Optimising follow-up for women after primary treatment for early breast cancer

Dr Taryn Bessen
MBBS, FRANZCR

School of Population Health
Faculty of Health Sciences
The University of Adelaide
Australia

A thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy
October 2013

……

If we knew what it was we were doing,

it would not be called research, would it?

Albert Einstein

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to gratefully acknowledge and thank my supervisors for their valued commitment to my PhD. To the late Konrad Jamrozik, thank you for setting me on my path. To Jon Karnon, the guru of modelling, without your incredible expertise this PhD would not have been possible. Thank you for being so generous with your time, for your instruction and your patience. To Dorothy Keefe, thank you being my clinical ally, for your pragmatism, and for your counsel. To Jackie Street, thank you for agreeing to become my supervisor late in the piece, your intellectual engagement, and for teaching me about the world of academia. I would also like to thank the National Breast Cancer Foundation (NBCF) for providing Doctoral Funding and recruitment support, which enabled this research to be both contemplated and conducted.

I have an incredible family to whom I am truly indebted. As a cancer survivor and a clinician, my husband valued what I aspired to do, and my hopes to help improve follow-up services for cancer patients. Sepehr gifted me the opportunity to reduce my clinical work as a Radiologist to embark on a PhD. He shared the trials and tribulations, kept me grounded, and carried the load while I was absorbed in my studies. My young children Alexander and Isabella have been exceptionally patient while mummy has been “writing her story”, and their hugs and jokes have kept me going in difficult times. I am truly blessed.

ABSTRACT

Title

Optimising follow-up for women after primary treatment for early breast cancer

Overview

Due to early diagnosis and improved treatment outcomes, there is a growing pool of breast cancer survivors who will require follow-up during their lifetime. International guidelines currently recommend routine annual mammography, but there is no randomised controlled trial evidence to support this frequency over any other. In addition, there are economic and workforce imperatives around the provision of cancer follow-up care. The current workload growth is unsustainable for breast cancer specialists who also provide care for women newly diagnosed or with a recurrence. If new models of care are to be developed, it is important that these are appropriate and acceptable, yet currently we know little about patient preferences for possible alternative modes of delivery of follow-up services.

Research Questions

1. What is the impact on survival of the method and timing of detection of a second breast cancer event within the breast?
2. Using a model based economic analysis, is it efficient to tailor mammographic follow-up according to risk of recurrence?
3. What do Australian breast cancer survivors prefer with respect to the provider, location, frequency, and method of delivery of routine follow-up care in years 3, 4 and 5 following diagnosis if existing specialist services were not available; and what is the perceived value of offering “drop-in” clinics providing additional support?

Structure of Thesis

Chapter 2: Systematic Review

The aim of this paper was to review the evidence around the effectiveness of mammographic follow-up. The review identified a complete absence of randomised clinical trials (RCTs) in this area. I conclude that we should embrace alternative research techniques, such as decision analytic modelling, to guide our practice in the likely continued absence of randomised controlled studies in this field.

Chapter 3: A patient-level calibration framework for evaluating surveillance strategies: a case-study of mammographic follow-up after early breast cancer

This paper describes the development, calibration, and cost-effectiveness analyses of an early breast cancer surveillance discrete event simulation (DES) model. The DES model was used to analyse three alternative mammographic follow-up schedules for postmenopausal women who were disease free following primary treatment for moderate prognosis early breast cancer; taking into account age and adherence to mammography. This study demonstrates the potential value of combining linked, retrospective data and decision analytic modelling to provide estimates of costs and health outcomes that are sufficiently robust to inform cancer clinical guidelines and individual patient decisions regarding appropriate follow-up schedules.

Chapter 4: One size does not fit all? Cost utility analyses of alternative mammographic follow-up schedules, by risk of recurrence

The aim of this paper is to report the full set of cost-effectiveness results from the model described in chapter 3, comparing alternative mammographic follow-up schedules for postmenopausal women with excellent, good, moderate and poor prognosis early breast cancer. Our results suggest that annual mammographic follow-up is not cost effective for most postmenopausal women, and that mammographic follow-up can be tailored based on the Nottingham Prognostic Index score of the primary breast cancer and age at diagnosis.

Chapter 5: Issues with data access and quality

This chapter describes the large gulf between the ideal dataset to inform the cost-effectiveness model and what is currently available in South Australia (and beyond).

Chapter 6: Discrete Choice Experiment (DCE)

The aim of this study was to explore the preferences of Australian breast cancer survivors for alternative modes of delivery of follow-up services if we could no longer offer long term specialist-led hospital based follow-up. This study provides important insights into what attributes of a breast cancer follow-up service women value most.

Chapter 7: Conclusion

This chapter summarises and synthesises the research findings, discusses the limitations of this research, describes the advances in the field during the period of this work, and suggests directions for future research and recommendations for policy.

THESIS DECLARATION

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

.....

Dr Taryn Bessen

31st October 2013

GLOSSARY OF ACRONYMS

AT	Adjuvant therapy
BCD	Breast cancer death
BCS	Breast-conserving surgery
BDM	Registry of Births, Deaths and Marriages
CE	Clinical examination
CI	Confidence Interval
CLBC	Contralateral breast cancer
CT	Chemotherapy
CXR	Chest x-ray
DCE	Discrete choice experiment
DCIS	Ductal carcinoma in-situ
DES	Discrete event simulation
DF	Disease-free
DM	Distant metastases
ER+	Estrogen receptor positive
ER-	Estrogen receptor negative
HER2+	HER2 receptor positive
HER2-	HER2 receptor negative
HR	Hazard rate, hazard ratio
IBTR	Ipsilateral breast tumor recurrence
ICER	Incremental cost effectiveness ratio
ILR	Impalpable local recurrence
ISAAC	Integrated South Australian Activity Collection
LABC	Locally advanced breast cancer
LCIS	Lobular carcinoma in-situ
LN	Lymph nodes
LR	Local recurrence
M	Metastases
MMG	Mammogram
MRI	Magnetic resonance imaging
MRN	Medical record number
Mx	Mastectomy
NPI	Nottingham Prognostic Index
OACIS	Open Architecture Clinical Information System
OCD	Other cause death
OR	Odds ratio
PBC	Primary breast cancer
PLR	Palpable local recurrence
PR+	Progesterone receptor positive
PR-	Progesterone receptor negative
QALY	Quality adjusted life year
RCT	Randomised controlled trial
RILR	Removed impalpable local recurrence
RT	Radiotherapy
SA	South Australia
SACR	South Australian Cancer Registry
T	Tumor (size)
URN	Unit record number
US	Ultrasound

LIST OF TABLES

Chapter 4 extension

Chapter 5

Chapter 6

Chapter 1

INTRODUCTION

1. OVERVIEW

1.1 Breast cancer in Australia

Clinical cancer services around Australia are facing a rapid expansion in workload through a combination of three factors: (a) the incidence of cancer is expected to increase with age,[1] and with unprecedented ageing of the population, there will be growing numbers in the most cancer-prone age groups; (b) increasing incidence of cancer that can only be partly explained by the ageing and increasing size of the population;[1] and (c) falling age-standardised mortality rate for all cancer combined,[1] expanding the pool of survivors. In Australia, the term "cancer survivor" is generally used to refer to people who are disease-free following completion of primary treatment for cancer.[2]

In October 2012, the Australian Institute of Health and Welfare & Cancer Australia released the most up-to-date statistical information on breast cancer in Australia. Breast cancer is the most common cancer diagnosed in Australian women, with 13,567 new cases (28% of all cancers in females) in 2008, and the second most common cancer death with 2,680 breast cancer deaths (15.5% of all cancers in females) in 2007.[3] It is expected that the incidence rate will increase moderately, with 17,210 new cases per year estimated for 2020.[3] However the breast cancer death rate continues to decline, with 2,730 women estimated to die from breast cancer in 2020.[3]

Survival from breast cancer has increased over the last 27 years for which national data is available. Between the periods 1982-1987 and 2006-2010, the 5 year relative survival from breast cancer increased from 72% to 89%.[3] While survival gains were seen for all age groups, the largest increase in survival was for those aged between 50-69 years, the target age group for Breast Screen Australia.[3] At the end of 2008, more than 57,300 women who had been diagnosed with breast cancer within the previous five years, were still alive.[3] The highest five year prevalence, as a proportion of the female population was in South Australia (59 per 100,000).[3]

Given the high incidence and declining mortality, it is evident that the number of Australian breast cancer survivors will continue to increase with time.

1.2 Breast Cancer

Breast cancer is a primary cancer that originates in breast tissue. The breast is a specialised sweat gland,that lies in the subcutaneous tissue of the anterior chest wall, extends toward the axilla and sits anterior to the pectoralis major muscle. It is present in both sexes, but in females the breast develops the capacity for milk production. The breast is divided into 15-20 lobes, radially arranged around the nipple. The lobes are comprised of smaller lobules, and separated by fat and supporting connective tissue. Each of the lobes has its own lactiferous duct which opens onto the nipple.[4] Blood supply is from branches of the internal thoracic and axillary arteries. Venous drainage is mainly by deep veins that run with these arteries. The breast has an extremely rich lymphatic supply arising both from the deep tissue of the breast (connective tissue and lactiferous ducts) and the overlying skin. Lymph from the lateral breast (upper outer quadrant, lower outer quadrant) drains laterally to axillary and infraclavicular nodes. Lymph from the medial breast (upper inner quadrant, lower inner quadrant) drains medially to the internal thoracic (parasternal lymph nodes).[5]

Breast cancer is not a single disease. Breast cancer most commonly arises from the ducts (ductal carcinoma, 77.6% of all breast cancer in 2008) or lobules (lobular carcinoma, 10.7% of all breast cancer in 2008). Less common forms of breast cancer include medullary carcinoma and atypical medullary carcinoma (0.5%); tubular carcinoma and invasive cribiform carcinoma (1.6%); mucinous carcinoma (2%), invasive papillary carcinoma (0.6%), inflammatory carcinoma (0.1%), other – specified (2.2%), and unspecified (4.6%). Each type of breast cancer is further subdivided; with the most common breast cancer, invasive ductal carcinoma, consisting of seven subtypes.[3]

Breast cancer can be either non-invasive and confined to the ducts or lobules (ductal carcinoma in situ, DCIS; or lobular carcinoma in situ, LCIS) or invasive and involve adjacent breast parenchyma (invasive ductal carcinoma or invasive lobular carcinoma).

Females diagnosed with DCIS or LCIS have an increased risk of developing invasive breast cancer.[3] Invasive breast cancers can spread to regional lymph nodes (e.g. axillary and internal mammary lymph nodes); and to distant sites. Breast cancer has a predilection for metastasizing to lung, liver, bone and brain. Non-invasive cancers are usually cured by surgical resection; invasive carcinoma confined to the breast ± axillary lymph nodes is potentially curable; distant metastatic disease is usually not.

Breast cancer can be diagnosed by mammography, whilst impalpable and asymptomatic. For women with breast symptoms or breast change, the triple test is recommended.[6] This consists of clinical breast examination, imaging tests (mammography and/or ultrasound) and biopsy (fine needle or core biopsy most commonly). If the results are positive, the woman is diagnosed with breast cancer. Surgery is then generally performed, to remove part (e.g. wide local excision) or all of the breast (mastectomy), with removal of one or more lymph nodes from the axilla. If the tumor is large, neo-adjuvant chemotherapy may be given to reduce tumor volume prior to surgery.

Once resected, a pathologist will examine the specimen and describe features of the primary breast cancer that are used to inform prognosis and to guide treatment planning. Important microscopic features include, but are not limited to:

- Histologic type: ductal, lobular, medullary etc. Some subtypes are comparatively indolent (e.g. tubular carcinomas), while other subtypes are more aggressive (e.g. inflammatory carcinoma)
- Grade: this refers to the degree of differentiation of the breast tumor (that is, how closely it resembles normal breast tissue), and the rate of growth.[7] Grade 1 tumors are well differentiated, are usually slow growing and likely to behave less aggressively. Grade 2 tumors are moderately differentiated and grow faster than normal cells. Grade 3 tumors are poorly differentiated, the cells grow more quickly and are more likely to be aggressive.[7]

- Stage: this refers to the extent of tumor spread, and includes a description of tumor size, lymph node involvement and distant metastases. Stages 0 to 4 reflect increasing tumor burden, and worsening prognosis (see below).
- Surgical margin: this refers to the completeness of surgical excision. The surgeon will remove the breast cancer with a rim of surrounding normal tissue, with the aim of ensuring complete resection of the cancer (negative margins).[7] If, when viewed under the microscope, the edge of normal tissue around the tumor contains cancer cells (positive margins), further surgery is often required.[7]
- Lymph nodes: the number of axillary lymph nodes which contain metastases is the single most powerful prognostic variable.[8] The larger the number of positive lymph nodes; fixed or matted lymph nodes and involvement of regional lymph node groups outside of the axilla (i.e. internal mammary, infra/supraclavicular lymph node groups), are all associated with worsening prognosis.
- Hormone receptors: the presence or absence of the estrogen (ER) and progesterone (PR) receptors will reflect whether tumor growth is affected by these hormones, and will determine whether endocrine therapy is recommended as part of adjuvant therapy.[7] Tumors may be either positive or negative for the receptors, creating four different receptor combinations (ER+PR+, ER+PR-, ER-PR+, ER-PR-). ERPR negativity is one high risk marker for lymph node negative tumors.[8]
- HER2 receptor status: overexpression of HER2 receptors will lead to abnormal cell division. Cancers with a high expression of the HER2 receptor (HER2 positive) are more aggressive than HER2 negative breast cancers, and have a poorer prognosis. Women with HER2 positive cancers are offered tailored treatment e.g. Herceptin.[7]

The stage of the tumor reflects the extent of spread. The most widely used classification for breast cancer is the TNM classification.[8] This refers to Tumor, Nodes and Metastases respectively. Increasing tumor size, involvement of lymph nodes and distant metastatic disease are associated with worsening prognosis. Clinical staging determines how much cancer there is based on the physical examination, imaging tests, and biopsies of the affected area.[9] Pathologic staging can only be determined postoperatively, and combines results of both the clinical staging with surgical results.[9] Once T,N and M are

determined, they are combined, and an overall stage is assigned.[9] A simplified version of the most recent TNM breast cancer classification is presented in table 1, and the stage groupings are presented in table 2.[10] Stage 1 cancers are the least advanced and often have a better prognosis. Higher stage cancers are more advanced and often have a worse prognosis.

Table 1: TNM classification of breast cancer[10]

T	Primary Tumor
Tx	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Carcinoma in situ
	Tis (DCIS) – Ductal carcinoma in situ
	Tis (LCIS) – Lobular carcinoma in situ
	Tis (Paget's) – Paget's disease of the nipple NOT associated with invasive carcinoma and/or carcinoma in situ in the underlying breast parenchyma
T1	Tumor ≤ 2cm in greatest dimension
	T1mi - ≤ 1mm
	T1a - >0.1cm but ≤ 0.5cm
	T1b - >0.5 but ≤1cm
	T1c - >1cm but ≤2cm
T2	Tumor >2cm but ≤5 cm in greatest dimension
T3	Tumor >5cm in greatest dimension
T4	Tumor of any size with direct extension to chest wall and/or skin(ulceration or skin nodules)
	T4a – Extension to the chest wall, not including pectoralis muscle adherence/invasion
	T4b – Ulceration and/or ipsilateral satellite nodules and/or edema (including peau d'orange) of the skin, which does not meet the criteria for inflammatory carcinoma
	T4c – Both T4a and T4b
	T4d – Inflammatory carcinoma

N	Node categories (clinical)
Nx	Regional lymph nodes cannot be assessed (e.g. previously removed)
N0	No regional lymph node metastasis
N1	Metastasis to movable ipsilateral level I,II axillary lymph node/s
N2	Metastasis in ipsilateral level I,II axillary lymph nodes that are clinically fixed or matted; or in clinically detected ipsilateral internal mammary nodes in the absence of clinically evident axillary lymph node metastases
	N2a – metastases in ipsilateral level I,II axillary lymph nodes fixed to one another (matter) or to other structures
	N2b – metastases only in clinically detected ipsilateral internal mammary nodes and in the absence of clinical evident level I,II axillary lymph node metastases
N3	Metastases in ispilateral infraclavicular (level III axillary) lymph nodes/s with or without level I,II axillary lymph node involvement; or in clinically detected ipsilateral internal mammoary lymph nodes/ with clinically evidnect level I,II axillary lymph node metastases; or metastases in ispialteral supraclavicular lymph hode/s with or without axillary or internal mammary lymph node involvement
	N3a – Metastases in ipsilateral infraclavicular lymph node/s
	N3b – Metastases in ipsilateral internal mammary lymph node/s and axillary lymph node/s
	N3c – Metastases in ipsilateral supraclavicular lymph node/s
N	**Node categories (pathologic)**
pNX	Regional lylmph nodes cannot be assessed (e.g. previously removed, or not removed for pathologic study)
pN0	No regional lymph node metastasis identified histologically
	pN0(i-)- No regional lymph node metastases histologically, negative IHC
	pN0(i+)- Malignant cells in regional lymph node/s no greater than 0.2mm (detected by H&E or IHC including ITC)
	pN0(mol-) – No regional lymph node metastases histologically, negative molecular findings (RT-PCR)
	pN0 (mol+) – Positive molecular findings (RT-PCR), but no regional lymph node metastases detected by histology or IHC)
pN1	Micrometastases; or metastases in 1-3 axillary lymph nodes; and/or internal mammary nodes with metastases detected by sentinel lymph node biopsy but not clinically detected
	pN1mi- Micrometastases (greater than 0.2mm and/or more than 200 cells, but none greater than 2.0mm)
	pN1a- Metastases in 1-3 axillary lymph nodes, at least one metastasis greater than 2.0mm
	pN1b – Metastases in internal mammary nodes with micrometastases or macrosmetastases detected by sentinel lymph node biopsy but not clinically detected

	pN1c – Metastases in 1-3 axillary lymph nodes and in internal mammary lymph nodes with microcmetastases or macrometastases detected by sentinel lymph node biopsy but not clinically detected
pN2	Metastases in 4-9 axillary lymph nodes; or in clinically detected internal mammary lymph nodes in the absence of axillary lymph node metastases
	pN2a – Metastases in 4-9 axillary lymph nodes (at least one tumor deposit greater than 2.0mm)
	pN2b – Metastases in clinically detected internal mammary lymph nodes in the absence of axillary lymph node metastases
pN3	Metastases in 10 or more axillary lymph nodes; or in infraclavicular (level III axillary) lymph nodes; or in clinically detected ipsilateral internal mammary lymph nodes in the presence of one or more positive level I, II axillary lymph nodes; or in more than three axillary lymph nodes and in internal mammary lymph nodes with micrometastases or macrometastases detected by sentinel lymph node biopsy but not clinically detected; or in ipsilateral supraclavicular lymph nodes
	pN3a – Metastases in 10 or more axillary lymph nodes (at least one tumor deposit greater than 2.0mm); or metastases to the infraclavicular (level III axillary lymph) nodes
	pN3b – Metastases in clinically detected ipsilateral internal mammary lymph nodes in the presence of one or more positive axillary lymph nodes; or in more than three axillary lymph nodes and in internal mammary lymph nodes with micrometastases or macrometastases detected by sentinel lymph node biopsy but not clinically detected
	pN3c – Metastases in ipsilateral supraclavicular lymph nodes
M	**Metastasis categories**
M0	No clinical or radiographic evidence of distant metastases
cM0(i+)	No clinical or radiographic evidence of distant metastases, but deposits of molecularly or microscopically detected tumor cells in circulating blood, bone marrow, or other nonregional nodal tissue that are no larger than 0.2mm in a patient within symptoms or signs of metastases
M1	Distant detectable metastases as determine by classic clinical and radiographic means and/or histologically proven larger than 0.2mm.

Table 2: Anatomic stage/prognostic groups[10]

Stage	T	N	M
Stage 0	Tis	N0	M0
Stage IA	T1	N0	M0
Stage IB	T0	N1mi	M0
	T1	N1mi	M0
Stage IIA	T0	N1	M0
	T1	N1	M0
	T2	N0	M0
Stage IIB	T2	N1	M0
	T3	N0	M0
Stage IIIA	T0	N2	M0
	T1	N2	M0
	T2	N2	M0
	T3	N1	M0
	T3	N2	M0
Stage IIIB	T4	N0	M0
	T4	N1	M0
	T4	N2	M0
Stage IIIC	Any T	N3	M0
Stage IV	Any T	Any N	M1

T= tumor, N= nodes, M= metastasis

Following surgery, women usually receive local radiotherapy and systemic adjuvant therapy, with the aim of treating undetectable disease and reducing the risk of local and distant relapse, thereby improving survival. Systemic treatment includes chemotherapy, hormonal therapy and/or targeted therapy. Choice of agent will depend on the age, general health and preferences of the patient, the particular pathological features of the

tumor, the stage of the tumor at diagnosis, and whether the cancer has a lower or higher risk of recurrence.[7] That is, while treatment for breast cancer could be considered fairly standardised, the specific treatment for any individual women is tailored to her prognosis and risk of recurrence.

Cancer Australia has separate clinical practice guidelines for management of early breast cancer and advanced breast cancer. Early breast cancer is defined as tumours of not more than five cm in diameter, with either impalpable or palpable but not fixed lymph nodes and with no evidence of distant metastases.[8] Advanced breast cancer includes locally advanced (has one or more of: peau d'orange, skin ulceration, or fixation to the underlying intercostal muscles, serratus anterior muscles or bones of the chest wall; or inflammatory breast cancer) and metastatic breast cancer (cancer that has spread to distant sites).[11] The goals of treatment are different for both: for early breast cancer the most effective treatment is offered with intent to cure,[8] for advanced breast cancer the primary goals are to improve the length and quality of life.[11]

All women will require follow-up after completion of their primary treatment. The prevalence of early and advanced breast cancer survivors is not available in Australia, as currently there is no recording of stage data in the National Cancer Registry. However, for women with early breast cancer, follow-up is likely to be for a much longer duration as many women with early breast cancer ultimately survive their disease. It is for this reason that the focus of this PhD is on early breast cancer.

1.3 Breast cancer follow-up

Cancer Australia has produced clear guidelines for follow-up of women with early breast cancer,[12] yet they acknowledge there is limited high level evidence on follow-up care, particularly with respect to follow-up procedures, and optimal frequency or duration of follow-up.[12] Cancer Australia has defined the aims and objectives of follow-up: to detect and treat local recurrence; to deal with adverse effects of treatment; to provide psychological support; to screen for a new primary breast cancer; review and update family history; observe outcomes of therapy, and review treatment including the potential for new therapies.[12]

Conceptually, follow-up for early breast cancer can be considered as having two components: mammography and the clinic visit. Follow-up mammography is used to detect local recurrence and to screen for a new primary breast cancer. The follow-up visit provides the opportunity for history taking regarding general wellbeing, treatment related side-effects or problems associated with living with breast cancer; and a clinical breast examination to assess for any new changes in the breast.

Follow-up mammography

The aim of follow-up mammography is to detect new disease in the treated or opposite breast at an early stage when it is potentially curable.[13] Women with a past history of breast cancer have an overall risk of local recurrence in the treated breast of 0.5-1% per annum (including new primaries, and is lifelong), [14] while the risk of developing a cancer in the opposite breast is estimated to be just under 0.03% per annum.[14]

Australian guideline recommendations for follow-up after early breast cancer state that breast cancer recurrences that are detected by mammography whilst impalpable tend to have a better prognosis than those that are detected when larger and palpable.[12] While more frequent surveillance mammography will detect more impalpable recurrent cancers, an optimal mammographic strategy will balance the financial costs, anxiety associated with increased frequency of mammography and/or false positive test results, and potential identification of cancers at minimal risk of further progression; with the benefits of detecting more impalpable progressive local recurrence.

However, not all women have the same risk of in-breast relapse, with rates per 1000 woman-years varying by patient, tumour and treatment characteristics at initial diagnosis.[15] Buist et al[15] (table 3) showed that recurrence rates decreased with increasing age of woman at diagnosis, with rates ranging from 9.7 per 1000 woman-years for women aged 18-39 years, to 4.6 for women aged 70-79 years.[15] Rates of second primaries were fairly similar across the same age spectrum (5.5-7), except for women aged ≥ 80 years whose rate was considerably lower (4.2).[15] Increasing stage at diagnosis of first invasive cancer was associated with increasing rates of both breast cancer recurrence and second primary breast cancers (7.1 and 7.8 respectively).[15] Tumors that were negative for estrogen receptors but positive for progesterone receptors (ER-PR+) had the highest rate of breast cancer recurrence, but lowest rate of second breast primaries (10.8 and 2.7 respectively).[15] Tumors that were negative for

both estrogen and progesterone receptors (ER-PR-) had not only a high breast cancer recurrence rate (9.0) but also the highest rate of second breast primaries (6.4).[15] Breast cancer recurrence rates were highest in women whose first cancer was treated with breast conserving surgery without radiation therapy (12.7), but the rate of second breast primaries did not differ by type of treatment (5.5-6.3) regardless of whether the woman was treated by breast conserving surgery with or without radiation therapy or treated with mastectomy.[15]

Table 3: Rates (per 1,000 woman years) of breast cancer recurrences and second primaries by characteristics at initial diagnosis, adapted from Buist et al[15]

Patient, tumor and treatment characteristics at initial diagnosis	Breast cancer recurrences			Second breast primaries		
	n	Rates	95% CI	n	Rates	95% CI
Age at 1st diagnosis						
18-39 years	25	9.7	(6.3,14.4)	17	6.6	(3.9,10.6)
40-49 years	87	7.4	(5.9,9.1)	71	6.0	(4.7,7.6)
50-59 years	78	5.0	(3.9,6.2)	86	5.5	(4.4,6.8)
60-69 years	57	4.3	(3.3,5.6)	75	5.7	(4.5,7.1)
70-79 years	50	4.6	(3.4,6.0)	77	7.0	(5.5,8.8)
≥80 years	17	4.0	(2.3,6.4)	18	4.2	(2.5,6.7)
Stage at 1st diagnosis						
0	92	8.5	(6.8,10.4)	80	7.4	(5.8,9.2)
I	110	3.9	(3.2,4.7)	167	5.9	(5.0,6.8)
IIA	68	5.3	(4.1,6.7)	49	3.8	(2.8,5.0)
IIB	44	7.1	(5.2,9.5)	48	7.8	(5.7,10.3)
Hormone receptor						
ER-PR-	60	9.0	(6.9,11.6)	43	6.4	(4.7,8.7)
ER-PR+	8	10.8	(4.7,21.3)	2	2.7	(0.3,9.8)
ER+PR-	24	5.5	(3.5,8.2)	16	3.7	(2.1,6.0)
ER+PR+	109	3.6	(2.9,4.3)	161	5.3	(4.5,6.2)
Treatment						
BCS without radiotherapy	120	12.7	(10.6,15.2)	59	6.3	(4.8,8.1)
BCS with radiotherapy	118	4.2	(3.5,5.1)	155	5.5	(4.7,6.5)
BCS with & without radiotherapy	238	6.4	(5.6,7.2)	214	5.7	(5.0,6.5)
Mastectomy	76	3.6	(2.8,4.5)	130	6.2	(5.2,7.3)

ER= estrogen receptor, PR = progesterone receptor, BCS= breast-conserving surgery

It is likely therefore that the costs and benefits of follow-up mammography may differ between different subsets of patients. It is possible that by defining different risk profiles, we could tailor mammographic schedules that are more efficient and better balance risks and benefits from a patient perspective.

Follow-up clinic visit

Follow-up clinic visits provide an opportunity to offer psychosocial support, manage treatment related side-effects, and provide care for patients who develop metastases.[16]

Due to early diagnosis and improved treatment outcomes, there is a growing pool of breast cancer survivors who will require follow-up during their lifetime. The current workload growth is unsustainable for breast cancer specialists who also provide care for women newly diagnosed or with a recurrence. In the face of rising absolute in-flow of cancer patients, one of the few strategies available to oncology units to relieve the pressures they face, is to rationalise follow-up of survivors beyond primary treatment in terms of numbers of contacts with clinical specialists.

Relevant research suggests that nurses may be better at identifying psychological concerns and side-effects of drug treatment than clinicians,[17] and shared cared models of breast cancer follow-up between cancer specialists and General Practitioners (GP) have recently been the focus of a demonstration project within Australia by Cancer Australia.[18]

Patient preferences for follow-up

Given limited resources, if we are considering designing new follow-up services that are both appropriate and acceptable to women, we need to understand which attributes of the service women value most. Brennan et al, recently published a body of qualitative research on breast cancer follow-up in Australia survivors: Australian women reported a high level of satisfaction with current specialist based care and an initial reluctance to consider models of care that would involve them moving away from the cancer specialist.[19] While recognising advantages to GP follow-up, there was a much stronger

level of support for shared care between the GP and a specialist, rather than a complete transfer to GP led care.[19] There was also a perceived need for additional training of GPs or breast cancer nurses if they were to have an increased role in follow-up care.[20] Patients highlighted psychosocial needs and menopausal symptoms as areas of the follow-up consultation that needed improving.[20]

Limitations in the evidence base

International guidelines report no randomised controlled trial (RCT) evidence that supports their recommendations of annual compared to other frequencies of follow-up mammography, and there is currently no tailoring of mammographic frequency according to risk. [12-14,16,21-23] All acknowledge the paucity of definitive evidence available on follow-up care after breast cancer, with much of the information from low level observational studies.[12] The main limitation of observational studies is that group allocation is not random, reducing the likelihood that the groups are alike in all relevant aspects other than the intervention of interest (i.e. follow-up). Whilst statistical analyses can be undertaken to adjust for observed confounders, there remains the possibility of unobserved confounding, which can create biased results. Observational studies of screening interventions are also subject to other potential biases, such as lead time bias and length bias (see chapter 2 for detail) that limit the reliability of their results. Thus, observational studies do not provide sufficient evidence on which to make informed decisions regarding the development of personalised surveillance strategies.

RCTs can provide an unconfounded estimate of effect, and are a common and relatively easy source for effectiveness and cost data; but there are significant potential barriers to performing RCTs in follow-up for cancer. These include difficulty with patient accrual[24] (for example, if patients have a 50% chance of receiving follow-up that is less frequent than current guideline recommendations they are unlikely to consent to be in a trial), and the large sample size and long follow-up (and thus cost implications) required to demonstrate a significant difference between alternative follow-up programs for different patient subgroups.[24] While patient level data on both outcomes and cost can

be collected alongside a clinical trial, there are disadvantages to relying on cost data from RCTs.[25] For example, the choice of comparator (and hence costing) in an RCT may not be applicable to a given local policy context;[25] the resource use in an RCT designed to determine efficacy of an intervention may be more intensive than in everyday clinical practice;[25] duration of follow-up and sample size may be inadequate if the patient is censored upon experiencing an "event of interest" especially if cost-effectiveness is determined by the patient's prognosis after the event (e.g. detection of a local recurrence or contralateral new primary breast cancer);[25] and the impact of an intervention on mortality and morbidity may not be demonstrated within the time frame of an RCT making decisions about resource allocation difficult.[25]

So while we clearly need better evidence for the value of follow-up mammography [26] and "we might in principle prefer trial based economic evaluations",[27] given that RCTs are unlikely to occur, how can we guide clinical practice in the meantime?

2. RESEARCH QUESTIONS

The aim of this PhD is to contribute towards an evidence base that will help optimise mammographic and clinical follow-up for women who are disease free after primary treatment for early breast cancer.

The first two questions are linked and address the clinical and cost-effectiveness of follow-up mammography. Specifically, does follow-up mammography improve patient outcomes, and is it efficient to move away from a "one size fits all" mammography schedule to a personalised schedule based on individual prognosis? A systematic review was performed to answer question 1 and to inform the model based economic evaluation used to answer question 2. The decision not to consider mammography with different strategies of clinical follow-up visits was essentially a pragmatic one. To inform such a model would require patient-level data describing the timing and frequency of clinical examination, and the mode of detection of recurrent tumors (whether by imaging or clinical examination). This would have required casenote review for the entire South Australian cohort over a 10 year period, which was beyond the scope of this study.

The third question addresses the follow-up visit. In the light of financial and workforce pressures that reduce the feasibility of the majority of follow-up services in Australia being provided by breast cancer specialists, what are the preferences of breast cancer survivors for possible alternative modes of delivery of follow-up services? The intent is that the results of the third study can help inform the design of new models of breast cancer follow-up that are both appropriate and acceptable to patients. This question was addressed using a discrete choice experiment.

The research questions are:

1. What is the impact on survival of the method and timing of detection of a second breast cancer event within the breast?
2. Is it efficient to tailor mammographic follow-up according to risk of recurrence, for women who are disease free following completion of primary treatment for early breast cancer?
3. What do Australian breast cancer survivors prefer with respect to the provider, location, frequency, and method of delivery of routine follow-up care in years 3, 4 and 5 following diagnosis if existing specialist services were not available; and what is the perceived value of offering “drop-in” clinics providing additional support?

3. METHODS

Systematic Review

While many international guidelines exist, all acknowledge that there is a paucity of definitive evidence on follow-up mammography after early breast cancer. [12-14,16,21-23]

A systematic literature (Chapter 2) review was performed to identify papers that estimated the impact on survival of follow-up aimed at early detection of a second cancer event within the breast, for women with a personal history of early invasive breast cancer. Studies were included in the review if they (1) enrolled women treated for early invasive breast cancer without evidence of distant metastasis at primary diagnosis, who were in complete remission following primary treatment; (2) present data on overall survival or breast cancer related survival for early vs late detection of a second malignancy in the breast OR present data that compare survival in groups receiving different frequencies of breast cancer surveillance.

Studies that appeared to meet the inclusion criteria on the basis of title and abstract were retrieved for full appraisal to determine suitability. Reviewed papers were categorised by study design, and data extracted according to study design and mortality outcome. The methodological quality of the included studies was assessed by two reviewers using the validated Downs and Black checklist.[28] Studies were given an overall score and a score for risk of bias and confounding. The threshold levels for overall score (OS) and bias and confounding score (BCS) used to define study quality were determined a priori, with papers with the higher BCS being deemed better papers.

Following identification and appraisal of the evidence around mammographic effectiveness, I suggest strategies for improving our research to help overcome identified limitations in the evidence base.

Health Economics

The ultimate purpose of health economics is to inform the efficient use of scarce health resources.[25] To inform the efficient allocation of resources with respect to the follow-up of breast cancer survivors, I selected two forms of economic analysis that are most relevant to this question: decision analytic modelling and a discrete choice experiment. Discrete event simulation modelling is used in chapter 3 and 4 to assess the cost-effectiveness of alternative mammographic follow-up schedules, based on risk of recurrence. A discrete choice experiment is used in chapter 6 to examine Australian women's preferences for breast cancer follow-up.

For each technique, the following sections provide general background, followed by issues relating to the two techniques that are specific to breast cancer follow-up.

Technique 1: Discrete event simulation modelling

General background

Economic evaluation is the most common form of health economic analysis, and "compares the costs and benefits of two or more alternative interventions or services".[29] The two most common forms of economic evaluation involve evaluations alongside clinical trials, and decision analytic model-based evaluations. Evaluations alongside clinical trials, sometimes described as piggyback evaluations, allow prospective patient-specific data on both costs and outcomes,[25] and have been described earlier. While randomised controlled trials are focussed on measurement from a single source, model based evaluations draw on evidence from a range of sources for the purpose of informing specific decisions.[25]

Modelling is a decision analytic tool that enables data to be synthesised to describe disease progression over the time horizon of the model, commonly over the remainder of patients' lifetime, and to analyse the costs and benefits of a wide range of alternative intervention strategies. Decision modelling is considered appropriate in situations where few data are available or where trials are difficult or impossible to conduct.[27] It can be used to extrapolate data beyond the period of a trial (e.g. to determine survival for trials

with intermediate clinical endpoints or short duration of follow-up); to generalise findings of trials to regular clinical practice or to other settings (e.g. other countries); to create head to head comparisons of interventions where trials do not currently exist; and to identify priorities for further research or data collection.[27]

Decision analytic models are designed to reflect the natural history of a disease, and represent possible patient pathways over a long time horizon. For any given patient, their pathway may be influenced by their individual characteristics (e.g. age, menopausal status), their tumor (e.g. size, nodal status, grade), or their treatment; which in turn impacts on health service costs, quality of life, and overall survival.

Conceptually, decision analysis consists of four main stages: [29]

1. Defining the decision problem, and building a model structure that accurately reflects reality without being overly complex
2. Identifying and obtaining relevant cost and clinical data, to populate the model
3. Input the data into the decision model framework, to estimate the costs and effects of the interventions under study, including calculating the incremental cost effectiveness ratio (ICER), or cost per additional unit of benefit obtained, of a given intervention compared with the next best alternative.
4. Sensitivity analysis to assess the robustness of the ICERs, and to examine how the results may vary due to uncertainties in the data, model structure or model analysis.

Model performance should also be evaluated, including an assessment of face validity (does the model make sense?), internal validity (consistency between model inputs and outputs), external validity (comparing outputs to observed estimates of those outputs), cross-model validity (between model comparisons), and ideally predictive validity (comparing model outputs to prospectively collected data).[30]

Markov models have been traditionally used for economic evaluation of health care technologies.[31] Markov models are state transition models, where events occur at

regular intervals. Markov models comprise a finite number of health states that a patient can occupy.[30] At a given time, the patient will be in one health state, and all patients in that health state have identical characteristics (cohort).[30] The model is run for multiple cycles of a predetermined length over a fixed time horizon, and patients move between health states or remain in the same health state, according to transition probabilities. A Markov model is "memoryless", that is, the transition probabilities only depend on the current health state, and not on past health states (Markovian assumption).[30] Costs and QALYs are assigned to each health state and calculated at the end of each cycle.[30] The aim is to provide expected costs and outcomes for alternative interventions of interest.

Discrete event simulation models (DES) have been more recently introduced in economic evaluation in health care.[31] In DES models, patients move through the model and experience events at any discrete time period after the previous event.[31] DES models allow more complicated representations of the system being modelled [32] and a more intuitive representation of the available data; but require greater time and expertise to develop and evaluate.[31]

Discrete evaluation simulation has been used to evaluate screening and treatment of several cancers. Most notably, screening for Helicobacter pylori to prevent gastric cancers;[33-35] choice of surgical approach[36] and radiotherapy modality[37] for localised prostate cancer; and cervical screening programs.[38] DES is also used in optimising health services delivery to cancer patients: chemotherapy scheduling;[39,40] radiotherapy planning or delivery;[41,42] performance of colonoscopy suites[43] and flow through mammography clinics.[44] Erenay et al used a DES model to mimic the progression of metachronous colorectal cancer (MCRC) under the surveillance protocol suggested by the American Gastroenterological Association.[45] They used observed outcomes from a single institution to "feed" the model and by calibrating some of the model parameters, were able to estimate 6 previously unknown parameters of the natural history of MCRC. They anticipate that their model could be used to perform a cost-effectiveness analysis for the evalution of different colonoscopy follow-up schedules for patients with a personal history of colorectal cancer.[45]

Whilst a model can be rapidly updated or extended in response to new knowledge, it has several limitations. Models depend on clinical data, and the better the quality and the greater the extent to which it reflects all appropriate evidence[25] the more confident we can be about the results of our model analysis. Models are simplifications of reality. Any process of simplification will require assumptions, and these judgements impact on the structure of the model.[25] The extent to which a particular assumption impacts on the results requires careful evaluation. There are also concerns regarding model transparency, in that the inner workings of the model are often difficult to understand. Model performance evaluation includes an assessment of whether the model structure and results make intuitive sense (face validity); internal "debugging" and comparison of model results with both the data used to construct the model (internal validity) and independent data (external validity); and determining whether different model types built for the same decision yield similar results (cross-model validity). [46]

Breast cancer follow-up

In breast cancer, discrete event simulation has been used in economic evaluations of screening;[47,48] staging[49] and adjuvant treatment,[50] but no applications to breast cancer follow-up were identified. Stout et al assessed the cost-effectiveness of U.S. mammography screening practices during 1990-2000 from a societal perspective, and compared it to 64 alternative mammographic screening scenarios based on age at first screen, age at last screen and screening interval.[47] They confirmed that a mammography screening program can be cost-effective compared with no screening, but as many of their alternative screening scenarios were more effective and less costly than the prevailing patterns from 1990-2000 it is likely that the mammography screening programs could be improved.[47] Stout et al identified that the cost-effectiveness of mammographic screening schedules were sensitive to quality of life effects associated with the test itself (and varied according to current age, and whether the mammogram result was negative or false positive), and to level of population participation in the screening program (which is analogous to adherence with follow-up mammography).[47]

Model calibration

As is the case for incident cancer in the general population, for breast cancer follow-up, we cannot observe the time at which patients develop impalpable recurrence. In the absence of empirical data, these parameters may either be estimated, for example, by clinicians, or calibrated. Calibration involves fitting the model's input parameters so that the model's outputs match some observed data as closely as possible.[51] Ideally, the two approaches are combined, i.e. experts provide a range of plausible parameter values, which inform a calibration process.

The ideal dataset for calibrating an early breast cancer follow-up model would therefore comprise a cohort of women with similar risk of recurrent breast cancer, details of follow-up received (i.e. timing, type, and results) and the timing and type of any recurrent events and deaths. The accuracy of the model predictions of recurrence events and deaths would be assessed against the observed values in the dataset. The calibrated model could then be analysed to estimate the cost-effectiveness of alternative mammographic follow-up schedules for early breast cancer survivors, according to risk of recurrence, as reflected by relevant characteristics of their primary tumour, including appropriate sensitivity analysis to represent uncertainty.[52]

However, data on breast cancer follow-up and recurrence is not routinely collected in Australia. As no single data source within Australia contains the information required, we would need to extract and link retrospective data from the Cancer Registries and clinical and administrative hospital databases in order to construct a patient level dataset of women with early breast cancer. We would need to stratify women according to age and prognosis based on collected pathological criteria. For each woman we would determine the follow-up she received and her disease course. These data would form the basis for calibrating a decision analytic model to evaluate the cost-effectiveness of alternative mammographic follow-up schedules for early breast cancer survivors.

Technique 2: Discrete Choice Experiment

General background

Discrete Choice Experiments (Chapter 6) are a choice-based stated preference method of economic analysis.[29] Respondents are provided with scenarios that describe two or more interventions or services, and asked to choose the option they most prefer. It assumes that respondents make trade-offs between the characteristics of each option, and choose the intervention or service that is of higher value to them.[29]

Discrete Choice Experiments (DCE) have become a commonly used technique in health economics.[53] The technique is an attribute-based measure of benefit, based on the assumption that interventions (e.g. breast cancer follow-up services) can be described by a number of salient or key attributes (characteristics) and that an individual's valuation of the intervention depends upon the levels of these attributes.[54] Important attributes may be identified from literature reviews, expert clinical opinion, focus groups or individual interviews. Patients are given hypothetical scenarios comprising different levels of attributes and asked to choose between two or more alternatives. An individual's preference for a given alternative will reflect their personal evaluation of the advantages and disadvantages of each choice, and the strength of their preference will be a measure of their willingness to accept that choice.[55] That is, the patients' stated preferences reflect their perceived benefit, where benefit is defined in terms of the economic concept of "utility" or value.[49] The results are analysed with regression models,[56,57] for example, a mixed logit regression model (also known as the random parameter logit model) to investigate the potential for preference heterogeneity.[58-60]

While DCEs were introduced into health economics to value patient experiences,[61,62] their use has broadened within health care. A recent literature review identified that in addition to patient or consumer experience factors, the key applications of DCEs in health care to date have been: valuing health outcomes; investigating trade-offs between health outcomes and patient experience factors; estimating utility weights within the QALY framework; eliciting job choices among health professionals; developing priority setting frameworks; and eliciting health professional's preferences for treatment or screening options for patients.[53]

In the context of cancer, DCEs have been used to evaluate preferences with respect to screening; diagnostic testing; treatment options; side-effects of chemotherapy and paediatric oncology.[63-82]

Breast cancer follow-up

Patient preferences can inform clinical decision-making and improve satisfaction and adherence to health programs.[83] If we are considering designing a new breast cancer clinical follow-up service we need to understand which features of the clinic visit and service are of most value to the women who are eligible to attend. This can help inform the design of alternative service pathways that are acceptable to patients, for which further assessments of costs and patient outcomes can be undertaken. After all, it is of little benefit designing new care pathways or evaluating their cost-effectiveness, if the new service is not acceptable to the end user.

The only published DCE addressing follow-up services is in breast cancer. In 2010, a group used DCE methodology to assess patient preferences for the first year of breast cancer follow-up in the Netherlands.[84] Attributes studied were attendance at an educational group program, frequency of visits, waiting time (after the set time of the appointment), contact mode (face-to-face or telephone) and health care provider (medical specialist, breast care nurse, general practitioner, breast care nurse and medical specialist). Kimman et al found preference heterogeneity for most of the attributes under study, "indicating that one strategy does not fit all" for breast cancer follow-up.[84] In Australia, alternative modes of delivery are being considered to reduce the burden upon the cancer specialist. Proposed new models of care include different providers and different locations of follow-up services.[85-87] There is also a growing recognition of unmet need relating to treatment and psychosocial sequelae in breast cancer survivors. [88]

Both model based evaluation and discrete choice experiments are techniques that could be used across a variety of cancer types to address similar issues around the follow-up of patients beyond primary treatment.

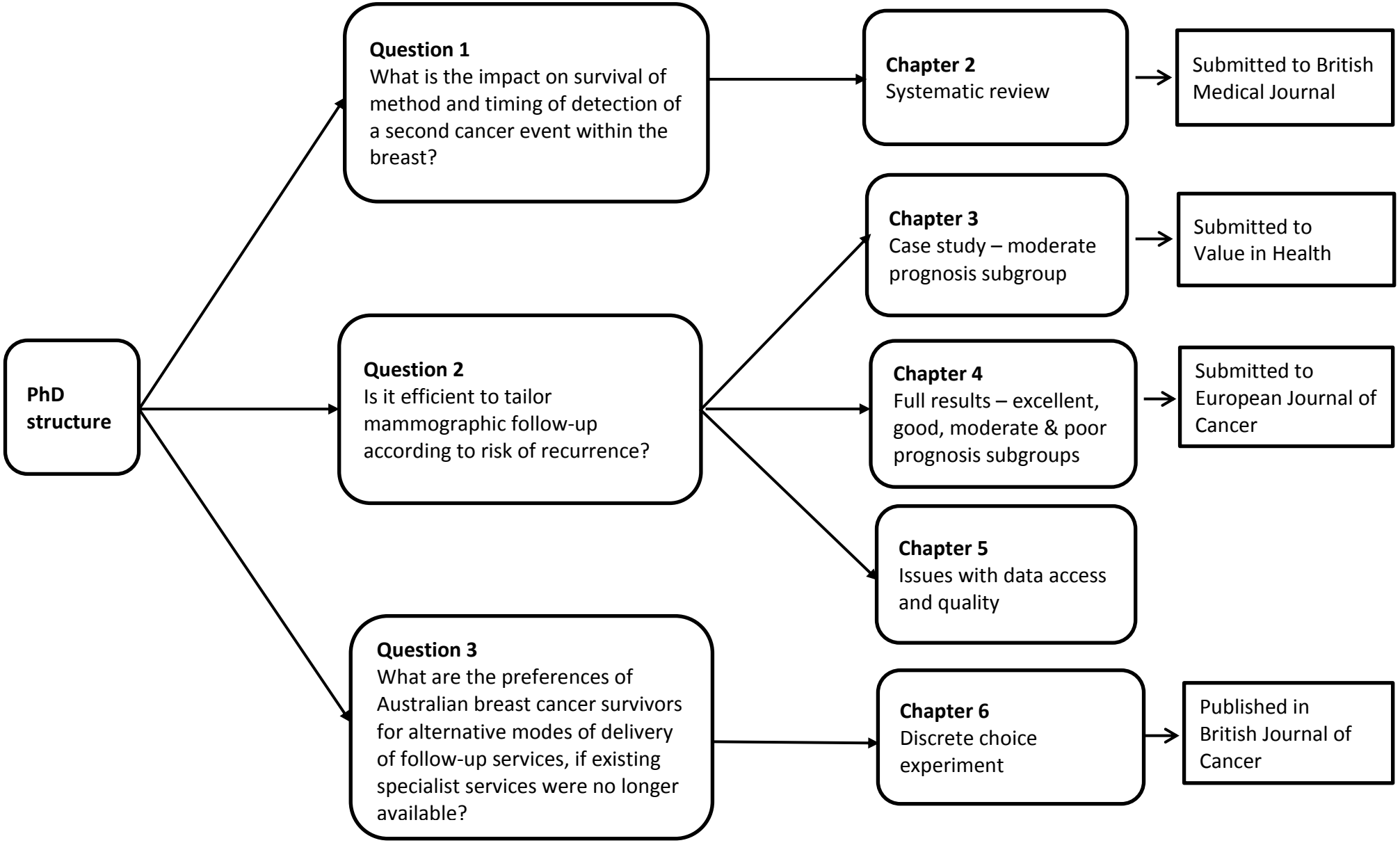
4. MY PhD

Three distinct research projects were undertaken within this thesis: a systematic review, discrete event simulation modelling, and a discrete choice experiment.

When I commenced my PhD in 2009, there was no similar work as yet published, for any of the studies in my projected program of work. However in late 2011, the National Health Service in the United Kingdom, published a commissioned Health Technology Assessment (HTA) titled "The clinical effectiveness and cost effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews, registry databases analyses and economic evaluation".[89] In this publication, the authors included a systematic review and a model based analysis of different mammographic schedules. At the time of the HTA publication, my systematic review had already been submitted to the British Medical Journal and been sent out for external review. It was eventually rejected, as a few days prior to my submission, the HTA had been published, and my review was no longer novel. In the thesis, I have kept the systematic review in the same form as the BMJ submission (but added Appendix 2 and 3), and have also compared my review with the HTA. As well as a systematic review similar to my own, the HTA also included a model based economic evaluation of the clinical and cost-effectiveness of different follow-up mammography regimens after the treatment for primary breast cancer. Whilst the HTA used a Markov model, I used a very different model structure (discrete event simulation) and modelling approach (calibration). In addition, in 2010 Kimman et al from The Netherlands published a discrete choice experiment (DCE) that explored patient preferences for follow-up after treatment for breast cancer.[84] The attributes and levels included in my study are however notably different (see chapter 6), and carefully chosen to reflect the Australian health system and the current policy context.

Below I diagrammatically represent the research questions, and how the research projects address each question. Chapters 2, 3, 4 and 6 were written as publications, and chapter 5 as a standard chapter. This is followed by a brief outline of the content of the respective chapters within this thesis. A comparison between my systematic review and economic evaluation, with that published in the Health Technology Assessment, is included at the end of chapters 2 and 4 respectively.

Figure 1: Structure of research program



Outline of Chapters

Chapter 2: Systematic Review

The aim of this paper was to review the evidence around the effectiveness of mammographic follow-up. The review identified a complete absence of randomised clinical trials in this area. I conclude that we should embrace alternative research techniques, such as decision analytic modelling, to guide our practice in the likely continued absence of randomised controlled trials (RCTs) in this field.

Chapter 3: A patient-level calibration framework for evaluating surveillance strategies: a case-study of mammographic follow-up after early breast cancer

This paper describes the development, calibration, and cost-effectiveness analyses of an early breast cancer surveillance discrete event simulation (DES) model. The DES model was used to analyse three alternative mammographic follow-up schedules for postmenopausal women who were disease free following primary treatment for moderate prognosis early breast cancer: taking into account age and adherence to mammography. This study demonstrates the potential value of combining linked, retrospective data and decision analytic modelling to provide estimates of costs and health outcomes that are sufficiently robust to inform cancer clinical guidelines and individual patient decisions regarding appropriate follow-up schedules.

Chapter 4: One size does not fit all? Cost utility analyses of alternative mammographic follow-up schedules, by risk of recurrence

The aim of this paper is to report the full set of cost-effectiveness results from the model described in chapter 3, comparing alternative mammographic follow-up schedules for postmenopausal women with excellent, good, moderate and poor prognosis early breast cancer. Our results suggest that annual mammographic follow-up is not cost effective for most postmenopausal women, and that mammographic follow-up can be tailored based on the Nottingham Prognostic Index score of the primary breast cancer and age at diagnosis.

Chapter 5: Issues with data access and quality

This chapter describes the large gulf between the ideal dataset to inform the cost-effectiveness model and what is currently available in South Australia (and beyond).

Chapter 6: Discrete Choice Experiment (DCE)

The aim of this study was to explore the preferences of Australian breast cancer survivors for alternative modes of delivery of follow-up services if we could no longer offer long term specialist-led hospital based follow-up. This study provides important insights into which attributes of a breast cancer follow-up service women value most.

Chapter 7: Conclusion

This chapter summarises and synthesises the findings of the systematic review, discrete event simulation model, and discrete choice experiment. I then discuss limitations of this research with respect to scope and data; describe relevant new research in the field of breast cancer follow-up over the period of this work; suggest potential directions for future research; and provide recommendations for policy that arise from this thesis.

Chapter 2

SYSTEMATIC REVIEW

TITLE

Mammographic follow-up after early breast cancer: why is the evidence so poor, and how can we improve it?

INTRODUCTION

Due to early diagnosis and improved treatment outcomes, there is a growing population of breast cancer survivors who will all require mammographic follow-up during their lifetime.[1] Women with a past history of breast cancer are at increased risk of relapse in the treated breast or a second malignancy in the opposite breast, but it is unclear how often, or for how long, we should continue to perform mammography in women who are disease-free following primary treatment for early breast cancer.

Mammographic follow-up is effectively screening of a high risk population. However not all women who have survived early breast cancer have the same risk of breast cancer recurrence or a second primary, with rates per 1000 woman-years varying by patient, tumour and treatment characteristics at initial diagnosis.[2]

Many international guidelines exist, and all acknowledge the paucity of definitive evidence available on follow-up care after breast cancer, with much of the information from low level observational studies.[3] International guidelines do not report randomised controlled trial evidence that supports their recommendations of annual compared to other frequency of follow-up mammography, and there is currently no tailoring of mammographic frequency according to risk.[3-7] In a recent guideline update, no primary studies were identified which addressed how long follow-up should continue after diagnosis or treatment.[3]

The purpose of this paper is to systematically review the existing evidence around mammographic follow-up after breast cancer, assess the quality of the evidence, and to suggest strategies for improving our primary research to help overcome identified limitations in the evidence base.

METHODS

A systematic literature review was performed to identify papers that estimated the impact on survival of follow-up aimed at early detection of a second cancer event within the breast. PubMed, Scopus, the Cochrane library, Ovid Medline and Web of Knowledge were searched for relevant studies. Studies in any language published between 1980 and August 21, 2010 were examined. A separate search string was written to identify literature assessing the method and timing of detection of a second breast cancer event within the breast. MESH terms used were "Breast Neoplasms", "Neoplasm recurrence, local" and "Follow-up studies", plus additional MESH terms of "Mammography", "Physical examination", "Diagnostic Tests, Routine" and "Time factors" for the method of detection search string , and "Disease-free survival" and "Time factors" for the timing of detection search string. For each MESH term, multiple related text words were included (Appendix 1). Early invasive breast cancer was defined as breast cancer that has not spread beyond the breast or the axillary lymph nodes, and includes stage I, IIA, IIB and stage IIIA breast cancers (http://www.cancer.gov).

Studies were included in the review if they (1) enrolled women treated for early invasive breast cancer without evidence of distant metastasis at primary diagnosis, who were in complete remission following primary treatment; (2) presented data on overall survival or breast cancer related survival for early vs late detection of a second malignancy in the breast OR presented data that compared survival in groups receiving different frequencies of breast cancer surveillance.

Studies that appeared to meet the inclusion criteria on the basis of title and abstract were retrieved for full appraisal to determine suitability. Reviewed papers were categorised by study design, and data extracted according to study design and mortality outcome. We operationalised relapse in the treated breast or axilla as "ipsilateral breast recurrence" (unless treated as separate entities within the paper), and a new primary or a metastasis in the opposite breast as a "contralateral breast cancer". We assumed "local recurrence" referred to an ispilateral recurrence, that "overall survival" was measured from the primary

breast cancer, and that reported 5 year survival rates were measured from time of diagnosis of an in-breast relapse, unless otherwise stated.

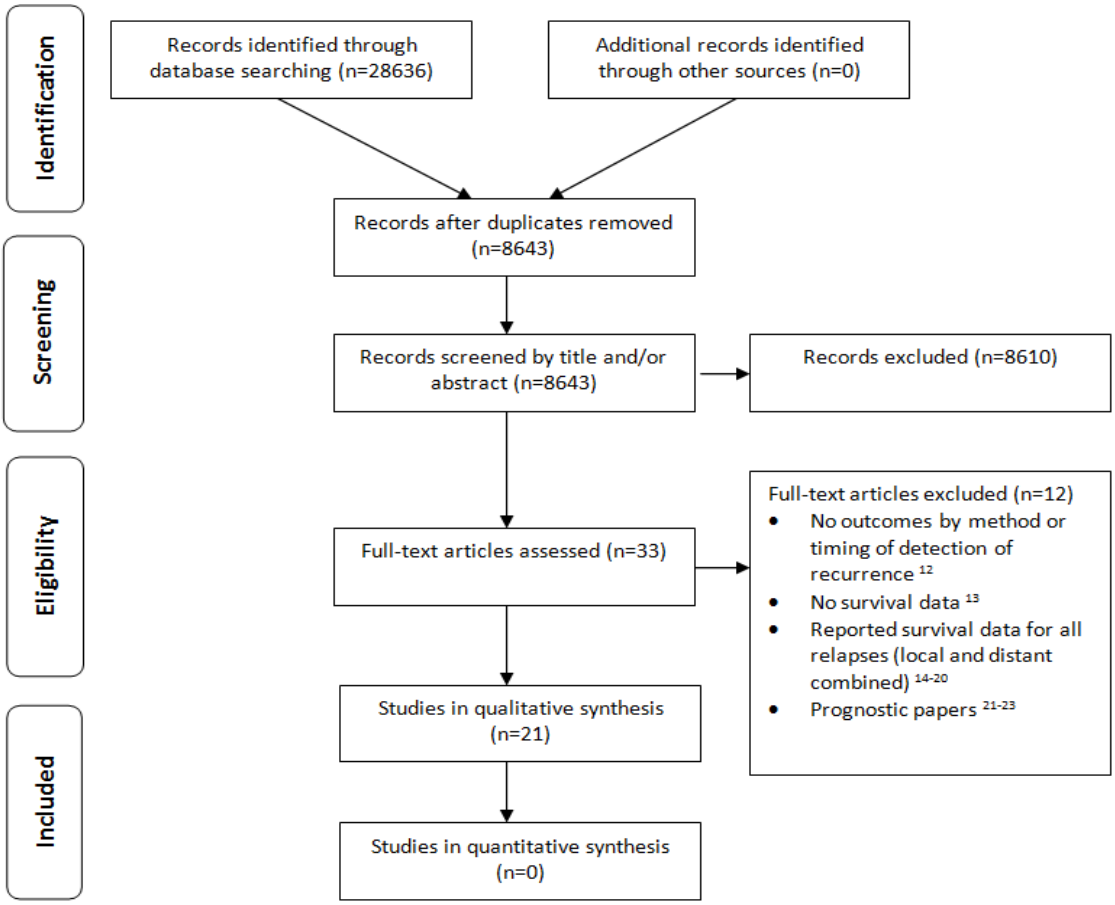
The methodological quality of the included studies was assessed using the validated Downs and Black checklist[8] (Appendix 2).This scale was chosen as it assesses the methodological quality of both randomised and nonrandomised studies of health care interventions, was specifically developed for use in systematic reviews, has been evaluated as one of the "best" critical appraisal tools for assessing quality of nonrandomised studies, is "easy to use" and the responses to individual questions can be used to contribute to an overall judgement of study quality using a summary score.[9] Studies were appraised in terms of "reporting" (10 items), "external validity" (3 items), "internal validity-bias" (7 items) and "internal validity - confounding (selection bias)" (7 items). Question 27 concerning statistical power was dropped from the analysis as there was no pre-specified description of what constituted a clinically important difference in outcome.

Each of the included papers was assessed by two reviewers using the checklist. Studies were given an overall score and a score for risk of bias and confounding. Based on a previous review,[10] the threshold levels for overall score (OS) and bias and confounding score (BCS) used to define study quality were determined a priori as follows: good quality = OS ≥20 and BCS ≥10; good-moderate=OS ≥20 and BCS 6-9 OR OS 16-19 and BCS ≥10; moderate quality = OS 16-19 and BCS 6-9; moderate-poor = OS 16-19 and BCS ≤5 OR OS ≤15 and BCS 6-9; and poor quality = OS ≤15 and BCS ≤5. For the good-moderate and moderate-poor quality ratings, the papers with the higher BCS are deemed better papers.

RESULTS

The process of article identification, screening, and exclusion are shown in Figure 1 below.

Figure 1 – PRISMA Flowchart[11]



A total of 21 papers were included; all were observational, and of level III-2 or III-3 evidence for intervention studies[24] (Appendix 3). Three nested case-control studies and eighteen cohort studies were analysed by mortality outcome (Tables 1&2) and critically appraised (Table 3).

Nested case-control studies

Lash and colleagues published three nested case-control studies focussing on the impact of follow-up mammography on all-cause mortality[25,26] and breast cancer specific mortality[27] in women previously treated for early (stage I-II) breast cancer (table 1).

Table 1: Nested case-control studies

Author, year, quality	Time frame, country	Population	Cases	Controls	Exposure	Covariates	Analysis	Results
Lash, 2005, moderate quality	1992-1994, USA	≥ 55 yrs old, previous stage I-II breast cancer, from single city (n=303)	Women who died (n=63). Women who died from breast cancer (n=27).	For each case, 5 controls were sampled with replacement from the cohort of women at risk of dying at the same time in their follow-up period (n=315).	Consecutive years of receipt of guideline surveillance (history, clinical examination, annual MMG). Women who received no guideline surveillance were the reference group.	Age at diagnosis, race, education, employment, number in house, marital status, cardiopulmonary co-morbidity status, tumor stage, primary surgical therapy, receipt of radiation therapy and adjuvant systemic therapy (chemotherapy, tamoxifen)	Multivariate conditional logistic regression model.	All cause mortality : per year of guideline surveillance, age adjusted OR[a] = 0.66 (95% CI 0.51,0.86). Other than age, no other covariates were significant predictors of outcome.
Lash, 2006, moderate quality	1996-1999, USA.	≥ 65 yrs old, previous stage I-II breast cancer, from 4 geographically different locations (n=865). Case-control analysis limited to women with at least 1 surveillance visit after completing their therapy (n=334).	Women who died (n=32). Women who died from breast cancer (n=13).	For each case, 8 controls were sampled with replacement from the cohort of women at risk of dying at the same time in their follow-up period (n=302).	Number of annual mammograms received during the follow-up period.	Enrollment site, age at diagnosis, Charlson comorbidity score, tumor stage, estrogen recepor status, primary surgical therapy, receipt of radiotherapy and adjuvant systemic therapy (chemotherapy, tamoxifen)	Multivariate conditional logistic regression model.	All cause mortality: age, stage and tamixofen adjusted OR (per additional MMG) = 0.77 (95% CI 0.53-1.1) (p for trend 0.007)
Lash, 2007, moderate quality	1990-1994, USA	≥ 65 yrs old, previous stage I or II breast cancer, from six integrated health care delivery systems	Women who died from breast cancer (n=178)	Matched controls observed at least as long as the case and matched on enrollment site, age category, AJCC stage, type of primary surgery and baseline co-morbidity index (n=634). 4 controls were randomly selected from the risk set when risk set contained >4 subjects, and selected the whole risk set to be the case's controls when the risk set contained < 4 subjects.	Number of annual mammograms received during the follow-up period	Year of diagnosis, age at diagnosis, race, co-morbidity index, tumor stage, estrogen receptor status, primary surgical therapy, receipt of radiotherapy and adjuvant systemic therapy (chemotherapy, tamoxifen).	Multivariate conditional logistic regression model.	Breast cancer mortality: OR (per additional MMG) = 0.69 (95% CI 0.52,0.92) adjusted for matched factors and change in Charlson co-morbidity index from baseline. All but breast cancer mortality : OR (per additional MMG) = 0.68 (95% CI 0.56,0.83), adjusted for age and change in Charlson co-morbidity index from baseline.

a OR = Odds Ratio

In the 2005 and 2006 studies, the cases were women who had died, and the controls were sampled from the cohort of women at risk of dying at the same time in their follow-up period. The two studies addressed, respectively, the receipt of guideline surveillance (history, clinical examination, annual mammography) and the receipt of annual mammography, on all-cause mortality. In the 2007 study, the cases were women who had died from breast cancer, controls were matched based on length of observation and five additional criteria (see below), and the study examined the effect of receipt of annual mammography on breast cancer mortality .

Using the critical appraisal tool, the three studies were all considered moderate quality (table 3, see below). Limitations of these studies primarily related to the integrity of information on exposure, matching and self-selection bias.

In the 2005 and 2006 papers by Lash and colleagues, there was **incomplete information on the exposure** (receipt of guideline surveillance and number of annual mammograms) within the community setting which, the authors recognised, may have biased the impact on survival of follow-up towards the null.[25,26] This limitation was addressed in their 2007 study where there was complete ascertainment of follow-up mammography.

Lash et al (2007) calculated the odds of breast cancer mortality associated with each surveillance mammogram (MMG), adjusted for matched factors (enrolment site, age category, AJCC stage, primary surgery type, and baseline Charlson co-morbidity index), and for change in Charlson co-morbidity index during the study.[27] **Matching** is defined as the process of selecting controls so that they are similar to cases in certain characteristics.[28] The aim is to control for an unequal distribution of potential confounders between groups, given these cannot be balanced through randomisation in an observational study design. It is possible that factors other than those matched by Lash et al (eg family history, tumour genetic profile) or other exposures (eg preventive health care) may have accounted for the observed effect. All three studies demonstrated a protective effect of mammography, but the magnitude of this effect in the largest and last study was similar for both all-cause and breast cancer mortality.[27] No analysis by baseline risk was reported. The benefit could be attributed to earlier detection of recurrence by surveillance mammography (with better

prognosis); receipt of preventive health care during follow-up visits for co-morbidities (eg obesity) and lifestyle factors (eg alcohol intake) that are thought to impact on breast cancer risk but which also affect other cause mortality.[26,27]Alternatively, the results may reflect **self-selection bias** in that women who attended follow-up may have been systematically different (ie. more health conscious) than non- attenders.

Given the potential bias associated with these studies, the reported effect of mammographic follow-up on survival would have to be interpreted cautiously (Table 4, see below).

Cohort studies

Eighteen cohort studies were identified, of which fourteen addressed early vs late detection of a second cancer event within the breast, and four addressed alternative levels of mammographic surveillance (table 2).

Table 2: Cohort studies

Author, year, quality	Time frame, country	Population : PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Ciatto 1984, moderate quality	1971-1978, Italy	Previous breast cancer treated with Mx ± RT/systemic AT (n=1139), relapse at **any site** prior to Nov 1983 (n=315)	Asymptomatic vs symptomatic detection	**IBR** = 104, asymptomatic = 54/104	Primary breast cancer	CE and CXR 6/12 for 3 years, then yearly. MMG frequency not reported.	Minimum follow-up of recurrent cases of 5 years.	5 year actuarial and mean survival	Better mean survival (75.5 vs 64.9 months) and 5 year actuarial survival (54% vs 40%) for cases detected asymptomatically vs symptomatically, but not statistically significant
Ciatto 1985, moderate quality	1970-1983, Italy	Previous breast cancer, relapse at **any site** (n=1120)	Asymptomatic vs symptomatic detection	**IBR** = 349, asymptomatic = 215/349	Primary breast cancer	Variable choice of tests and frequency between sites, CE 3-6/12 , CXR, WBBS, Liver US 6-12/12. MMG frequency not reported.	Vital status assessed April 1984	10 year actuarial and median survival	No significant difference in median survival (45 vs 46 months) or 10 year actuarial survival (31% vs 28%) for cases detected asymptomatically vs symptomatically.
Ciatto 1990, good-moderate quality	1969-1986, Italy	Previous breast cancer treated with radical surgery ± AT, second BCE in **opposite breast** (n=175). Mean age at PBC 54.3yrs (range 26-81, SD 10.4 yrs)	Asymptomatic vs symptomatic detection	CLBC = 175, asymptomatic = 89/175	Primary breast cancer	Not reported	Median from PBC 14.5 years and from CLBC 7.5 years	Kaplan- Meier method. Log rank test comparing survival. Multivariate analysis of correlation to survival (variables entered showed significant correlation on univariate analysis).	No difference in 10 year overall survival (0.72 vs 0.8, p=0.44) on univariate analysis, for cases detected asymptomatically vs symptomatically
Ciatto 2004, moderate quality	1970-2001, Italy	PBC pTx-T4, pNx-N+>10, second BCE in **opposite breast** (n=429).Mean age 54.7 yrs (range 30-90)	Asymptomatic vs symptomatic detection	Known symptom status at diagnosis of CLBC = 339, asymptomatic = 234/339	Primary breast cancer	CE 6/12 for 5 yrs then yearly, MMG yearly for 5 yrs then every 2 years	End of follow-up August 2003, last linkage with Mortality Registry December 2001	Paired data - McNemar's Chi square, unpaired data - Chi square and unpaired t test. Cox regression.	HR of breast cancer death = 0.49 (95% CI 0.29-0.83, p=0.008) for CLBC detected asymptomatically vs symptomatically.

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population: PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Doyle, 2001, moderate quality	1977 - 1997, USA	PBC stage I or II treated with BCS & RT ± AT, second BCE in **treated breast** (n=112).	MMG vs clinical detection	Invasive IBR = 93, MMG = 32/93, clinical = 40/93, both = 21/93	Not explicitly stated	Not reported	PBC to IBR : mean 63 months, median 56 months (range 2-186). IBR to study end: mean 49 months, median 44 months (range 1-149)	Kaplan-Meier method. Mantel Cox test comparing survival. Cox proportional hazards multivariable regression model.	Detection of recurrence by clinical examination vs MMG vs both revealed a significant difference in 5 year overall survival (73% vs 91% vs 93%, p=0.04), borderline significant difference in 5 year cause specific survival (p=0.06) and a nonsignificant difference in 5 year distant metastasis free survival (p=0.22) on multivariate analysis.
Grosse, 1997 moderate quality	1972-1995, Germany	PBC pT1 - pT2 treated with BCS & RT, ± AT (n=3072), second BCE in **treated breast** prior to Dec 1995 (n= 90, invasive = 85, DCIS = 5). Median age at recurrence = 41 yrs (21-74).	MMG vs clinical detection	MMG= 26/85, CE = 10/85, both = 35/85, incidental/ unknown = 4/85, additional diagnostics = 10/85	Recurrence	MMG postop and 3/52 post RT. In early years, CE and MMG every 3/12, later policy CE and MMG every 6/12 for 5 years, then annual CE and MMG. Opposite breast examined annually. US & MRI prn.	Median from BCS 83 months (19-250), from salvage Mx 27 months (0-204).	Kaplan-Meier method. Wilcoxon-Mann-Whitney and the Kruskal-Wallis test. "Multivariate analysis".	5 year survival rate - MMG detected 83.6% vs clinically detected 71.8% vs detected by both MMG and CE 40%. "In a multivariate analysis, however, the method of detection could not be shown to be significantly predictive." .
Houssami, 2009, good to moderate quality	1980-2005, Italy	PBC in situ or stageI-III invasive treated with BCS/Mx, second BCE in **either breast** (n=1044). Median age second BCE = 60 yrs (IQR 51-70 yrs)	Asymptomatic vs symptomatic detection	IBR = 455, asymptomatic = 314/455, CLBC = 589, asymptomatic = 385/589	Primary breast cancer	BCS: CE 6/12 and annual MMG for 5 yrs then CE yearly and annual/biennial MMG. Mx: CE annual and annual/biennial MMG	Median from PBC 13.7 yrs (IQR 9-18.1 yrs)	Chi square test, McNemar's chi square test. Kaplan-Meier method. Cox regression analysis separately for IBR and CBC then a global multivariate Cox proportional hazards model. Sensitivity analyses for length bias.	HR of breast cancer specific survival for asymptomatic vs symptomatic detection - 0.51 (95% CI 0.32,0.80, p=0.004) for IBR, 0.53 (95% CI 0.36.0.78, p<0.0001) for CLBC, and 0.53 (95% CI0.4,0.72, p<0.0001) for all subjects. Sensitivity analyses for length bias gave range of adjusted HR from 0.53-0.73.

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population: PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Hussain, 1995, poor quality	1980-1991, United Kingdom	PBC stage I or II treated with BCS & RT (n=354), second BCE in **treated breast** &/or axilla (n =33).	MMG vs clinical detection	LR= 24, LR + RR =6, RR = 3. MMG = 5/33, clinical by doctor = 24/33, clinical by patient = 4/33	Recurrence	CE 3/12 for 2 yrs, 6/12 for 3 years, then yearly up to 10 years, MMG at 6/12 and then yearly.	Median 72 mths (24-131)	Mann-Whitney U -test, chi-squared or Fisher's exact test.	14/30 died (42%) , none of the 5 women detected on MMG died (p=0.02)
Kaas, 2001, moderate quality. FREQUENCY	1976-1990, Netherlands	PBC stage 0-IIIA treated with surgery and/or RT(n=3900), second BCE in **opposite breast** prior to Dec 1997 (n=323). Mean age (SE) at PBC: 12/12 MMG group = 51 yrs (12), 15/12+ MMG group = 54 yrs (14). Mean age (SE) at CLBC: 12/12 MMG group = 56 yrs (12), 15/12 MMG group = 61 yrs (14)	Mean MMG interval 12 months vs 15+ months	Complete follow-up information for CLBC (n=275). MMG every 12 months = 141, MMG every 15+ months = 134	Recurrence	CE 3/12 for 2 years, 6/12 for 3 years then yearly. Prior to 1985, MMG not used routinely (hence CLBC detected in that period were more often palpable), and mean interval between MMGs was longer. Usually MMG at longer interval in patients > 60 yrs.	CLBC recruitment to Dec 1997, follow-up to Dec 1998	2 sample t test, chi-square test, logrank test, ANOVA, Cox proportional hazard regression model	5 year "disease-specific" survival from CLBC (SE) - MMG every 12 months vs MMG every 15+ mths = 75% (5) vs 75% (5), p=1.00.
Kaas, 2001, moderate to poor quality. DETECTION	1976-1990, Netherlands	PBC stage 0-IIIA treated with surgery and/or RT(n=3900), second BCE in **opposite breast** prior to Dec 1997 (n=323). Mean age (SE) at PBC: 12/12 MMG group = 51 yrs (12), 15/12+ MMG group = 54 yrs (14). Mean age (SE) at CLBC: 12/12 MMG group = 56 yrs (12), 15/12 MMG group = 61 yrs (14)	MMG vs clinical detection	Complete follow-up information for CLBC =275. MMG =109/275, clinical = 166/275	Recurrence	As above (Kaas 2001 freq).	CLBC recruitment to Dec 1997, follow-up to Dec 1998	2 sample t test, chi-square test, logrank test, ANOVA, Cox proportional hazard regression model	5 year disease specific survival (%) from CLBC - clinically detected 31% dead at 5 yrs vs MMG detected 15% dead at 5 years (p=0.015). Controlling for stage at PBC, p=0.015 remains.

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population : PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Krengli, 1993, poor quality	1964-1988, Italy	Previous nonmetastatic BC treated with Mx (n=1979) or BCS (n=503), relapse at **any site** (n=928). Mean age = 55 yrs (range 19-88yrs) , median age = 54 years.	Aymptomatic vs symptomatic detection	**LRR** (n=289), DM (n=639). Sample of 350 investigated [Note % of LRR in sample not stated, data provided for relapse in **opposite** breast but % of CLR within LRR and % of CLR within Mx vs BCS groups not stated]	Not explicitly stated	Medical checks every 3-4/12 for 2 years, 6/12 for 3 yrs, then annual. Each year, at least in first 5 years, patients received CXR, skeletal x-ray/WBBS, LUS (from 1983), MMG and 2 blood tests (from 1984).	Median 172 months (range 42-330)	p values for survival difference	"LRR were asymptomatic in 48%", "diagnostic anticipation of 3 months", "increased survival (p=0.04)"
Montgomery 2007, moderate quality	1991-1998, Scotland	PBC stage I- II treated with BCS & RT ± AT (n=1312), second BCE in **either breast** prior to Jan 2006 (n=110). Mean age at PBC = 54.28 yrs (range 24-83 yrs)	Routine vs interval	IBR = 36, IRR = 25, IBR+IRR = 10, CLBC =35, bilat=1, bilat+IRR =1, incidental excluded = 2. Routine MMG= 56/108, routine symptoms = 13/108, routine clinical = 15/108, interval symptoms = 24	Primary breast cancer and recurrence	Initially, CE 3-4/12 for 2 years, 6/12 for 3 years then annual visits until 10 years when discharged to national breast screening program. From 2000 changed to annual CE. Annual MMG throughout.	Median 10 years (range 1.5-15)	Log rank test comparing survival	Patients with symptomatic or MMG detected IBR had significantly longer survival from PBC (p=0.0002) and from recurrence (p=0.0014) compared with clinically detected. No association between method of detection of relapse and survival in patients who developed CLBC or isolated IRR (data not provided)
Orel, 1993, moderate to poor quality.	1977-1990, USA	PBC stage I or II treated with BCS & RT (n=1,636), second BCE in **treated breast** and underwent salvage Mx (n=83). Median age at second BCE: 49.5yrs (range 21-76).	MMG vs clinical detection	IBR = 72, MMG = 34/72, clinical = 24/72, both = 14/72	Recurrence	CE and MMG, but frequency not reported	Median from PBC 82.5 months (range 10-166 months), from second BCE 28 months (range 0-108 months).	Kaplan-Meier method. Mantel-Cox test comparing survival. Cox regression model.	Actuarial overall survival as a function of method of detection of recurrence- MMG vs CE ± abnormal MMG = 94% vs 84% (p=0.28). No statistically significant variables that could be used to predict improved overall survival in univariate or multivariate analysis

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population : PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Paszat, 2009, moderate quality	1991-1993, Canada	Population database of PBC stage I-II treated with surgery (n=12,279). Sample of 591/1200 women with and 310/400 women without subsequent breast surgery. Second BCE in **either breast** (treated conserved breast n= 84, opposite breast n= 49)	MMG vs no MMG	Receipt of ≥ 1 surveillance MMG (721/901) vs none (180/901).	Primary breast cancer	Surveillance MMG defined as MMG beginning ≥ 6 months following diagnosis of unilateral PBC, and if multiple ≥ 11 months apart and not prompted by clinical concern or symptom.	Median 8.36 yrs (IQR 1121-4928 days)	Analysed the study as a case series. Cox proportional hazards model for all women in the study population, and univariate analysis for IBR and CLBC separately.	Adjusted HR for time to breast cancer death associated with receipt of ≥ 1 vs 0 episodes of surveillance MMG was 0.28 (95% CI 0.22,0.37), adjusted for age, stage, type of surgery, adjuvant chemotherapy and tamoxifen. For 84/584 women with IBR, unadjusted HR = 0.36 (95% CI 0.13,1.00) and 49/901 women with CLBC, unadjusted HR = 0.86 (95% CI 0.2, 3.77).
Robinson, 1993, moderate-poor quality	1950-1989, Israel	Previous breast cancer at age ≤ 55, treated at single hospital, second BCE in **opposite breast** (n=167). Mean age at PBC = 53 +/- 12 yrs, mean age at CLBC = 60+/-12 yrs	Adherence vs non-adherence	CLBC =167. Adherence with routine follow-up = 73%. Detected by patient = 45%, GP= 35%, routine follow-up =38%, routine MMG = 8.2%, other= 0.6%.	Recurrence assumed (not explicity stated)	Routine follow-up = 1-3 visits for CE/year. MMG was not routinely done during the years of this study.	CLBC recruitment 1950-1989, follow-up to 2000.	Cox proportional hazards model.	"There was no difference in the 5 year survival rate between those patients who were under follow-up and those who were not" (data not shown). "Follow-up procedures for many years used mainly clinical examination and this was not enough to decrease mortality"
Te Boekhorst, 2001, moderate quality	1974-1990, Netherlands	PBC stage I-IIIB treated with BCS or Mx ± AT (n=1023), who subsequently developed relapse at **any site** (n=336). Mean age (SD) asymptomatic vs symptomatic = 53 yrs (13) vs 54 yrs(13).	Asymptomatic vs symptomatic detection	**IBR** alone = 68, asymptomatic =45/68, symptomatic (routine or interval) = 23/68 [Note IBR refers to treated breast and axilla, sites apart from treated breast and axilla are grouped as distant relapses, not explicit how second event in opposite breast is categorised]	Primary breast cancer and recurrence	CE 3/12 for 2 years, 6/12 for 3 years, then annual. MMG annual. Early part of study, CXR, WBBS, bloods done routinely, later only when indicated	Not reported	Log rank test comparing survival	5 year overall survival of asymptomatic vs symptomatic detected LRR 78% vs 61%, HR 1.3 (p=0.34). 5 year survival after recurrence of asymptomatic vs symptomatic detected LRR 51.1% vs 26.1%, HR 1.3 (p= 0.36)

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population : PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Tomin, 1987, moderate-poor quality. ADHERENCE	1967-1983, USA	Previous breast cancer treated with Mx (n = 1230), relapse at **any site** (n=324)	Adherence vs non-adherence	Adherence with follow-up to 5 years: 54/248 (22%)	Recurrence	CE 3/12 for 2 years, 6/12 for 3 years, then annual. MMG annual. Also blood tests every 6/12 and CXR annually.	Not reported	Kaplan-Meier method.	5 year survival : "survival was indistinguishable, with curves overlapping throughout their course" for patients who were compliant with follow-up and those who were not
Tomin, 1987, moderate-poor quality. DETECTION	1967-1983, USA	Previous breast cancer treated with Mx (n = 1230), relapse at **any site** (n=324)	Asymptomatic vs symptomatic detection	LR = 47, RR = 34, other sites = 167. Asymptomatic: **LR** = 16/47, **RR** = 17/34 [Note LR refers to Mx scar or skin/tissue flaps, RR refers to supraclavicular, internal mammary or axillary LN, other sites refers to lung, bone, visceral and multiple, not explicit how second BCE on opposite side is categorised]	Recurrence	As above (Tomin 1987 adherence)	Not reported	Kaplan-Meier method.	Significant difference in 5 year survival of asymptomatically detected LR (50% vs 10.8%, p=0.01), but not for RR (p=0.52)
Wagman, 1991, moderate-poor quality	1977-1983, USA	PBC stage I-III treated with Mx (n=208), relapse at **any site** prior to June 1989 (n=64)	Routine vs interval	**LRR**=11, **CLBC** =13, DM=40. Routinely detected LRR=9/11, CLBC=12/13	Primary breast cancer	Visits dictated by doctor preference or protocol on one of several NSABP breast protocols. Visits included CE, MMG, WBBS, liver US, CXR, bloods.	Observed to June 1989	Kaplan-Meier method. Mantel Cox test comparing survival. Adjustments for multiple comparisons using method of Bonferroni.	Survival following PBC: for locoregional recurrence – implied mean survival for interval vs routine detected = 87 months vs 82.1 months, for contralateral breast cancer – 1 interval case did not die, no comparison presented.

Table 2: Cohort studies (cont.)

Author, year, quality	Time frame, country	Population : PBC, second BC event, age	Exposure	Study	Survival from	Frequency of routine testing	Duration of follow-up	Analysis	Results
Yau, 2008, moderate quality	1994-2003, Hong Kong	PBC pT1-pT2 invasive breast cancer or DCIS treated with BCS & RT (n=507 women, 511 cancers), second BCE in **either breast** (n=36). Median age = 46 yrs (range 25-90)	MMG &/or US vs clinical detection	IBR = 23 (incl 4 DCIS), CLBC = 13. IBR : MMG = 10/23, US = 3/23, clinical = 10/23. CLBC : MMG = 8/13, US =1/13, clinical = 4/13	Primary breast cancer	CE 2-3/12 for 2 years, 4-6/12 for 3 years, then annual. Annual MMG in first 5 years, and then every 1-2 years in next 5 years. US prn.	Outcome data updated March 2007. Median follow-up = 5.9 yrs (range 0.2-13)	Kaplan-Meier method. Log rank test comparing survival.	Distant metastases free survival: no significant difference for MMG and/or US vs clinical detection (p=0.342)

PBC = primary breast cancer
BCS = breast-conserving surgery, Mx = mastectomy, RT = radiotherapy
BCE = breast cancer event, IBR = ipsilateral breast recurrence, CLBC = contralateral breast cancer
LR = local recurrence, LBR = locoregional recurrence, RR= regional recurrence, DM = distant metastases, LN = lymph nodes
MMG = mammography, CE = clinical examination, US = ultrasound, CXR = chest x-ray, WBBS = whole body bone scan, MRI = magnetic resonance imaging
GP = general practitioner
HR = hazard rate

The fourteen studies addressing early vs late detection, variously described comparisons between asymptomatic vs symptomatic detection[29,37,38,41,42,43,46], mammographic vs clinical detection[31,33-36,39], or routine vs interval detection[30,44] of a recurrence or new primary breast cancer.

Only one cohort study comparing different frequency of mammographic surveillance[39] met the inclusion criteria. Two papers comparing adherence vs nonadherence with mammography[40,45],and one paper comparing mammography vs no mammography[32] were thematically grouped with the frequency paper as all compared alternative levels of mammographic surveillance[32,39,40,45] (Table 2).

Two papers addressed both early vs late detection and alternative levels of mammographic surveillance.[39,45] For these papers, each outcome was analysed separately giving a total of 23 analyses for 21 papers.

The fourteen early vs late detection papers identified women at the point of recurrence, and compared survival from either primary breast cancer or recurrence. The four alternative levels of mammographic surveillance papers identified women at the point of surgery for primary breast cancer, and then compared survival with respect to frequency, adherence or receipt of mammography.

Important biases observed in the cohort studies include **selection bias** (a systematic difference in survival prognostic factors between women receiving one method of frequency of surveillance, relative to women receiving the comparison method of frequency of surveillance, which may affect the internal validity of the results[47]) and **attrition bias** (systematic differences between comparison groups in withdrawals or exclusion of participant from the study[47]). Selection bias affects all observational studies to varying degrees because of lack of randomisation. Selection bias may be apparent when the method or timing of an intervention is determined by the patient or health professional. Women who opt for frequent surveillance may have a higher baseline risk of recurrence (ie. family history of breast cancer, higher staging of primary tumor) than

women who opt for less frequent monitoring and thus bias results against a survival benefit.[48] Similarly, interpretation of survival rates between exposed and non-exposed groups are difficult to interpret if there is selective loss to follow-up or duration of follow-up is not well reported. For example, even if surveillance groups were similar at study commencement, over time biases may be introduced as lower risk women become less likely to comply with frequent surveillance and drop out of that study arm. Intention to treat analysis was not a common feature of these studies, and this was reflected in the fact that duration of follow-up itself was under-reported. The above biases were most relevant to the alternative levels of surveillance studies.

For the studies that compared survival in patients with early vs late detection of a second breast cancer event, 9 of 14 studies calculated survival from the detection of the primary breast cancer,[29-31,37,38,41-44] three studies calculated survival from the detection of tumor recurrence,[33,35,36] and the method of survival calculation was not explicitly stated in two studies.[34,46] Of the four studies that compared alternative levels of mammographic surveillance,[32,39,40,45]only one calculated survival from the detection of the primary breast cancer.[32]

Interpretation of survival differences requires consideration of the potential for lead and length time bias. **Lead time bias** may occur when survival is measured from the time of a recurrence. It is defined as the interval between when the cancer was detected by screening and when it would have been detected clinically without screening[49]. Apparent survival gains may reflect the patient living longer with the knowledge of a recurrence, rather than a true increase in longevity. This can be overcome by estimating overall survival from time of primary tumor rather than from relapse. Lead time bias is relatively easy to control for, and is relevant to both the early vs late detection and the alternative frequency of surveillance studies. Of the total of ten cohort studies that calculated survival from primary breast cancer, only one good-moderate[29] and three moderate quality studies demonstrated a survival benefit.[30,32,38]

Length bias reflects the possibility that screening will detect cancers that have a longer pre-clinical phase and better prognosis, whilst faster growing more aggressive tumors

will present clinically in the interval between screening examinations, and is much harder to address. Length bias is most relevant for early vs late detection studies, and is very difficult to control for in the absence of randomisation. Houssami 2009 alone used sensitivity analyses to test the effects of screening detecting a higher proportion of asymptomatic cancers with a smaller prior probability of proving fatal. The sensitivity analysis hypothesised two latent tumor populations, one of which is more prone to screen detection. The sensitivity analyses re-estimated the HRs of breast cancer death for asymptomatic versus symptomatic tumors, over a range of plausible magnitudes of length bias.[29] They reported a survival benefit from the asymptomatic detection of a second breast cancer event compared with symptomatic detection (HR_{adj} range 0.53-0.73).[29]

Only seven of the 18 cohort studies[29,32,34,35,37,38,45] attempted to adjust for **confounding**. A good example of adjustment for confounding is a large population based study by Paszat et al, who assessed the effect of surveillance mammography on time to breast cancer death.[32] They adjusted for age, stage, type of local treatment, and receipt of adjuvant treatment, and showed that receipt of surveillance MMG was associated with a reduced risk of breast cancer death (HR 0.28, 95% CI 0.22, 0.37). Other potential confounders that could have been included were histologic type, grade and HER2 receptor status of the primary tumour, lymphovascular invasion, extent of axillary staging, family history, and presence or absence of BRCA mutations. The most common confounders adjusted for in the other six studies were age, stage of primary breast cancer at diagnosis, and time interval between first and second breast cancer diagnoses.

Using the critical appraisal tool, the cohort studies were considered to be of variable quality with no study assessed as good quality (table 3). Quality ratings of the cohort studies and demonstrated survival benefit are shown in (table 4), which provide no definitive indication of the effects of follow-up mammography on survival.

Table 3: Quality rating of included studies

Author	Study type	Reporting	External validity	Internal validity - bias	Internal validity - confounding	Overall score[ł]	Bias and confounding score	Quality rating
Ciatto 1990	Cohort	9	3	5	4	21	9	Good-moderate
Houssami 2009	Cohort	9	3	4	4	20	8	Good-moderate
Lash 2005	Case-control	8	3	5	3	19	8	Moderate
Lash 2006	Case-control	8	3	4	3	18	7	Moderate
Lash 2007	Case-control	8	3	4	3	18	7	Moderate
Ciatto 2004	Cohort	9	2	4	4	19	8	Moderate
Grosse 1997	Cohort	8	3	4	3	18	7	Moderate
Montgomery 2007	Cohort	8	3	4	3	18	7	Moderate
Doyle 2001	Cohort	7	3	4	3	17	7	Moderate
te Boekhorst 2001	Cohort	8	3	4	2	17	6	Moderate
Yau 2008	Cohort	7	3	5	2	17	7	Moderate
Ciatto 1984	Cohort	7	3	3	3	16	6	Moderate
Ciatto 1985	Cohort	7	3	3	3	16	6	Moderate
Kaas 2001 FREQUENCY	Cohort	6	3	5	2	16	7	Moderate
Paszat 2009	Cohort	8	1	4	3	16	7	Moderate
Wagman 1991	Cohort	6	3	4	2	15	6	Moderate - poor
Orel 1993	Cohort	6	2	4	3	15	7	Moderate - poor
Kaas 2001 DETECTION	Cohort	4	3	5	2	14	7	Moderate - poor
Tomin 1987 DETECTION	Cohort	6	1	3	4	14	7	Moderate - poor
Tomin 1987 ADHERENCE	Cohort	6	1	3 *	3	14*	6.5*	Moderate - poor
Robinson 1993	Cohort	4	3	3**	2	12.5**	5**	Moderate - poor
Hussain 1995	Cohort	7	2	2	2	13	4	Poor
Krengli 1993	Cohort	2	0	1	0	3	1	Poor

*Qu 19 (on adherence) is not applicable as this was the focus of the paper, hence overall score is 13/25 bias score is 6/12, which we have rounded up to 14/26 and 6.5/13 to allow comparison between studies.

**Qu19 (on adherence) is not applicable as this was the focus of the paper, hence overall score is 12/25, bias score is 5/12, which we have rounded up to 12.5/26 and 5/13 to allow comparison between studies.

ł Full checklist has 27 questions. We have dropped question 27 from confounding section so denominator is 26 (see text)

Table 4: Quality of included papers and survival benefit

	Survival benefit		
Study design and quality	Yes	No	Not reported[a]
Case-control			
Moderate	Lash[25-27]	n/a[b]	n/a
Cohort			
Early vs late detection			
Good-moderate	Houssami[29] (BCSS)	Ciatto[37]	n/a
Moderate	Ciatto[38], Doyle[34], Montgomery[30]	Ciatto[42], Ciatto[43], Grosse[33], te Boekhorst[41]	Yau[31]
Moderate-poor	Kaas[39] DETECTION (BCSS), Tomin[45] DETECTION	Orel[35]	Wagman[44]
Poor	Hussain[36], Krengli[46]	n/a	n/a
Alternative levels of surveillance			
Moderate	Kaas[39] FREQUENCY (BCSS), Paszat[32] (BCSS)	n/a	n/a
Moderate-poor	n/a	Robinson[40], Tomin[45] ADHERENCE	n/a

[a] Not reported = no report of breast cancer specific survival or overall survival; [b] n/a= no papers in this category,
BCSS = breast cancer specific survival (rather than overall survival)

DISCUSSION

Limitations in the evidence base

RCTs have been published that examine the impact of different follow-up programs on mortality, morbidity and quality of life from early detection of distant metastases,[50] but we were unable to identify any RCTs that focussed on the impact on survival of early detection of relapse within the breast.

The main limitation of observational studies is that group allocation is not random. This reduces the likelihood that the groups are alike in all relevant aspects other than the intervention of interest (i.e. follow-up). There were no studies that met our two predetermined criteria for a good quality study. There was heterogeneity between studies in terms of patient, tumor and treatment characteristics, and different survival outcome measures. Limitations in the evidence base include small sample sizes in some studies, bias inherent to the observational study design, lead time and length bias, and lack of adjustment for confounding. Less than half of the cohort studies adjusted for potential confounders, and in many older studies, data on more recently recognised potential confounders were not routinely collected.

Overcoming limitations in the evidence base

Randomised controlled trials (RCTs) can provide an unconfounded estimate of effect, and are a common and relatively easy source for effectiveness and cost data. However clinical trials are limited to the patients, interventions and time span of the trial, and are sometimes not feasible. Potential barriers to performing RCTs in follow-up imaging in cancer include difficulty with patient accrual (for example, if patients have a 50% chance of receiving follow-up that is less frequent than current guideline recommendations they are unlikely to consent to be in a trial), the large sample size and long follow-up (and thus cost implications) required to demonstrate a significant difference between alternative follow-up programs (e.g. different frequencies of mammographic follow-up) or different patient sub-groups (e.g. high and low risk), and the possibility that all important differences may not be captured during the finite duration of the trial. So while we clearly need better evidence for the value of follow-up mammography,[32] and given that RCTs are unlikely to occur, how can we guide clinical practice in the meantime?

Improving the quality of our observational studies

We can strengthen our observational studies by addressing the key components of reporting, external validity and internal validity (bias and confounding). Firstly through better design of studies that attempt to mitigate the risk of bias, as discussed in Table 5. Secondly, often the study design is satisfactory but the reporting of study characteristics is not. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement is a 22 point checklist designed to improve the clarity and transparency of reporting of observational studies. There are 18 items common to cohort, case-control and cross-sectional study designs, and 4 items specific to each study type pertaining to description of patient selection, statistical methods, descriptive and outcome data.[51]

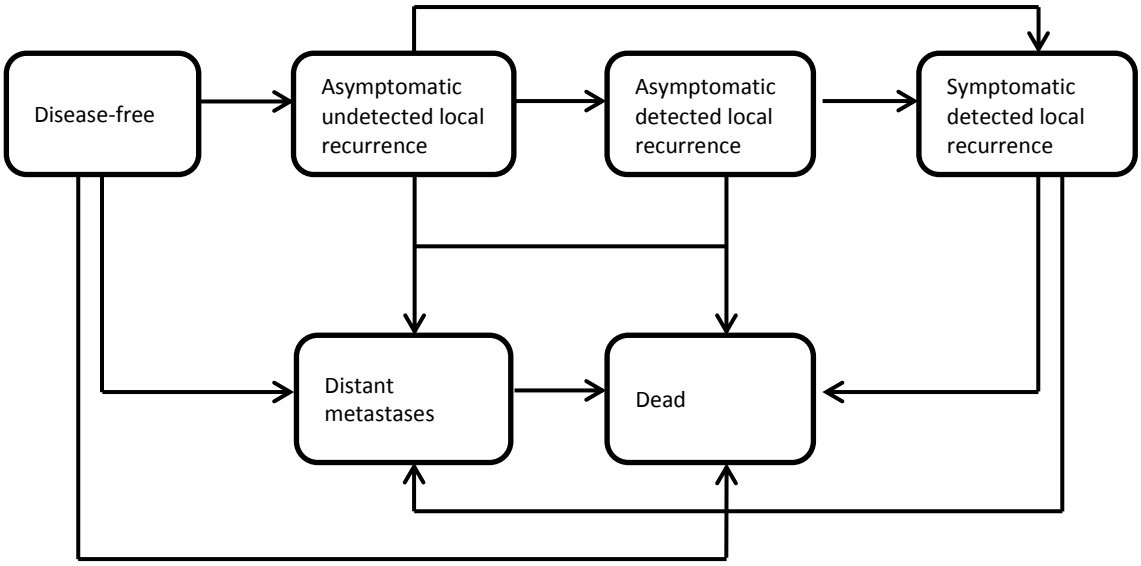
Table 5: Improving the design of observational studies

Study type	Bias identified in papers included in our appraisal	Possible methods to improve the design of observational studies.
Case-control	Integrity of information on exposure	Complete ascertainment of follow-up mammography eg. data linkage, mammography attendance records, GP casenotes etc.
	Matching	Baseline characteristics table include other relevant factors (eg. family history, tumor genetic profile) or other exposures (eg. preventive health care) which may account for the observed effect. Select controls that match cases on these factors.
	Self-selection bias	Information on patients who choose not to attend follow-up or participate in the study.
Cohort	Selection bias	Matching, stratification of known confounders, control for confounders in analysis
	Attrition bias	Minimise losses to follow-up
	Lead time bias	Estimate survival from time of primary breast cancer not from recurrence
	Length bias	Focus on alternative study designs to cohort studies that compare alternative methods of recurrence detection. Sensitivity analysis can provide a range of hazard ratios to estimate benefit of follow-up. Health economic modelling can be used to estimate clinical and cost-effectiveness of alternative follow-up strategies.

Decision analytic modelling

Decision analytic modelling is commonly used in health economic evaluation in situations where few data are available or where trials are difficult or impossible to conduct.[52] Modelling enables data to be synthesised to describe disease progression over the time horizon of the model, commonly over the remainder of patients' lifetime. The modelling process involves the specification of a model structure comprising a set of mutually exclusive health states that represent events that patients may experience (Figure 2). Parameters are estimated that describe the movement of patients between the health states over time, which can be estimated using experimental, observational or routinely collected data.

Figure 2: Potential model structure for evaluation of surveillance in breast cancer.



Costs and quality of life weights are applied to the time spent in each health state. Summing these weights over all of the models' health states over the time horizon of the model (e.g 40 years for a cohort of 60 year old patients), provides an estimate of the total costs and quality adjusted life years (QALYs) for the original patient cohort.

In the case of cancer surveillance, we cannot observe the movement of patients to and from asymptomatic recurrence. These parameters need to be calibrated. Calibration involves fitting the model's input parameters so that the model's outputs match some observed data as closely as possible.[53]

The ideal dataset for calibrating an early breast cancer surveillance model would comprise of women with early breast cancer, followed up with respect to their surveillance pathway (i.e. timing, type, and results of any surveillance) and the timing and type of any recurrent events and deaths. The accuracy of the model predictions of recurrence events and deaths would be assessed against the observed values in the dataset, as well as against measures of variation within groups.

The calibration model can then be analysed to estimate the cost-effectiveness of alternative surveillance strategies, by patient sub-group, including appropriate sensitivity analysis to represent uncertainty around data and model specifications.[54]

CONCLUSION

This review of the evidence around the effectiveness of follow-up in general, and mammographic surveillance in particular, has identified a complete absence of randomised clinical trials in this area. The range of observational studies that have been reported are subject to a range of potential biases that limit their interpretation, and thus do not provide sufficient evidence on which to make informed decisions regarding the development of personalised surveillance strategies. We should be striving to improve the quality of our observational studies and standardise their reporting, as well as to embrace alternative research techniques, such as decision analytic modelling, to guide our practice in the likely continued absence of RCTs in this field. Modelling can inform clinical decisions, but also identify areas for further primary research targeted at patient subgroups for whom the appropriate strategy is most uncertain.[55] As better prognostic factors are identified in breast cancer, the potential benefits of tailoring follow-up mammography to the risk of recurrence will increase and so we should be developing analytic methods that will guide clinical practice in an evidence-based but also pragmatic manner.

APPENDIX 1 - Search string for PubMed

1. (early breast cancer[tw] OR (breast neoplasms[mh] OR breast neoplasm*[tw] OR breast cancer*[tw] OR breast carcinoma*[tw] OR breast adenocarcinoma*[tw] OR breast malignanc*[tw] OR breast metastas*[tw] OR breast micrometastas*[tw] OR breast oncolog*[tw] OR paget*[tw] OR breast sarcoma*[tw] OR breast tumor*[tw] OR breast tumour*[tw] OR mammary tumor*[tw] OR mammary tumour*[tw] OR mammary cancer*[tw] OR mammary neoplasm*[tw] OR mammary adenocarcinoma*[tw] OR mammary carcinoma*[tw] OR mammary sarcoma*[tw]))
2. (neoplasm recurrence, local[mh] OR local recurrence[tw] OR loco-regional recurrence[tw] OR locoregional recurrence[tw] OR regional recurrence[tw] OR breast recurrence[tw] OR (axillary[tw] AND recurrence[tw]) OR second primary breast[tw] OR contralateral breast[tw] OR ipsilateral breast[tw])
3. (Clinical Trial[pt] OR Meta-Analysis[pt] OR Randomized Controlled Trial[pt] OR Review[pt] OR follow-up studies[mh] OR follow up[tw] OR follow-up[tw])
4. (mammography[mh] OR mammograph*[tw] OR mammogram*[tw] OR physical examination[mh] OR physical exam*[tw] OR Palpat*[tw] OR breast Self Examination*[tw] OR bse[tw] OR Diagnostic Tests, Routine[mh] OR Routine Diagnostic Test*[tw] OR surveillance*[tw] OR monitoring*[tw] OR follow up*[tw] OR follow-up*[tw] OR follow-up*[tw] OR diagnosis[sh] OR diagnos*[tw] OR radiograph*[tw] OR detect*[tw] OR symptomatic[tw] OR asymptomatic[tw] OR early[tw] OR late[tw] OR interval*[tw] OR time factor*[tw] OR time factors[mh])
5. (disease-free survival[mh] OR disease free[tw] OR event free[tw] OR progression free[tw] OR relapse free[tw] OR recurrence free[tw] OR time factors[mh] OR time[tw])
6. Method of detection of a second breast cancer event within the breast = 1 AND 2 AND 3 AND 4
7. Timing of detection of a second breast cancer event within the breast = 1 AND 2 AND 3 AND 5

APPENDIX 2 – Downs and Black checklist[8]

Appendix

Checklist for measuring study quality

Reporting

1. *Is the hypothesis/aim/objective of the study clearly described?*

yes	1
no	0

2. *Are the main outcomes to be measured clearly described in the Introduction or Methods section?*
If the main outcomes are first mentioned in the Results section, the question should be answered no.

yes	1
no	0

3. *Are the characteristics of the patients included in the study clearly described ?*
In cohort studies and trials, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.

yes	1
no	0

4. *Are the interventions of interest clearly described?*
Treatments and placebo (where relevant) that are to be compared should be clearly described.

yes	1
no	0

5. *Are the distributions of principal confounders in each group of subjects to be compared clearly described?*
A list of principal confounders is provided.

yes	2
partially	1
no	0

6. *Are the main findings of the study clearly described?*
Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions. (This question does not cover statistical tests which are considered below).

yes	1
no	0

7. *Does the study provide estimates of the random variability in the data for the main outcomes?*
In non normally distributed data the inter-quartile range of results should be reported. In normally distributed data the standard error, standard deviation or confidence intervals should be reported. If the distribution of the data is not described, it must be assumed that the estimates used were appropriate and the question should be answered yes.

yes	1
no	0

8. *Have all important adverse events that may be a consequence of the intervention been reported?*
This should be answered yes if the study demonstrates that there was a comprehensive attempt to measure adverse events. (A list of possible adverse events is provided).

yes	1
no	0

9. *Have the characteristics of patients lost to follow-up been described?*
This should be answered yes where there were no losses to follow-up or where losses to follow-up were so small that findings would be unaffected by their inclusion. This should be answered no where a study does not report the number of patients lost to follow-up.

yes	1
no	0

10. *Have actual probability values been reported(e.g. 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?*

yes	1
no	0

External validity

All the following criteria attempt to address the representativeness of the findings of the study and whether they may be generalised to the population from which the study subjects were derived.

11. *Were the subjects asked to participate in the study representative of the entire population from which they were recruited?*
The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source population, an unselected sample of consecutive patients, or a random sample. Random sampling is only feasible where a list of all members of the relevant

population exists. Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.

yes	1
no	0
unable to determine	0

12. *Were those subjects who were prepared to participate representative of the entire population from which they were recruited?*
The proportion of those asked who agreed should be stated. Validation that the sample was representative would include demonstrating that the distribution of the main confounding factors was the same in the study sample and the source population.

yes	1
no	0
unable to determine	0

13. *Were the staff, places, and facilities where the patients were treated, representative of the treatment the majority of patients receive?*
For the question to be answered yes the study should demonstrate that the intervention was representative of that in use in the source population. The question should be answered no if, for example, the intervention was undertaken in a specialist centre unrepresentative of the hospitals most of the source population would attend.

yes	1
no	0
unable to determine	0

Internal validity - bias

14. *Was an attempt made to blind study subjects to the intervention they have received ?*
For studies where the patients would have no way of knowing which intervention they received, this should be answered yes.

yes	1
no	0
unable to determine	0

15. *Was an attempt made to blind those measuring the main outcomes of the intervention?*

yes	1
no	0
unable to determine	0

16. *If any of the results of the study were based on "data dredging", was this made clear?*
Any analyses that had not been planned at the outset of the study should be clearly indicated. If no retrospective unplanned subgroup analyses were reported, then answer yes.

yes	1
no	0
unable to determine	0

17. *In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls ?*
Where follow-up was the same for all study patients the answer should yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.

yes	1
no	0
unable to determine	0

18. *Were the statistical tests used to assess the main outcomes appropriate?*
The statistical techniques used must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.

yes	1
no	0
unable to determine	0

19. *Was compliance with the intervention/s reliable?*
Where there was non compliance with the allocated treatment or where there was contamination of one group, the question should be answered no. For studies where the effect of any misclassification was likely to bias any association to the null, the question should be answered yes.

yes	1
no	0
unable to determine	0

20. *Were the main outcome measures used accurate (valid and reliable)?*

For studies where the outcome measures are clearly described, the question should be answered yes. For studies which refer to other work or that demonstrates the outcome measures are accurate, the question should be answered as yes.

yes	1
no	0
unable to determine	0

Internal validity - confounding (selection bias)

21. *Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population?*
For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case-control studies where there is no information concerning the source of patients included in the study.

yes	1
no	0
unable to determine	0

22. *Were study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time?*
For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.

yes	1
no	0
unable to determine	0

23. *Were study subjects randomised to intervention groups?*
Studies which state that subjects wererandomised should be answered yes except where method of randomisation would not ensure random allocation. For example alternate allocation would score no because it is predictable.

yes	1
no	0
unable to determine	0

24. *Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable?*
All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.

yes	1
no	0
unable to determine	0

25. *Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?*
This question should be answered no for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomised studies if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses the question should be answered as no.

yes	1
no	0
unable to determine	0

26. *Were losses of patients to follow-up taken into account?*
If the numbers of patients lost to follow-up are not reported, the question should be answered as unable to determine. If the proportion lost to follow-up was too small to affect the main findings, the question should be answered yes.

yes	1
no	0
unable to determine	0

Power

27. *Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?*
Sample sizes have been calculated to detect a difference of x% and y%.

	Size of *smallest* intervention group	
A	$<n_1$	0
B	n_1–n_2	1
C	n_3–n_4	2
D	n_5–n_6	3
E	n_7–n_8	4
F	n_9+	5

APPENDIX 3: Levels of evidence for intervention studies

Adapted from NHMRC Evidence Hierarchy by Merlin et al.[24]

Level	Intervention studies
I	A systematic review of level II studies
II	A randomised controlled trial
III-I	A pseudorandomised controlled trial (i.e. alternate allocation or some other method)
III-2	A comparative study with concurrent controls: • Non-randomised experimental trial • Cohort study • Case-control study • Interrupted time series with a control group
III-3	A comparative study without concurrent controls: • Historical control study • Two or more single arm study • Interrupted time series without a parallel control group
IV	Case series with either post-test or pre-test/post-test outcomes

REFERENCES

1. Elshaug A, Bessen T, Moss J, Hiller J. Addressing “waste” in diagnostic imaging: some implications of comparative effectiveness research. J Am Coll Radiol. 2010 Aug; 7(8):603-13
2. Buist D, Abraham L, Barlow W, Krishnaraj A, Holdridge R, Sickles E, Carney P, Kerlikowske K, Geller B; Breast Cancer Surveillance Consortium. Diagnosis of second breast cancer events after initial diagnosis of early stage breast cancer. Breast Cancer Res Treat. 2010 Dec; 124(3): 863-73
3. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
4. National Institute for Health and Clinical Excellence. Early and locally advanced breast cancer: diagnosis and treatment. National Collaborating Centre for Cancer, Cardiff, Wales, 2009.
5. Scottish Intercollegiate Guidelines Network. Management of breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network, Edinburgh, Scotland, 2005.
6. Khatcheressian J, Wolff A, Smith T, Grunfeld E, Muss H, Vogel V, Halberg F, Somerfield M, Davidson N; American Society of Clinical Oncology. American Society of Clinical Oncology 2006 update of the breast cancer follow-up and management guidelines in the adjuvant setting. J Clin Oncol 2006 Nov 1;24 (31) :5091-7.
7. Grunfeld E, Dhesy-Thind S, Levine M; Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment for breast cancer (summary of the 2005 update). CMAJ. 2005 May 10;172(10) :1319-20.
8. Downs S, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998 Jun; 52(6): 377-84

9. Deeks J, Dinnes J, D'Amico R, Sowden A, Sakarovitch C, Song F, Petticrew M, Altman D; International Stroke Trial Collaborative Group; European Carotid Surgery Trial Collaborative Group. Evaluating non-randomised intervention studies. Health Technol Assess. 2003; 7 (27): iii-x, 1-173
10. Doidge J, Merlin T, Liufu Z, Tamblyn D, Jia LY, Hiller J. Review of interim funded service: vertebroplasty and new review of kyphoplasty. Commonwealth of Australia, Canberra, ACT, 2011
11. Moher D, Liberati A, Tetzlaff J, Altman D; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21; 6(7):e1000097.
12. Jack W, Kerr G, Kunkler I. Long-term follow-up after breast conservation: the Edinburgh experience. Breast. 1998 April;7(2):80-4.
13. Kollias J, Evans A, Wilson A, Ellis I, Elston C, Blamey R. Value of contralateral surveillance mammography for primary breast cancer follow-up. World J Surg. 2000 Aug;24(8):983-7; discussion 988-9.
14. Kindler M, Steinhoff G. [Retrospective investigation of validity of different methods for early detection of recurrence and metastases of mammary cancer]. Archiv fur Geschwulstforschung. 1989; 59(4):289-297. German.
15. Stierer M, Rosen H. [Influence of early diagnosis on prognosis of recurrent breast cancer. Wiener Klinische Wochenschrift]. 1989; 101(4):138-142. German
16. Imoto S, Jitsuiki Y. Detection of the first recurrence during intensive follow-up of breast cancer patients. Jpn J Clin Oncol. 1998 Oct;28(10):597-600.
17. Broyn T, Froyen J. Evaluation of routine follow-up after surgery for breast carcinoma. Acta Chir Scand. 1982;148(5):401-4.
18. Perrone M, Musolino A, Michiara M, Di Blasio B, Bella M, Franciosi V, Cocconi G, Camisa R, Todeschini R, Cascinu S. Early detection of recurrences in the follow-up of primary breast cancer in an asymptomatic or symptomatic phase. Tumori. 2004 May-Jun;90(3):276-9.
19. Zwaveling A, Albers G, Felthuis W, Hermans J. An evaluation of routine follow-up for detection of breast cancer recurrences. J Surg Oncol. 1987 Mar;34(3):194-7.

20. Dewar J, Kerr G. Value of routine follow up of women treated for early carcinoma of the breast. Br Med J (Clin Res Ed). 1985 Nov 23; 291(6507):1464-7.
21. Tuli R, Christodouleas J, Roberts L, Deol S, Usuki K, Frassica D, Rosenberg A. Prognostic indicators following ipsilateral tumor recurrence in patients treated with breast-conserving therapy. Am J Surg. 2009 Oct;198(4):557-61.
22. Voogd A, van Tienhoven G, Peterse H, Crommelin M, Rutgers E, van de Velde C, van Geel B, Slot A, Rodrigus P, Jobsen J, von Meyenfeldt M, Coebergh J. Local recurrence after breast conservation therapy for early stage breast carcinoma: detection, treatment, and outcome in 266 patients. Dutch Study Group on Local Recurrence after Breast Conservation (BORST). Cancer. 1999 Jan 15;85(2):437-46.
23. Lu W, Schaapveld M, Jansen L, Bagherzadegan E, Sahinovic M, Baas P, Hanssen L, van der Mijle, Brandenburg J, Wiggers T, De Bock G. The value of surveillance mammography of the contralateral breast in patients with a history of breast cancer. Eur J Cancer. 2009 Nov;45(17):3000-7.
24. Merlin T, Weston A, Tooher R. Extending an evidence hierarchy to include topics other than treatment: revising the Australian 'levels of evidence'. BMC Med Res Methodol. 2009 Jun 11; 9:34.doi: 10.1186/1471-2288-9-34
25. Lash T, Clough-Gorr K, Silliman R. Reduced rates of cancer-related worries and mortality associated with guideline surveillance after breast cancer therapy. Breast Cancer Res Treat. 2005 Jan;89(1):61-7.
26. Lash T, Fox M, Silliman R. Reduced mortality rate associated with annual mammograms after breast cancer therapy. Breast J. 2006 Jan-Feb;12(1):2-6.
27. Lash T, Fox M, Buist D, Wei F, Field T, Frost F, Geiger A, Quinn V, Yood M, Silliman R. Mammography surveillance and mortality in older breast cancer survivors. J Clin Oncol. 2007 Jul 20;25(21):3001-6.
28. Gordis L. Epidemiology. 3rd ed. Philadelphia: Elsevier; 2004
29. Houssami N, Ciatto S, Martinelli F, Bonardi R, Duffy S. Early detection of second breast cancers improves prognosis in breast cancer survivors. Ann Oncol. 2009 Sep;20(9):1505-10.

30. Montgomery D, Krupa K, Jack W, Kerr G, Kunkler I, Thomas J, Dixon J. Changing pattern of the detection of locoregional relapse in breast cancer: the Edinburgh experience. Br J Cancer. 2007 Jun 18;96(12):1802-7.
31. Yau T, Sze H, Soong I, Wong W, Chan K, Chang A, Lau K, Lee A. Surveillance mammography after breast conservation therapy in Hong Kong: effectiveness and feasibility of risk-adapted approach. Breast. 2008 Apr;17(2):132-7.
32. Paszat L, Sutradhar R, Grunfeld E, Gainford C, Benk V, Bondy S, Coyle D, Holloway C, Sawka C, Shumak R, Vallis K, van Walraven C. Outcomes of surveillance mammography after treatment of primary breast cancer: a population-based case series. Breast Cancer Res Treat. 2009 Mar;114(1):169-78.
33. Grosse A, Schreer I, Frischbier H, Maass H, Loening T, Bahnsen J. Results of breast conserving therapy for early breast cancer and the role of mammographic follow-up. Int J Radiat Oncol Biol Phys. 1997 Jul 1 ;38(4):761-7.
34. Doyle T, Schultz D, Peters C, Harris E, Solin L. Long-term results of local recurrence after breast conservation treatment for invasive breast cancer. Int J Radiat Oncol Biol Phys. 2001 Sep 1; 51(1):74-80.
35. Orel S, Fowble B, Solin L, Schultz D, Conant E, Troupin R. Breast cancer recurrence after lumpectomy and radiation therapy for early-stage disease: prognostic significance of detection method. Radiology. 1993 Jul;188(1):189-94.
36. Hussain S, Gui G, Lee K, Plowman P, Gilmore O, Allum W. Detection of loco-regional recurrence after breast-conserving surgery and radiotherapy. J R Coll Surg Edinb 1995; 40(3):163-6.
37. Ciatto S, Ambrogetti D, Bonardi R, Bravetti P, Del Turco M. Prognostic impact of early detection of contralateral primary breast cancer. Tumori. 1990 Aug 31;76(4):370-3.
38. Ciatto S, Miccinesi G, Zappa M. Prognostic impact of the early detection of metachronous contralateral breast cancer. Eur J Cancer. 2004 Jul;40(10):1496-501.
39. Kaas R, Hart A, Besnard A, Peterse J, Rutgers E. Impact of mammographic interval on stage and survival after the diagnosis of contralateral breast cancer. Br J Surg. 2001 Jan;88(1):123-7.

40. Robinson E, Rennert G, Bar-Deroma R, Dori D, Neugut A. The pattern of diagnosis of a second primary tumor in the breast. Breast Cancer Res Treat. 1993;25(3):211-5.
41. te Boekhorst D, Peer N, van der Sluis R, Wobbes T, Ruers T. Periodic follow-up after breast cancer and the effect on survival. Eur J Surg. 2001 Jul;167(7):490-6.
42. Ciatto S, Rosselli Del Turco M, Pacini P, De Luca Cardillo C, Bastiani P, Bravetti P. Early detection of local recurrences in the follow-up of primary breast cancer. Tumori. 1984 Apr 30;70(2):179-83.
43. Ciatto S, Rosselli Del Turco M, Pacini P, Mustacchi G, Simonis M, Sismondi P, Giardina G, Belsanti V, Aristei C, Molino A, et al. Early detection of breast cancer recurrences through periodic follow-up - is it useless? Tumori. 1985 Aug 31 ;71(4):325-9.
44. Wagman L, Sanders R, Terz J, Beatty J, Kokal W, Riihimaki D. The value of symptom directed evaluation in the surveillance for recurrence of carcinoma of the breast. Surg Gynecol Obstet. 1991 Mar;172(3):191-6.
45. Tomin R, Donegan W. Screening for recurrent breast cancer – its effectiveness and prognostic value. J Clin Oncol. 1987 Jan;5(1):62-7.
46. Krengli M, Pastore G, Maffei S. [The importance of the follow-up in patients operated on for breast cancer. A retrospective analysis of 2482 cases]. Minerva Med. 1993 Jul-Aug;84(7-8):409-15. Italian.
47. Higgins J, Green S (editors).Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
48. Roder D, Houssami N, Farshid G, Gill G, Luke C, Downey P, Beckmann K, Iosifidis P, Grieve L, Willliamson L. Population screening and intensity of screening are associated with reduced breast cancer mortality: evidence of efficacy of mammography screening in Australia. Breast Cancer Res Treat. 2008 Apr; 108 (3): 409-416.
49. Habbema J, van Oortmarssen G, van Putten D. An analysis of survival differences between clinically and screen-detected cancer patients. Stat Med. 1983 Apr-Jun; 2 (2): 279-285

50. Rojas M, Telaro E, Moschetti I, Coe L, Fossati R, Liberati A, Del Turco Rosselli M. Follow-up strategies for women treated for early breast cancer. Cochrane Database Systematic Reviews. 2009: CD001768

51. von Elm E, Altman D, Egger M, Pocock S, Gotzsche P, Vandenbroucke J; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007 Oct 16; 4(10): e296.

52. Buxton M, Drummond M, Van Hout B, Prince R, Sheldon T, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. Health Econ. 1997 May-Jun; 6(3) : 217-27

53. Vanni T, Karnon J, Madan J, White R, Edmunds W, Foss A, Legood R. Calibrating models in economic evaluation: a seven step approach. Pharmacoeconomics. 2011 Jan: 29(1):35-49.

54. Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies, and convergence criteria. Pharmacoeconomics. 2011 Jan; 29(1):51-62

55. Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? Pharmacoeconomics. 2000 May;17(5):445-59.

CHAPTER 2 EXTENSION: Cross validation of our systematic review

Please note:

- References 55 and below refer to the papers included in our systematic review publication.
- References 90-92 refer to papers not included in our systematic review publication. These papers are included in the reference list at the end of the thesis.

The above paper was submitted to the British Medical Journal in 2011 and was sent out for peer review. It was eventually rejected, as the reviewers identified that a few days prior to my submission, the NHS had published a Health Technology Assessment (HTA) to address "The clinical effectiveness and cost effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews, registry databases analyses and economic evaluation".[90]

At the time of writing this extension, the 2011 HTA is the most recent and largest systematic review published on this topic. We present below the key features of the HTA systematic review and compare those with our own study. We have used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) appraisal tool to identify and comment on the key components of the HTA systematic review.[11]

The purpose of the HTA was to determine the clinical effectiveness of differing surveillance mammography regimens for detection of new disease within the breast after treatment for primary breast cancer, to inform a later economic evaluation. The participants, interventions and comparators were identical to our own, but the outcome measures were more extensive. The primary outcome measure was overall survival, and the secondary outcomes were ipsilateral breast tumor recurrence (IBTR) and/or metachronous contralateral breast cancer (MCBC) event rate. In addition, the HTA included studies that reported any of the following outcomes: quality of life, harm from mammography (described in the review as including radiation risk, pain, false-positive results), uptake of mammography or economic data (specifically costs to the UK National Health Service and to patients) The economic data to be reviewed included resource use, costs arising from investigating true-positives and false-negatives, as well as investigating incidental findings. Estimates of cost-effectiveness were planned to be

taken as defined in any given study, and if possible, the incremental cost per life year or quality adjusted life year would be estimated if this could be calculated from the data reported.

Both randomised (RCTs) and non-randomised comparative studies that compared differing surveillance mammography regimens and alternative breast cancer follow-up regimens, for women diagnosed with breast cancer from 1990 or later, were eligible for inclusion. Also included were retrospective or prospective cohort studies if they contained 100 or more participants with follow-up for at least a median of 5 years. Eight studies met their inclusion criteria, of which six were retrospective cohort studies, and two were prospective cohort studies (see below).

The authors searched for primary studies, ongoing trials, grey literature and relevant websites. Primary studies were identified through the electronic databases of MEDLINE, MEDLINE In-Process, EMBASE, BIOSIS, Science Citation Index, CANCERLIT and Cochrane Central Register of Controlled Trials. Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects and the HTA Database were searched for reports of evidence synthesis. Trials were sought from Current Controlled Trials, Clinical Trials, WHO International Clinical Trials Registry Platform, NCI Clinical Trials Database, National Research Register Archive, and NIHR Portfolio Database. Websites searched included National Cancer Institute, National Comprehensive Cancer Network, CancerWEB, Breast Cancer Surveillance Consortium and the National Library for Health, and selected websites for oncology and radiology professions. The searches were run from 1990 to June 2009, and were restricted to full text papers, but without any language restriction. In contrast, our search was limited to five electronic databases including PubMed, Scopus, the Cochrane Library, Ovid Medline and Web of Knowledge, but included studies from 1980 to August 2010.

In the HTA, two reviewers independently assessed papers to determine if they met the inclusion criteria. An economist evaluated economic evaluations or cost analysis studies. A 10% check of inclusion assessment was performed, with any disagreement resolved by

a third party. One reviewer extracted details of study design, participant characteristics, description of the intervention and outcome data using a data extraction form, and a second reviewer independently validated the data extraction. Any uncertainty was resolved by a third party. In our systematic review, one reviewer (TB) reviewed all titles and abstracts and selected studies that appeared to fulfil the pre-defined inclusion criteria. When in doubt, the decision to include or exclude a paper was made by a second reviewer (JK). The full text papers were then retrieved and appraised to determine suitability. The reviewed papers were categorised by one reviewer (TB) by study design, and data extracted according to study design and mortality outcome. The data was then independently assessed by a second reviewer (JK).

In the HTA, it was planned that the quality of the included studies would be assessed using three separate tools depending on the design of the studies identified. The authors planned to use an adapted version of the Cochrane Collaboration's tool for assessing risk of bias in individual RCTs; an adapted tool from the Review Body for Interventional Procedures for assessment of non-randomised studies; and the NHS Economic Evaluation Database Handbook for assessment of any economic evaluation. In our study two reviewers (TB, TM) assessed the included papers using the validated Downs and Black checklist.[8] This single tool assesses the methodological quality of both randomised and nonrandomised studies of health care interventions, was specifically developed for use in systematic reviews, and appraises reporting in addition to methodological quality (external validity, risk of bias and confounding) of individual studies.

The authors of the HTA planned to perform a meta-analysis of included studies, favouring intention to treat over per-protocol results for the analysis. They planned to derive a pooled HR for time-to-event outcomes (e.g. survival, recurrence); and to use a standardised mean difference to combine quality of life scores for data on harms of mammography, adverse events and quality of life. The authors state that no quantitative synthesis was planned for the economic outcome data. In our study, a narrative synthesis was planned where included papers would be divided into study

type, analysed by mortality outcome and critically appraised. The decision as to whether to perform a meta-analysis would be informed by the nature of the reported data in the included studies.

The HTA search yielded 2849 titles and abstracts, of which 422 were selected for full text assessment. Of these, 407 were excluded, 7 were unavailable, leaving only 8 studies which met the inclusion criteria for the review. The authors did not identify any RCTs, or any studies that reported data on quality of life, harm of mammography, uptake of mammography or economic data.

The 8 studies included in the HTA were all classified as "cohort" studies. We included 6 of these studies, but classified 3 of them as cohort studies[30-32] and the three papers by Lash et al[25-27] as nested case-control studies. The 2 studies included in the HTA that we did not include, were excluded because one provided local recurrence-free survival rather than overall or breast cancer specific survival,[91] and one did not provide comparative survival data.[92] We identified a further 15 cohort studies for our review, 10 of which included patients that were diagnosed with breast cancer before 1990[35,37,39-46] and thus would have been excluded by the HTA. The remaining 5 papers included patients diagnosed in time frames extending both before and after 1990,[29,33,34,36,38] of which three were excluded by the HTA team on the basis that the "study design was not met"[33,34,36], and two were not identified.[29,38]

In the HTA, the quality of the included studies were considered to be variably diminished by the nature of patient sampling. The study populations were considered unrepresentative of the target population in 4 studies and there was a lack of clarity regarding whether: patients were a consecutively treated series (5 studies), mortality data were adjusted for lead and length time bias (6 studies), or attrition bias had affected the results as information on patient loss to follow-up was not provided (4 studies). The majority of the studies were retrospective (6 studies). These issues are similar to those we described in our review.

The HTA data extraction template contained a section on outcomes, including two questions about confounding (HTA Appendix 13, page 210). The HTA used a quality assessment tool adapted from the Review Body for Interventional Procedures (ReBIP) checklist for non-randomised studies (HTA Appendix 14, page 219). In their results (page 228), the authors report that question 6 and 17 of the quality appraisal tool on baseline characteristics and adjustment for confounding factors respectively, were not applicable (N/A) to the included studies. It appears that while information on confounders was collected, that it was not analysed or appraised. In our assessment, the three case-control studies by Lash et al[25-27] did adjust for matched factors, but it is possible that factors other than those matched, other exposures or self-selection bias may have accounted for the observed effect. Of the 3 cohort studies in common with the HTA, we assessed them as scoring either two[31]or three[30,32] out of a possible 6 points for "internal validity- confounding". Of the additional 15 cohort studies we included in our systematic review, only 5 attempted to adjust for confounding.[29,34,35,37,38,45]

In the HTA it was reported that one study assessed only all-cause mortality,[91] four assessed all-cause mortality and breast cancer specific mortality,[25,26,27,32] one assessed all-cause, breast cancer specific and IBTR/MCBC specific mortality,[27] one assessed IBTR/MCBC specific mortality only,[30] and two papers provided no mortality data.[31,92] Two studies did not report detail of the relapse within the breast,[25,26] and the other six studies reported numbers of individual events within the cohort but no time to event data.[27,30-32,91,92] The HTA authors did not include any axillary recurrences or distant metastases in their analysis, but provided details in an appendix. With respect to method of detection of within the breast, three studies did not report this data,[25-27] of the other five studies, only one reported deaths in addition to number of IBTR and MCBC events.[30] No studies that reported data on quality of life, harm of mammography, uptake of mammography or economic data were identified.

In our study, three case-control studies (all by the same author) reported a survival benefit from follow-up, and all were appraised as moderate quality papers (table 4).Of the eighteen cohort studies, fourteen addressed early vs late detection, two addressed

alternative levels of surveillance, and two studies addressed both. In the latter papers, each outcome was separately analysed giving a total of twenty analyses for eighteen cohort papers. For early vs late detection: eight analyses reported a survival benefit, six analyses reported no survival benefit, and two analyses did not report any survival data. The quality of the sixteen analyses varied from good-moderate through to poor (table 4). With respect to alternative levels of surveillance: two moderate quality analyses reported a survival benefit, and two moderate-poor quality papers reported no survival benefit (table 4).

The HTA authors conclude that while there is a paucity of evidence, "there is a suggestion that" surveillance mammography offers a protective effect against all-cause mortality,[25,26] breast cancer specific death,[32] and against death for an IBTR detected by mammography (compared to one diagnosed by clinical examination).[30] All included papers were prone to lead and length time bias, and the "limited and variable nature of the data provided by the included studies precluded formal quantitative synthesis". As in the HTA, our review identified a complete absence of randomised controlled trials in this area, and concluded that the observational studies were subject to a range of biases that limit their interpretation. Limitations of the case-control studies primarily related to the integrity of information on exposure, matching and self-selection bias. Important biases observed in the cohort studies included selection bias, attrition bias, lead time and length bias. Similarly, we were unable to perform a meta-analysis due to the heterogenous nature of the data. However, in contrast to the HTA, we conclude that less than half of the cohort studies attempted to adjust for confounding; and while we agree that the evidence suggest an effect, as we cannot measure the magnitude of the effect, we interpreted the results as providing no definitive indication of the effects of follow-up mammography on survival.

As a result of their findings, the HTA authors subsequently performed a systematic review of diagnostic performance and analyses of existing registry datasets to inform their planned economic evaluation. Our review subsequently suggests ways of overcoming limitations in the evidence base. Specifically we conclude that we should be striving to improve the quality of our observation studies and standardise their

reporting, as well as to embrace alternative research techniques, such as decision analytic modelling, to guide our practice in the likely continued absence of RCTs in this field.

Our analyses were undertaken concomitantly and without prior knowledge of the commissioned research program in the United Kingdom. While approached differently, our review yielded broadly similar results to that of the HTA which was published only days before. As such, we were unable to publish our systematic review as it was not considered to be an original contribution to research.

Chapter 3

A PATIENT-LEVEL CALIBRATION FRAMEWORK FOR EVALUATING SURVEILLANCE STRATEGIES: A CASE STUDY OF MAMMOGRAPHIC FOLLOW-UP AFTER EARLY BREAST CANCER

PREAMBLE

The following two chapters and extension address research question 2: "Using a model based economic analysis, is it efficient to tailor mammographic follow-up according to risk of recurrence?"

Chapter 3: A patient-level calibration framework for evaluating surveillance strategies: a case-study of mammographic follow-up after early breast cancer

This paper describes the development, calibration, and cost-effectiveness analyses of an early breast cancer surveillance discrete event simulation (DES) model. The DES model was used to analyse three alternative mammographic follow-up schedules for postmenopausal women who were disease free following primary treatment for moderate prognosis early breast cancer: taking into account age and adherence to mammography. This study demonstrates the potential value of combining linked, retrospective data and decision analytic modelling to provide estimates of costs and health outcomes that are sufficiently robust to inform cancer clinical guidelines and individual patient decisions regarding appropriate follow-up schedules.

Chapter 4: One size does not fit all? Cost utility analyses of alternative mammographic follow-up schedules, by risk of recurrence

The aim of this paper is to report the full set of cost-effectiveness results from the model described in chapter 3, comparing alternative mammographic follow-up schedules for postmenopausal women with excellent, good, moderate and poor prognosis early breast cancer. Our results suggest that annual mammographic follow-up is not cost effective for most postmenopausal women, and that mammographic follow-up can be tailored based on the Nottingham Prognostic Index score of the primary breast cancer and age at diagnosis.

Chapter 4 Extension: Cross validation of our economic evaluation

The extension provides a comparison between our economic evaluation, with that published in the Health Technology Assessment, with respect to model structure, analyses and results. Although different modelling approaches were used, the results reported for the two studies are broadly similar. This increases confidence regarding the validity of our conclusions that a one-size-fits-all annual surveillance strategy may not be the optimal approach to monitoring breast cancer survivors.

TITLE

A patient-level calibration framework for evaluating surveillance strategies: a case study of mammographic follow-up after early breast cancer.

INTRODUCTION

Breast cancer is the most common incident form of malignancy in Australian women, accounting for 28% of cancer diagnoses in 2008.[1] Between the periods 1982-1987 and 2006-2010, in Australia the 5 year relative survival from breast cancer increased from 72% to 89%.[1] The largest survival gain was for women aged between 50-69 years. The 5 year relative survival increased from 70% to 91% for women aged 50-59 years, and from 72% to 93% for women aged 60-69 years.[1]

Given the ageing population, the number of women diagnosed with breast cancer is expected to increase,[1] and with improving survival, there is a growing population of breast cancer survivors who will all require mammographic follow-up during their lifetime.[2] Women with a past history of breast cancer are at increased risk of cancer within the breast (either recurrence or a new primary), but it is unclear how often, or for how long, we should continue to perform mammography in women who are disease-free following primary treatment for early breast cancer.[3] Mammographic follow-up is effectively screening of a high risk population. However, not all women who have survived early breast cancer have the same risk of breast cancer recurrence or a second primary, with rates per 1000 woman-years varying by patient, tumour and treatment characteristics at initial diagnosis.[4]

Many international guidelines exist, and all acknowledge the paucity of definitive evidence on follow-up care after breast cancer,[5] with much of the information from low level observational studies.[5] There is no randomised controlled trial evidence on which to base guideline recommendations of annual compared to other frequency of follow-up mammography, and there is currently no tailoring of mammographic frequency according to risk.[3,5-9] In a recent guideline update, no primary studies were identified which addressed how long follow-up should continue after diagnosis or treatment.[5]

While RCTs can provide an unconfounded estimate of effect, and are a common and relatively easy source for effectiveness and cost data, they are limited to the patients, interventions and time span of the trial, and are sometimes not feasible. Potential barriers to performing RCTs in follow-up imaging for cancer include difficulty with patient accrual[10] (for example, if patients have a 50% chance of receiving follow-up that is less frequent than current guideline recommendations they are unlikely to consent to be in a trial), and the large sample size and long follow-up (and thus cost implications) required to demonstrate a significant difference between alternative follow-up programs for different patient subgroups.[10] Reviews of the many observational studies that have investigated the clinical (but not economic) effects of mammographic surveillance have shown these studies to be of poor quality and prone to bias, particularly length bias.[11] As such, observational studies provide a limited basis for the direct, non-modelled evaluation of cost-effectiveness of alternative surveillance strategies.

To move away from a "one size fits all" mammographic follow-up schedule to a "personalised" mammographic schedule based on risk of recurrence; given that RCTs are unlikely to occur, we need to consider alternative research methodologies to guide clinical practice.

Decision analytic modelling facilitates data synthesis to describe disease progression over an extended time horizon in order to capture all important differences in costs and benefits between alternative strategies. Most models used to evaluate surveillance options include unobservable input parameters. In the case of cancer follow-up, we cannot observe the time at which patients develop asymptomatic recurrence. Calibration is often used to fit values for these parameters, such that models' outputs match some observed data for the population being evaluated.[12] This process is most common in models of population-based screening programs,[13] though previous models have not used patient-level data to represent observed surveillance pathways to inform the calibration of the underlying disease progression parameters. The use of longitudinal data that includes patient-level surveillance pathways and outcomes may provide a more robust basis for calibrating surveillance models.

Discrete evaluation simulation (DES) has been used to evaluate a wide range of screening and other surveillance activities.[14-19] For health technology assessment, DES is most commonly used to facilitate the representation of complex model structures,[20] which is often the case for surveillance models that represent the full disease course from incidence to death. However, in the context of calibrating surveillance models, the individual-level nature of DES may confer other advantages over the more commonly used cohort-based state transition model. Such individual-level models allow observed individual surveillance pathways to be specified as input parameters, which extracts greater value from individual-level data, where available.

This paper describes the development, calibration, and cost-effectiveness analyses of an early breast cancer surveillance DES model. The DES model was used to analyse three alternative mammographic follow-up schedules for postmenopausal women who were disease free following primary treatment for moderate prognosis early breast cancer; taking into account age, adherence to mammography, and baseline risk of recurrence.

METHODS

Model structure

The model represented disease pathways related to early breast cancer. DES was selected because the model was calibrated to individual-level patient data.[21]Patients move through the model and experience events at any discrete time period after the previous event. For any given patient, their pathway may be influenced by their individual characteristics (eg. age, menopausal status), their tumor (eg. size, nodal status, grade), or their treatment; which in turn impacts on health service costs, quality of life, and overall survival.

The DES model was developed in Simul8® (figure 1). The model uses the following health states to represent progression of breast cancer in women who are disease free after completion of primary treatment (surgery, radiotherapy, chemotherapy):

- Disease free – no recurrence of breast cancer
- Impalpable local recurrence – recurrence within the treated or opposite breast that is only detectable by follow-up mammography

- Palpable local recurrence – recurrence within the treated or opposite breast that is clinically detectable by a doctor or the patient
- Removed impalpable local recurrence – local recurrence detected by follow-up mammography and surgically removed (± adjuvant treatment), rendering the patient disease free.
- Distant metastases – disease relapse outside the breast/axilla
- Breast cancer death – death due to breast cancer
- Other cause death – death not caused by breast cancer

Mammographic follow-up is represented as an instantaneous event. Model assumptions are presented in Box 1. The following section describes the rationale for the pathways between the represented events.

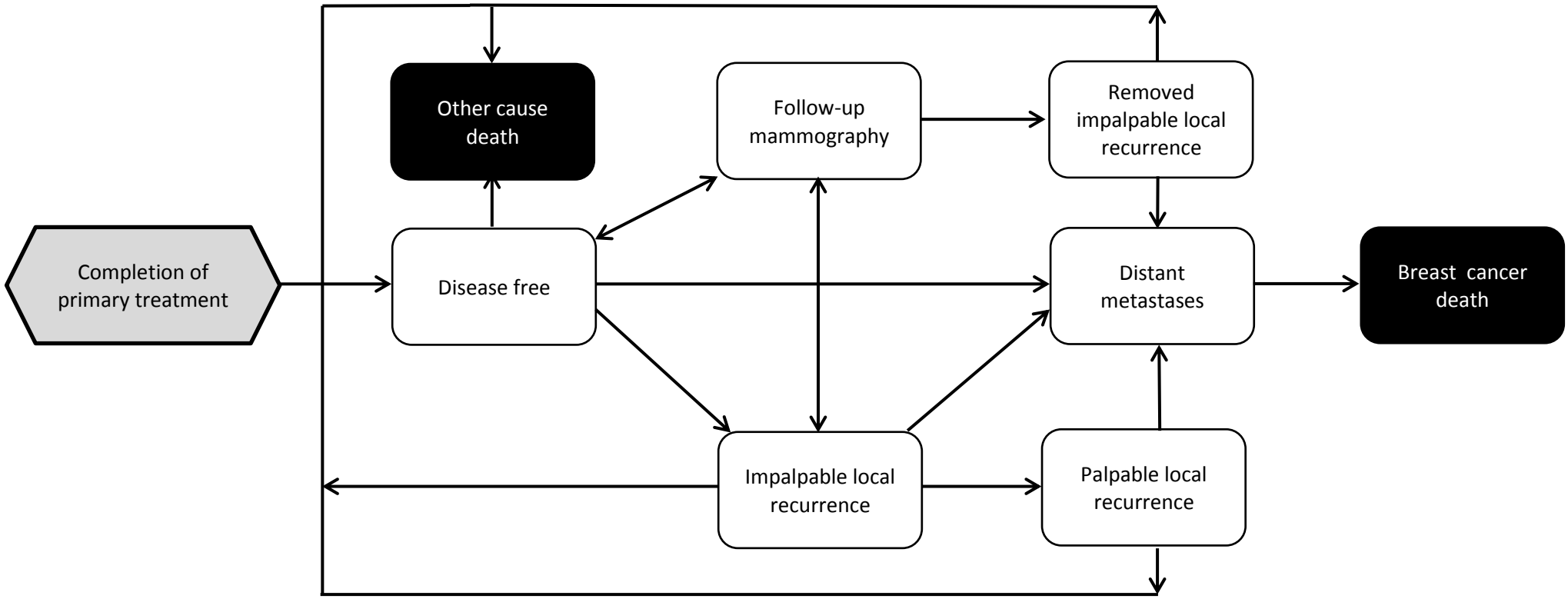
Model pathways

Women enter the model disease-free, but are at risk of developing a recurrence. A woman will leave the disease-free state if she develops a local recurrence. Initially this will be impalpable, and only detectable via mammographic surveillance. If an impalpable local recurrence is not detected, and the patient does not die in the intervening period, the recurrent tumour will either continue to grow locally and be detected clinically by the doctor or patient (palpable local recurrence) or metastasize to other parts of the body (distant metastases).

Women can develop distant metastases from any health state, which is typically incurable, and such women are assumed to die of causes related to breast cancer (breast cancer death). Prior to the development of distant metastases, women can die from causes unrelated to breast cancer at any time (other cause death).

For women with an impalpable local recurrence, a true positive follow-up mammogram results in the removal of the lesion (removed impalpable local recurrence). Following a false negative follow-up mammogram, the lesion will remain undiagnosed, and continue to grow. For women in the disease-free state, a false positive follow-up mammogram results in a biopsy and upon a negative result, such women return to the disease-free state.

Figure 1: Structure of DES model describing the possible progression of early breast cancer after completion of treatment, and the detection of impalpable disease by follow-up mammography



Box 1: Model assumptions

1. All patients are women
2. Women aged ≥ 50 years are defined as postmenopausal
3. Local recurrence refers to recurrence of breast cancer in the treated breast/axilla or new primary breast cancers in the contralateral breast
4. A local recurrence is curable, and women with local recurrence will not experience a breast cancer death unless they develop metastatic disease
5. A local recurrence will be initially impalpable and if untreated will continue to increase in size and eventually become palpable
6. Early detection of a local recurrence when impalpable reduces risk of metastatic disease compared to late detection of a local recurrence when palpable
7. Once a local recurrence (impalpable or palpable) is detected, it will be surgically removed (± adjuvant treatment) rendering the patient disease-free
8. Distant metastases include all systemic relapses outside the breast/axilla, and includes supraclavicular lymphadenopathy as well as visceral metastases in lung, liver, bone, brain and other sites
9. Women are at risk of developing distant metastases with and without the prior development of local recurrence
10. Distant metastases are incurable and result in death from breast cancer

Model scenarios

The DES model was used to analyse three alternative mammographic follow-up schedules for postmenopausal women who were disease-free following primary treatment for moderate prognosis early breast cancer; taking into account age at diagnosis and adherence with mammography. The twelve scenarios are described in table 1.

Table 1 – Model scenarios for moderate subgroup

Mammography schedules	Adherence %
50-69 years	
2 yearly	90
Mixed	90
Annual	90
2 yearly	75
Mixed	75
Annual	75
70-79 years	
2 yearly	90
Mixed	90
Annual	90
2 yearly	75
Mixed	75
Annual	75

The age cohorts were 50-69 years and 70-79 years, to reflect the current target age group for breast cancer screening in Australia, and an older target age group, respectively.

Mammographic schedules were chosen to reflect the current annual follow-up interval and two less intense schedules that were deemed feasible options within Australia as informed by expert consultation. The "mixed" schedule consisted of annual mammography for 5 years, then 2 yearly thereafter. This was designed to reflect a follow-up frequency intermediate between annual and 2 yearly surveillance. Two yearly mammography was the least intense follow-up schedule. This was chosen to investigate the impact of women with a personal history of breast cancer returning to a frequency of mammography identical to that of the screening population with no personal history of breast cancer. A "no surveillance" option was not considered to be clinically safe or efficient for women who have a personal history of breast cancer. This differs from early stage melanoma, for example, where the majority of recurrences are detected by the patient and "no follow-up" is currently being assessed as a viable option.

In the absence of South Australian or national data describing adherence with follow-up mammography by women with a personal history of breast cancer, we assumed adherence would be higher than with screening mammography of asymptomatic women with no personal history of breast cancer. In 2009-2010 in Australia, 55% of eligible women attended Breast Screen Australia.[22] We separately modelled 75% and 90% adherence with follow-up mammography to provide feasible upper and lower boundaries of mammographic attendance by Australian breast cancer survivors.

Model inputs

Data on breast cancer follow-up and recurrence is not routinely collected in South Australia. As no single data source within South Australia contained the information required, we obtained ethics approval from SA Health Human Research Ethics Committee to extract and link data across the South Australian Cancer Registry; public metropolitan hospital database (OACIS); state-wide inpatient database (ISAAC which contains administrative data on both public and private hospitals) and the Registry of Births, Deaths and Marriages.

We constructed a patient level dataset of women with early breast cancer, and stratified women according to prognosis based on pathological criteria that were routinely collected in South Australia. For each woman we determined the follow-up she received and her disease course. This included the frequency of her mammographic follow-up; whether or not she developed a recurrence (and if so, whether in the breast or distant, and when it was detected) and whether or not she died from breast cancer.

Eleven hundred postmenopausal women diagnosed with early breast cancer between 2000-2008, who had their primary breast cancer treatment and mammographic follow-up in the public health system, were identified and followed until 30/6/2011. The original treatment for all women included surgery ± radiotherapy ±chemotherapy. For each woman, we collected date of birth; date of surgery; tumor size, nodal status and grade of the primary breast cancer; date and type of first recurrence; date and cause of death; and date and result of each follow-up mammogram. From this data, we

calculated age and menopausal status at diagnosis, and the Nottingham Prognostic Index (NPI) for the primary breast cancer.[23] The Nottingham Prognostic Index is calculated as follows:

NPI= Tumor size(cm) x 0.2 + Lymph node stage (1,2,3) + Histologic grade (1,2,3).

Prognosis can be defined in terms of the NPI score. Tumors with excellent prognosis have a score of ≤ 2.4, good prognosis have a score of ≤ 3.4, moderate prognosis have scores of >3.4 and ≤ 5.4, and poor prognosis tumors have scores >5.4. The model calibration process and results are reported for the largest subgroup of our study cohort, the postmenopausal (age ≥ 50 years) moderate prognosis group (n=407) (table 2).

Table 2: Observed outcomes of postmenopausal women with moderate prognosis breast cancer in South Australia (n=407)

Event	Time horizon	Total event rate	Patient years at risk	Incidence rate	SD	Lower 95% CI	Upper 95% CI
Local recurrence	5 years	19	1768	0.011	0.002	0.006	0.016
	10 years	24	2407	0.010	0.002	0.006	0.014
Breast cancer death	5 years	17	1802.5	0.009	0.002	0.006	0.014
	10 years	28	2483	0.011	0.002	0.008	0.016
Other cause death	5 years	9	1802.5	0.005	0.002	0.002	0.009
	10 years	25	2483	0.010	0.002	0.007	0.014

SD = standard deviation

Cost and utility data was identified from the literature. Karnon et al provided mean annual costs with 95% confidence intervals, for patients remaining alive in the same contralateral, locoregional or distant recurrence health state for one year.[24] We converted the costs in £2004 to $A2004 and then uprated to Australian $2011. The United Kingdom and Australian Implied Purchasing Power Parity (PPP) Conversion Rates

were used to convert £2004 to $US2004 (0.655), and $A2004 to $US2004 (1.367) respectively.[25,26] This equated to a PPP £2004 to $A2004 of 2.087022901. The costs in $A2004 were then uprated to $A2011 using Australian health inflation figures.[27] This gave mean and 2.5th and 97.5th percentiles for costs related to contralateral recurrence (year 1, and year 2+), local recurrence (year 1, and year 2+), and distant recurrence (separately for years 1-5) post diagnosis. Assuming 70% of recurrence within the breast occurs in the treated breast and 30% in the contralateral breast, we calculated a weighted local recurrence cost for year 1 and year 2+ respectively. These costs were converted from annual to weekly costs. Year 1 costs were used as model inputs for "removed impalpable local recurrence surgery costs" and "palpable local recurrence surgery costs" health states. Year 2+ costs were used as model inputs for ""removed impalpable local recurrence ongoing costs" and "palpable local recurrence ongoing costs" health states. Separate mean annual costs of treating distant metastases in years 1-5 post diagnosis were used to calculate a single weighted mean annual cost for the distant metastases health state, based on the proportion of women with distant metastases surviving to the end of each year. Utility values for each health state were identified in a Health Technology Assessment published in 2007[28] whose primary sources were the CEA registry of the Harvard School of Public Health[29] and a paper by Tengs and Wallace[30](table 3). A quality of life decrement associated with mammography screening was identified in Tengs and Wallace,[30] based upon work by de Koning et al.[31]

Table 3: Baseline costs in $A2011 and utilities, for health states in the model

	Mean values		Best case values**		Worst case values**		
Health states	**Annual costs***	**Utilities**	**Annual costs***	**Utilities**	**Annual costs***	**Utilities**	**References**
Disease-free	$0	0.94	$0	0.94	$0	0.94	30
Surveillance*	$124	-0.01	$124	0	$124	-0.02	31
ILR	$0	0.94	$0	0.94	$0	0.94	30
RILR - surgery costs	$32,323	0.74	$45,773	0.7	$20,032	0.74	30
RILR - ongoing costs	$2,207	0.85	$3,915	0.85	$1,175	0.85	29
PLR - surgery costs	$32,323	0.74	$45,773	0.7	$20,032	0.74	30
PLR - ongoing costs	$2,207	0.85	$3,915	0.85	$1,175	0.85	29
Distant metastases	$17,263	0.5	$22,935	0.35	$12,715	0.65	30

ILR = impalpable local recurrence, RILR= removed impalpable local recurrence, PLR = palpable local recurrence

*Surveillance = costs of follow-up reflect the weighted costs of a single MMG encounter (assuming unilateral mammography in 30% and bilateral mammography in 70% of women) plus a single attendance fee, per year, based on 2011 MBS rates.

**Best case from the perspective of increased surveillance frequency

Other cause death was calculated by subtracting the proportion of Australian women who died from breast cancer from age-specific mortality rates derived from Australian life-tables.[32,33]

Model calibration

Probabilistic model calibration was performed according to the seven step approach described by Vanni and Karnon,[12] which integrates model calibration with probabilistic sensitivity analysis:

1. Selection of calibration targets - calibration targets were selected for the 407 women for whom a linked dataset had been compiled. Five and ten year incidence rates (with lower and upper 95% confidence limits), were sought for recurrence within the breast (breast recurrence), distant metastases, breast cancer death and other cause death. Breast recurrences were identified from pathology and mammography data available for all women. Mortality data, including cause of death, were extracted from South Australian Cancer Registry. On review, other cause death was discarded as a calibration target due to the small numbers of observed events. Distant metastases incidence was sought from available inpatient data, though these data were deemed insufficient to reliably identify distant metastases, and so this calibration target was also discarded.

2. Input parameters included in the calibration process – all disease progression input parameters, plus the test characteristics of the mammography were calibrated. To reduce the parameter space to be searched, prior probability functions were specified for these parameters. Beta distributions were used to represent reported values for the sensitivity and specificity of mammography.[34] For the time to event parameters, Weibull distributions were specified to facilitate time varying event rates. Uniform ranges for the alpha and beta parameters for each Weibull function were specified to represent the uncertainty around the event rates reported in the following data sources:
   (a) 10 year event rates for disease progression from the disease-free state to impalpable local recurrence, and from disease-free to distant metastases were extracted from the Early Breast Cancer Trialists' Collaborative Group (EBCTCG).[35]

(b) 1 year event rates for disease progression from impalpable to palpable local recurrence; and from impalpable local recurrence, palpable local recurrence and removed impalpable local recurrence respectively to distant metastases, were based on expert consultation.

(c) 1 year event rate for progression from distant metastases to breast cancer death were derived for published modelled evaluations of systemic treatments for early breast cancer.[36]

3. Convergence criteria – convergence describes the identification of sets of input parameter values that predict the observed model output parameter values with sufficient accuracy. Sampled input parameter sets were defined as convergent if all output parameters lay within the 95% confidence intervals for the observed value for every calibration target (5 year and 10 year event rates for both in-breast recurrence and breast cancer death).

4. Parameter search strategy – a random search strategy was used to sample sets of input parameter values from the defined probability functions. Each sampled set was used to populate the model, and the associated predictions of the calibration targets were compared to the observed values to determine convergence.

5. Stopping rule – we repeated steps 3 and 4 until we had at least 1000 convergent sets of input parameter. Testing confirmed non-significant variation in the mean output values across 1,000 and 2,000 convergent parameter sets.

6. Goodness-of-fit measures – to represent the relative accuracy of the alternative convergent sets of input parameter values with respect to the observed calibration targets, we used the chi-squared goodness of fit method. This measure reflects the uncertainty around the observed calibration targets,[12] and has been found to be more sensitive to variations in model accuracy than other measures.[37]

7. Integrating the results of the calibration and model sensitivity analysis –probability weights were attached to all convergent input parameter sets based on the chi-squared goodness of fit measure. A probabilistic sensitivity analysis was undertaken using weighted sampling of the convergent sets of input parameters.[38]

Model analysis

The model was run for the same 2000 sampled sets of convergent input parameter values for each scenario. Model outputs were total costs (health state costs + costs of surveillance), and total QALYs across the model health states. Within each age and adherence group (e.g. 50-69 year olds assuming 75% adherence), mean incremental cost-effectiveness ratios (ICERs) were estimated between mammographic schedules arranged in increasing order of effectiveness. Estimates of the probability of each follow-up scenario maximising net benefits were also generated for a range of QALY threshold values for each age and adherence group.

Deterministic sensitivity analyses were undertaken around the health state cost and utility weight input parameters, including the utility decrement associated with follow-up mammography. Best and worst case scenarios were defined from the perspective of favouring more intensive surveillance. For the best case scenario, higher costs for removed impalpable local recurrence ($45,773 and $3,915 in the first and following years, respectively), palpable local recurrence ($45,773 and $3,915) and distant metastases ($22,935) were used; with lower utility values for the RILR – surgery (0.7), PLR – surgery (0.7) and distant metastases (0.35) health states; compared to the base case (Table 3). The reverse is tested for the worst case scenario. That is, lower costs for removed impalpable local recurrence ($20,032 and $1,175), palpable local recurrence ($20,032 and $1,175) and distant metastases ($12,715); with identical utilities for the RILR – surgery (0.74) and PLR – surgery (0.74) health states, but higher utilities associated with distant metastases (0.65) than the base case (Table 3). The utility values for the surgery states in the worst case scenario are unchanged from the base case, but the utility decrement associated with surveillance is increased (see below).

In the base-case the applied QALY decrement associated with follow-up mammography of 0.01 is equivalent to a 0.5 utility decrement for one week, or a 0.25 utility decrement for two weeks, reflecting the heightened anxiety around the surveillance period. The QALY decrement also captures false positive effect for the small proportion of patient who undergo further investigations to rule out recurrence. For the best case scenario,

the utility decrement for surveillance was decreased to zero, and in the worst case scenario, the utility decrement for surveillance is increased (-0.02).

RESULTS

Base-case analysis

Table 4 reports the results of the base case scenarios for the alternative mammographic follow-up schedules in women with moderate prognosis tumors, by age at diagnosis, adherence with follow-up mammography and assuming a 0.01 QALY decrement associated with a surveillance episode. The incremental cost-effectiveness ratios (ICERs) varied by assumed adherence levels.

Table 4: Mean incremental cost-effectiveness results for postmenopausal women with moderate prognosis early breast cancer*

Mammography schedules	Adherence %	Breast cancer deaths %	Mean costs	Mean QALYs	Cost difference	QALY difference	ICER
50-69 years							
2 yearly	90	0.330	$6,302	11.258			
Mixed	90	0.329	$6,535	11.263	$234	0.005	$48,583
Annual	90	0.328	$6,985	11.266	$450	0.003	$121,227
2 yearly	75	0.331	$6,090	11.260			
Mixed	75	0.330	$6,311	11.267	$222	0.008	$29,241
Annual	75	0.329	$6,676	11.269	$364	0.001	$289,910
70-79 years							
2 yearly	90	0.188	$4,264	7.978			
Mixed	90	0.187	$4,467	7.979	$203	0.002	$113,078
Annual	90	0.187	$4,739	7.980	$272	0.001	$254,030
2 yearly	75	0.189	$4,187	7.970			
Mixed	75	0.188	$4,375	7.973	$189	0.003	$62,405
Annual	75	0.188	$4,589	7.975	$214	0.001	$150,911

*Based on the mean output values from the probabilistic analysis of the weighted sampling of the convergent sets of input parameters
QALY = Quality adjusted life year, ICER = Incremental cost-effectiveness ratio
Mixed = follow-up mammography every year for 5 years then 2 yearly thereafter,2 yearly = follow-up mammography every 2 years

Assuming 90% adherence, the cost-effectiveness of mixed follow-up compared to two-yearly follow-up range from $48,583 per QALY gained for women aged 50 to 69 to $113,078 for women aged 70-79. The corresponding values for annual screening compared to mixed follow-up range from $121,227 for women aged 50-69 to $254,030 for women aged 70-79.

If adherence is reduced to 75%, the cost-effectiveness of mixed follow-up compared to two-yearly follow-up reduce to $29,241 per QALY gained for women aged 50 to 69 to $62,405 for women aged 70-79. The corresponding values for annual screening compared to mixed follow-up range from $289,910 for women aged 50-69 to $150,911 for women aged 70-79.

The results are sensitive to the assumed adherence rate due to the small absolute gain in QALYs between surveillance options, which means that changes in adherence result in larger relative changes in QALY gains compared to cost differences. For example, in 50-69 year old women, moving from a mixed schedule to an annual schedule results in an incremental cost difference of $450 assuming 90% adherence, and $364 assuming 75% adherence. The corresponding QALY differences are 0.003 and 0.001.

Deterministic sensitivity analyses

Best and worst case scenario analyses were undertaken around the cost and utility input parameters, as presented in table 5. Compared to the base case results, the best case scenario results in substantially lower ICERs in both analyses, primarily due to the best case assumption of differential costs and utility weights for local recurrence that is detected by mammography, and that detected clinically. The worst case scenario was less divergent from the base case, and so the effects on the ICER are smaller.

Table 5: Best and worse case cost and utility input parameter scenario analyses*

	Base case			Best case scenario			Worse case scenario		
	Cost difference	QALY difference	ICER	Cost difference	QALY difference	ICER	Cost difference	QALY difference	ICER
50-69 years, 90% adherence									
Mixed MMG - 2 yearly MMG	$234	0.0048	$48,583	$129	0.0074	$17,446	$238	0.0045	$52,446
Annual MMG - Mixed MMG	$450	0.0037	$121,227	$358	0.0068	$52,884	$454	0.0031	$146,605
70-79 years, 75% adherence									
Mixed MMG - 2 yearly MMG	$189	0.0030	$62,405	$106	0.0047	$22,598	$191	0.0028	$68,265
Annual MMG - Mixed MMG	$214	0.0014	$150,911	$176	0.0023	$76,617	$216	0.0011	$191,194

ICER – incremental cost-effectiveness ratio

*Best case is defined from the perspective of favouring more intensive surveillance, please see text for further details.

Cost and QALY differences are presented between the ordered options, e.g. the $234 cost difference is the additional cost of the mixed program relative to the 2 yearly program.

Probabilistic sensitivity analyses

The results of the probabilistic sensitivity analysis are shown in table 6, and illustrate the cumulative probabilities of cost-effectiveness for alternative follow-up schedules, by QALY threshold and assumed adherence with follow-up schedule, for each age cohort.

Table 6: Probabilistic sensitivity analysis

Cohort	Threshold l adherence	Annual	Mixed	2 yearly
50-69 years	$25k | 90%	0.052	0.365	0.583
	$50k | 90%	0.208	0.391	0.401
	$75k | 90%	0.295	0.371	0.334
	$25k | 75%	0.114	0.394	0.492
	$50k | 75%	0.267	0.402	0.331
	$75k | 75%	0.345	0.383	0.272
70-79 years	$25k | 90%	0.055	0.248	0.697
	$50k | 90%	0.175	0.338	0.487
	$75k | 90%	0.244	0.359	0.397
	$25k | 75%	0.076	0.302	0.622
	$50k | 75%	0.208	0.359	0.433
	$75k | 75%	0.263	0.361	0.376

Threshold = QALY threshold of $25,000, $50,000 or $75,000 (the value attached to the gain of additional QALYs)
Adherence = % of women adherent with given mammography schedule
Annual = annual mammography, mixed = annual mammography for 5 years, then 2 yearly thereafter, 2 yearly= mammography every 2 years.

For women aged 50 to 69, at a QALY threshold of $25,000 and assuming 90% adherence, 2 yearly follow-up has a 58% probability of being cost-effective, which decreases to 27% when the threshold is increased to $75,000 and adherence falls to 75%. The corresponding figures for annual surveillance are 5% and 35%, respectively. The probability of cost-effectiveness for the mixed follow-up schedule remains fairly constant at around 40% across the scenarios.

For women aged 70 to 79, there is greater certainty around the cost-effectiveness of two yearly surveillance, though the probability of cost-effectiveness still drops to 38% when assuming a $75,000 threshold and 75% adherence. There is more uncertainty around the probability of cost-effectiveness for the mixed option, which ranges from 25% to 36%.

DISCUSSION

For women with moderate prognosis breast cancer aged 50-69 years at diagnosis, our modelled analysis suggests that compared to mammographic surveillance once every two years, annual follow-up for 5 years, with 2 yearly visits thereafter has a mean incremental cost-effectiveness ratio that falls within or around commonly implied cost-effectiveness thresholds for Australia (depending on assumptions regarding adherence).[39] For women aged 70-79 years at diagnosis, a similar surveillance frequency to general population screening (2 yearly) appears to be most cost-effective based on the mean results.

As expected, the likelihood of annual follow-up schedules being the most cost-effective option is positively associated with increasing QALY thresholds (from $25,000 to $50,000 to $75,000), and decreasing adherence with mammography (90% to 75%). However, in the evaluated, moderate risk sub-group, the results suggest that it is unlikely that the incremental benefits of ongoing annual surveillance compensate for the incremental costs associated with annual surveillance beyond five years.

The relative cost-effectiveness of the mixed and the mixed and two yearly schedules is sensitive to uncertainty around the cost and utility parameters, for which a contemporary Australian cost and utility study would provide useful information to provide greater assurance the value of these schedules.

A recent Health Technology Assessment of surveillance strategies for women with early breast cancer in the UK used an alternative modelling approach.[11] Using a cohort Markov model, Robertson et al represented disease progression in more detail, but they did not calibrate (or externally validate) their model. Different frequencies (12, 18, 24 and 36 month intervals) and different methods of surveillance (no surveillance, mammography, mammography + clinical examination, MRI + clinical examination) were tested, including sub-group analyses of two hypothetical women aged 40 years and 70 years, representing respectively, higher and lower likelihoods of relapse within the breast (based on increased risk of recurrence, shorter time for an undetected cancer to progress to worse risk profile, and more aggressive and costly treatment of recurrence in the younger woman). Their results suggested that more intensive follow-up of women judged to be at high risk, and less intensive follow-up of women judged to be at low risk, may be cost-effective.[11] This is consistent with findings reported in this paper, which reports results for a moderate risk group.

The ISPOR Task Force on Good Research Practices in Modeling Studies states that model validation is a key requirement for high quality models.[40] We followed a recently published reporting framework for evaluating model performance (validation and calibration) suggested by Haji Ali Afzali et al.[41] Expert opinion was obtained to ensure that model, data sources and results made intuitive sense; code within the model was subject to extensive debugging/verification (JK); and the model was calibrated to local South Australian data prior to performing the model analysis. We cross-validated the study by comparing and contrasting the outcomes of our discrete event simulation model with those of an alternative, cohort-based state transition model that was built to inform the same decision.[11] Ultimately external validation of our model will occur as new evidence becomes available.[40]

A further strength of our study is the use of linked, longitudinal individual level data that represented the observed follow-up mammography schedule received by each of 407 patients. This approach precluded the secondary parameterisation of surveillance frequency, an important parameter that is generally subject to significant uncertainty. Taylor et al illustrated the importance of reflecting actual surveillance, comparing the effect of conventional calibration (assume all women had been exposed to current screening processes in the past) and an historically accurate calibration that reflected the fact that US women ≥ 65 years of age had not received currently available screening practices for cervical cancer at younger ages.[42] They concluded that calibrating longitudinal models to cross-sectional data without accounting for temporal changes in clinical practice may lead to bias in evaluating the effectiveness of interventions.[42]

Our study does have several limitations. Firstly, the quality and quantity of the collected input and calibration target data placed constraints on our intended analysis: the South Australian Cancer Registry did not collect hormone receptor status of the primary breast cancer prior to June 2012, and HER2 receptor status is not yet collected which limits the precision of our risk stratification technique; our eligible population was limited to patients receiving their primary treatment in the public sector as privacy laws prevented access to mammography and pathology reports from the private sector; and as recurrence and follow-up data is not routinely collected in South Australia, the need to link four separate data sources (each built for different purposes) resulted in loss of data due to an inability to match all patients across all databases.

Secondly, although prior probability functions were used to reduce the parameter space searched as part of the calibration process, guided, algorithmic search strategies are likely to have improved the efficiency and accuracy of the process compared to the applied random sampling approach.[43,44]

Thirdly, there is the potential to build more comprehensive models that could include premenopausal women; additional features of the primary tumor that impact on prognosis (eg. receptor status); treatment regimens; and a range of alternative surveillance strategies that include incorporating other imaging modalities (eg. MRI or ultrasound) as an adjunct to mammography, imaging both with and without clinical

examination, and other frequencies of mammographic surveillance. Our analysis does not represent clinical examination as an alternative or supplementary method of surveillance, rather the detection of a recurrent tumour by the patient or physician is reflected as an exogenous event. Without patient-level data describing the timing and frequency of clinical examination, and the mode of detection of recurrent tumours, there was too little data to inform such an extension to the model. To inform this option would have required casenote review for the entire South Australian cohort over a 10 year period, which was beyond the scope of this study.

Our results suggest that mammographic follow-up can potentially be tailored according to risk of recurrence based on the NPI score of the primary breast cancer and age at diagnosis. If our current model is validated with a larger dataset, this could potentially provide the foundations towards a significant change to current imaging practice in breast cancer follow-up. The model-based cost-effectiveness analysis could inform subsequent activities, including elicitation of patient preferences, stakeholder engagement, and deliberations and research that explicitly consider all salient social, ethical, political, legal, economic, and clinical consequences[45] of tailored mammography schedules.

CONCLUSION

The reported analyses suggest that the current 'one size fits all' guideline of annual follow-up mammography for all women who are disease free following completion of primary treatment for early breast cancer may not be cost-effective. Alternative, less frequent mammographic surveillance appears to be more cost-effective for women aged between 50-79 years, with moderate risk of recurrent breast cancer.

Moreover, this study in breast cancer follow-up has demonstrated the potential value of using linked, longitudinal, individual-level data within a calibration-based decision analytic modelling framework to evaluate the cost-effectiveness of alternative surveillance interventions. Further effort may be required to collect more complete data that are sufficiently robust to inform policy and practice, but given the difficultly in

conducting high quality clinical studies, model-based analyses are of particular relevance to the evaluation of surveillance.

As better prognostic factors are identified across a wide range of cancers, the potential benefits of tailoring follow-up to the risk of recurrence will increase and so we should be developing analytic methods that will guide clinical practice in an evidence based, but also pragmatic manner.

REFERENCES

1. Australian Institute of Health and Welfare & Cancer Australia 2012. Breast cancer in Australia: an overview. Cancer series no. 71. Cat.no. CAN 67. Canberra: AIHW.
2. Elshaug A, Bessen T, Moss J, Hiller J. Addressing “waste” in diagnostic imaging: some implications of comparative effectiveness research. J Am Coll Radiol. 2010 Aug; 7(8):603-613
3. National Institute for Health and Clinical Excellence. Early and locally advanced breast cancer: diagnosis and treatment. National Collaborating Centre for Cancer, Cardiff, Wales, 2009.
4. Buist D, Abraham L, Barlow W, Krishnaraj A, Holdridge R, Sickles E, Carney P, Kerlikowske K, Geller B; Breast Cancer Surveillance Consortium. Diagnosis of second breast cancer events after initial diagnosis of early stage breast cancer. Breast Cancer Res Treat. 2010 Dec; 124(3): 863-873
5. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
6. Khatcheressian J, Hurley P, Bantug E, Esserman L, Grunfeld E, Halberg F, Hantel A, Henry N, Muss H, Smith T, Vogel V, Wolff A, Somerfield M, Davidson N; American Society of Clinical Oncology. Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol. 2013 Mar 1; 31(7): 961-5
7. Aebi S, Davidson T, Gruber G, Cardoso F; ESMO Guidelines Working Group. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2011 Sep; 22 Suppl 6: vi12-24
8. Scottish Intercollegiate Guidelines Network. Management of breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network, Edinburgh, Scotland, 2005.
9. Grunfeld E, Dhesy-Thind S, Levine M; Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment for breast cancer (summary of the 2005 update). CMAJ. 2005 May 10;172(10) :1319-1320

10. Grunfeld E, Noorani H, McGahan L, Paszat L, Coyle D, van Walraven C, Joyce J, Sawka C. Surveillance mammography after treatment of primary breast cancer: a systematic review. Breast. 2002 Jun; 11(3): 228-35
11. Robertson C, Arcot Ragupathy S, Boachie C, Dixon J, Fraser C, Hernandez R, Heys S, Jack W, Kerr G, Lawrence G, MacLennan G, Maxwell A, McGregor J, Mowatt G, Pinder S, Ternent L, Thomas R, Vale L, Wilson R, Zhu S, Gilber F. The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews registry database analyses and economic evaluation. Health Technol Assess. 2011 Sep; 15(34): v-vi, 1-322.
12. Vanni T, Karnon J, Madan J, White R, Edmunds W, Foss A, Legood R. Calibrating models in economic evaluation: a seven step approach. Pharmacoeconomics. 2011 Jan: 29(1): 35-49.
13. Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, Madan J. A review and critique of modelling in prioritising and designing screening programmes. Health Technol Assess. 2007 Dec; 11(52):iii-iv, ix-xi, 1-145.
14. Davies R, Crabbe D, Roderick P, Goddard J, Raftery J, Patel P. A simulation to evaluate screening for Helicobacter pylori infection in the prevention of peptic ulcers and gastric cancers. Health Care Manag Sci. 2002 Nov; 5(4): 249-58.
15. Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Bhandari P, Patel P. The cost-effectiveness of screening for Helicobacter pylori to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model. Health Technol Assess. 2003; 7(6): 1-86.
16. Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Patel P, Bhandari P. Cost-effectiveness of population screening for Helicobacter pylori in preventing gastric cancer and peptic ulcer disease, using simulation. J Med Screen. 2003; 10(3): 148-56.
17. Willis B, Barton P, Pearmain P, Bryan S, Hyde C. Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK. Health Technol Assess. 2005 Mar; 9(13): 1-207, iii.

18. Stout N, Rosenberg M, Trentham-Dietz A, Smith M, Robinson S, Fryback D. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006 Jun 7; 98(11): 774-82
19. Tosteson A, Stout N, Fryback D, Acharyya S, Herman B, Hannah L, Pisano E; DMIST Investigators. Cost-effectiveness of digital mammography breast cancer screening. Ann Intern Med. 2008 Jan 1; 148(1): 1-10.
20. Karnon J, Stahl J, Brennan A, Caro J, Mar J, Möller J; ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--4. Value Health. 2012 Sep-Oct;15(6):821-7
21. Karnon J, Brown J. Selecting a decision model for economic evaluation: a case study and review. Health Care Manag Sci. 1998 Oct; 1(2): 133-40
22. Australian Institute of Health and Welfare & Australasian Association of Cancer Registries 2012. Cancer in Australia: an overview, 2012. Cancer series no. 74. Cat. no. CAN 70. Canberra: AIHW.
23. Galea M, Blamey R, Elston C, Ellis I. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992; 22(3): 207-219.
24. Karnon J, Kerr G, Jack W, Papo N, Cameron D. Health care costs for the treatment of breast cancer recurrent events: estimates from a UK-based patient-level analysis. Br J Cancer. 2007 Aug 20; 97(4): 479-485.
25. United Kingdom Implied Purchasing Power Parity (PPP) Conversion Rate [Internet]. http://www.tradingeconomics.com/united-kingdom/implied-purchasing-power-parity-ppp-conversion-rate-imf-data.html (last accessed 30th Sept 2012)
26. Australia Implied Purchasing Power Parity (PPP) Conversion Rate [Internet]. http://www.tradingeconomics.com/australia/implied-purchasing-power-parity-ppp-conversion-rate-imf-data.html (last accessed 30th Sept 2012)
27. Australian Institute of Health and Welfare 2012. Health expenditure Australia 2010-11. Health and welfare expenditure series no.47. Cat.no. HWE 56. Canberra: AIHW
28. Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L. Hormonal therapies for early breast cancer: systematic review and economic evaluation. Health Technol Assess. 2007 Jul; 11(26): iii-iv, ix-xi, 1-134

29. CEA Registry Website. https://research.tufts-nemc.org/cear4/default.aspx
30. Tengs T, Wallace A. One thousand health-related quality-of-life estimates. Med Care. 2000 Jun; 38(6):583-637
31. de Koning H, van Ineveld B, van Oortmarssen G, de Haes J, Collette H, Hendriks J, van der Maas P. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer. 1991 Oct 21; 49(4) :531-7. 31.
32. Australian Institute of Health and Welfare (AIHW) 2012. ACIM (Australian Cancer Incidence and Mortality) Books. AIHW: Canberra. Breast cancer (ICD10 C50), Australia, 1968-2007. Age specific mortality rates. http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129542435 (last accessed 6 August 2013)
33. Australian Bureau of Statistics. 3302.0.55.001 - Life Tables, Australia, 2003-2005. http://abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/6A75A50AE791F95FCA25738D000F462B (last accessed 6 August 2013)
34. Robertson C, Arcot Ragupathy S, Boachie C, Fraser C, Heys S, MacLennan G, Mowatt G, Thomas R, Gilbert F and the Mammographic Surveillance Health Technology Assessment Group. Surveillance mammography for detecting ipsilateral breast tumour recurrence and metachronous contralateral breast cancer: a systematic review. Eur Radiol. 2011; 21: 2484-2491
35. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. Lancet 2005 May 14-20; 365 (9472):1687-717
36. Karnon J, Delea T, Barghout V. Cost utility analysis of early adjuvant letrozole or anastrazole vs tamoxifen in postmenopausal women with early invasive breast cancer: the UK perspective. Eur J Health Econ. 2008 May; 9(2): 171-83.
37. Carlton J, Karnon J, Czoski-Murray C, Smith K, Marr J. The clinical effectiveness and cost-effectiveness of screening programmes for amblyopia and strabismus in children up to the age of 4-5 years: a systematic review and economic evaluation. Health Technol Assess. 2008 Jun; 12(25): iii, ix-194.

38. Karnon J, Czoski-Murray C, Smith K, Brand C. A hybrid cohort individual sampling natural history model of age-related macular degeneration: assessing the cost-effectiveness of screening using probabilistic calibration. Med Decis Making. 2009 May-Jun; 29(3): 304-316
39. Lopert R. Pharmacoeconomics and drug subsidy in Australia. Evidence, economics, and politics: Australia's experiment in evidence-based medicine; December 2006; St Regis Hotel, San Francisco.
40. Weinstein M, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, Luce B; ISPOR Task Force on Good Research Practices – Modeling Studies. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices – Modeling Studies. Value Health. 2003 Jan-Feb; 6(1): 9-17
41. Haji Ali Afzali H, Gray J, Karnon J. Model performance evaluation (validation and calibration) in model-based studies of therapeutic interventions for cardiovascular diseases: A review and suggested reporting framework. Applied Health Econ Health Policy. 2013 April; 11 (2):85-93
42. Taylor D, Pawar V, Kruzikas D, Gilmore K, Pandya A, Iskandar R, Weinstein M. Calibrating longitudinal models to cross-sectional data: the effect of temporal changes in health practices. Value Health. 2011 Jul-Aug; 14(5): 700-4
43. Whyte S, Walsh C, Chilcott J. Bayesian calibration of a natural history model with application to a population model for colorectal cancer. Med Decis Making. 2011 Jul-Aug; 31(4), 625-641.
44. Taylor D, Pawar V, Kruzikas D, Gilmore K, Sanon M, Weinstein M. Incorporating calibrated model parameters into sensitivity analyses: deterministic and probabilistic approaches. Pharmacoeconomics. 2012 Feb 1;30(2):119-26
45. Elshaug A, Bessen T, Moss J, Hiller J. Addressing "waste" in diagnostic imaging: some implications of comparative effectiveness research. J Am Coll Radiol. 2010 Aug; 7(8):603-13

Chapter 4

ONE SIZE DOES NOT FIT ALL? COST UTILITY ANALYSES OF ALTERNATIVE MAMMOGRAPHIC FOLLOW-UP SCHEDULES, BY RISK OF RECURRENCE

TITLE

One size does not fit all? Cost utility analyses of alternative mammographic follow-up schedules, by risk of recurrence

INTRODUCTION

After completion of primary treatment for early breast cancer, the aim of follow-up mammography is to detect new disease in the treated or opposite breast at an early stage when it is potentially curable.[1] Overall the risk of local recurrence in the treated breast is 0.5-1% per annum (including new primaries), and the risk of developing a cancer in the opposite breast is estimated to be just under 0.03% per annum.[2]

Breast cancers that are detected by mammography whilst impalpable tend to have a better prognosis than those that are detected when larger and palpable,[3] and more frequent surveillance mammography will detect more impalpable recurrent cancers. However, an optimal mammographic strategy will balance the financial costs, and patient concerns (e.g. anxiety) of increased frequency of mammography with the benefits of detecting more impalpable local recurrence. Given that not all women have the same risk of recurrence, the costs and benefits of follow-up mammography may differ between different subsets of patients. It is possible that by defining different risk profiles, we can tailor mammographic schedules that are more effective and efficient.

International guidelines recommend annual follow-up mammography for all women, but there report no randomised controlled trial evidence to support this schedule over any other.[2-7] While randomised controlled trials (RCTs) can provide an unconfounded estimate of effect, there are barriers to performing RCTs in follow-up imaging in cancer. These include difficulties with patient accrual (for example, if patients have a 50% chance of receiving follow-up that is less frequent than current guideline recommendations they are unlikely to consent to be in a trial), and the large sample size and long follow-up (and thus cost implications) required to demonstrate significant differences between alternative programs for different patient subgroups.[8] Many observational studies have investigated the clinical (but not economic) effects of

mammographic surveillance. Reviews have shown these studies to be of poor quality and prone to bias, particularly length bias.[9] As such, observational studies provide a limited basis for the direct, non-modelled evaluation of cost-effectiveness of alternative surveillance strategies.[9]

Decision analytic modelling facilitates data synthesis to describe disease progression over an extended time horizon in order to capture all important differences in costs and benefits between alternative strategies. Benefits are commonly represented as gains in quality adjusted life years (QALYs), where one QALY is equivalent to one additional year of life in perfect health.

In chapter 3 we described the development and calibration of a cost-effectiveness model for the analysis of alternative mammographic follow-up schedules for postmenopausal women who were disease free following primary treatment for moderate prognosis early breast cancer.[10] The aim of this paper is to report the full set of cost-effectiveness results from that model, comparing alternative follow-up mammography schedules for postmenopausal women across four different risk profiles; taking into account age and adherence to mammography

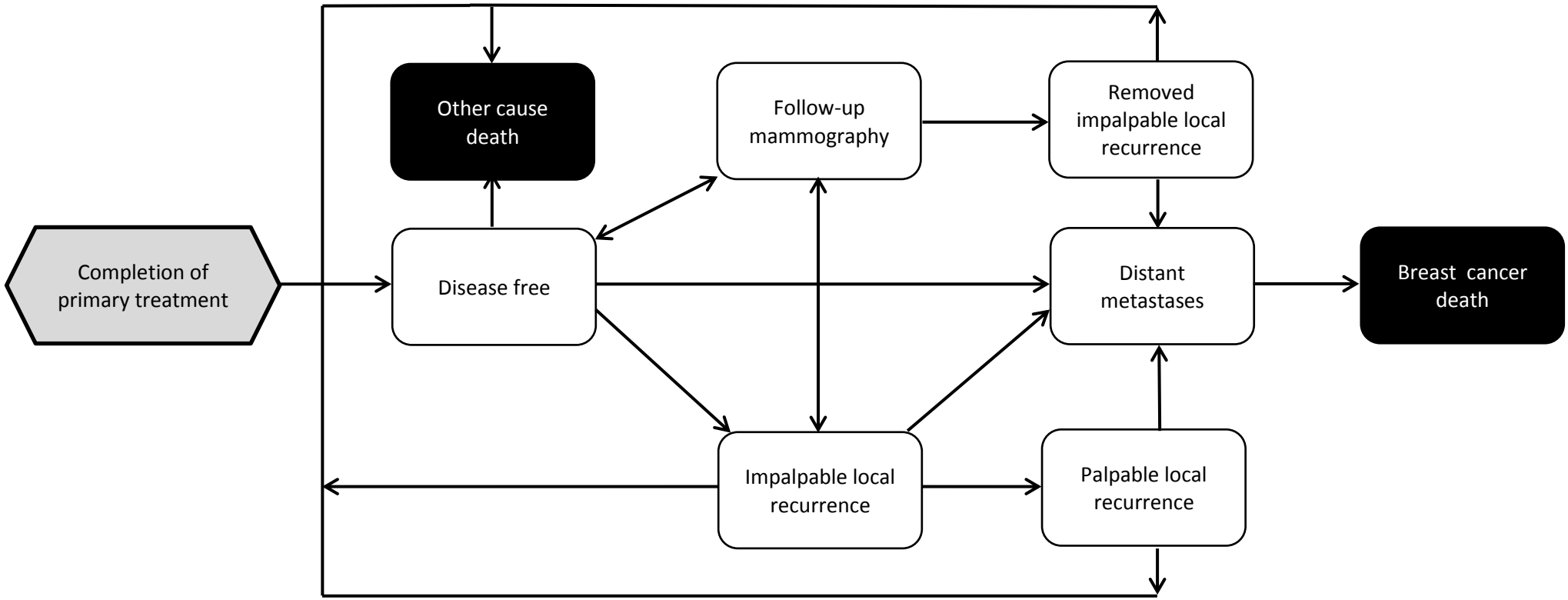
METHODS

The development and population of the model is described in detail in chapter 3. Here a brief summary of the methods is provided.

Model structure

A discrete event simulation (DES) model was developed in Simul8® (figure 1), to represent progression of early breast cancer in women who are disease free after completion of primary treatment. Mammographic follow-up is represented as an instantaneous event. Model assumptions are presented in Box 1.

Figure 1: Structure of DES model describing the possible progression of early breast cancer after completion of treatment, and the detection of impalpable disease by follow-up mammography



Box 1: Model assumptions

1. All patients are women
2. Women aged ≥ 50 years are defined as postmenopausal
3. Local recurrence refers to recurrence of breast cancer in the treated breast/axilla or new primary breast cancers in the contralateral breast
4. A local recurrence is curable, and women with local recurrence will not experience a breast cancer death unless they develop metastatic disease
5. A local recurrence will be initially impalpable and if untreated will continue to increase in size and eventually become palpable
6. Early detection of a local recurrence when impalpable reduces the risk of metastatic disease compared to late detection of a local recurrence when palpable
7. Once a local recurrence (impalpable or palpable) is detected, it will be surgically removed (± adjuvant treatment) rendering the patient disease-free.
8. Distant metastases include all systemic relapses outside the breast/axilla, and includes supraclavicular lymphadenopathy as well as visceral metastases in lung, liver, bone, brain and other sites
9. Women are at risk of developing distant metastases with and without the prior development of local recurrence
10. Distant metastases are incurable and result in death from breast cancer.

Model pathways

Women enter the model disease-free, but are at risk of developing a recurrence. A woman will leave the disease-free state if she develops a local recurrence. Initially this will be an impalpable local recurrence, and only detectable via mammographic surveillance. If an impalpable local recurrence is not detected, and the patient does not die in the intervening period, the recurrent tumour will either continue to grow locally and be detected clinically by the doctor or patient (palpable local recurrence) or metastasize to other parts of the body (distant metastases).

Women can develop distant metastases from any health state, which is typically incurable, and such women are assumed to die of causes related to breast cancer (breast cancer death). Prior to the development of distant metastases, women can die from causes unrelated to breast cancer at any time (other cause death).

For women with an impalpable local recurrence, a true positive follow-up mammogram results in the removal of the lesion (removed impalpable local recurrence). Following a false negative follow-up mammogram, the lesion will remain undiagnosed, and continue to grow. For women in the disease-free state, a false positive follow-up mammogram results in a biopsy and upon a negative result, such women return to the disease-free state.

Model scenarios

Cost-utility analyses of alternative mammographic follow-up schedules were performed across four different risk profiles, based on the Nottingham Prognostic Index of the primary breast cancer. The rationale for the scenarios and the results for the moderate subgroup, are presented in chapter 3. For the excellent, good and poor prognosis subgroups, we also compared the cost-effectiveness of three different mammographic schedules, in two different age groups. Analyses were performed assuming two different levels of adherence to mammographic follow-up (90% and 75%). The analyses performed are presented in Table 1.

Table 1 – Model scenarios for all risk subgroups

Age group	NPI subgroup	Mammography schedule	Adherence with mammography %
50-69 years	Excellent	2 yearly	90
		Mixed	90
		Annual	90
	Good	2 yearly	90
		Mixed	90
		Annual	90
	Moderate	2 yearly	90
		Mixed	90
		Annual	90
	Poor	2 yearly	90
		Mixed	90
		Annual	90
70-79 years	Excellent	2 yearly	90
		Mixed	90
		Annual	90
	Good	2 yearly	90
		Mixed	90
		Annual	90
	Moderate	2 yearly	90
		Mixed	90
		Annual	90
	Poor	2 yearly	90
		Mixed	90
		Annual	90

*Identical analysis also performed assuming 75% adherence with mammography

Model Inputs

As described in chapter 3, we linked routinely collected data sources, to construct a patient level dataset of eleven hundred postmenopausal women diagnosed with early breast cancer between 2000-2008, who had their primary breast cancer treatment and mammographic follow-up in the public health care system.

Women were followed up to death, or 30 June 2011 with respect to the timing of mammographic follow-up, recurrence events, and mortality (breast cancer or other cause mortality). Women were categorised by age and Nottingham Prognostic Index (NPI) category of the primary breast cancer. [11]

Estimates of the health service costs and utility values for each health state; and a quality of life decrement associated with mammography screening; were identified from the literature (Table 2). [12-16] Other cause death rates were calculated by subtracting the proportion of Australian women who died from breast cancer from age specific mortality rates derived from Australian life-tables.[17,18]

Table 2: Baseline costs in $A2011 and utilities, for health states in the model

	Mean values		Best case values**		Worst case values**		
Health states	**Annual costs***	**Utilities**	**Annual costs***	**Utilities**	**Annual costs***	**Utilities**	**References**
Disease-free	$0	0.94	$0	0.94	$0	0.94	15
Surveillance*	$124	-0.01	$124	0	$124	-0.02	16
ILR	$0	0.94	$0	0.94	$0	0.94	15
RILR - surgery costs	$32,323	0.74	$45,773	0.7	$20,032	0.74	15
RILR - ongoing costs	$2,207	0.85	$3,915	0.85	$1,175	0.85	14
PLR - surgery costs	$32,323	0.74	$45,773	0.7	$20,032	0.74	15
PLR - ongoing costs	$2,207	0.85	$3,915	0.85	$1,175	0.85	14
Distant metastases	$17,263	0.5	$22,935	0.35	$12,715	0.65	15

ILR = impalpable local recurrence, RILR= removed impalpable local recurrence, PLR = palpable local recurrence

*Surveillance = costs of follow-up reflect the weighted costs of a single MMG encounter (assuming unilateral mammography in 30% and bilateral mammography in 70% of women) plus a single attendance fee, per year, based on 2011 MBS rates.

**Best case from the perspective of increased surveillance frequency

Model calibration

Model calibration was performed to identify the best fitting sets of input parameter values, according to the seven step approach described by Vanni and Karnon.[19] Four calibration targets (five and ten year in-breast recurrence, and breast cancer mortality rates) were specified for postmenopausal women in each of the four NPI prognostic groups (excellent, good, moderate and poor) (appendix 1). Of the original 1100 postmenopausal women, we were unable to calculate the NPI for 113 women due to missing data for one or more of tumor, node or grade status. Rates of recurrence within the breast were determined from pathology and mammography data for the remaining 987 women, and breast cancer death rates were extracted from the SA Cancer Registry.

Model Analysis

The model was run for the same 2000 sampled sets of convergent input parameter value sets for each scenario. Model outputs were total costs (health state costs + costs of surveillance), and total QALYs across the model health states. Within each age and adherence scenario (e.g. 50-69 year olds assuming 75% adherence), the mean incremental cost effectiveness ratios (ICERs) were estimated between mammographic schedules arranged in increasing order of effectiveness. A probabilistic sensitivity analysis informed probabilities that each follow-up strategy is the most cost-effective for each scenario, at alternative assumed monetary values for the gain of additional QALYs. We created three QALY thresholds to reflect the implied current Australian funding threshold ($50,000), and alternative thresholds below ($25,000) and above ($75,000) the current funding threshold.

Deterministic sensitivity analyses were undertaken around the health state cost and utility weight input parameters, including the utility decrement associated with follow-up mammography. Best and worst case scenarios were defined from the perspective of favouring more intensive surveillance. For the best case scenario, higher costs for removed impalpable local recurrence ($45,773 and $3,915 in the first and following years, respectively), palpable local recurrence ($45,773 and $3,915) and distant metastases ($22,935) were used; with lower utility values for the RILR – surgery (0.7),

PLR – surgery (0.7) and distant metastases (0.35) health states; compared to the base case (Table 3). The reverse is tested for the worst case scenario. That is, lower costs for removed impalpable local recurrence ($20,032 and $1,175), palpable local recurrence ($20,032 and $1,175) and distant metastases ($12,715); with identical utilities for the RILR – surgery (0.74) and PLR – surgery (0.74) health states, but higher utilities associated with distant metastases (0.65) than the base case (Table 2). The utility values for the surgery states in the worst case scenario are unchanged from the base case, but the utility decrement associated with surveillance is increased (see below).

In the base-case the applied QALY decrement associated with follow-up mammography of 0.01 is equivalent to a 0.5 utility decrement for one week, or a 0.25 utility decrement for two weeks, reflecting the heightened anxiety around the surveillance period. The QALY decrement also captures false positive effect for the small proportion of patient who undergo further investigations to rule out recurrence. For the best case scenario, the utility decrement for surveillance was decreased to zero, and in the worst case scenario, the utility decrement for surveillance is increased (-0.02).

RESULTS

Table 3 reports the mean ICERs for increasing frequencies of mammographic surveillance, by age cohort and prognostic subgroup of primary tumor; assuming 90% adherence with surveillance mammography.

Table 3: ICER for increasing frequency of mammographic surveillance, by age cohort and prognostic subgroup of primary tumour, assuming 90% adherence and 0.01 quality of life decrement with surveillance mammography

Cohort	NPI subgroup	Mammography schedule	%Breast cancer deaths	Mean costs	Mean QALYs	Cost difference	QALY difference	ICER
50-69 years	excellent	2 yearly	0.12	$2,613	12.509			
		mixed	0.12	$2,807	12.510	$193	0.001	$166,413
		annual	0.12	$3,333	12.511	$526	0.001	$656,957
	good	2 yearly	0.18	$3,470	12.205			
		mixed	0.18	$3,669	12.207	$199	0.002	$88,498
		annual	0.18	$4,188	12.210	$519	0.003	$170,989
	moderate	2 yearly	0.33	$6,302	11.258			
		mixed	0.33	$6,535	11.263	$234	0.005	$48,583
		annual	0.33	$6,985	11.266	$450	0.004	$121,227
	poor	2 yearly	0.54	$10,327	8.982			
		mixed	0.54	$10,551	8.993	$224	0.011	$20,121
		annual	0.54	$10,922	9.000	$371	0.007	$50,336
70-79 years	excellent	2 yearly	0.06	$1,704	8.535			
		mixed	0.06	$1,888	8.535	$185	0.000	$460,756
		annual	0.06	$2,178	8.535	$290	0.000	Dominated
	good	2 yearly	0.10	$2,268	8.409			
		mixed	0.09	$2,457	8.411	$189	0.001	$139,898
		annual	0.09	$2,747	8.412	$290	0.001	$363,074
	moderate	2 yearly	0.19	$4,264	7.978			
		mixed	0.19	$4,467	7.979	$203	0.002	$113,078
		annual	0.19	$4,739	7.980	$272	0.001	$254,030
	poor	2 yearly	0.39	$8,329	6.625			
		mixed	0.39	$8,537	6.633	$209	0.008	$26,857
		annual	0.39	$8,755	6.635	$217	0.002	$99,632

NPI = Nottingham Prognostic Index, QALY = Quality adjusted life year, ICER = Incremental cost-effectiveness ratio, Mixed schedule = annual mammography for 5 years, then 2 yearly thereafter

For women aged 50-69 years with excellent prognosis tumors, moving from a 2 yearly to a mixed schedule gains additional QALYs at an incremental cost of $166,413. For women aged 70-79 years, the ICER is very high, at $460,756.

In the good prognosis, 50-69 year age group, the mixed surveillance strategy has a mean ICER of $88,498, which increases to $139,898 for women aged 70-79 years.

For women with moderate prognosis tumors, the mixed mammographic schedule has an ICER of $48,583 for women aged 50-69 years, and $113,078 for women aged 70-79 years.

For women with poor prognosis tumors, the ICER for the mixed mammographic schedule ranges from $20,121 for women 50-69 years, to $26,857 for women 70-79 years.

For the excellent, good, and moderate prognostic groups, the ICER of annual surveillance compared to a mixed schedule is very unlikely to be considered cost-effective. However, in the poor group, annual mammographic surveillance may be potentially cost-effective for women aged 50-69 years, with an ICER of $50,336, though for women aged 70-79 years, the ICER remains high at $99,632.

Assuming 75% adherence with mammography (Table 4) improves the cost-effectiveness of the more intensive follow-up schedules. In some cases, the assumed level of adherence may influence the decisions around the cost-effectiveness of the alternative schedules. For example, the ICER for the mixed mammographic schedule for women aged 50-69 years with moderate prognosis tumors reduces to $29,241 (from $48,583), and for 50-69 year old women with poor prognosis tumors, the ICER for annual mammography reduces to $40,087 (from $50,336).

Table 4: ICER for increasing frequency of mammographic surveillance, by age cohort and prognostic subgroup of primary tumour, assuming 75% adherence and 0.01 quality of life decrement with surveillance mammography

Cohort	NPI subgroup	Mammography schedule	%Breast cancer deaths	Mean costs	Mean QALYs	Cost difference	QALY difference	ICER
50-69 years	excellent	2 yearly	0.12	$2,479	12.508			
		mixed	0.12	$2,666	12.509	$187	0.001	$132,512
		annual	0.12	$3,076	12.510	$410	0.001	$496,847
	good	2 yearly	0.18	$3,330	12.202			
		mixed	0.18	$3,523	12.205	$193	0.003	$64,589
		annual	0.18	$3,931	12.207	$408	0.003	$157,267
	moderate	2 yearly	0.33	$6,090	11.260			
		mixed	0.33	$6,311	11.267	$222	0.008	$29,241
		annual	0.33	$6,676	11.269	$364	0.001	$289,910
	poor	2 yearly	0.54	$10,162	8.949			
		mixed	0.54	$10,419	8.970	$257	0.021	$12,128
		annual	0.54	$10,753	8.979	$335	0.008	$40,087
70-79 years	excellent	2 yearly	0.06	$1,611	8.534			
		mixed	0.06	$1,786	8.535	$175	0.001	$335,523
		annual	0.06	$2,005	8.534	$218	0.000	Dominated
	good	2 yearly	0.10	$2,173	8.408			
		mixed	0.10	$2,354	8.410	$181	0.002	$109,761
		annual	0.09	$2,571	8.411	$217	0.001	$330,244
	moderate	2 yearly	0.19	$4,187	7.970			
		mixed	0.19	$4,375	7.973	$189	0.003	$62,405
		annual	0.19	$4,589	7.975	$214	0.001	$150,911
	poor	2 yearly	0.39	$8,221	6.619			
		mixed	0.39	$8,430	6.628	$210	0.008	$25,037
		annual	0.39	$8,592	6.630	$162	0.002	$65,774

NPI = Nottingham Prognostic Index, QALY = Quality adjusted life year, ICER = Incremental cost-effectiveness ratio, Mixed schedule = annual mammography for 5 years, then 2 yearly thereafter

The absolute ICER estimates were not very sensitive to the assumed utility decrement associated with surveillance episodes, though, assumptions around the utility (or disutility) effects of surveillance may influence some decisions. Assuming 90% adherence, removing the utility decrement reduces the ICER for the mixed mammographic schedule for women aged 50-69 years with moderate prognosis tumors from $48,583 to $39,909, and the ICER for annual mammography for 50-69 year old women with poor prognosis tumors reduces from $50,336 to $47,342.

The results of the probabilistic sensitivity analysis are shown in Table 5. They represent the probability that each surveillance strategy is the most cost-effective strategy for each prognostic sub-group, assuming alternative monetary values for the gain of additional QALYs and alternative levels of adherence. The results show that we can be most certain that a 2 yearly mammographic schedule is cost-effective for the excellent prognosis sub-group. Assuming a $50,000 value of a QALY, and 75% adherence, 2 yearly mammography has a 60% probability of being cost-effective for excellent prognosis 50-69 year old women, and a 70% probability for 70-79 year old women.

Table 5: Probabilistic sensitivity analysis – probability that given mammographic surveillance schedule is cost-effective*

	90% adherence with given mammography schedule				75% adherence with given mammography schedule			
Prognostic subgroup	**QALY threshold**	**Annual**	**Mixed**	**2 yearly**	**QALY threshold**	**Annual**	**Mixed**	**2 yearly**
Excellent								
50-69 years	$25k	0.00	0.17	0.83	$25k	0.00	0.20	0.80
	$50k	0.04	0.32	0.64	$50k	0.07	0.33	0.60
	$75k	0.10	0.35	0.54	$75k	0.13	0.38	0.49
70-79 years	$25k	0.00	0.08	0.92	$25k	0.00	0.11	0.89
	$50k	0.03	0.24	0.74	$50k	0.05	0.25	0.70
	$75k	0.07	0.31	0.62	$75k	0.11	0.32	0.57
Good								
50-69 years	$25k	0.01	0.24	0.75	$25k	0.02	0.28	0.70
	$50k	0.11	0.35	0.53	$50k	0.14	0.38	0.48
	$75k	0.22	0.36	0.42	$75k	0.24	0.39	0.38
70-79 years	$25k	0.01	0.16	0.83	$25k	0.02	0.18	0.8
	$50k	0.09	0.31	0.61	$50k	0.02	0.18	0.8
	$75k	0.18	0.35	0.47	$75k	0.02	0.18	0.8
Moderate								
50-69 years	$25k	0.05	0.37	0.58	$25k	0.11	0.39	0.49
	$50k	0.21	0.39	0.40	$50k	0.27	0.40	0.33
	$75k	0.30	0.37	0.33	$75k	0.35	0.38	0.27
70-79 years	$25k	0.06	0.25	0.70	$25k	0.08	0.30	0.62
	$50k	0.18	0.34	0.49	$50k	0.21	0.36	0.43
	$75k	0.24	0.36	0.40	$75k	0.26	0.36	0.38
Poor								
50-69 years	$25k	0.20	0.42	0.38	$25k	0.25	0.46	0.29
	$50k	0.39	0.39	0.22	$50k	0.40	0.43	0.17
	$75k	0.48	0.35	0.18	$75k	0.46	0.40	0.14
70-79 years	$25k	0.15	0.39	0.46	$25k	0.22	0.38	0.41
	$50k	0.31	0.41	0.28	$50k	0.36	0.38	0.26
	$75k	0.36	0.40	0.24	$75k	0.40	0.37	0.23

*Probability a given surveillance strategy is cost-effective, by prognostic subgroup, age, adherence and assumed monetary value of a QALY
QALY threshold = $25,000, $50,000, or $75,000 (the value attached to the gain of additional QALYs).
Annual = annual mammography, Mixed = annual mammography for 5 years, then 2 yearly thereafter, 2 yearly = mammography every 2 years

DISCUSSION

Our incremental economic analysis using a discrete event simulation model suggests that, for most postmenopausal women, annual mammographic follow-up may not be cost-effective. Our results suggest that for women with excellent prognosis tumors, two yearly follow-up mammograms are cost-effective regardless of age (that is, for women aged 50-69 years and 70-79 years). For women with good and moderate prognosis tumors, there is some uncertainty regarding the most cost-effective schedule. If high adherence rates can be achieved, 2 yearly surveillance for good prognosis women is likely to be cost-effective. For moderate prognosis women, a mixed surveillance strategy may be most cost-effective, especially in the younger age group. For women with poor prognosis tumors, annual mammography is potentially cost-effective for women aged 50-69 years, but the mixed schedule is the most cost-effective schedule for women aged 70-79 years.

A recent Health Technology Assessment of surveillance strategies for women with early breast cancer in the UK, used an alternative modelling approach that represents the natural history in greater detail, and facilitates analyses of a broader set of follow-up options.[9] Using a cohort Markov model, Robertson et al, represented disease progression in more detail, but they did not calibrate (or externally validate) their model. Different frequencies (12,18, 24 and 36 month intervals) and different methods of surveillance (no surveillance, mammography, mammography + clinical examination, MRI + clinical examination) were tested, including sub-group analyses of two hypothetical women aged 40 years and 70 years, representing respectively, higher and lower likelihoods of relapse within the breast (based on increased risk of recurrence, shorter time for an undetected cancer to progress to worse risk profile, and more aggressive and costly treatment of recurrence in the younger woman). Their results suggested that more intensive follow-up of women judged to be at high risk, and less intensive follow-up of women judged to be a low risk, may be cost-effective.[9] This is consistent with findings reported in this paper, which reports results for the early, good, moderate and poor risk groups. Although different modelling approaches were used, their results are broadly similar, which reinforces our findings that a one-size-fits-all,

annual surveillance strategy may not be the optimal approach to monitoring breast cancer survivors.

There are no randomised controlled trials to date, that examine the optimal frequency of mammographic follow-up in women who are disease free following primary treatment for early breast cancer. In the likely continued absence of RCTs in this field, we need to embrace alternative research techniques, such as decision analytic modelling.

The key strengths of our DES model include the use of probabilistic model calibration and the use of longitudinal data that includes patient level surveillance pathways and outcomes. In cancer follow-up, we cannot observe the time at which patients develop asymptomatic recurrence. Calibration is often used to fit values for these parameters, such that the models' outputs match some observed data for the population being evaluated.[19] This process is most common in models of population based screening programs,[20] though previous models have not used patient-level data to represent observed surveillance pathways to inform the calibration of the underlying disease progression parameters. The use of longitudinal data that includes patient-level surveillance pathways and outcomes, may provide a more robust basis for calibrating surveillance models, and increase confidence that the model reflects reality. We used linked, longitudinal individual patient level data that represented the observed follow-up mammography schedule received by each of the 987 patients. This approach precluded the secondary parameterisation of surveillance frequency, an important parameter that is generally subject to significant uncertainty. In addition, the methods described in this paper could readily be applied to optimising follow-up schedules for other cancer types.

The main limitations of the reported study relate to the data sources. Recurrence and follow-up data is not routinely collected by the South Australian Cancer Registry. This meant that data had to be extracted and linked across four separate data sources (each built for different purposes), which resulted in some uncertainties regarding the capture

of all relevant follow-up and recurrent events. Hormonal receptor data was not collected in the South Australian Cancer Registry prior to June 2012, and HER2 receptor status is not yet collected. Breast density was not routinely included in mammography reports in South Australia during the period 2000-2008. In addition, our data was limited to the public health sector, as privacy laws prevented access to mammography reports from the private health sector.

We chose to combine ipsilateral recurrence in the treated breast/axilla (IBTR), and new contralateral breast cancer (CLBC) as "local recurrence". This was due to the much smaller annual event rate for CLBC[1,2] and the findings of an economic evaluation in the 2011 Health Technology Assessment (HTA) by Robertson et al.[9] In the HTA model analysis, the effect of an IBTR and MCBC on risk of breast cancer death were assumed to be identical. Subsequently IBTR was modelled alone to take into account the substantially higher mortality associated with IBTR as identified in their datasets. The results of both analyses were essentially the same; that is, not including CLBC in the analysis made no difference to the results as the results were driven by IBTR.

The model could be expanded to include premenopausal women and women over the age of 80 years; to study a range of adherences between 55% and 100%, and to subdivide the primary breast cancer into six NPI prognostic subgroups.[21] We did not model alternative surveillance strategies incorporating other imaging modalities (e.g. MRI, ultrasound) as these are not currently recommended in international guidelines for routine breast cancer follow-up. Other frequencies of mammographic surveillance could also be explored, but less than two yearly follow-up is unlikely to be accepted by most breast cancer survivors.

To maximise the potential of decision analytic modelling to guide clinical practice, we need to improve the quality of the data that inform such models. We need to ensure we have robust data on primary tumor stage at diagnosis, and ideally to extend the data that we collect to include receptor status and details of the primary treatment received (both local and adjuvant). More importantly, we need more complete capture of

recurrence data in our population based cancer registries. We need to know the timing and site of first relapse (treated breast, opposite breast, distant relapse) and the method of detection of relapse (imaging detection of an impalpable lesion or clinical detection of a palpable lesion). We also need easily accessible granular data on the frequency and results of clinical examinations and diagnostic imaging to assess adherence with follow-up, and to better inform the calibration of cost-effectiveness models.

Avenues for further exploration could include validating our current model with a larger more complete dataset; incorporating additional prognostic detail from the primary tumor (eg. receptor status, histologic subtype, lymphovascular invasion etc), or using other prognostic indices (e.g.Adjuvant! Online);[22,23] and to explore the impact of primary treatment on the cost-effectiveness of surveillance.

CONCLUSION

Our results suggest that annual mammographic follow-up is not cost-effective for most postmenopausal women, and that mammographic follow-up can be tailored according to risk of recurrence based on the NPI score of the primary breast cancer and age at diagnosis. Whilst models rely on existing data, their strength lies in the ability to be rapidly updated and extended in response to new knowledge. If validated with a more robust dataset, our model could potentially provide the foundations towards a significant change to current mammographic and diagnostic imaging practice in breast cancer follow-up.

Appendix 1: Observed event rates for 987 postmenopausal women treated for early breast cancer in South Australia between 2000-2008, by Nottingham Prognostic Index of their primary breast cancer

Prognostic subgroup	Women	Event	Duration	Events	Patient years at risk	Annual incidence	Standard deviation	lower 95% CI	Upper 95% CI
Excellent	176	Recurrence	5 year	1	515.5	0.0019	0.002	0.00005	0.00714
			10 year	2	1180.5	0.0017	0.001	0.00021	0.00471
		Death	5 year	1	515.5	0.0019	0.002	0.00005	0.00714
			10 year	2	1182.5	0.0017	0.001	0.00021	0.00471
Good	250	Recurrence	5 year	3	732.5	0.0041	0.002	0.00085	0.00984
			10 year	9	1656	0.0054	0.002	0.00249	0.00950
		Death	5 year	1	735.5	0.0014	0.001	0.00003	0.00501
			10 year	5	1678.5	0.0030	0.001	0.00097	0.00609
Moderate	407	Recurrence	5 year	19	1768	0.0107	0.002	0.00649	0.01605
			10 year	24	2407	0.0100	0.002	0.00640	0.01431
		Death	5 year	17	1802.5	0.0094	0.002	0.00551	0.01438
			10 year	28	2483	0.0113	0.002	0.00751	0.01579
Poor	154	Recurrence	5 year	9	435	0.0207	0.007	0.00953	0.03600
			10 year	12	780	0.0154	0.004	0.00798	0.02513
		Death	5 year	22	443	0.0497	0.010	0.03145	0.07171
			10 year	32	829.5	0.0386	0.007	0.02657	0.05269

Note – Complete data for 987 of 1100 postmenopausal women, recurrence = relapse within the treated breast or axilla, or a new primary or metastasis in the opposite breast

REFERENCES

1. New Zealand Guidelines Group. Management of early breast cancer. Wellington: New Zealand Guidelines Group; 2009.
2. National Institute for Health and Clinical Excellence. Early and locally advanced breast cancer: diagnosis and treatment. National Collaborating Centre for Cancer, Cardiff, Wales, 2009.
3. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
4. Khatcheressian J, Hurley P, Bantug E, Esserman L, Grunfeld E, Halberg F, Hantel A, Henry N, Muss H, Smith T, Vogel V, Wolff A, Somerfield M, Davidson N; American Society of Clinical Oncology. Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol. 2013 Mar 1; 31(7): 961-5
5. Aebi S, Davidson T, Gruber G, Cardoso F; ESMO Guidelines Working Group. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2011 Sep; 22 Suppl 6: vi12-24
6. Scottish Intercollegiate Guidelines Network. Management of breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network, Edinburgh, Scotland, 2005.
7. Grunfeld E, Dhesy-Thind S, Levine M; Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment for breast cancer (summary of the 2005 update). CMAJ. 2005 May 10;172(10) :1319-1320
8. Grunfeld E, Noorani H, McGahan L, Paszat L, Coyle D, van Walraven C, Joyce J, Sawka C. Surveillance mammography after treatment of primary breast cancer: a systematic review. Breast. 2002 Jun; 11(3): 228-35
9. Robertson C, Arcot Ragupathy S, Boachie C, Dixon J, Fraser C, Hernandez R, Heys S, Jack W, Kerr G, Lawrence G, MacLennan G, Maxwell A, McGregor J, Mowatt G, Pinder S, Ternent L, Thomas R, Vale L, Wilson R, Zhu S, Gilbert F.

The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews registry database analyses and economic evaluation. Health Technol Assess. 2011 Sep; 15(34): v-vi, 1-322.

10. Bessen T, Karnon J. A patient-level calibration framework for evaluating surveillance strategies: a case study on mammographic follow-up after early breast cancer. Value in Health, under review.
11. Galea M, Blamey R, Elston C, Ellis I. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992; 22(3): 207-219.
12. Karnon J, Kerr G, Jack W, Papo N, Cameron D. Health care costs for the treatment of breast cancer recurrent events: estimates from a UK-based patient-level analysis. Br J Cancer. 2007 Aug 20; 97(4): 479-485.
13. Hind D, Ward S, De Nigris E, Simpson E, Carroll C, Wyld L. Hormonal therapies for early breast cancer: systematic review and economic evaluation. Health Technol Assess. 2007 Jul; 11(26): iii-iv, ix-xi, 1-134
14. CEA Registry Website. https://research.tufts-nemc.org/cear4/default.aspx
15. Tengs T, Wallace A. One thousand health-related quality-of-life estimates. Med Care. 2000 Jun; 38(6):583-637
16. de Koning H, van Ineveld B, van Oortmarssen G, de Haes J, Collette H, Hendriks J, van der Maas P. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer. 1991 Oct 21; 49(4):531-7.
17. Australian Institute of Health and Welfare (AIHW) 2012. ACIM (Australian Cancer Incidence and Mortality) Books. AIHW: Canberra. Breast cancer (ICD10 C50), Australia, 1968-2007. Age specific mortality rates. http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=60129542435 (last accessed 6 August 2013)
18. Australian Bureau of Statistics. 3302.0.55.001 - Life Tables, Australia, 2003-2005. http://abs.gov.au/AUSSTATS/abs@.nsf/allprimarymainfeatures/6A75A50AE791F95FCA25738D000F462B (last accessed 6 August 2013)

19. Vanni T, Karnon J, Madan J, White R, Edmunds W, Foss A, Legood R. Calibrating models in economic evaluation: a seven step approach. Pharmacoeconomics. 2011 Jan: 29(1):35-49

20. Karnon J, Goyder E, Tappenden P, McPhie S, Towers I, Brazier J, Madan J. A review and critique of modelling in prioritising and designing screening programmes. Health Technol Assess. 2007 Dec; 11(52):iii-iv, ix-xi, 1-145.

21. Blamey R, Ellis I, Pinder S, Lee A, Macmillan R, Morgan D, Robertson J, Mitchell M, Ball G, Haybittle J, Elston C. Survival of invasive breast cancer according to the Nottingham Prognostic Index in cases diagnosed in 1990-1999. Eur J Cancer. 2007 Jul; 43(10): 1548-55.

22. Ravdin P, SiminoffL, Davis G et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. J Clin Oncol. 2001 Feb 15; 19(4): 980-91

23. Hajage D, de Rycke Y, Bollet M, Savignoni A, Caly M, Pierga J, Horlings H, Van de Vijver M, Vincent-Salomon A, Sigal-Zafrani B, Senechal C, Asselain B, Sastre X, Reyal F. External validation of Adjuvant! Online breast cancer prognosis tool. Prioritising recommendations for improvement. PLoS One. 2011; 6(11): e27446

CHAPTER 4 EXTENSION: Cross validation of our economic evaluation

Please note:

- References 16 and below refer to the papers included in our model results paper.
- References 93-100 refer to papers not included in model results paper. These papers are included in the reference list at the end of the thesis.

In 2011, the NHS published a Health Technology Assessment to address "The clinical effectiveness and cost effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews, registry databases analyses and economic evaluation".[9] At the time of writing, this is the only other model-based economic evaluation of alternative surveillance strategies following primary treatment for early breast cancer. The key features of the HTA model are presented below, and compared with my study. Specifically I compare the model structure, data used to populate the model, follow-up schedules examined, analyses and results. Certain aspects of the methodology were difficult to assess, in these areas the uncertainty is described, and my best interpretation of the applied methods is presented.

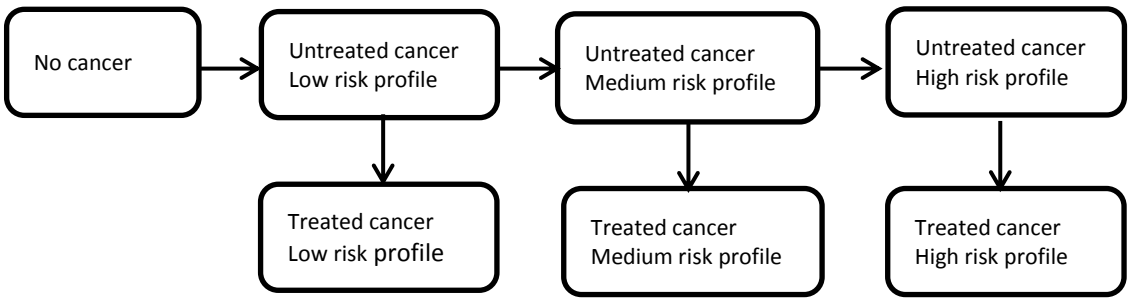
HTA MODEL

MODEL STRUCTURE

The HTA reports the use of a Markov model to describe possible paths of individuals who are disease free after treatment of their primary breast cancer. Women enter the model in the "no cancer" health state. Women can develop a second cancer in the treated or opposite breast. The model represents recurrent cancer from the point at which it becomes detectable by mammography. It is assumed that recurrent cancers differ in terms of risk of death, and that all begin with a low-risk profile. Figure 2 is a simplified model structure, which shows that over time, an undetected lesion will increase in size and severity, progressing from low-risk to medium-risk to high-risk (note that within the text, the authors describe five risk profiles, but these are not all represented in their figure). Women undergo periodic surveillance and the recurrent lesions may be either detected and treated (true positive), or undetected and untreated (false negative).

Based upon advice of the clinical members of the Advisory Group, the authors estimated that 14% of women with a personal history of breast cancer who have not been diagnosed with a recurrent cancer will present to a GP every 6 months. Following clinical examination, women may be referred for an unscheduled mammogram, which may result in the detection of recurrent breast cancer in the interval between surveillance episodes.

Figure 2: simplified HTA model structure



Risk profile refers to the mortality risk for a given recurrent cancer within the breast.

Natural History

Using data from the West Midlands Cancer Intelligence Unit (WMCIU) Breast Cancer Registry database, the estimated 10 year failure probability for ipsilateral breast tumor recurrence (IBTR) and metachronous contralateral breast cancer (MCBC) was determined following breast conserving surgery or mastectomy. An exponential curve was fitted to estimate the 6 month cycle failure probability. The model assumed IBTR and MCBC were independent and the individual rates were summed to provide a net rate of cancer incidence.

From earlier analysis of the WMCIU Breast Cancer Registry database it was identified that the incidence of second cancers within the breast was in part predicted by features of the primary breast cancer (tumor size, nodal status, grade of tumor and vascular invasion). To represent varying level of recurrence risk, the 10 year failure probability was estimated for a reference patient group (age 50-64 years, with tumor size <1cm, grade 1, no evidence of vascular invasion and lymph node negative). A Cox proportional

hazard model was used to generate hazard ratios for a range of alternative patient groups, which were applied to the reference case event rates.

Follow-up strategy

The HTA model facilitated analysis of a broader set of follow-up options than in our model, assessing four different strategies and four different frequencies of follow-up.

Follow-up strategies for women treated with breast conserving surgery included no surveillance, mammography alone, mammography + clinical examination, and MRI + clinical examination. Women treated with mastectomy were not offered the MRI + clinical examination strategy. Frequencies of follow-up examined were 12, 18, 24 and 36 month intervals.

For women who develop an ipsilateral breast tumor recurrence, the no surveillance option was replaced by GP only for the 18 month, 24 month and 36 month intervals, but otherwise the strategies were identical to the breast conserving surgery group.

The mammography result was considered as a true positive or a false negative. Uptake of surveillance mammography was assumed to be 100%. The authors also specified the sensitivity and specificity of clinical examination when performed by a general practitioner, and assumed it to be slightly lower than the values reported in their systematic review of diagnostic performance in routine surveillance.

Breast cancer death rates

Recurrent cancers were grouped into risk profiles according to predicted 10 year mortality rates as determined by the Advisory Group, based on mortality data within Adjuvant! Online.[93] (HTA Appendix 30, page 305). The mortality estimates assume that breast cancer is detected and treated, that the risk of death is independent of women's age, and that death occurs at a constant rate over time.

It was hypothesized that a woman with untreated cancer would face a higher risk of death than an identical woman with treated cancer. This increased risk of death was proxied by comparing expected 5 year mortality rates for women diagnosed between 1980-1984 with women diagnosed from 2000-2004. Across all women aged 15-74 years at diagnosis, 5 year mortality was 34.9% if the cancer was detected between 1980 and 1984 and 14.9% mortality rate if the cancer was detected between 2000 and 2004. The ratio of deaths in the earlier time period to the later time period (2.34) was used to inflate the mortality rates derived from Adjuvant! Online[93] for women with undetected cancer.

There are two areas relating to breast cancer mortality that are unclear:

- Risk profile - while figure 1 shows three risk profiles, it is not clear how the five risk profiles used in the evaluation fit into the model structure, nor how the eight risk profiles in Adjuvant! Online[93] have been condensed down to five (that is, which combination of prognostic factors fall into each profile). While the risk profiles are based on mortality risk, they are not explicitly defined. Based on the utility weights applied to each of the five stages, we assume that risk profiles 3&4 are regional and risk profile 5 is metastatic disease, and hence that risk profiles 1&2 are local disease.
- The model does not explicitly describe a pathway for women whose first site of relapse is distant. We assume the model does not allow for the incidence of metastatic disease in the absence of prior local recurrence. The rationale for this assumption is not stated, and is at variance with the reporting of early breast cancer clinical trials, which report the incidence of metastatic disease in the absence or presence of local disease.[94,95] Most published models of early breast cancer included a direct transition between disease-free and metastatic disease.[96]

Disease progression

The HTA model assumes that IBTR and MCBC began in risk profile 1 at the point at which they are detectable; that lesions are detectable at 0.75cm; and if a lesion is detectable but undetected it will continue to grow and increase in severity over time.

A structured literature review relating to doubling time of a breast cancer informed plausible estimates of the range of rates of growth at which an undiagnosed cancer may move to a worse risk profile. The data used was based on work by Peer and colleagues, taking the mean doubling time in tumor volume to be 157 days for women aged 50-70 years. The HTA team then calculated that it would take 540 days for a 0.75cm tumor to progress to the next risk profile group. This equated to a six monthly rate of 0.1555 for increasing risk profile in this age group. A sensitivity analysis explored the consequences of changing the risk of unmanaged women progressing to higher risk profiles by altering the base-case estimate of 0.1555 to 0.2623 and 0.0923 per six month cycle.[97]

Health state utility values

Utilities were assumed to decrease with increasing severity of breast cancer, and be sensitive to use of chemotherapy. For each risk profile, the percentage of women on chemotherapy was determined, and utility values for “on treatment”, “after treatment” and “unmanaged” were estimated.

Using published literature,[98] the utility values for each of the five risk profiles were defined and adjusted to include a decrement[99] for those women who received chemotherapy. Utility values for risk profiles 3 & 4 were based on women aged 50-59 years of age with regional cancer[100] and for risk profile 5 on women aged 50-59 years with distant cancer.

ANALYSES

The base-case analysis assessed the impact of four different surveillance strategies (no surveillance, mammography alone, mammography + clinical examination, and MRI + clinical examination) at four different intervals (12,18,24 and 36 months), by type of

initial surgery for primary breast cancer (breast conserving surgery or mastectomy). The effect of substituting MRI for mammography was only assessed in the group that received breast conserving surgery. Net benefits (£) were calculated and ranked at thresholds of £20,000 and £30,000. For the purpose of these analyses, the effect of an IBTR and MCBC on risk of breast cancer death were assumed to be identical.

Modelling of IBTR alone was then performed to take into account the substantially higher mortality associated with IBTR as identified in analyses of WMCIU and Edinburgh datasets.

The HTA performed a large range of sensitivity analyses. Specifically, sensitivity analyses around the breast conserving model included both one-way and multi-way sensitivity analyses. One-way sensitivity analyses included varying the probability of recurrence, inflating the risk of death from cancer, inflating the risk of cancer progression in undiagnosed or untreated cancers, varying accuracy of diagnostic tests, and varying costs (treatment, imaging tests, clinical examination) and utility estimates. A multi-way sensitivity analysis was performed to explore the impact of age at the time surveillance commenced. Surveillance regimens were evaluated for a hypothetical (assumed high risk) 40 year old woman and a hypothetical (assumed low risk)70 year old woman. For the younger woman, the incidence of recurrence was increased, the time taken for undetected cancer to progress to a more severe profile was reduced, and treatment costs were increased (treated more aggressively). For the older woman, all parameters were assumed to be as base-case assumptions apart from starting age of 70 years, and the incidence of cancer was based on lowest hazard estimates. There was less need for separate sensitivity analyses of clinical parameters in our model due to the calibration process. We reported sensitivity analyses around costs and utilities, as these were not included in the calibration process.

RESULTS

Incremental cost-effectiveness ratios are presented within follow-up intervals, rather than between follow-up intervals, eg. for a 12 month interval, the cost-effectiveness of

no surveillance (or GP only), MMG alone, MMG + clinical, and MRI + clinical are assessed. Separate results are reported for 12, 18, 24 and 36 month intervals.

In the base-case analysis, for all active surveillance strategies, the ICER falls as the surveillance interval increases, with the reduction in QALYs "more than compensated for" by a reduction in costs.

Mammography alone is the most cost-effective option within each time interval for the base case and most sensitivity analyses, with an ICER of less than £5,000 in most cases. The only exception is for the multi-way sensitivity analysis based on surveillance of a 40 year old woman. In this patient group, the ICER for MRI + clinical is less than $30,000 for a 12-months interval, and less than $20,000 for all other intervals.

Net benefits are also presented, which enables a comparison across the full set of 16 alternative surveillance options. At thresholds of £20,000 and £30,000 per QALY, mammography alone at 12 month intervals has the highest net benefit in the base case and most sensitivity analyses. At a threshold of $20,000 MMG alone even has the highest net benefits for the high risk 40 year old patient. However, at a threshold of £30,000, MRI + clinical examination at 12 month intervals has the highest net benefit for the 40 year old patient. , followed by mammography + clinical examination at 12 month intervals. For a 70 year old low risk woman, mammography alone has the highest net benefit at 36 month intervals if the threshold is £20,000, and at 24 month intervals if the threshold is increased to £30,000.

No probabilistic sensitivity analyses were reported.

OUR MODEL

MODEL STRUCTURE

The structure of the model I have developed (page 120) is significantly different to the HTA model (page 142).

In terms of the represented natural history of breast cancer, my model represents impalpable (asymptomatic) and palpable (symptomatic) local recurrence states, and a distant (metastatic) recurrence state. Impalpable local disease can be detected via surveillance mammography, otherwise local recurrence is assumed to be detected and treated if and when it moves into the palpable state.

The model structure is less complex than the HTA model, which represents five disease states, and a pathway of detection via GP referrals. The HTA model implies a constant likelihood of non-mammographic detection across the disease states, an assumption which seems unrealistic.

The other key structural difference concerns the assumption in my model that patients can move directly from disease-free to distant recurrence, whilst the HTA model implies all patients experience a local recurrence initially. The former assumption would seem to be supported by available evidence.[94,95]

Model implementation

Despite the reduced complexity of my model, it was implemented as a DES model, whilst the HTA model was implemented as a cohort-based state transition model. DES models are generally used to represent more complex model structures, but in my case DES was used due to the data and processes used to populate the model. This factor introduces the most important difference between the models, that my model was calibrated to observed model output parameter values, whilst the HTA model was not subject to any form of calibration or external validation.

Model population

The natural history model parameters in my model were calibrated to input (surveillance) and output (recurrence and mortality) data derived from a linked dataset of routinely collected data sources. Model output parameters for recurrence and breast

cancer death were compared to observed values, and sets of input parameters that predicted the observed outputs with sufficient accuracy were selected for the main analysis of the model.

The HTA model was populated using direct estimates of each input parameter. Some of the input parameters were subject to bias. For example, incident recurrence rates were informed by rates of observed (or detected) recurrence, but the recurrence may have been incident but not necessarily detected. Other input parameters were subject to high levels of uncertainty, for example, the proportion patients visiting a GP every six months, and the sensitivity of GP referrals for an interval mammography.

Follow-up strategy

The follow-up strategy assessed was mammography, as this the only imaging modality currently recommended in international guidelines for routine follow-up after breast cancer. The decision not to consider mammography with different strategies of clinical follow-up was essentially a pragmatic one. Without patient-level data describing the timing and frequency of clinical examination, and the mode of detection of recurrent tumours, there was too little data to inform such an extension to the model. To inform this option would have required casenote review for the entire South Australian cohort over a 10 year period, which was beyond the scope of this study.

Follow-up mammography was assessed at three different frequencies of follow-up, chosen to reflect the current annual follow-up interval and two less intense schedules that were deemed feasible options within Australia as informed by expert consultation. For the analysis of the calibrated model, 75% and 90% adherence rates with the specified surveillance strategies were tested, to provide feasible upper and lower boundaries of mammographic attendance by Australian breast cancer survivors.

Health state utility values

Utility values for each health state were identified in a Health Technology Assessment published in 2007.[13] This publication sourced its utility values primarily from the CEA Registry of the Harvard School of Public Health (https://research.tufts-

nemc.org/cear4/default.aspx)[14] and from a paper by Tengs and Wallace.[15] A quality of life decrement associated with mammography screening was identified in Tengs and Wallace,[15] based upon work by de Koning et al.[16]

Table 1: Our health state utility values

Health states	Utilities
Disease-free	0.94
Surveillance (single MMG + clinical visit per year)	-0.01
Impalpable local recurrence	0.94
Removed impalpable local recurrence – surgery costs	0.74
Removed impalpable local recurrence – ongoing costs	0.85
Palpable local recurrence – surgery costs	0.74
Palpable local recurrence – ongoing costs	0.85
Distant metastases	0.5

Table 2: HTA health state utility values

		Managed				Unmanaged	
		On treatment		After treatment			
Severity	% on chemotherapy	Low	High	Low	High	Low	High
1	0.240	0.71165	0.80654	0.8	0.95	0.8	0.95
2	0.759	0.59673	0.67629	0.8	0.95	0.8	0.95
3	0.769	0.52693	0.60376	0.63	0.68	0.8	0.95
4	0.991	0.49716	0.60807	0.63	0.68	0.8	0.95
5	1	0.39362	0.50394	0.5	0.55	0.8	0.95

The health states described in my study and those in the HTA differ quite markedly. I describe relapse within the breast as either impalpable or palpable, and do not differentiate as to whether it is isolated to the breast parenchyma or involves both breast and axillary lymph nodes (severity 3&4). If we assume severity 1 after treatment is equivalent to our disease-free state, their high value is similar to our point estimate, but their low value is substantially less than our value. The low utility value after treatment reported for distant metastases in the HTA (severity 5) is identical to our

utility value for the same health state. The other health states are not directly comparable between our study and those of the HTA.

ANALYSES

Our analyses were designed to determine the most cost-effective mammographic follow-up schedule based on the age of the patient, her risk of recurrence and adherence with mammography. We determined the probabilities of cost-effectiveness for alternative follow-up schedules, by QALY threshold and assumed adherence with follow-up schedule, for each age cohort. We also assessed the effect of a utility decrement associated with surveillance on the ICER.

In the HTA, analyses were designed to determine the incremental cost effectiveness ratio for each surveillance strategy per follow-up interval, and were performed separately for women based on the nature of the primary surgery (breast conserving surgery or mastectomy). In these base-case analyses, the outcomes of an IBTR and MCBC were assumed to be identical. An additional analysis was performed on IBTR alone, to reflect the increased risk of death from tumor recurrence in the treated breast. A number of one-way sensitivity analyses were performed; and a multi-way sensitivity analysis to assess the effect of age at the time follow-up starts; but no probabilistic sensitivity analysis. The HTA model assumed 100% adherence and no quality of life decrement associated with surveillance.

RESULTS

The two early breast cancer surveillance models have been shown to have significant differences with respect to model structure, population, and analysis. My model had the advantage of using patient-level data that links surveillance and outcomes for each individual woman. These direct estimates inform a formal calibration process that provides confidence in the realism of the model, and that the model results represent the associated uncertainty around the input parameter values. The HTA model represents the natural history in more detail, and facilitates analyses of a broader

set of follow-up options. While the HTA model uses extensive outcome data from cancer registries, the model has no direct estimates of actual follow-up received in each surgical group.

It is only in the analysis based on age that there is any substantial overlap between the two models. In our model, analysis is stratified by age into 50-69 years, and 70-79 years, with four different levels of risk in each age group. In the HTA model, the effect of age is addressed by an analysis comparing cost-effectiveness of surveillance based on a hypothetical (assumed high risk) 40 year old woman and a hypothetical (assumed low risk) 70 year old woman.

Our results demonstrate that for postmenopausal women (50-79 years), 2 yearly mammography is the most cost-effective strategy for women of any age with excellent or good prognosis tumors, and 70-79 year old women with moderate prognosis tumors. A mixed schedule is most cost-effective for 50-69 year old women with moderate prognosis tumors, and 70-79 year old women with poor prognosis tumors. Annual mammography is only cost-effective for 50-69 year old women with poor prognosis tumors. In the HTA model, annual mammography had the highest net benefit for the 40 year old woman, whereas mammography alone every 2 or 3 years was determined to be more cost-effective for the 70 year old woman. The authors conclude that “a more intensive follow-up of women judged to be at high risk may be more effective, conversely for women at lower risk it may be more cost-effective for surveillance to be performed less often….” If we assume the young high risk woman in the HTA model is roughly equivalent to our 50-69 year old woman with a poor prognosis tumor; and the low risk woman in the HTA model is roughly equivalent to our 70-79 year old woman with an excellent or good prognosis tumor; then the conclusions between the two models are broadly similar.

The two models can potentially inform different decisions. From a policy perspective, if a particular frequency of follow-up was to be funded, the HTA model could inform which regimen all women should receive based on the extent of their primary surgery.

Our model could be used to personalise the mammographic strategy for a woman based on her age and risk of relapse within the breast.

CONCLUSION

Although different modelling approaches were used, the results reported for the two studies are broadly similar. This increases confidence regarding the validity of our conclusions that a one-size-fits-all, annual surveillance strategy may not be the optimal approach to monitoring breast cancer survivors.

Chapter 5

ISSUES WITH DATA ACCESS AND QUALITY

INTRODUCTION

In the preceding two chapters I have attempted to answer research question 2: "Using a model based economic analysis, is it efficient to tailor mammographic follow-up according to risk of recurrence?" I have described the modelling methods, which involved calibration of times between events in the model to observed surveillance pathways for individual women within each prognostic group and outcomes for each group (i.e. proportions recurring and dying in each year post-primary); followed by analysis of alternative surveillance strategies using the calibrated time to event parameters. My results suggest that annual mammographic follow-up is not cost-effective for most postmenopausal women, and that mammographic follow-up can be tailored according to risk of recurrence based on the NPI score of the primary breast cancer and age at diagnosis.

This chapter describes the large gulf between the ideal dataset to inform the cost-effectiveness model and what is currently available in South Australia (and beyond). The process and problems encountered in collecting data on the study cohort are described in terms of sourcing, extracting and linking data. The final study cohort is described, and the implications of the data quality on the interpretation of the model analysis is reported.

The intention of the chapter is to highlight the difficulty in obtaining data for health services research within South Australia, even for common conditions like breast cancer. This chapter is deliberately process driven. Identification of ideal data, data gaps and data roadblocks is a necessary first step in guiding and generating change within the existing data infrastructure, to help guide clinical care and to improve the quality of health services research in the future.

IDEAL DATA

To analyse the calibration model, we firstly need to define the relevant prognostic subgroups of women with early breast cancer, based on validated clinical and pathology criteria that are routinely collected in South Australia.

For each identified woman within each prognostic group, we then need to determine the follow-up she received and her disease course. This includes the frequency of her clinical and mammographic follow-up; whether or not she developed a recurrence (and if so, where, when and how it was detected); and whether or not she died from breast cancer.

To inform the calibration model, an ideal data set would include a large cohort of women with early breast cancer who were disease free following their primary treatment, and who were followed over several years. Routinely collected data on this cohort of women would include:

Primary breast cancer (PBC)

- Patient: age, menopausal status
- Tumor: size, nodes, grade, receptor status, ± other prognostic indicators
- Treatment: type of surgery, radiotherapy, adjuvant therapy received

Follow-up clinical examination & mammography

- Frequency for each year post diagnosis
- Results of diagnostic tests

Recurrences

- Method of detection - impalpable or palpable
- Timing - interval between detection of PBC and recurrence
- Location - treated breast, opposite breast or distant relapse

Mortality

- Timing: interval from detection of PBC and recurrence
- Cause: breast cancer, other primary cancer, other cause

Cost

- Of follow-up
- Of treating recurrence
- Quality of life/utility data

SOURCING THE DATA

INTERNATIONAL TRIAL DATA

We approached the Australian and New Zealand Breast Cancer Trials Group (ANZBCTG) for access to clinical trial data. We met with the Chief Investigator of ANZBCTG who identified the clinical trial most likely to fulfil our needs as BIG 01-98. The Chief Investigator encouraged us to apply, and championed our request to the international committee. After considering our request for 6 months, the international committee eventually declined access on the basis of the small number of recurrences that had been detected in the trial.

ROYAL AUSTRALASIAN COLLEGE OF SURGEONS DATABASE

The National Breast Cancer Audit, managed by the Royal Australasian College of Surgeons, collects data on surgical practice with the aim of improving quality of surgical care to women with early breast cancer. Whilst there are more than 100,000 episodes of breast cancer care in the database, voluntarily contributed by over 250 surgeons, discussion with local surgeons participating in the database revealed that follow-up information is poorly recorded. In order to collect robust data on follow-up, we would need to identify and then contact each of the women in the database requesting further information. Not only would this breach confidentiality, the completeness of follow-up information obtained from patient recollection/provision of reports may be insufficiently robust.

SOUTH AUSTRALIAN DATA

I then investigated routinely collected data within South Australia, and concluded that this would be challenging for four main reasons:

1. No single data source contains the information that we require.

We would need to extract and link data across four different data custodians:

- South Australian Cancer Registry (SACR) – includes all women diagnosed with breast cancer, and contains their demographic information and data

describing size, nodal status and grade of the index cancer. In 2009 it did not routinely collect recurrence data or receptor status. As of June 2012, the SACR commenced collecting hormonal receptor status (ER, PR) but not HER2 status.

- Open Architecture Clinical Information System (OACIS) database - contains data describing all hospital presentations (both inpatient and outpatient) and diagnostic investigations (including mammography and pathology) for public metropolitan hospitals within South Australia.
- Integrated South Australian Activity Collection (ISAAC) database - contains data describing inpatient separations using ICD 10-AM diagnostic and procedure codes, for all public and private hospitals within South Australia.
- Registry of Births, Deaths and Marriages (BDM) – this registry is part of the Department of the Attorney General, and owns the mortality data that is recorded within the SA Cancer Registry.

2. Follow-up data and recurrences are not routinely collected.

Follow-up and recurrence data would need to be manually extracted from pathology and mammography reports identified within the OACIS database. All pathology reports relating to the index surgical procedure, and all subsequent mammography reports would be reviewed, and salient data points entered into a bespoke database.

3. Electronically available registry data was too limited for our purposes.

- South Australian Cancer Registry (SACR): due to mandatory notification of new cancer diagnoses, SACR receives all pathology reports from the public and private sector. From each pathology report, only a few key data points are entered into an electronic database before the paper pathology reports are filed. In 2009, the SA Cancer Registry could provide age, tumour, nodal status and grade, but no information on receptor status or recurrences. The electronic data point “multifocal” was used to encompass both synchronous

tumours and metachronous tumours in the breast, and also distant recurrence. As such it could not be used to reliably differentiate between more than one tumor in the same breast, concurrent tumors in both breasts, a recurrent tumor in the treated breast , a contralateral new primary breast cancer or distant relapse within lung, liver, bone or brain. As such we decided not to use the variable "multifocal" as a means of identifying a second breast cancer in either breast.

- The Royal Adelaide Hospital Cancer Registry: could provide detailed information on patient, tumour and treatment characteristics of the primary tumour but had sparse data on follow-up. For this hospital based registry, recurrences were only entered into the database if the registry was notified by ad hoc provision of a pathology report identifying the recurrence, or more commonly when the patient died. The RAH Cancer Registry, due to resource constraints, was 2-3 years behind in updating their database with recurrence data.

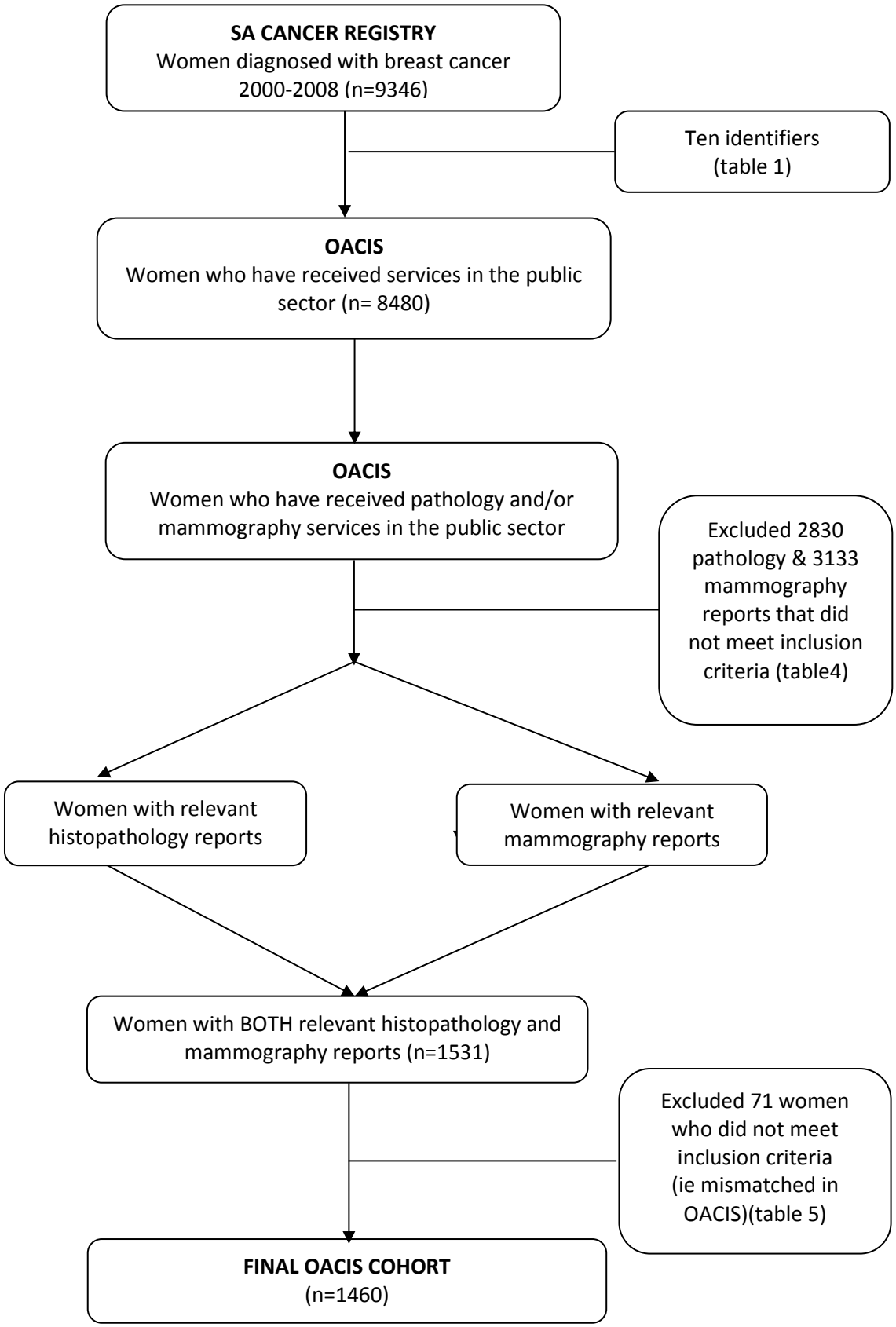
4. Women with breast cancer may be treated in public, private or both.

Women can choose to be treated for breast cancer and followed up entirely in the public sector, entirely in the private sector, or cross between both sectors. Due to privacy laws, we would not be able to access pathology or mammography reports for women who received these services in the private sector. We could access data from the OACIS system for women who had their primary breast cancer treated in the public health system, but the problem would be identifying within the ISAAC database, which of these women subsequently developed a recurrence and sought treatment for the recurrence in the private sector. Women who had their primary breast cancer treated in private and their recurrence treated in public would be excluded as we would not be able to access any pathology reports related to their primary tumour. Women who received their care entirely in the private sector would be excluded from the study as we would not be able to access pathology or mammography reports.

EXTRACTING THE SOUTH AUSTRALIAN DATA

The process of data collection described in the following section is represented in Figure 1.

Figure 1: Data collection to define the final OACIS cohort



Ethics approval was provided by SA Health Human Research Ethics Committee for the data extraction and linkage process described below. Only de-identified data was provided for analysis.

SOUTH AUSTRALIAN CANCER REGISTRY (SACR)

9346 women were diagnosed with breast cancer between 1/1/2000 and 31/12/2008. The SACR provided 10 identifying variables per woman (Table 1), and OACIS were able to match 8480 females to unique patient identifiers (PIDs) in the OACIS database, but were unable to match 866 women.

Table 1: Identifying variables provided by SACR for matching in OACIS

Variable	Explanation
Surname	Surname of patient
Given name 1	First given name of patient
Given name 2	Second given name of patient
Sex	Sex of patient
DOB	Date of birth
Postcode	Postcode of patient
DOD	Date of death – if applicable, blank if patient is alive
Age at diagnosis	Age of patient at first diagnosis
Site	Site of invasive cancer 1749 = female breast, 1759=male breast
In situ	In situ cancers only, value of "1" in this field
TOTAL	9417 cases (9346 female, 71 male)

OPEN ARCHITECTURE CLINICAL INFORMATION SYSTEM (OACIS)

No unique patient identifier

Any person presenting to a public metropolitan hospital will be given their own unit record number (URN) or medical record number (MRN). This URN/MRN will be specific for each hospital. If the person attends multiple different public hospitals, they will have multiple different URN/MRNs. To identify all presentations across the metropolitan area for a given woman, each URN/MRN had to be identified, and a unique Patient Identifier Number created by OACIS.

SA Health use a program called EMPI that compares patient demographic data when a new URN is created at a hospital to all the active patients in SA Health's patients table and performs probabilistic matching. The program looks at: first name, middle name surname, DOB, address and sex. Each item is given a weight and for every item matched a score is given and if the total score is above 22.9 it gets automatically matched by the program, if the score is between 18.1 and 22.9, the details get sent to a central team (consisting of representatives from Medicare and the hospitals) that makes the decision whether or not to manually match the patient, based on a set of rules and procedures. While this matching process has matured over the years, it is not perfect and some people are still matched when they shouldn't have been. In theory a new patient would only complete a registration form on the first encounter at a new hospital, but in reality they may be asked more than once. For example, if a registration clerk cannot identify the patient within the hospital administration system due to errors in completing the registration form during a previous encounter (name or birth date entered incorrectly, or due to language difficulties), then it is possible that a patient could have more than one URN, and therefore more than one Study ID. Matching accuracy is also complicated by different hospitals having different data quality standards, and using different patient administration systems.

No unique test identifier

We needed to build a query to identify the breast surgical histopathology and follow-up mammography tests received by patients with breast cancer, identified from the SACR, that were provided in public metropolitan hospitals within SA. We needed test IDs for searching within OACIS as text is not indexed or sorted in the database, however there are no unique test IDs for breast surgical histopathology or follow-up mammography.

To overcome this problem, we needed to find text descriptors that could potentially identify these services and map the test IDs ascribed to each service. We started with a small number of women with known breast cancer who had been matched in OACIS, searched within OACIS using an exhaustive list of text descriptors and established a list of multiple test IDs for histopathology and mammography services at the largest

metropolitan teaching hospital in South Australia. We then repeated the process for the other metropolitan teaching hospitals to capture any other tests IDs that might exist for these same services.

We were unable to drill down to the level of breast surgical histopathology. We had to identify "histopathology" services and then sift through all histopathology reports for each woman. "Histopathology" reports could arise from any organ system in the body, and did not differentiate between core biopsies and surgical specimens.

There is no unique test code that identifies a follow-up mammogram. Searching by the term "Mammography" or "Mammogram" will yield mammography reports, but does not differentiate between mammography performed pre-operatively (diagnostic), for a family history of breast cancer (screening), during neo-adjuvant chemotherapy (monitoring) or post-operative follow-up. As such, we had to sift through all mammography (MMG) reports for each woman with breast cancer who had been matched within OACIS, read the clinical indications for mammography from the report, and compare the date of mammography with date of surgery (if known) to determine if the mammogram post-dated surgery for the index breast cancer.

Missing data

During data extraction, it was discovered that due to resource issues within OACIS, there were 409 histopathology reports for the 2003/2004 financial year that were missing from the data warehouse. No resources were planned by SA Health to remedy this deficit during the timeframe of my PhD.

Data extraction

The list of 8480 PIDs and the relevant test IDs were used to extract histopathology and mammography results for each patient from the OACIS database. 5362 women had received histopathology or mammography services in the public sector. The reports for these services were de-identified by the OACIS team, by ascribing a Study ID to each patient (created by combining the patient's date of birth and their PID) and removing

the patient's name in the first line of the body of the report and replacing it with the text "Line removed due to privacy". All reports were then provided on a CD for analysis.

All de-identified histopathology and mammography reports for these 5362 women were reviewed by another medically qualified doctor employed specifically for this purpose. The doctor worked with a data entry assistant who entered the dictated findings immediately into a bespoke database.

The database was designed and built using Auditmaker (http://auditmaker.net/) a generic online tool for clinical audit and research. Individual users can create their own online databases for collection of research data, in a secure password protected environment, and the study data can either be analysed using the online tools, or by exporting the data to a spreadsheet format. The data fields for breast surgical histopathology included all data items that were identified within pathology reports from both public and private providers. We used the date that the surgical specimen was received at the pathology department to represent the date of primary surgery, and we added the data field recurrence if we had histopathological confirmation of a recurrence within either breast that post-dated the index surgical procedure. The data fields for mammography included date of test and laterality of the examination, and we simplified the report findings to no significant abnormality (accepting expected postoperative appearances of the breast as normal) or suspicious if further imaging or biopsy was recommended. Data points extracted from the reports for each Study ID are shown in Table 2.

Table 2: Data points extracted from OACIS reports

Pathology of primary breast cancer	
Date of birth	DD/MM/YYYY
Date of primary surgery	DD/MM/YYYY
Surgery	Side, breast conserving or mastectomy
Diagnosis	Invasive or in-situ
Tumor	Number, histological subtype, total size, size of invasive component, size of in-situ component, grade
DCIS	Present/absent, proportion within and adjacent tumor, EIC status, margins of clearance, nuclear grade, patterns, presence or absence of comedo necrosis, calcification
Lymph nodes	Number of positive lymph nodes, total lymph nodes resected
Receptor status	Estrogen, progesterone and HER2 receptors
Recurrence before July1, 2011	Yes/no, date
Mammography	
Date of test	DD/MM/YYYY
Side	Left, right or bilateral
Result	No significant abnormality or suspicious

Defining the OACIS study cohort

Following data entry, the results were exported into Excel. All pathology and mammography entries for the 5362 women were reviewed to determine if they met the inclusion criteria for the study (Table 3). The number of women excluded from the study, and reasons for exclusion are shown in Table 4.

Table 3: OACIS Inclusion Criteria

Pathology
Index primary breast cancer diagnosed between 2000-2008
Surgery occurred in public metropolitan hospital with a report on OACIS
Surgical histopathology report is complete, legible and includes date of surgery
Primary breast cancer is early stage
Primary breast cancer is of ductal or lobular subtype (or a mixture of the two)
Mammography
Mammography post-dates the index surgical admission
Mammography occurred in public metropolitan hospital with a report on OACIS
Mammography report is complete, legible and includes date of test

Table 4: OACIS Exclusion Criteria

Pathology	
Duplicate patient	1
Primary breast cancer prior to 2000	3
No pathology on OACIS	230
No surgery	1
Missing date of surgery	2
Pathology on OACIS but not related to breast	1088
Breast pathology but not malignant	32
Breast pathology but not surgical histopathology[a]	739
Breast pathology but not related to primary breast cancer[b]	157
Breast pathology but not early stage invasive[c]	142
Breast pathology but not ductal or lobular[d]	10
Histopathology report missing	409
Histopathology report incomplete, illegible or missing date of surgery	16
TOTAL pathology excludes	2830
Mammography	
No mammography on OACIS	2884
No record of surgery in clinical history or report	11
Mammography but pre-operative only	195
Mammography but for family history	11
Mammography but for neo-adjuvant chemotherapy	23
No record of surgery in clinical history or report	11
No date on MMG report	1
Some MMG reports in patient series missing	8
TOTAL mammography excludes	3133

[a] FNA, core, skin/punch biopsies only
[b] recurrence of DCIS, local invasive recurrence, metastatic recurrence, secondary cosmetic breast surgery
[c] DCIS, LABC
[d] Occult, phyllodes, sarcoma, or undifferentiated carcinoma

Relevant breast surgical histopathology reports were identified for 2532 women and relevant mammography reports were identified for 2229 women. Women with breast surgical histopathology reports pertaining to the index primary breast cancer AND follow-up mammography reports were identified within OACIS for 1531 women. We then matched these women back with SACR, to confirm that our index cases fell within the study period, and included women with invasive cancer resident within SA. We identified 71 additional exclusions (table 5). The final study cohort was 1460 women.

Table 5: SACR Exclusion Criteria

Study exclusion criteria identified in SACR	
Index case occurred prior to 2000	6
Index case occurred after 2008	47
Male patient	1
No record in SACR (ie couldn't be matched back from OACIS)	6
In-situ disease	2
Treated here but not SA resident	9
TOTAL	71

Risk stratification

Tailoring mammographic surveillance strategies according to risk of recurrence requires risk stratification, based on proven prognostic factors that were routinely collected within South Australia during the study period.

Menopausal status

Clinical trial participants are often classified according to menopausal status,[101] "younger" women are premenopausal and/or aged < 50 years, and "older" women are postmenopausal and/or aged 50 years or older.[102] I therefore chose to define the women in our study cohort as <50 years of age (premenopausal) or ≥50 years of age (postmenopausal). This decision does not allow separation of women who may be <50

years of age but already postmenopausal. However accurate ascertainment of menopausal status by casenote review was beyond the scope of this study.

Nottingham Prognostic Index (NPI)

The SACR routinely collect tumor, node and grade information on all women with breast cancer, which we could convert to prognostic subgroups using the Nottingham Prognostic Index.[103] The Nottingham Prognostic Index is calculated as follows:

NPI= Tumor size(cm) x 0.2 + Lymph node stage (1,2,3) + Histologic grade (1,2,3).

Lymph-node stage is defined in terms of number of positive lymph nodes resected, with stage 1 = 0 ("lymph node negative"), stage 2 = 1-3 and stage 3 = > 3 positive lymph-nodes resected. Prognostic subgroups are identified in terms of the overall NPI score where excellent prognosis is a NPI score ≤ 2.4, good prognosis is ≤ 3.4, moderate prognosis is >3.4 and ≤ 5.4 and poor prognosis is a NPI score >5.4.

Receptor status

It is possible to distinguish prognostic breast cancer subgroups on the basis of combined hormone and human epidermal growth factor receptor 2 status.[104] While estrogen (ER), progesterone (PR) and HER2 receptor status were not routinely collected within the SACR prior to June 2012, receptor status was increasingly reported by pathology providers in the public sector during the study period. The intention was to categorise into 8 subgroups according to whether the resected breast tissue was positive (+) or negative (-) for the estrogen, progesterone and HER2 receptors respectively: ER+PR+HER2+, ER+PR+HER2-, ER+PR-HER2+ , ER+PR-HER2-, ER-PR+HER2+, ER-PR+HER2-, ER-PR-HER2+, ER-PR-HER2-. During data extraction it was appreciated that HER2 status was not consistently reported, and when reported could be one of two different techniques (immunohistochemistry or gene amplification), the results of which were not directly comparable. Therefore we decided not to use HER2 status in our subsequent analysis, but initially intended to retain the ERPR subgroupings. However, as the final study cohort was small, further stratification by hormonal receptor status resulted in very small risk subgroups (table 6). As such we reluctantly decided to limit our modelling analysis within premenopausal and postmenopausal groups to NPI category only.

Table 6: ERPR status of the primary breast cancer of women in the OACIS study cohort

	Premenopausal				Postmenopausal			
		Recurrences				Recurrences		
NPI Status	Women	Yes	Uncertain	No	Women	Yes	Uncertain	No
Excellent								
ER+PR+	19	0	1	18	129	3	5	121
ER+PR-	0	0	0	0	14	0	1	13
ER-PR+	1	0	0	1	2	0	0	2
ER-PR-	0	0	0	0	3	0	0	3
Missing data	4	0	0	4	30	0	0	30
Total	24	0	1	23	178	3	6	169
Good								
ER+PR+	41	0	4	37	193	8	7	178
ER+PR-	1	0	0	1	13	0	1	12
ER-PR+	1	0	0	1	5	0	0	5
ER-PR-	3	0	0	3	4	1	0	3
Missing data	6	0	0	6	41	0	4	37
Total	52	0	4	48	256	9	12	235
Moderate								
ER+PR+	96	5	2	89	243	10	7	226
ER+PR-	7	0	0	7	28	1	0	27
ER-PR+	3	1	0	2	10	0	0	10
ER-PR-	23	2	1	20	71	7	5	59
Missing data	21	1	2	18	68	1	4	63
Total	150	9	5	136	420	19	16	385

Table 6: ERPR status of the primary breast cancer of women in the OACIS study cohort (cont.)

	Premenopausal				Postmenopausal			
		Recurrences				Recurrences		
NPI Status	Women	Yes	Uncertain	No	Women	Yes	Uncertain	No
Poor								
ER+PR+	43	2	2	39	103	8	4	91
ER+PR-	7	0	0	7	12	1	1	10
ER-PR+	2	0	0	2	4	0	0	4
ER-PR-	18	2	0	16	25	1	1	23
Missing data	7	2	0	5	13	2	0	11
Total	77	6	2	69	157	12	6	139
Cant calculate								
ER+PR+	20	2	2	16	66	3	3	60
ER+PR-	0	0	0	0	9	1	0	8
ER-PR+	3	0	1	2	0	0	0	0
ER-PR-	3	0	1	2	12	1	0	11
Missing data	5	2	0	3	28	2	1	25
Total	31	4	4	23	115	7	4	104
TOTAL	**334**	**19**	**16**	**299**	**1126**	**50**	**44**	**1032**

NPI = Nottingham Prognostic Index
Recurrence = relapse within the treated breast or axilla, or a new primary or metastasis in the opposite breast
ER+ = estrogen receptor positive, ER-= estrogen receptor negative
PR+ = progesterone receptor positive, PR-= progesterone receptor negative

INTEGRATED SOUTH AUSTRALIAN ACTIVITY COLLECTION (ISAAC)

Due to privacy laws, we were unable to access pathology or mammography reports from the private sector. In order to determine if women in the final OACIS study cohort had developed a recurrence which had been treated in the private sector, we needed to identify a proxy measure of recurrence that is routinely collected within the larger ISAAC database, which could then be linked back to OACIS. This process is described in the section below.

BIRTHS, DEATHS AND MARRIAGES

Births, Deaths and Marriages (BDM) are the only database in South Australia that can provide access to cause of death data in addition to date of death.

OACIS receives regular monthly mortality updates from Births, Deaths and Marriages, but are supplied only with date of death and not cause of death.

ISAAC records date of death only if a patient dies while in hospital.

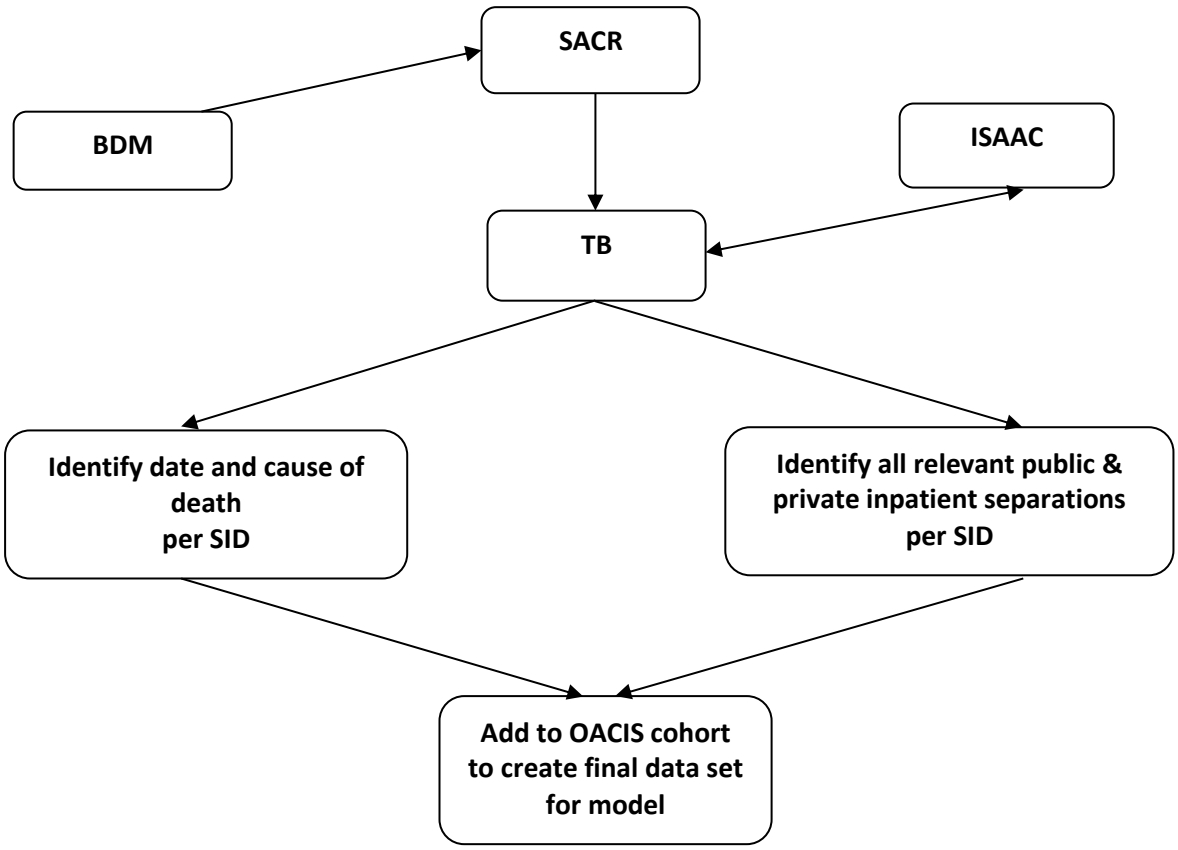
SACR receives regular monthly mortality updates from Births, Deaths and Marriages, and is supplied with date of death and cause of death, but is not permitted to release this information as the data is “owned” by Births, Deaths and Marriages (BDM). It is necessary to apply to and receive approval from the Registrar of BDM to access mortality data contained within the SACR.

We recognise that the mortality data obtained from the BDM does not take into account women who may have moved interstate. However to obtain complete mortality data from the National Death Index costs a minimum of $5000 for data extraction which was beyond our budget.

LINKING DATA AND QUALITY CHECKS

The process of data linkage and quality checks described in the following section is represented in Figure 2.

Figure 2: Data linkage process



SACR South Australian Cancer Registry
OACIS Open Architecture Clinical Information System
ISAAC Integrated South Australian Activity Collection
BDM Register of Births, Deaths and Marriages
SID Study Identification (unique per patient)
TB Taryn Bessen

1. BDM to SACR

Permission was obtained from the Registrar of Births, Deaths and Marriages for access to the mortality data of all women with breast cancer between 2000-2008 kept within the SACR. Mortality data was provided for the entire SA cohort as of 16th August 2010, to correspond with the date that the OACIS data was extracted for the study. For those who had died, date of death and cause of death was available. Within the SACR, cause of death for women with a diagnosis of breast cancer is coded as breast cancer death, death from another primary cancer or other cause death.

2. OACIS vs SACR

The final OACIS study cohort of women with early breast cancer who had breast surgical histopathology and follow-up mammography available (n=1460 of 9362 women), is highly likely to be a biased sample due to methodological issues relating to identifying and extracting the data as described above.

To assess the extent to which our sample was/was not representative of all women diagnosed with breast cancer in SA between 2000-2008, we ascertained if there were differences in age, tumour characteristics, date of death (if applicable) and cause of death between those women in the study cohort and the total cohort of women diagnosed with breast cancer in SA during the study period.

We obtained age at diagnosis and electronically recorded pathology data (tumor size, number of positive lymph nodes resected, grade) of all 9346 women diagnosed with breast cancer from 2000-2008 from the SACR, and converted this data into a NPI category for each woman (table 7).

Table 7: Comparison of prognostic subgroups between SACR and final OACIS study cohort

	Premenopausal						Postmenopausal					
	Women		BCD		OCD		Women		BCD		OCD	
NPI Status	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
South Australian Cancer Registry (n=9346)												
Excellent	127	1.36	0	0	1	0.01	885	9.47	12	0.13	40	0.43
Good	282	3.02	4	0.04	3	0.03	1252	13.40	25	0.27	60	0.64
Moderate	598	6.40	53	0.57	3	0.03	1693	18.11	158	1.69	107	1.14
Poor	272	2.91	63	0.67	3	0.03	549	5.87	156	1.67	29	0.31
Cant calculate	742	7.94	113	1.21	8	0.09	2946	31.52	622	6.66	390	4.17
TOTAL	2021	21.62	233	2.49	18	0.19	7325	78.38	973	10.41	626	6.70
OACIS study cohort (n=1460)												
NPI Status	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**	**n**	**%**
Excellent	24	1.64	1	0.07	0	0	178	12.19	0	0	12	0.82
Good	52	3.56	1	0.07	0	0	256	17.53	7	0.48	11	0.75
Moderate	150	10.27	4	0.27	2	0.14	420	28.77	31	2.12	26	1.80
Poor	77	5.27	10	0.68	1	0.07	157	10.75	34	2.33	11	0.75
Cant calculate	31	2.12	3	0.21	1	0.07	115	7.88	11	0.75	16	1.10
TOTAL	334	22.88	19	1.30	4	0.27	1126	77.12	83	5.68	76	5.20

3. OACIS to ISAAC

Linkage

The objective of linking the OACIS to ISAAC data was to identify public patients in our study cohort who had treatment for recurrences in the private sector.

This process involved identifying potential variables common to both databases upon which to link; the quantity and quality of the data this would yield and whether this was sufficient for our purposes; how up to date the available data was; whether the data extraction from each custodian and the subsequent linkage could be performed within a feasible PhD timeframe; and relevant ethical issues pertaining to linkage.

Discussions on how to best do this began in December 2010. This involved several meetings with the data custodians, and a meeting with data coders. The proposed methodology went through several successive iterations including those below, before we finally were left with option 6:

1. South Australia & Northern Territory Data Linkage Unit (SANTDLU) - we considered using this newly formed entity to perform the linkage. Whilst this would provide better linkage within ISAAC for the public patients, they would not be able to identify private patients. Also, due to resourcing issues, the SANTDLU would not be able to commence the data extraction/linkage until early 2012.
2. Medicare numbers - OACIS would provide medicare numbers for all of the women identified in its database to ISAAC for matching. However a small pilot test within OACIS demonstrated that medicare numbers were recorded for < 50% of all patients in OACIS.
3. Sex, date of birth, postcode - OACIS would provide a list of all of the women identified in their database, and ISAAC would match on sex, date of birth and postcode. ISAAC would then determine if they had Medicare numbers for the matched women, within their database. If they did, they would identify both their public and private inpatient separations. This method did not progress due to resourcing issues within ISAAC.
4. OACIS Linking Variable - while listed under OACIS specifications, ISAAC stated that this variable is not in their inpatient data collection.

5. Hospital, URN, separation date and time - OACIS performs linkage with ISAAC inpatient data using hospital, URN, separation date and time, but again this would only be for public inpatient separations.
6. Perform the linkage ourselves - Prof Karnon already had ISAAC data with encrypted medicare numbers for the financial years of 2001/2002 through to 2009/2010. We requested additional data from January 2000 to encompass the entire study period, and also the 2010/2011 financial year to provide additional follow-up.

From initial discussions in December 2010, it took 11 months to receive the additional ISAAC data. The data extracts received from ISAAC were only available from July 2000 so we have no information on patients diagnosed with breast cancer from January to June 2000 (the first 6 months of the study period).

Linking new ISAAC data with pre-existing ISAAC data in Prof Karnon's collection involved multiple steps as follows:

- Prepare the additional data (pre and post 2001-2008) into the same format as the 2001-2008 data. Then append the additional data to the 2001-2008 dataset.
- Apply the linkage algorithm developed for the 2001-2008 data to the expanded dataset to link hospital separations across patients over the period 2000-2011. The linkage was based on the following identifiers: nine digit Medicare number, date of birth, and gender. It was assumed that separations with the same nine digit Medicare number (change in the 10^{th} character occur when Medicare card is renewed), same month and year of birth (to account for data entry errors around day of birth) and same gender involved the same person
- Check and remove duplicate separations.
- Identify eligible breast cancer patients within ISAAC using hospital and hospital-specific URN as recorded in OACIS.
- Restrict ISAAC search to relevant inpatient separation types, using the ICD-10-AM V6 diagnostic code for breast cancer (C50) and multiple procedure codes relating to breast surgery, chemotherapy and radiotherapy as shown in (table 8),as advised by medical coders in SA Health. This would provide a list of all breast cancer related inpatient separations per patient, at both public and private hospitals.

Table 8: Breast cancer related ICD 10-AM V6 codes

Variable	Codes	Descriptions
Principal Diagnosis	C50	Breast cancer
Procedure Codes Surgery	31500-00; 31500-01; 31515-00 31524-00; 31524-01 31518-00; 31518-01	Excision of lesion of breast; open biopsy of breast; re-excision of lesion of breast Subcutaneous mastectomy - unilateral,bilateral Simple mastectomy - unilateral, bilateral
Procedure Codes Radiotherapy	15000-00; 15003-00 15100-00; 15103-00 15224-00; 15239-00 15254-00; 15269-00 15600-00; 15600-01; 90766-00	Radiation treatment, superficial - 1 field; ≥2 fields Radiation treatment, orthovoltage - 1 field; ≥2 fields Radiation treatment, megavoltage, single modality linear accelerator – 1 field; ≥ 2fields Radiation treatment, megavoltage, dual modality linear accelerator - 1 field; ≥ 2fields Stereotactic radiation treatment - single dose; fractionated Brachytherapy using surface applicators, other sites
Procedure Codes Chemotherapy	1920-00 1920-03 1920-09 96196-XX to 96199-XX 96209- XX	Administration of pharmacotherapy – antineoplastic agent Adminstration of pharmacotherapy – steroid Administration of pharmacotherapy – other and unspecified pharmacological agents Intra-arterial; intramuscular; intrathecal; intravenous Loading of drug delivery device

Linkage analysis

Using the OACIS generated Study IDs, we were able to identify records in ISAAC for 1114 women. We were unable to identify any ISAAC record for 346 women (23.7%)

The 1114 linked women had a total of 5627 inpatient separations. The number of inpatient separations per woman varied as follows: 440 (39.5%) had one inpatient separation, 515 (46.2%) had 2-10, 128 (11.5%) had 11-20 and 31 (2.8%) had 21 or more admissions for breast cancer related procedures.

For each woman, every procedure (up to a maximum of 25) for each of her admissions was reviewed. The patient's date of birth as recorded in ISAAC, total number of inpatient separations, time periods in which the admissions occurred, and the clinical purpose for each inpatient separation was documented.

All inpatient separations per woman were then reviewed, and given the absence of pathology reports in ISAAC, each woman's disease course was interpreted and summarised based on my clinical experience (described below):

- Primary breast cancer treatment was considered generally to be surgery (excision ± re-excision/completion mastectomy OR primary mastectomy, AND axillary sentinel lymph node biopsy ± regional ± radical lymph node resection), radiotherapy and chemotherapy; and to take 6-9 months to complete (longer duration if 2 stage index surgical procedure OR wide local excision followed by mastectomy, as this delayed commencement of radiotherapy or chemotherapy by approximately 6 weeks for post-operative recovery).
- Subsequent surgeries were interpreted as either for malignant disease, for benign disease or prophylactic, based on the principal diagnosis created for a given inpatient separation. Subsequent breast surgeries for malignancy could be for either a local recurrence in the treated breast or a new primary breast cancer in the opposite breast, but laterality of the index breast cancer and relapse within the breast, cannot be determined within ISAAC.

- Second rounds of chemotherapy were interpreted as those initiated beyond the time course of the primary breast cancer treatment. Second rounds of chemotherapy, without second surgeries, were deemed to be for distant metastases. For women who only had surgery for their index cancer (e.g for a very small tumor), subsequent events of surgery with first round of chemotherapy were deemed local recurrence/new primary breast cancer. As of 1/7/2007, chemotherapy was no longer administered as an inpatient procedure in the public sector, but became an outpatient procedure. The implications for data capture and interpretation are discussed later in the chapter.
- The status was determined as uncertain if there was insufficient information in ISAAC to determine the woman's disease course (single presentation coded as either administration of packed cells, or chemotherapy only), or for whom there was no record in ISAAC.
- Review of ISAAC separations identified a small group of women who had received neo-adjuvant chemotherapy prior to their index surgery (this is usually given for locally advanced breast cancer, which is an exclusion criteria for our study), which we had not detected during our review of OACIS reports.

Within OACIS, presence/absence of documented recurrence from the pathology reports were classified as yes/no respectively. If pathology reports were missing, incomplete or illegible, the status was determined as uncertain.

A patient level dataset was created in excel, which combined the results of the OACIS, SACR and ISAAC findings. For every OACIS generated study ID; the date of birth, date of surgery, age at surgery, menopausal status, calculated NPI category, and presence/absence of documented recurrence from the pathology reports were listed. Date and cause of death obtained from the SACR, and summarised disease course from ISAAC, were added to the dataset for each woman.

The final determination of whether or not a recurrence had occurred, was a combination of separate assessments of results obtained from both the OACIS and ISAAC databases as shown in table 9.

Table 9: Decision process for recurrence status

OACIS	ISAAC	Number	Decision
No	Exclude – neoadjuvant chemotherapy	13	Exclude
No	Exclude – wrong woman matched	3	Exclude
No	No	925	No
No	Uncertain - admin packed cells	1	No
No	Uncertain - chemo only	10	No
No	Uncertain - no record in ISAAC	325	No
No	Yes	54	Yes
Uncertain	Exclude – neoadjuvant chemotherapy	2	Exclude
Uncertain	No	38	No
Uncertain	Uncertain - chemo only	1	Exclude
Uncertain	Uncertain - no record in ISAAC	13	Exclude
Uncertain	Yes	6	Yes
Yes	No	9	Yes
Yes	Uncertain - chemo only	1	Yes
Yes	Uncertain - no record in ISAAC	8	Yes
Yes	Yes	51	Yes
TOTAL		1460	

OACIS Open Architecture Clinical Information System
ISAAC Integrated South Australian Activity Collection

The OACIS cohort was originally 1460 women, but following review of ISAAC data, another 32 women (table 9) were excluded based on receipt of neoadjuvant chemotherapy (n=15), wrong woman being matched (n=3), and recurrence assessment deemed uncertain in both OACIS and ISAAC (n=14).This brought the final study cohort down to 1428, with a total of 129 recurrences. We then proceeded to determine the type and date of recurrence. Whilst we had collected data from OACIS as to whether a relapse within the breast was in the treated or opposite breast, we decided to group them together as a breast relapse to keep it consistent with our model structure. Of the 129 recurrences, 80 were breast recurrences and 49 were distant recurrences. Date of relapse was derived from date of subsequent surgery for breast recurrences, and for distant recurrences, date of relapse was derived from date of commencement of second round of chemotherapy OR date of inpatient radiotherapy. If the date of relapse was derived from subsequent surgery (the only item that was identified in both OACIS and

ISAAC), and this date differed between the two databases, the earlier documented recurrence was chosen as date of recurrence.

For all 1428 women, we then added their surveillance mammogram dates and reports to the patient level dataset.

Within our study cohort of 1460 women with an OACIS histopathology report confirming surgical removal of an invasive breast cancer, there were 346 women for whom there was no record in ISAAC. This suggests that either there have been considerable coding errors in the examined data, our linking process was not sufficiently robust, or a combination of the two. If recurrence status in OACIS was deemed "no" or "uncertain" these women were excluded. If recurrence status in OACIS was deemed "yes" these women were included (table 9).

There are also a number of assumptions we made in reviewing the ISAAC data, primarily relating to coding of the principal diagnosis and procedure codes as follows:

Surgery

- If in the one inpatient separation, procedure code 1 is for an excision of breast and procedure code 2 is also for an excision of breast, is this for two different lesions (and if so, are they in the same or opposite breast), or is this a coding error (either only one excision, or an excision followed by re-excision in the same operation)? We assumed that a single lesion had been completely resected in a single operation.
- If for two inpatient separations within the primary treatment period, procedure code 1 is for an excision of breast, is this for a two different lesions (and if so, are they in the same or opposite breast), or is this a coding error (an excision for the first surgical admission and re-excision of the same lesion for the second surgical admission)? We assumed a single lesion required two operative interventions on different dates (an excision and re-excision) to ensure complete surgical clearance.
- If for two inpatient separations within the primary treatment period, procedure code 1 of the first admission is for an excision of breast and procedure code 1 of the second admission is for mastectomy, is this the same or different breast? We

assume that a single lesion required two operative interventions on different dates (an excision and subsequent mastectomy) to ensure complete surgical clearance.

- If for the index cancer, procedure code 1 is coded as "re-excision", is this a true re-excision or a coding error? If a true re-excision, this would mean that the OACIS pathology report is not that of the index tumor. Alternatively this could be a coding anomaly where large core biopsies done as outpatient procedures (eg. at Breast Screen, or an ABBI biospy) may have been considered "primary surgery" and the inpatient surgical procedure as a "re-excision" OR a coding error where procedure code 1 should have been coded as "excision". We assume the latter reason.
- If a laparoscopic bilateral salpingo-oophorectomy is performed in a pre-menopausal woman, we assume that this is performed for ovarian prophylaxis rather than for distant metastatic disease, in the absence of any related chemotherapy.

Radiotherapy

- Radiotherapy is usually an outpatient procedure in the public sector, unless given overnight (uncommon) when it is coded as an inpatient admission. Radiotherapy services in the private sector are less well described, but usually outpatient based. This will result in an undercount of radiotherapy services provided, and means that we have very likely missed some distant metastases for which radiotherapy is offered (eg. bone metastases).
- If a principal diagnosis was a non-cancer presentation, but procedure code 1 was for radiotherapy, we assumed that the principal diagnosis was incorrect and that radiotherapy occurred as an inpatient. If this occurred after the primary treatment period, in the absence of prior second surgery or second round of chemotherapy, we assumed this represented the date of distant metastatic recurrence.

Chemotherapy

- Chemotherapy is always an inpatient procedure in the private sector. However, in the public sector, chemotherapy changed from an inpatient to an outpatient procedure on 1st July 2007. Therefore chemotherapy in the public sector from 1st

July 2007 will not be captured in ISAAC, resulting in an undercount of chemotherapy, unless given overnight (uncommon) when it is coded as an inpatient admission. It is possible to infer that chemotherapy occurred as an outpatient procedure if there are inpatient separations for insertion/removal of a vascular access device under general anaesthesia (central line placement) OR administration of packed cells/platelets indicating bone marrow suppression usually from chemotherapy. This group of patients were assumed to have had distant metastases, but we were unable to ascribe an accurate date of recurrence as we did not know the exact date on which the chemotherapy was commenced. We therefore assumed chemotherapy commenced within one month of a suspicious mammogram, or four months prior to an inpatient admission for bone marrow suppression.

- We did not know if a metastasis developed during initial chemotherapy unless the principal diagnosis code changes during a chemotherapy course to suggest that this has occurred. We did not know the exact date of recurrence in this group of patients, so we therefore assumed that the inpatient separation corresponding to the change in principal diagnosis reflected the date of distant recurrence.
- It was often unclear for women who had very long chemotherapy courses without interruption, exactly why this has occurred. Was this a poor prognosis tumor in a young woman; was the primary breast cancer metastatic at diagnosis (and we didn't detect this in OACIS); did a metastasis develop during a standard chemotherapy course but was not coded as such; did the patient have co-morbidities that necessitated less toxic chemotherapy regimens over a longer time period; was the patient on a clinical trial of a new agent; was the type of chemotherapy changed for some reason which extended the treatment length; did the patient have a preference for chemotherapy with different side-effect profiles to what was normally provided, or were there other reasons? We assumed that this was a long course of primary chemotherapy without development of metastatic disease during its duration.
- If there was a second round of chemotherapy, we assumed that a principal diagnosis of "Pharmacotherapy session for neoplasm", in the absence of another principal diagnosis to suggest a different primary cancer (eg AML, Multiple Myeloma etc), referred to chemotherapy for distant metastases for breast cancer.

In the presence of a new principal diagnosis, we assume that the chemotherapy was for treatment of the new primary cancer, and that the woman remained disease free from her breast cancer.

- After 1/7/2007 (date when chemotherapy became an outpatient in the public sector), if one or two cycles of a second round of chemotherapy were given as an inpatient, it is possible that the chemotherapy was started as an outpatient but the patient developed toxicities necessitating an inpatient admission where chemotherapy was continued OR it could be a second round of chemotherapy for distant metastases commenced as an inpatient to be continued upon discharge OR chemotherapy was started as an inpatient and discontinued before completion for some reason. We assumed that all women received a complete course of chemotherapy. We estimated date of distant recurrence for women who were deemed to have started chemotherapy as an outpatient (as previously described). If we decided that the second round of chemotherapy had commenced as an inpatient, we used the date of commencement to reflect the date of distant recurrence.

Other issues

- Date of birth: seemed to vary between OACIS and ISAAC for some Study IDs. The difference was often 2-3 weeks, but we were unable to explain why.
- Date of recurrence: we chose to ascribe the date of recurrence to the date of chosen therapeutic intervention eg surgery, chemotherapy, radiotherapy, as within ISAAC we had no access to the pathology reports to determine the exact date of diagnosis (unlike OACIS).
- No procedure: an inpatient separation was ascribed but the only annotation for presenting complaint was “No procedure”. It is unclear what this means.
- Inpatient separations for lumbar back pain: we assume this is for a co-morbidity eg. facet joint disease, disc prolapse etc, rather than for metastatic disease.
- Inpatient imaging: we assumed that inpatient imaging during primary chemotherapy is for a non-breast cancer indication, rather than reflecting development of distant metastases during initial treatment. This is based on our

assumption that our cohort of women were correctly identified as having early breast cancer, were staged pre or perioperatively with no evidence of spread beyond the breast/axilla, had their disease completely removed by surgery, and that the adjuvant treatment is to prevent recurrence. In this setting, imaging is not required for monitoring effect of breast cancer treatment, but could be for assessment of treatment complications eg. chest infection, or for a concurrent non-breast cancer problem.

- Inability to determine laterality of breast lesion in ISAAC: if ISAAC data only is used, it is impossible to determine if a relapse within the breast is in the treated breast (true local recurrence) or opposite breast (new primary breast cancer).
- Linkage error: we assume that the wrong women had been linked in ISAAC when we identified a pathology report in OACIS that indicates surgical removal of a malignant breast lesion, but not only was there was no corresponding inpatient separation recorded in ISAAC, but later breast related inpatient separations identified in OACIS for these women (n=3) were for excision of benign breast disease. These women were excluded from our study cohort.
- Insufficent information from 2 data systems: for 14 women we were unable to determine their recurrence status. They were deemed uncertain to have a recurrence in OACIS, and there was either no record in ISAAC (n=13) or the only record in ISAAC was for chemotherapy (n=1). These women were excluded from our study cohort.
- Decision regarding final recurrence status: any woman deemed to have had a recurrence in OACIS or ISAAC or both, was called “recurrence”. For women deemed not to have a recurrence in OACIS, all assessments in ISAAC called uncertain (n=336) were generously given the benefit of the doubt and called “no recurrence”(table 9).

RESULTS

The patient level dataset for all 1428 women contained all relevant surgical histopathology of the index primary breast cancer and surveillance mammography. A further 3 women were removed when it was noticed that their OACIS date of birth was missing (1 postmenopausal woman moderate NPI, no recurrence, alive; 1 postmenopausal woman moderate NPI, breast recurrence, breast cancer death; and 1 postmenopausal woman excellent NPI, no recurrence, alive). The final study cohort of 1425 women, had 79 breast recurrences and 49 distant recurrences.

The final study cohort of 1425 had a similar proportion of premenopausal (22.53%) women to the SACR cohort (21.62%) (table 10). There were however considerably larger numbers of premenopausal women in the SACR cohort for whom the NPI category could not be calculated (7.94%) compared to the study cohort (~1.96%). Breast cancer mortality was higher in the SACR cohort (2.49%) than in the study cohort (1.26%), which is likely to reflect the inclusion of advanced breast cancers within the SACR. It is possible that the women in the SACR cohort for whom we had insufficient information to calculate the NPI, would have been in the moderate or poor prognostic subgroups.

Table 10: Prognostic subgroups and mortality in premenopausal women with invasive breast cancer – SACR vs final study cohort

Premenopausal								
NPI Category	Women				BCD		OCD	
	SACR		Study		SACR	Study	SACR	Study
	No.	%	No.	%	%	%	%	%
Excellent	127	1.36	23	1.61	0.00	0.07	0.01	0
Good	282	3.02	52	3.64	0.04	0.07	0.03	0
Moderate	598	6.40	146	10	0.57	0.28	0.03	0.14
Poor	272	2.91	76	5.32	0.67	0.63	0.03	0.07
Cannot calculate	742	7.94	28	1.96	1.21	0.21	0.09	0.07
TOTAL	2021	21.62	325	22.53	2.49	1.26	0.19	0.28

NPI = Nottingham Prognostic Index
SACR = South Australian Cancer Registry
BCD = breast cancer death, OCD = other cause death

The final study cohort of 1425 also had a similar proportion of postmenopausal (77.12%) to the SACR cohort (78.38%) (table 11). There was a four-fold difference in the number of postmenopausal women in the SACR cohort for whom the NPI category could not be calculated (31.52%) compared to the study cohort (7.92%). Breast cancer mortality was also higher in the SACR cohort (10.41%) than in the study cohort (5.32%), and this is likely to be for similar reasons as in the premenopausal group (see above).

Table 11: Prognostic subgroups and mortality in postmenopausal women with invasive breast cancer – SACR vs final study cohort

Postmenopausal								
NPI Category	Women				BCD		OCD	
	SACR		Study		SACR	Study	SACR	Study
	No.	%	No.	%	%	%	%	%
Excellent	885	9.47	176	12.35	0.13	0	0.43	0.84
Good	1252	13.40	250	17.51	0.27	0.35	0.64	0.63
Moderate	1693	18.11	407	28.56	1.69	1.96	1.14	1.75
Poor	549	5.87	154	10.78	1.67	2.24	0.31	0.77
Cannot calculate	2946	31.52	113	7.92	6.66	0.77	4.15	1.12
TOTAL	7325	78.38	1100	77.12	10.41	5.32	6.70	5.11

NPI = Nottingham Prognostic Index
SACR = South Australian Cancer Registry
BCD = breast cancer death, OCD = other cause death

Within our premenopausal group (n=325), 1.68% of women developed breast recurrences and 1.12% developed distant recurrence. Breast cancer mortality in the premenopausal group was 1.26%. Within our postmenopausal group (n=1100), 3.85% of women developed breast cancer recurrences, 2.24% developed distant recurrence, and 5.32% died from breast cancer (table 12).

Table 12: Proportion of breast and distant recurrence in pre and post menopausal women in the final study cohort.

NPI Category	Premenopausal				Postmenopausal			
	Women	Breast	DM	BCD	Women	Breast	DM	BCD
	%	%	%	%	%	%	%	%
Excellent	1.61	0	0	0.07	12.35	0.14	0	0
Good	3.64	0.07	0.07	0.07	17.51	0.63	0	0.35
Moderate	10	0.77	0.35	0.28	28.56	1.68	0.77	1.96
Poor	5.32	0.56	0.56	0.63	10.78	0.84	1.12	2.24
Cannot calculate	1.96	0.28	0.14	0.21	7.92	0.56	0.35	0.77
TOTAL	22.53	1.68	1.12	1.26	77.12	3.85	2.24	5.32

NPI = Nottingham Prognostic Index
Recurrence = relapse within the treated breast or axilla, or a new primary or metastasis in the opposite breast
DM = distant metastases
BCD = breast cancer death

SUMMARY

The purpose of this chapter was to describe the process by which we sourced, extracted and linked data in order to determine recurrence status for women diagnosed with early breast cancer in South Australia between 2000-2008, given the absence of routinely collected recurrence data.

Despite the difficulties experienced, we were able to construct a patient level dataset for 1425 women with early breast cancer. In our dataset, women were grouped into prognostic subgroups (based on patient and tumor characteristics at diagnosis) and the mammographic surveillance schedule for each woman was defined. Recurrence data was extracted or estimated from OACIS and ISAAC data systems, and linked with date and cause of death as of August 2010, from the South Australian Cancer Registry.

Limitations of our methods reflect four key issues:

1. Information available within accessible data systems - key limitations relating to content and quality of routinely collected data within the SACR, OACIS and ISAAC systems, as they pertain to our study, are outlined in table 13.
2. Accuracy of linking data systems - we were unable to link almost 25% of our study cohort, despite using a linkage algorithm previously developed and successfully used by Prof Karnon and his team. The reason for this remains unclear.
3. Assumptions made in interpreting the data – as previously described, we have had to make a number of assumptions when interpreting the OACIS and ISAAC data. The extent to which our assumptions match the reality of the patient's disease course cannot be confirmed, and will affect the accuracy of our findings.
4. Information to which we had no access - we did not have access to mammography reports from private radiology practices. For women with fewer than expected mammogram reports on OACIS (assuming one mammogram per year for every year following treatment), it is possible that our records of their surveillance mammography are incomplete, rather than that women missed or stopped having mammograms. We also do not have any mortality data for women who were treated in South Australia but subsequently moved interstate.

Table 13: Major data issues within the South Australian Health system

SACR	OACIS	ISAAC
Recurrence data not routinely collected	No unique patient identifier	No access to MMG or pathology reports
Hormonal receptor data not routinely collected prior to June 2012	No unique test identifier	Need to use surrogate measure of recurrence
HER2 receptor data not routinely collected	Missing data from July 2003-June 2004	Data unavailable from January to June 2000
Have date and cause of death but require permission from Registrar of Births, Deaths and Marriages to release the data	Vital status but not cause of death	No mortality data
Only a small number of data points from pathology reports are available electronically	Medicare number inconsistently recorded	Chemotherapy changed from inpatient to outpatient procedure in mid 2007, radiotherapy mainly provided as an outpatient procedure

SACR = South Australian Cancer Registry
OACIS= Open Architecture Clinical Information System
ISAAC= Integrated South Australian Activity Collection

Four key issues for data development have become apparent during this research:

1. The need for routinely capturing stage information in population based cancer registries. We were unable to ascertain tumor, node or grade (or a combination) for 40% of women diagnosed with breast cancer from 2000-2008 in the SACR. This suggests that either they were not reported by pathologists, or were incompletely transcribed from the pathology reports into the electronic database within the SACR. From a population perspective, this makes it very difficult to identify mortality trends over time by stage of the index breast cancer at diagnosis.
2. The need for routinely capturing recurrence information in population based cancer registries. From a population perspective, this makes it very difficult to identify recurrence rates over time by stage of index breast cancer at diagnosis, and the effect of timing of detection of recurrence on overall survival.

3. The ability to be able to access imaging reports for cancer surveillance across both public and private sectors. From a population perspective, this makes it very difficult to assess the effectiveness of our current mammographic surveillance strategies. That is, we are unable to compare mortality between women who have impalpable recurrence detected on mammography with women who have a palpable recurrence detected between surveillance mammograms.
4. To update the prognostic indicators routinely recorded within the cancer registries to more accurately reflect contemporary clinical knowledge and practice. Current breast cancer treatment is guided not only by tumor, node and grade information, but also by receptor status (ER, PR, HER2). Addition of hormonal receptor information has only recently (June 2012) been added to data items routinely collected by the SACR. HER2 status is not yet collected.

Our patient level dataset was used to populate our health economic model, with the purpose of developing cost-effective mammographic surveillance strategies that are tailored to risk of recurrence. However, the implications of the poor data quality on the model analysis and interpretation of our modelling results cannot be underestimated. The analyses that could be performed were severely restricted by the nature of the available data (e.g. receptor status). In fact, we decided not to use the extracted distant metastases recurrence data as a calibration target for the model because they were deemed too unreliable for reasons mentioned previously in this chapter. Given the difficulties in sourcing, extracting and linking the data it is highly likely that the final study cohort is a biased sample, and any findings of the modelling analysis must therefore be interpreted with caution. However, data quality does not detract from the potential value of the methodology developed for the use of patient-level data to inform a simulation-based approach to evaluating surveillance programs.

Chapter 6

DISCRETE CHOICE EXPERIMENT

TITLE

What sort of follow-up services would Australian breast cancer survivors prefer if we could no longer offer long-term specialist based follow-up? – a Discrete Choice Experiment

INTRODUCTION

Breast cancer is the most common incident form of malignancy in Australia accounting for 28% of cancer diagnoses in 2008.[1] It is a well characterised disease with clear guidelines available through Cancer Australia for follow-up of women with early breast cancer.[2] Cancer Australia has defined the aims and objectives of follow-up: to detect and treat local recurrence; to deal with adverse effects of treatment; to provide psychological support; to screen for a new primary breast cancer; review and update family history; observe outcomes of therapy, and review treatment including the potential for new therapies.[2] In addition, Cancer Australia identifies that there is a lack of clarity about the "optimal duration and frequency of follow-up", and it is apparent that this is a contentious area.[3,4]

Clinical cancer services around Australia are facing a rapid expansion in workload through a combination of three factors: (a) the incidence of cancer is expected to increase with age,[5] and with unprecedented ageing of the population, there will be growing numbers in the most cancer-prone age groups; (b) increasing incidence of cancer that can only be partly explained by the ageing and increasing size of the population;[5] and (c) falling age-standardised mortality rate for all cancer combined,[5] expanding the pool of survivors. There are economic and workforce imperatives to changing models of care and reconfiguring service provision within the public health sector. Relevant research suggests that nurses may be better at identifying psychological concerns and side-effects of drug treatment than clinicians,[3] and shared cared models of breast cancer follow-up between Cancer specialists and General Practitioners have recently been the focus of a demonstration project within Australia by Cancer Australia.[6]

Patient preferences can also inform clinical decision-making and improve satisfaction and adherence to health programs.[7] Discrete Choice Experiments (DCE) have become a commonly used technique in health economics.[8] The technique is an attribute-based measure of benefit, based on the assumption that interventions (e.g. breast cancer follow-up services) can be described by a number of salient or key attributes (characteristics) and that an individual's valuation of the intervention depends upon the levels of these attributes.[9] Important attributes may be identified from literature reviews, expert clinical opinion, patient focus groups or individual patient interviews. Patients are given hypothetical scenarios comprising different levels of attributes and asked to choose between two or more alternatives. Patients' stated preferences reflect their perceived benefit, where benefit is defined in terms of the economic concept of "utility" or value.[8]

Given limited resources, if we are considering designing new follow-up services that are both appropriate and acceptable to women, we need to understand which attributes of the service women value most. The aim of this study was to determine the preferences of breast cancer survivors for possible alternative modes of delivery of follow-up services in years 3, 4, and 5 after diagnosis.

METHODS

Producing scenarios

Salient attributes and levels for inclusion within the DCE were established from a literature review. We searched in PubMed from inception through to December 31, 2011, limiting to original papers, literature reviews or systematic reviews published in English. Key search terms were text words "breast cancer", "follow*", "surveillance", "monitoring", "survivor*"; and MESH terms "patient satisfaction" and "patient preference*". Papers that focussed on either the experiences of women undergoing breast cancer follow-up, or on women's preferences for alternative modes of delivery of follow-up services were sought. In reviewing the identified papers, greater weight was given to recent Australian publications.

A key set of papers included a recently published body of qualitative research on breast cancer follow-up in Australia. Australian women reported a high level of satisfaction with their current specialist based care and an initial reluctance to consider models of care that would involve them moving away from the cancer specialist.[10] There was also a perceived need for additional training of General Practitioners (GPs) or Breast Cancer Nurses if they were to have an increased role in follow-up care.[11] Patients highlighted psychosocial needs and menopausal symptoms as areas of the follow-up consultation that needed improving[11] While recognising advantages to GP follow-up, there was a stronger level of support for shared care between the GP and a specialist, rather than a complete transfer to GP led care.[10]We explicitly excluded the shared care option as the objective of the study was to identify alternatives to specialist follow-up and previous research has shown that respondents tend to prefer what they know best or have experienced.[12,13]

Four attributes relating to service organisation were identified: type of provider, location, frequency, and method of delivery of routine follow-up care. A fifth attribute was created from "gaps" described in the literature relating to existing service provision, with the intent to evaluate the perceived benefit of offering "drop-in" clinics to provide additional support to women living with breast cancer. Based on the literature and expert consultation, three levels for each of the five attributes were developed to examine characteristics of interest to the Australian health care context:

1. Clinician-
    (a) Breast Physician:
    Usually General Practitioners who have undergone specialised training in breast medicine[14]
    (b) General Practitioner:
    Local doctor providing whole person health care to individuals and families in the community[15]
    (c) Breast Cancer Nurse:
    Registered nurse with additional training and certification in breast cancer care[16]
2. Frequency-every 6, 9 or 12 months
3. Location-hospital clinic, general practice, local breast cancer follow-up clinic

4. Method-face-to-face, telephone, alternate between face-to-face and telephone
5. Drop-in clinics-treatment side-effects, psychosocial support, secondary prevention

The 5 x 3 design resulted in 254 (3^5) possible scenarios. We used a fractional factorial design and the techniques developed by Street and Burgess to reduce this to a more pragmatic 18 binary choice sets which were 100% efficient for the estimation of main effects.[17] This design was then divided into 3 blocks or versions, each containing a total of 6 choice sets for presentation to participants. The detailed definition of the chosen attributes and attribute levels for the DCE are presented in Table 1.

Table 1: Attributes and levels

Attributes	Levels	Description
CLINICIAN	Breast Physician	Most are General Practitioners who have undergone additional training in breast assessment, planning and co-ordinating breast cancer treatment, and counselling.*
	General Practitioner	Your local doctor who provides person centred, continuing, comprehensive and coordinated whole person health care to individuals and families in their communities.* This is your own local doctor who you normally see when you get sick (eg. coughs & colds, blood pressure, diabetes, infections etc)
	Breast Cancer Nurse	A registered nurse who has completed extra study (a Graduate Diploma or higher at University) in the specialty of cancer nursing or its equivalent.*
FREQUENCY	Every 6 months	Follow-up visit scheduled every 6 months
	Every 9 months	Follow-up visit scheduled every 9 months
	Every 12 months++	Follow-up visit scheduled every 12 months
LOCATION	Hospital clinic	A breast cancer follow-up clinic at the hospital where you had your treatment, but you would no longer be seen by the cancer specialist/s who provided your surgery, chemotherapy or radiotherapy
	General practice++	Your local general practice
	Local breast cancer follow-up clinic	A breast cancer follow-up clinic in your local area
METHOD	Face-to-face	A visit to a clinician where you would have a clinical breast examination, and discussion about your wellbeing and issues related to your breast cancer care
	Telephone++**	At a scheduled time, a clinician telephones you and you have a discussion about your well-being and issues related to your breast cancer care
	Alternate between face-to-face and telephone	Your first appointment is a clinic visit, your second appointment is by telephone, your third appointment is a clinic visit and so on
DROP-IN CLINICS	Treatment side-effects clinic	For management of lymphedema, menopausal symptoms, sexual dysfunction etc that relate to your surgery, chemotherapy or radiotherapy for breast cancer
	Psychosocial support clinic++	For identification, referral and management of depression, anxiety, and issues relating to self-image, relationships and return to work
	Secondary prevention clinic	For advice about diet, alcohol and exercise, to both develop and maintain good health and reduce the risk of breast cancer returning

++ Reference group in modelling

* adapted from sources outlined in Methods section.

**It is important to note that if your care was provided by telephone but it becomes clear that you needed a face-to-face appointment, this would be arranged within 1 week

The questionnaire started with a short introduction on the background to the research study, and an explanation of what the questionnaire would involve (including a practice example). The first section of the questionnaire contained the Discrete Choice Experiment. In the preamble, participants were given a hypothetical context (Figure 1); followed by a comprehensive description of each of the attributes and levels. Women were then asked to complete six questions. Each question contained a pair-wise choice, and participants were asked to choose the option they most preferred for their routine follow-up appointments in years 3, 4 and 5 following diagnosis of breast cancer (Figure 2). The second section consisted of background information regarding demographics, family history, features related to breast cancer diagnosis, treatment and follow-up, and self-assessed level of anxiety regarding cancer and other health concerns[18] (Appendix A). The questionnaire was designed principally for on-line administration. However a paper version was made available upon request for women who expressed a preference for a paper version and/or who did not have easy access to the internet.

Figure 1 – Preamble to DCE

For the purpose of this study, please **imagine** the following:

Your breast cancer care is provided by a cancer specialist at the hospital for the first 2 years after your diagnosis and you are then discharged to follow-up care with a different clinician in a different location. This medical practitioner would be responsible for your routine follow-up appointments in years 3, 4 and 5, however if a serious problem arose, you would be referred immediately back to the cancer specialist who provided your initial treatment (please note that your follow-up would still continue beyond 5 years, but for the purpose of this study we would like you to focus on years 3, 4 and 5). You would be expected to attend your routine follow-up appointments at this new location. This new location also runs weekly "drop-in" clinics for further advice and support about living with breast cancer. These "drop-in" clinics are **in addition** to your routine follow-up appointments, require no booking, and you can use them as little or as often as you wish.

You will be asked to answer 6 questions about **hypothetical** breast cancer follow-up programs. Each question contains a pair of options for you to choose between. The features of the follow-up programs will differ in five ways:

1. Which clinician would provide your care
2. How often your appointments would be
3. Where your follow-up appointment would occur
4. Type of routine follow-up appointment
5. Type of additional drop-in clinics offered

Figure 2 – Example of a DCE question*

1.1 For your routine follow-up appointments in years 3, 4 and 5, which program do you prefer?

Program A	Program B
Breast Physician	General Practitioner
Every 6 months	Every 12 months
Local breast cancer follow-up clinic	Hospital clinic
Face-to-face	Alternate between face-to-face and telephone
Psychosocial support clinics available	Secondary prevention clinics available

◯ Program A
◯ Program B

*Scenarios varied over 18 pair wise choices

A small pilot study (n=10) was administered by one author (TB) in a follow-up breast clinic at a major public hospital in Adelaide. The main objectives of the pilot study were to establish face validity and identify any problems in understanding the requirements of the questionnaire or individual questions. Feedback from the pilot study resulted in some minor changes in question wording to improve clarity.

Recruitment

Patients with a personal history of breast cancer, who had completed their primary treatment (surgery, radiotherapy and chemotherapy) were eligible for the study.

Women were recruited from a variety of settings including Breast Surgical Oncology clinics in both the public and private health sector (SA only), local and national print media, and patient support and advocacy groups including Cancer Voices SA, Cancer Council Australia, Breast Cancer Network of Australia, National Breast Cancer Foundation and Register 4, over a 6 month period from May to October 2012.

Patients recruited through their treating clinician, were provided with a study envelope at the end of their consultation if they met the inclusion criteria. The patient, if willing to participate, was asked to complete the questionnaire at home and return the questionnaire within the stamped addressed envelope provided, or alternatively, access the survey online using the URL provided in the Letter of Invitation within the study envelope.

Patient support and advocacy groups were asked to disseminate a short description of the study to their members through their usual communication channels. Patients were also recruited through print media (a local newspaper, magazine for senior citizens, and a state-based journal of the Australian Medical Association). Women who chose to participate could either complete the survey online, or contact the research team requesting a study envelope to be posted to them.

Block randomisation occurred at the online survey site, and by ensuring that each clinician received an equal number of all three versions of the survey to provide to their patients.

Ethics approval

Approval was obtained from the Human Research Ethics Committees of the University of Adelaide, and each of the three 3 public and 3 private Adelaide hospitals participating in the study.

Socio-economic status

Socio-economic status was classified using the Socio-economic Indexes for Areas, 2006 (SEIFA 2011 had not been released at the time of writing), from the Australian Bureau of Statistics.[19]

The Postal Area (POA) Index of Relative Socio-economic Advantage and Disadvantage (IRSAD) was applied to derive national deciles from individual postcodes, with decile 1 indicating the most disadvantaged and decile 10 the most advantaged areas.

Data analysis

The DCE data was analysed within a random utility maximisation framework.[20] The empirical model to be estimated is specified as:

$$U_{itj} = x'_{itj}\beta_i + \varepsilon_{itj}$$

where U_{itj} is the utility individual i derives from choosing alternative j in choice scenario t, x_{itj} is a vector of observed attributes of alternative j, β_i is a vector of individual specific coefficients reflecting the desirability of the attributes, and ε_{itj} is a random error term. In order to investigate the potential for preference heterogeneity, the mixed logit model, also known as the random parameter logit model, was utilised.[21-23] Within the mixed logit model the β_i is expressed as $\beta + \eta_i$, where β constitutes the vector of average preferences of the population for each attribute and η_i is the individual's specific preference components. In this study, it is assumed that all coefficients of attribute levels are random with normal distribution and are freely correlated. The utility function for the DCE was of the following form:

$$U_{itj} = (\beta_1 + \eta_{1i})Physician_i + (\beta_2 + \eta_{2i})Nurse_i + (\beta_3 + \eta_{3i})SixMonths_i + (\beta_4 + \eta_{4i})NineMonths_i + (\beta_5 + \eta_{5i})FollowupClinic_i + (\beta_6 + \eta_{6i})HospitalClinic_i + (\beta_7 + \eta_{7i})Face2face_i + (\beta_8 + \eta_{8i})Alternate_i + (\beta_9 + \eta_{9i})Sideffects_i + (\beta_{10} + \eta_{10i})Prevention_i + \varepsilon_{itj}$$

where *Physician* and *Nurse* are dummy variables for Breast Physician and Breast Cancer Nurse in the clinician attribute, *SixMonths* and *NineMonths* are dummy variables indicating every 6 or 9 months in the follow-up frequency attribute, *Follow-upClinic* and *HospitalClinic* are dummy variables representing local breast cancer follow-up clinic and hospital clinic in the location attribute; *Face2face* and *Alternate* are dummy variables indicating face-to-face and alternate between face-to-face and telephone for the method attribute, *Sideffects* and *Prevention* are dummy variables for treatment side-effects clinic and secondary prevention clinic for the Drop-in clinics attribute. All attributes were dummy coded, with one level for each attribute coded as the referent (refer to Table 1 and 2 for referent levels).

From the fitted model, a statistically significant parameter estimate for an attribute level indicates it's importance in influencing the preferences of respondents. A positive (negative) sign indicates that this attribute level is preferred (not preferred) to the base level of the attribute. Internal validity (i.e. the extent to which results are consistent with *a priori* expectations) was tested by examining the sign and significance of parameter estimates. We expected all attributes to have a significant influence on preferences for follow-up. Based on our literature review and expert opinion, the *a priori* assumptions were that more frequent face-to-face visits with a General Practitioner (GP)/specialist GP (Breast Physician) in a local community setting close to home which provided additional psychosocial support clinics, would be the most preferred follow-up options for years 3-5 in our hypothetical scenarios.

The main outputs from the mixed logit model are estimates of the proportions of respondents who prefer each attribute level, compared to the reference level for each attribute. For

example, for the attribute 'clinician', the proportion of the population with a preference for Breast Physician compared to a General Practitioner can be estimated. Following Kimman et al[12] a series of interaction terms between the attribute levels and respondents' characteristics (reflecting socio-economic status, education and clinical characteristics including the number of years since initial diagnosis, age at diagnosis and the type of breast cancer diagnosed) were included into the utility function to further explore possible observable sources of preference heterogeneity.

Preference scores (V_j - also called "indirect utility scores"[8]) were generated as the sum of the model coefficients for every combination of attribute levels. The probability that each combination of attribute levels is the most preferred scenario (P_j) is further calculated as:

$$P_j = \frac{\exp(V_j)}{\sum_{k=1}^{J} \exp(V_k)}$$

where j = 1, …, J. In this paper, only the 10 most highly ranked scenarios are considered.

RESULTS

Characteristics of respondents

A total of 836 women accessed the survey, of whom 722 (86.4%) completed the DCE and were included in the analysis. Of the participants, 75.6% were born in Australia and 98.2% were non-indigenous. The largest number of responses was from the most densely populated states (New South Wales 26.7%, South Australia 26.5%, Victoria 19.8% and Queensland 17.2%) but all states and territories were represented. At diagnosis, the majority of women had early invasive breast cancer (74.1%), and were between 40 and 60 years of age (67.8%). The largest group of respondents was 2-5 years post diagnosis (40.7%), followed by women 5-10 years post diagnosis (26.6%). Most study participants (83.5%) saw a cancer specialist (breast surgeon, medical oncologist, radiation oncologist, and specialist registrar in training) most frequently for their breast cancer follow-up. Complete characteristics of the respondents are presented in Appendix A.

Discrete choice experiment results

The mixed logit estimates for the total sample are reported in Table 2.

Table 2: Mixed logit estimates for total sample

Attributes	Levels	Reference level for each attribute	Mean coefficient		Standard deviation	
			Coefficient	SE	Coefficient	SE
CLINICIAN	Breast Physician	Compared with routine follow-up provided by General Practitioner	2.692**	(0.331)	1.911**	(0.279)
	Breast Cancer Nurse		1.167**	(0.207)	2.191**	(0.315)
FREQUENCY	Every 6 months	Compared with routine follow-up every 12 months	0.375*	(0.166)	1.615**	(0.298)
	Every 9 months		0.178	(0.149)	1.049**	(0.284)
LOCATION	Local breast cancer follow-up clinic	Compared with routine follow-up at local general practice	1.118**	(0.172)	0.805*	(0.349)
	Hospital clinic		0.016	(0.169)	1.232**	(0.341)
METHOD	Face-to-face	Compared with routine follow-up appointment by phone	2.711**	(0.311)	1.855**	(0.334)
	Alternate face-to-face & telephone		2.145**	(0.256)	1.476**	(0.295)
DROP IN CLINICS	Treatment side-effects clinics	Compared with drop-in clinics providing psychosocial support	0.898**	(0.196)	1.240**	(0.377)
	Secondary prevention clinics		0.404*	(0.186)	1.567**	(0.326)
Log likelihood			-2218.025			
AIC			4566.050			
N			722			
Observations			4332			

Standard errors (SE) in parentheses. ** p<0.01, * p<0.05. Random parameters correlated. AIC: Akaike Information Criterion. Number of Halton draws: 500.

As hypothesized *a priori*, all attributes were found to have a significant influence on preferences for follow-up. The results indicate that women demonstrated strongest positive preferences for a Breast Physician (followed by a Breast Cancer Nurse), 6-monthly visits, local breast cancer clinic, face-to-face attendance (followed by alternate face-to-face and telephone), and drop-in clinics for treatment side-effects (followed by secondary prevention) (all $p<0.05$). The parameter estimates relating to face-to-face follow-up and the Breast Physician were in the expected direction, providing evidence of internal validity. The statistically significant standard deviation of all random coefficients ($p<0.05$) also confirmed the existence of preference heterogeneity. Since all coefficients of attribute levels are assumed to be normally distributed, the mixed logit estimates relating to the mean coefficient and standard deviation for each attribute level were applied to calculate the distribution of preference heterogeneity. For example, the coefficient (standard deviation) of Breast Physician is 2.692 (1.911), indicating that 92% of the respondents exhibited a preference for the breast cancer follow-up service to be provided by a Breast Physician. Similarly, the results indicate that 70% of respondents would prefer to see a Breast Cancer Nurse over a General Practitioner and 59% of respondents would prefer to have appointments every 6 months over every 12 months.

The extent to which preference heterogeneity was related to observable characteristics was explored by including the interaction terms between respondents' characteristics and attribute levels. No interaction terms were found to be statistically significant suggesting that preference heterogeneity was largely unobservable and was not systematically related to respondents' characteristics.

Predicting choice probabilities for different breast cancer follow-up scenarios

To illustrate respondent preferences for the factors in combination, Table 3 presents the 10 most valued surveillance strategy scenarios, all of which included a Breast Physician as the provider of follow-up care. The most preferred scenario is a face-to-face local breast cancer follow up clinic held every 6 months and led by a Breast Physician, where additional clinics focused on the side effects of treatment are also provided.

Table 3: Preference scores and predicted probabilities within the top 10 breast cancer follow-up scenarios

CLINICIAN	FREQUENCY	LOCATION	METHOD	DROP IN CLINICS	Pref. score	Prob.	Rank
Breast Physician	Every 6 months	Local breast cancer follow-up clinic	Face-to-face	Treatment side-effects	7.794	0.192	1
Breast Physician	Every 9 months	Local breast cancer follow-up clinic	Face-to-face	Treatment side-effects	7.597	0.158	2
Breast Physician	Every 6 months	Local breast cancer follow-up clinic	Face-to-face	Secondary prevention	7.300	0.117	3
Breast Physician	Every 6 months	Local breast cancer follow-up clinic	Alternate face-to-face & telephone	Treatment side-effects	7.228	0.109	4
Breast Physician	Every 9 months	Local breast cancer follow-up clinic	Face-to-face	Secondary prevention	7.103	0.096	5
Breast Physician	Every 9 months	Local breast cancer follow-up clinic	Alternate face-to-face & telephone	Treatment side-effects	7.031	0.090	6
Breast Physician	Every 6 months	Local breast cancer follow-up clinic	Alternate face-to-face & telephone	Secondary prevention	6.734	0.067	7
Breast Physician	Every 6 months	Hospital clinic	Face-to-face	Treatment side-effects	6.692	0.064	8
Breast Physician	Every 9 months	Local breast cancer follow-up clinic	Alternate face-to-face & telephone	Secondary prevention	6.537	0.055	9
Breast Physician	Every 9 months	Hospital clinic	Face-to-face	Treatment side-effects	6.495	0.052	10

Pref.score = preference score, Prob = probability

The data indicate that women would prefer to reduce the frequency of follow-up from 6 to 9 months rather than alter the location or method of delivery, but they would accept alternating methods of delivery, and follow-up to be based at a hospital based clinic, to a decrease in follow-up frequency to 12 monthly. The rankings also show that women would be prepared to accept alternating methods of delivery (face-to-face and telephone-based visits) before switching the location of the clinic to a hospital setting.

DISCUSSION

Our results demonstrate clear preferences of Australian breast cancer survivors for delivery of their follow-up care in the absence of specialist follow-up in years 3, 4 and 5. Breast Physicians were the most preferred provider, followed by Breast Cancer Nurses and then General Practitioners. Breast Physicians bring a specialist level of knowledge to survivorship care with many of the benefits of a GP approach.[24]The observed preference for Breast Cancer Nurses to local GPs suggests that specialised training in breast medicine is valued by respondents. Additional reasons for this order of preferences could be explored by qualitative research, but was beyond the scope of this study. A follow-up service located in the community had broad appeal to women, and while face-to-face visits were preferred, women would consider alternating face to face visits with telephone contact if this meant that the frequency of contact with the follow-up service was not reduced.

To our knowledge this is only the second discrete choice experiment investigating patient preferences for breast cancer follow-up, and the first in an Australian context. Kimman et al[12] used DCE methodology to assess patient preference for the first year of breast cancer follow-up in the Netherlands, described by attributes of attendance at an educational group program, frequency of visits, waiting time, contact mode and type of health care provider, for 331 women. The authors demonstrated that “overall patient satisfaction would be similar if patients were followed up by a medical specialist alternating with a Breast Cancer Nurse compared to follow up by a medical specialist only”. Kimman et al also found preference heterogeneity for most attributes, “indicating that one strategy does not fit all”.[12]

It is important to note that our study differs from Kimman's study in a number of key areas. Firstly, we elicited the views and preferences of a large sample of Australian women previously treated for breast cancer, with no fixed time since completion of treatment. Secondly, the attributes and levels included in our study were also notably different from those selected by Kimman et al. The attributes and levels were carefully chosen to reflect the Australian health system and the current policy context whereby alternative modes of delivery of follow-up services are being considered to reduce the burden upon cancer specialist. Hence, the cancer specialist was removed from the scenario. A community based site for follow-up was offered as an alternative to hospital based care, and the opportunity to attend drop-in clinics to address areas of unmet need was also provided. By including a Breast Physician, a Breast Cancer Nurse and a General Practitioner, we were able to examine how women's preferences might be influenced by specialised knowledge of breast medicine (specialist GP vs general GP) and clinician craft group (specialist GP vs specialist nurse). By deliberately removing the cancer specialist (breast surgeon, medical oncologist, radiation oncologist, and registrar in training) from the hypothetical scenario, we attempted to reduce the effect of people tending to prefer what they know best or have experienced[13] which was confirmed in the Kimman study.[12] We also chose not to offer a shared care option, as we sought to identify the single key clinician women would prefer to oversee their follow-up care. By offering a hypothetical local breast cancer follow-up clinic we could assess whether a publicly funded community location for follow-up would be acceptable to breast cancer survivors.

Brennan et al identified a perceived need for additional training of GPs or Breast Cancer Nurses if they were to have an increased role in follow-up care.[11]Our study has confirmed that if patients cannot see a breast cancer specialist, the specialised knowledge of breast medicine is of more importance to women than the clinician craft group. Brennan also highlighted psychosocial needs and menopausal symptoms as areas of the follow-up consultation that needed improving.[11] Our study showed that patients valued drop-in clinics for treatment related side-effects (which includes menopausal symptoms) most strongly.

The strengths of our study include the large sample size (n=722) and national sample frame. The attributes and levels in the study were informed by international literature and contemporary qualitative research on the experiences, needs and preferences of breast cancer survivors in Australia.[10,11,25-27] The limitations of this study include lack of recruitment by breast clinicians in states and territories other than South Australia, that women had to be literate in English to participate, and we were unsuccessful in recruiting breast cancer survivors of Aboriginal or Torres Strait Islander descent (ATSI comprised <2% of the respondents). The results may not be representative of women from non-English speaking backgrounds or indigenous communities, but addressing these limitations was beyond the scope of the current study. Nonetheless, our study sample is large and diverse, and is likely to be generalisable to English-speaking women in Australia.

Tele-health or videoconferencing is potentially an attractive option for women in rural and remote areas of Australia, who could be seen by their local General Practitioner and have regular case conferences with multiple care providers simultaneously from their own home or local general practice clinic. Further research could apply DCE methodology to investigate the preferences of women who live in rural and remote areas, for this new mode of technology.

Identification of training requirements, and careful workforce planning and modelling would be required to ensure that “specialist GP” led breast cancer follow-up clinics avoid exacerbating current GP workforce shortages.[28] It would also be important to explore barriers to specialists discharging their patients to follow-up by a different clinician; and how this may be influenced by characteristics of the individual professional, the patient (eg. age, co-morbid anxiety or depression, social support, urban vs rural/remote etc), the primary breast cancer (eg. low, medium or high risk of recurrence) or other factors.

CONCLUSION

In the absence of cancer specialists, in years 3, 4 and 5 following diagnosis, Australian women would prefer to have their routine breast cancer follow-up provided by a Breast Physician (or a Breast Cancer Nurse) in a dedicated local breast cancer clinic, rather than with their local General Practitioner. Drop-in clinics for the management of treatment related side-effects and to provide advice to both develop and maintain good health are also highly valued by breast cancer survivors.

This study provides important insights into what attributes of a breast cancer follow-up service women value most, in the light of financial pressures that reduce the feasibility of all follow-up services being provided by breast cancer specialists. Our results can help inform the design of alternative service pathways that are acceptable to patients, for which further assessments of costs and patient outcomes can be undertaken.

Appendix A – Respondent characteristics (n=722)

DEMOGRAPHICS	No.	%
Country of Birth		
Africa	6	0.8
Asia	14	1.9
Australia	546	75.6
Europe	18	2.5
New Zealand	13	1.8
North America	5	0.7
Oceania	2	0.3
South America	1	0.1
United Kingdom	101	14.0
No response	16	2.2
Aboriginal or Torres Strait Islander descent		
Yes	5	0.7
No	709	98.2
No response	8	1.1
State or territory of current residence		
ACT	6	0.8
NSW	193	26.7
NT	4	0.6
QLD	124	17.2
SA	191	26.5
TAS	12	1.7
VIC	143	19.8
WA	41	5.7
No response	8	1.1
National SEIFA decile		
1 (most disadvantaged)	28	3.9
2	32	4.4
3	47	6.5
4	45	6.2
5	48	6.6
6	62	8.6
7	100	13.9
8	102	14.1
9	121	16.8
10 (most advantaged)	129	17.9
No response	8	1.1
Education		
≤ Year 11	127	17.6
Year 12	78	10.8
Certificate	88	12.2
Diploma	121	16.8
Degree	305	42.2
No response	3	0.4
Employment		
Full time	204	28.3
Part time	221	30.6
Retired	212	29.4
Unemployed	17	2.4
Home duties	64	8.9
No response	4	0.6

FAMILY	No.	%
Family status		
Live with adult/s	427	59.1
Live with adult/s + children ≤ 18years	147	20.4
Live alone	124	17.2
Live alone + children ≤ 18 years	21	2.9
No response	3	0.4
Family history of breast cancer		
Yes	278	38.5
No	439	60.8
Not sure	5	0.7
BRCA mutation in family		
Yes	26	3.6
No	452	62.6
Not sure	239	33.1
No response	5	0.7
BREAST CANCER DIAGNOSIS	**No.**	**%**
Time since diagnosis		
≤ 2years	154	21.3
>2years but <5 years	294	40.7
5-10 years	192	26.6
>10 years	77	10.7
No response	5	0.7
Age at diagnosis		
<40 years	103	14.3
40-49 years	272	37.7
50-59 years	217	30.1
60-69 years	111	15.4
>70 years	12	1.7
No response	7	1.0
Type of breast cancer		
Early invasive breast cancer	535	74.1
DCIS/LCIS	130	18.0
Advanced	41	5.7
Not sure	12	1.7
No response	4	0.6
BREAST CANCER TREATMENT	**No.**	**%**
Type of hospital where treatment received		
Public	182	25.2
Private	363	50.3
Both public and private	171	23.7
No response	6	0.8

Treatment/s received		
Surgery	708	98.1
Radiotherapy	519	71.9
Chemotherapy	453	62.7
Tamoxifen	530	73.4
Herceptin	115	15.9
None	1	0.1
Not sure	0	0
No response	4	0.7
Persistent treatment side-effects		
Yes	490	67.9
No	223	30.9
No response	9	1.2
BREAST CANCER FOLLOW-UP	**No.**	**%**
Recurrence		
Yes	46	6.4
No	667	92.4
No response	9	1.2
Clinician seen most frequently for follow-up		
Breast Surgeon	336	46.5
Medical Oncologist	224	31.0
Radiation Oncologist	17	2.4
Registrar in training	26	3.6
General Practitioner	80	11.1
Breast care nurse	6	0.8
Other	25	3.5
No response	8	1.1
Have written Survivorship Care Plan		
Yes	18	2.5
No - know about SCP	34	4.7
No - don't know about SCP	657	91.0
No response	13	1.8
ASC QUESTIONNAIRE [18]	**No.**	**%**
Future tests		
Not at all	105	14.5
A little bit	286	39.6
Somewhat	209	28.9
Very much	105	14.5
No response	17	2.4
New cancer		
Not at all	86	11.9
A little bit	283	39.2
Somewhat	218	30.2
Very much	117	16.2
No response	18	2.5

Recurrence		
Not at all	36	5.0
A little bit	231	32.0
Somewhat	235	32.5
Very much	207	28.7
No response	13	1.8
Death		
Not at all	176	24.4
A little bit	280	38.8
Somewhat	148	20.5
Very much	95	13.2
No response	23	3.2
Health		
Not at all	48	6.6
A little bit	254	35.2
Somewhat	246	34.1
Very much	154	21.3
No response	20	2.8
Children's health		
Not at all	144	19.9
A little bit	157	21.7
Somewhat	161	22.3
Very much	195	27.0
No response	65	9.0

REFERENCES

1. Australian Institute of Health and Welfare & Cancer Australia 2012. Breast cancer in Australia: an overview. Cancer series no. 71. Cat. no. CAN 67. Canberra: AIHW
2. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
3. Dixon J, Montgomery D. Extended follow-up of breast cancer patients in clinic wastes time for both patients and doctors: the case for. Breast Cancer Res. 2008; 10 Suppl 4:S7. doi: 10.1186/bcr2167
4. Cameron D. Extended follow-up of breast cancer patients in clinic wastes time for both patients and doctors: the case against. Breast Cancer Res. 2008; 10 Suppl 4:S8. doi: 10.1186/bcr2168.
5. Australian Institute of Health and Welfare & Australasian Association of Cancer Registries 2012. Cancer in Australia: an overview, 2012. Cancer series no. 74. Cat. no. CAN 70. Canberra: AIHW.
6. Cancer Australia. Shared Care. Surry Hills, NSW: Cancer Australia; 2013. http://canceraustralia.gov.au/affected-cancer/cancer-types/breast-cancer/life-after-breast-cancer/shared-care (last accessed August 9, 2013)
7. Lancasar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008; 26(8):661-77.
8. Ryan M, Gerard K, Amaya-Amaya M. Using Discrete Choice Experiments to Value Health and Health Care. Dordrecht: Springer; 2006.
9. De Bekker-Grob E, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012 Feb; 21(2):145-72.
10. Brennan M, Butow P, Spillane A, Marvin M, Boyle F. Follow up after breast cancer - views of Australian women. Aust Fam Physician 2011 May; 40(5):311-6.
11. Brennan M, Butow P, Marven M, Spillane A, Boyle F. Survivorship care after breast cancer treatment – experiences and preferences of Australian women. Breast. 2011 Jun; 20(3):271-7.
12. Kimman M, Dellaert B, Boersma L, Lambin P, Dirksen C. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncol. 2010 Apr; 49(3):328-37

13. Salkeld G, Ryan M, Short L. The veil of experience: do consumers prefer what they know best? Health Econ. 2000 Apr; 9 (3):267-70.
14. The Australasian Society of Breast Physicians [Internet]. What is a Breast Physician? Nambour: The Australasian Society of Breast Physicians; 2010. http://breastphysicians.org/index.php?option=com_content&view=article&id=1&Itemid=2 (last accessed August 9, 2013)
15. The Royal Australian College of General Practitioners [Internet]. Becoming a GP in Australia. What is a GP? What is General Practice? Melbourne: The Royal Australian College of General Practitioners; 2013. http://www.racgp.org.au/becomingagp/what-is-a-gp/what-is-general-practice/ (last accessed August 9, 2013)
16. National Breast Cancer Centre. Specialist Breast Nurse Competency Standards and Associated Educational Requirements. Camperdown, NSW: National Breast Cancer Centre; 2005.
17. Burgess L, Street D. The optimal size of choice sets in choice experiments. Statistics 2006 Dec 1; 40(6):507-15.
18. Gotay C, Pagano I. Assessment of Survivor Concerns (ASC): a newly proposed brief questionnaire. Health Qual Life Outcomes. 2007 Mar 13; 5:15.
19. Australian Bureau of Statistics [Internet]. 2033.0.55.001 - Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia - Data only, 2006. Belconnen: Commonwealth of Australia; 2008. http://abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2033.0.55.0012006?OpenDocument (last accessed August 9, 2013)
20. McFadden D. Conditional logit analysis of qualitative choice behaviour. In: Zarembka P, editor. Frontiers in Econometrics. New York: Academic Press; 1973.
21. McFadden D, Train K. Mixed MNL models for discrete response. J Appl Econ. 2000 Sept;15(5):447-70.
22. Hole A. Fitting mixed logit models by using maximum simulated likelihood. Stata J. 2007; 7(3):388-401.
23. Eberth B, Watson V, Ryan M, Hughes J, Barnett G. Does one size fit all? Investigating heterogeneity in men's preferences for benign prostatic hyperplasia treatment using mixed logit analysis. Med Decis Making. 2009 Nov-Dec; 29(6):707-15.

24. Brennan M, Jefford M. General Practitioner-based models of post-treatment follow up. Cancer Forum. 2009 http://cancerforum.org.au/Issues/2009/November/Forum/General_practitionerbased_models_of_posttreatment_follow_up.htm (last accessed August 9, 2013)
25. Jiwa M, Halkett G, Deas K, Ward P, O'Connor M, O'Driscoll C, O'Brien E, Wilson L, Boyle S, Weir J. How do specialist breast nurses help breast cancer patients at follow-up? Collegian 2010; 17(3):143-9.
26. Jiwa M, Halkett G, Deas K, Meng X. Women with breast cancers' preferences for surveillance follow-up. Collegian 2011; 18(2):81-6.
27. Lawler S, Spathonis K, Masters J, Adams J, Eakin E. Follow-up care after breast cancer treatment: experiences and perceptions of service provision and provider interactions in rural Australian women. Support Care Cancer. 2011 Dec; 19(12):1975-82.
28. Jiwa M, Ee H, Beilby J. Will promoting general practitioners with special interests threaten access to primary care? Med J Aust. 2007 Jul 16; 187(2):71-2.

Chapter 7

CONCLUSIONS

SUMMARY OF RESEARCH FINDINGS

The key findings of the respective research projects designed to answer the three research questions are described below.

Question 1: What is the impact on survival of the method and timing of detection of a second breast cancer event within the breast?

A systematic literature review was performed to identify papers that estimated the impact on survival of follow-up aimed at early detection of a second cancer event within the breast. Five electronic databases were searched. Studies were included in the review if they: (1) enrolled women treated for early invasive breast cancer without evidence of distant metastasis at primary diagnosis, who were in complete remission following primary treatment; (2) presented data on overall survival or breast cancer related survival for early vs late detection of a second malignancy in the breast OR presented data that compared survival in groups receiving different frequencies of breast cancer surveillance. Included papers were analysed by mortality outcome and critically appraised using the Downs and Black checklist.[105] Twenty one papers were included, all were level III-2 or III-3 evidence for intervention studies. No randomised controlled trials (RCTs) were identified. No good quality studies were identified, and only two studies were of good-moderate quality. Twelve papers reported a survival advantage associated with mammographic surveillance, but potential biases limit the values of these findings. Limitations of the case-control studies primarily related to the integrity of information on exposure, matching and self-selection bias. Important limitations observed in the cohort studies included selection bias, attrition bias, lead time and length bias; and lack of adjustment for confounding. The features of decision analytic modelling are described, and the strengths of the process in overcoming identified limitations in the evidence base demonstrated. These techniques could be used to guide clinical practice in the likely continued absence of RCTs in this field.

In 2011, the NHS published a commissioned Health Technology Assessment (HTA) to address “The clinical effectiveness and cost effectiveness of different surveillance

mammography regimens after the treatment for primary breast cancer: systematic reviews, registry databases analyses and economic evaluation".[106] A comparison of my systematic review with that performed as part of the HTA, confirmed that while there were methodological differences between the two studies, neither identified any RCTs and both reported biases present in the included observational studies. The HTA reports that the quality of the included studies were variably diminished by the nature of patient sampling, lead and length time bias, and attrition bias; that the majority of the studies were retrospective; and that the heterogenous nature of the data precluded meta-analysis. These issues are similar to those described in my review. The HTA authors conclude that while there is a paucity of evidence, the findings "suggest" that surveillance mammography offers a protective effect against all-cause mortality,[107,108] breast cancer specific death,[109] and death for an IBTR detected by mammography (compared to one diagnosed by clinical examination).[110] However, in contrast to the HTA, I conclude that while 12 studies report a survival benefit associated with mammographic surveillance, potential biases limit the value of these findings and as such provide no definitive indication of the effects of follow-up mammography on survival.

Question two: Is it efficient to tailor mammographic follow-up according to risk of recurrence, for women who are disease free following completion of primary treatment for early breast cancer?

Cost-effectiveness analysis is a systematic method of identifying the additional costs incurred to achieve additional units of benefit, when comparing interventions with differing outcomes.[111] Benefits are commonly represented as gains in quality adjusted life years (QALYs), where one QALY is equivalent to one additional year of life in perfect health. The aim of this research project was to determine the incremental cost per QALY gained of varying the frequencies of mammographic follow-up in postmenopausal women with varying risk of relapse within the breast; by age, QALY threshold and adherence with mammography.

A discrete event simulation model that described the progression of early breast cancer after completion of primary treatment, representing impalpable and palpable

recurrence and the detection of impalpable disease via follow-up mammography, was developed. Retrospective data from the South Australian Cancer Registry, and clinical and administrative hospital databases were linked for 407 postmenopausal women diagnosed with moderate prognosis early breast cancer from 2000-2008. These data form the basis of a patient-level probabilistic calibration process. Our results demonstrated that the most cost-effective mammographic schedule for moderate prognosis women aged between 50-69 years at diagnosis is annual follow-up for 5 years, with 2 yearly visits thereafter. For women aged 70-79 years at diagnosis, the base case results show that 2 yearly mammography is cost-effective. Probabilistic sensitivity analyses show that ongoing annual surveillance is unlikely to be cost-effective for either age group.

The full set of cost-effectiveness analyses for the model was then reported. Retrospective data for 1100 postmenopausal women diagnosed with early breast cancer in South Australia from 2000-2008 were used to calibrate the model across four mutually exclusive risk groups. Women were divided into four prognostic subgroups based on the Nottingham Prognostic Index score of their primary tumor. For each subgroup, we compared the cost-effectiveness of three different mammographic schedules for two different age groups. Annual mammographic follow-up was not cost-effective for most postmenopausal women. Two yearly mammography was cost-effective for all women with excellent prognosis tumors; and for women with good prognosis tumors if high adherence rates can be achieved. A mixed schedule (annual for 5 years, then 2 yearly thereafter) may be cost-effective for 50-69 year old women with moderate prognosis tumors, and for women aged 70-79 years with poor prognosis tumors. For younger women with poor prognosis tumors, annual mammography is potentially cost-effective. Our results suggest that mammographic follow-up tailored according to risk of recurrence may improve efficiency. That is, result in a reallocation of health care services that improves outcomes across the health care system.

Comparison with the economic model performed as part of the HTA demonstrates that the two early breast cancer surveillance models have significant differences with respect to model structure, population, and analysis. My model had the advantage of using

patient-level data that links surveillance and outcomes for each individual woman. These direct estimates inform a formal calibration process that provides confidence in the realism of the model and that the model results represent the associated uncertainty around the input parameter values. The HTA model represents the natural history in more detail, and facilitates analyses of a broader set of follow-up options, but did not use observed mammography data to populate the model. Although different modelling approaches were used, the results from the two studies are broadly similar, which increases confidence regarding the validity of our conclusions that a one-size-fits-all, annual surveillance strategy may not be the optimal approach to monitoring breast cancer survivors.

Chapter 5 describes the process by which I developed a local patient level dataset to populate and calibrate the health economic model. I describe problems with sourcing, extracting and linking data in order to determine recurrence status for women diagnosed with early breast cancer in South Australia from 2000-2008, given the absence of routinely collected recurrence data. Despite the difficulties experienced, I was able to construct a patient level dataset for 1425 women with early breast cancer. In the dataset, each woman was assigned a prognostic subgroup based on Nottingham Prognostic Index score of her primary tumor at diagnosis, and her individual mammographic surveillance schedule was defined. Recurrence data was extracted from hospital and administrative databases and linked with date and cause of death from the South Australian Cancer Registry for each woman. This patient level dataset was then used to populate and calibrate the health economic model. The methodological limitations of the process and the implications of the poor data quality on interpretation of our modelling results are described (pages 191-192). Issues for future data development are also discussed (pages 192-193).

Question three: What do Australian breast cancer survivors prefer with respect to the provider, location, frequency, and method of delivery of routine follow-up care in years 3, 4 and 5 following diagnosis if existing specialist services were not available; and what is the perceived value of offering "drop-in" clinics providing additional support?

A self-administered questionnaire (on-line or paper) was developed that contained a discrete choice experiment (DCE) designed to explore patient preferences with respect to the above question. The attributes and levels were carefully chosen to reflect the Australian health system and the current policy context whereby alternative modes of delivery of follow-up services are being considered to reduce the burden upon cancer specialist. Women were recruited through breast surgeons in both the public and private health sector (SA only), local and national print media, and state and national patient support and advocacy groups. A total of 836 women participated from across Australia, of whom 722 (86.4%) completed the DCE. Our results demonstrated clear preferences of Australian breast cancer survivors for delivery of their follow-up care in the absence of cancer specialists, in years 3, 4 and 5 following diagnosis. Beyond the first two years post-diagnosis, the 10 most valued surveillance strategy scenarios all included a Breast Physician as the provider of follow-up care. Australian women would prefer to have their routine breast cancer follow-up provided by a Breast Physician (or a Breast Cancer Nurse) in a dedicated local breast cancer clinic, rather than with their local General Practitioner. Drop-in clinics for the management of treatment related side-effects and to provide advice to both develop and maintain good health are also highly valued by breast cancer survivors.

The only other DCE investigating patient preference for breast cancer follow-up was published by Kimman et al in 2010.[112] The DCE assessed patient preferences for the first year of breast cancer follow-up in the Netherlands. The attributes and levels of the Kimman DCE were notably different from those used in our study. They included attendance at an educational group program (yes,no), frequency of visits (3,4,6,12 months), waiting time (5,30,60,90 minutes), contact mode (face-to-face, telephone), and health care provider (medical specialist, breast care nurse/nurse practitioner, general practitioner, breast care nurse and medical specialist). The authors

demonstrated that “overall patient satisfaction would be similar if patients were followed up by a medical specialist alternating with a Breast Cancer Nurse compared to follow-up by a medical specialist only.” Kimman et al also found preference heterogeneity for most attributes “indicating that one strategy does not fit all”.

SYNTHESIS OF RESEARCH FINDINGS

My systematic review identified no randomised controlled trials that estimated the impact on survival of follow-up mammography aimed at early detection of a second cancer event within the breast. The included observational studies were prone to bias and confounding that limit their reliability, and thus do not provide sufficient evidence on which to make informed decisions regarding the development of personalised surveillance strategies.

We developed a discrete event simulation (DES) model to determine if it is efficient to move from a "one size fits all" mammographic follow-up schedule, to one that is tailored according to risk of recurrence. Our results suggest that annual mammographic follow-up was not cost-effective for most postmenopausal women, and that more efficient mammographic schedules based on patient age and Nottingham Prognostic Index score of the primary tumor, could potentially be offered for breast cancer surveillance. The results of our model analysis must however be interpreted with caution. The South Australian data used to populate and calibrate the discrete event simulation model was of poor quality, and it is highly likely that the final study cohort is a biased sample. However, if our results are validated with larger better quality datasets, this work could potentially set the stage for personalised mammographic follow-up after breast cancer.

Given the financial and workforce pressures that reduce the feasibility of the majority of follow-up services being provided by breast cancer specialists, we designed a discrete choice experiment to what sort of follow-up services Australian breast cancer survivors would prefer if we could no longer offer long-term specialist based follow-up. We explored women's preferences for provider, location, frequency and method of delivery of routine follow-up; as well as the perceived value of "drop-in" clinics providing additional support. Our results demonstrate clear preferences of Australian breast cancer survivors for delivery of their follow-up care in the absence of specialist follow-up in years 3, 4 and 5. Breast Physicians were the most preferred provider, followed by Breast Cancer Nurses and the General Practitioners. This suggests that specialised

training in breast medicine is valued by respondents. A follow-up service located in the community had broad appeal to women, and while face-to-face visits were preferred, women would consider alternating face-to-face visits with telephone contact if this meant that the frequency of contact with the follow-up service was not reduced. This study provides important insights into which attributes of a breast cancer follow-up service women value most, and can help inform the design of alternative service pathways that are acceptable to patients, for which further assessment of costs and patient outcomes can be undertaken.

LIMITATIONS OF THIS RESEARCH

The specific methodological limitations of each research project are described in detail in the respective chapter/s. Here I will focus on the general limitations of this research.

Scope

- Data sources: the modelling study was limited to the public sector, as the privacy laws prevented access to pathology and mammography reports from the private sector. It is possible that the results of the modelling may differ, if we had access to data from both the public and private sector in South Australia, and it would be interesting to perform separate analyses of both public and private data.
- Premenopausal women: whilst pathology and imaging data was collected and linked for women aged <50 years, due to time constraints I had to limit my planned modelling analysis. I therefore elected to focus only on the larger group of postmenopausal women.
- Women only: from 2000-2008, there were only 71 documented cases of breast cancer in males in South Australia. This number was too small to consider running separate modelling analyses for females and males.
- Input parameters: targeted searches for relevant literature were performed to identify input parameters for the model. Ideally, full literature reviews to inform each input parameter would be performed.
- Single imaging modality: there is the potential to build a more comprehensive model including incorporating other imaging modalities (e.g. MRI or ultrasound) as an adjunct to mammography, and other frequencies of mammographic surveillance.
- Clinical examination: there is the potential to build a more comprehensive model, with a variety of imaging modalities both with and without clinical examination.
- Treatment effect: it was beyond the scope of the data to examine the effects of varying treatment regimens within the modelling analysis.

Data

Recurrence and follow-up data are not routinely collected in Australia. Chapter 5 describes the numerous difficulties in sourcing, extracting, linking and analysing the local South Australian data. After many failed attempts, accessible data was only identified by early 2010 (the start of my second research year). After careful selection of data points; designing extraction techniques with the data custodians that were feasible in each of the four separate databases and met the ethical requirements of each data custodian; my requests for data were made by the middle of 2010. OACIS data was first extracted Aug 16, 2010, but required two further runs to obtain the data I had requested, due to errors made during the extraction. OACIS data was finally available in late October 2010. After receipt of OACIS data it took in excess of 700 hours of a clinician paired with a typist, to identify and extract the relevant data from each pathology and mammography record for each of the women in the OACIS study cohort (see chapter 5), and enter it into a bespoke database. Following application to the Registrar of Births, Deaths and Marriages, approval was granted to access mortality data from the SA Cancer Registry which was then added to the OACIS data. It took 11 months for the ISAAC team to provide the promised data, which arrived November 2011. It then took a further 3 months to link the OACIS and ISAAC data, a substantial process in its own right. The entire data extraction and linkage process took 2 years and 11 months, before analysis of the patient level dataset could commence.

THE FIELD OF BREAST CANCER FOLLOW-UP OVER THE PERIOD OF THIS WORK

Relevant new research

When I commenced my PHD in 2009, there was no published literature on my proposed research program. However, as described earlier in this thesis, over the course of this work two key papers were published. Most importantly, a commissioned Health Technology Assessment[106] was published in 2011 which included a systematic review and an economic evaluation. In addition, a discrete choice experiment was published by Kimman et al in 2010 examining breast cancer survivors' preferences for follow-up in The Netherlands.[112] In 2012, the Netherlands group published a methodological approach to analyse the cost-effectiveness of customised care, using breast cancer follow-up as a case study. The authors proposed an approach of combining individual-specific parameter estimates from a DCE with cost data, as a means of comparing the costs and benefits of different care programs.[113] They argue that this methodological approach can provide information required by decision makers who are considering investing resources in customised care; and that their approach may help find ways to save costs and increase patient satisfaction.[113] This methodological approach could potentially add value to my analyses, by enabling an extension of my results to inform the costs of customised care in Australia.

Relevant new research in melanoma follow-up has occurred during the period of this work. In 2011, a paper titled "Optimising the frequency of follow-up visits for patients treated for localised primary cutaneous melanoma," was published in the Journal of Clinical Oncology.[114] Kaplan-Meier curves and multivariable Cox proportional hazard models were used to characterise the time course and predictors for melanoma recurrence and new melanoma primaries, based on data from the Melanoma Institute Australia on patients diagnosed with early stage melanoma (stage I or II) between 1985-2009.[114] The authors modelled the delay in diagnosis of recurrence or new primary tumours, as well as the number of monitoring visits, associated with two different monitoring schedules over 10 years of follow-up.[114] Schedule one was based on the 2008 Australian and New Zealand guidelines: follow-up every 6 months for 5 years, then annually for 5 years (stage I disease) or every 3 months for 5 years, then annually for 5 years (stage II disease). Schedule two had fewer visits for both stage I and II disease, but

also provided a separate strategy for stage IIA disease compared to the more advanced stages IIB and IIC: annual follow-up for 10 years (stage I); every 6 months for 2 years, then annually for 8 years (stage IIA); or every 4 months for 2 years, every 6 months during year 3, then annually for 5 years (stages IIB and IIC). The study found only a small difference in modelled delay in diagnosis using schedule two; that AJCC substage was an important predictor of recurrence; and age and date of primary diagnosis were important predictors of developing a new primary.[114] The authors conclude that by providing less frequent monitoring, more efficient follow-up strategies are possible.[114] While the methodology of the study differs to mine, the objective and findings of the study are similar to my work.

Guidelines

At last review in April 2013, there had been no change in the recommendations for breast cancer follow-up by the major international guidelines.[115-120]

FUTURE RESEARCH & RECOMMENDATIONS FOR POLICY

Model extension

Future research could include validating our current model with a larger dataset with high quality pathology on the primary tumor and reliable data on mammographic follow-up. Modelling the frequency of clinical visits would also be required, as this is a core component of international guidelines for breast cancer follow-up.

Our breast cancer model could also potentially be extended in many different directions: age (premenopausal and postmenopausal women, smaller age cohorts); mammographic breast density; primary tumor characteristics (receptor status, histologic subtype, lymphovascular invasion etc); treatment (stratify by type of surgery, radiotherapy or adjuvant systemic therapy), recurrence (model IBTR and CLBC separately), and other factors (eg. BRCA 1&2 mutations, results of Oncotype DX analyses etc).

The methods described in this thesis could readily be applied to optimising follow-up schedules for other cancer types.

Discrete Choice Experiment

Low risk breast cancer follow-up in the community, by breast physicians or breast care nurses, is a pragmatic solution to reducing the burden on hospital based specialist contact.

Potential locations for cancer follow-up services in the community could include large existing general practices or specialist centres, or the GP Superclinics that are currently opening around Adelaide. These large locations could potentially run “drop-in” clinics to address unmet needs for breast cancer survivors. These clinics would also need to provide imaging services (mammography and ultrasound) to ensure that both the follow-up clinic visit and mammogram can occur at the one visit in the one location. This would require an investment in high quality mammography and ultrasound equipment for each community location. The imaging service would ideally be staffed by mammographers, ultrasonographers and radiologists with expertise in the post-treatment breast, to minimise the need for referral back to the tertiary setting due to

suboptimal image acquisition or interpretation. The capacity to offer image guided biopsy on-site, will in part depend on the demand for this service in the low-risk setting.

For Breast Physicians or Breast Care Nurses to assume the role of key clinician in follow-up care, proactive workforce development must occur. This would require funding, willing involvement of the General Practice and Nursing fraternity and appropriate training and remuneration.

Patient expectations would also need to be addressed. Currently women with breast cancer in South Australia, expect to be followed up by a breast cancer specialist (usually a breast surgeon) for the remainder of their lifetime. If existing specialist services were no longer available after completion of primary treatment (or after a predetermined time post-diagnosis), patients must be made aware that this new paradigm is now routine patient care. Education plays a key role. For patients who are newly diagnosed, setting up the expectation from the outset will hopefully diminish difficulties in transitioning from medical cancer specialist to an alternative clinician. Patients whose primary breast cancer is in the remote past, may be reluctant to part with specialist care, and education regarding the safety and effectiveness of proposed alternative models of care would be required.

Pilot programs on alternative models of care, based on the results of the DCE could be performed. Alternatively, a randomised trial of alternative models of care (that involve breast physicians or breast cancer nurses) may be required, before a change in current clinical practice is accepted and funded.

Data collection

In order to conduct rigorous research in breast cancer follow-up (or follow-up of any cancers for that matter), routine capture of stage and recurrence information in population based registries is required. In addition prognostic indicators (eg. hormonal and HER2 receptors) routinely recorded within the cancer registries need to be updated to more accurately reflect contemporary clinical knowledge and practice. From a population perspective, it is currently very difficult to identify recurrence rates, time to recurrence from diagnosis, and mortality trends by stage at diagnosis, and we are

currently unable to quantify the hormonal and HER2 status of breast cancers in Australia. In addition, we need to be able to access imaging reports for cancer surveillance across both public and private sectors. This would enable a more accurate assessment of the effect of our current mammographic intervals, and adherence with mammography, on patient outcomes. Cancer Council Victoria is currently undertaking a research and development project designed to capture stage and recurrence data for population based cancer registries.[121] They are trialling and evaluating the use of natural language processing (NLP) solutions to capture, code, process and store stage and recurrence information from existing information systems, including imaging information systems, to augment information in population based cancer registries. The provision of accurate and high quality information will provide "better evidence for research, policy development and action in cancer prevention, screening and treatment".[121] I would argue that this information could also be used to inform follow-up after completion of primary treatment.

Data access

Three main issues need to be addressed:

1. Private sector data: Data custodians currently invoke privacy laws as a barrier to data access in the private sector. A new framework needs to be established which allows access to patient level data from the private health care system. This could take a variety of forms, but could include instituting an individual "opt-in/opt-out" consent process that grants access to private data for research; or the capacity to obtain de-identified retrospective data from the private health sector where individual patient consent has not been obtained, but an Ethics Committee has assessed the research has potential benefit for the greater public good.
2. Data linkage: If future data collection is to continue in separate "silos", then the capacity for data linkage across the separate databases needs to be enhanced. This will require adequate resourcing to enable timely completion of requests.
3. State-wide health data repository: An alternative to the "silo" approach is to create a single health data "repository" within South Australia. This should include collection of contemporary data points; and strong governance regarding data provision, access and use. This could be a cancer specific database that incorporates health

outcomes post treatment (e.g. recurrence, metastases) and patterns of follow-up care (both clinic visits and diagnostic testing). Ideally "whole of health" database could be developed, that could sit in or be linked with the new electronic medical record (Enterprise Patient Activity System, EPAS) which is currently being implemented across SA Health.

FINAL NOTE

Given the growing pool of breast cancer survivors, there will be an increasing need to determine how we can best provide follow-up services to these women. In 2010 Cancer Australia[116] updated the follow-up chapter of their breast cancer guidelines, which confirmed that the questions that I have tried to address in this PHD are at present unanswered. This highlights that my research is both topical and relevant, and offers an important original contribution to informing the organisation and delivery of breast cancer follow-up services. The methods applied in this thesis could also potentially be used to help optimise follow-up for other forms of cancer treated with curative intent.

APPENDICES

QUESTIONNAIRE

Introduction

Background to the research study

Due to early diagnosis and improved treatment, the number of women who are surviving breast cancer is increasing. This has led to an increasing demand on breast cancer specialists for follow-up. The current workload is not sustainable for cancer specialists (breast surgeon, medical oncologist, radiation oncologist), who also care for women when they are first diagnosed with breast cancer or develop a recurrence.

We are interested in understanding what women who have completed their primary treatment for breast cancer would prefer for follow-up, if we were to design a different service.

We would be grateful if you could take 30 minutes to complete the questionnaire. Before you start, please read the Letter of Invitation and Patient Information Sheet, and then sign the Consent Form.

What does the questionnaire involve?

Section A

In Section A you will be provided with six questions. Each question contains a pair of options. You will need to consider the features of each option and then choose the one that you prefer. Below is a practice example, where you are asked to choose between two different restaurants. You will need to consider the features of restaurant A and restaurant B, and then choose the restaurant you would prefer.

Practice example:

Which restaurant do you prefer?

Restaurant A	Restaurant B
5 km away	10km away
Little variety on menu	Lots of variety on menu
Large serving sizes	Small serving sizes
$30 cost	$15 cost

◯ **Restaurant A**

◯ **Restaurant B**

Section B

In Section B, you will be asked a series of simple questions about you and your care. This section is important because it allows us to understand the range of women participating so that we will know whether the answers we collect reflect the preferences of most women undergoing breast cancer follow-up. All answers will be treated in complete confidence and used for purposes of this research only.

Section A: Discrete Choice Experiment

For the purpose of this study, please **imagine** the following:

Your breast cancer care is provided by a cancer specialist at the hospital for the first 2 years after your diagnosis and you are then discharged to follow-up care with a different clinician in a different location. This medical practitioner would be responsible for your routine follow-up appointments in years 3, 4 and 5, however if a serious problem arose, you would be referred immediately back to the cancer specialist who provided your initial treatment (please note that your follow-up would still continue beyond 5 years, but for the purpose of this study we would like you to focus on years 3,4 and 5). You would be expected to attend your routine follow-up appointments at this new location. This new location also runs weekly "drop-in" clinics for further advice and support about living with breast cancer. These "drop-in" clinics are **in addition** to your routine follow-up appointments, require no booking, and you can use them as little or as often as you wish.

You will be asked to answer 6 questions about **hypothetical** breast cancer follow-up programs. Each question contains a pair of options for you to choose between. The features of the follow-up programs will differ in five ways:

1. Which clinician would provide your care
2. How often your appointments would be
3. Where your follow-up appointment would occur
4. Type of routine follow-up appointment
5. Type of additional drop-in clinics offered

Section A: Discrete Choice Experiment

1. Which clinician would provide your care :
    - **Breast Physician**: most are General Practitioners who have undergone additional training in breast assessment, planning and co-ordinating breast cancer treatment, and counselling.
    - **General Practitioner**: your local doctor who provides person centred, continuing, comprehensive and coordinated whole person health care to individuals and families in their communities. This is your own local doctor you normally seen when you get sick (eg coughs and colds, blood pressure, diabetes, infections etc).
    - **Breast cancer nurse**: a registered nurse who has completed extra study (a Graduate Diploma or higher at University) in the specialty of cancer nursing or its equivalent.

2. How often your appointments would be:
    - **Every 6 months**
    - **Every 9 months**
    - **Every 12 months**

3. Where your follow-up appointment would occur:
    - **Hospital clinic**: a breast cancer follow-up clinic at the hospital where you had your treatment, but you would no longer be seen by the cancer specialist/s who provided your surgery, chemotherapy or radiotherapy.
    - **General Practice**: your local general practice
    - **Local breast cancer follow-up clinic**: a breast cancer follow-up clinic in your local area

Section A: Discrete Choice Experiment

4. Type of routine follow-up appointment:
    - **Face-to-face**: a visit to a clinician where you would have a clinical breast examination, and discussion about your well-being and issues related to your breast cancer care.
    - **Telephone**: at a scheduled time, a clinician telephones you and you have a discussion about your well-being and issues related to your breast cancer care.
    - **Alternate between face-to-face and telephone**: your first appointment is a clinic visit, your second appointment is by telephone, your third appointment is a clinic visits and so on.

    It is important to note that if your care was provided by telephone but it became clear that you needed a face-to-face appointment this would be arranged within one week.

5. Type of additional drop-in clinics offered:
    - **Treatment side-effects clinic**: for management of lymphedema, menopausal symptoms, sexual dysfunction etc that relate to your surgery, chemotherapy or radiotherapy for breast cancer.
    - **Psychosocial support clinic**: for identification, referral and management of depression, anxiety, and issues relating to self-image, family, relationships and return to work
    - **Secondary prevention clinic**: for advice about diet, alcohol and exercise, to both develop and maintain good health and reduce the risk of breast cancer returning.

Section A: Discrete Choice Experiment

1.1 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
Breast Physician	General Practitioner
Every 6 months	Every 12 months
Local breast cancer follow-up clinic	Hospital clinic
Face-to-face	Alternate between face-to-face and telephone
Psychosocial support clinics available	Secondary prevention clinics available

◯ **Program A**

◯ **Program B**

Section A: Discrete Choice Experiment

1.2 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
Breast Physician	General Practitioner
Every 12 months	Every 9months
General practice	Local breast cancer follow-up clinic
Face-to-face	Alternate between face-to-face and telephone
Secondary prevention clinics available	Treatment side-effects clinics available

◯ **Program A**

◯ **Program B**

Section A: Discrete Choice Experiment

1.3 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
General Practitioner	Breast Cancer Nurse
Every 6 months	Every 12 months
Local breast cancer follow-up clinic	Hospital clinic
Telephone	Face-to-face
Treatment side-effects clinics available	Psychosocial support clinics available

◯ **Program A**

◯ **Program B**

Section A: Discrete Choice Experiment

1.4 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
General Practitioner	Breast Cancer Nurse
Every 9 months	Every 6 months
Hospital clinic	General practice
Telephone	Face-to-face
Secondary prevention clinics available	Treatment side-effects clinics available

◯ **Program A**

◯ **Program B**

Section A: Discrete Choice Experiment

1.5 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
Breast Cancer Nurse	Breast Physician
Every 9 months	Every 6 months
Hospital clinic	General practice
Alternate between face-to-face and telephone	Telephone
Psychosocial support clinics available	Secondary prevention clinics available

◯ **Program A**

◯ **Program B**

Section A: Discrete Choice Experiment

1.6 For your routine follow-up appointments in years 3,4 and 5, which program would you prefer?

Program A	Program B
Breast Cancer Nurse	Breast Physician
Every 12 months	Every 9 months
General practice	Local breast cancer follow-up clinic
Alternate between face-to-face and telephone	Telephone
Treatment side-effects clinics available	Psychosocial support clinics available

◯ **Program A**

◯ **Program B**

Section B: About you

In what country were you born?

Do you identify as being of Aboriginal or Torres Strait Islander descent?

- Yes
- No

In which state or territory of Australia were you living when your breast cancer was diagnosed?

What is your current postcode?

What is the highest level of education you have completed?

- Year 11 or lower
- Year 12
- Certificate
- Diploma or advanced diploma
- Degree or higher

What is the most accurate description of your current employment status?

- Employed full time
- Employed part time
- Unemployed
- Retired
- Home duties

Section B: About your family

Family status

- Live alone
- Live alone with children less than 18 years of age
- Live with an adult/s
- Live with an adult/s and children less than 18 years of age

Has any member of your family been diagnosed with breast cancer

- Yes
- No

Is there any member of your family with a known BRCA 1 or BRCA2 genetic mutation?

- Yes
- No
- Not sure

Section B: About your breast cancer diagnosis

How long ago was your breast cancer diagnosed?

- Less than 2 years
- More than 2 years but less than 5 years
- 5 -10 years
- More than 10 years

How old were you when your breast cancer was first diagnosed?

- Less than 40 years
- 40-49 years
- 50-59 years
- 60-69 years
- More than 70 years

What is the most accurate description of your breast cancer when it was diagnosed?

- Non-invasive breast cancer (ductal carcinoma in-situ or lobular carcinoma in- situ)
- Invasive breast cancer contained within the breast that may/may not have spread to lymph nodes in the breast or armpit (early breast cancer)
- Invasive breast cancer that affects the blood vessels in the skin of the breast, causing the breast to become red and inflamed (inflammatory breast cancer)
- Invasive breast cancer that has spread to areas near the breast such as the chest wall (locally advanced breast cancer)
- Invasive breast cancer that has spread from the breast to other parts of the body (secondary, metastatic or advanced breast cancer)
- Not sure

Section B: About your breast cancer treatment

At what type of hospital did you receive treatment?

- Public
- Private
- Both public and private
- Not sure

What treatments did you receive? (circle all that apply)

- Surgery
- Radiotherapy
- Chemotherapy
- Tamoxifen, Faslodex, Arimidex, Femara, Aromasin, Zoladex, Megace, Provera
- Herceptin, Tykerb
- None
- Not sure

Were you told at the end of your treatment, that there was no obvious or visible breast cancer in your body?

- Yes
- No

Are you still experiencing side-effects from your treatment?

- Yes
- No

Section B: About your follow-up

Has your breast cancer returned?

- Yes
- No

Are you currently having treatment for your breast cancer returning?

- Yes
- No

Who do you see <u>most frequently</u> for your breast cancer follow-up?

- Breast surgeon (doctor in charge of your surgery)
- Medical oncologist (doctor in charge of your chemotherapy)
- Radiation oncologist (doctor in charge of your radiotherapy)
- Specialist in training (resident or registrar at the hospital)
- General practitioner (your local doctor)
- Breast nurse
- Other

Do you have a written Survivorship Care Plan?

- Yes
- No – but I know about written Survivorship Care Plans
- No – I don't know about written Survivorship Care Plans

Please select the answer that best describes how much you worry about each of the following:

	Not at all	A little bit	Somewhat	Very much
Future diagnostic tests				
Another type of cancer				
My cancer coming back				
Dying				
My health				
My children's health				

Section B: Any final comments?

Is there anything else important about follow-up care that you think we should know about?

We may conduct a future study where we would talk to some participants. If you would be willing to participate in this kind of research, please provide your contact details below.

Please note that we will only keep your contact details in order to provide you with information about the planned research. It is okay if you decide later to change your mind about being involved.

Thank you so much for participating in this survey.

Please place your completed consent form and questionnaire in the stamped addressed envelope provided and post back to us!

REFERENCES

REFERENCES

References for chapter 1, 2 extension, 4 extension, 5 & 7 only.

All other references included as component of publication chapters

Chapter 1 – Introduction

1. Australian Institute of Health and Welfare & Australasian Association of Cancer Registries 2012. Cancer in Australia: an overview, 2012. Cancer series no. 74. Cat. no. CAN 70. Canberra: AIHW.
2. Jefford M. Overview: Improving outcomes for cancer survivors in Australia. Cancer Forum. 2009. http://cancerforum.org.au/Issues/2009/November/Forum/Overview_Improving_outcomes_for_cancer_survivors_in_Australia.htm (last accessed September 7, 2013)
3. Australian Institute of Health and Welfare & Cancer Australia 2012. Breast cancer in Australia: an overview. Cancer series no. 71. Cat. no. CAN 67. Canberra: AIHW
4. Clemente C. Anatomy: A Regional Atlas of the Human Body. 3rd ed. Munich: Urban & Schwarzenberg; 1987.
5. Last's Anatomy : Regional and Applied. 9th ed. Edinburgh: Churchill Livingstone; 1994.
6. Cancer Australia. Tests for breast cancer. Surry Hills, NSW: Cancer Australia; 2012. http://canceraustralia.gov.au/affected-cancer/cancer-types/breast-cancer/diagnosis/tests-breast-cancer
7. Breast Cancer Network Australia [Internet]. Breast Cancer Pathology: Fact Sheet. Sydney: Breast Cancer Network Australia; c2007 – Version 3, September 2011. http://www.bcna.org.au/store/products/breast-cancer-pathology-fact-sheet (last accessed 26/8/2013)
8. Cancer Australia [Internet]. Clinical practice guidelines for the management of early breast cancer (2nd ed). Surry Hills: Cancer Australia; 2001. http://canceraustralia.gov.au/publications-resources/cancer-australia-publications/clinical-practice-guidelines-management-early (last accessed 26/8/2013)

9. American Joint Committee on Cancer [Internet]. What is Cancer Staging? (7th ed). Chicago: American Joint Committee on Cancer; 2010. https://cancerstaging.org/references-tools/Pages/What-is-Cancer-Staging.aspx
10. International Union Against Cancer. TNM Classification of Malignant Tumors. 7th ed. West Sussex: Wiley-Blackwell; 2010. p181-193
11. Cancer Australia [Internet]. Clinical practice guidelines for the management of advanced breast cancer. Surry Hills: Cancer Australia; 2000. http://canceraustralia.gov.au/publications-resources/cancer-australia-publications/clinical-practice-guidelines-management-advanced (last accessed 26/8/2013)
12. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
13. New Zealand Guidelines Group. Management of early breast cancer. Wellington: New Zealand Guidelines Group; 2009
14. National Institute for Health and Clinical Excellence. Early and locally advanced breast cancer: diagnosis and treatment. National Collaborating Centre for Cancer, Cardiff, Wales, 2009.
15. Buist D, Abraham L, Barlow W, Krishnaraj A, Holdridge R, Sickles E, Carney P, Kerlikowske K, Geller B; Breast Cancer Surveillance Consortium. Diagnosis of second breast cancer events after initial diagnosis of early stage breast cancer. Breast Cancer Res Treat. 2010 Dec; 124(3): 863-873
16. Scottish Intercollegiate Guidelines Network. Management of breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network, Edinburgh, Scotland, 2005.
17. Dixon J, Montgomery D. Extended follow-up of breast cancer patients in clinic wastes time for both patients and doctors: the case for. Breast Cancer Res. 2008; 10 Suppl 4:S7. doi: 10.1186/bcr2167
18. Cancer Australia. Shared Care. Surry Hills, NSW: Cancer Australia; 2013. http://canceraustralia.gov.au/affected-cancer/cancer-types/breast-cancer/life-after-breast-cancer/shared-care (last accessed August 9, 2013)

19. Brennan M, Butow P, Spillane A, Marvin M, Boyle F. Follow up after breast cancer - views of Australian women. Aust Fam Physician. 2011 May; 40(5):311-6.
20. Brennan M, Butow P, Marven M, Spillane A, Boyle F. Survivorship care after breast cancer treatment – experiences and preferences of Australian women. Breast. 2011 Jun; 20(3):271-7.
21. Khatcheressian J, Hurley P, Bantug E, Esserman L, Grunfeld E, Halberg F, Hantel A, Henry N, Muss H, Smith T, Vogel V, Wolff A, Somerfield M, Davidson N; American Society of Clinical Oncology. Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol. 2013 Mar 1; 31(7): 961-5
22. Grunfeld E, Dhesy-Thind S, Levine M; Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment for breast cancer (summary of the 2005 update). CMAJ. 2005 May 10;172(10) :1319-1320
23. Aebi S, Davidson T, Gruber G, Cardoso F; ESMO Guidelines Working Group. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2011 Sep; 22 Suppl 6: vi12-24
24. Grunfeld E, Noorani H, McGahan L, Paszat L, Coyle D, van Walraven C, Joyce J, Sawka C. Surveillance mammography after treatment of primary breast cancer: a systematic review. Breast. 2002 Jun; 11(3): 228-35
25. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes. 3rd ed. New York: Oxford University Press; 2005
26. Paszat L, Sutradhar R, Grunfeld E, Gainford C, Benk V, Bondy S, Coyle D, Holloway C, Sawka C, Shumak R, Vallis K, van Walraven C. Outcomes of surveillance mammography after treatment of primary breast cancer: a population-based case series. Breast Cancer Res Treat. 2009 Mar;114(1):169-78.
27. Buxton M, Drummond M, Van Hout B, Prince R, Sheldon T, Szucs T, Vray M. Modelling in economic evaluation: an unavoidable fact of life. Health Econ. 1997 May-Jun; 6(3) : 217-27

28. Downs S, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998 Jun; 52(6): 377-84
29. Elliott R, Payne K. Essentials of Economic Evaluation in Healthcare. London: Pharmaceutical Press; 2005.
30. Gray A, Clarke P, Wolstenholme J, Wordsworth S. Applied Methods of Cost-effectiveness Analysis in Health Care. New York: Oxford University Press; 2011
31. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes vs discrete event simulation. Health Econ. 2003 Oct; 12(10): 837-48
32. Campbell H, Karnon J, Dowie R. Cost analysis of a hospital-at-home initiative using discrete event simulation. J Health Serv Res Policy. 2001 Jan; 6(1):14-22
33. Davies R, Crabbe D, Roderick P, Goddard J, Raftery J, Patel P. A simulation to evaluate screening for Helicobacter pylori infection in the prevention of peptic ulcers and gastric cancers. Health Care Manag Sci. 2002 Nov; 5(4): 249-58
34. Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Bhandari P, Patel P. The cost-effectiveness of screening for Helicobacter pylori to reduce mortality and morbidity from gastric cancer and peptic ulcer disease: a discrete-event simulation model. Health Technol Assess. 2003; 7(6): 1-86.
35. Roderick P, Davies R, Raftery J, Crabbe D, Pearce R, Patel P, Bhandari P. Cost-effectiveness of population screening for Helicobacter pylori in preventing gastric cancer and peptic ulcer disease, using simulation. J Med Screen. 2003; 10(3): 148-56.
36. Ramsay C, Pickard R, Robertson C, Close A, Vale L, Armstrong N, Barocas D, Eden C, Fraser C, Gurung T, Jenkinson D, Jia X, Lam T, Mowatt G, Neal D, Robinson M, Royle J, Rushton S, Sharma P, Shirley M, Soomro N. Systematic review and economic modelling of the relative clinical benefit and cost-effectiveness of laparoscopic surgery and robotic surgery for removal of the prostate in men with localised prostate cancer. Health Technol Assess. 2012; 16(41): 1-313
37. Hummel S, Stevenson M, Simpson E, Staffurth J. A model of the cost-effectiveness of intensity-modulated radiotherapy in comparison with three-dimensional conformal

radiotherapy for the treatment of localised prostate cancer. Clin Oncol (R Coll Radiol). 2012 Dec; 24(10): e159-67.

38. Willis B, Barton P, Pearmain P, Bryan S, Hyde C. Cervical screening programmes: can automation help? Evidence from systematic reviews, an economic analysis and a simulation modelling exercise applied to the UK. Health Technol Assess. 2005 Mar; 9(13): 1-207, iii.
39. Ahmed Z, Elmekkawy T, Bates S. Developing an efficient scheduling template of a chemotherapy treatment unit: A case study. Australas Med J. 2011; 4(10): 575-88
40. Santibanez P, Aristizabal R, Puterman M, Chow V, Huang W, Kollmannsberger C, Nordin T, Runzer N, Tyldesley S. Operations research methods improve chemotherapy patient appointment scheduling. Jt Comm J Qual Patient Saf. 2012 Dec; 38(12): 541-53.
41. Werker G, Saure A, French J, Shechter S. The use of discrete-event simulation modelling to improve radiation therapy planning processes. Radiother Oncol. 2009 Jul; 92(1): 76-82.
42. Corazza U, Filippini R, Setola R. Discrete event simulation of a proton therapy facility: a case study. Comput Methods Programs Biomed. 2011 Jun; 102(3): 305-16
43. Berg B, Denton B, Nelson H, Balasubramanian H, Rahman A, Bailey A, Lindor K. A discrete-event simulation model to evaluate operational performance of a colonoscopy suite. Med Decis Making. 2010 May-Jun; 30(3): 380-7
44. Coelli F, Ferreira R, Almeida R, Pereira W. Computer simulation and discrete-event models in the analysis of a mammography clinic patient flow. Comput Methods Programs Biomed. 2007 Sep; 87(3): 201-7.
45. Erenay F, Alagoz O, Banerjee R, Cima R. Estimating the unknown parameters of the natural history of metachronous colorectal cancer using discrete-event simulation. Med Decis Making. 2011 Jul-Aug; 31(4): 611-24.
46. Haji Ali Afzali H, Gray J, Karnon J. Model performance evaluation (validation and calibration) in model-based studies of therapeutic interventions in cardiovascular diseases. A review and suggested reporting framework. Appl Health Econ Health Policy. 2013 Apr; 11(2): 85-93

47. Stout N, Rosenberg M, Trentham-Dietz A, Smith M, Robinson S, Fryback D. Retrospective cost-effectiveness analysis of screening mammography. J Natl Cancer Inst. 2006 Jun 7; 98(11): 774-82
48. Tosteson A, Stout N, Fryback D, Acharyya S, Herman B, Hannah L, Pisano E; DMIST Investigators. Cost-effectiveness of digital mammography breast cancer screening. Ann Intern Med. 2008 Jan 1; 148(1): 1-10.
49. Meng Y, Ward S, Cooper K, Harnan S, Wyld L. Cost-effectiveness of MRI and PET imaging for the evaluation of axillary lymph node metastases in early stage breast cancer. Eur J Surg Oncol. 2011 Jan; 37(1): 40-6
50. Karnon J, Brown J; Adjuvant Breast Cancer (ABC) Steering Committee. Tamoxifen plus chemotherapy vs tamoxifen alone as adjuvant therapies for node-positive postmenopausal women with early breast cancer: a stochastic economic evaluation. Pharmacoeconomics. 2002; 20(2): 119-37
51. Vanni T, Karnon J, Madan J, White R, Edmunds W, Foss A, Legood R. Calibrating models in economic evaluation: a seven step approach. Pharmacoeconomics. 2011 Jan: 29(1):35-49.
52. Karnon J, Vanni T. Calibrating models in economic evaluation: a comparison of alternative measures of goodness of fit, parameter search strategies, and convergence criteria. Pharmacoeconomics. 2011 Jan; 29(1):51-62
53. De Bekker-Grob E, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ. 2012 Feb; 21(2):145-72.
54. Ryan M, Gerard K, Amaya-Amaya M. Using Discrete Choice Experiments to Value Health and Health Care. Dordrecht: Springer; 2006.
55. Blinman P, King M, Norman R, Viney R, Stockler M. Preferences for cancer treatments: an overview of methods and applications in oncology. Ann Oncol. 2012 May; 23(5): 1104-1110
56. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. BMJ. 2000 Jun 3; 320(7248):1530-3
57. Ryan M. Discrete choice experiments in health care. BMJ. 2004 Feb 14; 328(7436): 360-1.

58. McFadden D, Train K. Mixed MNL models for discrete response. J Appl Econ. 2000 Sept;15(5):447-70.
59. Hole A. Fitting mixed logit models by using maximum simulated likelihood. Stata J. 2007; 7(3):388-401.
60. Eberth B, Watson V, Ryan M, Hughes J, Barnett G. Does one size fit all? Investigating heterogeneity in men's preferences for benign prostatic hyperplasia treatment using mixed logit analysis. Med Decis Making. 2009 Nov-Dec; 29(6):707-15.
61. Ryan M. Using conjoint analysis to go beyond health outcomes and take account of patient preferences: an application to in vitro fertilisation. Soc Sci Med. 1999 Feb; 48(4): 535-46
62. Ryan M, Hughes J. Using conjoint analysis to assess women's preferences for miscarriage management. Health Econ. 1997 May-Jun; 6(3):261-73
63. Pignone M, Howard K, Brenner A, Crutchfield T, Hawley S, Lewis C, Sheridan S. Comparing 3 techniques for eliciting patient values for decision making about prostate-specific antigen screening: a randomized controlled trial. JAMA Intern Med. 2013 Mar 11; 173(5): 362-8
64. de Bekker-Grob E, Rose J, Donkers B, Essink-Bot M, Bangma C, Steyerberg E. Men's preferences for prostate cancer screening: a discrete choice experiment. Br J Cancer. 2013 Feb 19;108(3):533-41
65. Cheng J, Pullenayegum E, Marshall D, Marshall J, Thabane L. An empirical comparison of methods for analyzing correlated data from a discrete choice survey to elicit patient preference for colorectal cancer screening. BMC Med Res Methodol. 2012 Feb 20; 12:15. doi:10.1186/1471-2288-12-15
66. Hol L, de Bekker-Grob E, van Dam L, Donkers B, Kuipers E, Habbema J, Steyerberg E, van Leerdam M, Essink-Bot M. Preferences for colorectal cancer screening strategies: a discrete choice experiment. Br J Cancer. 2010 Mar 16 ;102(6):972-80
67. Nayaradou M, Berchi C, Dejardin O, Launoy G. Eliciting population preferences for mass colorectal cancer screening organization. Med Decis Making. 2010 Mar-Apr; 30(2): 224-33

68. van Dam L, Hol L, de Bekker-Grob E, Steyerberg E, Kuipers E, Habbema J, Essink-Bot M, van Leerdam M. What determines individuals' preferences for colorectal cancer screening programmes? A discrete choice experiment. Eur J Cancer. 2010 Jan;46(1):150-9

69. Oteng B, Marra F, Lynd L, Ogilvie G, Patrick D, Marra C. Evaluating societal preferences for human papillomavirus vaccine and cervical smear test screening programme. Sex Transm Infect. 2011 Feb; 87(1): 52-7

70. Wordsworth S, Ryan M, Skatun D, Waugh N. Women's preferences for cervical cancer screening: a study using a discrete choice experiment. Int J Technol Assess Health Care. 2006 Summer; 22(3): 344-350

71. Howard K, Salkeld G, Pignone M, Hewett P, Cheung P, Olsen J, Clapton W, Roberts-Thomson I. Preferences for CT colonography and colonoscopy as diagnostic tests for colorectal cancer: a discrete choice experiment. Value Health. 2011 Dec;14(8):1146-52

72. de Bekker-Grob E, Bliemer M, Donkers B, Essink-Bot M, Korfage I, Roobol M, Bangma C, Steyerberg E. Patients' and urologists' preferences for prostate cancer treatment: a discrete choice experiment. Br J Cancer. 2013 Aug 6;109(3):633-40

73. King M, Viney R, Smith D, Hossain I, Street D, Savage E, Fowler S, Berry M, Stockler M, Cozzi P, Stricker P, Ward J, Armstrong B. Survival gains needed to offset persistent adverse treatment effects in localised prostate cancer. Br J Cancer. 2012 Feb 14;106(4):638-45

74. Lloyd A, Penson D, Dewilde S, Kleinman L. Eliciting patient preferences for hormonal therapy options in the treatment of metastatic prostate cancer. Prostate Cancer Prostatic Dis. 2008; 11(2):153-9

75. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. BMJ. 2004 Feb 14; 328(7436):382.

76. Salkeld G, Solomon M, Butow P, Short L. Discrete-choice experiment to measure patient preferences for the surgical management of colorectal cancer. Br J Surg. 2005 Jun;92(6):742-7

77. Tinelli M, Ozolins M, Bath-Hextall F, Williams H. What determines patient preferences for treating low risk basal cell carcinoma when comparing surgery vs imiquimod? A discrete choice experiment survey from the SINS trial. BMC Dermatol. 2012 Oct 4;12:19. doi: 10.1186/1471-5945-12-19
78. Essers B, Dirksen C, Prins M, Neumann H. Assessing the public's preference for surgical treatment of primary basal cell carcinoma: a discrete-choice experiment in the south of the Netherlands. Dermatol Surg. 2010 Dec;36(12):1950-5
79. Sung L, Alibhai S, Ethier M, Teuffel O, Cheng S, Fisman D, Regier D. Discrete choice experiment produced estimates of acceptable risks of therapeutic options in cancer patients with febrile neutropenia. J Clin Epidemiol. 2012 Jun;65(6):627-34
80. Ossa D, Briggs A, McIntosh E, Cowell W, Littlewood T, Sculpher M. Recombinant erythropoietin for chemotherapy-related anaemia: economic value and health-related quality-of-life assessment using direct utility elicitation and discrete choice experiment methods. Pharmacoeconomics. 2007; 25(3):223-37.
81. Sung L, Regier D. Decision making in pediatric oncology: evaluation and incorporation of patient and parent preferences. Pediatr Blood Cancer. 2013 Apr;60(4):558-63
82. Goodall S, King M, Ewing J, Smith N, Kenny P. Preferences for support services among adolescents and young adults with cancer or a blood disorder: a discrete choice experiment. Health Policy. 2012 Oct;107(2-3):304-11.
83. Lancasar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics. 2008; 26(8):661-77.
84. Kimman M, Dellaert B, Boersma L, Lambin P, Dirksen C. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncol. 2010 Apr; 49(3):328-37
85. Cancer Australia. Shared Care. Surry Hills, NSW: Cancer Australia; 2013. http://canceraustralia.gov.au/affected-cancer/cancer-types/breast-cancer/life-after-breast-cancer/shared-care (last accessed August 9, 2013)
86. Brennan M, Jefford M. General Practitioner-based models of post-treatment follow up. Cancer Forum. 2009
http://cancerforum.org.au/Issues/2009/November/Forum/General

practitionerbased models of posttreatment follow up.htm (last accessed August 9, 2013)

87. Jiwa M, Halkett G, Deas K, Ward P, O'Connor M, O'Driscoll C, O'Brien E, Wilson L, Boyle S, Weir J. How do specialist breast nurses help breast cancer patients at follow-up? Collegian 2010; 17(3):143-9.
88. Brennan M, Butow P, Marven M, Spillane A, Boyle F. Survivorship care after breast cancer treatment – experiences and preferences of Australian women. Breast. 2011 Jun; 20(3):271-7.
89. Robertson C, Arcot Ragupathy S, Boachie C, Dixon J, Fraser C, Hernandez R, Heys S, Jack W, Kerr G, Lawrence G, MacLennan G, Maxwell A, McGregor J, Mowatt G, Pinder S, Ternent L, Thomas R, Vale L, Wilson R, Zhu S, Gilber F. The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews registry database analyses and economic evaluation. Health Technol Assess. 2011 Sep; 15(34): v-vi, 1-322.

Chapter 2 extension- comparison with HTA systematic review

90. Robertson C, Arcot Ragupathy S, Boachie C, Dixon J, Fraser C, Hernandez R, Heys S, Jack W, Kerr G, Lawrence G, MacLennan G, Maxwell A, McGregor J, Mowatt G, Pinder S, Ternent L, Thomas R, Vale L, Wilson R, Zhu S, Gilber F. The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews registry database analyses and economic evaluation. Health Technol Assess. 2011 Sep; 15(34): v-vi, 1-322.
91. Churn M, Kelly V. Outpatient follow-up after treatment for early breast cancer: updated results after 5 years. Clin Oncol (R Coll Radiol). 2001; 13 (3): 187-94
92. Montgomery D, Krupa K, Cooke T. Locoregional relapse after breast cancer: most relapses occur late and are not clinically detected. Breast J. 2009 Mar-Apr; 15(2): 163-7

Chapter 4 extension – comparison with HTA economic evaluation

93. Adjuvant! Online website. Adjuvant! Inc., 2008. www.adjuvantonline.com/index.jsp

94. Baum M, Budzar A, Cuzick J, Forbes J, Houghton J, Klijn J, Sahmoud T; ATAC Trialists' Group. Anastrozole alone or in combination with tamoxifen vs tamoxifen alone for adjuvant treatment of postmenopausal women with early breast cancer: first results of the ATAC randomised trial. Lancet. 2002 Jun 22; 359 (9324): 2131-9
95. Breast International Group (BIG) 1-98 Collaborative Group, Thurlimann B, Keshaviah A, Coates A, Mouridsen H, Mauriac L, Forbes J, Paridaens R, Castiglione-Gertsch M, Gelber R, Rabaglio M, Smith I, Wardley A, Price K, Goldhirsch A. A comparison of letrozole and tamoxifen in postmenopausal women with early breast cancer. N Engl J Med. 2005 Dec 29; 353 (26): 2747-57
96. Karnon J. Cost-effectiveness of letrozole, anastrazole and exemestane for early adjuvant breast cancer. Expert Rev Pharmacoecon Outcomes Res. 2007 Apr; 7(2): 143-153
97. Peer P, van Dijck J, Hendriks J, Holland R, Verbeek A. Age-dependent growth rate of primary breast cancer. Cancer. 1993 Jun; 71(11):3547-51
98. Brennan V, Wolowacz S. A systematic review of breast cancer utility weights. ISPOR 13th Annual International Meeting, Toronto, 3-7 May 2008.
99. Jansen S, Stiggelbout A, Wakker P, Vliet Vlieland T, Leer J, Nooy M, Kievit J. Patients' utilities for cancer treatments: a study of the chained procedure for the standard gamble and time trade-off.Med Decis Making. 1998 Oct-Dec; 18(4): 391-9
100. Tosteson A, Stout N, Fryback D, Acharyya S, Herman B, Hannah L, Pisano E; DMIST Investigators. Cost-effectiveness of digital mammography breast cancer screening. Ann Intern Med. 2008 Jan 1; 148(1): 1-10.

Chapter 5- Issues with data access and quality

101. National Breast Cancer Centre. Clinical practice guidelines for the management and support of younger women with breast cancer. National Breast Cancer Centre, Camperdown, NSW, 2004
102. Early Breast Cancer Trialists' Collaborative Group. Ovarian ablation in early breast cancer: overview of the randomised trials. Lancet 1996 Nov 2; 348 (9036): 1189-96
103. Galea M, Blamey R, Elston C, Ellis I. The Nottingham Prognostic Index in primary breast cancer. Breast Cancer Res Treat. 1992; 22(3): 207-219.

104. Brouckaert O, Pintens S, Van Belle V, Van Huffel S, Camerlynck E, Amant F, Leunen K, Smeets A, Berteloot P, Van Limbergen E, Decock J, Hendrickx W, Weltens C, Van den Bogaert W, Vanden Bempt I, Drijkoningen M, Paridaens R, Wildiers H, Vergote I, Christiaens M, Neven P. Short-term outcome of primary operated early breast cancer by hormone and HER-2 receptors. Breast Cancer Res Treat. 2009 May; 115 (2): 349-358.

Chapter 7- Conclusions

105. Downs S, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998 Jun; 52(6): 377-84

106. Robertson C, Arcot Ragupathy S, Boachie C, Dixon J, Fraser C, Hernandez R, Heys S, Jack W, Kerr G, Lawrence G, MacLennan G, Maxwell A, McGregor J, Mowatt G, Pinder S, Ternent L, Thomas R, Vale L, Wilson R, Zhu S, Gilber F. The clinical effectiveness and cost-effectiveness of different surveillance mammography regimens after the treatment for primary breast cancer: systematic reviews registry database analyses and economic evaluation. Health Technol Assess. 2011 Sep; 15(34): v-vi, 1-322.

107. Lash T, Clough-Gorr K, Silliman R. Reduced rates of cancer-related worries and mortality associated with guideline surveillance after breast cancer therapy. Breast Cancer Res Treat. 2005 Jan;89(1):61-7.

108. Lash T, Fox M, Silliman R. Reduced mortality rate associated with annual mammograms after breast cancer therapy. Breast J. 2006 Jan-Feb;12(1):2-6.

109. Paszat L, Sutradhar R, Grunfeld E, Gainford C, Benk V, Bondy S, Coyle D, Holloway C, Sawka C, Shumak R, Vallis K, van Walraven C. Outcomes of surveillance mammography after treatment of primary breast cancer: a population-based case series. Breast Cancer Res Treat. 2009 Mar;114(1):169-78.

110. Montgomery D, Krupa K, Jack W, Kerr G, Kunkler I, Thomas J, Dixon J. Changing pattern of the detection of locoregional relapse in breast cancer: the Edinburgh experience. Br J Cancer. 2007 Jun 18;96(12):1802-7.

111. Elliott R, Payne K. 2005. Essentials of Economic Evaluation in Healthcare. London: Pharmaceutical Press; 2005.

112. Kimman M, Dellaert B, Boersma L, Lambin P, Dirksen C. Follow-up after treatment for breast cancer: one strategy fits all? An investigation of patient preferences using a discrete choice experiment. Acta Oncol. 2010 Apr; 49(3):328-37
113. Benning T, Kimman M, Dirksen C, Boersma L, Dellaert B. Combining individual-level discrete choice experiment estimates and costs to inform health care management decisions about customized care: the case of follow-up strategies after breast cancer treatment. Value Health. 2012 Jul-Aug; 15(5): 680-9
114. Turner R, Bell K, Morton R, Hayen A, Francken A, Howard K, Armstrong B, Thompson J, Irwig L. Optimizing the frequency of follow-up visits for patients treated for localized primary cutaneous melanoma. J Clin Oncol. 2011 Dec 10: 29(35):4641-6
115. National Institute for Health and Clinical Excellence. Early and locally advanced breast cancer: diagnosis and treatment. National Collaborating Centre for Cancer, Cardiff, Wales, 2009.
116. National Breast and Ovarian Cancer Centre. Recommendations for follow-up of women with early breast cancer. National Breast and Ovarian Cancer Centre, Surry Hills, NSW, 2010.
117. Khatcheressian J, Hurley P, Bantug E, Esserman L, Grunfeld E, Halberg F, Hantel A, Henry N, Muss H, Smith T, Vogel V, Wolff A, Somerfield M, Davidson N; American Society of Clinical Oncology. Breast cancer follow-up and management after primary treatment: American Society of Clinical Oncology clinical practice guideline update. J Clin Oncol. 2013 Mar 1; 31(7): 961-5
118. Aebi S, Davidson T, Gruber G, Cardoso F; ESMO Guidelines Working Group. Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2011 Sep; 22 Suppl 6: vi12-24
119. Scottish Intercollegiate Guidelines Network. Management of breast cancer in women. A national clinical guideline. Scottish Intercollegiate Guidelines Network, Edinburgh, Scotland, 2005.
120. Grunfeld E, Dhesy-Thind S, Levine M; Steering Committee on Clinical Practice Guidelines for the Care and Treatment of Breast Cancer. Clinical practice guidelines for the care and treatment of breast cancer: follow-up after treatment

for breast cancer (summary of the 2005 update). CMAJ. 2005 May 10; 172(10) :1319-1320

121. Cancer Council Victoria [Internet]. Capturing cancer stage & recurrence for population-based cancer registries. SUBURB: Cancer Council Victoria; YEAR. www.cancervic.org.au/about-our-research/registry-statistics/capture-stage-recurrence (last accessed 26/8/2013)