010;28(15 SUPPL. 1).
452. Ruddy KJ, Meyer ME, Giobbie-Hurder A, Emmons KM, Weeks JC, Winer EP, et al. Long-term risk perceptions of women with ductal carcinoma in situ. The oncologist. 2013;18(4):362-8.

453. Department of Health and Social Care. PHE priorities in health and social care: 2019 to 2020. In: Public Health England, editor. London, UK: Department of Health and Social Care; 2019.
454. Loeb S, Curnyn C, Walter D, Fagerlin A, Siebert U, Muhlberger N, et al. Health state utilities among contemporary prostate cancer patients on active surveillance. Translational andrology and urology. 2018;7(2):197-202.
455. Kattan MW, Cowen ME, Miles BJ. A decision analysis for treatment of clinically localized prostate cancer. Journal of general internal medicine. 1997;12(5):299-305.
456. 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.
457. Venkatesh S, Pasternak JD, Beninato T, Drake FT, Kluijfhout WP, Liu C, et al. Cost-effectiveness of active surveillance versus hemithyroidectomy for micropapillary thyroid cancer. Surgery. 2017;161(1):116-26.
458. Paci E, Warwick J, Falini P, Duffy S. Overdiagnosis in screening: is the increase in breast cancer incidence rates a cause for concern? Journal of Medical Screening. 2004;11(1):23-7.
459. Puliti D, Miccinesi G, Zappa M, Manneschi G, Crocetti E, Paci E. Balancing harms and benefits of service mammography screening programs: a cohort study. Breast Cancer Research. 2012;14(1):9.
460. Baker R, Robinson A. Responses to standard gambles: are preferences 'well constructed'? Health economics. 2004;13(1):37-48.
461. Schoemaker PJH. The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations. Journal of Economic Literature. 1982;20(2):529-63.
462. Keogh LA, McClaren BJ, Apicella C, Hopper JL. How do women at increased, but unexplained, familial risk of breast cancer perceive and manage their risk? A qualitative interview study. Hereditary cancer in clinical practice. 2011;9(1):7.
463. 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. New England Journal of Medicine. 2016;375(15):1415-24.

464. Public Health England. UK NSC: Evidence review process. In: Department of Health and Social Care, editor. London, England 2017.
465. Roberts T, Barton P, Auguste P, Middleton L, Furmston A, Ewer A. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a cost-effectiveness analysis. Archives of disease in childhood. 2012;97(3):221-6.
466. Ewer A, Furmston A, Middleton L, Deeks J, Daniels J, Pattison H, et al. Pulse oximetry as a screening test for congenital heart defects in newborn infants: a test accuracy study with evaluation of acceptability and cost-effectiveness. NIHR Health Technology Assessment programme: Executive Summaries: NIHR Journals Library; 2012.
467. Sweeting MJ, Masconi KL, Jones E, Ulug P, Glover MJ, Michaels JA, et al. Analysis of clinical benefit, harms, and cost-effectiveness of screening women for abdominal aortic aneurysm. Lancet. 2018;392(10146):487-95.
468. Manchanda R, Patel S, Gordeev VS, Antoniou AC, Smith S, Lee A, et al. Cost-effectiveness of Population-Based BRCA1, BRCA2, RAD51C, RAD51D, BRIP1, PALB2 Mutation Testing in Unselected General Population Women. JNCI: Journal of the National Cancer Institute. 2018;110(7):714-25.