

Epidemiology of Ductal Carcinoma in Situ

Gurdeep S. Mannu

Green Templeton College, University of Oxford

Doctor of Philosophy
Michaelmas Term 2017



Nuffield Department of Population Health,
Richard Doll Building,
Old Road Campus,
Oxford
OX3 7LF


National Institute for
Health Research



The author was supported by the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) fellowship, the Nuffield Department of Population and by Cancer Research UK.

**Life is short,
and the art long,
opportunity fleeting,
experimentations perilous,
and judgment difficult.**

Hippocrates

Epidemiology of Ductal Carcinoma in Situ

Gurdeep S. Mannu, Green Templeton College, Doctor of Philosophy, Michaelmas 2017.

Abstract

Introduction:

Almost 7,000 people are diagnosed with ductal carcinoma in situ (DCIS) in the United Kingdom each year, but there remains uncertainty regarding its natural history and optimal management. The aim of this thesis was to evaluate factors contributing to the epidemiology of DCIS and its outcomes.

Methods:

1) A cohort study comparing risk factors for DCIS and invasive breast cancer (IBC) using UK Biobank; 2) A cohort study examining the accuracy of preoperative biopsy in DCIS using clinical records from the Netherlands Cancer Institute; 3) A cohort study examining the rate of invasive breast cancer following treatment for screen-detected DCIS in England using the National Health Service Breast Screening Programme (NHSBSP) audit; 4) A methodological study to develop an algorithm to identify invasive breast cancer recurrences, which in the future may be used to identify DCIS recurrences, using all relevant routinely collected data stored within Public Health England (PHE).

Results:

(1) For both DCIS and IBC, postmenopausal BMI was associated with an increased risk of developing disease, and the number of live births was associated with a decreased risk of developing disease. However, the magnitude of the effect differed between DCIS and IBC. The increased risk from postmenopausal BMI ≥ 35 kg/m² was larger for DCIS than for IBC (RR 2.35, 95% CI 1.14 – 4.82), and the trend of reduction in risk with each additional live birth was greater for DCIS than for IBC (p for trend = 0.03).

(2) Consideration of mammographic lesion size and the absence of necrosis on biopsy may be helpful in selecting low-risk women for non-operative management of DCIS in the future, as may use of the 9G vacuum-assisted method of biopsy.

(3) The cumulative risks of IBC at 5, 10 and 15 years after screen-detected DCIS in England were 3.5%, 7.1%, and 9.4% respectively. Women with clear surgical margins of 1 – 2 mm had a higher IBC rate than women with clear margins of 5+ mm (RR 1.85, 95% CI 1.20 – 2.84). Women given breast-conserving surgery (BCS) without radiotherapy had a higher ipsilateral IBC rate than women given BCS with radiotherapy (RR 1.63, 95% CI 1.27 – 2.10). Women given hormone therapy had a lower rate of any IBC compared with oestrogen receptor (ER) positive women not given hormone therapy (RR 0.76, 95% CI 0.63 – 0.93).

(4) There was good agreement between the number of recurrences indicated by the developed algorithm using routinely collected data sources and the number of recurrences recorded in the test dataset. This finding supports the potential value of compiling recurrence information on a nationwide basis from routinely collected data, for use in future descriptive and epidemiological studies and in follow-up for randomised trials.

Conclusions:

Using a variety of methods these studies have all succeeded in adding to knowledge about the epidemiology of DCIS (see Figure A.1). This knowledge can be used to help the future

management of women with DCIS. In addition, each of the studies has planned extensions and will continue to contribute further knowledge periodically into the future.

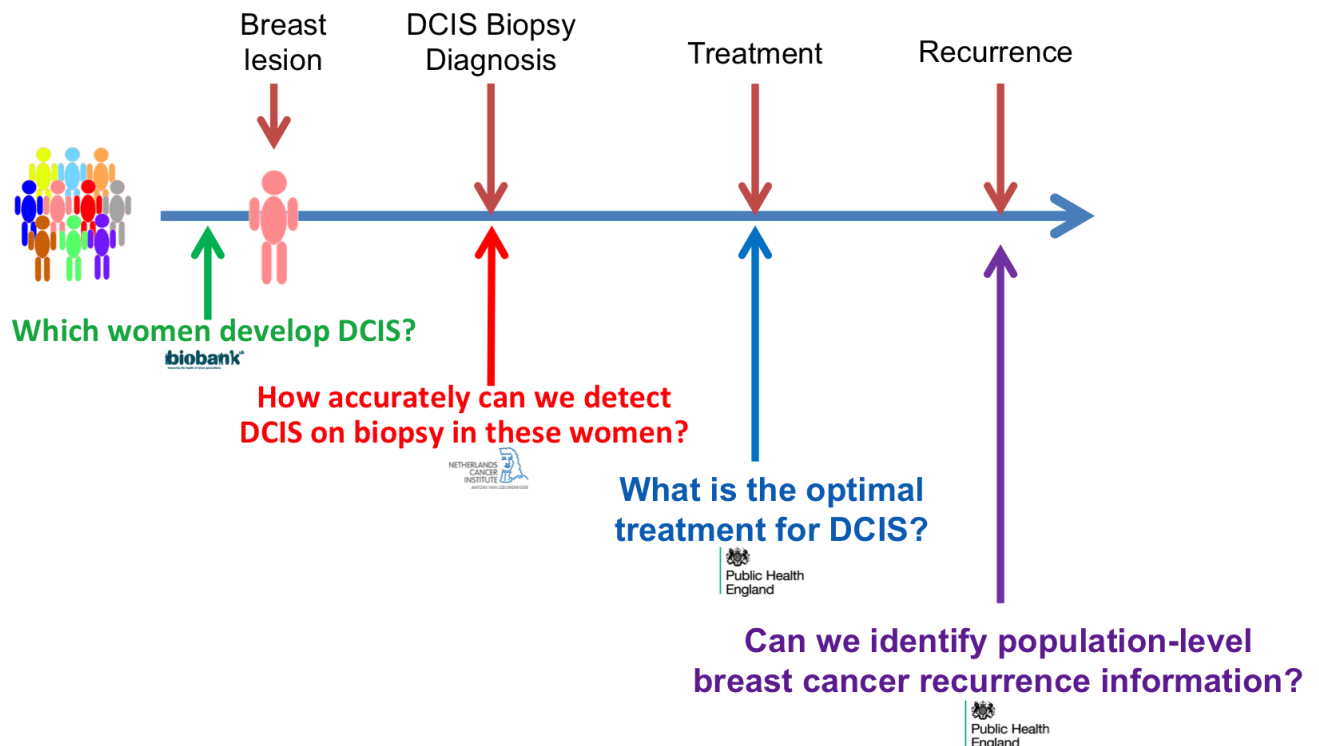


Figure A.1: Schematic diagram summarising the medical pathway of patients with DCIS (top row), and the related areas of clinical uncertainty and corresponding work-streams contributing to this thesis (bottom row).

Contents

Abstract	i
Contents	iii
List of Figures	vi
List of Tables	viii
Abbreviations	ix
Research outputs during this work.....	x
Acknowledgements.....	xiii
Dedication	xiv
<u>Chapter 1: Contemporary uncertainties in the epidemiology and management of Ductal Carcinoma in Situ</u>	<u>1</u>
1.1. Introduction.....	2
1.2. Methods.....	4
1.3. The natural history of DCIS.....	4
1.4. National Breast Screening and the evolving incidence of DCIS	7
1.5. Theories of disease progression in DCIS	12
1.6. Risk factors for DCIS incidence and recurrence.....	16
1.7. Genetic risk factors for DCIS occurrence	17
1.8. Biomarkers of prognosis in DCIS.....	22
1.9. Developing prognostic indices for women with DCIS.....	23
1.10. Contemporary treatment strategies for DCIS	25
1.10.1. Breast Surgery.....	25
1.10.2 Core-Biopsy in DCIS.....	27
1.10.3 Axillary staging.....	28
1.10.4 Adjuvant radiotherapy	29
1.10.5 Adjuvant hormonal therapy	30
1.11. Conclusion	31
1.12. References.....	32
<u>Chapter 2: Motivation and Aims.....</u>	<u>45</u>
2.1. Motivation.....	46
2.2. Aims.....	46
<u>Chapter 3: Differences in Risk Factors for Ductal Carcinoma in Situ and Invasive Breast Cancer: Prospective Cohort Study</u>	<u>49</u>
3.1. Background	50
3.2. Methods.....	51
3.2.1 Participants.....	51
3.2.2 Variables	52
3.2.3 Statistical methods	52
3.3. Results.....	54
3.3.1 Risk factors for incident IBC	54
3.3.2 Risk factors for incident DCIS.....	58
3.3.3 Comparison of IBC and DCIS	61
3.4. Discussion	64
3.4.1 Body mass index	64

3.4.2 Number of live births	65
3.4.3 Dietary factors.....	66
3.4.4 Limitations	66
3.4.5 Conclusion	67
3.5. References.....	68
<u>Chapter 4: The accuracy of preoperative biopsy in Ductal Carcinoma in Situ</u>	<u>70</u>
4.1. Background	71
4.2. Methods.....	73
4.2.1 Data sources	73
4.2.2 Participants.....	73
4.2.3 Variables	73
4.2.4 Statistical methods	74
4.3. Results.....	75
4.4. Discussion	80
4.4.1 Conclusion	82
4.5. References.....	83
<u>Chapter 5: Population-based Study of Invasive Breast Cancer in Women Diagnosed with DCIS</u>	<u>84</u>
5.1. Background	85
5.2. Methods.....	87
5.2.1 Introduction to data sources	87
5.2.2 Participants.....	87
5.2.3 Variables	88
5.2.4 Statistical methods	88
5.3. Results.....	90
5.3.1 Risk of any IBC or breast cancer death by time since diagnosis	91
5.3.2 Risk of IBC in women with unilateral DCIS who received surgery.....	92
5.3.3 Patient and tumour related factors and iIBC rate.....	94
5.3.4 Treatment related factors and IBC rate	99
5.4. Discussion	101
5.4.1 Calendar year	101
5.4.2 Tumour related factors and iIBC rate	102
5.4.3 Margin distance.....	103
5.4.4 Radiotherapy and DCIS	104
5.4.5 Hormone therapy and DCIS.....	105
5.4.6 Limitations	106
5.4.7 Conclusion	107
5.5. References.....	108
<u>Chapter 6: The Use of Routine Data Sources to Identify Recurrence in Women Registered with Invasive Breast Cancer in England</u>	<u>111</u>
6.1. Introduction.....	112
6.2. Methods.....	114
6.2.1 Collating all relevant routinely collected datasets	114
6.2.2 Development of the algorithm	114
6.2.3 Identifying a population in which to test the algorithm	115
6.2.4 Inclusion criteria	115
6.2.5 Exclusion criteria	116
6.2.6 Application of the algorithm to the test population	116
6.2.7 Testing the algorithm using the GRACE data	117

6.2.8 Unexpected challenges in conducting the study	118
6.3. Results.....	120
6.3.1 Recurrence results from the algorithm executed using routinely collected data .	120
6.3.2 Evaluation of algorithm using the GRACE data.....	122
6.3.3 Evaluation of recurrence dates identified by the algorithm using the GRACE data	127
6.4. Discussion	130
6.4.1 The growing need for information on recurrence	130
6.4.2 Relation between local recurrence and subsequent breast cancer death.....	131
6.4.3 Validity of the algorithm.....	132
6.4.4 Conclusion	133
6.5. References.....	134
<u>Chapter 7: Discussion and Conclusion</u>	<u>135</u>
7.1. Discussion	136
7.2. Conclusions.....	141
7.3. References.....	143
<u>Appendices.....</u>	<u>144-182</u>

List of Figures

Abstract

Figure A.1: Schematic diagram summarising the medical pathway of patients with DCIS (top row), and the related areas of clinical uncertainty and corresponding work-streams contributing to this thesis (bottom row).....	ii
--	----

Chapter 1

Figure 1.1: Schematic diagram illustrating the pathology of ductal carcinoma-in-Situ. Copyright © Gurdeep S. Mannu 2016.....	3
Figure 1.2: Incidence of DCIS by age. Incidence rates of ductal carcinoma in situ (DCIS) by age group per 100,000 women in Great Britain between 1979-2010.....	8
Figure 1.3: The main theories of disease progression in ductal carcinoma in situ (DCIS).	15

Chapter 3

Figure 3.1: The association between different factors and the incidence of invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following basic stratification	55
Figure 3.2: The association between different factors on the incidence of invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following full stratification	56
Figure 3.3: The association between different factors on the incidence of ductal carcinoma in situ, as first cancer, in women enrolled into UK Biobank, following basic stratification	59
Figure 3.4: The association between different factors on the incidence of ductal carcinoma in situ, as first cancer, in women enrolled into UK Biobank, following full stratification.....	60
Figure 3.5: The association between different factors on the incidence of ductal carcinoma in situ compared with invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following basic stratification.	62
Figure 3.6: The association between different factors on the incidence of ductal carcinoma in situ compared with invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following full stratification.....	63

Chapter 4

Figure 4.1: Derivation of the population used in the analysis.	75
--	----

Chapter 5

Figure 5.1: Flow diagram showing the selection of participants for inclusion in the study.	90
Figure 5.2: Cumulative risk of any (ipsilateral or contralateral) IBC and breast cancer death by time since diagnosis of DCIS.	91

Figure 5.3: Ipsilateral invasive breast cancer after screen-detected DCIS by different patient, tumour and treatment factors, in women who did not receive hormone therapy. No adjustment applied.	95
Figure 5.4: Ipsilateral invasive breast cancer after screen-detected DCIS by different patient, tumour and treatment factors, in women who did not receive hormone therapy. Each variable is adjusted by all other variables in the categories shown....	96
Figure 5.5: Comparison of cumulative risk of ipsilateral invasive breast cancer between high and low/intermediate grade DCIS, by time since DCIS diagnosis. Adjusted for surgery and radiotherapy.....	98
Figure 5.6: Hormone therapy and any (i.e. ipsilateral or contralateral) invasive breast cancer after DCIS.....	100

Chapter 6

Figure 6.1: Comparison of cumulative risk of any recurrence by time since cancer registration identified by the algorithm using the routinely collected data (red) and the GRACE data (green)	127
Figure 6.2: Comparison of recurrence hazard rate by time since cancer registration identified by the algorithm using the routinely collected data (red) and the GRACE data (green).....	128
Figure 6.3: Comparison of the recurrence rate of different types of recurrence, by time since cancer registration between routinely collected data and the GRACE data	129

List of Tables

Chapter 1

Table 1.1: Demonstrating all available data on the natural history of Ductal Carcinoma in Situ (DCIS).....	5
Table 1.2: Comparison of the risk factors for developing DCIS and IBC.....	16
Table 1.3: Factors associated with a poor prognosis following treatment for DCIS.....	17
Table 1.4: GWAS studies studying genetic risk factors for invasive breast cancer.	20-21
Table 1.5: Biomarkers for the presence of IBC in biopsies showing DCIS..	22
Table 1.6: Demonstrating classification systems commonly used in DCIS.	24
Table 1.7: Showing all randomized trials comparing breast surgery alone with breast surgery and radiotherapy for DCIS.....	30

Chapter 4

Table 4.1: Characteristics of women diagnosed with ductal carcinoma in situ on initial biopsy.....	76
Table 4.2: Final diagnoses in women with ductal carcinoma in situ on biopsy by various characteristics.....	78

Chapter 5

Table 5.1: Participant characteristics for all women with unilateral screen-detected DCIS treated with surgery and recorded in the NHSBSP audit (2000 – 2015).	93
--	----

Chapter 6

Table 6.1: Number of women classified according to first recorded outcome event using information from the routinely collected data only, without the addition of the GRACE dataset. Women with an outcome event <3 months from date of diagnosis are excluded..	121
Table 6.2: Cross Tabulation of the numbers of women with non-breast cancer related first events, any recurrences and no recorded recurrences recorded in the GRACE data, with the same categories identified by the algorithm using routinely collected data.	123
Table 6.3: Cross Tabulation of the numbers of women with different types of recurrences, who were removed from follow-up or who did not have a recurrence, recorded in the GRACE data, with the same categories identified by the algorithm using routinely collected data.	126

Abbreviations

ABS	Association of Breast Surgery
BCS	Breast Conserving Surgery
Bx	Biopsy
CT	Chemotherapy
DCIS	Ductal Carcinoma In Situ
DFI	Disease Free Interval
DFS	Disease Free Survival - time without disease prior to relapse or last follow-up
Dx	Diagnosis
EBM	Evidence-Based Medicine
EFS	Event Free Survival - time from diagnosis to defined events (eg relapse or death)
FNA	Fine Needle Aspiration - a type of biopsy using a thin needle (or FNAB)
FU	Follow up
GCP	Good Clinical Practice (guidelines)
Gy	Grays (units of radiation)
H&E	Hematoxylin and Eosin (stain)
HRT	Hormone replacement therapy
IBC	Invasive Breast Cancer
ICD	International Classification of Diseases (coding system)
IDC	Invasive Ductal Carcinoma
ILC	Invasive Lobular Carcinoma
LCIS	Lobular Carcinoma In Situ - type of breast cancer
LN	Lymph Node
Lx	Lumpectomy
Mx	Mastectomy
NAD	No Abnormality Detected
NK	Not known
OS	Overall Survival
PD	Progressive disease
PFS	Progression Free Survival
PHE	Public Health England
QALY	Quality-Adjusted Life Year
RCT	Randomised Controlled Trial
SLNB	Sentinel Lymph Node Biopsy
TNM	Staging system - primary tumour
US	Ultrasound (scan)
WHO	World Health Organisation

Research outputs during this work

Publications during DPhil:

Mannu GS and Darby S. Cancer Epidemiology, *Textbook of Surgical Oncology*, Second Edition, CRC publishing, edited by Graeme Poston, Lynda Wyld and Riccardo A. Audisio. 2016, p. 3-16.

Mannu GS, Dodwell D, Making progress in early breast cancer: Taking time or accepting risk? *Eur J Cancer*. 2017, 12;78:24-27.

McGale P, Dodwell D, Gray R, Mannu GS, Peto R, Taylor C, Wang Z, Wang Y, Bergh J, Pritchard KI, Clarke M, Davies C, Bradley R, Braybrooke J, Pan H, Gianni L, Lohrisch C, Anderson S, von Minckwitz G, on behalf of Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Neo-adjuvant versus adjuvant chemotherapy in early breast cancer: patient-level meta-analysis of long-term outcomes among 4756 women in 10 randomised trials. *Lancet Oncology*. 2018, 19;1:27-39.

National and International Presentations during DPhil:

Mannu GS, Groen E, Wang Z, Schaapveld M, Lips E, Chung M, Joore I, van Leeuwen F, Teerstra J, Winter-Warnars GAO, Darby SC, Wesseling J. 'Risk Factors for Upgrading and Upstaging of Pre-operative Biopsies in Ductal Carcinoma in Situ', San Antonio Breast Cancer Symposium, USA, 5-9 December 2017.

Mannu GS, Wang Z, Cheung S, Kearins O, Broggio J, Charman J, Darby SC. 'Population-based Study of Factors Influencing Invasive Breast Cancer Risk after Screen-detected Ductal Carcinoma in Situ: First results from the Non-Invasive Breast Cancer in England (NINBE) Study', the National Cancer Research Institute (NCRI) Cancer Conference in Liverpool, 5-8 November 2017.

Mannu GS, McGale P, Wang Z, Darby SC. 'Differences in Risk Factors for Ductal Carcinoma in Situ and Invasive Breast Cancer: A Prospective Cohort Study', the National Cancer Research Institute (NCRI) Cancer Conference in Liverpool, 5-8 November 2017.

Mannu GS, Dodwell D, McGale P, Taylor C. 'Trials of surgery versus not after sentinel node dissection', the Early Breast Cancer Trialists' Collaborative Group Steering Committee meeting, Said Business School, Park End Street, Oxford, OX1 1HP, 18-19 September 2017.

Mannu GS, Wang Z, Broggio J, Charman J, Kearins O, Cheung S, Darby SC. 'Population-based Study of Long-term Outcomes in Women Diagnosed with Ductal Carcinoma in Situ of the Breast (DCIS)', Public Health England National Cancer Registration and Analysis Service (NCRAS) and United Kingdom and Ireland Association of Cancer Registries (UKIACR) Cancer Data and Outcomes Conference 2017: Using data to drive services at the University of Manchester, 13-14 June 2017.

Mannu GS, Wang Z, Broggio J, Charman J, Kearins O, Cheung S, Darby SC. ‘Nationwide Study of Outcomes after DCIS, 1971-2014’, Association of Breast Surgeons Conference at Belfast Waterfront, 2 Lanyon Pl, Belfast BT1 3WH, 15-16 May 2017.

Mannu GS, Broggio J, Charman J, Darby SC. ‘Identifying progressive cancers using routine health datasets in England’, World Cancer Congress, Paris, France, 31 October-3 November 2016.

Mannu GS, Broggio J, Charman J, Watters A, Darby SC. ‘Using routinely collected data in England to identify recurrences in breast cancer patients’, Public Health England National Cancer Registration and Analysis Service (NCRAS) and United Kingdom and Ireland Association of Cancer Registries (UKIACR) Cancer Data and Outcomes Conference 2016: Using data to drive services at the University of Manchester, 13-14 June 2016.

Mannu GS, Broggio J, Charman J, Darby SC. ‘Identifying recurrence in breast cancer patients from routinely collected data in England’, Association of Breast Surgeons Conference at Manchester Central, Windmill St, Manchester, 16-17 May 2016.

Prizes during DPhil:

BJS Proffered Prize Paper Award at the National Cancer Research Institute (NCRI) Cancer Conference in Liverpool, L3 4BX, 6 November 2017.

Nuffield Department of Population Health DPhil Poster Prize, The Big Data Institute (BDI), Old Road Campus, Oxford, OX3 7LF, 13 October 2017.

Award for leading the winning team at the Seventh National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) Infrastructure Doctoral Research Training Camp, Ashridge Business School, Ashridge, Berkhamsted, HP4 1NS, 8 July 2016.

Green Templeton College Nautilus Award for Academic Achievement, University of Oxford, 43 Woodstock Road, Oxford, OX2 6HG, 27 June 2015.

Research Funding/Grants during DPhil:

The Royal College of Surgeons of Edinburgh (RCSEd) Small Research Pump Priming Grants, 1st April 2015, RCSEd, Nicolson Street, Edinburgh, EH8 9DW, £2,700.

Nuffield Department of Population Health Pump-Priming Grant for Feasibility study of High Intensity Focused Ultrasound (HIFU) in the treatment of Breast cancer (The HIFUB study) award ID: H6D00410-H601.06, 24th March 2017, University of Oxford, Richard Doll Building, Old Road Campus, Oxford, OX3 7LF. £41,097.

Chief Investigator (CI) or Principle Investigator (PI) experience during DPhil:

“Study to investigate the accuracy with which breast cancer recurrence can be identified in women registered with invasive breast cancer using routinely collected data compared with recurrence information collected by the AZURE trial”, IRAS project ID: 232895, Research ethics committee approval reference: 17/WM/0347, 15 September 2017.

“Population based study of outcomes and long-term trends in women diagnosed with non-invasive breast cancer”, IRAS project ID: 192243, Research ethics committee approval reference: 16/YH/0209, 18 May 2016.

‘Study investigating the lifestyle and biological risk factors for developing Ductal Carcinoma in Situ (DCIS) and its progression to invasive breast cancer in women’, CI of UK Biobank Application Number 11091, 10 July 2015.

Acknowledgements

I would like to thank my supervisors, Sarah Darby and Freddie Hamdy for their guidance and support during this work. I would like to thank Zhe Wang, Paul McGale and David Kennedy for their statistical and programming support.

I would also to thank my collaborators from Public Health England (John Broggio and Jackie Charman, Olive Kearins and Shan Cheung) and the Netherlands Cancer Institute (Jelle Wessling, Michael Schaapveld, Esther Lips and Emilie Groen), without whom, this work would not be possible.

I would like to extend my gratitude to the National Institute for Health Research (NIHR) Biomedical Research Centre (BRC) fellowship for funding the first two years of my DPhil, to the Nuffield Department of Population Health for funding my third year, to Cancer Research UK for funding my final year and to the Royal College of Surgeons of Edinburgh Small Research Pump Priming Grants committee for funding my application to UK Biobank.

I would like to thank SK Farid Ahmed and Giles Cunnick for the experience I have gained in their breast surgery clinics. These clinics gave me the opportunity to speak with patients newly diagnosed with DCIS, who provided my personal motivation for this research.

I am grateful for the moral support I received from my friend, Joao H Bettencourt-Silva, which was invaluable throughout this work.

I would also like to thank my parents, Piara S Mannu and Gurbax K Mannu, for their backing and my brother, Kuldeep S. Mannu, for his steadfast support.

Last but not least, I would like to thank Angeline and Simran, for their understanding and encouragement during this time.

Dedication

I would like to dedicate this thesis to my wife, Angeline and to my daughter, Simran.

Chapter 1: Contemporary uncertainties in the epidemiology and management of Ductal Carcinoma in Situ

1.1. Introduction

Ductal carcinoma in situ (DCIS) is amongst the earliest forms of breast cancer. It is a heterogeneous entity, encompassing a broad spectrum of disease but it is characterised by the proliferation of abnormal epithelial cells lining the milk ducts within the breast (Figure 1.1) (1). These abnormal cells do not yet have the capacity to invade outside their basement membrane into the surrounding stroma so DCIS is, by definition, a non-invasive cancer (2). However, if left untreated, it can progress to invasive breast cancer (IBC) over a number of years. Since not all women with DCIS will develop IBC within their lifetime, it is sometimes considered as a non-obligate precursor of IBC and it is not yet clear which women with DCIS will develop IBC and over what period of time (1). In light of this uncertainty, most women with DCIS are treated with surgical excision to minimise their risk of developing IBC, even though some may be inadvertently over-treated as a result. However, even after adequate treatment of DCIS, some women will develop a recurrence of DCIS or a local occurrence of IBC within the same breast. The reason for this is not entirely clear and although it is not yet possible to accurately predict which women will suffer a recurrence or subsequent IBC, there is currently a great deal of interest in developing predictive indices to help to identify those at increased risk.

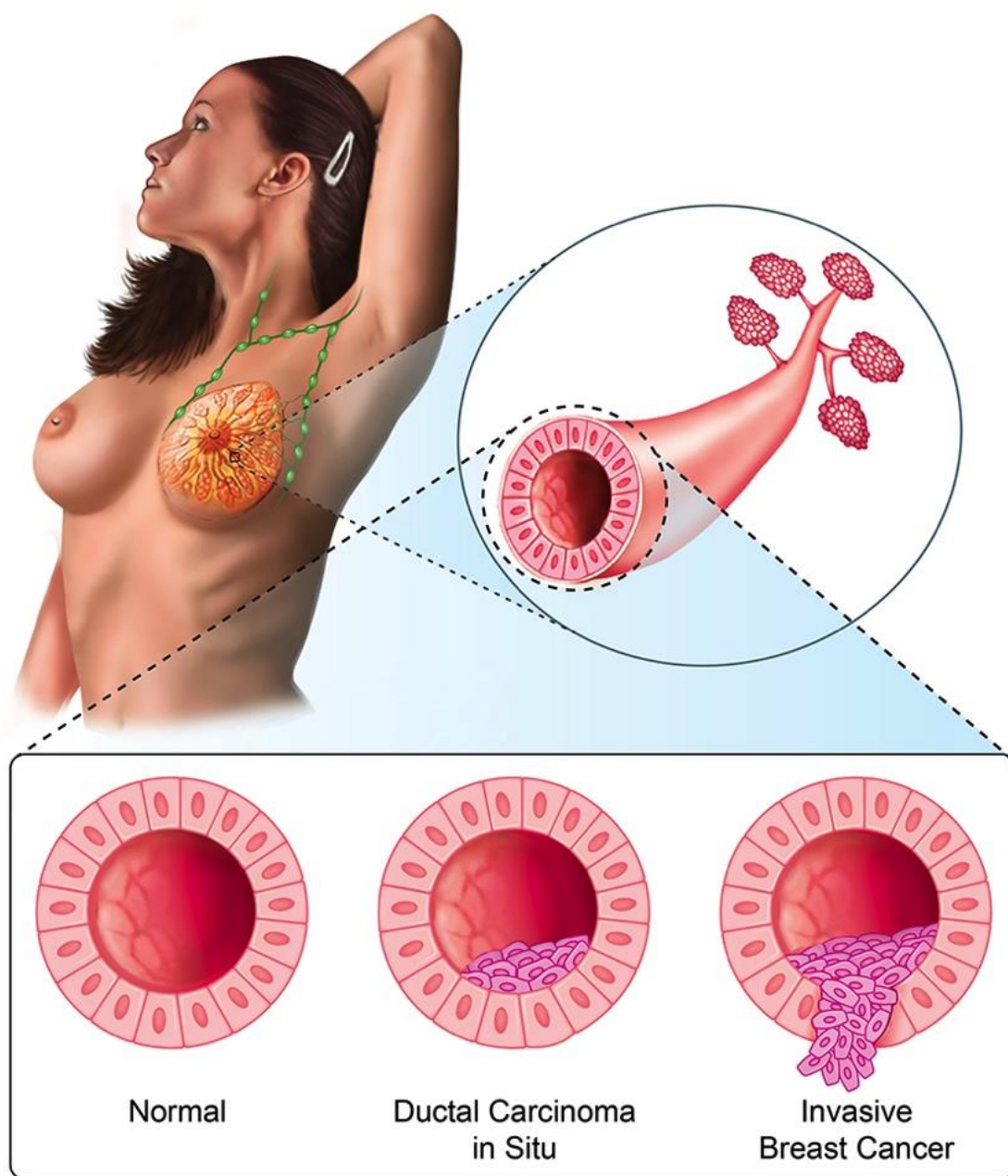


Figure 1.1: Schematic diagram illustrating the pathology of ductal carcinoma-in-Situ. Copyright © Gurdeep S. Mannu 2016.

1.2. Methods

This chapter is a systematic review of the published literature regarding the epidemiology and management of DCIS. The purpose was to identify areas of uncertainty for further research. In October 2014, a systematic search was undertaken using PubMed/Medline MeSH terms: ("Carcinoma, Intraductal, Noninfiltrating"[Mesh]) AND "Prognosis"[Mesh]. The abstracts of the resultant 2064 papers were reviewed and 210 relevant papers were retrieved for detailed study. The reference lists of relevant papers were further reviewed to identify additional pertinent publications. Articles published since this date were identified and screened prospectively using weekly Google Scholar alerts for terms: 'DCIS' and 'ductal carcinoma in situ', over the three-year period between 1/12/2014 – 1/12/2017. The findings from this search are summarised in the following sections of this chapter.

1.3. The natural history of DCIS

The management of DCIS has been the source of uncertainty ever since its malignant potential was first described in the early part of the twentieth century (3). Prior to the introduction of mammographic screening programmes, DCIS was detected symptomatically, such as by a palpable lump or nipple discharge. In this setting, cases of DCIS were often larger, higher grade, more extensive and more often accompanied at the same time by IBC (4, 5). The majority of women with DCIS were offered primary surgery and the roles of margin clearance and adjuvant therapy were less well developed during this time. In exceptional circumstances women did not receive surgical treatment and the current published literature on the natural history of DCIS arises from sequential follow-up of these women (Table 1.1). Although case-series of women with DCIS receiving non-operative management confirm the high risk of IBC developing after DCIS, underlying poor health, advanced age or other

factors that were responsible for the non-operative approach to treatment mean that the observed outcomes from these women are unlikely to represent the expected outcomes from the wider population of women with non-screen detected DCIS today, if surgical treatment is omitted. More importantly, the majority of diagnoses of DCIS today are not diagnosed symptomatically, but from mammographic screening, and this in itself is a significant prognostic factor (6), limiting the validity of these older case series.

Table 1.1: Demonstrating all available data on the natural history of Ductal Carcinoma in Situ (DCIS) without surgery. Modified from (2, 7).

Study ID	Location of study	DCIS type	No. untreated DCIS patients	Follow up duration (years)	% IBC at end of f/up	% all-cause mortality at end of f/up
Farrow 1970 (8)	USA	NR	25	1–8	5 (20%)	NR
Eusebi et al 1994 (9)	Italy	C, NC	80	5–28 ^a	11 (14%)	5 (6.3%)
Kraus & Neubecker 1962 (10)	USA	P	4	5-13	2 (50%)	NR
Lewis & Geschickter 1938 (11)	USA ^b	C	8 ^c	4	6 (75%)	1 (12.5%)
Millis & Thynne 1975 (12)	U.K.	C, P, CR	9 ^d	0–19	4 (44%)	3 (33%)
Page et al 1995 (13-15)	USA	NC	28	3–46	11 (39%)	5 (17.9%)
Rosen et al 1980 (16, 17)	USA	P	15	1-24 ^e	8 (53%)	3 (20%) ^e
Sagara et al 2015 ^f (18)	USA	NR	1169	6 ^g	NR	78 (6.7%)

Abbreviations: C=Comedo, CR=Cribiform, IBC=Invasive Breast Cancer, P=Papillary, NC-Non-comedo, NR=Not reported.

Superscript: ^aAverage follow-up period was 17.5 years. ^bCohort study of DCIS patients with varying degrees of operative excision. ^c8 patients were treated with excision biopsy alone with no analysis of operative margin. ^d8 treated with wide local excision with no report of margin thickness and possibility margins involvement suspected and 1 women treated with radiotherapy alone. ^eAverage follow-up period was 9.7 years. ^fRetrospective review of records previously unrecognized DCIS. ^gMedian follow-up was 72 months.

Due to the inherent confounding from selection bias in these retrospective case series, the natural history of DCIS without treatment can only be reliably examined in the setting of prospective randomised trials. There are two such trials underway, namely the LOw RiSk (LORIS) Trial and the LOw RiSk Dcis (LORD) study. The LORIS trial is a UK-based phase III multicentre study that has randomised vacuum biopsy diagnosed low and intermediate grade DCIS women to surgery or to active monitoring (19). The main outcome measured in this study is ipsilateral IBC free survival rate at 5 years. It will aim to investigate whether patients with newly diagnosed low risk DCIS can safely avoid surgery and whether those patients who do require surgery can be accurately identified. The study started in 2014 and should be completed by June 2020 (20). Simultaneously, the LORD trial is a similar randomised, phase III non-inferiority trial but is an international multicentre study organised by the European Organization for Research and Treatment of Cancer and the Dutch Breast Cancer Research Group and is due to complete in 2029. These two randomised trials will provide invaluable contributions to understanding the natural history of untreated low and intermediate grade DCIS (21). There are currently on-going efforts to undertake a third trial using similar methodology in the United States of America, called the ‘Comparison of Operative to Monitoring and Endocrine Therapy (COMET) trial for low risk DCIS’ (22).

None of the on-going or planned randomised trials will investigate outcomes in high grade DCIS in the absence of surgery and it is likely that a proportion of women within the active monitoring arm of each study will develop ipsilateral IBC due to limitations in the methods used in obtaining a biopsy diagnosis of DCIS grade, irrespective of whether it is undertaken by central pathology review or locally at randomising institutions. However, these studies will provide insights into the scale of the problem of over-treatment in DCIS as well as helping to better understand the natural history of low and intermediate grade DCIS.

1.4. National Breast Screening and the evolving incidence of DCIS

DCIS is the most common type of non-invasive breast tumour and with the advent of breast screening the number of women diagnosed with DCIS each year has been increasing, so that it now accounts for approximately one in five tumours detected by national screening programmes (23). In 2009 the United States National Institutes of Health (NIH) convened a State-of-the-Science Conference on DCIS which included a systematic review of the literature available at the time and also presentations by several dozen investigators who were carrying out research on DCIS (24). A panel of experts were invited to identify the most critical research questions for the diagnosis and management of DCIS in the light of the evidence they had heard. In particular, they noted the importance of improving the risk stratification of DCIS patients in order to identify the optimal therapy for each woman.

Although some additional studies have been performed since the NIH Conference they have mostly included only a small number of women and there is still a paucity of evidence on which to base clinical management decisions in DCIS. Furthermore, few randomised trials have been published since and it is unclear whether the results found in these trials can be translated to a population setting (25-27). What is clear is that there is still marked variation in the management of DCIS both between different countries (28, 29), and between treatment centres within the same country (30).

The majority of women diagnosed with DCIS in England today are between 50-70 years of age, corresponding to the age group eligible for breast cancer screening (Figure 1.2). The incidence of DCIS at ages 50-64 rose dramatically from 1988 onwards, coinciding with the introduction of the breast screening programme in that age-group, while the incidence at ages 65-69 rose rapidly from 2001, when the screening programme was extended to cover this age

group. In contrast, the incidence in the non-screened age groups (≤ 49 years and >70 years of age) has risen much more gradually. As a result, there is little doubt that the overall increase in DCIS incidence in the England observed over the last few decades is largely a direct result of the NHS breast-screening programme.

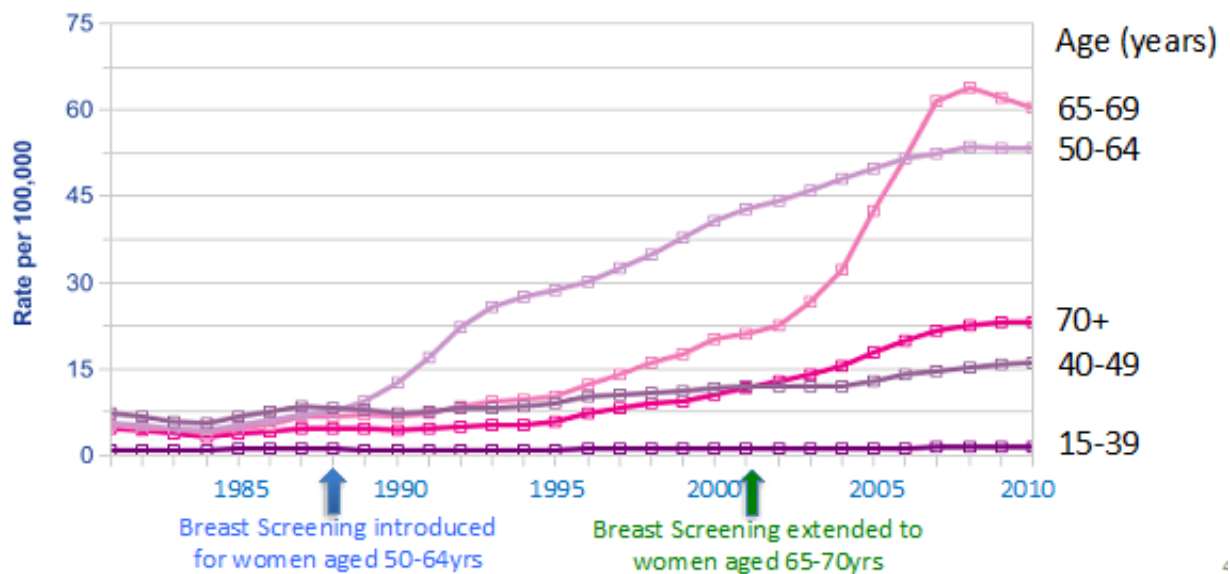


Figure 1.2: Incidence of DCIS by age. Incidence rates of ductal carcinoma in situ (DCIS) by age group per 100,000 women in Great Britain between 1979 - 2010. Modified from (31).

Whilst the introduction of the breast-screening programme has been the major factor for the overall increase in DCIS incidence in recent decades, it is not the only factor responsible. In the USA, the incidence of DCIS has increased over time from 1.87 per 100,000 women in 1975, to 32.5 per 100,000 women in 2005, even when the rate of screening mammography has remained constant (32, 33). In addition to this absolute increased in DCIS incidence over time, there is evidence to suggest that the incidence of different histological subtypes of DCIS has also changed over time (34). This has chiefly been driven by increases in non-comedo and prognostically favourable low-grade DCIS. In contrast, the incidence of the prognostically less favourable comedo-type DCIS has remained largely stable over that

period (35). Consequently, there may be influences on DCIS incidence in addition to the role of mammographic screening.

Since the advent of breast screening, several new questions have arisen regarding the treatment of DCIS. Some have argued that many cases of screen detected disease would never have become clinically apparent or caused mortality, and that diagnosing and treating such women simply causes psychological and physical harm through unnecessary surgery. Others argue that early detection of DCIS is important because it prevents the future development of IBC. What is known is that progression to IBC or death can take many years to occur. Hence critical examination of long-term data is important to help formulate appropriate patient selection for different management options.

Although it is clear that women with DCIS are at a higher risk of invasive breast cancer than other women (36), and that more intensive treatment reduces this risk, as yet no study has found a clear relationship between intensity of treatment and subsequent survival. Indeed there is epidemiological evidence to suggest that even complete omission of treatment in certain women may not negatively impact on survival. An example of this evidence comes from Sagara and colleagues who undertook a retrospective cohort study using the Surveillance, Epidemiology, and End Results (SEER) database and evaluated all women diagnosed with DCIS between 1988 and 2011 (18). Of 57, 222 women diagnosed with DCIS, 1169 cases (2.0%) did not receive surgery. There was a clear difference in the 10-year breast cancer-specific survival between the non-operative group (93.4%) and those that were treated with surgery (98.5%) ($p < 0.001$). However this difference was amongst women with high and intermediate grade disease and was not evident amongst women with low-grade disease

where the 10-year breast cancer-specific survival of the non-operative group was 98.8% and that of the surgery group was 98.6% ($p = 0.95$).

The findings from Sagara and colleagues demonstrate that the survival benefit from breast surgery for low-grade DCIS was lower than that for intermediate- or high-grade DCIS. Thus a stratified approach to treatment is warranted, taking into account a woman's individual risk based on patient and tumour-related factors. As a consequence there is uncertainty over the optimal level of treatment for women with DCIS. One consequence of this uncertainty is that there is wide variation in the degree of treatment of DCIS geographically, suggesting that many women may be being over- or under-treated (30, 37).

Although there is little information on long-term outcomes of women diagnosed with screen-detected DCIS in the UK, it is apparent that a small number of women diagnosed with DCIS do not survive their first 15 years following their diagnosis and have IBC as the cause of death on their death certificate. Physiologically, since DCIS is by definition a non-invasive disease, women do not die as a result of their diagnosis of DCIS *per se*, but instead as a consequence of occult or subsequent IBC. Thus an argument in favour of population screening has been that by improving the detection and treatment of DCIS, IBC (and therefore breast cancer mortality) would be decreased. In the UK, a significant reduction in breast cancer mortality from breast screening has yet to be proven, but in 2016 Duffy and colleagues provided the first evidence of a clear reduction in IBC following DCIS detection at the screening unit level. They compared DCIS detection frequencies with invasive interval cancer screen detection rates amongst 84 local screening units and found that amongst 90% of units in England, for every three screen-detected cases of DCIS diagnosed, there was one fewer invasive interval cancer in the subsequent 3 years at that unit (38).

While the findings from Duffy and colleagues' study do support the premise that the detection and treatment of DCIS helps prevent future invasive disease at a unit level, 3 years is a relatively short period of time in the natural history of DCIS or IBC and there is little longer-term evidence published to date on outcomes following DCIS in the UK (39). Furthermore, by grouping women with different risk profiles at the unit level in this way, such an analysis does little to contribute towards the discussion of personalising treatment according to individual risk profiles. A more meaningful approach would be to examine differences and similarities amongst patient, tumour and treatment factors in women diagnosed with DCIS and who suffer subsequent IBC with those who do not, and over a longer period of follow-up using the national screening audit and cancer registries. However, as yet this has not been undertaken in the UK.

By quantifying how the risk of DCIS recurrence or IBC occurrence following DCIS changes over time since the date of original diagnosis, population level data would also provide evidence to indicate the optimal duration and frequency of follow-up required for women diagnosed with DCIS. In the years after completion of their initial treatment, current follow-up for women with DCIS is similar to that for women with IBC. However, since DCIS is not invasive, it remains unclear whether this is truly warranted or whether follow-up should be less frequent than for women with IBC. This in turn, would have important economic and clinical implications.

1.5. Theories of disease progression in DCIS

Understanding the biological mechanism of DCIS progression is important for two main reasons. Firstly, it aids interpretation of epidemiological trends in DCIS incidence, outcomes over time and the elucidation of potential causal relationships from associations seen on population level data. Secondly, it may allow the development and clinical utility of molecular tests to predict the likelihood of disease progression, thus enabling the separation of women at high and low risk of progression, and paving the way for strategies that reduce overtreatment (40). Unfortunately, the pathway of disease progression in DCIS remains poorly understood and the underlying biological mechanisms have not yet been fully identified. Since the early 1970s, several probable models of disease progression have been formulated from the available evidence. This evidence includes: the presence/absence of molecular, morphological and immunohistochemical biomarkers of IBC and DCIS, computational simulation experiments of ductal cell division and genetic expression profiling with genomic analyses of DCIS and invasive tumour tissue.

In 1973, Wellings and Jensen, building on work by previous authors, examined 60 whole breast specimens and published the hypothesis of step-wise progression of DCIS into IBC (41). This model has since been further developed and in subsequent years has commonly been described as the ‘linear’ pathway whereby consecutive genetic and morphological changes mark the progression from normal ductal epithelium through the proliferative changes of epithelial atypia to atypical hyperplasia, and then to DCIS. This in turn progresses from grade 1 DCIS, through grade 2 to grade 3 DCIS followed by a milestone transition from non-invasive disease to grade 1 IBC, through to grade 2 and then finally grade 3 IBC. Credence for this model has arisen from the widely accepted and generally applied ‘multiple hit’ hypothesis of cancer whereby sequential aberrations in cell signalling pathways may

result in more aggressive cancer features. Evidence from transcriptomic and genomic studies have lent further support to this model (42).

In 2005, several decades following Wellings and Jensen's publication, Sontag and Axelrod undertook a computer simulation study and compared several mathematical models of disease progression with clinically observed frequencies of grades of DCIS and IBC that were reported to occur in heterogeneous tumours that contained both non-invasive and invasive disease of varying grades. They found that in tumours where DCIS and IBC were both present, the model of a 'parallel' pathway fitted the observed findings better than a linear model of disease progression (43). In the parallel model, it is assumed that both DCIS and invasive breast tumours have arisen from a common DCIS progenitor (which could be at DCIS grade 1, 2 or 3), and that both cell lines have continued to progress at similar rates through respective grades. Further support for this model comes from the observation that DCIS and IBC cells from the same breast have similar cellular copy numbers (with high genomic similarity), indicating a common clonal origin (44).

The linear and parallel models provide orderly descriptions of disease progression and are in contrast to the other contemporary 'nonlinear' or 'branched' model of disease progression from DCIS, whereby the transition to IBC can occur at any grade of DCIS towards any grade of IBC (45). This model of disease progression is analogous to a Darwinian tree with the trunk representing a common ancestor and the branches representing heterogeneous and novel mutations shifting the phenotype of the tumour from low to high grade and from non-invasive to invasive disease (46). This model is supported by genomic and mutational profiling experiments demonstrating that the progression of DCIS to IBC is analogous with the evolution of DCIS cell lines showing recurrent patterns of genomic imbalances,

consistent with a biological continuum (47, 48). It is likely that a proportion of DCIS encompass a heterogeneous spectrum of cancer cells with distinct genetic aberrations and so progression from DCIS to IBC in the absence of treatment may follow a Darwinian evolutionary model (40). This explains the common clinical finding of concomitant DCIS and IBC within the same breast. The three main theories of disease progression in DCIS are summarised in Figure 1.3.

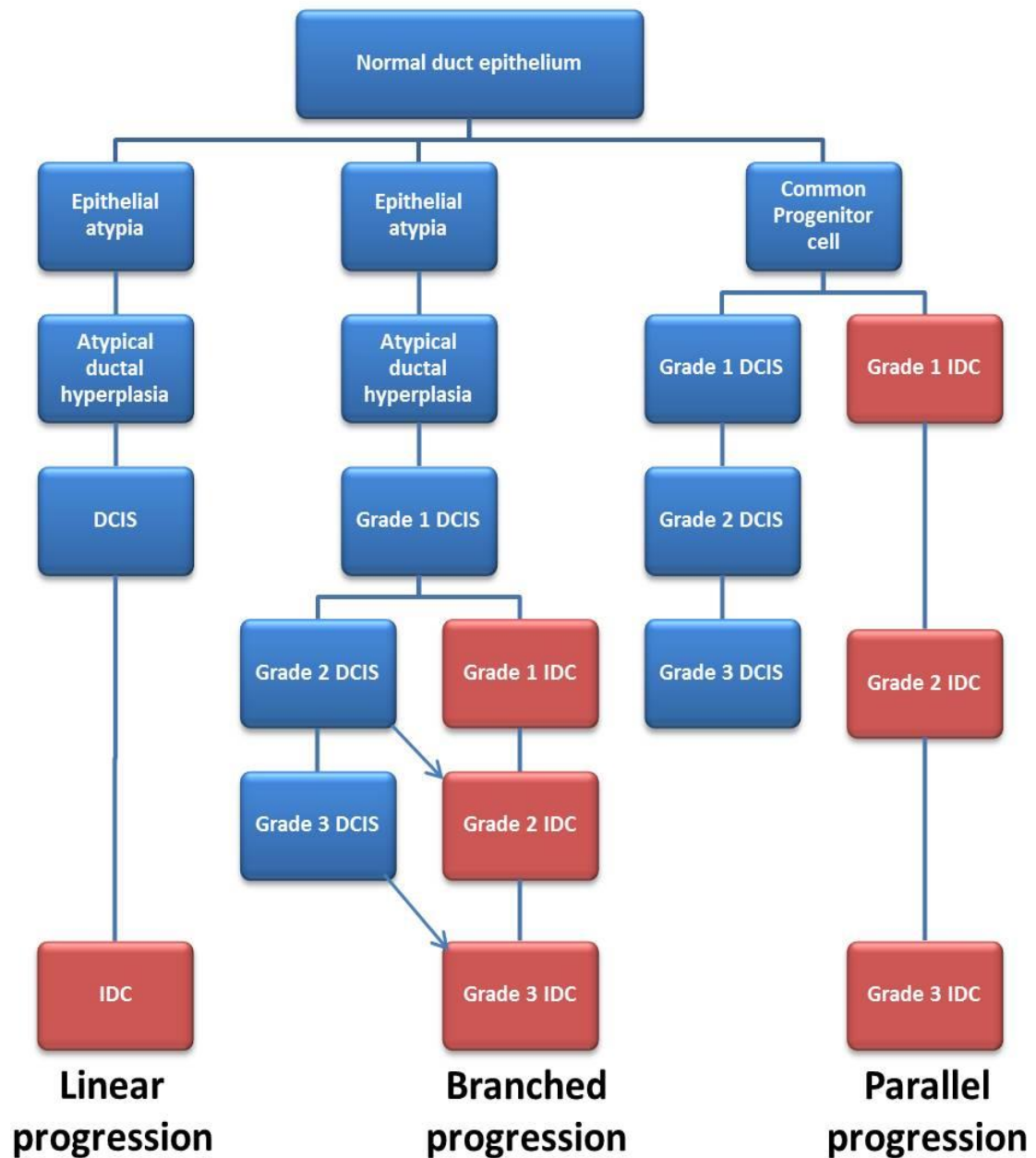


Figure 1.3: The main theories of disease progression in ductal carcinoma in situ (DCIS). IDC: invasive ductal carcinoma. Modified from (42).

1.6. Risk factors for DCIS incidence and recurrence

Comparing and contrasting the risk factors for DCIS and IBC can provide insights into why some women are diagnosed with IBC, and others with DCIS as their first cancer. It is clear that there are many similarities between these two conditions and this has been the principle support for the conventional view that DCIS and IBC are extremes of the same spectrum of disease (Table 1.2). However, there are not only epidemiological differences, such as the role of BMI on the incidence of the two conditions, but also evolving evidence to suggest DCIS and IBC have distinct genetic aberrations that can be modified to different degrees by modifiable lifestyle factors (49).

Table 1.2: Comparison of the risk factors for developing DCIS and IBC. Breast cancer 1 and 2 genes (BRCA 1 and 2), HRT (Hormone Replacement Therapy). *Whereas the incidence rate of invasive breast cancer increases with age, the increased incidence of DCIS in the age-group 60-74 years is due to the breast-screening programme. Modified from (50).

Factor	DCIS	Invasive cancer
Breast density	Women with high breast density have increased risk	
Family History	BRCA1/2 gene positive or positive family history increases risk	
HRT	Increased risk with HRT	
Mammographic screening	Increased risk	
Parity	Increased risk in women with who give birth at an older age or are nulliparous	
Race	Less common in Asian, African American, and Hispanic women compared with white women	
Tamoxifen	Decreased risk	
Age*	Peak 60-74yrs	Peak 75-79yrs
Body mass index (BMI)	No clear association	Increased risk with higher BMI in post-menopausal women

Whilst there is evidence internationally to show that recurrences rates following DCIS have fallen over time (51, 52), reliable information showing this in the UK is lacking. Well recognised risk factors for DCIS recurrence or occurrence of IBC after breast conserving surgery include resected margin width, histological subtype, nuclear grade, molecular phenotype and the absence of adjuvant therapy (53). With the introduction of national breast screening programmes in many countries, the method of DCIS detection and age at diagnosis

have emerged as further prognostic factors (54). Risk factors associated with poorer prognosis following treatment for DCIS are shown in Table 1.3.

Table 1.3: Factors associated with a poor prognosis following treatment for DCIS. A systematic search of the literature was undertaken to identify factors linked with poor prognosis (defined as DCIS recurrence, invasive breast cancer occurrence or mortality from breast cancer after DCIS). Factors are placed into three categories by a modified format of Greenhalgh's hierarchy whereby factors with supporting evidence from two or more randomised controlled trials were classified having 'high' evidence.

Risk factor	Evidence	RCTs
Positive, close, or unknown surgical margins	High	(55-57)
High tumour grade	High	(57, 58)
Comedo necrosis histological subtype	High	(55, 57)
Age <40 years at diagnosis	High	(55-59)
Symptomatic presentation	High	(55, 57)
High Van Nuys prognostic index	Moderate	(58)

1.7. Genetic risk factors for DCIS occurrence

Phenotypic and genetic features of DCIS associated with concomitant IBC, or with IBC occurrence following treatment for DCIS, have been of substantial interest in recent years. In addition to mutations in BRCA 1 and BRCA 2 genes associated with DCIS risk (60), a key gene mutation thought to underlie the progression from DCIS to IBC is in the gene TP53 which has been associated with high grade DCIS and HER2 positivity (61-63). Two further genetic drivers of clinical progression to IBC are mutations in PIK3CA and GATA 3 (64, 65), which are linked to oestrogen receptor positivity, transducing the promotor effect of circulating oestrogen (66, 67). There have been a number of studies investigating genetic risk factors for incidental IBC but considerably fewer studies examining genetic risk factors for

DCIS or for diagnosis of IBC after DCIS treatment. Furthermore, the genetic research in DCIS that has been published has focused largely on genetic features of the tumour tissue with less research on germ-line genetic susceptibility factors that may predispose some women to develop DCIS, or to develop IBC following DCIS (68-70).

The few studies that have examined genetic influences on outcomes following DCIS have investigated a limited number of single nucleotide polymorphisms (SNPs) that have previously been associated with breast cancer (71). One exception to this is the ICICLE study, which is a multicentre trial investigating genetic biomarkers in blood and tumour samples from women diagnosed with DCIS and from healthy volunteers (72). As accrual was completed in 2013, the ICICLE study is currently in its follow-up stage and focuses on inherited familial genetic variations that may predispose women to develop DCIS. Enrolled women completed questionnaires assessing risk factors. However, these were completed after DCIS diagnosis, meaning the information collected on risk factors is susceptible to recall bias. Furthermore, genetic analyses from the ICICLE study lack detailed lifestyle information and so cannot take lifestyle-associated factors into account when interpreting genetic data.

In addition to the studies cited above, many genome wide association studies have identified novel SNPs associated with IBC in recent years that had not previously been studied in the context of DCIS (61-67, 73-86) (Table 1.4). Evaluation of these newly identified SNPs in women with DCIS and in women with IBC, after taking known environmental risk factors into account, could further help our understanding of the aetiology and pathophysiology of DCIS as well as its subsequent prognosis. More importantly, such studies could help to go beyond what is known about the ‘non-obligate precursor’ concept of DCIS discussed in section 1.1.1, and identify the genetic characteristics associated with women with DCIS who

subsequently develop IBC and those who do not. This understanding is not only important to women with DCIS who are in the process of making treatment decisions to help reduce their future risk of a second breast cancer diagnosis [6], but also to clinicians aiming to quantify this risk in order to advise on optimal management and follow up frequency as well as duration. The current drive for such research stems from the goal of constructing a clinically relevant prognostic score that combines genetic and environmental risk factors to give a personalised risk profile for each woman in order to allow clinicians to provide a tailored treatment regime. By virtue of the heterogeneous nature of DCIS, it is clear that no single treatment strategy is ideal in all circumstances. Instead, treatment should be personalised in accordance with each woman's individual risk profile (1).

Table 1.4: GWAS studies studying genetic risk factors for invasive breast cancer. Results of a systematic review of genome wide association studies in breast cancer to date. Studies identified through PubMed/Medline Mesh search: ("Genome-Wide Association Study"[Majr]) AND "Breast Neoplasms"[Mesh] and through review of subsequent study abstracts. Data crosschecked with (87-91). NA: Not available, NR: Not reported, SNPs: Single nucleotide polymorphisms.

Citation	Initial Sample Size and ethnicity	Replication Sample Size and ethnicity	SNPs associated with breast cancer	Platform
Ahsan H (61)	6225 European	8945 European	18	Illumina
Purrington KS (62)	4928 European	3457 European	4	Illumina
Rudolph A (63)	2,920 European	16955 European	4	Illumina
Low SK (64)	4741 Japanese	6280 Japanese	4	Illumina
Chung S (65)	1183 Japanese	1183 Japanese	30	Illumina
Garcia-Closas M (66)	39387 European	47969 European	17	Illumina
Michailidou K (67)	22627 European	87170 European	63	Illumina & Affymetrix
Couch FJ (73)	2727 European	11964 European	6	Illumina
Liu M (74)	772 international	NA	18	Illumina
Song C (75)	5761 African American	NA	21	Illumina
Hein R (76)	731 European	16093 European & East Asian	NR	Illumina
Rinella ES (77)	1001 Ashkenazi Jewish	466 Ashkenazi Jewish	3	Affymetrix
Rafiq S (78)	536 European	1,516 European	5	Illumina
Orr N (79)	3618 European	912 European	3	Illumina
Siddiq A (80)	36278 European & African American	9350 international	4	Illumina
Elgazzar S (81)	2902 Japanese	4450 Japanese	2	Illumina
Chen F (82)	5761 African American	14579 African	1	Illumina
Kim HC (83)	4325 Korean	7894 Korean	1	Affymetrix
Long J (84)	5242 Chinese	34455 East Asian	3	Affymetrix
Shu XO (85)	1,950 Chinese	4,160 Chinese	3	Affymetrix
Kiyotani K (86)	240 Japanese	222 Japanese	1	Illumina
Haiman CA (92)	9137 African American and European	19193 European	1	Illumina
Cai Q (93)	4128 East Asian	29968 East Asian	2	Affymetrix
Sehrawat B (94)	623 European	2368 European	2	Affymetrix
Fletcher O (95)	6346 European	18021 European	9	Illumina
Spraggs CF (96)	323 European	179 European	NR	Illumina

Table 1.4: Continued from previous page.

Citation	Initial Sample Size and ethnicity	Replication Sample Size and ethnicity	SNPs associated with breast cancer	Platform
Li J (97)	5200 European	8615 European	NR	Illumina
Gaudet MM (98)	1703 European	2486 European	1	Affymetrix
Li J (99)	8428 European	14962 European	4	Illumina
Antoniou AC (100)	2383 European	5986 European	1	Illumina
Long J (101)	4157 Chinese	20618 East Asian & European controls	1	Affymetrix
Turnbull C (102)	8556 European	24799 European	13	Illumina
Azzato EM (103)	1,145 European	4,335 European	NR	Illumina
Thomas G (104)	2287 Europe	18282 European	6	Illumina
Zheng W (105)	3027 Chinese	10559 Chinese	1	Affymetrix
Kibriya MG (106)	60 international	NA	1	Affymetrix
Gold B (107)	548 Ashkenazi Jewish	2359 Ashkenazi Jewish	1	Affymetrix
Murabito JM (108)	723	NA	9	Affymetrix
Easton DF (109)	754	51535	6	Perlegen
Hunter DJ (110)	2287 European	3248 European	1	Illumina
Stacey SN (111)	13145 European	12198 international	2	Illumina

1.8. Biomarkers of prognosis in DCIS

The key driving force towards overtreatment in DCIS is the uncertainty about which DCIS cases may have occult IBC present and which cases may go on and develop IBC if not treated. Biomarkers predicting adverse outcomes following DCIS are by definition predictors of the occult presence of, or risk of development of, IBC in women with DCIS. The former is more important in the short-term and is a consequence of limitations of radiology, technical biopsy methods and techniques for pathological diagnosis. The term biomarker, in this context, refers to a characteristic that can be evaluated as an indicator of the presence of IBC and this term has come into vogue in the context of personalised medicine and patient-centred prognostic indices (112, 113). A number of biomarkers for the presence of IBC in the setting of DCIS have been published, of which there are two main types: immunohistochemical, and genetic markers. These are summarised in Table 1.5 below. In addition, positive expression of ERBB2 and negative expression of bcl-2, oestrogen receptor (ER) and progesterone receptor (PR) have each been associated with increased risk of clinical recurrence (114).

Table 1.5: Biomarkers for the presence of IBC in biopsies showing DCIS. Established biomarkers for the presence IBC in DCIS that could be compared between initial biopsies that are upstaged with those that are not, in order to help provide insights into improving the accuracy of DCIS biopsies and also form the basis of future studies examining invasive transformation of ductal cells.

Biomarker	Relevant supporting citations
ER/PR	(115-121)
HER2	(116-120, 122-124)
P16	(120, 125-129)
Cox-2	(125, 126, 129-132)
Ki-67	(116, 120, 125, 126, 129)
p53	(127, 128, 133)
bcl-2	(134-136)
Hsp-27	(137-139)
EGFR	(115, 140)

1.9. Developing prognostic indices for women with DCIS.

The historical classification of DCIS was pragmatic, and based on histological appearance rather than prognostic value (141). Several classification systems have arisen in the last few decades, but each is hampered by inter-rater variability amongst the pathologists using them (Table 1.6). This inter-rater variability arises from a subjective component in the interpretation of cellular appearances. The ability to provide a uniform categorisation of disease type is paramount to risk prediction and subsequent treatment decision-making. One approach to reduce inter-rater variability could be to use pattern recognition image analysis software combined with artificial intelligence and deep neural networks (142-144). Another approach could be to shift analysis techniques towards more objective, molecular marker based classification systems (145-147). Proof of principle supporting the use of molecular classification systems arises from the successful implementation of one-step nucleic acid (OSNA) amplification in the intraoperative detection of lymph node metastasis in breast cancer patients (148, 149). However, both artificial intelligence pattern recognition and molecular classification assays, whilst showing promising results, are yet to overcome pragmatic hurdles in achieving mainstream use in tumour analysis. Thus, irrespective of their limitations from inter-rater variability, the traditional histopathological classification systems are likely to remain the primary way of describing tumours for the foreseeable future, at least for use in epidemiology studies.

Table 1.6: Demonstrating classification systems commonly used in DCIS. Agreement is defined as percentage agreement in studies assessing inter-rater variability amongst pathologists when examining pathology slides: Poor agreement (less than or equal to 50% agreement), Moderate agreement (51-80% agreement), Good agreement (more than 81% agreement). *Table modified from and data obtained from (150).

Classification	Group	Necrosis	Nuclear grade	Cell polarization	Percentage agreement in each classification grouping amongst pathologists examining pathology slides*
Van Nuys (151)	Group 1	Absent	Low	N/A	47
	Group 2	Absent	Intermediate		36
	Group 3	Present	High		93
Modified Lagios (152)	Low grade	Absent	Low	N/A	50
	Intermediate grade	Focal	Intermediate		73
	High grade	Extensive	High		63
Holland (153)	Well differentiated	N/A	Well differentiated	Absent or very focal	88
	Intermediately differentiated		Intermediately differentiated	Present and not prominent	92
	Poorly differentiated		Poorly differentiated	Prominent	76

The Van Nuys Prognostic Index (VNPI) aims to predict the risk of further ipsilateral breast events in women with DCIS based on the factors shown in Table 1.6, in addition to: tumour size, margin width and age (154). It is accompanied by recommendations of excision alone, excision plus radiotherapy, or mastectomy for women with a low, medium or high VNPI score respectively. The Van Nuys classification has gained popularity over other classification systems as it is easy to use, simplifies a difficult decision making process and is more consistently reported amongst pathologists with lower inter-rater variation amongst cases of higher grade disease than other histopathological systems (155). However, it was derived from the experience of a single centre and from data analysed retrospectively (156). The most recent version of the VNPI was based on observational data from 949 women (604 treated with excision alone and 345 treated with excision plus radiotherapy), and followed for

a median time of 86 months. Several studies have attempted to validate the VNPI (156-160), but they have had only moderate success. As a result the index is not uniformly used, although its constituent factors form the basis of clinical management decisions (161).

1.10. Contemporary treatment strategies for DCIS

The management of DCIS can be considered under the categories of surgery, radiotherapy and endocrine therapy. However, there are unanswered clinical questions within each of these treatment modalities and these, alongside the evidence supporting each of these modalities, are described below.

1.10.1. Breast Surgery

In the past many women with DCIS underwent mastectomy. However, in recent years with developments in advanced oncoplastic procedures, even relatively large lesions can now be treated with breast conserving surgery (BCS). The trend towards less extensive surgery for women with DCIS has consistently tended to lag behind that for IBC. Currently, approximately 75% of women with DCIS have BCS compared to 78% of women with IBC (162). The decision on which of these two procedures to carry out is based on several factors, including the size of the DCIS in relation to the patient's breast, multifocality and patient choice. BCS alone has lower surgery-related morbidity than mastectomy, but it has a higher rate of further ipsilateral breast events compared with mastectomy (163, 164).

When BCS is given without radiotherapy, surgical margin width is possibly the most important factor in predicting the rate of further ipsilateral breast events (165). Lower risks have been observed in women with negative margins than in women with involved margins

(166), and among women with negative margins lower risks have been reported with increasing margin width up to around 6 mm. Some studies suggest that there is little additional benefit in a margin width of more than 6 mm (17) while others have suggested that a 10 mm margin may be needed to achieve risks comparable with those from mastectomy (167). There is also considerable evidence that margin width is a determinant of further ipsilateral breast events even in the presence of radiotherapy. In a review of 4660 patients from 22 studies of radiotherapy after breast-conserving surgery, patients with negative margins had lower event rates than those with either positive margins (odds ratio [OR] 0.36; 95% confidence interval [CI] 0.27-0.47) or close margins (OR 0.59; 95% CI 0.42-0.83), while among those with negative margins a 2 mm margin was superior to a margin of <2 mm (OR 0.53, 95% CI 0.26-0.96) (168).

The uncertainty regarding optimal margin width in the published literature is reflected in the variation in professional guidance on the topic between clinical bodies. For example, the Society of Surgical Oncology (SSO), American Society for Radiation Oncology (ASTRO) and American Society of Clinical Oncology (ASCO) Consensus, as well as the National Institute of Clinical Excellence (NICE) Guidelines recommend a 2 mm margin (169, 170). However, the Association of Breast Surgery (ABS) recommends 1 mm and the Sloane questionnaire of participating units in 2008 concluded that local margin protocols ranged from 1 mm to 10 mm (171). Generally, many surgeons agree that a margin width less than 2 mm warrants a discussion with the patient regarding re-excision, although there is little evidence documenting the benefit of re-excision. Long term follow up of a large group of DCIS patients with varying margin sizes, such as would be available from the breast screening audit, could make a substantial contribution to the existing evidence-base on optimal margin distance.

1.10.2. Core-Biopsy in DCIS

Core-needle breast biopsy is usually performed under ultrasound guidance for a breast lesion suspicious for cancer. The biopsy sample is analysed by a histopathologist and a preliminary diagnosis made prior to the management plan being discussed at a multidisciplinary team meeting and before surgical intervention. However, a core biopsy diagnosis of DCIS is sometimes upstaged to a diagnosis of IBC following the definitive histological analysis of the resected specimen. When IBC is identified in these circumstances, patients will often undergo a second surgical procedure to evaluate the status of their axillary nodes. Therefore quantifying the accuracy of core biopsy in DCIS patients and investigating the risk factors for missing IBC in DCIS core biopsies is of paramount importance in avoiding the potential psychological, financial and medical complications of further surgery. The importance of an accurate initial diagnosis is even more important in light of recent interest in treating low grade DCIS with observation alone and in the absence of surgical excision (172).

A number of studies have investigated the accuracy of core biopsy in DCIS patients and have reported from 17.2% to 37.2% of surgically excised specimens for core biopsies showing DCIS actually contain IBC (173-181). Whilst, there has been evidence to show that more tissue obtained from vacuum assisted biopsies may improve the sampling accuracy, there has been conflicting information in the literature regarding general risk factors for upstaging DCIS biopsy specimens to IBC, since such studies have largely been small scale, single-centre and retrospective in nature. Meta-analysis of the published reports from these studies is thereby limited by their differences in populations, local diagnostic and treatment protocols. As a result, there remains a need for large scale, standardised data on this topic in order to guide future management decisions. This is particularly pertinent in recent years with the start of randomised trials of non-operative treatment for low risk disease (see section 1.1.2 above).

1.10.3. Axillary staging

Sentinel lymph node biopsy (SLNB) has become the standard approach in evaluating the presence of axillary metastasis in the clinically and radiologically node negative axilla for women with IBC in many centres worldwide. Although introduced as a ‘gatekeeping’ procedure to minimise the morbidity resulting from conventional axillary lymph node dissection, it has since evolved as the sole treatment modality for managing axillary isolated tumour cell (ITC) metastasis, and also for micrometastatic axillary involvement (182, 183). Although controversial, removal of macrometastatic nodes by SLNB has started to be considered adequate surgery for the axilla in some circumstances (183, 184).

During the past 10 years, SLNB has been rapidly adopted into clinical practice, but long-term outcomes are uncertain. In the UK, NICE guidelines recommend SLNB for DCIS patients undergoing mastectomy, but not for those undergoing BCS unless there is a palpable mass or extensive microcalcifications (185), and in the US, the ASCO guidelines advise against SNLB in patients undergoing BCS unless the lesion is >5 cm or there is suspected or proven microinvasion (186). Under both of these guidelines, only a small percentage of patients undergoing BCS would receive SNLB, as many lesions are detected when smaller and prior to microinvasion. However, a study of 62,935 women registered in the US SEER cancer registries as having received BCS for DCIS from 2000 to 2008 showed that 12% of them had received SLNB, and a further 3% had one or more lymph node removed and examined, although SLNB was not specifically recorded (187)

Recent advances in the ability to analyse SLNB histology have given rise to new questions regarding the long-term outcome of women with DCIS where a SLNB has shown isolated tumour cells or micro-metastases, but where the surgically excised specimen has not revealed

IBC. When this happens, it raises doubt regarding the original diagnosis of DCIS, which is by definition a non-invasive cancer, and also on whether further axillary treatment is required. Examination of outcomes using population level data could help to reduce some of the uncertainty regarding appropriate treatment for such women.

1.10.4. Adjuvant radiotherapy

Five randomised trials comparing BCS plus adjuvant radiotherapy with BCS alone have been carried out in women with DCIS (55, 59, 160, 166, 188-191). These are shown in Table 1.7. With over 4000 women randomised, the trials clearly showed that radiotherapy after BCS reduced both IBC occurrence and non-invasive recurrence of DCIS. These trials were meta-analysed at the individual patient level in the 2010 Early Breast Cancer Trialists' Collaborative Group (EBCTCG) overview, which showed radiotherapy reduced the absolute 10-year risk of any ipsilateral breast event (i.e. either recurrent DCIS or invasive cancer) by 15.2% (SE 1.6%, 12.9% vs 28.1% 2 P <.00001) (161). The trials did not, however, find that radiotherapy had a significant impact on breast cancer mortality and at present it is unclear whether this was simply due to the limited number of high-risk women who were randomised, so that only a small number of IBCs were observed in the trials, or whether there is no such effect.

Unfortunately, detailed data on several potentially important factors, including margin width, were lacking in the trials as these were not routinely recorded at the time the women were randomised. As a result, it is unclear whether all women with DCIS would benefit from radiotherapy after BCS, or whether there are groups of women for whom radiotherapy need not be given. This uncertainty may, to some degree, explain why there is such variation

between clinicians regarding the delivery of adjuvant radiotherapy following BCS for these patients across the UK (192).

Table 1.7: Showing all randomized trials comparing breast surgery alone with breast surgery and radiotherapy for DCIS. Modified from (161).

Study	Year completed accrual	No. women randomised*	Median f/up (yr)	Surgery	Radiotherapy	Events in surgery + RT (%)	Events in surgery alone (%)	RR recurrence (SE)
NSABP B17 (57, 188)	1990	818	16.5	BCS (+/- AC)	50 Gy	19.5	34.9	0.49 (0.10)
EORTC 10853 (55, 58, 193, 194)	1996	1010	10.4	BCS (+/- AC)	50 Gy	13.9	25.0	0.52 (0.11)
SWeDCIS (190, 195, 196)	1999	1067	8.4	BCS (+/- AC)	48/50 /54 Gy	11.5	26.2	0.41 (0.10)
UK/ANZ DCIS (59)	1998	1030	4.8	BCS	50 Gy/f	5.5	13.5	0.41 (0.14)
RTOG (25)	2006	636	7.2	BCS	42.5/50/50.4 Gy	10.8	11.5	NA

AC = Axillary Clearance; EORTC = European Organisation for Research and Treatment of Cancer; f = fraction; Gy = Gray; NA = Data not available at time of EBCTCG meta-analysis; NSABP = National Surgical Adjuvant Breast and Bowel Project; RTOG = Radiation Therapy Oncology Group; RR recurrence= Ratio for annual ipsilateral breast cancer recurrence rates; SE = Standard Error. Breast event is defined as ipsilateral recurrence of DCIS, ipsilateral occurrence of invasive breast cancer, contralateral occurrence of DCIS or invasive breast cancer, or regional or distant metastasis of breast cancer. *Number excludes drop-outs.

1.10.5. Adjuvant hormonal therapy

Oestrogen receptor (ER) and human epidermal growth factor receptor (HER) status are being measured in an increasing number of DCIS patients but the relevance of these hormone receptors in DCIS is unclear, as the association between receptor status and long-term outcome is uncertain. Two randomised trials, involving over 3000 randomised patients, have reported that tamoxifen reduces the rate of further ipsilateral and contralateral breast events (189, 197). However, a recent Cochrane meta-analysis of tamoxifen use after local excision for DCIS, with or without adjuvant radiotherapy, showed that although use of tamoxifen reduced the risk of recurrent DCIS, it did not result in a significant reduction in all-cause

mortality (198). Consequently there has not been an update to the 2009 NICE guidance, which advises against the routine use of tamoxifen therapy in DCIS (199). As yet, no trials have reported on the benefit of Trastuzumab in DCIS, although at least one trial is on-going (200). Analyses of observational data from the UK cancer registries could contribute to the question of the risks and benefits from hormonal therapy use in DCIS by quantifying the breast-cancer specific and all-cause mortality rates amongst women with DCIS who have received hormone therapy with those who have not, on a population level after taking all other relevant factors into account.

1.11. Conclusion

Despite its increasing incidence, there remain a number of fundamental and clinically important uncertainties relating to the aetiology, diagnosis and treatment of DCIS. This chapter has highlighted the unanswered questions that would benefit most from further study. Research into each of these areas has the potential to benefit large numbers of women each year. This thesis will aim to focus on these research questions and thereby contribute to the literature upon which contemporary clinical practice is based.

1.12. References

1. Mokbel K, Cutuli B. Heterogeneity of ductal carcinoma in situ and its effects on management. *The Lancet Oncology*. 2006;7(9):756-65.
2. Silverstein MJ. Ductal Carcinoma In Situ of the Breast. Second edition ed. Lagios ARMD, editor. Philadelphia, USA: Lippincott Williams & Wilkins; 2002 15/08/2002.
3. Muir R. The evolution of carcinoma of the mamma. *The Journal of Pathology and Bacteriology*. 1941;52(2):155-72.
4. Barnes NLP, Dimopoulos N, Williams KE, Howe M, Bundred NJ. The frequency of presentation and clinico-pathological characteristics of symptomatic versus screen detected ductal carcinoma in situ of the breast. *European Journal of Surgical Oncology*. 40(3):249-54.
5. Shin HJ, Kim HH, Kim SM, Kwon GY, Gong G, Cho OK. Screening-Detected and Symptomatic Ductal Carcinoma in Situ: Differences in the Sonographic and Pathologic Features. *American Journal of Roentgenology*. 2008;190(2):516-25.
6. Cheung S, Booth ME, Kearins O, Dodwell D. Risk of subsequent invasive breast cancer after a diagnosis of ductal carcinoma in situ (DCIS). *The Breast*. 23(6):807-11.
7. Leonard GD, Swain SM. Ductal carcinoma in situ, complexities and challenges. *Journal of the National Cancer Institute*. 2004;96(12):906-20.
8. Farrow JH. Current concepts in the detection and treatment of the earliest of the early breast cancers. *Cancer*. 1970;25(2):468-77.
9. Eusebi V, Feudale E, Foschini MP, Micheli A, Conti A, Riva C, et al. Long-term follow-up of in situ carcinoma of the breast. *Seminars in diagnostic pathology*. 1994;11(3):223-35.
10. Kraus FT, Neubecker RD. The differential diagnosis of papillary tumors of the breast. *Cancer*. 1962;15:444-55.
11. Lewis D, Geschickter CF. Comedo carcinoma of the breast. *Archives of Surgery*. 1938;36(2):225-44.
12. Millis RR, Thynne GS. In situ intraduct carcinoma of the breast: a long term follow-up study. *Br J Surg*. 1975;62(12):957-62.
13. Page DL, Dupont WD, Rogers LW, Landenberger M. Intraductal carcinoma of the breast: follow-up after biopsy only. *Cancer*. 1982;49(4):751-8.
14. Page DL, Dupont WD, Rogers LW, Jensen RA, Schuyler PA. Continued local recurrence of carcinoma 15-25 years after a diagnosis of low grade ductal carcinoma in situ of the breast treated only by biopsy. *Cancer*. 1995;76(7):1197-200.
15. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer*. 2005;103(12):2481-4.
16. Rosen PP, Braun DW, Jr., Kinne DE. The clinical significance of pre-invasive breast carcinoma. *Cancer*. 1980;46(4 Suppl):919-25.
17. Betsill WL, Jr., Rosen PP, Lieberman PH, Robbins GF. Intraductal carcinoma. Long-term follow-up after treatment by biopsy alone. *JAMA*. 1978;239(18):1863-7.
18. Sagara Y, Mallory MA, Wong S, Aydogan F, DeSantis S, Barry WT, et al. Survival Benefit of Breast Surgery for Low-Grade Ductal Carcinoma In Situ: A Population-Based Cohort Study. *JAMA Surg*. 2015;150(8):739-45.
19. Francis A, Fallowfield L, Rea D. The LORIS Trial: Addressing Overtreatment of Ductal Carcinoma In Situ. *Clinical oncology*. 2015;27(1):6-8.
20. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *European Journal of Cancer*. 2015;51(16):2296-303.

21. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, Dif N, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ – The LORD study. *European Journal of Cancer*. 2015;51(12):1497-510.
22. Youngworth LM, Boughey JC, Hwang ES. Surgery versus monitoring and endocrine therapy for low-risk DCIS: The COMET Trial. *Bulletin of the American College of Surgeons*. 2017.
23. NHSBSP. Screening for breast cancer in England: PAST AND FUTURE. NHSBSP Publication 2006(No 61):24-5.
24. National Cancer Institute. National Institutes of Health State-Of-The-Science Meeting on Ductal Carcinoma In Situ: Management and Diagnosis. *Journal of the National Cancer Institute Monographs*. 2009(41):111-223.
25. McCormick B, Winter K, Hudis C, Kuerer HM, Rakovitch E, Smith BL, et al. RTOG 9804: a prospective randomized trial for good-risk ductal carcinoma in situ comparing radiotherapy with observation. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2015;33(7):709-15.
26. Lee O, Page K, Ivancic D, Helenowski I, Parini V, Sullivan ME, et al. A randomized phase II presurgical trial of transdermal 4-hydroxytamoxifen gel versus oral tamoxifen in women with ductal carcinoma in situ of the breast. *Clin Cancer Res*. 2014;20(14):3672-82.
27. McCaffery K, Nickel B, Moynihan R, Hersch J, Teixeira-Pinto A, Irwig L, et al. How different terminology for ductal carcinoma in situ impacts women's concern and treatment preferences: a randomised comparison within a national community survey. *BMJ Open*. 2015;5(11):e008094.
28. Dick AW, Sorbero MS, Ahrendt GM, Hayman JA, Gold HT, Schiffhauer L, et al. Comparative effectiveness of ductal carcinoma in situ management and the roles of margins and surgeons. *Journal of the National Cancer Institute*. 2011;103(2):92-104.
29. NHSBSP. Uncertainties in the management of screen-detected ductal carcinoma in situ. *NHS Cancer Screening Programmes*. 2008(66):1-28.
30. Mammary Fold A, Research C. Variation in the management of ductal carcinoma in situ in the UK: Results of the Mammary Fold National Practice Survey. *Eur J Surg Oncol*. 2016;42(8):1153-61.
31. UK CR. In situ breast carcinoma incidence statistics London: Cancer Research UK; 2017 [Available from: <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-in-situ - heading-Zero>].
32. LaG R, D M, M K. SEER Cancer Statistics Review, 1975–2005: National Cancer Institute; 2008 [Available from: http://seer.cancer.gov/csr/1975_2005/].
33. Virnig BA, Shamliyan T, Tuttle TM, Kane RL, Wilt TJ. Diagnosis and management of ductal carcinoma in situ (DCIS). *Evidence report/technology assessment*. 2009(185):1-549.
34. Virnig BA, Wang SY, Shamliyan T, Kane RL, Tuttle TM. Ductal carcinoma in situ: risk factors and impact of screening. *Journal of the National Cancer Institute Monographs*. 2010;2010(41):113-6.
35. Li CI, Daling JR, Malone KE. Age-specific incidence rates of in situ breast carcinomas by histologic type, 1980 to 2001. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2005;14(4):1008-11.
36. Page DL, Dupont WD, Rogers LW, Jensen RA, Schuyler PA. Continued Local Recurrence of Carcinoma 15-25 Years after a Diagnosis of Low Grade Ductal Carcinoma In Situ of the Breast Treated Only by Biopsy. *Cancer-Paper Edition*. 1995;76(7):1197-200.

37. Dodwell D, Clements K, Lawrence G, Kearins O, Thomson CS, Dewar J, et al. Radiotherapy following breast-conserving surgery for screen-detected ductal carcinoma in situ: indications and utilisation in the UK. Interim findings from the Sloane Project. *British journal of cancer*. 0000;97(6):725-9.
38. Duffy SW, Diben A, Michalopoulos D, Offman J, Parmar D, Jenkins J, et al. Screen detection of ductal carcinoma in situ and subsequent incidence of invasive interval breast cancers: a retrospective population-based study. *The Lancet Oncology*.17(1):109-14.
39. Thomas J, Hanby A, Pinder SE, Ball G, Lawrence G, Maxwell A, et al. Adverse surgical outcomes in screen-detected ductal carcinoma in situ of the breast. *European Journal of Cancer*. 2014;50(11):1880-90.
40. Cowell CF, Weigelt B, Sakr RA, Ng CK, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Molecular oncology*. 2013;7(5):859-69.
41. Wellings SR, Jensen HM. On the origin and progression of ductal carcinoma in the human breast. *Journal of the National Cancer Institute*. 1973;50(5):1111-8.
42. Kaur H, Mao S, Shah S, Gorski DH, Krawetz SA, Sloane BF, et al. Next-generation sequencing: a powerful tool for the discovery of molecular markers in breast ductal carcinoma in situ. *Expert Rev Mol Diagn*. 2013;13(2):151-65.
43. Sontag L, Axelrod DE. Evaluation of pathways for progression of heterogeneous breast tumors. *Journal of Theoretical Biology*. 2005;232(2):179-89.
44. Johnson CE, Gorringer KL, Thompson ER, Opeskin K, Boyle SE, Wang Y, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast cancer research and treatment*. 2012;133(3):889-98.
45. Sontag L, Axelrod DE. Evaluation of pathways for progression of heterogeneous breast tumors. *J Theor Biol*. 2005;232(2):179-89.
46. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. *Sci Transl Med*. 2012;4(127):127ps10.
47. Heselmeyer-Haddad K, Berroa Garcia LY, Bradley A, Ortiz-Melendez C, Lee WJ, Christensen R, et al. Single-cell genetic analysis of ductal carcinoma in situ and invasive breast cancer reveals enormous tumor heterogeneity yet conserved genomic imbalances and gain of MYC during progression. *Am J Pathol*. 2012;181(5):1807-22.
48. Hernandez L, Wilkerson PM, Lambros MB, Campion-Flora A, Rodrigues DN, Gauthier A, et al. Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. *J Pathol*. 2012;227(1):42-52.
49. Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, et al. Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet*. 2013;9(3):e1003284.
50. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal carcinoma in situ of the breast: a systematic review of incidence, treatment, and outcomes. *Journal of the National Cancer Institute*. 2010;102(3):170-8.
51. Subhedar P, Olcese C, Patil S, Morrow M, Van Zee KJ. Decreasing Recurrence Rates for Ductal Carcinoma In Situ: Analysis of 2996 Women Treated with Breast-Conserving Surgery Over 30 Years. *Ann Surg Oncol*. 2015;22(10):3273-81.
52. Habel LA, Achacoso NS, Haque R, Nekhlyudov L, Fletcher SW, Schnitt SJ, et al. Declining recurrence among ductal carcinoma in situ patients treated with breast-conserving surgery in the community setting. *Breast cancer research : BCR*. 2009;11(6):R85.
53. Benson JR, Wishart GC. Predictors of recurrence for ductal carcinoma in situ after breast-conserving surgery. *The Lancet Oncology*. 2013;14(9):e348-57.

54. Shamliyan T, Wang SY, Virnig BA, Tuttle TM, Kane RL. Association between patient and tumor characteristics with clinical outcomes in women with ductal carcinoma in situ. *Journal of the National Cancer Institute Monographs*. 2010;2010(41):121-9.
55. Group EBCC, Group ER, Bijker N, Meijnen P, Peterse JL, Bogaerts J, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma-in-situ: ten-year results of European Organisation for Research and Treatment of Cancer randomized phase III trial 10853--a study by the EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2006;24(21):3381-7.
56. Holmberg L, Garmo H, Granstrand B, Ringberg A, Arnesson LG, Sandelin K, et al. Absolute risk reductions for local recurrence after postoperative radiotherapy after sector resection for ductal carcinoma in situ of the breast. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2008;26(8):1247-52.
57. Fisher ER, Dignam J, Tan-Chiu E, Costantino J, Fisher B, Paik S, et al. Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) eight-year update of Protocol B-17: intraductal carcinoma. *Cancer*. 1999;86(3):429-38.
58. Bijker N, Peterse JL, Duchateau L, Julien JP, Fentiman IS, Duval C, et al. Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2001;19(8):2263-71.
59. Houghton J, George WD, Cuzick J, Duggan C, Fentiman IS, Spittle M, et al. Radiotherapy and tamoxifen in women with completely excised ductal carcinoma in situ of the breast in the UK, Australia, and New Zealand: randomised controlled trial. *Lancet*. 2003;362(9378):95-102.
60. Yang RL, Mick R, Lee K, Graves HL, Nathanson KL, Domchek SM, et al. DCIS in BRCA1 and BRCA2 mutation carriers: prevalence, phenotype, and expression of oncogenes C-MET and HER3. *J Transl Med*. 2015;13:335.
61. Ahsan H, Halpern J, Kibriya MG, Pierce BL, Tong L, Gamazon E, et al. A genome-wide association study of early-onset breast cancer identifies PFKFB3 as a novel breast cancer gene and supports a common genetic spectrum for breast cancer at any age. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2014;23(4):658-69.
62. Purrington KS, Slager S, Eccles D, Yannoukakos D, Fasching PA, Miron P, et al. Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer. *Carcinogenesis*. 2014;35(5):1012-9.
63. Rudolph A, Hein R, Lindstrom S, Beckmann L, Behrens S, Liu J, et al. Genetic modifiers of menopausal hormone replacement therapy and breast cancer risk: a genome-wide interaction study. *Endocr Relat Cancer*. 2013;20(6):875-87.
64. Low SK, Takahashi A, Ashikawa K, Inazawa J, Miki Y, Kubo M, et al. Genome-wide association study of breast cancer in the Japanese population. *PLoS One*. 2013;8(10):e76463.
65. Chung S, Low SK, Zembutsu H, Takahashi A, Kubo M, Sasa M, et al. A genome-wide association study of chemotherapy-induced alopecia in breast cancer patients. *Breast cancer research : BCR*. 2013;15(5):R81.
66. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet*. 2013;45(4):392-8, 8e1-2.

67. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet.* 2013;45(4):353-61, 61e1-2.
68. Muggen AA, Hallett M, Johnsen H, Kleivi K, Zhou W, Tahmasebpour S, et al. Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Molecular oncology.* 2010;4(4):357-68.
69. Hannemann J, Velds A, Halfwerk JB, Kreike B, Peterse JL, van de Vijver MJ. Classification of ductal carcinoma in situ by gene expression profiling. *Breast cancer research : BCR.* 2006;8(5):R61.
70. Castro NP, Osorio CA, Torres C, Bastos EP, Mourao-Neto M, Soares FA, et al. Evidence that molecular changes in cells occur before morphological alterations during the progression of breast ductal carcinoma. *Breast cancer research : BCR.* 2008;10(5):R87.
71. Reeves GK, Pirie K, Green J, Bull D, Beral V, Million Women Study C. Comparison of the effects of genetic and environmental risk factors on in situ and invasive ductal breast cancer. *International journal of cancer Journal international du cancer.* 2012;131(4):930-7.
72. Institute NC. Study of Genetics in Patients With Ductal Breast Carcinoma In Situ and Healthy Volunteers.: National Cancer Institute; 2009 [updated 14/07/2009. Available from: <http://www.cancer.gov/clinicaltrials/search/view?cdrid=629681&version=HealthProfessional>
73. Couch FJ, Wang X, McGuffog L, Lee A, Olswold C, Kuchenbaecker KB, et al. Genome-wide association study in BRCA1 mutation carriers identifies novel loci associated with breast and ovarian cancer risk. *PLoS Genet.* 2013;9(3):e1003212.
74. Liu M, Ingle JN, Fridley BL, Buzdar AU, Robson ME, Kubo M, et al. TSPYL5 SNPs: association with plasma estradiol concentrations and aromatase expression. *Mol Endocrinol.* 2013;27(4):657-70.
75. Song C, Chen GK, Millikan RC, Ambrosone CB, John EM, Bernstein L, et al. A genome-wide scan for breast cancer risk haplotypes among African American women. *PLoS One.* 2013;8(2):e57298.
76. Hein R, Flesch-Janys D, Dahmen N, Beckmann L, Lindstrom S, Schoof N, et al. A genome-wide association study to identify genetic susceptibility loci that modify ductal and lobular postmenopausal breast cancer risk associated with menopausal hormone therapy use: a two-stage design with replication. *Breast cancer research and treatment.* 2013;138(2):529-42.
77. Rinella ES, Shao Y, Yackowski L, Pramanik S, Oratz R, Schnabel F, et al. Genetic variants associated with breast cancer risk for Ashkenazi Jewish women with strong family histories but no identifiable BRCA1/2 mutation. *Hum Genet.* 2013;132(5):523-36.
78. Rafiq S, Tapper W, Collins A, Khan S, Politopoulos I, Gerty S, et al. Identification of inherited genetic variations influencing prognosis in early-onset breast cancer. *Cancer Res.* 2013;73(6):1883-91.
79. Orr N, Lemnrau A, Cooke R, Fletcher O, Tomczyk K, Jones M, et al. Genome-wide association study identifies a common variant in RAD51B associated with male breast cancer risk. *Nat Genet.* 2012;44(11):1182-4.
80. Siddiq A, Couch FJ, Chen GK, Lindstrom S, Eccles D, Millikan RC, et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet.* 2012;21(24):5373-84.
81. Elgazzar S, Zembutsu H, Takahashi A, Kubo M, Aki F, Hirata K, et al. A genome-wide association study identifies a genetic variant in the SIAH2 locus associated with hormonal receptor-positive breast cancer in Japanese. *J Hum Genet.* 2012;57(12):766-71.

82. Chen F, Chen GK, Stram DO, Millikan RC, Ambrosone CB, John EM, et al. A genome-wide association study of breast cancer in women of African ancestry. *Hum Genet.* 2013;132(1):39-48.
83. Kim HC, Lee JY, Sung H, Choi JY, Park SK, Lee KM, et al. A genome-wide association study identifies a breast cancer risk variant in ERBB4 at 2q34: results from the Seoul Breast Cancer Study. *Breast cancer research : BCR.* 2012;14(2):R56.
84. Long J, Cai Q, Sung H, Shi J, Zhang B, Choi JY, et al. Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer. *PLoS Genet.* 2012;8(2):e1002532.
85. Shu XO, Long J, Lu W, Li C, Chen WY, Delahanty R, et al. Novel genetic markers of breast cancer survival identified by a genome-wide association study. *Cancer Res.* 2012;72(5):1182-9.
86. Kiyotani K, Mushiroda T, Tsunoda T, Morizono T, Hosono N, Kubo M, et al. A genome-wide association study identifies locus at 10q22 associated with clinical outcomes of adjuvant tamoxifen therapy for breast cancer patients in Japanese. *Hum Mol Genet.* 2012;21(7):1665-72.
87. Hindorff LA MJEBI, Morales J (European Bioinformatics Institute), Junkins HA, Hall PN, Klemm AK, and Manolio TA. A Catalog of Published Genome-Wide Association Studies. 2014 [Available from: <http://www.genome.gov/gwastudies>].
88. International HapMap C. The International HapMap Project. *Nature.* 2003;426(6968):789-96.
89. Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* 2014.
90. COSMIC. cosmic catalogue of somatic mutations in cancer (COSMIC): Wellcome Trust Sanger Institute; 2014 [Available from: cancer.sanger.ac.uk].
91. International HapMap C, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM, Gibbs RA, et al. Integrating common and rare genetic variation in diverse human populations. *Nature.* 2010;467(7311):52-8.
92. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet.* 2011;43(12):1210-4.
93. Cai Q, Long J, Lu W, Qu S, Wen W, Kang D, et al. Genome-wide association study identifies breast cancer risk variant at 10q21.2: results from the Asia Breast Cancer Consortium. *Hum Mol Genet.* 2011;20(24):4991-9.
94. Sehwat B, Sridharan M, Ghosh S, Robson P, Cass CE, Mackey JR, et al. Potential novel candidate polymorphisms identified in genome-wide association study for breast cancer susceptibility. *Hum Genet.* 2011;130(4):529-37.
95. Fletcher O, Johnson N, Orr N, Hosking FJ, Gibson LJ, Walker K, et al. Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. *Journal of the National Cancer Institute.* 2011;103(5):425-35.
96. Spraggs CF, Budde LR, Briley LP, Bing N, Cox CJ, King KS, et al. HLA-DQA1*02:01 is a major risk factor for lapatinib-induced hepatotoxicity in women with advanced breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology.* 2011;29(6):667-73.
97. Li J, Humphreys K, Darabi H, Rosin G, Hannelius U, Heikkinen T, et al. A genome-wide association scan on estrogen receptor-negative breast cancer. *Breast cancer research : BCR.* 2010;12(6):R93.

98. Gaudet MM, Kirchhoff T, Green T, Vijai J, Korn JM, Guiducci C, et al. Common genetic variants and modification of penetrance of BRCA2-associated breast cancer. *PLoS Genet.* 2010;6(10):e1001183.
99. Li J, Humphreys K, Heikkinen T, Aittomaki K, Blomqvist C, Pharoah PD, et al. A combined analysis of genome-wide association studies in breast cancer. *Breast cancer research and treatment.* 2011;126(3):717-27.
100. Antoniou AC, Wang X, Fredericksen ZS, McGuffog L, Tarrell R, Sinilnikova OM, et al. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nat Genet.* 2010;42(10):885-92.
101. Long J, Cai Q, Shu XO, Qu S, Li C, Zheng Y, et al. Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. *PLoS Genet.* 2010;6(6):e1001002.
102. Turnbull C, Ahmed S, Morrison J, Pernet D, Renwick A, Maranian M, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet.* 2010;42(6):504-7.
103. Azzato EM, Pharoah PD, Harrington P, Easton DF, Greenberg D, Caporaso NE, et al. A genome-wide association study of prognosis in breast cancer. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology.* 2010;19(4):1140-3.
104. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet.* 2009;41(5):579-84.
105. Zheng W, Long J, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet.* 2009;41(3):324-8.
106. Kibriya MG, Jasmine F, Argos M, Andrulis IL, John EM, Chang-Claude J, et al. A pilot genome-wide association study of early-onset breast cancer. *Breast cancer research and treatment.* 2009;114(3):463-77.
107. Gold B, Kirchhoff T, Stefanov S, Lautenberger J, Viale A, Garber J, et al. Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc Natl Acad Sci U S A.* 2008;105(11):4340-5.
108. Murabito JM, Rosenberg CL, Finger D, Kreger BE, Levy D, Splansky GL, et al. A genome-wide association study of breast and prostate cancer in the NHLBI's Framingham Heart Study. *BMC Med Genet.* 2007;8 Suppl 1:S6.
109. Easton DF, Pooley KA, Dunning AM, Pharoah PD, Thompson D, Ballinger DG, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature.* 2007;447(7148):1087-93.
110. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet.* 2007;39(7):870-4.
111. Stacey SN, Manolescu A, Sulem P, Rafnar T, Gudmundsson J, Gudjonsson SA, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet.* 2007;39(7):865-9.
112. Steinman S, Wang J, Bourne P, Yang Q, Tang P. Expression of cytokeratin markers, ER-alpha, PR, HER-2/neu, and EGFR in pure ductal carcinoma in situ (DCIS) and DCIS with co-existing invasive ductal carcinoma (IDC) of the breast. *Ann Clin Lab Sci.* 2007;37(2):127-34.
113. Mylonas I, Makovitzky J, Jeschke U, Briese V, Friesen K, Gerber B. Expression of Her2/neu, steroid receptors (ER and PR), Ki67 and p53 in invasive mammary ductal

carcinoma associated with ductal carcinoma In Situ (DCIS) Versus invasive breast cancer alone. *Anticancer Res.* 2005;25(3A):1719-23.

114. Hussein MR, Abd-Elwahed SR, Abdulwahed AR. Alterations of estrogen receptors, progesterone receptors and c-erbB2 oncogene protein expression in ductal carcinomas of the breast. *Cell Biol Int.* 2008;32(6):698-707.

115. Liu C, Zhang H, Shuang C, Lu Y, Jin F, Xu H, et al. Alterations of ER, PR, HER-2/neu, and P53 protein expression in ductal breast carcinomas and clinical implications. *Med Oncol.* 2010;27(3):747-52.

116. Azizun N, Bhurgri Y, Raza F, Kayani N. Comparison of ER, PR and HER-2/neu (C-erb B 2) reactivity pattern with histologic grade, tumor size and lymph node status in breast cancer. *Asian Pac J Cancer Prev.* 2008;9(4):553-6.

117. Pape-Zambito D, Jiang Z, Wu H, Devarajan K, Slater CM, Cai KQ, et al. Identifying a highly-aggressive DCIS subgroup by studying intra-individual DCIS heterogeneity among invasive breast cancer patients. *PLoS One.* 2014;9(6):e100488.

118. Hanley K, Wang J, Bourne P, Yang Q, Gao AC, Lyman G, et al. Lack of expression of androgen receptor may play a critical role in transformation from in situ to invasive basal subtype of high-grade ductal carcinoma of the breast. *Human pathology.* 2008;39(3):386-92.

119. Tamimi RM, Baer HJ, Marotti J, Galan M, Galaburda L, Fu Y, et al. Comparison of molecular phenotypes of ductal carcinoma in situ and invasive breast cancer. *Breast cancer research : BCR.* 2008;10(4):R67.

120. Wiechmann L, Kuerer HM. The molecular journey from ductal carcinoma in situ to invasive breast cancer. *Cancer.* 2008;112(10):2130-42.

121. Allred DC, Clark GM, Molina R, Tandon AK, Schnitt SJ, Gilchrist KW, et al. Overexpression of HER-2/neu and its relationship with other prognostic factors change during the progression of in situ to invasive breast cancer. *Human pathology.* 1992;23(9):974-9.

122. Bartova M, Ondrias F, Muy-Kheng T, Kastner M, Singer C, Pohlodek K. COX-2, p16 and Ki67 expression in DCIS, microinvasive and early invasive breast carcinoma with extensive intraductal component. *Bratisl Lek Listy.* 2014;115(7):445-51.

123. Kerlikowske K, Molinaro AM, Gauthier ML, Berman HK, Waldman F, Bennington J, et al. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma in situ diagnosis. *Journal of the National Cancer Institute.* 2010;102(9):627-37.

124. Shan M, Zhang X, Liu X, Qin Y, Liu T, Liu Y, et al. P16 and p53 play distinct roles in different subtypes of breast cancer. *PLoS One.* 2013;8(10):e76408.

125. Chae SW, Sohn JH, Kim DH, Choi YJ, Park YL, Kim K, et al. Overexpressions of Cyclin B1, cdc2, p16 and p53 in human breast cancer: the clinicopathologic correlations and prognostic implications. *Yonsei Med J.* 2011;52(3):445-53.

126. Radisky DC, Santisteban M, Berman HK, Gauthier ML, Frost MH, Reynolds CA, et al. p16(INK4a) expression and breast cancer risk in women with atypical hyperplasia. *Cancer Prev Res (Phila).* 2011;4(12):1953-60.

127. Glover JA, Hughes CM, Cantwell MM, Murray LJ. A systematic review to establish the frequency of cyclooxygenase-2 expression in normal breast epithelium, ductal carcinoma in situ, microinvasive carcinoma of the breast and invasive breast cancer. *British journal of cancer.* 2011;105(1):13-7.

128. Oliveira VM, Piatto S, Silva MA. Correlation of cyclooxygenase-2 and aromatase immunohistochemical expression in invasive ductal carcinoma, ductal carcinoma in situ, and adjacent normal epithelium. *Breast cancer research and treatment.* 2006;95(3):235-41.

129. Leo C, Faber S, Hentschel B, Hockel M, Horn LC. The status of cyclooxygenase-2 expression in ductal carcinoma in situ lesions and invasive breast cancer correlates to cyclooxygenase-2 expression in normal breast tissue. *Ann Diagn Pathol.* 2006;10(6):327-32.

130. Miki Mori GK, Loretta Chan, Denise Wolf, Yunn-Yi Chen, Karla Kerlikowske, Thea D. Tlsty, Jean-Philippe Coppe, Naoko Matsuda, Koyu Suzuki, Seigo Nakamura, Laura van't Veer. Predictive biomarkers of invasive carcinoma on excision in patients with a diagnosis of ductal carcinoma in situ of the breast by needle biopsy. 2014 American Society of Clinical Oncology Annual Meeting: J Clin Oncol 2014.
131. Quinn CM, Ostrowski JL, Harkins L, Rice AJ, Loney DP. Loss of bcl-2 expression in ductal carcinoma in situ of the breast relates to poor histological differentiation and to expression of p53 and c-erbB-2 proteins. *Histopathology*. 1998;33(6):531-6.
132. Van Slooten H, Van de Vijver M, Van de Velde C, Van Dierendonck J. Loss of Bcl-2 in invasive breast cancer is associated with high rates of cell death, but also with increased proliferative activity. *British journal of cancer*. 1998;77(5):789.
133. Zhang G, Kimijima I, Tsuchiya A, Abe R. The role of bcl-2 expression in breast carcinomas (Review). *Oncology reports*. 1998;5(5):1211-7.
134. Storm FK, Gilchrist KW, Warner TF, Mahvi DM. Distribution of Hsp-27 and HER-2/neu in in situ and invasive ductal breast carcinomas. *Ann Surg Oncol*. 1995;2(1):43-8.
135. O'Neill P, Shaaban A, West C, Dodson A, Jarvis C, Moore P, et al. Increased risk of malignant progression in benign proliferating breast lesions defined by expression of heat shock protein 27. *British journal of cancer*. 2004;90(1):182-8.
136. Ioachim E, Tsanou E, Briasoulis E, Batsis C, Karavasilis V, Charchanti A, et al. Clinicopathological study of the expression of hsp27, pS2, cathepsin D and metallothionein in primary invasive breast cancer. *Breast*. 2003;12(2):111-9.
137. Tang P, Wang X, Schiffhauer L, Wang J, Bourne P, Yang Q, et al. Expression patterns of ER- α , PR, HER-2/neu, and EGFR in different cell origin subtypes of high grade and non-high grade ductal carcinoma in situ. *Annals of Clinical & Laboratory Science*. 2006;36(2):137-43.
138. Biomarkers Definitions Working G. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95.
139. Biomedical F. NIH definition of biomarker. *Clin Pharmacol Ther*. 2001;69:89-95.
140. Provenzano E, Hopper JL, Giles GG, Marr G, Venter DJ, Armes JE. Biological markers that predict clinical recurrence in ductal carcinoma *in situ* of the breast. *European Journal of Cancer*. 39(5):622-30.
141. Page DL, Anderson TJ. *Diagnostic histopathology of the breast*: Churchill Livingstone; 1987.
142. Kayser K, Görtler J, Bogovac M, Bogovac A, Goldmann T, Vollmer E, et al. AI (artificial intelligence) in histopathology-from image analysis to automated diagnosis. *Folia Histochemica et Cytobiologica*. 2009;47(3):355.
143. Ripley PM, Rose IG, Kelley CA, Bigio IJ, Pickard D, Briggs G, et al. A comparison of Artificial Intelligence techniques for spectral classification in the diagnosis of human pathologies based upon optical biopsy. *Journal of Optical Society of America*. 2000:217-9.
144. Cireşan DC, Giusti A, Gambardella LM, Schmidhuber J, editors. Mitosis detection in breast cancer histology images with deep neural networks. *International Conference on Medical Image Computing and Computer-assisted Intervention*; 2013: Springer.
145. Tang P, Skinner KA, Hicks DG. Molecular classification of breast carcinomas by immunohistochemical analysis: are we ready? *Diagnostic Molecular Pathology*. 2009;18(3):125-32.
146. Porter D, Lahti-Domenici J, Keshaviah A, Bae YK, Argani P, Marks J, et al. Molecular Markers in Ductal Carcinoma in Situ of the Breast1 1 National Cancer Institute Cancer Genome Anatomy Project and Specialized Program in Research Excellence in Breast Cancer at Dana-Farber/Harvard Cancer Center (CA89393) and Johns Hopkins University

- (CA88843); Department of Defense Breast Cancer Center of Excellence Grants. Molecular Cancer Research. 2003;1(5):362-75.
147. Clark S, Warwick J, Carpenter R, Bowen R, Duffy SW, Jones JL. Molecular subtyping of DCIS: heterogeneity of breast cancer reflected in pre-invasive disease. British journal of cancer. 2011;104(1):120-7.
 148. Tsujimoto M, Nakabayashi K, Yoshidome K, Kaneko T, Iwase T, Akiyama F, et al. One-step nucleic acid amplification for intraoperative detection of lymph node metastasis in breast cancer patients. Clinical Cancer Research. 2007;13(16):4807-16.
 149. Tamaki Y, Akiyama F, Iwase T, Kaneko T, Tsuda H, Sato K, et al. Molecular detection of lymph node metastases in breast cancer patients: results of a multicenter trial using the one-step nucleic acid amplification assay. Clinical Cancer Research. 2009;15(8):2879-84.
 150. Wells WA, Carney PA, Eliassen MS, Grove MR, Tosteson AN. Pathologists' agreement with experts and reproducibility of breast ductal carcinoma-in-situ classification schemes. The American journal of surgical pathology. 2000;24(5):651-9.
 151. Silverstein MJ, Lagios MD, Craig PH, Waisman JR, Lewinsky BS, Colburn WJ, et al. A prognostic index for ductal carcinoma in situ of the breast. Cancer. 1996;77(11):2267-74.
 152. Scott MA, Lagios MD, Axelsson K, Rogers LW, Anderson TJ, Page DL. Ductal carcinoma in situ of the breast: reproducibility of histological subtype analysis. Human pathology. 1997;28(8):967-73.
 153. Holland R, Peterse JL, Millis RR, Eusebi V, Faverly D, van de Vijver MJ, et al. Ductal carcinoma in situ: a proposal for a new classification. Seminars in diagnostic pathology. 1994;11(3):167-80.
 154. Silverstein MJ, Lagios MD. Choosing treatment for patients with ductal carcinoma in situ: fine tuning the University of Southern California/Van Nuys Prognostic Index. Journal of the National Cancer Institute Monographs. 2010;2010(41):193-6.
 155. Sloane JP, Amendoeira I, Apostolikas N, Bellocq JP, Bianchi S, Boecker W, et al. Consistency achieved by 23 European pathologists in categorizing ductal carcinoma in situ of the breast using five classifications. European Commission Working Group on Breast Screening Pathology. Human pathology. 1998;29(10):1056-62.
 156. MacAusland SG, Hepel JT, Chong FK, Galper SL, Gass JS, Ruthazer R, et al. An attempt to independently verify the utility of the Van Nuys Prognostic Index for ductal carcinoma in situ. Cancer. 2007;110(12):2648-53.
 157. Boland GP, Chan KC, Knox WF, Roberts SA, Bundred NJ. Value of the Van Nuys Prognostic Index in prediction of recurrence of ductal carcinoma in situ after breast-conserving surgery. Br J Surg. 2003;90(4):426-32.
 158. de Mascarel I, Bonichon F, MacGrogan G, de Lara CT, Avril A, Picot V, et al. Application of the van nuys prognostic index in a retrospective series of 367 ductal carcinomas in situ of the breast examined by serial macroscopic sectioning: practical considerations. Breast cancer research and treatment. 2000;61(2):151-9.
 159. Asjoe FT, Altintas S, Huizing MT, Colpaert C, Marck EV, Vermorken JB, et al. The value of the Van Nuys Prognostic Index in ductal carcinoma in situ of the breast: a retrospective analysis. The breast journal. 2007;13(4):359-67.
 160. Early Breast Cancer Trialists' Collaborative G, Correa C, McGale P, Taylor C, Wang Y, Clarke M, et al. Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. Journal of the National Cancer Institute Monographs. 2010;2010(41):162-77.
 161. Dodwell D, Clements K, Lawrence G, Kearins O, Thomson CS, Dewar J, et al. Radiotherapy following breast-conserving surgery for screen-detected ductal carcinoma in

situ: indications and utilisation in the UK. Interim findings from the Sloane Project. *British journal of cancer*. 2007;97(6):725-9.

162. England PH. NHS breast screening programme and Association of Breast Surgery: An audit of screen detected breast cancers for the year of screening April 2014 to March 2015. London: Public Health England; 2016.

163. Lee LA, Silverstein MJ, Chung CT, Macdonald H, Sanghavi P, Epstein M, et al. Breast cancer-specific mortality after invasive local recurrence in patients with ductal carcinoma-in-situ of the breast. *Am J Surg*. 2006;192(4):416-9.

164. Schouten van der Velden AP, van Vugt R, Van Dijck JA, Leer JW, Wobbes T. Local recurrences after different treatment strategies for ductal carcinoma in situ of the breast: a population-based study in the East Netherlands. *Int J Radiat Oncol Biol Phys*. 2007;69(3):703-10.

165. MacDonald HR, Silverstein MJ, Mabry H, Moorthy B, Ye W, Epstein MS, et al. Local control in ductal carcinoma in situ treated by excision alone: incremental benefit of larger margins. *American journal of surgery*. 2005;190(4):521-5.

166. Goodwin A, Parker S, Ghera D, Wilcken N. Post-operative radiotherapy for ductal carcinoma in situ of the breast--a systematic review of the randomised trials. *Breast*. 2009;18(3):143-9.

167. Silverstein MJ, Lagios MD, Groshen S, Waisman JR, Lewinsky BS, Martino S, et al. The influence of margin width on local control of ductal carcinoma in situ of the breast. *N Engl J Med*. 1999;340(19):1455-61.

168. Dunne C, Burke JP, Morrow M, Kell MR. Effect of margin status on local recurrence after breast conservation and radiation therapy for ductal carcinoma in situ. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009;27(10):1615-20.

169. Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR, et al. Society of Surgical Oncology-American Society for Radiation Oncology-American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery with Whole-Breast Irradiation in Ductal Carcinoma In Situ. *Ann Surg Oncol*. 2016;23(12):3801-10.

170. NICE guidelines [CG80]. Early and locally advanced breast cancer: Diagnosis and treatment. NICE guidance. 2009.

171. Baso AoBSa. Surgical guidelines for the management of breast cancer. *European journal of surgical oncology: the journal of the European Society of Surgical Oncology and the British Association of Surgical Oncology*. 2009;35:1.

172. Soumian S, Verghese ET, Booth M, Sharma N, Chaudhri S, Bradley S, et al. Concordance between vacuum assisted biopsy and postoperative histology: implications for the proposed Low Risk DCIS Trial (LORIS). *Eur J Surg Oncol*. 2013;39(12):1337-40.

173. Dillon MF, McDermott EW, Quinn CM, O'Doherty A, O'Higgins N, Hill AD. Predictors of invasive disease in breast cancer when core biopsy demonstrates DCIS only. *J Surg Oncol*. 2006;93(7):559-63.

174. King TA, Farr GH, Jr., Cederbom GJ, Smetherman DH, Bolton JS, Stoller AJ, et al. A mass on breast imaging predicts coexisting invasive carcinoma in patients with a core biopsy diagnosis of ductal carcinoma in situ. *Am Surg*. 2001;67(9):907-12.

175. Fuhrman GM, Cederbom GJ, Bolton JS, King TA, Duncan JL, Champaign JL, et al. Image-guided core-needle breast biopsy is an accurate technique to evaluate patients with nonpalpable imaging abnormalities. *Ann Surg*. 1998;227(6):932-9.

176. Walters LL, Pang JC, Zhao L, Jorns JM. Ductal carcinoma in situ with distorting sclerosis on core biopsy may be predictive of upstage on excision. *Histopathology*. 2014.

177. Park SH, Kim MJ, Kim SJ, Kim EK. Ductal carcinoma in situ diagnosed using an ultrasound-guided 14-gauge core needle biopsy of breast masses: can underestimation be predicted preoperatively? *Ultrasonography*. 2014;33(2):128-35.
178. Chang WC, Hsu HH, Yu JC, Ko KH, Peng YJ, Tung HJ, et al. Underestimation of invasive lesions in patients with ductal carcinoma in situ of the breast diagnosed by ultrasound-guided biopsy: a comparison between patients with and without HER2/neu overexpression. *Eur J Radiol*. 2014;83(6):935-41.
179. Park AY, Gweon HM, Son EJ, Yoo M, Kim JA, Youk JH. Ductal carcinoma in situ diagnosed at US-guided 14-gauge core-needle biopsy for breast mass: preoperative predictors of invasive breast cancer. *Eur J Radiol*. 2014;83(4):654-9.
180. Goto M, Yuen S, Akazawa K, Nishida K, Konishi E, Kajihara M, et al. The role of breast MR imaging in pre-operative determination of invasive disease for ductal carcinoma in situ diagnosed by needle biopsy. *Eur Radiol*. 2012;22(6):1255-64.
181. Brennan ME, Turner RM, Ciatto S, Marinovich ML, French JR, Macaskill P, et al. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology*. 2011;260(1):119-28.
182. Galimberti V, Cole BF, Zurrada S, Viale G, Luini A, Veronesi P, et al. Axillary dissection versus no axillary dissection in patients with sentinel-node micrometastases (IBCSG 23-01): a phase 3 randomised controlled trial. *The Lancet Oncology*. 2013;14(4):297-305.
183. Donker M, van Tienhoven G, Straver ME, Meijnen P, van de Velde CJ, Mansel RE, et al. Radiotherapy or surgery of the axilla after a positive sentinel node in breast cancer (EORTC 10981-22023 AMAROS): a randomised, multicentre, open-label, phase 3 non-inferiority trial. *The Lancet Oncology*. 2014;15(12):1303-10.
184. Giuliano AE, Ballman KV, McCall L, Beitsch PD, Brennan MB, Kelemen PR, et al. Effect of Axillary Dissection vs No Axillary Dissection on 10-Year Overall Survival Among Women With Invasive Breast Cancer and Sentinel Node Metastasis: The ACOSOG Z0011 (Alliance) Randomized Clinical Trial. *JAMA*. 2017;318(10):918-26.
185. Claus EB, Stowe M, Carter D. Breast carcinoma in situ: risk factors and screening patterns. *Journal of the National Cancer Institute*. 2001;93(23):1811-7.
186. Lyman GH, Giuliano AE, Somerfield MR, Benson AB, 3rd, Bodurka DC, Burstein HJ, et al. American Society of Clinical Oncology guideline recommendations for sentinel lymph node biopsy in early-stage breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2005;23(30):7703-20.
187. Shah DR, Canter RJ, Khatri VP, Bold RJ, Martinez SR. Utilization of lymph node assessment in patients with ductal carcinoma in situ treated with lumpectomy. *J Surg Res*. 2012;177(1):e21-6.
188. Fisher B, Costantino J, Redmond C, Fisher E, Margolese R, Dimitrov N, et al. Lumpectomy compared with lumpectomy and radiation therapy for the treatment of intraductal breast cancer. *N Engl J Med*. 1993;328(22):1581-6.
189. Wapnir IL, Dignam JJ, Fisher B, Mamounas EP, Anderson SJ, Julian TB, et al. Long-term outcomes of invasive ipsilateral breast tumor recurrences after lumpectomy in NSABP B-17 and B-24 randomized clinical trials for DCIS. *Journal of the National Cancer Institute*. 2011;103(6):478-88.
190. Emdin SO, Granstrand B, Ringberg A, Sandelin K, Arnesson LG, Nordgren H, et al. SweDCIS: Radiotherapy after sector resection for ductal carcinoma in situ of the breast. Results of a randomised trial in a population offered mammography screening. *Acta Oncol*. 2006;45(5):536-43.
191. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Eval Health Prof*. 2002;25(1):76-97.

192. Group AaNBSA. NHS Breast Screening Programme and Association of Breast Surgery. An Audit of Screen Detected Breast Cancers for the Year of Screening April 2012 to March 2013. Public Health England. 2014.
193. Julien JP, Bijker N, Fentiman IS, Peterse JL, Delledonne V, Rouanet P, et al. Radiotherapy in breast-conserving treatment for ductal carcinoma in situ: first results of the EORTC randomised phase III trial 10853. EORTC Breast Cancer Cooperative Group and EORTC Radiotherapy Group. *Lancet*. 2000;355(9203):528-33.
194. Bijker N, Peterse JL, Duchateau L, Robanus-Maandag EC, Bosch CA, Duval C, et al. Histological type and marker expression of the primary tumour compared with its local recurrence after breast-conserving therapy for ductal carcinoma in situ. *British journal of cancer*. 2001;84(4):539-44.
195. Ringberg A, Nordgren H, Thorstensson S, Idvall I, Garmo H, Granstrand B, et al. Histopathological risk factors for ipsilateral breast events after breast conserving treatment for ductal carcinoma in situ of the breast—results from the Swedish randomised trial. *European Journal of Cancer*. 2007;43(2):291-8.
196. Holmberg L, Garmo H, Granstrand B, Ringberg A, Arnesson L-G, Sandelin K, et al. Absolute Risk Reductions for Local Recurrence After Postoperative Radiotherapy After Sector Resection for Ductal Carcinoma In Situ of the Breast. *Journal of Clinical Oncology*. 2008;26(8):1247-52.
197. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *The Lancet Oncology*. 2011;12(1):21-9.
198. Staley H, McCallum I, Bruce J. Postoperative Tamoxifen for ductal carcinoma in situ: Cochrane systematic review and meta-analysis. *Breast*. 2014;23(5):546-51.
199. NICE. Early and locally advanced breast cancer: diagnosis and treatment. NICE. 2009(CG80):60-1.
200. Cobleigh MA, SJ.; Julian, TB.; Siziopikou, KP.; Arthur, DW.; Rabinovitch, R.; Zheng, P.; Mamounas, EP.; Wolmark, N., . NSABP B-43: A phase III clinical trial to compare trastuzumab (T) given concurrently with radiation therapy (RT) to RT alone for women with HER2+ DCIS resected by lumpectomy (Lx) Cancer Research. 2012;72(24).

Chapter 2: Motivation and Aims

2.1. Motivation

The previous chapter has highlighted the areas of uncertainty in the natural history and clinical management of DCIS. The general aim of this thesis is to provide insight into some of these uncertainties, particularly regarding risk factors for DCIS, its optimal treatment and its long-term prognosis. Specifically, this thesis has the following four aims:

2.2. Aims

(i) Examine to what extent the risk factors for DCIS are similar to, or differ from, those for IBC.

To this end, chapter 3 will:

(a) Investigate how DCIS incidence varies with different modifiable and non-modifiable factors when compared with women without cancer;

(b) Investigate how IBC incidence varies with the same risk factors compared with women without cancer;

(c) Compare the risk factors for IBC directly with those for DCIS.

(ii) Investigate factors influencing the likelihood of a preoperative biopsy diagnosis of DCIS being upgraded to a higher grade of DCIS, or upstaged to IBC, upon pathological examination of the surgically excised specimen.

To this end, chapter 4 will:

(a) Identify the pathological and radiological factors associated with upstaging a biopsy of DCIS to IBC upon pathological analysis of the resected specimen.

(b) Evaluate to what extent these factors differ from those for upgrading a biopsy of DCIS to higher grade of DCIS upon pathological analysis of the resected specimen.

(iii) Explore the long-term risk of ipsilateral IBC in women diagnosed with screen-detected DCIS for the whole of England.

To this end, chapter 5 will:

(a) Characterise the risk of developing IBC among women diagnosed with screen-detected DCIS;

(b) Characterise the risk of death from breast cancer in women diagnosed with screen-detected DCIS;

(c) Identify the patient, tumour and treatment related factors influencing the rate of subsequent IBC in women diagnosed with screen-detected DCIS.

Whereas the previous three aims examine the risk factors for developing DCIS, accuracy of its initial diagnosis, and factors associated with the occurrence of IBC following treatment for DCIS respectively, they have not referred to examining rates of recurrence of DCIS. The reason for this is that much less information is available on DCIS recurrence, or on recurrence more generally even after IBC. In light of this, this thesis will examine whether recurrence information can be identified using the wealth of information stored in routinely collected datasets by using IBC as a model, with the intention that the process may be expanded to include DCIS at a later stage, in future studies. Specifically, this thesis will:

(iv) Investigate whether routinely collected data sources can be used to accurately identify breast cancer recurrences in women registered with IBC in England.

To this end, chapter 6 will:

(a) Collate all routinely collected information on breast cancer that is stored within Public Health England;

(b) Use this information to develop an algorithm to identify which women have had a recurrence of their breast cancer and, for those who did, the date of the recurrence;

(c) Examine the accuracy of this algorithm in terms of its sensitivity and specificity.

Chapter 3: Differences in Risk Factors for Ductal Carcinoma in Situ and Invasive Breast Cancer: Prospective Cohort Study

3.1. Background

Historically, the risk factors for DCIS have been thought to be the same as for invasive breast cancer (IBC). However, there has been little evidence supporting this assumption, and only a few studies have examined directly whether the magnitude of the effect from shared risk factors is similar in DCIS and IBC or whether it differs. Previous studies that have aimed to compare risk factors between DCIS and IBC have been either constrained by data on a limited number of variables (71, 201) or confounded by recall bias (202).

Underlying genetic changes are responsible for both DCIS and IBC, and recent research into identifying women at risk of developing either disease has focused on the presence or absence of various genetic factors (203, 204). However, emerging evidence suggests that the risk of disease development from common cancer related genetic variants is substantially influenced by modifiable risk factors (49). Understanding similarities and differences in these risk factors between women who are diagnosed with DCIS or IBC with those without either condition may provide insights into the process of invasive transformation from normal ductal epithelium.

This study aimed: firstly, to investigate how DCIS incidence varies with different modifiable and non-modifiable factors when compared with women without cancer; secondly, to investigate how IBC incidence varies with the same risk factors compared with women without cancer; and thirdly, to compare the risk factors for IBC directly with those for DCIS.

3.2. Methods

This study was carried out using data from UK Biobank, which is a prospective cohort study of 502,649 participants, recruited between 2006 and 2010 at ages 40–69 years. Baseline assessments were conducted at 22 research centres located across the United Kingdom (205). Participants completed touch-screen questionnaires to provide information on lifestyle, socioeconomic and medical factors. Detailed information on the methodology, available variables and data completeness is available at (206). Through linkage to national cancer registries and Hospital Episodes Statistics (HES) data, UK Biobank currently has information on the incidence of both DCIS and IBC amongst participants up to September 2016, and this is further updated on an annual basis. The present study has been conducted under UK Biobank Application Number 11091 (Appendix A), and is covered by the generic ethical approval for UK Biobank studies from the NHS National Research Ethics Service (approval dated 17th June 2011, Ref 11/NW/0382).

3.2.1. Participants

Data were available for all women enrolled into UK Biobank. However, the overwhelming majority of diagnoses of DCIS occur as a result of screening and are in women aged 50-69 years. Therefore the study was restricted to women aged 50-69 years. Women with diagnoses of DCIS, IBC or any other invasive cancer, except non-melanoma skin cancer (NMSC, ICD-10: C44 (207)), prior to enrolment into UK Biobank were excluded.

3.2.2. Variables

Variables examined included: age at menarche (≤ 11 , 12-13, 14+ years), number of live births (nulliparous, 1, 2, 3, 4+), age at menopause (<45, 45-49, 50-54, 55+ years) or pre-menopausal, age at first live birth (<19, 20-29, 30-39, 40+ years) or nulliparous, body mass index (BMI) (<25, 25-30, 30-35, 35+ kg/m²), smoking status (never, previous, current), alcohol intake (never, less than daily, daily), ever used hormone replacement therapy (HRT) (yes, no), night shift work (regular, sometimes, never), daily vegetable consumption ($\frac{1}{4}$, $\frac{1}{2}$, 1, 2 or 3+), daily fruit consumption (0, <1, 1, 2, 3+) and oral contraceptive pill (OCP) use (never, previous, current).

3.2.3. Statistical methods

Three types of analyses were carried out: firstly, the relationship between the rate of development of DCIS and the variables listed above was examined; secondly, the relationship between the rate of development of IBC and these same variables was examined; and thirdly the effect of each variable on the rate of development of DCIS and the rate of development of IBC was compared directly. For all three analyses, rate ratios were estimated using Poisson regression with participant age as the underlying time variable. Women contributed person years from the date of consent to UK Biobank and were followed-up until the earliest of date of diagnosis of either IBC or DCIS (ICD-10 codes D051 and C50, respectively), date of diagnosis of another type of invasive cancer (except for NMSC, ICD-10 C44), death, or 31st December 2015. The endpoints of interest in the three analyses were: (i) first diagnosis of DCIS censoring for IBC, other cancers and deaths; (ii) first diagnosis of IBC, censoring for DCIS, other cancers and deaths; (iii) first diagnosis of either IBC or DCIS, censoring for other cancers and deaths.

Analyses were stratified by quintiles of socioeconomic group based on the Townsend deprivation index (208), current age (5 year categories), number of years since last attended mammographic screening, and oral contraceptive pill (OCP) use (never, previous, current). These four stratification variables formed the ‘basic stratification’. We also studied the effect of risk factors using full stratification, which included, as well as the basic stratification factors, all the other variables being considered in the categories described above. Women with unknown values for any of the stratification variables were assigned to a separate category for that variable. Due to the previously documented influence of HRT use on the association between BMI on breast cancer risk (209), analyses of BMI and age at menopause were undertaken only on women who had never taken HRT. The effect of BMI on cancer risk was also studied separately in pre- and postmenopausal women.

3.3. Results

Overall, 273,632 women were included in the study population. The mean length of follow-up for all participants was 5.1 years (SD 1.01). A total of 2772 women were diagnosed during follow-up with IBC as their first cancer. A further 456 women were diagnosed with DCIS as their first cancer. The mean age at diagnosis amongst women with incident IBC was 54.8 years (SD 8.85) and amongst women with incident DCIS was 56.5 years (SD 8.85).

3.3.1. Risk factors for incident IBC

The rate ratios for developing IBC according to the modifiable and non-modifiable risk factors for which information was available following basic stratification are shown in Figure 3.1. Figure 3.2 gives the same analysis but including full stratification. For several variables, similar trends were seen in both analyses. There was strong evidence of an increased IBC rate with increasing BMI amongst postmenopausal women who had not used HRT (p for trend 0.00074) (Figure 3.2). Compared with women with BMI $<25 \text{ kg/m}^2$, women with BMI of $>35 \text{ kg/m}^2$ had a substantially increased rate of IBC (RR 1.57, 95% CI 1.17–2.12) (Figure 3.2). There was no such trend between IBC rate and high BMI ($>35 \text{ kg/m}^2$) amongst premenopausal women (p for trend 0.38) (Figure 3.2).

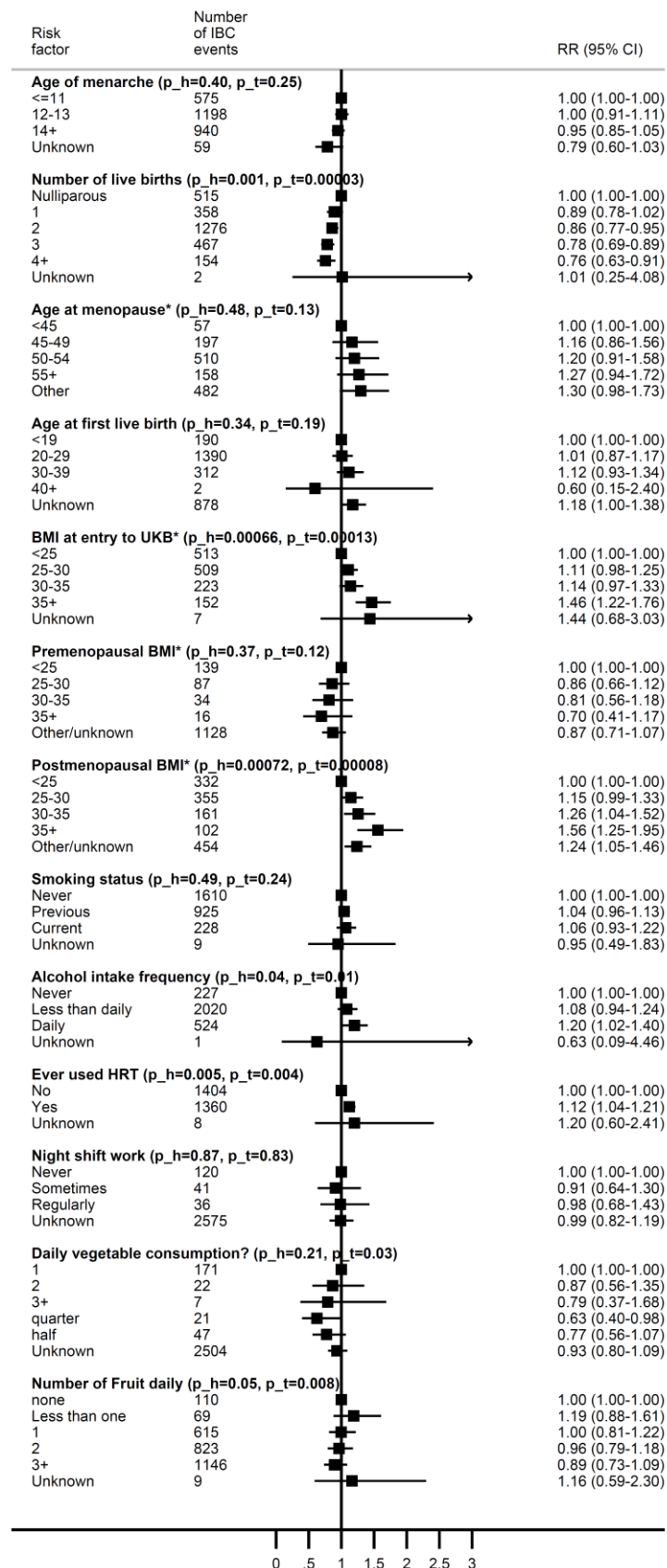


Figure 3.1: The association between different factors and the incidence of invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following basic stratification (see Methods Section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m^2), CI = Confidence Interval, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

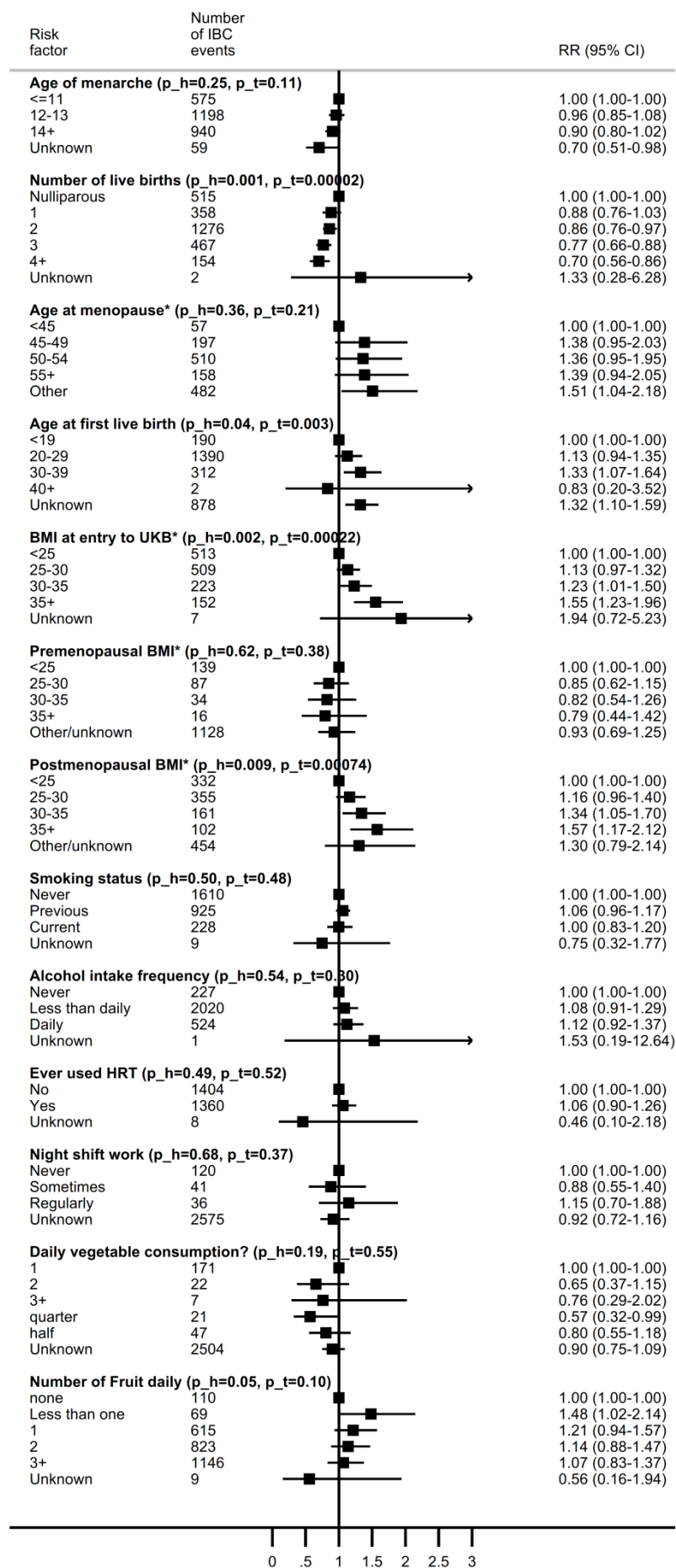


Figure 3.2: The association between different factors on the incidence of invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following full stratification (see methods section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m^2), CI = Confidence Interval, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

The IBC rate also increased with increasing age at first live birth ($p = 0.003$) and the IBC rate amongst women giving birth when aged 30-39 years was increased by a third compared to women giving birth aged ≤ 19 years (RR 1.33, 95% CI 1.07–1.64) (Figure 3.2). Compared to women giving birth aged ≤ 19 years, the IBC rate was lower in women aged 40+ at the time of first live birth (RR 0.83, 95% CI 0.20–3.52). However, there were only 2 women with IBC in the 40+ group and the confidence intervals for this estimate were wide suggesting this estimate is consistent with random variation. There was a reduction in the IBC rate with the number of live births (p for trend = 0.00002). Biparous women were at a 14% lower IBC rate when compared to nulliparous women (RR 0.86, 95% CI 0.76–0.97), whereas the rate for women with four or more children was lower by 30% compared with nulliparous women (RR 0.70, 95% CI 0.56–0.86) (Figure 3.2).

Several further risk factors were associated with a woman's rate of IBC in the analysis using the basic stratification, but not with the full stratification. Women who drank alcohol daily had IBC rate 20% higher than compared to those who did not drink alcohol (RR 1.20, 95% CI 1.02–1.40) and those who had used HRT had a 12% higher IBC rate compared to those who had not, using basic stratification (RR 1.12, 95% CI 1.04–1.21) (Figure 3.1). The magnitude of both of these factors was attenuated and the association became non-significant with full stratification (RR 1.12, 95% CI 0.92–1.37 and RR 1.06, 95% CI 0.90–1.26 respectively) (Figure 3.2). Similarly, there was a trend in the IBC rate with the consumption of fruits ($p < 0.01$) and vegetables ($p = 0.03$) using the basic stratification (Figure 3.1), however these too became non-significant upon full stratification ($p = 0.10$ and $p = 0.55$, respectively) (Figure 3.2).

No significant trend in IBC rate was observed with age at menarche ($p = 0.11$), age at menopause ($p = 0.21$), smoking status ($p = 0.48$) or with night shift work ($p = 0.37$), with either the basic or full stratification (Figure 3.2).

3.3.2. Risk factors for incident DCIS

The influence of the factors discussed above on the DCIS rate following basic and full stratification is shown in Figures 3.3 and 3.4 respectively. For several variables, similar trends were seen in both analyses. The DCIS rate was lower in women with more live births (p for trend < 0.00004) (Figure 3.4). When compared with nulliparous women, the DCIS rate amongst women with four or more live births was more than halved (RR 0.40, 95% CI 0.21–0.76) (Figure 3.4).

There was strong evidence of an increased DCIS rate with increasing BMI amongst postmenopausal women who had not used HRT (p for trend 0.00021) (Figure 3.4). Compared with women with BMI $< 25 \text{ kg/m}^2$, women with BMI of $> 35 \text{ kg/m}^2$ had more than a threefold increase in DCIS rate (RR 3.70, 95% CI 1.92–7.11) (Figure 3.4). There was no such trend between DCIS rate and high BMI ($> 35 \text{ kg/m}^2$) amongst pre-menopausal women (p for trend 0.13). No significant associations were seen in any of the other factors studied (Figure 3.4).

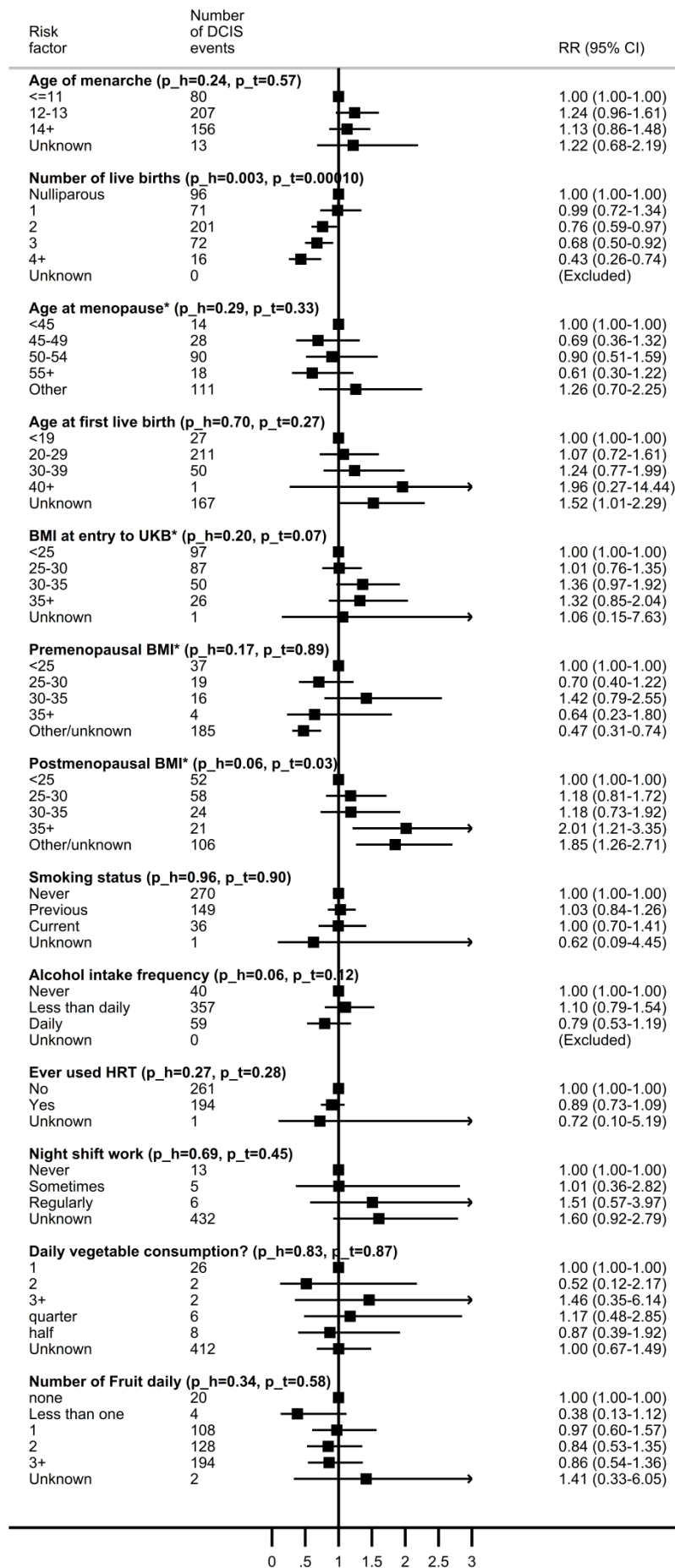


Figure 3.3: The association between different factors on the incidence of ductal carcinoma in situ, as first cancer, in women enrolled into UK Biobank, following basic stratification (see methods section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m²), CI = Confidence Interval, DCIS = Ductal Carcinoma in Situ, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

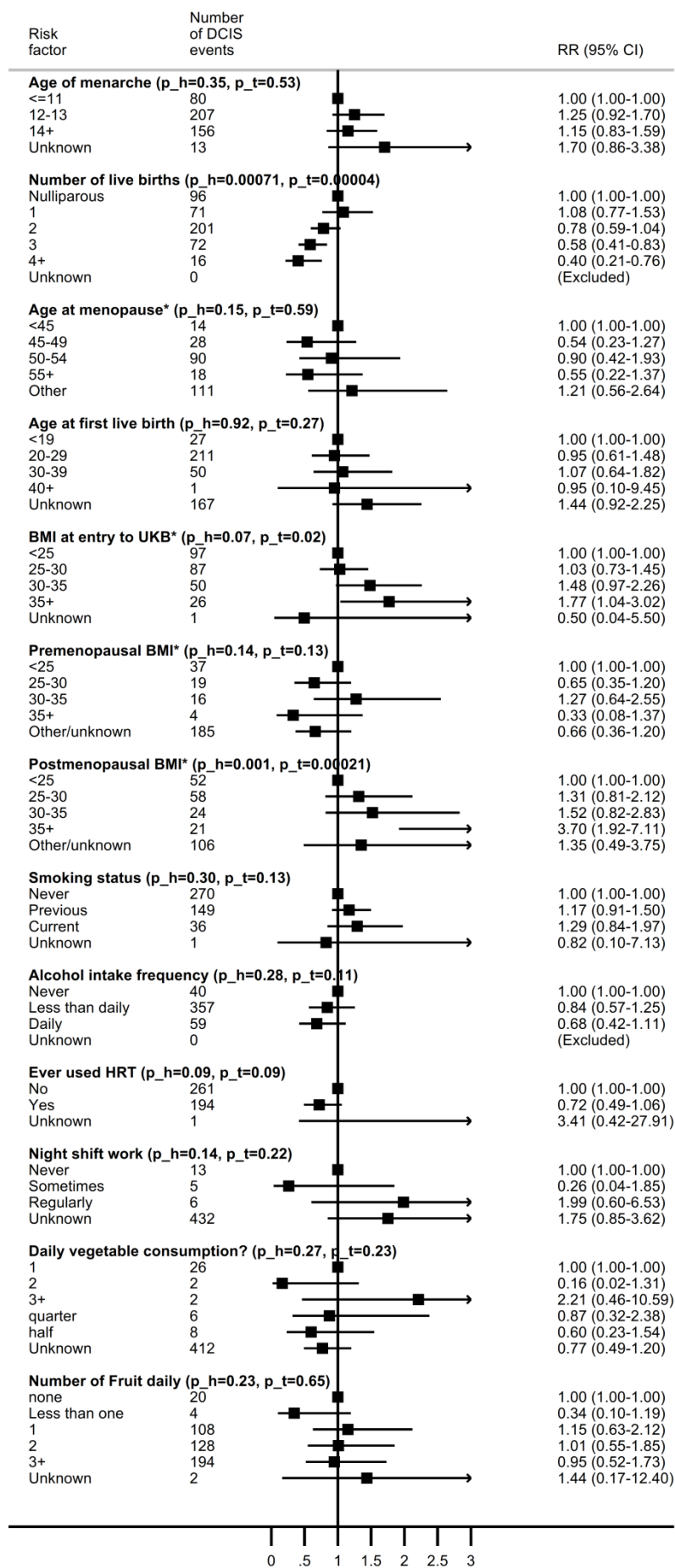


Figure 3.4: The association between different factors on the incidence of ductal carcinoma in situ, as first cancer, in women enrolled into UK Biobank, following full stratification (see methods section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m^2), CI = Confidence Interval, DCIS = Ductal Carcinoma in Situ, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

3.3.3. Comparison of IBC and DCIS

Figure 3.5 compares the effect of the different factors in DCIS and IBC with basic stratification, and Figure 3.6 shows the same analyses with full stratification. For most reproductive and environmental risk factors shown in Figure 3.6, the pattern of association did not differ between DCIS and IBC. However, for postmenopausal BMI the association was much stronger for DCIS than for IBC. Compared with women with BMI $<25\text{kg/m}^2$, the increase in DCIS rate for women with BMI $\geq 35\text{ kg/m}^2$ was more than twice the increase seen in IBC rate (RR 2.35, 95% CI 1.14–4.82). There was also a larger reduction in disease rate associated with each additional live birth live birth for DCIS than for IBC ($p=0.03$).

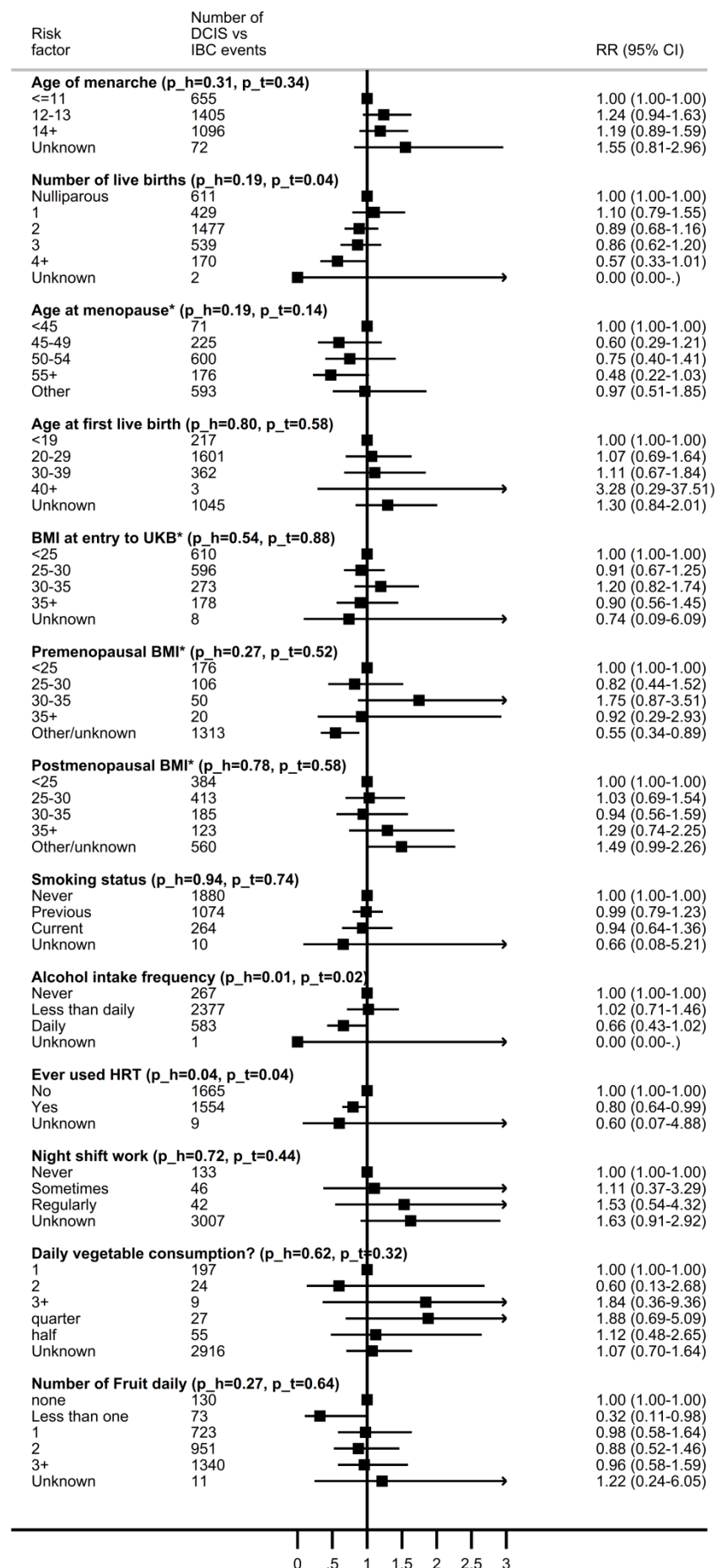


Figure 3.5: The association between different factors on the incidence of ductal carcinoma in situ compared with invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following basic stratification (see methods section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m^2), CI = Confidence Interval, DCIS = Ductal Carcinoma in Situ, IBC = Invasive Breast Cancer, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

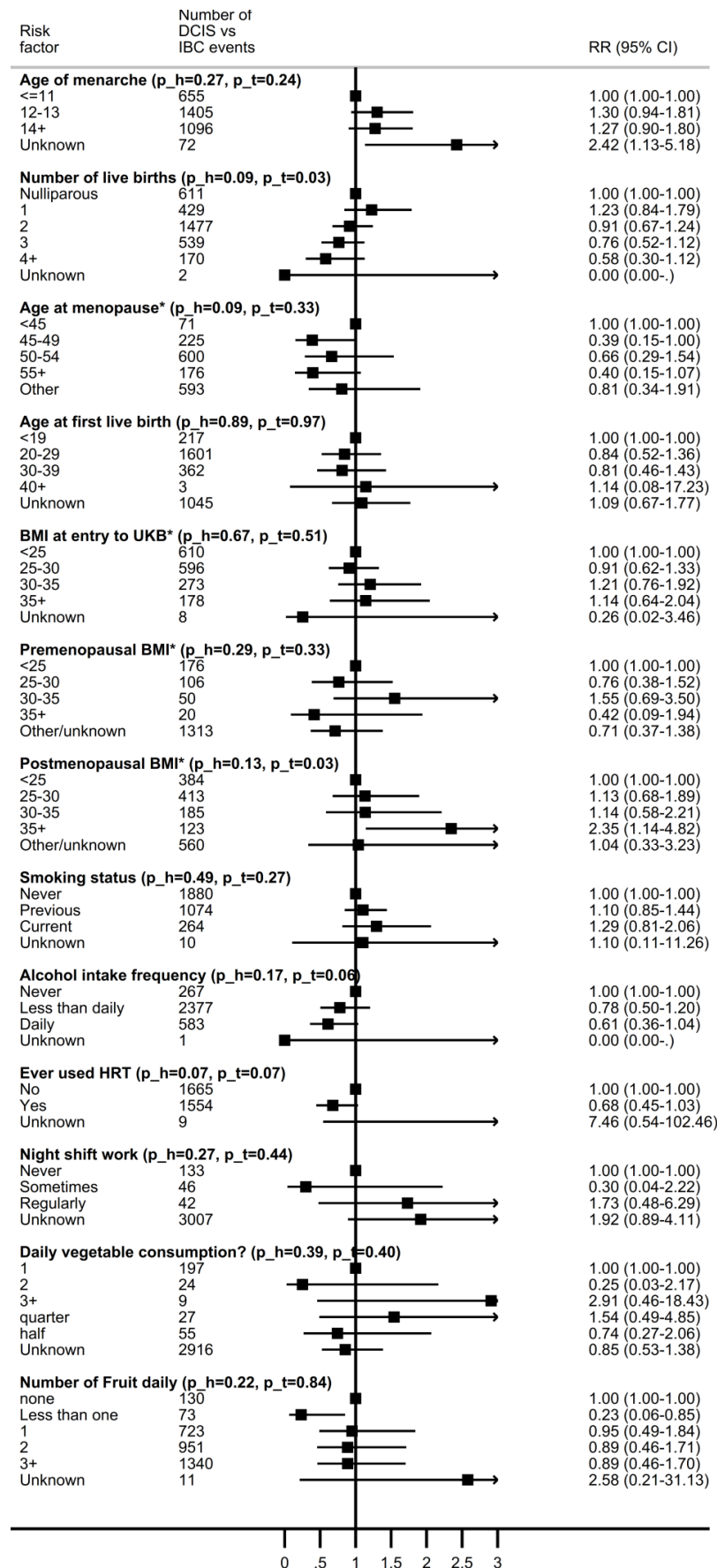


Figure 3.6: The association between different factors on the incidence of ductal carcinoma in situ compared with invasive breast cancer, as first cancer, in women enrolled into UK Biobank, following full stratification (see methods section for stratification variables). Abbreviations: BMI = Body Mass Index (kg/m^2), CI = Confidence Interval, DCIS = Ductal Carcinoma in Situ, IBC = Invasive Breast Cancer, HRT = Hormone Replacement Therapy, IBC = Invasive Breast Cancer, p_h = p value for heterogeneity, p_t = p value for trend, RR = Rate Ratio, UKB = UK Biobank, *Amongst never users of HRT.

3.4. Discussion

This study is amongst the most detailed comparisons of risk factors for DCIS and IBC to date and contributes to the literature in this area. For both conditions, postmenopausal BMI was associated with an increased risk of developing disease, and the number of live births was associated with a decreased risk of developing disease. However, the magnitude of the effect differed between DCIS and IBC. The increased rate from postmenopausal BMI ≥ 35 kg/m² was larger for DCIS than for IBC (RR 2.35, 95% CI 1.14–4.82), and the trend of reduction in rate with each additional live birth was greater for DCIS than for IBC (p for trend = 0.03).

3.4.1. Body mass index

The identification of risk factors for cancer is of limited relevance unless individuals have the ability to influence their risk of developing disease and its subsequent progression by modifying such factors. Obesity is an example of one such modifiable risk factor that this study has identified as having a strong relationship with the development of both diseases amongst postmenopausal women. It has long been hypothesised that obesity may increase levels of circulating oestradiol due to its biosynthesis within adipose tissue, which in turn may promote cell growth in susceptible tumours. To date, over a hundred studies have demonstrated an association between BMI and IBC in postmenopausal women (209-211). However, few studies have investigated the risk of high BMI specifically on DCIS risk. Furthermore, those studies that have examined this have reported conflicting findings, with some authors finding no association (71, 212, 213), others concluding a negative association (214) and yet others reporting that DCIS risk increases with increasing BMI (215). Many of these studies use data from SEER, and thereby acknowledge the limitations of this data source (216). The results of the present study show strong evidence of an increased DCIS rate

with increasing BMI amongst postmenopausal women who had not used HRT (p for trend 0.00021) (Figure 3.4).

To date, only one previous study has directly compared the influence of high BMI on the rate of developing DCIS with the rate of developing IBC (71). It showed that increasing BMI was positively associated with the risk of IBC, but not with the risk of DCIS (RRs per 5 kg/m² increase in BMI = 1.20 for IBC and 1.01 for DCIS, respectively; p-value for heterogeneity = 0.002). This finding led the study's authors to conclude that the lack of association between BMI and DCIS suggested a greater influence of BMI on disease progression to invasive potential than in its initiation (71). However, findings from the present study challenge this assertion. In the present study, the effect of high BMI (BMI >35 kg/m²) on the rate of disease amongst postmenopausal women was two-fold larger for incident DCIS than it was for incident IBC (ratio of relative risks 2.35 (95% CI 1.14 – 4.82)). If this effect is causal, this suggests that BMI may, conversely, have more of a role in the initiation of the carcinogenic process from ductal atypia to *in situ* disease, than it does on its subsequent progression to IBC.

3.4.2. Number of live births

The association between number of live births and breast cancer has been suspected as early as the 18th century (217). There is now strong evidence that multiparous women have a lower risk of breast cancer than their nulliparous counterparts (218), and that increased exposure to endogenous oestrogen levels increases the risk of breast cancer and also drives breast cancer progression (219). However, the hormonal mechanisms underlying this process are not yet fully understood. Our results show higher numbers of live births decreases the risk of both DCIS and IBC and is consistent with previously published studies examining this relationship

(213, 220). However, our findings show, for the first time, that there is a stronger association with DCIS than IBC ($p=0.03$). The reason for this is not yet clear. Whilst the possibility of this being attributable to a chance finding cannot be completely excluded, it is possible that the hormonal milieu during pregnancy may be more influential in protecting against the initial carcinogenic *in situ* changes amongst ductal cells than in their subsequent invasive transformation.

3.4.3. Dietary factors

Although previous studies have found interesting associations between certain dietary factors and breast cancer incidence (221-225), the influence of dietary and related lifestyle factors has been difficult to study and therefore reliably reported dietary associations poses considerable challenges (226). There are problems with recall bias in case-control studies, while cohort studies often suffer from short-term follow-up, imperfect quantification of dietary intake or exercise as well as selection bias (227). In the present study, alcohol intake, HRT use, as well as fruit and vegetable consumption were all associated with IBC rate in the analysis using the basic stratification (Figure 3.1), but not with the full stratification (Figure 3.2). The reason for this finding is likely due to strong correlation between dietary and lifestyle risk factors. As a result, whilst individually they may significantly influence IBC rate, this effect is attenuated when correlated factors are stratified for.

3.4.4. Limitations

The present study had the advantage of avoiding recall bias by establishing risk factors at entry into the cohort. Whilst this is less of a limitation when considering factors such as age at first live birth or parity, it is of vital importance when considering dietary and lifestyle

factors such as smoking or alcohol consumption. Whilst the depth of information collected within UK Biobank allows the study of many different exposures with disease, it should be noted that the low response rate (5.5%) at recruitment suggests selection bias, which should be given thought when considering whether findings presented in this study are generalizable to the general population.

Most DCIS cases are treated with surgical excision to minimise the risk of progression to invasive cancer. However, even after adequate treatment of DCIS, some women will develop a local occurrence of IBC within the same breast. Well recognised risk factors of DCIS recurrence or occurrence of IBC after surgery include resected margin width, histological subtype, nuclear grade, molecular phenotype and the absence of adjuvant radiotherapy (53). Whether the risk of local IBC occurrence can be influenced by modifying certain lifestyle or behavioural factors following DCIS treatment has been a question of keen interest amongst patients and clinicians alike, and one which has been poorly addressed so far. Our study was underpowered to evaluate risk factors for subsequent IBC diagnosis amongst the 456 women with DCIS. However, it is expected that sufficient numbers of diagnoses in the coming years, as these data continue to accrue, will permit such analyses.

3.4.5. Conclusion

This study is one of the most detailed comparisons of risk factors for DCIS and IBC to date. Our results confirm similarities in the risk factors for these two conditions, but reveal differences in the magnitude of their effect. By doing so, these findings provide unique insights into the contribution of BMI and endocrine factors associated with parity at different stages in the process of carcinogenesis.

3.5. References

1. Reeves GK, Pirie K, Green J, Bull D, Beral V, Million Women Study C. Comparison of the effects of genetic and environmental risk factors on in situ and invasive ductal breast cancer. *International journal of cancer Journal international du cancer*. 2012;131(4):930-7.
2. Barclay J, Ernster V, Kerlikowske K, Grady D, Sickles EA. Comparison of Risk Factors for Ductal Carcinoma In Situ and Invasive Breast Cancer. *JNCI: Journal of the National Cancer Institute*. 1997;89(1):76-82.
3. Veal CT, Hart V, Lakoski SG, Hampton JM, Gangnon RE, Newcomb PA, et al. Health-related behaviors and mortality outcomes in women diagnosed with ductal carcinoma in situ. *J Cancer Surviv*. 2017;11(3):320-8.
4. Cowell CF, Weigelt B, Sakr RA, Ng CKY, Hicks J, King TA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. *Molecular oncology*. 2013;7(5):859-69.
5. Petridis C, Brook MN, Shah V, Kohut K, Gorman P, Caneppele M, et al. Genetic predisposition to ductal carcinoma in situ of the breast. *Breast cancer research*. 2016;18(1):22.
6. Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, et al. Evidence of gene-environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet*. 2013;9(3):e1003284.
7. Allen N, Sudlow C, Downey P, Peakman T, Danesh J, Elliott P, et al. UK Biobank: Current status and what it means for epidemiology. *Health Policy and Technology*. 1(3):123-6.
8. Biobank U. UK Biobank 2018 [Available from: <http://www.ukbiobank.ac.uk>].
9. WHO. ICD-10 Version: 2016 World Health Organisation; 2016 [Available from: <http://apps.who.int/classifications/icd10/browse/2016/en>].
10. Townsend P PP, Beattie A. Health and deprivation: Inequality and the North. London: Croom Helm; 1987.
11. Lahmann PH, Hoffmann K, Allen N, van Gils CH, Khaw KT, Tehard B, et al. Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). *International journal of cancer Journal international du cancer*. 2004;111(5):762-71.
12. Reeves GK, Pirie K, Beral V, Green J, Spencer E, Bull D, et al. Cancer incidence and mortality in relation to body mass index in the Million Women Study: cohort study. *BMJ*. 2007;335(7630):1134.
13. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371(9612):569-78.
14. Gapstur SM, Morrow M, Sellers TA. Hormone replacement therapy and risk of breast cancer with a favorable histology: Results of the iowa women's health study. *JAMA*. 1999;281(22):2091-7.
15. Kerlikowske K, Barclay J, Grady D, Sickles EA, Ernster V. Comparison of risk factors for ductal carcinoma in situ and invasive breast cancer. *Journal of the National Cancer Institute*. 1997;89(1):76-82.
16. Weiss H, Brinton L, Brogan D, Coates R, Gammon M, Malone K, et al. Epidemiology of in situ and invasive breast cancer in women aged under 45. *British journal of cancer*. 1996;73(10):1298.
17. Kerlikowske K, Walker R, Miglioretti DL, Desai A, Ballard-Barbash R, Buist DS. Obesity, mammography use and accuracy, and advanced breast cancer risk. *Journal of the National Cancer Institute*. 2008;100(23):1724-33.

18. Park HS, Lloyd S, Decker RH, Wilson LD, Yu JB. Limitations and Biases of the Surveillance, Epidemiology, and End Results Database. *Current Problems in Cancer*.36(4):216-24.
19. Franco G. Bernardino Ramazzini and women workers' health in the second half of the XVIIth century. *J Public Health (Oxf)*. 2012;34(2):305-8.
20. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50 302 women with breast cancer and 96 973 women without the disease. *The Lancet*. 2002;360(9328):187-95.
21. Stewart. B.W.; Wild CP. *World Cancer Report 2014*. Lyon, France: International Agency for Research on Cancer; 2014.
22. Wohlfahrt J, Rank F, Kroman N, Melbye M. A comparison of reproductive risk factors for CIS lesions and invasive breast cancer. *International Journal of Cancer*. 2004;108(5):750-3.
23. Hamajima N, Hirose K, Kungu A. Alcohol, tobacco and breast cancer--collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. 2002.
24. Gandini S, Merzenich H, Robertson C, Boyle P. Meta-analysis of studies on breast cancer risk and diet. *European Journal of Cancer*.36(5):636-46.
25. Liu XO, Huang YB, Gao Y, Chen C, Yan Y, Dai HJ, et al. Association between dietary factors and breast cancer risk among Chinese females: systematic review and meta-analysis. *Asian Pac J Cancer Prev*. 2014;15(3):1291-8.
26. Do MH, Lee SS, Jung PJ, Lee MH. Intake of fruits, vegetables, and soy foods in relation to breast cancer risk in Korean women: a case-control study. *Nutr Cancer*. 2007;57(1):20-7.
27. Gao Y, Huang YB, Liu XO, Chen C, Dai HJ, Song FJ, et al. Tea consumption, alcohol drinking and physical activity associations with breast cancer risk among Chinese females: a systematic review and meta-analysis. *Asian Pac J Cancer Prev*. 2013;14(12):7543-50.
28. Mannu GSD, S. *Cancer Epidemiology*. In: Poston GJ, Wyld L, Audisio RA, editors. *Surgical Oncology: Theory and Multidisciplinary Practice*, Second Edition: CRC Press; 2016.
29. Willett WCK, T.; Romieu, I. Diet, obesity, and physical activity. In: Stewart_B.W._and_Wild_C.P., editor. *World Cancer Report 2014*. Lyon, France: International Agency for Research on Cancer, World Health Organization; 2014. p. 124-34.
30. Benson JR, Wishart GC. Predictors of recurrence for ductal carcinoma in situ after breast-conserving surgery. *The Lancet Oncology*. 2013;14(9):e348-57.

Chapter 4: The accuracy of preoperative biopsy in Ductal Carcinoma in Situ

4.1. Background

Chapter 1 has described how the introduction of breast screening programmes in many countries has led to a dramatic increase in the numbers of women diagnosed with DCIS. Overdiagnosis, and thus overtreatment, are inevitable consequences of screening programmes and result in some women with DCIS potentially receiving treatment for a condition that, in the absence of screening, may never have come to light or become clinically significant during their lifetime (228). Clinical research in the field of DCIS management has focused on finding ways to help separate those screen-detected DCIS lesions that require active treatment from those that can be safely left under surveillance, and only treated if there is a change in their characteristics (229). Whereas it is widely accepted that high grade DCIS falls into the former category (230), the likelihood of low and intermediate grade DCIS to progress to IBC is less clear and therefore is the focus of trials evaluating a non-operative approach to management of women with lower risk DCIS (231, 232).

Clinical trials randomising women with low and intermediate grade DCIS to surgery or to observation alone rely upon an accurate biopsy diagnosis and determination of pathological grade in order to exclude women with a higher risk of disease progression from inclusion in the trial. This is clearly of vital importance since any occult IBC or high grade DCIS that is missed at the time of biopsy may subsequently go undetected, and therefore untreated until later, by which time the patient's prognosis may be poorer. However, much of the existing literature on the accuracy of pre-operative biopsy in DCIS has focused on identifying factors related to missing the presence of IBC on biopsy, rather than factors related to missing the presence higher grade DCIS on biopsy. The reason for this is that, prior to the introduction of trials of non-operative treatment for low-risk DCIS, DCIS grade has not influenced clinical management in the same way that the presence of IBC has done. In cases where DCIS is

detected by biopsy but IBC is subsequently detected following pathological analysis of the surgically excised breast specimen, women may need to undergo a second surgical procedure to evaluate the status of their axillary nodes (170, 233). This has clear financial cost, carries risk of medical complications and is also accompanied by psychological distress.

Whereas previous studies have reported the proportion of DCIS biopsies missing the presence of IBC as approximately one in four (181), few of them have been careful to exclude biopsies that were performed because IBC was already strongly suspected and little is known regarding the underestimation rate of DCIS grade. In addition, despite its importance to ongoing clinical research, no studies have directly compared the factors associated with biopsies that are upstaged to IBC with those that are upgraded to a higher grade of DCIS. The present study aimed to evaluate the proportion of preoperative biopsies showing DCIS that were subsequently upstaged to IBC and upgraded to higher grade DCIS following pathological analysis of the surgically excised specimen at a large tertiary referral centre, and to directly compare the factors associated with upstaging or upgrading a biopsy diagnosis of DCIS following surgery.

4.2. Methods

4.2.1. Data sources

Women were selected through the nationwide registry of histology and cytopathology records in the Netherlands (PALGA) and through the regional tumour registry at the Netherlands Cancer Institute (NKI). The NKI is a large specialist cancer centre. It has a large pathology database, which contains information on the initial biopsy diagnosis of every woman referred to its breast cancer service, as well as the final pathological diagnosis of each woman with DCIS or IBC treated at the NKI, allowing a retrospective cohort study to be conducted.

4.2.2. Participants

All women diagnosed with DCIS who either had their initial biopsy performed at the NKI or were referred for a second opinion at the NKI, between 2000 and 2014 were identified. Women who had evidence of IBC on initial biopsy, received neoadjuvant chemotherapy or where there was documented clinical suspicion of the presence of (ipsilateral or contralateral) IBC or lymph node invasion at the time of biopsy were excluded. Additionally, women who had encapsulated papillary carcinoma, LCIS or other non-DCIS pathology on biopsy or who had subsequent surgical excision performed at a different hospital were also excluded.

4.2.3. Variables

Radiology and pathology data from the time of their initial biopsy, and pathology data from the analysis of their surgically excised specimen were abstracted, in order to identify women whose biopsy result was subsequently upstaged to IBC or upgraded to a higher grade of DCIS following surgery. The NKI uses an integrated electronic patient record system.

Detailed information on clinical, radiological and pathological variables on all women was

collected directly from the patient's medical records using this system (Appendices B-D). The variables collected from the clinical data included whether the lesion was symptomatic or screen-detected lesion, whether symptoms were present at time of diagnosis, menopausal status and date of birth.

The variables collected from the pathology data included presence of invasive cancer, age at diagnosis, laterality of tumour, pathological grade of biopsy, presence of necrosis on biopsy, type of surgery performed and pathological size of DCIS. If more than one biopsy showing DCIS was taken from the same area, the highest grade was recorded. Missing values were replaced by blinded re-examination of a sample of the raw pathology slides by two expert pathologists (JW and EG, Appendix M). Variables collected from the radiology data included size of the lesion on mammogram and method of biopsy (9G vacuum-assisted or 14G core needle biopsy). Missing values were replaced by blinded re-examination of a sample of the original mammograms slides by two experienced radiologists (GW and JT, Appendix M)

4.2.4. Statistical methods

Exact chi-squared (χ^2) test was used to compare patient, tumour and radiological characteristics between women upstaged or upgraded with those with no change (or downgraded). Further analyses were conducted to directly compare risk factors for upstaging and upgrading. All analyses were undertaken using IBM Stata version 13.0 (Texas, USA).

4.3. Results

The clinical, radiological and pathological information from 714 women diagnosed with DCIS between 2000 and 2014 was abstracted in this study. Of these, 125 were removed due to meeting one of the exclusion criteria for the study (Figure 4.1), leaving information on 589 women available for analysis.

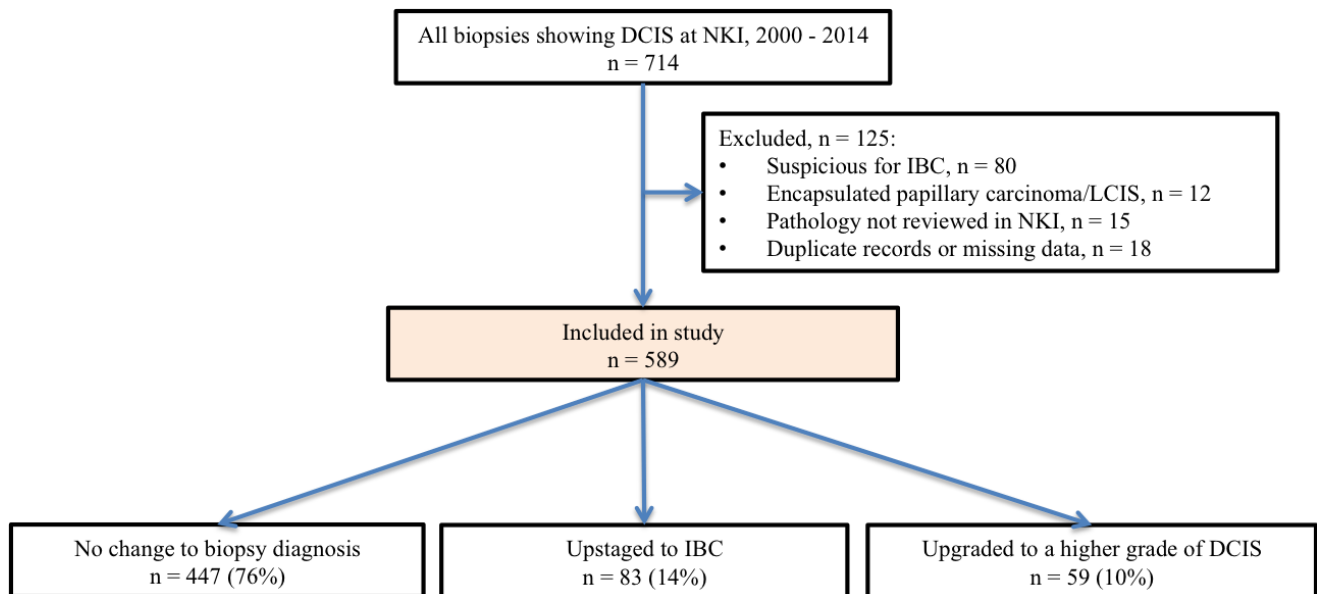


Figure 4.1: Derivation of the population used in the analysis. Abbreviations: NKI = Netherlands Cancer Institute Antoni van Leeuwenhoek hospital, DCIS = Ductal carcinoma in situ, LCIS: Lobular carcinoma in situ, IBC = Invasive Breast Cancer

Participant characteristics are shown in Table 4.1. The number of women diagnosed with DCIS on biopsy at the NKI/AVL increased during the study period from 73 women diagnosed between 2000 – 2004, to 197 women diagnosed between 2005 – 2009, and increased again to 319 women diagnosed between 2010 – 2014, as service capacity has increased. Overall, 32.1% women diagnosed with DCIS were aged 20 – 49 years, 40.8% were 50 – 59 years, 19.9% were 60 – 69 years, and 7.3% were aged 70+ years. The majority of DCIS cases were screen-detected (54.2%). A further 23.4% were detected during routine follow-up for genetic predisposition/family history or another reason, 11.7% presented symptomatically and 10.4% were referred for a second opinion from a different hospital.

Method of presentation was unknown for 2 (0.3%) women. 39.2% women were diagnosed using a 9G vacuum-assisted biopsy and 21.2% were diagnosed using a 14G core needle biopsy. The method of biopsy was unknown for 39.6% women. 25.5% women had low grade on biopsy, 42.8% had intermediate grade and 29.9% had high grade. DCIS grade on biopsy was unknown in 1.9% of cases. All women underwent surgical excision with approximately half receiving breast conserving surgery (52.8%) and approximately half receiving mastectomy (47.2%) (Table 4.1).

Table 4.1: Characteristics of women diagnosed with ductal carcinoma in situ on initial biopsy.

Characteristics	N	(%)
Calendar year of diagnosis		
2000-2004	73	(12.4)
2005-2009	197	(33.5)
2010-2014	319	(54.1)
Age at diagnosis (years)		
20-49	189	(32.1)
50-59	240	(40.8)
60-69	117	(19.9)
70+	43	(7.3)
Method of presentation		
Symptomatic	69	(11.7)
Screening	319	(54.2)
Follow-up	138	(23.4)
Second opinion	61	(10.4)
Unknown	2	(0.3)
Method of biopsy		
Core needle biopsy (14G)	125	(21.2)
Vacuum Assisted Biopsy (9G)	231	(39.2)
Unknown	233	(39.6)
Grade on initial biopsy		
Low	150	(25.5)
Intermediate	252	(42.8)
High	176	(29.9)
Unknown	11	(1.9)
Type of curative surgery		
BCS	219	(37.2)
BCS+SLNB	92	(15.6)
Mastectomy	75	(12.7)
Mastectomy+SLNB	199	(33.8)
Mastectomy+SLNB+AC	4	(0.7)
Total number of women:	589	(100)

Abbreviations: BCS: breast conserving surgery, SLNB: sentinel lymph node biopsy, AC: axillary clearance.

The biopsy results from 142 (24.1%) women were found to have underestimated disease severity following examination of the surgically excised specimen (Table 4.2). Of these, 83 (14.1%) women had IBC present in the final excised specimen that was not present on the biopsy (i.e. were upstaged) and 59 (10%) women had a higher grade of DCIS present in the final excised specimen compared with their initial biopsy (i.e. were upgraded).

Table 4.2: Final diagnoses in women with ductal carcinoma in situ on biopsy by various characteristics.

Characteristic	Final diagnosis following evaluation of surgically excised specimen										
	Total number of women		No Change		Upstage to IBC		Upstage vs no change	DCIS with higher grade		Upgrade vs no change	Upgrade vs upstage
	n	(%)	n	(%)	n	(%)	P value	n	(%)	P value	P value
Age at diagnosis (years)											
20-49	189	(100)	136	(72.0)	32	(16.9)	0.45	21	(11.1)	0.11	0.24
50-59	240	(100)	180	(75.0)	31	(12.9)		29	(12.1)		
60-69	117	(100)	96	(82.1)	16	(13.6)		5	(4.3)		
70+	43	(100)	35	(81.4)	4	(9.3)		4	(9.3)		
Method of Biopsy											
9G VAC	231	(100)	195	(84.4)	24	(10.4)	0.08	12	(5.2)	0.003	0.32
14G Core needle	125	(100)	87	(69.6)	20	(16)		18	(14.4)		
Unknown	233	-	165	-	39	-		29	-		
Pathological lesion size (mm)											
0-19	195	(100)	171	(87.7)	11	(5.6)	<0.001	13	(6.7)	0.003	0.67
20-50	175	(100)	115	(65.7)	34	(19.4)		26	(14.9)		
50+	104	(100)	69	(66.3)	19	(18.3)		16	(15.4)		
Unknown	115	-	92	-	19	-		4	-		
Mammographic lesion size (mm)											
0-19	175	(100)	146	(83.4)	14	(8.0)	0.01	15	(8.6)	0.04	0.20
20-50	140	(100)	103	(73.6)	25	(17.8)		12	(8.6)		
50+	95	(100)	63	(66.4)	16	(16.8)		16	(16.8)		
Unknown	179	-	135	-	28	-		16	-		
Method of presentation											
Symptomatic	69	(100)	53	(76.8)	12	(17.4)	0.43	4	(5.8)	0.70	0.33
Screen-detected	319	(100)	247	(77.4)	38	(11.9)		34	(10.7)		
Follow-up	138	(100)	101	(73.2)	24	(17.4)		13	(9.4)		
Second opinion	61	(100)	45	(73.8)	9	(14.7)		7	(11.5)		
Unknown	2	-	1	-	0	-		1	-		
Presence of symptoms at presentation											
Symptoms	143	(100)	103	(72.0)	29	(20.3)	0.02	11	(7.7)	0.52	0.03
No symptoms	219	(100)	171	(78.1)	23	(10.5)		25	(11.4)		
Unknown	227	-	173	-	31	-		23	-		
Presence of palpable lump											
Yes	112	(100)	80	(71.4)	21	(18.8)	0.10	11	(9.8)	0.89	0.47
No	439	(100)	343	(78.1)	54	(12.3)		42	(9.6)		
Unknown	38	-	24	-	8	-		6	-		
Menopausal status											
Premenopausal	207	(100)	152	(73.4)	32	(15.5)	0.49	23	(11.1)	0.50	1
Postmenopausal	295	(100)	229	(77.6)	39	(13.2)		27	(9.2)		
Unknown	87	-	66	-	12	-		9	-		
Lesion laterality											
Left	300	(100)	230	(76.7)	47	(15.7)	0.46	23	(7.7)	0.10	0.06
Right	289	(100)	217	(75.0)	36	(12.5)		36	(12.5)		
Unknown	0	-	0	-	0	-		0	-		
Immunohistochemistry performed on biopsy											
No	281	(100)	223	(79.4)	40	(14.2)	1	18	(6.4)	0.18	0.28
Yes	107	(100)	81	(75.7)	14	(13.1)		12	(11.2)		
Unknown	201	-	143	-	29	-		29	-		
Presence of necrosis on biopsy											
No	357	(100)	282	(79.0)	47	(13.2)	0.12	28	(7.8)	0.008	0.40
Yes	204	(100)	139	(68.1)	35	(17.2)		30	(14.7)		
Unknown	28	-	26	-	1	-		1	-		
Total	589	(100)	447	(75.9)	83	(14.1)		59	(10.0)		

Abbreviations: VAC: vacuum assisted biopsy, DCIS: ductal carcinoma in situ, IBC: invasive breast cancer, Hereditary: family history related screening mammography, SLNB: sentinel lymph node biopsy, AC: axillary clearance.

There were clinical and radiological differences between the biopsies that were upstaged or upgraded and those that were not. Fewer 9G vacuum-assisted biopsies than 14G core biopsies were upgraded (5.2% vs 14.4%, p for difference: 0.003), while the effect of biopsy method on upstaging was not significant (10.4% vs 16%, p for difference: 0.08) (Table 4.2). Lesions with pathological size 20+ mm were more likely to be both upstaged (p for difference: <0.001) and upgraded (p for difference: 0.003). Lesions with mammographic size 20+ mm were also more likely to be both upstaged (p for difference: 0.01) and upgraded (p for difference: 0.04).

The presence of symptoms increased the risk of upstaging (20.3% vs 10.5%, p for difference: 0.02), but not of upgrading (7.7% vs 11.4%), p for difference: 0.52) and the difference between the presence of symptoms on upstaging and upgrading was significantly different (p: 0.03). The presence of necrosis on biopsy increased the risk of upgrading (14.7% vs 7.8%, p for difference: 0.008) but the association between necrosis and upstaging was not significant (17.2% vs 13.2%, p for difference: 0.12). However, age at diagnosis of DCIS, method of presentation, the presence of a palpable lump, menopausal status, lesion laterality, and use of immunohistochemistry on biopsy were not significantly associated with upstaging or upgrading (Table 4.2).

4.4. Discussion

The accuracy of biopsy in determining DCIS grade is the single most important factor upon which the contemporary paradigm shift towards non-operative management of low-risk DCIS is based. The present study is one of the largest to date to evaluate risk factors for upgrading and upstaging preoperative biopsies in women diagnosed with DCIS. It has demonstrated that larger lesion size is significantly associated with both upstaging and upgrading of DCIS on preoperative biopsy. The presence of symptoms was significantly associated with upstaging but not with upgrading, whereas the use of core biopsy instead of vacuum-assisted biopsy and the presence of necrosis on biopsy were significantly associated with upgrading of DCIS, but not with upstaging. Consequently, this study highlights factors to aid in the risk stratification of women poised towards a non-operative approach.

The association between larger DCIS size on mammography and upstaging to IBC or upgrading to a higher grade of DCIS may be due to the heterogeneous growth patterns seen in larger DCIS lesions (70). Whereas invasive disease spreads through the ductal basement membrane, DCIS lesions grow along the milk duct with a branched process of growth, with emerging and competing cell-lines of low, intermediate and high-grade disease often resulting in a diverse cell-line environment (120, 234) (see section 1.1.3). Thus, the larger the area of disease, the higher likelihood of IBC or more than one grade of disease being present that may not have been sampled during biopsy and only subsequently detected following pathological analysis of the surgically excised specimen (235). The presence of clinical symptoms was significantly associated with upstaged biopsies, but not with upgraded biopsies. The reason for this was unclear and there was little information on the nature of symptoms available to allow this finding to be examined further.

There was strong evidence that fewer 9G vacuum-assisted biopsies than 14G core biopsies were upgraded (p for difference: 0.003), while the effect of biopsy method on upstaging was not significant (p for difference: 0.08). Pre-operative biopsy for breast cancer has conventionally been undertaken by radiologically guided core-needle (14G) biopsy, and in more recent years, by vacuum-assisted (9G) breast biopsy. The accuracy of either of these two methods is dependent on both human factors and technical factors. Human factors represent not only the level of skill on behalf of the radiologist responsible for taking an appropriate sample that is representative of the entire lesion, but also the experience of the pathologist in accurately interpreting the biopsied tissue. These human factors, whilst difficult to objectively quantify, have been previously discussed in the literature (236). In addition to human factors, certain technical factors also play a substantial role in influencing the overall accuracy of the breast biopsy (181). These include both the resolution of imaging-guidance and the type of biopsy device used. 9G vacuum-assisted biopsy devices obtain more tissue than 14G core-biopsy devices, increasing the likelihood of a representative tissue specimen being obtained (237). This is the likely reason for the lower number of 9G vacuum-assisted biopsies upgraded in the present study, compared with 14G core biopsies.

Upgraded lesions also were more likely to have necrosis present on biopsy when compared with lesions with no change in diagnosis following surgery. Necrosis in the DCIS biopsy tissue indicates that the cancer cells in the sampled area have died. The pathological finding of necrosis often accompanies rapid cell turnover, where the rate of growth may have outstripped the blood supply to a central portion of the cancer. The presence of this finding on biopsy, similar to a large tumour size, should raise the suspicion of a higher grade of DCIS being present.

Although our study was limited to a single centre, work is currently on going to validate our findings using data from other centres. If the results of the present study can be validated in other cohorts, they could provide recommendations for refining the selection criteria for studies evaluating a non-operative approach (Appendices E and F). In recent years, rapid developments in genomics and candidate gene analysis of tumour tissue have meant that focus has started to move away from the phenotypic characteristics such as nuclear grade as the chief determining factors of DCIS prognosis, and towards biomarkers for cell proliferation instead. A summary of these candidate biomarkers is shown in Table 1.5 (page 22). On-going research building on the present study is exploring whether the presence of these biomarkers may further improve the diagnostic accuracy of pre-operative biopsy to detect occult high-risk lesions.

4.4.1. Conclusion

Over the years there has been a trend towards less invasive surgical treatment of breast cancer. The coming decade will herald the publication of findings from on-going randomised trials of non-operative treatment in DCIS and by doing so will determine whether a further paradigm shift in its treatment is necessary. Subsequent interpretation and implementation of the findings from these trials must take into account the underlying factors determining the accuracy of pathological grade on biopsy in order to consider the safest transition towards possible non-operative management for a low risk group.

4.5. References

1. Jones JL. Overdiagnosis and overtreatment of breast cancer: Progression of ductal carcinoma in situ: the pathological perspective. *Breast Cancer Research*. 2006;8(2):204.
2. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *European Journal of Cancer*.51(16):2296-303.
3. Wallis MG, Clements K, Kearins O, Ball G, Macartney J, Lawrence GM. The effect of DCIS grade on rate, type and time to recurrence after 15 years of follow-up of screen-detected DCIS. *British journal of cancer*. 2012;106(10):1611-7.
4. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, Dif N, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ – The LORD study. *European Journal of Cancer*.51(12):1497-510.
5. Francis A, Fallowfield L, Rea D. The LORIS Trial: Addressing Overtreatment of Ductal Carcinoma In Situ. *Clinical oncology*.27(1):6-8.
6. NICE guidelines [CG80]. Early and locally advanced breast cancer: Diagnosis and treatment. NICE guidance. 2009.
7. van Roozendaal LM, Goorts B, Klinkert M, Keymeulen KB, De Vries B, Strobbe LJ, et al. Sentinel lymph node biopsy can be omitted in DCIS patients treated with breast conserving therapy. *Breast cancer research and treatment*. 2016;156(3):517-25.
8. Brennan ME, Turner RM, Ciatto S, Marinovich ML, French JR, Macaskill P, et al. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology*. 2011;260(1):119-28.
9. Castro NP, Osorio CA, Torres C, Bastos EP, Mourao-Neto M, Soares FA, et al. Evidence that molecular changes in cells occur before morphological alterations during the progression of breast ductal carcinoma. *Breast cancer research : BCR*. 2008;10(5):R87.
10. Thomson JZ, Evans AJ, Pinder SE, Burrell HC, Wilson AR, Ellis IO. Growth pattern of ductal carcinoma in situ (DCIS): a retrospective analysis based on mammographic findings. *Br J Cancer*. 2001;85(2):225-7.
11. Pape-Zambito D, Jiang Z, Wu H, Devarajan K, Slater CM, Cai KQ, et al. Identifying a highly-aggressive DCIS subgroup by studying intra-individual DCIS heterogeneity among invasive breast cancer patients. *PLoS One*. 2014;9(6):e100488.
12. Lee CW, Wu HK, Lai HW, Wu WP, Chen ST, Chen DR, et al. Preoperative clinicopathologic factors and breast magnetic resonance imaging features can predict ductal carcinoma in situ with invasive components. *Eur J Radiol*. 2016;85(4):780-9.
13. Wells WA, Carney PA, Eliassen MS, Grove MR, Tosteson AN. Pathologists' agreement with experts and reproducibility of breast ductal carcinoma-in-situ classification schemes. *The American journal of surgical pathology*. 2000;24(5):651-9.
14. Philpotts LE, Shaheen NA, Carter D, Lange RC, Lee CH. Comparison of rebiopsy rates after stereotactic core needle biopsy of the breast with 11-gauge vacuum suction probe versus 14-gauge needle and automatic gun. *American Journal of Roentgenology*. 1999;172(3):683-7.

Chapter 5: Population-based Study of Invasive Breast Cancer in Women Diagnosed with DCIS

5.1. Background

Approximately 4,000 women with screen-detected DCIS undergo surgical treatment each year in the UK (238). However, little is known about the long-term prognosis of these women following their treatment. Although formal reports from the National Health Service Breast Screening Programme (NHSBSP) audit are published on an annual basis, these reports focus on describing the tumour and treatment characteristics of women diagnosed only in a single year. To date there has been no population-level publication of the long-term risk of IBC in women diagnosed with screen-detected DCIS in England, nor of how it has changed over time or how it depends on the characteristics of the patient, the tumour, or the treatment received.

Whereas randomised trials of radiotherapy and hormone therapy have identified the benefits of these treatments in women with DCIS (160, 197, 239), it remains unclear as to what extent these benefits are carried over to women diagnosed with screen-detected DCIS in the general population. Critical examination of long-term data, such as that held by Public Health England (PHE) and the NHSBSP audit could help provide valuable insights into the prognosis of women with DCIS under different treatment strategies. Furthermore, describing how the rate of development of IBC varies with the time since treatment for DCIS could help to identify the optimal frequency and period of surveillance following treatment. In addition, comparing the rate of development of IBC for different tumour characteristics could contribute to the current understand of risk factors for IBC after DCIS (50, 53, 58), and thereby also help to shed light on which women should be considered for less extensive treatment, whilst enabling those at higher risk also to be identified.

This study examined women diagnosed with DCIS via the NHSBSP and aimed to: (i) characterise the risk of developing subsequent IBC among women diagnosed with screen-detected DCIS; (ii) characterise the risk of death from breast cancer in women diagnosed with screen-detected DCIS; and (iii) identify the patient, tumour and treatment related factors influencing the rate of subsequent IBC in women diagnosed with screen-detected DCIS.

5.2. Methods

5.2.1. Introduction to data sources

The NHSBSP was introduced in the UK in 1988 and currently provides three-yearly mammograms for women between 50 and 70 years of age. Prior to 2000, screening records were collected and stored locally at individual screening centres. From 2000 onwards, information for all women diagnosed with screen-detected breast cancer was collected centrally on an annual basis as part of the NHSBSP audit and it is currently stored within PHE. The resulting dataset contains information on all women diagnosed with screen-detected DCIS throughout England. By linking the dataset with the Cancer Analysis System (CAS), it was possible to identify which women with screen-detected DCIS went on to be diagnosed with a subsequent IBC and which women subsequently died with breast cancer as the underlying cause on their death certificate. This enabled a retrospective cohort study to be conducted. Study linkage was undertaken within Public Health England and this study was undertaken under research ethics approval (reference: 16/YH/0209) (Appendices G-H).

5.2.2. Participants

Records of all women who have been diagnosed with screen-detected DCIS (ICD 10: D05 (207)) in England during the period 1/1/2000 – 31/12/2015 were identified. Following linkage with CAS, any individuals who, prior to or within 6 months of their DCIS diagnosis that qualified them for entry into this study, had been were registered with invasive cancer of any type, other than non-melanoma skin cancer (ICD-10 code: C44 (207); ICD-9 code 173 (240)), were excluded. Similarly, any women recorded as having received chemotherapy as treatment for her DCIS, and any women with no date of birth recorded, were excluded.

5.2.3. Variables

Information on the following variables was obtained from the NHSBSP audit: age at diagnosis (<55, 55-60, 60-64, 65+ years), calendar year of diagnosis (2000-2004, 2005-2009 and 2010-2015), DCIS size (≤ 10 , 11-20, 21-50, 51+ mm), pathological grade of DCIS (low or intermediate, high), laterality of DCIS (left, right), surgery and radiotherapy (breast conserving surgery (BCS) with radiotherapy (RT), BCS no RT, mastectomy), final margin distance (involved, 1-2, 3-4, 5+ mm), oestrogen receptor (ER) status, (positive, negative) and hormone therapy (received, not received). Any missing information from the NHSBSP audit was supplemented, if possible, using data stored in CAS. If this was not possible, missing information was treated as an additional category for each variable.

5.2.4. Statistical methods

Three types of analyses were carried out: first, the cumulative risks of any (i.e. ipsilateral or contralateral) IBC and of breast cancer death by time since diagnosis of DCIS were estimated; second, factors associated with the development of ipsilateral invasive breast cancer (iIBC) in women who had not received hormone therapy were estimated; and third, the rates of developing any IBC in ER+ women with and without hormone therapy were compared. For all three analyses, women were followed from six months after their DCIS diagnosis to the earliest of IBC diagnosis, death, loss to follow-up, or 31st December 2015. Cumulative risks were estimated using the Kaplan-Meier method. Variability in the IBC rate with each of the variables listed above and with years of follow up (<2, 2-3, 4-5, 6-7, 8-9, 10+ years) was examined using Poisson regression. Initial analyses considered each variable on its own. The analyses were then repeated considering each variable after adjustment for all other variables. Confidence intervals were based on the standard error. Significance tests for

the effect of each variable, and for all pairwise interactions between them, were based on the likelihood ratio.

5.3. Results

The NHSBSP audit held information on a total of 196,294 women. 164,263 women were registered with a cancer prior to their diagnosis of DCIS, 218 women were diagnosed with IBC within 6 months of their diagnosis of DCIS, 1 woman died with breast cancer recorded on her death certificate within 6 months of her diagnosis of DCIS, 8 women did not have a date of birth recorded and 140 women were recorded as having received chemotherapy as treatment for their DCIS, and were not considered further. After applying the exclusion criteria as described above, a total of 31,664 women diagnosed with screen-detected DCIS in England were available for analysis (see Figure 5.1).

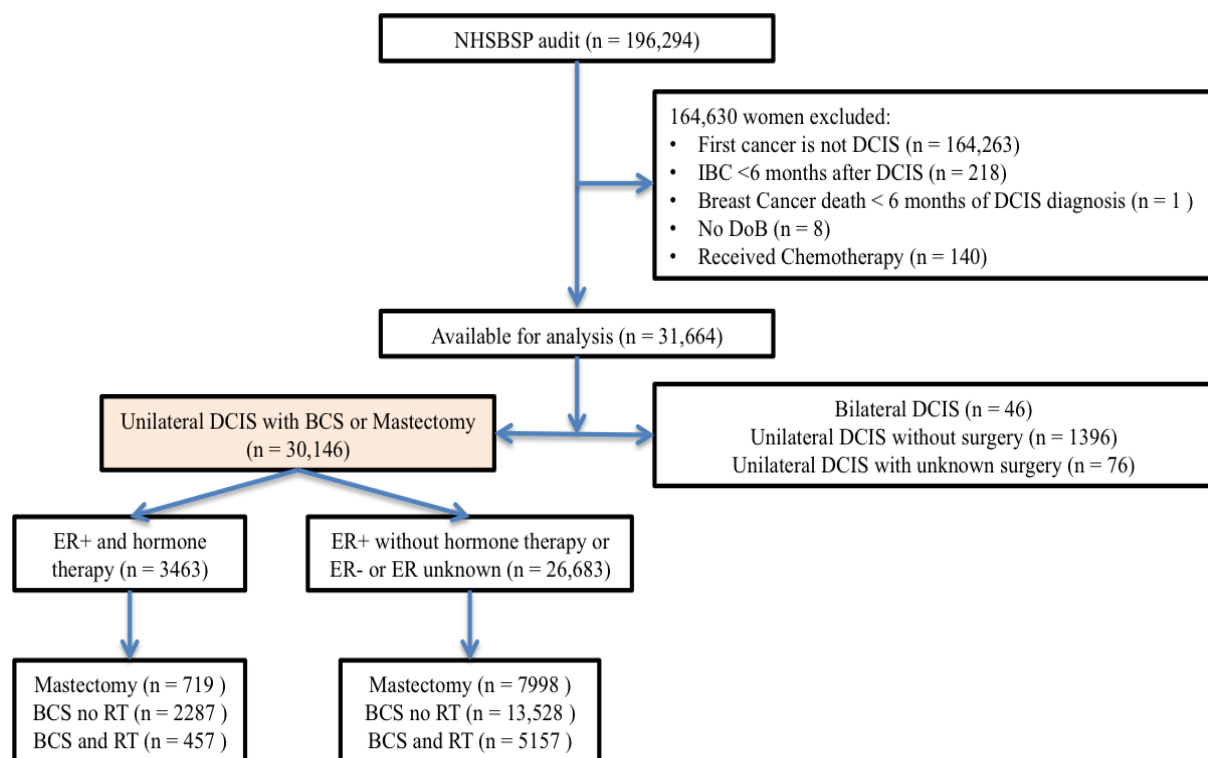


Figure 5.1: Flow diagram showing the selection of participants for inclusion in the study from the 196,294 women recorded in the National Health Service Breast Screening Programme (NHSBSP) audit. Abbreviations: DoB: date of birth, BCS: breast conserving surgery, DCIS: ductal carcinoma in situ, ER: oestrogen receptor, IBC: invasive breast cancer, RT: radiotherapy.

5.3.1. Risk of any IBC or breast cancer death by time since diagnosis

Of the 31,664 women available for analysis and diagnosed with screen-detected DCIS as their first cancer, 1416 women were subsequently diagnosed with any (i.e. ipsilateral or contralateral) IBC and 234 died with breast cancer as the underlying cause of death. The cumulative risks of any IBC at 5, 10 and 15 years after DCIS diagnosis were 3.5%, 7.1%, and 9.4% and the cumulative risks of dying from breast cancer after DCIS diagnosis were 0.3%, 1.1%, and 2.3% respectively (Figure 5.2).

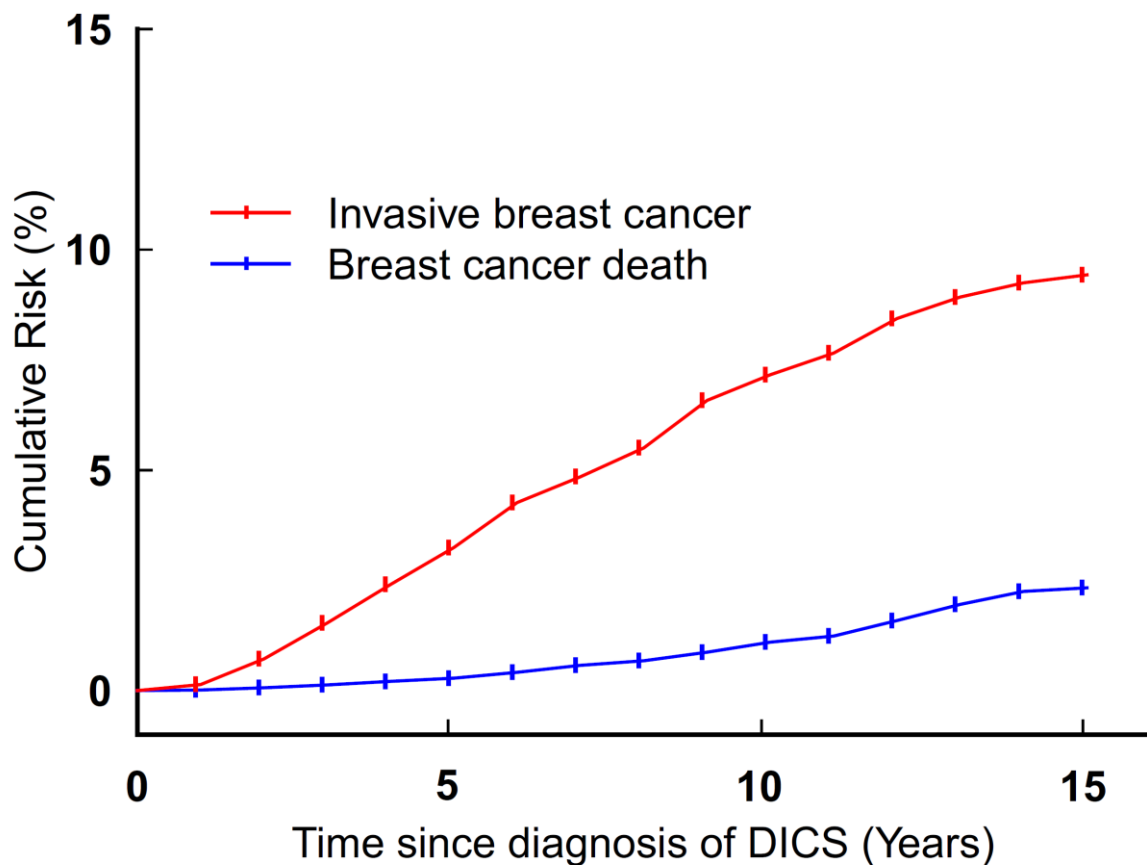


Figure 5.2: Cumulative risk of any (ipsilateral or contralateral) IBC and breast cancer death by time since diagnosis of DCIS among 31,664 women registered with screen-detected DCIS as their first cancer.

5.3.2. Risk of IBC in women with unilateral DCIS who received surgery

Of the 31,664 women diagnosed with screen-detected DCIS as their first cancer, 46 women had bilateral DCIS, 1396 women had unilateral DCIS without surgical treatment and 76 women had unilateral DCIS with unknown surgery and therefore were not considered further, leaving a total of 30,146 women with unilateral screen-detected DCIS and treated with surgery for study (Figure 5.1). The majority of women were aged between 50 and 70 years of age, corresponding with eligibility criteria for breast cancer screening in England during the study period. 9427 (31.3%) women were aged <55 year, 6181 (20.5%) women were aged 55-60 years, 6622 (22.0%) were aged 60-64 years and 7916 (26.3%) were aged 65+ years (Table 5.1).

The number of women diagnosed with screen-detected DCIS in each five-year period increased during the study period, from 5989 (19.9%) in 2000 – 2004 to 10,419 (34.6%) in 2005 – 2009, and again to 13,738 (45.6%) in 2010 – 2015. As expected with a screened population, the majority of women had small lesions ($\leq 10\text{mm}$: 32.6%, 11-20mm: 29.6%), whereas 30.0% of women had tumours measuring 21-50mm, and 7.7% had tumours measuring more than 51mm. Amongst those for whom grade was known, 59.6% of women had high grade tumours and 40.4% had low or intermediate grade tumours. Among those for whom ER status was known, 78.8% had tumours that were ER positive. Overall, 5614 (18.6%) women with DCIS were treated with BCS alone, 14,485 (52.3%) were treated with BCS and RT and 8717 (28.9%) with mastectomy. A total of 4459 (15.3%) women with DCIS were recorded as having received hormone therapy (Figure 5.1).

Table 5.1: Participant characteristics for all 30,146 women with unilateral screen-detected DCIS in England during 2000 – 2015 and treated with surgery.

Variable	Category	n	Percent
Age at DCIS diagnosis (years)	<55	9427	31.3%
	55-60	6181	20.5%
	60-64	6622	22.0%
	65+	7916	26.3%
Calendar year of diagnosis	2000-2004	5989	19.9%
	2005-2009	10,419	34.6%
	2010-2015	13,738	45.6%
Years of follow-up	<2	2893	9.6%
	2-3	5908	19.6%
	4-5	5136	17.0%
	6-7	4436	14.7%
	8-9	4123	13.7%
	10+	7650	25.4%
Tumour size (mm)	≤10	9293	32.6%
	11-20	8427	29.6%
	21-50	8543	30.0%
	51+	2204	7.7%
	Unknown	1679	
Grade	Low/intermediate	11,981	40.4%
	High	17,669	59.6%
	Unknown	496	
ER status	Negative	2935	21.2%
	Positive	10,891	78.8%
	Unknown	16,320	
Laterality of DCIS	Left	15,516	51.7%
	Right	14,485	48.3%
	Unknown	145	
Local treatment	BCS	5614	18.6%
	BCS+RT	15,815	52.3%
	Mastectomy	8717	28.9%
Final margin distance (mm)	5+	7846	58.0%
	3-4	2188	16.2%
	1-2	3057	22.6%
	Involved	443	3.3%
	Unknown	16,612	
Hormone therapy recorded	Yes	4459	15.3%
	No	25,687	84.7%
Total with DCIS as first event		30,146	100%

Abbreviations: mm = millimetres, ER = oestrogen receptor

5.3.3. Patient and tumour related factors and iIBC rate

Of the 30,146 women with unilateral screen-detected DCIS and treated with surgery, 3463 women were ER+ DCIS received hormone therapy and therefore were not considered further, leaving a total of 26,683 women for study of patient and tumour related factors and iIBC rate. There was a significant variation in iIBC rate by calendar year of diagnosis (p for heterogeneity: < 0.001) (Figure 5.3). Women diagnosed with screen-detected DCIS between 2010 – 2016 had almost half the rate of iIBC than women diagnosed between 2000-2004 (RR 0.54 (95% CI 0.42 – 0.69)). This finding was only slightly attenuated following adjustment for years since diagnosis and all other factors (RR 0.69 (95% CI 0.50 – 0.95)) (Figure 5.4).

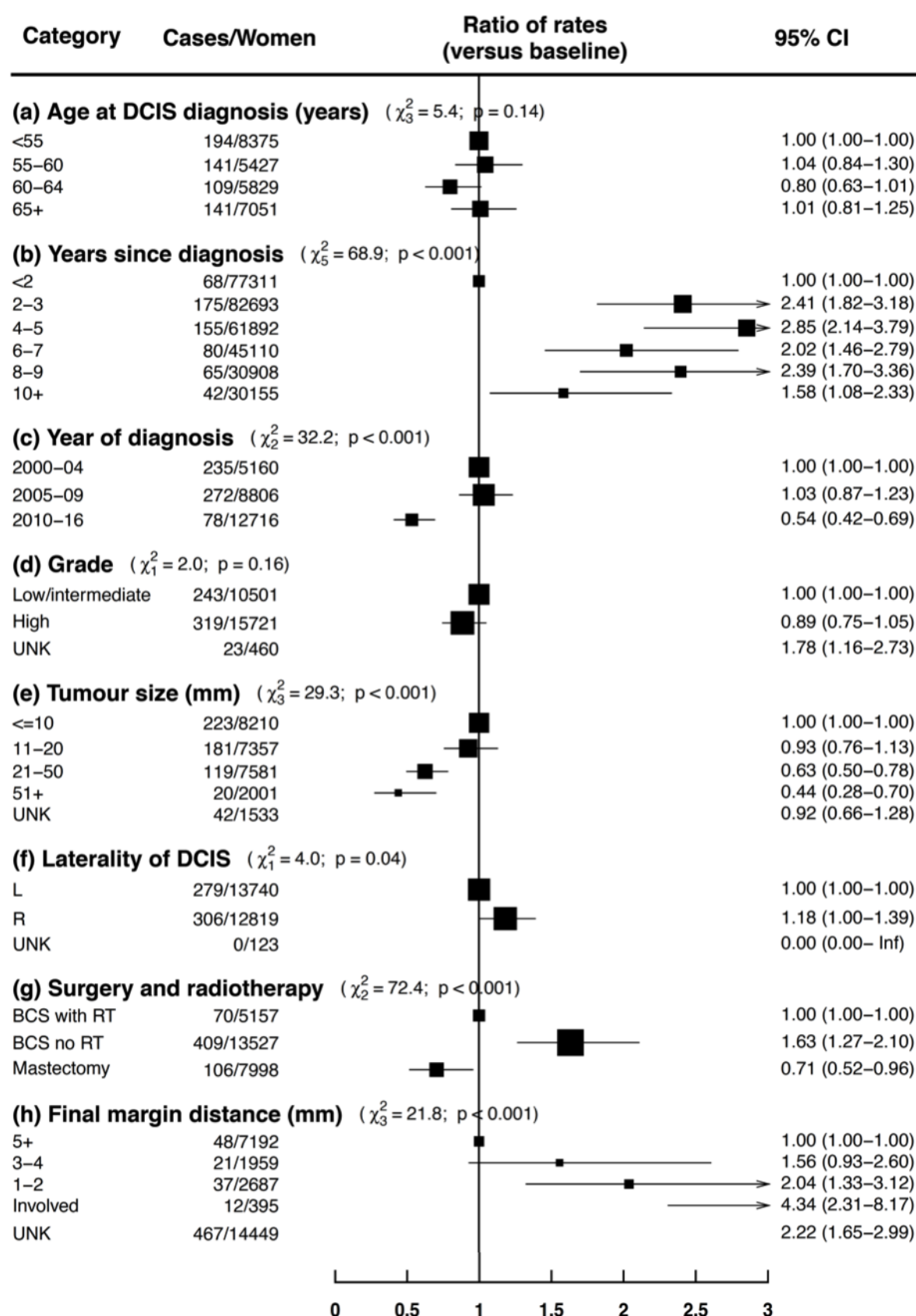


Figure 5.3: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 26,683 women who were not recorded as receiving hormone therapy. No adjustment applied. Baseline is the first row for each variable (a) – (h). Abbreviations: CI = Confidence Interval, UNK = Unknown, mm = millimeters, BCS = Breast Conserving Surgery, RT = Radiotherapy.

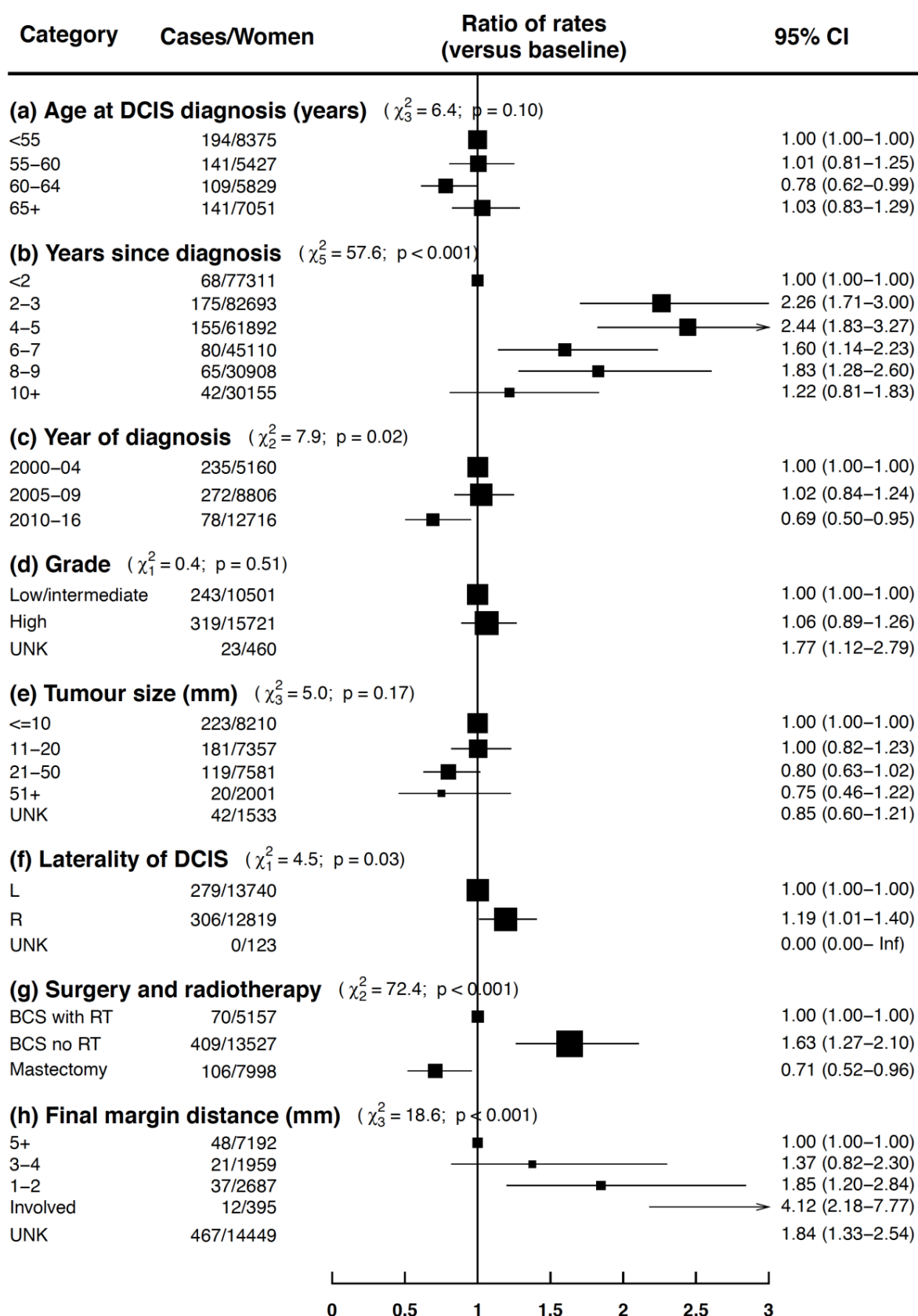


Figure 5.4: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 26,683 women who were not recorded as receiving hormone therapy. Each variable is adjusted by all other variables in the categories shown. Baseline is the first row for each variable (a) – (h). Abbreviations: CI = Confidence Interval, UNK = Unknown, mm = millimeters, BCS = Breast Conserving Surgery, RT = Radiotherapy.

In the unadjusted analysis, larger DCIS size resulted in a lower iIBC rate (p for heterogeneity: 0.005). Women with DCIS lesions measuring 21-50 mm and 50 mm+ had lower iIBC rates than women with tumours ≤ 10 mm in size (RR 0.63 (95% CI 0.50 – 0.78) and RR 0.44 (95% CI 0.28 – 0.70), respectively) (Figure 5.3). However, after adjusting for all the other factors, the size of the effect was much diminished and no significant heterogeneity in the iIBC rate with tumour size remained (p for heterogeneity: 0.17) (Figure 5.4). There was no significant heterogeneity in the iIBC rate by age at DCIS diagnosis in either the unadjusted or the adjusted analysis (p for heterogeneity: 0.14 and 0.10, respectively) although in the adjusted analysis the rate of iIBC for women aged 60-64 years compared with women aged <55 years was 0.78 (95% CI 0.62 – 0.99). However, it was consistent with random variation (p for heterogeneity = 0.10). There was an increase in the rate of iIBC in women diagnosed with DCIS in the right breast compared with the left breast (RR 1.18 (95% CI 1.00 – 1.39), p for difference = 0.04), which persisted following adjustment (RR 1.19 (95% CI 1.01 – 1.40), p for difference = 0.03).

There was no significant difference in iIBC rate between high and low/intermediate grade tumours either before or after adjustment for the other variables (p for heterogeneity: 0.16 and 0.51, respectively). There was, however, a significant interaction between DCIS grade and years of follow up after adjustments for the other variables (p for heterogeneity: 0.02), indicating that the difference between the iIBC rate in high and in low/intermediate grade DCIS differed by time since DCIS diagnosis. After adjustment for surgery and radiotherapy, there was a higher cumulative risk of iIBC with high grade DCIS in the first 5 years following DCIS diagnosis compared with low/intermediate grade DCIS (Figure 5.4). However in the intermediate period, during 5-10 years following DCIS diagnosis, the iIBC rate in low/intermediate grade DCIS increased, so that by 10 years, the cumulative risk of

IBC in low/intermediate grade DCIS was 2%, matching that in the high grade DCIS. This pattern persisted into the 10+ year period following DCIS diagnosis resulting in an overall higher cumulative risk from low or intermediate grade DCIS (2.7%) at 14 years, compared with high grade DCIS (2.4%) (Figure 5.5).

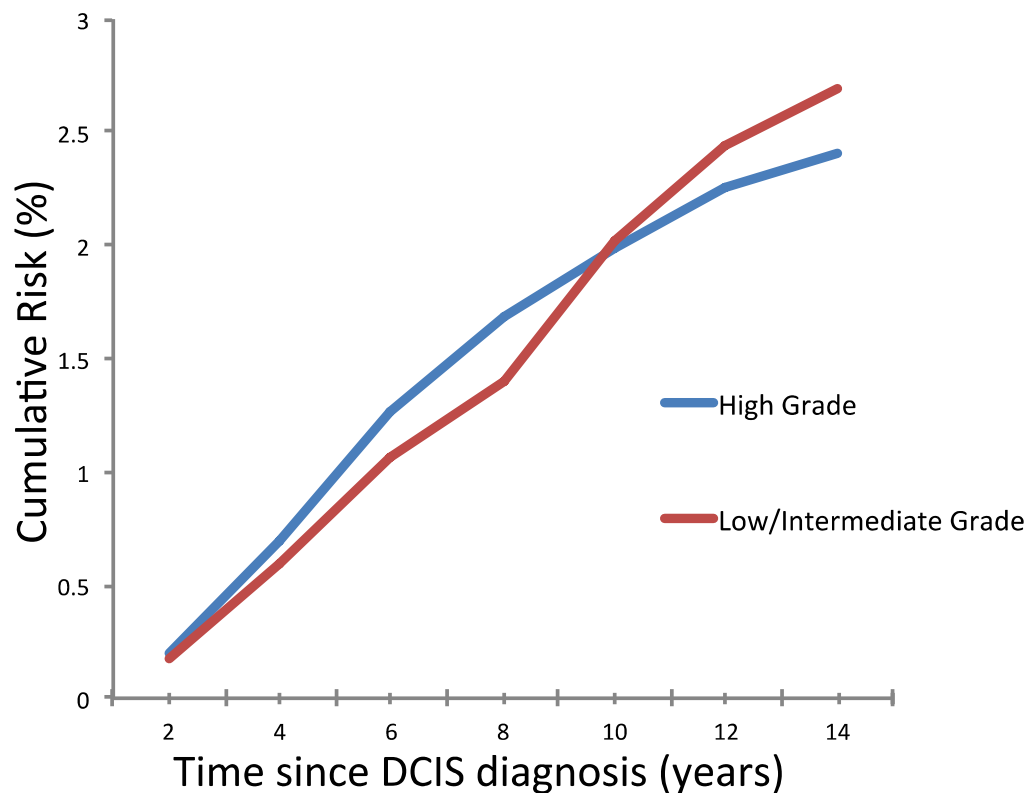


Figure 5.5: Cumulative risk of ipsilateral invasive breast cancer (iIBC) in high and low/intermediate grade DCIS. Adjusted for surgery and radiotherapy.

5.3.4. Treatment related factors and IBC rate

There was strong evidence of heterogeneity in the iIBC rate according to the local treatment the women received (p for heterogeneity <0.001 both before and after adjustment for the other variables). Without adjustment, women treated with mastectomy had a lower rate of iIBC than women treated with BCS with radiotherapy (RR 0.71 (95% CI 0.52–0.96). Whilst women who received BCS with no radiotherapy had a higher iIBC rate than women treated with BCS with radiotherapy (RR 1.63, 95% CI 1.27–2.10) (Figure 5.3). Rate ratios after adjustment for all other variables were identical considered (Figure 5.4).

As expected women who still had involved margins at the end of their surgery had higher iIBC rates than women with clear margins (p for heterogeneity: < 0.001 , both with and without adjustment), however among women with clear margins, iIBC rate increased with decreasing surgical margin distance (p for trend < 0.001 both before and after adjustment). Women with clear surgical margins of 1–2 mm had an increased iIBC rate compared with women with clear margins of 5+ mm both analyses before adjustment (RR 2.04, 95% CI 1.33 – 3.12) (Figure 5.3) and also with adjustment (RR 1.85, 95% CI 1.20 – 2.84) (Figure 5.4). The evidence for a relationship between margin distance and iIBC was strongest in women treated with BCS alone (p for heterogeneity: 0.003) (Appendix I). The trends were not significant in women who received radiotherapy after BCS (p for heterogeneity: 0.30) (Appendix J), or in women who received mastectomy (p for heterogeneity = 0.19) (Appendix K).

Due to its systemic, rather than local mechanism of action, analyses of hormone therapy after screen-detected DCIS considered any (both ipsilateral and contralateral) IBC, instead of solely iIBC. Analyses of hormone therapy were conducted in all 10,891 women recorded as having ER+ DCIS (Table 5.1) and compared women recorded as having received hormone therapy with women with no such record of having received hormone therapy.

In the unadjusted analysis, ER+ women given hormone therapy had a somewhat lower rate of any IBC compared with ER+ women who were not given hormone therapy (RR 0.84, 95% CI 0.69–1.02). The difference became statistically significant following adjustment for all other factors (RR 0.76, 95% CI 0.63–0.93) (Figure 5.6).

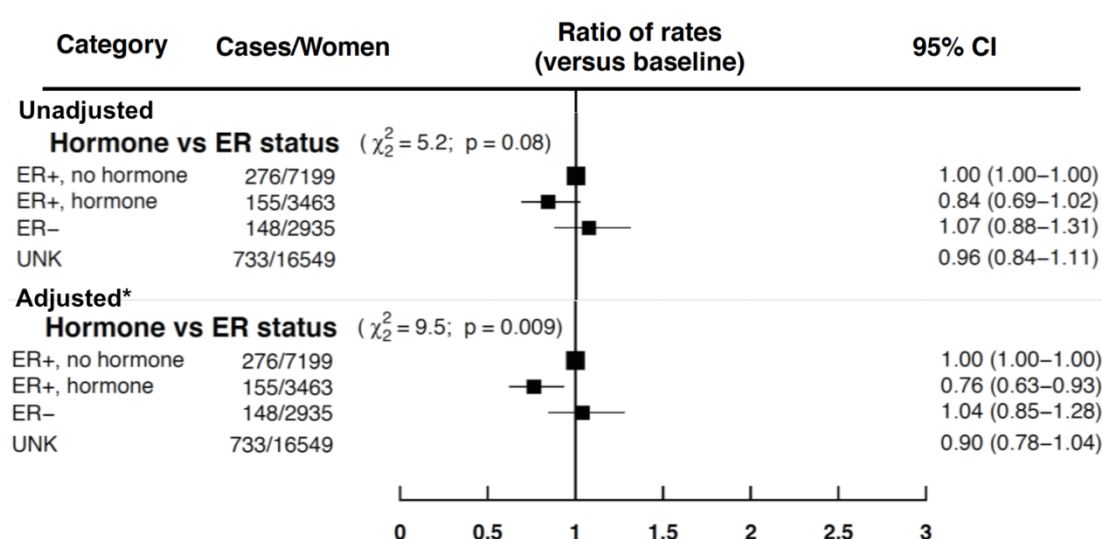


Figure 5.6: Hormone therapy and any (i.e. ipsilateral or contralateral) invasive breast DCIS in 10,891 women recorded with ER+ DCIS. *Adjusted for all variables listed in methods, section 2.3. Abbreviations: CI = Confidence Interval, UNK = Unknown, Hormone = Hormone therapy, ER = Oestrogen Receptor.

5.4. Discussion

This is the first nationwide study to examine the long-term risk of IBC in individual women diagnosed with screen-detected DCIS in England, and it has provided important insights into its epidemiology. This study has shown that the cumulative risk of any IBC after screen-detected DCIS is 9.4% at 15 years and that women diagnosed in more recent years (2010 – 2015) have lower rates of iIBC than those diagnosed a decade earlier. This study also contributes to the discussion of optimal resection margin distance in DCIS and it provides novel evidence that the pattern of risk for high and lower/intermediate grade DCIS changes over time since DCIS diagnosis, with higher IBC rates in women with low/intermediate grade DCIS than high grade DCIS more than 10 years after treatment. This work also provides important evidence that the benefits of hormone therapy and radiotherapy identified from randomised clinical trials are being carried over to women diagnosed with screen-detected DCIS in the general population (160, 197, 239). These results, together with evidence from randomised trials, can help clinicians to discuss the risks and benefits of different treatment strategies with women diagnosed with screen-detected DCIS.

5.4.1. Calendar year

Women diagnosed with screen-detected DCIS between 2010 – 2016 had lower iIBC rates than women diagnosed between 2000 – 2004 even after adjustment for all other factors considered (RR 0.69 (95% CI 0.50 – 0.95)) (Figure 5.4). The reasons for this are likely to be multifactorial. A number of national breast cancer initiatives were introduced shortly before 2010, such as the 2007 Cancer Reform Strategy and updated treatment recommendations from professional bodies such as the National Institute for Clinical Excellence and the Association of Breast Surgery, in order to spread best practice and to help minimise variation

in treatment strategies (170, 241, 242). The improvement in iIBC rates amongst women diagnosed with screen-detected DCIS in 2010 – 2016 may represent a beneficial consequence of these initiatives. The consequent changes in clinical pathways and cancer infrastructure from these initiatives may have influenced the rates of iIBC after DCIS in ways not accounted for by the patient, tumour and treatment factors adjusted for in the present study.

5.4.2. Tumour related factors and iIBC rate

High grade DCIS has traditionally been considered a marker of poor prognosis when compared with low or intermediate grade DCIS and this has led to it being incorporated into several prognostic indices for DCIS. However, the evidence these prognostic indices are derived from and validated against comprise small numbers of patients, who present both symptomatically and via screening, and who often have less than 5 years of subsequent follow up (151-153, 157). In this study, with larger size and longer follow-up duration, a higher cumulative risk from low or intermediate grade DCIS (2.7% at 15 years) was evident when compared with high grade DCIS (2.4% at 15 years) (Figure 5.5). This finding may apply only to the screened population but nevertheless it raises important implications when selecting women for less extensive treatment, or for non-operative management based on DCIS grade.

Some studies have found no association between DCIS size and disease recurrence or metastasis (243), whereas others show that it significantly influences outcomes following DCIS (244, 245). Due to irregular ductal growth patterns, differences in tumour sample processing between centres and tumour sampling limitations, accurate pathological determination of lesion size can be challenging (246). These factors are difficult to take into account when studying DCIS size, and in our analyses, we did not find a significant

difference in tumour size and rate of iIBC after taking all other factors into account (Figure 5.4). Surprisingly, there was an increase in the rate of iIBC in women diagnosed with DCIS in the right breast compared with the left breast (RR 1.18 (95% CI 1.00 – 1.39), *p* for difference = 0.04), which persisted following adjustment (RR 1.19 (95% CI 1.01 – 1.40), *p* for difference = 0.03). The reason for this finding is not clear and published literature supporting, or disputing, a role of tumour laterality on outcomes after DCIS is lacking.

5.4.3. Margin distance

In clinical practice, an involved margin with zero millimetres thickness, often referred to in pathology terms as ‘ink on tumour’, is a marker of inadequate local therapy and such cases often warrant further surgery. When clear margins are present, the larger the radial margin, the greater the likelihood of complete excision and the lower the likelihood of missing occult foci of disease (247). However, the optimal margin distance is contentious, since insufficient margins may result in treatment failure with high local DCIS recurrence or IBC occurrence rates, but disproportionate margins cause overtreatment with potentially poorer cosmetic outcome in the absence of any tangible oncologic benefit (248). Previously published studies show no significant difference in the rate of iIBC with margins of 2 mm compared with 5 mm+ margins (168). However, prognosis amongst women with clear margins, but with a margin distance of less than 2 mm, remains less clear.

There have been conflicting views for many years on the optimal margin distance for women with DCIS (168, 249), and this is evidenced both by the variation in clinical practice internationally (250-252), and between professional organisations within the UK. The Association of Breast Surgery is the professional body of breast surgeons in the UK and although the organisation’s formal guidance defers the decision of optimal margin distance to

local multidisciplinary team discussion, many members accept a 1 mm margin distance to be adequate (241). The National Institute of Clinical Excellence (NICE) guidance, however recommends a minimum 2 mm margin for women treated with BCS and advises a discussion of the risks and benefits for re-excision should the margin distance be less than 2 mm (170). The issue of optimal margin distance is further compounded by the technical difficulties pathologists face in reliably ascertaining the anatomy of the breast ductal system on excised specimens and from sampling error during specimen preparation (252).

The results of the present study contribute towards the discussion of optimal margin distance by showing a significant trend of increasing iIBC rate with decreasing surgical margin distance ($p < 0.001$) (Figure 5.3). After adjustment for other factors, women with screen-detected DCIS with clear surgical margins of 1-2 mm had nearly twice the iIBC rate of women with clear margins of 5+ mm (RR 1.85, 95% CI 1.20– 2.84). As with other recent studies (253, 254), the present study found no significant difference in iIBC rate with 1-2 mm margin distance when compared with 5+ mm margins in women who received radiotherapy following their BCS (RR 3.19 95% CI 0.86-11.92, p for heterogeneity = 0.30) (Appendix J), likely due to the beneficial effect of radiotherapy in treating any remaining occult tumour burden from close margins.

5.4.4. Radiotherapy and DCIS

Women treated with mastectomy had an iIBC rate of around 30% lower than women treated with BCS with radiotherapy, whereas women treated with BCS without radiotherapy had around 60% higher iIBC rate than women treated with BCS with radiotherapy. Our findings are comparable with the results from a recently published population-based cohort of 10,090 women with DCIS treated in the Netherlands (255). Together, these studies show that the

beneficial reduction in iIBC rate from radiotherapy seen in the highly selected participants enrolled into randomised trials are carried over to women diagnosed with screen-detected DCIS in the general population (25, 57, 59, 160, 190).

In clinical practice, women receiving BCS with no radiotherapy often have characteristics that have traditionally been considered as favourable, such as smaller lesions, lower grade of disease and larger margin distances, when compared with women treated with BCS with radiotherapy (161, 256). In addition, there remains wide geographical variation in the use of adjuvant RT by individual breast units across the UK (161). It is possible that this variation may, in the past, have been due to regional limitations in radiotherapy capacity, resulting in clinicians prioritising patients with IBC or high risk DCIS over those with low or intermediate risk DCIS, based upon on these supposedly favourable characteristics (161, 257). However, the variation in radiotherapy use after BCS for DCIS may also arise from differing views amongst clinicians regarding the absolute risks and benefits of radiotherapy in the setting of low or intermediate risk DCIS (161), particularly as professional bodies do not provide definitive recommendations on this issue (170).

5.4.5. Hormone therapy and DCIS

A meta-analysis of 3375 women from two randomised trials of tamoxifen versus placebo in women following surgical resection of DCIS showed no benefit in all-cause mortality from tamoxifen (RR 1.11, 95% CI 0.89-1.39) (189, 197, 198). However, there was a reduction in the risk of ipsilateral DCIS recurrence (hazard ratio [HR] 0.75, 95% CI 0.61 - 0.92), and a trend towards a lower risk of iIBC, but that did not reach statistical significance (HR 0.79, 95% CI 0.61 - 1.01). this meta-analysis did show that Tamoxifen reduced the risk of contralateral DCIS (relative risk [RR] 0.50, 95% CI 0.28 - 0.87) and contralateral IBC (RR

0.57, 95% CI 0.39 - 0.83). Only one of the two trials reported on adverse events and there was a non-significant trend towards more endometrial cancer in the tamoxifen group but no overall difference in adverse events when compared with placebo (258). These data were not available at the time NICE guidance advised against the routine use of hormone therapy in DCIS. Whilst the findings from the present study support the benefit of hormone therapy in reducing the rate of any IBC after screen-detected DCIS (RR 0.75, 95% CI 0.60–0.92) (Figure 5.6), information regarding the type of hormone therapy women received or any information on side-effects or adverse events that women may have suffered whilst taking hormone therapy, was not available. Consequently, it was not possible to weigh the benefits with the risks from taking hormone therapy in the present study.

5.4.6. Limitations

The NHSBSP audit and the CAS are invaluable sources of population-level information on outcomes in women diagnosed with DCIS in the UK, as well as in helping to inform how patterns of care have changed over time. However, care must be taken when inferring the effects of cancer treatments based solely on non-randomised data, even when efforts had been made to remove confounding as selection bias may remain. Therefore non-randomized comparisons can provide inaccurate estimates of treatment effects and careful justification is required every time they are used for this purpose (259). In the case of this study, DCIS is a common disease but with many areas of uncertainty regarding its epidemiology and optimal treatment that have not fully been addressed by randomised trials. It is in this context that population-based, non-randomised datasets can not only be hypothesis-generating, but can also provide an invaluable contribution to clinical decision-making, both for issues such as optimal margin distance, which are not amenable to study by means of a randomised trial, and for issues that have not yet been studied by appropriate randomised trials.

5.4.7. Conclusion

This study is the largest of its kind in the UK and for the first time has described the prognosis of women diagnosed with screen-detected DCIS and treated under different management strategies. The risk of IBC following a diagnosis of DCIS varies with patient, tumour and treatment factors. The findings from this study will help in the discussion between clinicians and patients regarding the optimal management of screen-detected DCIS by quantifying outcomes after different margin distances and after different adjuvant treatments such as radiotherapy or hormone therapy.

5.5. References

1. Public Health England. NHS breast screening programme and Association of Breast Surgery: An audit of screen detected breast cancers for the year April 2015 to March 2016: Public Health England; 2016 [Available from: <http://associationofbreastsurgery.org.uk/audit/nhs-breast-screening-programme-audit/>].
2. Early Breast Cancer Trialists' Collaborative G, Correa C, McGale P, Taylor C, Wang Y, Clarke M, et al. Overview of the randomized trials of radiotherapy in ductal carcinoma in situ of the breast. *Journal of the National Cancer Institute Monographs*. 2010;2010(41):162-77.
3. Cuzick J, Sestak I, Pinder SE, Ellis IO, Forsyth S, Bundred NJ, et al. Effect of tamoxifen and radiotherapy in women with locally excised ductal carcinoma in situ: long-term results from the UK/ANZ DCIS trial. *The Lancet Oncology*. 2011;12(1):21-9.
4. Fisher B, Dignam J, Wolmark N, Wickerham DL, Fisher ER, Mamounas E, et al. Tamoxifen in treatment of intraductal breast cancer: National Surgical Adjuvant Breast and Bowel Project B-24 randomised controlled trial. *Lancet*. 1999;353(9169):1993-2000.
5. Bijker N, Peterse JL, Duchateau L, Julien JP, Fentiman IS, Duval C, et al. Risk factors for recurrence and metastasis after breast-conserving therapy for ductal carcinoma-in-situ: analysis of European Organization for Research and Treatment of Cancer Trial 10853. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2001;19(8):2263-71.
6. Benson JR, Wishart GC. Predictors of recurrence for ductal carcinoma in situ after breast-conserving surgery. *The Lancet Oncology*. 2013;14(9):e348-57.
7. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal carcinoma in situ of the breast: a systematic review of incidence, treatment, and outcomes. *Journal of the National Cancer Institute*. 2010;102(3):170-8.
8. WHO. ICD-10 Version: 2016 World Health Organisation; 2016 [Available from: <http://apps.who.int/classifications/icd10/browse/2016/en>].
9. WHO. ICD-9 Version: 1979: World Health Organisation; 1979 [Available from: <https://www.cdc.gov/nchs/icd/icd9.htm>].
10. NICE guidelines [CG80]. Early and locally advanced breast cancer: Diagnosis and treatment. NICE guidance. 2009.
11. Association of Breast Surgery at B. Surgical guidelines for the management of breast cancer. *Eur J Surg Oncol*. 2009;35 Suppl 1:1-22.
12. Department of Health. Cancer Reform Strategy. 2007.
13. Boland GP, Chan KC, Knox WF, Roberts SA, Bundred NJ. Value of the Van Nuys Prognostic Index in prediction of recurrence of ductal carcinoma in situ after breast-conserving surgery. *Br J Surg*. 2003;90(4):426-32.
14. Holland R, Peterse JL, Millis RR, Eusebi V, Faverly D, van de Vijver MJ, et al. Ductal carcinoma in situ: a proposal for a new classification. *Seminars in diagnostic pathology*. 1994;11(3):167-80.
15. Silverstein MJ, Lagios MD, Craig PH, Waisman JR, Lewinsky BS, Colburn WJ, et al. A prognostic index for ductal carcinoma in situ of the breast. *Cancer*. 1996;77(11):2267-74.
16. Scott MA, Lagios MD, Axelsson K, Rogers LW, Anderson TJ, Page DL. Ductal carcinoma in situ of the breast: reproducibility of histological subtype analysis. *Human pathology*. 1997;28(8):967-73.
17. Bijker N, Peterse JL, Duchateau L, Julien J-P, Fentiman IS, Duval C, et al. Risk Factors for Recurrence and Metastasis After Breast-Conserving Therapy for Ductal

Carcinoma-In-Situ: Analysis of European Organization for Research and Treatment of Cancer Trial 10853. *Journal of Clinical Oncology*. 2001;19(8):2263-71.

18. Narod SA, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. *JAMA Oncol*. 2015;1(7):888-96.

19. Virnig BA, Tuttle TM, Shamliyan T, Kane RL. Ductal Carcinoma In Situ of the Breast: A Systematic Review of Incidence, Treatment, and Outcomes. *JNCI: Journal of the National Cancer Institute*. 2010;102(3):170-8.

20. Benson JR, Wishart GC. Predictors of recurrence for ductal carcinoma in situ after breast-conserving surgery. *The Lancet Oncology*. 14(9):e348-e57.

21. Holland R, Veling SH, Mravunac M, Hendriks JH. Histologic multifocality of Tis, T1-2 breast carcinomas. Implications for clinical trials of breast-conserving surgery. *Cancer*. 1985;56(5):979-90.

22. Singletary SE. Surgical margins in patients with early-stage breast cancer treated with breast conservation therapy. *The American Journal of Surgery*. 2002;184(5):383-93.

23. Dunne C, Burke JP, Morrow M, Kell MR. Effect of margin status on local recurrence after breast conservation and radiation therapy for ductal carcinoma in situ. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009;27(10):1615-20.

24. Groot G, Rees H, Pahwa P, Kanagaratnam S, Kinloch M. Predicting local recurrence following breast-conserving therapy for early stage breast cancer: the significance of a narrow (≤ 2 mm) surgical resection margin. *J Surg Oncol*. 2011;103(3):212-6.

25. Morrow M, Van Zee KJ, Solin LJ, Houssami N, Chavez-MacGregor M, Harris JR, et al. Society of Surgical Oncology-American Society for Radiation Oncology-American Society of Clinical Oncology Consensus Guideline on Margins for Breast-Conserving Surgery With Whole-Breast Irradiation in Ductal Carcinoma In Situ. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2016;34(33):4040-6.

26. Morrow M, Harris JR, Schnitt SJ. Surgical margins in lumpectomy for breast cancer--bigger is not better. *N Engl J Med*. 2012;367(1):79-82.

27. Ministry of Health. Management of Early Breast Cancer. 2009.

28. Tadros AB, Smith BD, Shen Y, Lin H, Krishnamurthy S, Lucci A, et al. Ductal Carcinoma In Situ and Margins <2 mm: Contemporary Outcomes With Breast Conservation. *Ann Surg*. 2017.

29. Thompson A. Close margins, radiotherapy and outcomes in screen-detected ductal carcinoma in situ (DCIS). *European Journal of Surgical Oncology*. 43(5):S3-S4.

30. Elshof LE, Schaapveld M, Schmidt MK, Rutgers EJ, van Leeuwen FE, Wesseling J. Subsequent risk of ipsilateral and contralateral invasive breast cancer after treatment for ductal carcinoma in situ: incidence and the effect of radiotherapy in a population-based cohort of 10,090 women. *Breast cancer research and treatment*. 2016;159(3):553-63.

31. McCormick B, Winter K, Hudis C, Kuerer HM, Rakovitch E, Smith BL, et al. RTOG 9804: a prospective randomized trial for good-risk ductal carcinoma in situ comparing radiotherapy with observation. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2015;33(7):709-15.

32. Houghton J, George WD, Cuzick J, Duggan C, Fentiman IS, Spittle M, et al. Radiotherapy and tamoxifen in women with completely excised ductal carcinoma in situ of the breast in the UK, Australia, and New Zealand: randomised controlled trial. *Lancet*. 2003;362(9378):95-102.

33. Emdin SO, Granstrand B, Ringberg A, Sandelin K, Arnesson LG, Nordgren H, et al. SweDCIS: Radiotherapy after sector resection for ductal carcinoma in situ of the breast.

Results of a randomised trial in a population offered mammography screening. *Acta Oncol.* 2006;45(5):536-43.

34. Fisher ER, Dignam J, Tan-Chiu E, Costantino J, Fisher B, Paik S, et al. Pathologic findings from the National Surgical Adjuvant Breast Project (NSABP) eight-year update of Protocol B-17: intraductal carcinoma. *Cancer.* 1999;86(3):429-38.

35. Dodwell D, Clements K, Lawrence G, Kearins O, Thomson CS, Dewar J, et al. Radiotherapy following breast-conserving surgery for screen-detected ductal carcinoma in situ: indications and utilisation in the UK. Interim findings from the Sloane Project. *British journal of cancer.* 2007;97(6):725-9.

36. Cutuli B, Lemanski C, Fourquet A, de Lafontan B, Giard S, Meunier A, et al. Breast-conserving surgery with or without radiotherapy vs mastectomy for ductal carcinoma in situ: French Survey experience. *British journal of cancer.* 2009;100(7):1048-54.

37. Dodwell D, Crellin A. Waiting for radiotherapy. *BMJ.* 2006;332(7533):107-9.

38. Staley H, McCallum I, Bruce J. Postoperative Tamoxifen for ductal carcinoma in situ: Cochrane systematic review and meta-analysis. *Breast.* 2014;23(5):546-51.

39. Wapnir IL, Dignam JJ, Fisher B, Mamounas EP, Anderson SJ, Julian TB, et al. Long-term outcomes of invasive ipsilateral breast tumor recurrences after lumpectomy in NSABP B-17 and B-24 randomized clinical trials for DCIS. *Journal of the National Cancer Institute.* 2011;103(6):478-88.

40. Staley H, McCallum I, Bruce J. Postoperative tamoxifen for ductal carcinoma in situ. *Cochrane Database Syst Rev.* 2012;10:CD007847.

41. McGale P, Cutter D, Darby SC, Henson KE, Jagsi R, Taylor CW. Can Observational Data Replace Randomized Trials? *J Clin Oncol.* 2016;34(27):3355-7.

Chapter 6: The Use of Routine Data Sources to Identify Recurrence in Women Registered with Invasive Breast Cancer in England

6.1. Introduction

Invasive breast cancer (IBC) incidence rates have been increasing since the mid-1970s and it now accounts for almost 1 in 3 of all newly diagnosed cancers in women (260, 261). Although high quality population-based information is available on breast cancer incidence and on survival after a diagnosis of breast cancer in the UK, much less information is available on breast cancer recurrence. Reliable information on breast cancer recurrence would be helpful for many purposes, including healthcare policy-making (e.g. by guiding decisions on optimal follow-up strategies for women treated with breast cancer), guiding clinical management (e.g. by inputting into clinical decision aids), facilitating follow-up in randomised trials and for descriptive and analytical epidemiology.

The reason for the paucity of information on breast cancer recurrence is that it has not traditionally been part of the standard National Health Service (NHS) information flow. To rectify this, Public Health England (PHE) mandated that from the 1st January 2013, all women newly diagnosed with breast cancer and who subsequently suffer a recurrence of their cancer, should be ‘flagged’ in the new Cancer Outcomes and Services Dataset (COSD) (262). Data entry for the recurrence flag is, however, labour intensive and there is wide geographical variation in its data completeness. Furthermore, on the 1st January 2013 there were approximately half a million women previously diagnosed with breast cancer alive in the UK for whom recurrence reporting was not included within the new COSD mandate. Little is known about how many of these women had a recurrence before COSD was introduced, and at present there is no mechanism in place to record any recurrences that they have developed since then, or may develop in the future. If an automated system using the wealth of data routinely collected by PHE for other purposes could be used to identify breast cancer

recurrences reliably and accurately, then it could be used for all of the purposes listed above and without the drawbacks associated with manual data collection.

The aims of this study were threefold: firstly, to collate all routinely collected information potentially relevant to breast cancer recurrence stored in datasets available within PHE; secondly, to use this information to develop an algorithm to identify which women have had a recurrence of their breast cancer and, for those who did, the date of the recurrence; and thirdly, to examine the accuracy of this algorithm in terms of its sensitivity and specificity.

6.2. Methods

This study involved five stages: (i) identifying and collating all sources of data relevant to breast cancer recurrence and held within PHE; (ii) developing an algorithm to identify recurrence using these data; (iii) identifying a population in which to test the proposed algorithm; (iv) applying the algorithm to the test population and (v) examining the performance of the algorithm. These stages are described below:

6.2.1. Collating all relevant routinely collected datasets

Data items that were considered potentially informative as to whether a woman has had a recurrence of her breast cancer were selected from the routinely collected data sources that are currently held within PHE. These sources included the Cancer Analysis System (CAS), Hospital Episode Statistics on inpatients (HES) (from approximately 1997), the Radiotherapy Treatment Dataset (RTDS) (from approximately 2009), the Digital Imaging Dataset (DID) (from approximately 2014) and the Cancer Waiting Times data (CWT) (from approximately 2009).

6.2.2. Development of the algorithm

Descriptions of all possible eventualities following a diagnosis of IBC were created. These descriptions were then used to create a list of different outcome categories into which the women could be classified based on the routinely collected data items. Examples included locoregional recurrence, distant recurrence, contralateral breast cancer, death from breast cancer, as well as death from an unrelated cause and diagnosis of another primary cancer of a different site. Using this list of possible outcomes, combined with clinically relevant data

items, it was possible to undertake a process of clinical pathway visualisation (263), in order to develop coding rules to classify women into their respective outcome categories. The data items used in these coding rules included information on treatment events, such as type of surgery and adjuvant therapy, the dates and durations of these treatments, and the pathological and radiological factors underlying these treatment decisions (Appendix L).

6.2.3. Identifying a population in which to test the algorithm

The former West Midlands Cancer Intelligence Unit (WMCIU) was the lead cancer registry for breast cancer in England for many years. In this capacity, it had traditionally collected detailed recurrence information through a meticulous process of contacting individual hospitals for women living within its catchment area and diagnosed with IBC. This information was recorded in a system referred to as ‘GRACE’.

The GRACE recurrence dataset is unique in the country in that it provides detailed recurrence information for women diagnosed with IBC in the West Midlands region and has reliable data quality from 1997 until 2011, at which time it was superseded by CAS and no longer collected detailed recurrence information. It is also readily accessible as a testing dataset for the algorithm and has the ability to provide a substantial yet manageable cohort to help identify and mitigate any problems arising prior to expanding work to a nationwide level. For these reasons the population of the West Midlands was selected for the testing stage of this project.

6.2.4. Inclusion criteria

All women registered with IBC: (ICD-10 code C50 (207) during the period 1st January 1997 – 31st December 2011 and living in the West Midlands area at the time of their cancer

registration were included, unless one of the exclusion criteria described below applied. Patients were given anonymised identifiers prior to their inclusion in the study.

6.2.5. Exclusion criteria

The following exclusion criteria were applied: male breast cancer cases; anyone not living in the West Midlands area at the time of their first registration of IBC; anyone with a diagnosis of IBC prior to 1st January 1997 (even if they had another IBC registered after 1st January 1997); and anyone who had another invasive cancer of a site other than IBC prior to their initial diagnosis of IBC, with the exception of non-melanoma skin cancer (ICD-10 code: C44), which was ignored. Previously diagnosed cancers that were not invasive, including in situ breast cancers, benign breast tumours, and breast tumours of uncertain or unknown behaviour (ICD-10: D05, D25, D48.6) were not grounds for exclusion. However, women who, at the time of their initial diagnosis of breast cancer, or within three months of it, were diagnosed with an invasive cancer in the other breast (i.e. had bilateral IBC) were excluded, as were women who were registered with a primary invasive cancer of another site at the same time as their initial diagnosis of IBC, or within three months of it, (with the exception of non-melanoma skin cancer, as specified above).

6.2.6. Application of the algorithm to the test population

The algorithm was then applied to the test population using only routinely collected data. Treatment events recorded in multiple datasets were combined. During this process a number of ‘working rules’ were developed in which events of the same type and occurring within 7 days of each other from different datasets were counted only once and allocated an event source by taking datasets in descending order of national accessibility and expected

completeness: i.e. CAS, HES, CWT, RTDS. Tables were created within CAS to hold patient/tumour data for the cohort, event or treatment related data, and recurrence or outcome related data.

Care was taken to ensure that the routinely collected data did not contain any data items that may have been transferred from GRACE into the CAS, as this would have caused the performance of the algorithm to be over-estimated. For each woman included in the test population, the period of time considered began from the date of initial diagnosis of breast cancer until the earliest of: date of death, date of emigration, date of a classification event of interest or the latest date for which both routinely collected data and the GRACE data from the West Midlands were thought to be complete, i.e. 31st December 2011.

6.2.7. Testing the algorithm using the GRACE data

Data items that had been especially collected in the West Midlands and stored in the GRACE system were also collated. These data provide information on the stage of the cancer at initial diagnosis, recurrences, contralateral breast cancers and deaths, together with the dates of these events, to act as a testing dataset for the algorithm. Using the GRACE dataset, women were classified according to the same list of recurrence categories as before, allowing the numbers of recurrences recorded in the GRACE data to be cross-tabulated with the numbers of recurrences identified from the algorithm using routinely collected data. The cumulative risk of any recurrence recorded in the GRACE data (defined as any of: locoregional recurrence, distant recurrence, recurrence with type unknown, contralateral breast cancer or death from breast cancer with no record of a previous recurrence) were plotted alongside the cumulative risk of any recurrence identified by the algorithm using the routinely collected data. A further comparison was undertaken by comparing the rate of developing the different

types of recurrences recorded in the GRACE data with that identified by the algorithm using the routinely collected data.

6.2.8. Unexpected challenges in conducting the study

This study has involved overcoming four key challenges prior to its completion. Firstly, it was discovered that, unknown to PHE, information on causes of death were missing in the PHE data for large numbers of cancer patients, including over 1500 women in the present study's cohort. Rectifying this deficiency involved elucidating and overcoming previously unidentified barriers to information flow within PHE and between the Office for National Statistics (ONS) and the CAS.

Secondly, early descriptive analyses revealed that almost one in five (17.8%) women within our cohort were recorded as having stage IV disease at presentation, or distant metastasis within 3 months of diagnosis. This was not realistic based on the published literature or reconcilable with clinical experience. A closer examination of this variable revealed that most women recorded as having stage IV disease had in fact unknown stage (referred to as 'stage VI' within PHE) and had been mistakenly classified as Stage IV. Once this was corrected, consideration was given to the remaining women with unknown stage recorded in CAS. By developing simple rules based on data items recorded within HES data, it was possible to infer the cancer stage at diagnosis for a substantial number of these women, leaving only 2.8% recorded as having stage IV disease at presentation, or distant metastasis within 3 months of diagnosis.

Thirdly, during the process of algorithm development, a number of important errors in the recording of cancer related data items within the cancer registries were identified and

required correction before this study could continue. An example was the variable ‘curative surgery’ within CAS, which was intended to describe the breast cancer related surgical procedure each woman had received. Since this variable had historically been completed within PHE in the absence of clinical expertise, this variable included non-breast cancer procedures such as ‘total cholecystectomy’, ‘ileectomy and anastomosis of Ileum to Ileum’ and ‘unspecified operations on aberrant adrenal tissue’. A process of manual reclassification of this variable was required before progress on the algorithm could be made.

Fourthly, the study involved establishing a consensus definition of breast cancer recurrence with the support of the National Cancer Registration and Analysis Service (NCRAS) Breast Cancer Site Specific Cancer Reference Group (SSCRG). In this study, a broad and pragmatic definition of recurrence, appropriate for epidemiological purposes, was used (11, 12). This definition described breast cancer recurrence as any subsequent IBC diagnosis, following 3 months from the date of registration of the incident breast cancer, recorded in the ipsilateral or contralateral breast or axilla, as a distant metastasis, as a record of breast cancer recurrence with unknown location, or as death with breast cancer recorded as the underlying cause.

Only after the resolution of these four challenges was it possible to successfully develop, test and refine an algorithm to reliably identify breast cancer recurrences.

6.3. Results

A total of 52,446 women living in the West Midlands, and registered with IBC during 1997-2011, were included in the population studied. Of these, 4164 women had either breast cancer related events, death or another cancer recorded within 3 months of their incident IBC diagnosis, and therefore were not considered further, leaving a total of 48,282 women for study.

6.3.1. Recurrence results from the algorithm executed using routinely collected data

Of the 48,282 women, 5428 (11.2%) had an event not related to breast cancer prior to any further recorded breast cancer event. These included a new primary cancer (1563 women, 3.2%), death from a cause other than breast cancer (3745 women, 7.8%), death with no cause recorded in the cancer registry (44 women, 0.1%) or embarkation (76 women, 0.2%) (Table 6.1).

Table 6.1: Number of women classified according to first recorded outcome event using information from the routinely collected data only, without the addition of the GRACE dataset. Women with an outcome event <3 months from date of diagnosis are excluded.

First recorded outcome event		Routinely collected data (excluding data items from GRACE)	
		Number of women	%
Not related to breast cancer	New primary non-breast cancer	1563	3.2%
	Death from cause other than cancer	3745	7.8%
	Death with no cause recorded in the registry	44	0.1%
	Embarkation	76	0.2%
	Subtotal	5428	11.2%
Locoregional recurrence	Recurrence in ipsilateral breast	3570	7.4%
	Recurrence in ipsilateral regional lymph nodes	995	2.1%
	Recurrence in regional lymph nodes with no laterality recorded	469	1.0%
	Subtotal	5034	10.4%
Distant metastasis	Confirmed distant metastases to liver, lung, bone or brain	2588	5.4%
	Inferred distant metastasis to liver, lung, bone or brain	2062	4.3%
	Distant metastasis - other locations	378	0.8%
	Record of palliative care	14	0.0%
	Subtotal	5042	10.4%
Recurrence, type unknown		3436	7.1%
Contralateral breast cancer		541	1.1%
Death from breast cancer with no record of previous recurrence		1793	3.7%
No events recorded by end of follow-up		27,008	55.9%
All women		48,282	100.0%

A further 5034 women (10.4%) had a locoregional recurrence. Of these, 3570 women (7.4%) had a recurrence in the ipsilateral breast and 995 women (2.1%) had a recurrence in the ipsilateral regional lymph nodes. A further 469 women (1.0%) also had a recurrence in the regional lymph nodes but no laterality was recorded, and thus were assumed to be ipsilateral recurrences.

A total of 5042 women (10.4%) had a distant metastasis. 2588 women (5.4%) had a confirmed metastasis to one of the four common sites of breast cancer metastases (liver, lung,

bone or brain). An additional 2062 women (4.3%) had a suspected distant metastasis to one of these four sites but histological tissue confirmation was not recorded. 378 women (0.8%) had a distant metastasis recorded in an anatomical location other than the four common sites listed above. A further 14 women (<0.1%) had a record of palliative care, and these were assumed to have had a distant metastasis although no confirmed site of metastasis was available. 3436 women (7.1%) had data items suggesting a recurrence, however the type could not be identified. A further 541 women (1.1%) had a contralateral breast cancer. 1794 women (3.7%) died with breast cancer as the underlying cause of death recorded on their death certificate, but with no record of a previous recurrence. In total, 27,008 women (55.9%) remained under follow-up with no recurrence-related events at the end of the study period.

6.3.2. Evaluation of algorithm using the GRACE data

There was complete agreement between the routinely collected data and the GRACE data as to which women were removed from follow up due to the diagnosis of a new primary cancer other than breast cancer, death from other causes (or with no cause recorded) or embarkation as the two datasets were based on the same sources of information for these outcomes (Table 6.2). 97.6% (15,846 out of 16,237) women recorded as having any type of recurrence (either locoregional, distant metastasis, recurrence with type unknown, contralateral breast cancer or death from breast cancer) in the GRACE data were identified as such from the routinely collected data. Of the 391 women with a recurrence recorded in the GRACE data that was not identified by the algorithm using the routinely collected data, 320 women were classified by the algorithm as having no recurrence, 55 women as having died from a cause other than breast cancer, 14 women as having new primary (non-breast) cancer diagnosis, 1 woman as having died with no cause recorded in the cancer registry, and 1 woman as having embarked.

Table 6.2: Cross Tabulation of the numbers of women with non-breast cancer related first events, any recurrences and no recorded recurrences recorded in the GRACE data (rows), with the same categories identified by the algorithm using routinely collected data (columns).

		Without GRACE					Total
		New primary non-breast cancer	Death from cause other than cancer	Death with no cause recorded in the registry	Embarkation	Any recurrence	
With GRACE	New primary non-breast cancer	1549 (100.0%)	—	—	—	—	1549 (3.2%)
	Death from cause other than cancer	—	3690 (100.0%)	—	—	—	3690 (7.6%)
	Death with no cause recorded in the registry	—	—	43 (100.0%)	—	—	43 (0.1%)
	Embarkation	—	—	—	75 (100.0%)	—	75 (0.2%)
	Any recurrence	14 (0.1%)	55 (0.3%)	1 (0.0%)	1 (0.0%)	15,846 (97.6%)	16,237 (33.6%)
	No recurrence	—	—	—	—	—	26,688 (100.0%)
	Total	1563 (3.2%)	3745 (7.8%)	44 (0.1%)	76 (0.2%)	15846 (32.8%)	48,282 (100.0%)

Further detail as to the level of agreement between the GRACE data and the algorithm using routinely collected data by the type of recurrence identified is given in Table 6.3. Contralateral breast cancer and distant recurrence had excellent agreement whereby 527/528 (99.8%) and 4784/4882 (98.0%) of events recorded in the GRACE data were correctly identified by the algorithm using the routinely collected data respectively. 1 woman classified as having contralateral breast cancer in GRACE was classified as having had a locoregional recurrence by the routinely collected data. This was likely due to a discrepancy in data recorded for the laterality variable. 98 women had a distant recurrence classification in GRACE that was not detected as such by the routinely collected data. Of these, 14 were classified by the algorithm using the routinely collected data as having had a recurrence of unknown type, 1 was classified as having a locoregional recurrence, 51 were classified by the algorithm as having died from breast cancer prior to a diagnosis of recurrence and 8 were identified with a new (non-breast primary cancer) or non-breast cancer death as their first event and for the remaining 24 women, no events were identified by the algorithm using routinely collected data before the end of follow up.

2717/2966 (91.6%) women had a recurrence recorded in the GRACE data, but with type unknown, and were identified as such by the algorithm using the routinely collected data. 250 women had a recurrence of unknown type recorded in the GRACE data that was not detected as such by the routinely collected data. Of these, 56 women were classified by the algorithm using the routinely collected data as having locoregional recurrence, 38 as having distant metastasis, 3 as having contralateral breast cancer, 17 were classified by the algorithm as having died from breast cancer prior to a diagnosis of recurrence, 11 were identified with a new (non-breast primary cancer) or non-breast cancer death as their first event and for the 125 women, no events were identified by the algorithm using routinely collected data before the end of follow up.

For locoregional recurrence the agreement between the two datasets was poorer with only 4976/6222 women (80%) recorded as having a locoregional recurrence in the GRACE data identified as such from the routinely collected data. 1246 women had a locoregional recurrence recorded in the GRACE data that was not detected as such by the routinely collected data. Of these, 706 women were classified by the algorithm using the routinely collected data as having 'recurrence of unknown location', 220 were identified as distant metastasis, 11 as contralateral breast cancer, 86 were classified by the algorithm as having died from breast cancer prior to a diagnosis of recurrence, 52 were identified with a new (non-breast primary cancer) or non-breast cancer death as their first event and for the remaining 171 women, no events were identified by the algorithm using routinely collected data before the end of follow up.

Table 6.3: Cross Tabulation of the numbers of women with different types of recurrences, who were removed from follow-up or who did not have a recurrence, recorded in the GRACE data (rows), with the same categories identified by the algorithm using routinely collected data (columns).

		Without GRACE						Total patients
		Locoregional recurrence	Distant metastasis	Recurrence, type unknown	Contralateral breast cancer	Death from breast cancer	Removed from follow up	
With GRACE	Locoregional recurrence	4976 (80.0%)	220 (3.5%)	706 (11.3%)	11 (0.2%)	86 (1.4%)	52 (0.8%)	6222 (12.9%)
	Distant metastasis	1 (0.0%)	4784 (98.0%)	14 (0.3%)	—	51 (1.0%)	8 (0.2%)	4882 (10.1%)
	Recurrence, type unknown	56 (1.9%)	38 (1.3%)	2716 (91.6%)	3 (0.1%)	17 (0.6%)	11 (0.4%)	2966 (6.1%)
	Contralateral breast cancer	1 (0.2%)	—	—	527 (99.8%)	—	—	528 (1.1%)
	Death from breast cancer	—	—	—	—	1639 (100.0%)	—	1639 (3.4%)
	Removed from follow up	—	—	—	—	—	5357 (100.0%)	5357 (11.1%)
	Under follow up with no recurrence	—	—	—	—	—	26,688 (100.0%)	26,688 (55.3%)
	TOTAL	5034 (10.4%)	5042 (10.4%)	3436 (7.1%)	541 (1.1%)	1793 (3.7%)	5428 (11.2%)	48,282 (100.0%)

6.3.3. Evaluation of recurrence dates identified by the algorithm using the GRACE data

The cumulative risk of any recurrence, both recorded in GRACE data and identified by the algorithm using the routinely collected data, increases rapidly within the first 5 years following diagnosis and rises to 7% and 6.8% respectively, at 15 years (Figure 6.1). In the intermediate term, between 5 and 15 years following diagnosis, the difference in cumulative risk between the GRACE data and algorithm using the routinely collected data remains similar, differing by only 0.2% at 15 years.

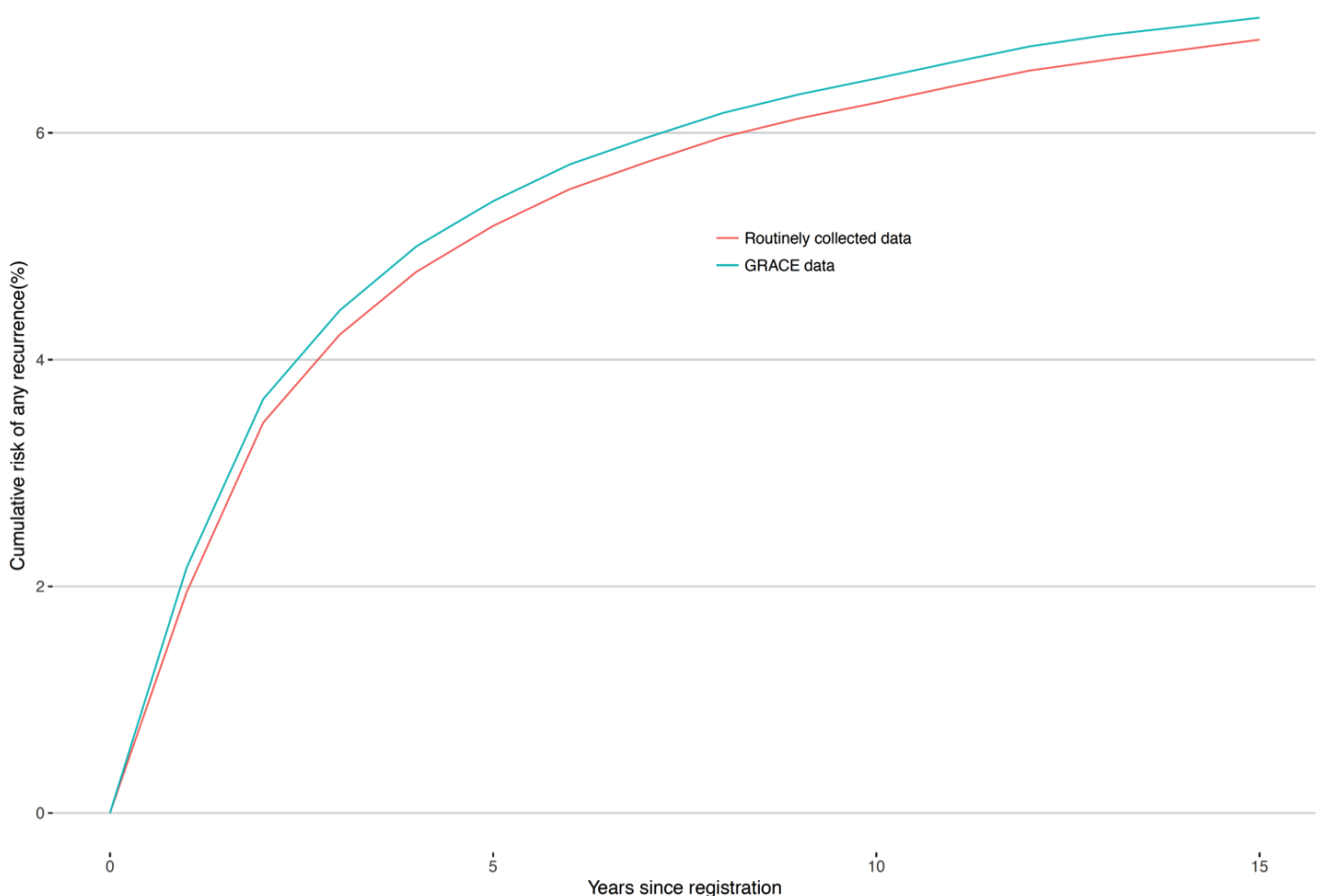


Figure 6.1: Comparison of cumulative risk of any recurrence by time since cancer registration identified by the algorithm using the routinely collected data (red) and the GRACE data (green).

Agreement as to the rate of any recurrence between the algorithm using routinely collected data and the GRACE data was generally good during each year following registration of breast cancer (Figure 6.2). However, the routinely collected data missed more recurrences in the earlier period following cancer registration (<2 years) than in the later period. Exploratory analyses of this finding revealed that this discrepancy is due to the difficulty in separating recurrence related treatment events from primary tumour related treatments using routinely collected data sources, particularly since this earlier period is when much of the primary cancer related treatment is undertaken. From year 5 onwards, however, the annual recurrence hazard rates from both data sources were almost identical.

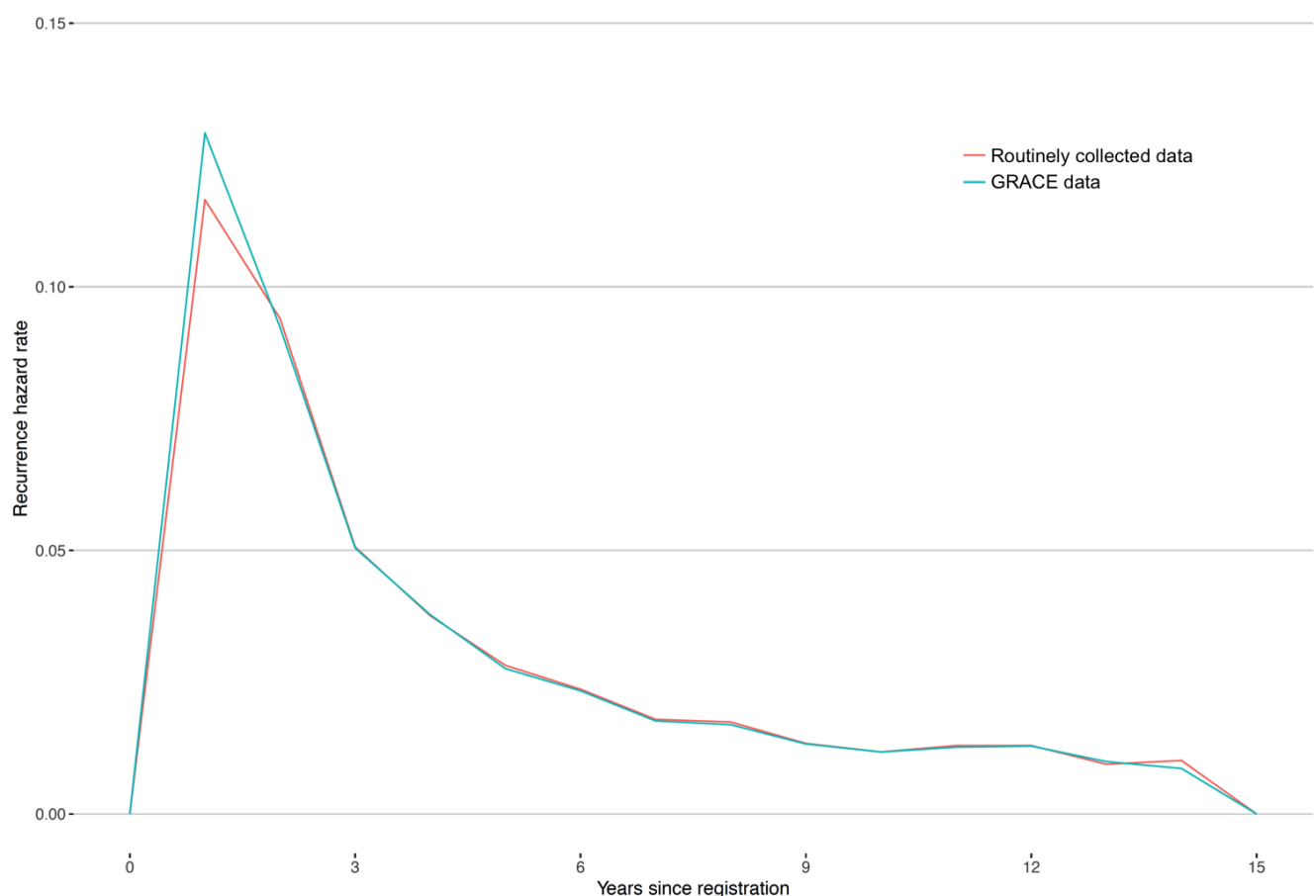


Figure 6.2: Comparison of recurrence hazard rate by time since cancer registration identified by the algorithm using the routinely collected data (red) and the GRACE data (green).

Further detail as to the level of agreement in the recurrence rate between the GRACE data and the algorithm using routinely collected data, by the type of recurrence identified, is given in Figure 6.3. There was good agreement in the rate of locoregional recurrence, distant metastasis, contralateral breast cancer and recurrence of unknown type between the GRACE data and the routinely collected data. The rate of locoregional recurrence in the GRACE data was slightly higher than in the routinely collected data between years 2 and 6 following diagnosis, whereas the rate of recurrence of unknown type in the GRACE data was lower than in the routinely collected data during the same period. This was due to the 706 women with locoregional recurrence recorded in the GRACE data that were identified as having recurrence of unknown type in the routinely collected data (Table 6.3)

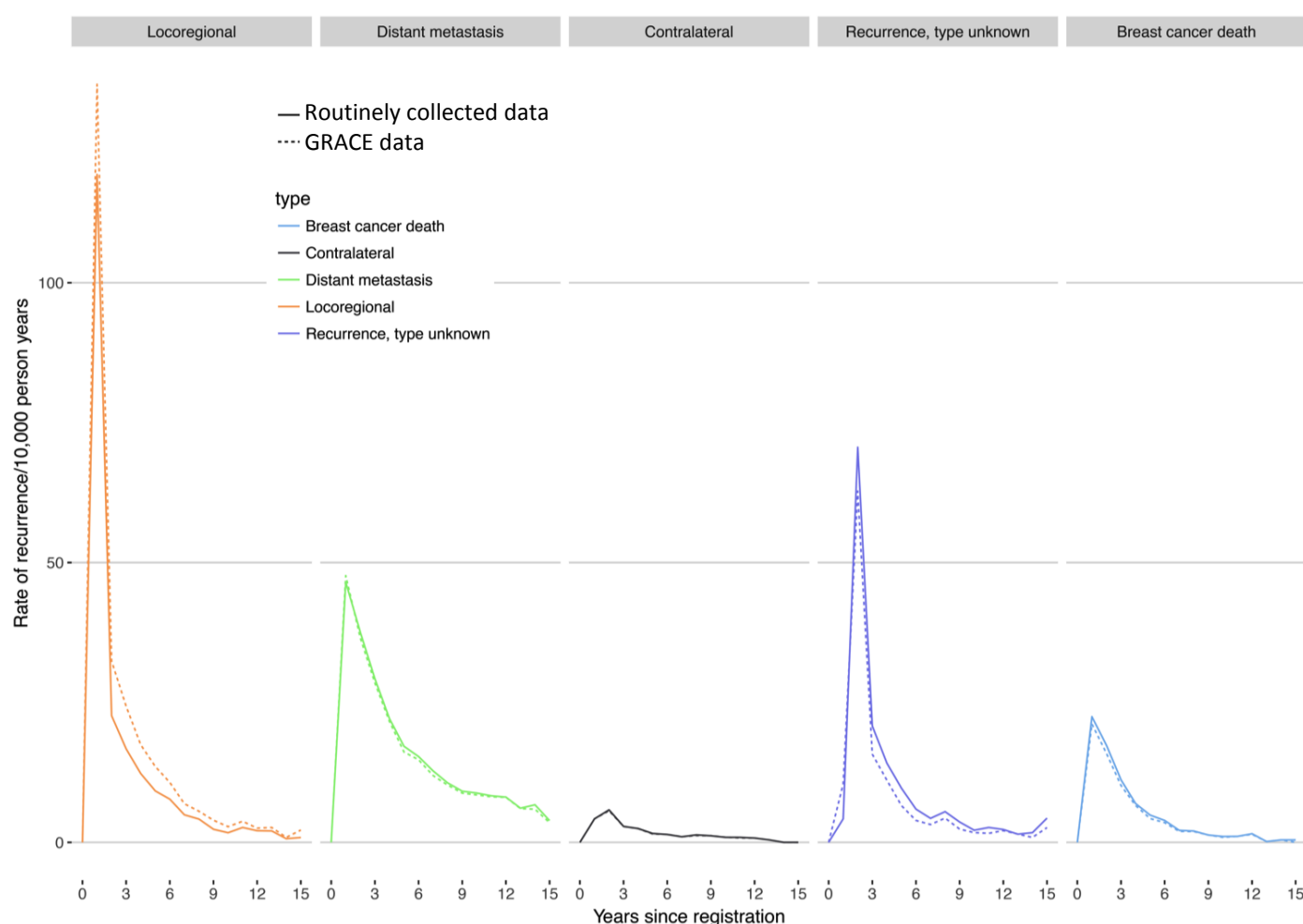


Figure 6.3: Comparison of the recurrence rate of different types of recurrence, by time since cancer registration between routinely collected data and the GRACE data

6.4. Discussion

This study represents the most comprehensive approach to the use of routinely collected data to identify recurrences of breast cancer at the population level in the UK. In the test population, 97.6% of all women with a breast cancer recurrence recorded in the GRACE data were identified as such from the routinely collected data. In addition, the recurrence rate by time since diagnosis in the GRACE data was similar to that identified using the routinely collected data. This study has, therefore, demonstrated the feasibility of using this approach and its potential in providing breast cancer recurrence information on a nationwide level both for women diagnosed with breast cancer in the past and for women diagnosed in the future.

6.4.1. The growing need for information on recurrence

Four important factors have influenced the landscape of clinical research, particularly randomised trials, in breast cancer in the UK. Firstly, as a beneficial consequence of advances in cancer treatments overall, women entering modern clinical trials are surviving for longer than in earlier trials, which in turn means follow-up is needed for a longer period of time than ever before. Secondly, some important side-effects of breast cancer treatments only become clinically important after more than a decade following cancer diagnosis, so that in modern clinical trials, follow-up needs to be continued over a longer period of time than previously (264). Whereas these first two points have encouraged the need for longer follow-up to examine the benefits and risks of treatments, a third important consideration is that progress in breast cancer has been made due to studies evaluating incremental modifications in treatment strategies and this continues to be likely for the foreseeable future. As a result, large trials are needed to provide sufficient power to detect small differences between treatments being compared, meaning more women require follow-up. Fourthly, changes in

data protection, confidentiality legislature and information governance regulations have meant conducting clinical trial follow-up is more resource intensive than in the past. As a result of these four points, trialists and research funding organisations now face the challenge of planning detailed, reliable follow-up of women randomised into modern trials whilst minimising the expense in time, manpower and bureaucracy resulting from the factors above (265). The use of routinely collected data for the purpose of clinical trial follow-up offers a viable solution to these problems by allowing clinical trials to be conducted more efficiently and cost-effectively than ever before.

6.4.2. Relation between local recurrence and subsequent breast cancer death

Epidemiological research into outcomes after a diagnosis of breast cancer in the UK has traditionally focused on evaluating factors affecting overall survival or breast cancer specific mortality. The absence of reliable recurrence information has been the major reason why recurrence has been used less frequently than these other outcome measures, but another reason has been uncertainty regarding the importance of recurrence in terms of health economic and clinical considerations. Recent research has affirmed the influence of local recurrence on subsequent breast cancer death (266), as well as the substantial health economic cost attributable to both local and distant recurrences (267).

Evidence suggests that a shorter disease-free interval between the date of primary diagnosis and date of local recurrence and contralateral breast cancer are determinants of poor prognosis (268-270). The presence of locoregional recurrence may also act as a marker of occult metastatic disease, tumour aggressiveness, large tumour burden, or depleted host defences (271, 272). Although the underlying mechanisms are unclear, the presence of

locoregional recurrence has equally important repercussions for post-treatment surveillance and follow-up strategies (271).

6.4.3. Validity of the algorithm

Whilst able to detect 97.6% of all recurrences, 2.4% of recurrences recorded in the GRACE data were not detected by the algorithm when using data items from routinely collected data sources. This raises a number of important points: Firstly, it raises the question of to what extent the GRACE data is an appropriate validation cohort. Whereas there is no doubt that the GRACE data represents the most complete source of breast cancer recurrence information within PHE, without comparison with an independently collected recurrence dataset, it is impossible to confirm its reliability. This could mean that the algorithm may perform better than has been reported here.

The method of validation used in this study may represent another reason for underestimation of the algorithm's performance as when the algorithm was used on the routinely collected data, care was taken to ensure that these data did not contain any data items that may have originated from GRACE. This stringent criterion would not be applied when examining the performance of the algorithm on a population level, where the algorithm may again, perform better than has been reported here. To investigate these issues further, the next step in this research is to repeat the evaluation exercise using data from a UK based randomised controlled trial, the AZURE trial, as the testing dataset (273). Ethical approvals are in place for this additional validation and this research is currently underway.

An additional issue is that this work raises the question of what degree of accuracy is acceptable for epidemiological use and for follow-up in randomised trials. The algorithm developed during this study is not aimed for clinical purposes, where reliable individual

patient level outcomes are critical, but instead for use in clinical research and epidemiological analyses at a population level, where findings of differences in outcomes by patient, tumour or treatment factors can accept a wider degree of inaccuracy, particularly in the early period i.e. <3 years from date of diagnosis, provided that missed recurrences are not closely associated with particular treatment modalities. In any case, there is promising potential for improving the accuracy of the algorithm even further, such as by the addition of more data items from newer datasets, such as DID and the Systemic Anti-Cancer Therapy (SACT) dataset.

6.4.4. Conclusion

In the test population, there was good agreement between recurrences indicated by our algorithm using routinely collected data and those recorded by the former WMCIU in the GRACE dataset. This finding supports the potential value of compiling recurrence information on a nationwide basis from routinely collected data, for use in future descriptive and epidemiological studies and in randomised trials. Additional work is needed to confirm the accuracy of information on recurrences identified from routinely collected data, and, if possible, to improve it further before considering its use on a nationwide basis.

6.5. References

1. ONS. Breast Cancer: Incidence, Mortality and Survival, 2010 London: Office for National Statistics; 2012 [Available from: <http://www.ons.gov.uk/ons/rel/cancer-unit/breast-cancer-in-england/2010/sum-1.html>].
2. CRUK. Breast cancer incidence statistics: Cancer Research UK; 2014 [updated 29/5/2014. Available from: <http://www.cancerresearchuk.org/cancer-info/cancerstats/types/breast/incidence/uk-breast-cancer-incidence-statistics>].
3. Intelligence WMC. Recurrent and Metastatic Breast Cancer Data Collection Project - Pilot report 2012 [Available from: <http://www.ncin.org.uk/view?rid=1043>].
4. Bettencourt-Silva JH, Mannu GS, de la Iglesia B. Visualisation of Integrated Patient-Centric Data as Pathways: Enhancing Electronic Medical Records in Clinical Practice. In: Holzinger A, editor. Machine Learning for Health Informatics: State-of-the-Art and Future Challenges. Cham: Springer International Publishing; 2016. p. 99-124.
5. WHO. ICD-10 Version: 2016 World Health Organisation; 2016 [Available from: <http://apps.who.int/classifications/icd10/browse/2016/en>].
6. Early Breast Cancer Trialists' Collaborative G. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365(9472):1687-717.
7. Kilburn LS, Banerji J, Bliss JM, Group NBCS. The challenges of long-term follow-up data collection in non-commercial, academically-led breast cancer clinical trials: the UK perspective. *Trials*. 2014;15:379.
8. Vicini FA, Kestin L, Huang R, Martinez A. Does local recurrence affect the rate of distant metastases and survival in patients with early-stage breast carcinoma treated with breast-conserving therapy? *Cancer*. 2003;97(4):910-9.
9. Stokes ME, Thompson D, Montoya EL, Weinstein MC, Winer EP, Earle CC. Ten-year survival and cost following breast cancer recurrence: estimates from SEER-medicare data. *Value Health*. 2008;11(2):213-20.
10. Anderson SJ, Wapnir I, Dignam JJ, Fisher B, Mamounas EP, Jeong JH, et al. Prognosis after ipsilateral breast tumor recurrence and locoregional recurrences in patients treated by breast-conserving therapy in five National Surgical Adjuvant Breast and Bowel Project protocols of node-negative breast cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*. 2009;27(15):2466-73.
11. Tanis E, van de Velde CJ, Bartelink H, van de Vijver MJ, Putter H, van der Hage JA. Locoregional recurrence after breast-conserving therapy remains an independent prognostic factor even after an event free interval of 10 years in early stage breast cancer. *Eur J Cancer*. 2012;48(12):1751-6.
12. Vichapat V, Garmo H, Holmberg L, Fentiman IS, Tutt A, Gillett C, et al. Prognosis of metachronous contralateral breast cancer: importance of stage, age and interval time between the two diagnoses. *Breast cancer research and treatment*. 2011;130(2):609-18.
13. Melvin JC, Purushotham AD, Garmo H, Pinder SE, Fentiman IS, Gillett C, et al. Progression of breast cancer following locoregional ipsilateral recurrence: importance of interval time. *British journal of cancer*. 2016;114(12):e23.
14. Botteri E, Bagnardi V, Rotmensz N, Gentilini O, Disalvatore D, Bazolli B, et al. Analysis of local and regional recurrences in breast cancer after conservative surgery. *Ann Oncol*. 2010;21(4):723-8.
15. Coleman R, Cameron D, Dodwell D, Bell R, Wilson C, Rathbone E, et al. Adjuvant zoledronic acid in patients with early breast cancer: final efficacy analysis of the AZURE (BIG 01/04) randomised open-label phase 3 trial. *The Lancet Oncology*. 2015(9):997-1006.

Chapter 7: Discussion and Conclusion

7.1. Discussion

Whilst almost 7,000 people are diagnosed with DCIS in the United Kingdom each year, many aspects of its natural history and optimal management remain unclear. The general aim of this thesis was to study the epidemiology of DCIS, including both the risk factors for the disease and the outcomes following its diagnosis. This has been achieved in four ways: first by quantifying the similarities and differences in the modifiable and non-modifiable risk factors for developing DCIS with those for developing IBC; second by evaluating pathological and radiological factors influencing the accuracy of DCIS diagnosis on pre-operative biopsy; third by evaluating the patient, tumour and treatment factors contributing to the rate of IBC following treatment for DCIS; and finally by developing a mechanism to permit the identification of breast cancer recurrences for use in future randomised trials, epidemiological studies and other purposes.

Chapters 1 and 2 have summarised the relevant literature on the pathophysiology and epidemiology of DCIS and have highlighted that, despite its increasing incidence, there remain a number of fundamental and clinically important uncertainties relating to its aetiology, diagnosis and treatment. Research into each of these areas has the potential to benefit large numbers of women each year.

Chapter 3 presented one of the most detailed comparisons of risk factors for DCIS and IBC to date, and contributes to the previously scant literature on this issue. This study confirmed similarities in the risk factors for both conditions, but revealed differences in the magnitude of their effect. Postmenopausal BMI was associated with an increased rate of developing both DCIS and IBC, and the number of live births was associated with a decreased rate of developing both conditions. However, the magnitude of the effect of these two factors

differed between DCIS and IBC. The effect of postmenopausal BMI ≥ 35 kg/m² on the rate of developing DCIS when compared with BMI < 25 kg/m², was more than double its effect on developing IBC (RR 2.35, 95% CI 1.14–4.82). In addition, increasing number of live births had a larger protective effect for DCIS than for IBC (p for trend = 0.03). These findings provide insights into the contribution of BMI and endocrine factors associated with parity at different stages in the process of carcinogenesis.

Over the years there has been a general trend towards less invasive surgical treatment of breast cancer. Contemporary studies evaluating the management of low-risk DCIS have pioneered the concept of non-operative treatment for the condition. The coming decade will herald the publication of findings from two such randomised trials comparing non-operative treatment with surgery in DCIS and by doing so, will determine whether a further paradigm shift in its treatment would be appropriate (20, 21). However, these trials rely upon an accurate biopsy determination of invasive status and DCIS grade. In light of this, **Chapter 4** evaluated 589 women diagnosed with DCIS on preoperative biopsy and is amongst the largest studies examining factors associated with an upstaged diagnosis of IBC following surgery and also with an upgraded diagnosis to a higher grade of DCIS. This study demonstrated that both lesion size and the presence of clinical symptoms at the time of biopsy were significantly associated with upstaging of biopsy diagnoses of DCIS to IBC following surgery, and that lesion size, the method of biopsy (14G core biopsy) and the presence of necrosis on biopsy were all significantly associated with upgrading of biopsy diagnoses to a higher grade of DCIS following surgery. By doing so, these findings provide important considerations for the selection of women with a biopsy diagnosis of DCIS for non-operative management.

Whereas the clinical trials of non-operative treatment for DCIS will provide invaluable information on the natural history of DCIS in the absence of surgery, there still remain uncertainties on the long-term prognosis of women with DCIS who are treated with surgery. To date there has been no nationwide study of long-term outcomes in individual women diagnosed with screen-detected DCIS in the UK, nor of how these outcomes have changed over time or vary with differences in the characteristics of the patient, the tumour, or the treatment received. To this end, **Chapter 5** presented the largest study of its kind in the UK and has described the long-term prognosis of women diagnosed with screen-detected DCIS and treated under different management strategies. This study has shown that the cumulative risk of any IBC after screen-detected DCIS is 9.4% at 15 years and that women diagnosed in more recent years (2010 - 2015) have lower rates of ipsilateral IBC than those diagnosed even a decade earlier.

The results presented in **Chapter 5** contribute towards the discussion on optimal margin distance in DCIS by showing that women with clear surgical margins of 1-2 mm had a significantly higher ipsilateral IBC rate than women with clear margins of 5+ mm (RR 1.85, 95% CI 1.20– 2.84). Furthermore, they provided novel evidence that the pattern of risk for high grade DCIS, and for low or intermediate grade DCIS, changes over the period of follow-up. Furthermore, this work provides evidence that the benefits of hormone therapy and radiotherapy identified from randomised clinical trials are being carried over to women diagnosed with screen-detected DCIS in the general population. These results, together with evidence from randomised trials, can help clinicians to discuss the risks and benefits of different treatment strategies with women diagnosed with screen-detected DCIS.

Whilst **Chapters 3, 4 and 5** have described risk factors for developing DCIS, accuracy of its initial diagnosis, and factors associated with the occurrence of IBC following treatment for DCIS respectively, recurrence of DCIS has not been discussed. The reason for this is that much less information is available on DCIS recurrence, or on recurrence more generally even after IBC. Reliable information on breast cancer recurrence would be useful for many purposes, including follow-up in randomised trials, healthcare policy-making, guiding clinical management, and for epidemiological research. **Chapter 6** presented a methodological study that identified and collated all sources of data relevant to breast cancer recurrence that are held within PHE, developed an algorithm to identify recurrence using these data, identified a population in which to test the proposed algorithm, and examined the performance of the algorithm when applied to this test population. Whilst this work has focused on IBC as a proof-of-concept model, the goal is to extend it to be to identify DCIS recurrences in future research.

Chapter 6 represents the most comprehensive approach to the use of routinely collected data to identify recurrences of IBC at the population level in the UK. In the test population, 97.6% of all women with any breast cancer recurrence recorded in the test dataset were identified as such from the routinely collected data. In addition, the recurrence rate by time since diagnosis in the test data was similar to that identified by the algorithm using the routinely collected data. This study has, therefore, demonstrated the feasibility of using this approach and its potential in providing breast cancer recurrence information for research purposes on a nationwide level both for women diagnosed with breast cancer in the past and for women diagnosed in the future.

The findings from **Chapter 6** support the potential value of compiling recurrence information on a nationwide basis from routinely collected data, for use in future descriptive and epidemiological studies and in randomised trials. Additional work is needed to confirm the accuracy of information on recurrences identified from routinely collected data, and, if possible, to improve it further before considering using it routinely on a nationwide basis. Once this has been completed, the intention is for this work to then be extended to identify IBC recurrences after DCIS, as well as to identify DCIS recurrences.

There remains a pressing need amongst clinicians and patients for more information concerning the risks and benefits of treatments for DCIS and the prognosis of low-risk DCIS in the absence of treatment. Whilst randomised trials can answer some of the outstanding questions, epidemiological analyses of observational data, such as those presented in this thesis, also have an important role to play in investigating many of the uncertainties around optimal management that cannot be evaluated by randomised trials.

7.2. Conclusions

(1) DCIS and IBC have similar risk factors, however the magnitude of the effects differed between DCIS and IBC. The effect of postmenopausal BMI ≥ 35 kg/m² was larger for DCIS than for IBC (RR 2.35, 95% CI 1.14–4.82), and the decrease in risk with each additional live birth was greater for DCIS than for IBC ($p = 0.03$).

(2) Consideration of lesion size and the presence of necrosis on biopsy may be helpful in selecting low-risk women for non-operative management of DCIS in the future, as may use of the 9G vacuum-assisted method of biopsy.

(3) In the general population of England, the cumulative risks of any (ipsilateral or contralateral) IBC at 5, 10 and 15 years after screen-detected DCIS were 3.5%, 7.1%, and 9.4% respectively and the cumulative risks of dying from breast cancer after DCIS were 0.3%, 1.1%, and 2.3% respectively.

(4) In England, women given breast-conserving surgery (BCS) with no radiotherapy had a higher ipsilateral IBC rate than women given BCS with radiotherapy (RR 1.63, 95% CI 1.27–2.10), and women with clear surgical margins of 1–2 mm had a higher ipsilateral IBC rate than women with clear margins of 5+ mm (RR 1.85, 95% CI 1.20–2.84).

(5) In England, ER+ women given hormone therapy had a lower rate of any (ipsilateral or contralateral) IBC compared with ER+ women who were not given hormone therapy (RR 0.76, 95% CI 0.63–0.93).

(6) Routinely collected data sources can be used to identify breast cancer recurrences on a nationwide basis, for use in future descriptive and epidemiological studies and in randomised trials.

7.3. References

1. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *European Journal of Cancer*. 2015;51(16):2296-303.
2. Elshof LE, Tryfonidis K, Slaets L, van Leeuwen-Stok AE, Skinner VP, Dif N, et al. Feasibility of a prospective, randomised, open-label, international multicentre, phase III, non-inferiority trial to assess the safety of active surveillance for low risk ductal carcinoma in situ – The LORD study. *European Journal of Cancer*. 2015;51(12):1497-510.

Appendices

Contents of appendices

Appendix A: Permissions and lay summary of approved UK Biobank study protocol. Taken from (38), Chapter 3	166-167
Appendix B: Study protocol for Chapter 4.	168-173
Appendix C: Pathology data collection form	174
Appendix D: Clinical data collection form.....	175
Appendix E: A summary of the respective inclusion and exclusion criteria for the two ongoing trials randomising women with DCIS to surgical excision or active surveillance.	176
Appendix F: Suggestions to the criteria used to identify low risk DCIS cases shown in supplementary table 3 which may help to minimise the risk of DCIS grade underestimation amongst women appropriate for non-operative management.....	177
Appendix G: NHS Breast Cancer Screening Programme (BCSP) Research Committee approval letter for Chapter 5.	178
Appendix H: Research Ethics Committee approval letter for Chapter 5.	179-183
Appendix I: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 13,528 women who were treated with breast conserving surgery alone and were not recorded as receiving hormone therapy.....	184
Appendix J: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 5157 women who were treated with breast conserving surgery and radiotherapy and not recorded as receiving hormone therapy.	185
Appendix K: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 7998 women who were treated with mastectomy and not recorded as receiving hormone therapy.	186
Appendix L: Description of algorithm for Chapter 6.	187-199
Appendix M: Contributions to and Planned Extensions of Studies in the Thesis.....	200-202

Appendix A: Permissions and lay summary of approved UK Biobank study protocol. Taken from (206), Chapter 3.



Principal Investigator

Dr Gurdeep Mannu

Address

University of Oxford, CTSU, Nuffield Department of Population Health,
L1/40 Richard Doll Building, Old Road Campus, Roosevelt Drive, Oxford,
OX3 7LF

Summary of research

Key words: Cancer, Breast, Diet, Lifestyle

Application lay Summary:

1a: Although ductal carcinoma in situ (DCIS) is considered a pre-invasive condition; it can progress to invasive breast cancer or death over a number of years. It is unclear which women with DCIS will develop invasive cancer within their lifetime and over what period of time. This project will investigate the risk factors for developing DCIS and its subsequent progression to invasive cancer and aim to quantify a woman's risk of disease progression based on her lifestyle, dietary, socioeconomic, medical and genetic factors in order to help clinicians tailor diagnosis/screening, follow-up and treatment options according to individual patient risk-profiles.

1b: The objective of the UK Biobank is to undertake research in improving the prevention, diagnosis and treatment of a wide range of serious illnesses including breast cancer. By investigating the risk factors for developing DCIS or having a poorer outcome after DCIS treatment, it may be possible to quantify a woman's risk based on her individual risk factors and so tailor prevention, diagnosis, and treatment options accordingly. Hence the aim of this project is in

synergy with that of the UK Biobank and could help to provide innovative new directions to improve healthcare for DCIS patients.

1c: Of the women within the UK Biobank diagnosed with DCIS, analyses will be conducted of patient, medical, socioeconomic, dietary, lifestyle and genetic factors to investigate possible risk factors for DCIS when compared against age matched controls without DCIS or invasive breast cancer and with those women who were diagnosed with invasive breast cancer without a prior diagnosis of DCIS. Further analyses of the women with DCIS will be conducted according to whether or not they were subsequently diagnosed with invasive breast cancer.

1d: Three groups ((i) DCIS patients (n=2000 women), (ii) age matched controls (n=8000 women), and (iii) primary invasive breast cancer patients (n=2000), ensuring at least 4 women as controls for each woman with DCIS or invasive breast cancer will be included in this study with a total of approximately 12,000 women. After 12-18 months a further data request will be made in order to increase the sample size of the cohort further by including new incident DCIS cases.

Appendix B: Study protocol for Chapter 4.

A study to investigate the accuracy of initial biopsy in DCIS and risk factors for upstaging a diagnosis of DCIS on biopsy to one of invasive cancer on surgical specimen analysis

Gurdeep Mannu, Sarah Darby, Flora van Leeuwen, Michael Schaapveld, Jelle Wesseling

Background: Core-needle (14G) or vacuum-assisted (9G) breast biopsy is usually performed under radiological guidance for a breast lesion suspicious for cancer. The biopsy sample is analysed by a pathologist and a histological diagnosis is made prior to a multidisciplinary team meeting where a treatment plan based on this diagnosis is discussed for each woman. However, a biopsy diagnosis of Ductal Carcinoma-In-Situ (DCIS) is sometimes upstaged to a diagnosis of invasive breast cancer (IBC) following the definitive histological analysis of the resected specimen, which is only available after surgery.

Quantifying the accuracy of the initial biopsy in women with DCIS and investigating the risk factors for missing invasive cancer in DCIS biopsies are of importance in two ways:

- (i) When IBC is identified only after surgery, patients usually need to undergo a second surgical procedure to evaluate the status of their axillary nodes. This has a financial cost and carries risks of medical complications and psychological distress;
- (ii) There is recent interest in treating low grade DCIS with observation alone without surgical excision (1). Under this scenario, any IBC that is missed at the time of biopsy would go undetected, and therefore untreated until later, by which time the patient's prognosis may be poorer.

There has been large variation in the reported accuracy of initial biopsy amongst studies focussing on DCIS patients (2-11). Furthermore many of these have been small scale, single-centre, retrospective studies which have provided little information on the relationship between radiological features and core-needle biopsy findings. Although there have been relatively larger scale studies comparing vacuum assisted biopsy with core needle biopsy in recent years, these again have not included detailed information on pathological, radiological and clinical factors (12-14). Despite the unsatisfactory nature of the available information, the initial breast biopsy result remains amongst the most influential factors in determining a woman's surgical management. There is, therefore, a need to carry out large-scale studies to obtain detailed and uniform data on this topic in order to guide future management decisions.

In light of this, Dr Pieter Westenend at the Albert Schweitzer hospital in the Netherlands has recently completed a study using the nationwide pathology register, PALGA, with the aim of providing further insights into this field. Although including large numbers of biopsies, this study has not re-assessed pathology samples, clinical records or radiological information and is restricted to the limited number of variables stored within PALGA. Furthermore, it also cannot compare the accuracy of vacuum assisted biopsy with core needle biopsy. As a result, there remains a need for a uniform, large-scale study on the topic which includes the detailed clinical, radiological and pathological data required to accurately and comprehensively investigate risk factors for underestimation of an initial biopsy in DCIS. The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital (NKI/AVL) could provide the ideal setting in which to conduct such a study.

Aims: To investigate:

- (i) the accuracy of initial biopsy in DCIS in the NKI/AVL;
- (ii) the risk factors for upstaging a biopsy of DCIS to IBC (or higher grade DCIS) upon pathological analysis of the resected specimen in the NKI/AVL;
- (iii) the relationship between radiological features and biopsy findings in women presenting symptomatically or via the national screening program in the NKI/AVL;
- (iv) the relationship between the above factors and the accuracy of core needle biopsy (14G) versus vacuum assisted biopsy (9G) in DCIS in the NKI/AVL.

Methodology: The NKI/AVL is a large specialist cancer centre with 180-beds, 5-operating theatres, 11 radiotherapy suites and receives approximately 27,000 new patients to its clinics every year. It has a large pathology database which contains information on the initial biopsy diagnosis as well as the final pathological diagnosis for each woman diagnosed with DCIS or IBC treated at the NKI/AVL. By comparing the biopsy histology with the final resected specimen histology, it will be possible to identify which women diagnosed with DCIS on initial biopsy were upstaged to IBC or to higher grade DCIS on final histology. This database includes biopsy results for approximately 1200 patients with DCIS, treated between 2008 and 2014. These raw data show that between 10% and 20% of biopsies for DCIS have underestimated the tumour when compared with the histopathological evaluation of the surgically resected specimen. By comparing differences in clinical, radiological and histopathological factors between those who were upstaged with those who were not, it will be possible to investigate risk factors for upstaging DCIS biopsies to IBC.

Whereas many previous studies have relied on retrospective datasets, this study will involve detailed re-examination and integration of pathological, clinical and radiological data. For biopsy specimens showing DCIS, a dedicated breast radiologist at the NKI/AVL will review the initial mammography and specimen images and assess radiological risk factors for under-estimation of DCIS biopsies. Furthermore a team of expert breast pathologists will review the histology for all women in this study, in order to assess pathological risk factors for biopsy under-estimation. By linking these radiological and pathological data to clinical records, it will be possible to further investigate patient or clinical presentation-related factors that might contribute to under-estimation of a biopsy in these cases. Furthermore it will be possible to examine these risk factors in relation to both vacuum-assisted (9G) and core needle biopsies (14G) in order to identify technique-specific risk profiles for each woman. The variables to be analysed will include the following:

Variables from pathology:

- Date of birth
- Presence of invasive cancer
- Age at diagnosis
- Location of tumour (laterality and quadrant)
- Pathological grade (1-3)
- Calcifications present in biopsy
- Stromal changes (i.e. concentric fibrosis in DCIS)
- Presence and amount of lymphocytic infiltrate
- Number of samples taken per biopsy
- Size of area with DCIS

- DCIS outside invasive tumour
- Growth pattern
- Presence of multifocality
- Presence of comedo necrosis
- Presence of microinvasion
- Paget's disease
- Minimum distance of DCIS to final margin
- Number of lymph nodes examined
- Number of nodes with macro- or micro-metastasis or isolated tumour cells
- ER status
- PGR status
- Her-2 status
- Type of surgery performed
- Adjuvant therapy given

From the radiological data:

- Breast density
- Site of lesion
- Predominant radiological feature (s)
- Microcalcifications present on mammogram, pattern, extensiveness (multifocality and multicentricity) and description according to the ACR BIRADS breast imaging atlas
- Visibility of mammographically (MMG) detected lesion at ultrasound (USS)
- Size of lesion on MMG and USS (radiological bi-dimensional & maximum diameter)
- Breast Imaging-Reporting and Data System category
- Method of biopsy: 9G vacuum assisted or 14G core needle biopsy
- Method of guided biopsy (clinical, USS, MMG, MRI)
- Additional imaging findings such as the absence of mass-forming lesions

From the clinical notes:

- Presentation type (whether symptomatic or screen-detected lesion)
- Symptoms at presentation
- If palpable, then clinical size and location

Method of analysis: Women will be placed into 4 groups (i-iv) as shown in red in figure 1, and detailed below:

- i. DCIS only on initial biopsy and same grade DCIS on final histology
- ii. DCIS only on initial biopsy and higher grade DCIS on final histology
- iii. DCIS only on initial biopsy and invasive cancer on final histology
- iv. Both DCIS and invasive cancer on initial biopsy

Analyses will compare women in group (i) with groups (ii) and (iii). Groups (ii) and (iii) will also be compared with group (iv). Regression techniques will be used to identify patient, tumour and radiological characteristics which are risk factors for under-staged biopsies (groups (ii) and (iii)). Further analyses will be conducted to compare the accuracy of core

needle (14G) and vacuum-assisted (9G) biopsies in DCIS. We will validate our findings in two series: (i) a nation-wide cohort of DCIS cases with or without concurrent IBC and (ii) a DCIS cohort from the Albert Schweitzer hospital (a large community hospital in the Netherlands), which also contains both detailed pathology and radiology data. Both validation series will be available from our collaborating breast pathologist, Dr. Pieter Westenend, MD, PhD.

Scientific significance: The percentages of core-needle biopsies showing DCIS that actually contain IBC based on the surgically excised specimen range from 17% to 48% in different studies (2-12). There are also several studies that have attempted to relate under-staging of core-needle biopsy in DCIS to possible risk factors. However, these studies have often examined only a limited number of variables or have been restricted by small numbers of patients. As a result they have produced heterogeneous results. Brennan and colleagues have published the largest meta-analysis to date on this question including 7350 cases of DCIS from 52 studies with 1736 under-staged samples (10). However, their work was based only on published data and so they were unable to undertake detailed analyses considering more than one variable at a time. Furthermore, few of the available studies reported all of the variables of interest or used comparable classification systems. In light of this, confounding factors could not be distinguished from true associations. As a result, there remains a need for a large-scale study with uniform high quality, detailed data to help to provide clearer information. The NKI/AVL is ideally suited to address this issue.

Clinical significance: By using data from the NKI/AVL to investigate (i) the accuracy of core needle biopsies (14G) and vacuum-assisted (9G) in DCIS and (ii) factors associated with underestimation of IBC in DCIS biopsies demonstrating DCIS, it will be possible to carry out by far the most detailed study to date on this issue. This will: (i) quantify the extent of the problem within the Netherlands and (ii) provide a platform from which more sophisticated comparisons using biomarkers to identify molecular risk factors for the presence of IBC alongside DCIS, can be undertaken. In addition to providing a substantial contribution to the existing evidence base, this study may help to influence contemporary clinical practice in the management of biopsy proven DCIS.

A large number of women currently undergo a second surgical procedure due to under-staged biopsy reports. By investigating risk factors for upstaging DCIS biopsies to IBC, risk-specific stratification of management options may be possible. A pragmatic example of the future implications of risk-stratifying treatment in this way could be by performing a sentinel lymph node biopsy (SLNB) at the same time as wide local excision of a DCIS lesion in a woman at very high risk for under-staged core-biopsy. This could alleviate the potential psychological, financial and medical complications of an unexpected IBC on final histology whilst balancing the risks from complications of SLNB. Furthermore, in recent years, vacuum-assisted biopsy has begun to be more widely used in place of core-needle biopsy in clinical and research settings (1, 11, 15). By comparing the accuracy of these two techniques in a large number of women with DCIS from the NKI/AVL, alongside radiological and clinical factors, clinically important information can be obtained that could help guide future management decisions.

Conclusion: This is an important clinical research question and information from the NKI/AVL could make a substantial contribution in helping to address it.

References:

1. Soumian S, Verghese ET, Booth M, Sharma N, Chaudhri S, Bradley S, Umranikar S, Millican-Slater RA, Hanby AM, Francis A. Concordance between vacuum assisted biopsy and postoperative histology: implications for the proposed Low Risk DCIS Trial (LORIS). *Eur J Surg Oncol.* 2013;39(12):1337-40.
2. Dillon MF, McDermott EW, Quinn CM, O'Doherty A, O'Higgins N, Hill AD. Predictors of invasive disease in breast cancer when core biopsy demonstrates DCIS only. *J Surg Oncol.* 2006;93(7):559-63.
3. King TA, Farr GH, Jr., Cederbom GJ, Smetherman DH, Bolton JS, Stoler AJ, Fuhrman GM. A mass on breast imaging predicts coexisting invasive carcinoma in patients with a core biopsy diagnosis of ductal carcinoma in situ. *Am Surg.* 2001;67(9):907-12.
4. Fuhrman GM, Cederbom GJ, Bolton JS, King TA, Duncan JL, Champaign JL, Smetherman DH, Farr GH, Kuske RR, McKinnon WM. Image-guided core-needle breast biopsy is an accurate technique to evaluate patients with nonpalpable imaging abnormalities. *Ann Surg.* 1998;227(6):932-9.
5. Walters LL, Pang JC, Zhao L, Jorns JM. Ductal carcinoma in situ with distorting sclerosis on core biopsy may be predictive of upstage on excision. *Histopathology.* 2014.
6. Park SH, Kim MJ, Kim SJ, Kim EK. Ductal carcinoma in situ diagnosed using an ultrasound-guided 14-gauge core needle biopsy of breast masses: can underestimation be predicted preoperatively? *Ultrasonography.* 2014;33(2):128-35.
7. Chang WC, Hsu HH, Yu JC, Ko KH, Peng YJ, Tung HJ, Chang TH, Hsu GC. Underestimation of invasive lesions in patients with ductal carcinoma in situ of the breast diagnosed by ultrasound-guided biopsy: a comparison between patients with and without HER2/neu overexpression. *Eur J Radiol.* 2014;83(6):935-41.
8. Park AY, Gweon HM, Son EJ, Yoo M, Kim JA, Youk JH. Ductal carcinoma in situ diagnosed at US-guided 14-gauge core-needle biopsy for breast mass: preoperative predictors of invasive breast cancer. *Eur J Radiol.* 2014;83(4):654-9.
9. Goto M, Yuen S, Akazawa K, Nishida K, Konishi E, Kajihara M, Shinkura N, Yamada K. The role of breast MR imaging in pre-operative determination of invasive disease for ductal carcinoma in situ diagnosed by needle biopsy. *Eur Radiol.* 2012;22(6):1255-64.
10. Brennan ME, Turner RM, Ciatto S, Marinovich ML, French JR, Macaskill P, Houssami N. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology.* 2011;260(1):119-28.
11. Suh YJ, Kim MJ, Kim EK, Moon HJ, Kwak JY, Koo HR, Yoon JH. Comparison of the underestimation rate in cases with ductal carcinoma in situ at ultrasound-guided core biopsy: 14-gauge automated core-needle biopsy vs 8- or 11-gauge vacuum-assisted biopsy. *The British journal of radiology.* 2012;85(1016):e349-56.
12. Ciatto S, Houssami N, Ambrogetti D, Bianchi S, Bonardi R, Brancato B, Catarzi S, Risso GG. Accuracy and underestimation of malignancy of breast core needle biopsy: the Florence experience of over 4000 consecutive biopsies. *Breast cancer research and treatment.* 2007;101(3):291-7.
13. Li JL, Wang ZL, Su L, Liu XJ, Tang J. Breast lesions with ultrasound imaging-histologic discordance at 16-gauge core needle biopsy: can re-biopsy with 10-gauge vacuum-assisted system get definitive diagnosis? *Breast.* 2010;19(6):446-9.
14. Wang ZL, Liu G, Li JL, Su L, Liu XJ, Wang W, Tang J. Breast lesions with imaging-histologic discordance during 16-gauge core needle biopsy system: would vacuum-assisted removal get significantly more definitive histologic diagnosis than vacuum-assisted biopsy? *The breast journal.* 2011;17(5):456-61.
15. Bianchi S, Caini S, Renne G, Cassano E, Ambrogetti D, Cattani MG, Saguatti G, Chiaramondia M, Bellotti E, Bottiglieri R, Ancona A, Piubello Q, Montemezzi S, Ficarra G, Mauri C, Zito FA, Ventrella V, Baccini P, Calabrese M, Palli D, Group VS. Positive predictive value for malignancy on surgical excision of breast lesions of uncertain malignant potential (B3) diagnosed by stereotactic vacuum-assisted needle core biopsy (VANCb): a large multi-institutional study in Italy. *Breast.* 2011;20(3):264-70.

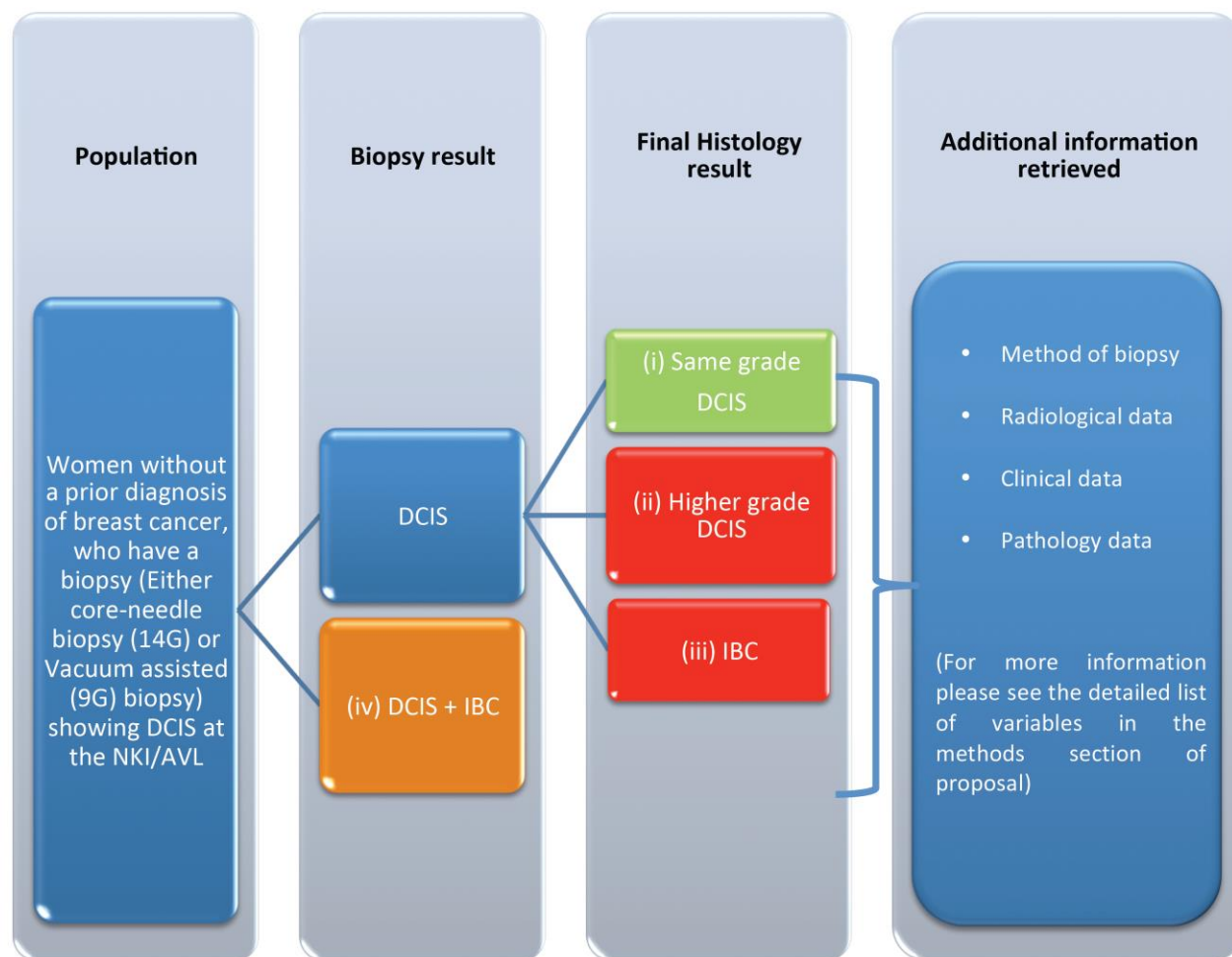


Figure 1: Flowchart of study methodology. DCIS: Ductal carcinoma-in-situ, IBC: Invasive breast cancer NKI/AVL: Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital. The information retrieved (right hand column) will be used to investigate risk factors for biopsy ‘underestimation’ of invasive cancer by comparing the groups (i-iv, shown in red) (as detailed in the ‘method of analysis’ section of the proposal) and to compare the accuracy of vacuum assisted biopsy (9G) with core needle biopsy (14G).

Appendix C: Pathology data collection form

PA score form		version 141124
1. DCIS scoring + case selection		
T-number:	Click to reset	
No. of representative slides available:		
DCIS ☐ Yes ☐ No		
DCIS morphology Present growth patterns: ☐ Clinging/ FEA ☐ (Micro)papillary ☐ Cribriform ☐ Solid Dominant growth pattern: Grade: Necrosis: Microcalcification: Environment Periductal fibrosis: Lymphocytic infiltrate: Normal tissue Normal ductal tissue slide selection:	Other lesions ☐ Non-proliferative, without atypia: (RR 1.27) ☐ Cyst ☐ Columnar cell change ☐ Apocrine metaplasia ☐ Epithelium related calcifications ☐ Limited ductal hyperplasia ☐ Proliferative, without atypia: (RR 1.80) ☐ Moderate/florid hyperplasia ☐ Intraduct./ sclerosing papillomas ☐ Sclerosing adenosis ☐ Radial scars ☐ Fibroadenomas ☐ Proliferative, with atypia: (RR 4.24) ☐ Atypical papillomas ☐ ALH/LCIS Potential exclusion: ☐ No (in case of RR ≤1.80 or atypical papillomas) ☐ Yes (in case of ALH/ LCIS) Slide selection Selected slides: (max. 5) Representativity selected slides: Remarks:	
Click to submit form		

Appendix D: Clinical data collection form

Version 9th April 2015

Study to investigate biopsy accuracy in DCIS Clinical data collection form

Name of person completing form:

Code: initials

Date form completed:

Code: (dd/mm/yyyy)

Patient Hospital number:

Code: (numerical text)

Date of Birth:

Code: (dd/mm/yyyy)

Age at diagnosis:

Code: (years)

Sex:

Code: 1= male, 2= female

Hormone replacement therapy:

Code: 0= never used, 1= current user, 2= previous user, 3= does not use, 4= history of former use not stated, 888= male/NA, 998= unknown. See code-book for definition.

Menopausal status:

Code: 1= premenopausal, 2= peri-menopausal, 3= postmenopausal, 888= male/not applicable 998= unknown. See code-book for definition.

Presentation type:

Code: 1= Symptomatic, 2= Screen-detected lesion, 998= unknown

Presence of palpable lump

Code: 1= yes, 2= no, 998= unknown

Clinical size of lump:

Code: (centimetres. 888=not palpable/not applicable)

Height:

Width:

Location of tumour:

Code: As shown below:

Laterality:

Code: 1= Left, 2= Right

Quadrant:

Code: 1= Upper inner quadrant, 2= upper outer quadrant, 3= lower inner quadrant, 4= lower outer quadrant, 5= retroareolar

Symptoms at presentation:

Code: Check-box (multiple options can be selected)

1= Palpable lump, 2= Nipple discharge, 3= Mastalgia, 4= Skin-tethering, 5= Other symptoms, 6= No symptoms, 998= Unknown. See code-book for definition.

Date of presentation:

Code: (dd/mm/yyyy)

Date of core biopsy:

Code: (dd/mm/yyyy)

Date of surgery:

Code: (dd/mm/yyyy)

See code-book for definition of 'date of presentation'

Type of breast surgery:

Code: 1= Breast conserving surgery (BCS) alone, 2= BCS with sentinel lymph node biopsy (SLNB), 3= BCS with axillary clearance (AC), 4= Mastectomy, 5= Mastectomy with SLNB, 6= Mastectomy with AC, 7= Other, 998= unknown

Notes:

For further help in completing this form, please see the relevant coding guidance form, or contact gurdeep.mannu@ctu.ox.ac.uk

Appendix E: A summary of the respective inclusion and exclusion criteria for the two ongoing trials randomising women with DCIS to surgical excision or active surveillance.

	LORIS trial	LORD trial
Inclusion Criteria:	1. Female, aged 46 years or above 2. Screen-detected or incidental microcalcification (unilateral or bilateral) 3. Histologically confirmed diagnosis of non high grade DCIS confirmed by local pathologist on either small volume core biopsy or VACB (in accordance with the current NHSBSP Guidelines for Pathology Reporting in Breast Cancer Screening) 4. DCIS diagnosed =90 days before registration 5. Able to give informed consent and comply with the trial schedule and completion of Patient Reported Outcome 6. Patient fit to undergo surgery 7. Written informed consent obtained questionnaires	1. Women of age > or = 45 years. 2. Any menopausal status. Calcifications only lesions, detected by population-based or opportunistic screening mammography. 3. Representative vacuum-assisted core biopsies with pure low-grade DCIS 4. Any size DCIS. 5. Marker placement at biopsy site(s) in the breast. 6. Good correlation between pathological and radiological findings i.e. both findings confirm low-grade DCIS and no suspicion of intermediate or high-grade DCIS or invasive breast cancer 7. Prior surgery of the ipsilateral breast because of a benign lesion allowed. 8. ASA score 1 or 2 9. Before patient registration/randomization, written informed consent must be given according to ICH/GCP, and national/local regulations
Exclusion Criteria:	1. Previous diagnosis of invasive cancer or ipsilateral DCIS (previous surgically treated contralateral DCIS is permitted) 2. A mass lesion clinically on mammogram or on ultrasound scan (if performed) at the site of the microcalcification before biopsy 3. Any serious and/or unstable pre-existing medical, psychiatric, or other condition that would prevent compliance with the trial or consent process 4. Recent onset ipsilateral bloodstained nipple discharge, unless cytology and/or Ultrasound Scan (USS) confirmed concomitant duct ectasia 5. High risk group for developing breast cancer (as defined in current NICE guidelines for familial breast cancer, or due to prior exposure to mantle field radiotherapy)	1. Presence of mass, increased density around calcifications, architectural distortion, stellate lesion on mammography. 2. Presence of bilateral DCIS. 3. Presence of Paget's disease, invasive breast cancer, or lobular carcinoma in situ on cytology or histology. 4. Presence of symptomatic DCIS e.g. DCIS detected by palpation or nipple discharge. 5. Presence of synchronous invasive carcinoma in the contralateral breast. 6. Prior history of invasive breast cancer or DCIS. 7. Prior history of other cancer except carcinoma in situ of the cervix or basocellular carcinoma of the skin. 8. Presence of serious disease that precludes definitive surgical treatment (e.g. cardiovascular/pulmonary/renal disease). 9. Individual from a family with a known BRCA1/2 mutation 10. Presence of pregnancy or breast-feeding. 11. Presence of any psychological, familial, sociological or geographical condition potentially hampering compliance with the study protocol and follow-up schedule; those conditions should be discussed with the patient before registration in the trial

Appendix F: Suggestions to the criteria used to identify low risk DCIS cases shown in supplementary table 3 which may help to minimise the risk of DCIS grade underestimation amongst women appropriate for non-operative management.

Recommendations to minimise grade discrepancy for non-operative management of low-risk DCIS
• Exclusion of women with the presence of necrosis on biopsy from consideration for non-operative management • Exclusion of lesions which have the presence of a larger tumour area (>20mm) on pre-operative imaging from consideration for non-operative management

Appendix G: NHS Breast Cancer Screening Programme (BCSP) Research Committee approval letter for Chapter 5.



Screening Programmes

C/o PHE Screening Programmes
Fulwood House
Old Fulwood Road
Sheffield
S10 3TH

Tel: 0114 2013040

rachel.crowther9@nhs.net

Gurdeep Mannu
CTSU, Nuffield Department of Population Health,
Richard Doll Building,
Old Road Campus,
Roosevelt Drive, Oxford,
OX3 7LF.

Gurdeep.Mannu@ndph.ox.ac.uk

19th July 2016

Dear Gurdeep Mannu,

The Breast Cancer Screening Programme (BSP) have discussed your project plans: *Population based study of outcomes and long-term trends in women diagnosed with non-invasive breast cancer (ODR1516_225)*

The Committee gave their support to the project.

As a condition of support, the BSP Research Committee requires you to keep them informed of developments with the project, including any changes of status, any significant adverse events, when completed, and when written up.

Any applications requiring patient identifiable data and/or potentially identifiable data from the Screening Programme will also require PHE ODR (Office of Data Release) approval. The ODR was established in January 2014 as a cross-agency service to manage the release of explicitly identifiable or potentially identifiable data from PHE. The ODR provides systematic, consistent and coherent approach to the review of data release requests for PHE datasets, to ensure that there is a legal and legitimate purpose for processing personal confidential data and that any risk of disclosure is minimised. Requests for data can be made by emailing ODR@PHE.gov.uk

The Screening Programme requires you to notify them promptly of any incidents that would be recorded on the Health Research Authority (HRA) Breaches Register. Undertaking research within the Screening Programme following receipt of this letter of support assumes your agreement to fulfil this obligation. HRA has the potential to share information with the BSP regarding any breaches of ethics related to projects involving the BSP.

The Committee wishes you well with your project.

Yours sincerely

A handwritten signature in blue ink, appearing to read "R. Crowther".

Rachel Crowther MSc BMedSci
On behalf of the NHS BCSP Research Committee.

Appendix H: Research Ethics Committee approval letter for Chapter 5.



Health Research Authority **Yorkshire & The Humber - Leeds East Research Ethics Committee**

Room 001
Jarrow Business Centre
Rolling Mill Road
Jarrow
Tyne and Wear
NE32 3DT

Telephone: 0207 1048081

18 May 2016

Dr Gurdeep S. Mannu
Clinical Research Fellow in Cancer Surgery
University of Oxford
CTSU, Richard Doll Building
Old Road Campus, Roosevelt Drive
OXFORD
OX3 7LF

Dear Dr Mannu

Study title:	Population based study of outcomes and long-term trends in women diagnosed with non-invasive breast cancer
REC reference:	16/YH/0209
Protocol number:	Version 1.
IRAS project ID:	192243

The Proportionate Review Sub-committee of the Yorkshire & The Humber - Leeds East Research Ethics Committee reviewed the above application on 13 May 2016.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager Katy Cassidy, nrescommittee.yorkandhumber-leedseast@nhs.net. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

A Research Ethics Committee established by the Health Research Authority

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA Approval (England)/ NHS permission for research is available in the Integrated Research Application System, www.hra.nhs.uk or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see “Conditions of the favourable opinion”).

Summary of discussion at the meeting

The Sub-Committee raised no concerns regarding the study.

Approved documents

The documents reviewed and approved were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
IRAS Checklist XML [Checklist_03052016]		03 May 2016
REC Application Form [REC_Form_29042016]		29 April 2016
Research protocol or project proposal [Protocol]	1.0	14 April 2016
Summary CV for Chief Investigator (CI) [CV Gurdeep Mannu]	1.0	29 April 2016
Summary CV for supervisor (student research) [Prof Sarah Darby Supervisor CV]	1.0	25 March 2016

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document “After ethical review – guidance for researchers” gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:

<http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/>

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at <http://www.hra.nhs.uk/hra-training/>

With the Committee's best wishes for the success of this project.

16/YH/0209

Please quote this number on all correspondence

Yours sincerely

pp



Dr Deborah Jane Fox
Chair

Email: nrescommittee.yorkandhumber-leedseast@nhs.net

Enclosures: *List of names and professions of members who took part in the review*
"After ethical review – guidance for researchers" [SL-AR2]

Copy to: *Ms Heather House, University of Oxford*

Yorkshire & The Humber - Leeds East Research Ethics Committee

Attendance at PRS Sub-Committee of the REC meeting on 13 May 2016

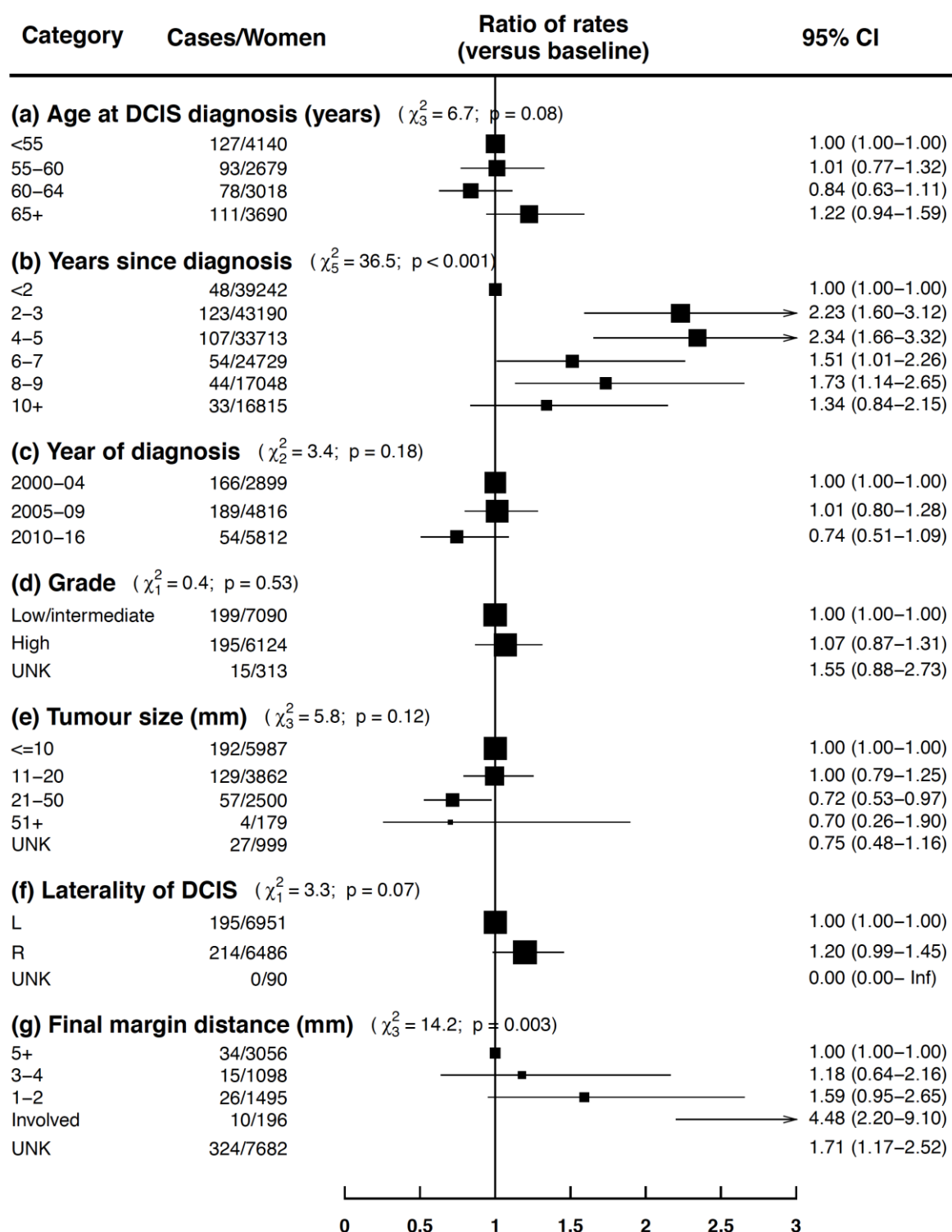
Committee Members:

<i>Name</i>	<i>Profession</i>	<i>Present</i>	<i>Notes</i>
Professor Kenneth Brodlie	Retired Professor of Visualization	Yes	
Dr Deborah Jane Fox	Senior Lecturer in Nursing	Yes	
Mr Tom Wilson	Consultant ENT Surgeon	Yes	

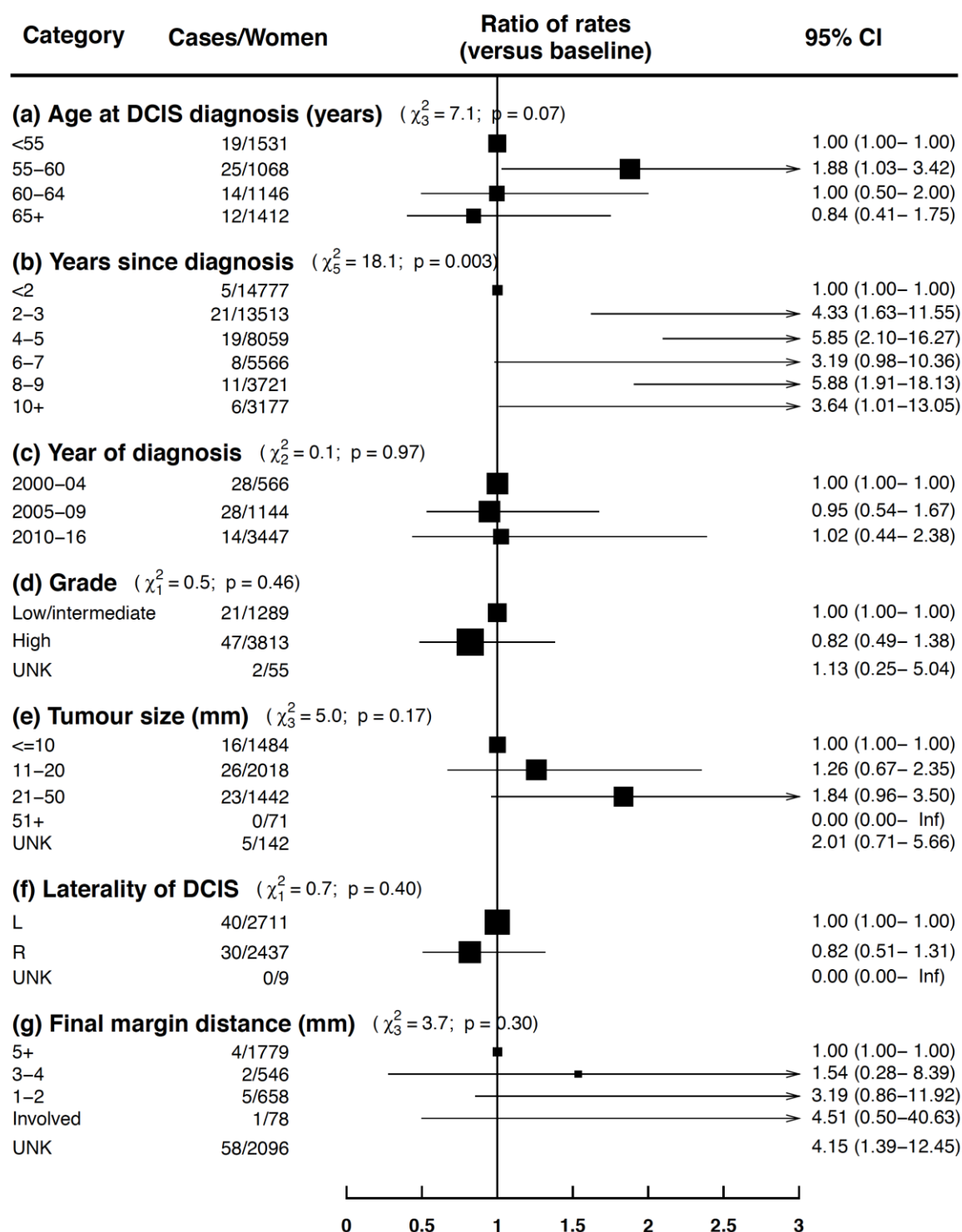
Also in attendance:

<i>Name</i>	<i>Position (or reason for attending)</i>
Ms Katy Cassidy	REC Manager

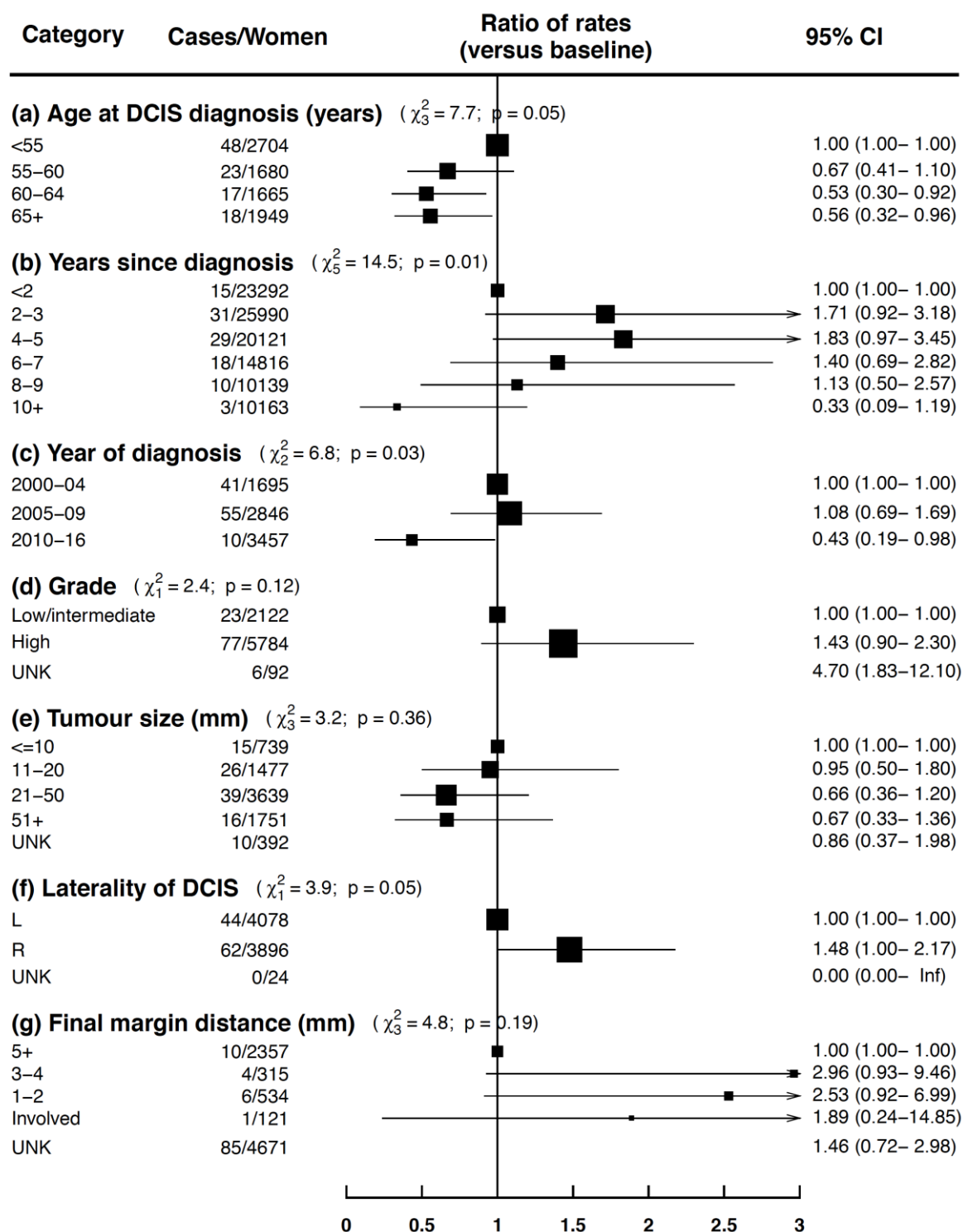
Appendix I: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 13,528 women who were treated with breast conserving surgery alone and were not recorded as receiving hormone therapy. Each variable is adjusted by all other variables in the categories shown. Baseline is the first row for each variable (a) – (g). Abbreviations: CI = Confidence Interval, UNK = Unknown, mm = millimeters.



Appendix J: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 5157 women who were treated with breast conserving surgery and radiotherapy and not recorded as receiving hormone therapy. Each variable is adjusted by all other variables in the categories shown. Baseline is the first row for each variable (a) – (g). Abbreviations: CI = Confidence Interval, UNK = Unknown, mm = millimeters.



Appendix K: Ipsilateral invasive breast cancer after screen-detected DCIS by patient, tumour and treatment factors, in 7998 women who were treated with mastectomy and not recorded as receiving hormone therapy. Each variable is adjusted by all other variables in the categories shown. Baseline is the first row for each variable (a) – (g). Abbreviations: CI = Confidence Interval, UNK = Unknown, mm = millimeters.



Appendix L: Description of algorithm for Chapter 6.

Draft of 31 January 2017

WORKING DOCUMENT

Classifications used in the Process of Identifying Recurrences in Invasive Breast Cancer

Any algorithm that is used to identify recurrences and other outcomes in invasive breast cancer will need rules that can be applied to patients with different tumour characteristics and treatment pathways. Therefore we are developing a 'classification system' based on known recurrence indicators, timing of treatment events and known deaths.

1. Method

A cohort of patients was extracted from the PHE's Cancer Analysis System (CAS). All women registered with invasive breast cancer (ICD10 C50, ICD9 174) during the time-period under consideration (1 Jan 1997-31 Dec 2011) and living in the West Midlands area at the time of their registration were included in the initial dataset, unless one of the exclusion criteria applied. The exclusion criteria are listed in Appendix A.

Hospital Episode Statistics (HES), Cancer Waiting Times (CWT), Radiotherapy Dataset (RTDS), Digital Imaging Dataset (DID) and the West Midlands Cancer Registry dataset (GRACE) were interrogated for data relating to these patients/tumours, and all events and information relating to possible recurrences (local, regional, distant, or of unknown location) were extracted for the time-period under consideration. A greater volume of recurrence data had been collated by the West Midlands Cancer Registry, acting as lead breast cancer registry, than had been recorded in other regions. Therefore, GRACE was included as a source of enriched event and recurrence data.

Events recorded in multiple datasets were combined where possible using a 'working rule' whereby events of the same type and occurring within 7 days of each other from different datasets would be counted once and allocated an event source by taking datasets in order of national accessibility and expected completeness: CAS, HES, CWT, RTDS, DID, GRACE i.e. GRACE data would only be used if no matching data was found elsewhere. Tables were created within CAS to hold patient/tumour data for the cohort, event data and recurrence data. Patients and tumours were given anonymised identifiers.

Initial analysis was performed twice, (i) on all datasets and (ii) excluding both GRACE data and also excluding CAS recurrence data transferred from GRACE to CAS in 2012. Classifications as defined in in section 2 (below) were allocated to each patient/tumour.

2. Definitions of Classifications (updated from the Project Description)

A. *Women who should have been excluded from the initial data set (see Appendix A), but who were not initially identified as such.*

If any of these exist, they are likely to be discovered during analysis of extracted data at later stages of the project.

B. *Women not in A and who already had metastatic breast cancer on date of cancer registration.*

B1. Women who were registered as having stage IV breast cancer, or who were described as having distant metastases at time of registration or within 3 months of registration.

-For women registered as stage VI (unknown stage), ICD-10 codes C77-C79 from HES data will be used to determine their cancer stage in accordance with the criteria listed in appendix B. All women who are not stage IV at time of, or within 3 months of registration will be classified into C1-11 below.

-Women receiving palliative care within 3 months of registration are B1

B2. *Women not in B1 and who were reported to have breast cancer only from their death certificate (ie DCO) and other women whose date of cancer registration was the same as their date of death.*

Patients/tumours recorded as DCO='Y' or where diagnosis date = date of death

C. *Women not in A or B and who had further cancer-related (or possibly cancer-related) events or who died during the time-period under consideration (ie 1 Jan 1997-31 Dec 2011). Women may have more than one event, but the main focus at present is on the earliest such event. In cases where one or more of the classifications titled 'C1'- 'C11' occur on the same date or very close together, all events will be recorded but the order of precedence is described in Appendix C.*

C1. Death from breast cancer

Breast cancer (code C50) entered in any cause of death field 1a, 1b, 1c or 'underlying'. Breast cancer (C50), 'Breast unspecified' or 'malignant neoplasm' that only appears in cause of death field 2 is not included.

C1a. Death from breast cancer and has had breast surgery

Breast surgery is defined as:

- '01' in CWT

- the following OPCS4 codes:

Breast conserving surgery: ('B281', 'B282', 'B283', 'B285', 'B286', 'B287', 'B288', 'B289', 'B34', 'B341', 'B342', 'B343', 'B344', 'B345', 'B348', 'B349', 'B35')

Mastectomy: ('B27', 'B271', 'B272', 'B27', 'B274', 'B275', 'B276', 'B278', 'B279')

SLNB: ('O142', 'T87', 'T873', 'T878', 'T879', 'T91', 'T911', 'T918', 'T919', 'T928', 'Z613')

Sampling: ('T86', 'T862', 'T869')

Axillary Dissection: ('Z613', 'T852', 'T859', 'T862')

C1b. *Death from breast cancer but has never had breast surgery (i.e. no evidence of surgery from CWT and none of the OPCS4 codes listed in C1a)*

[C2. Not in use]

C3. Distant metastasis

C3a (i): Distant metastases of liver/lung/bone/brain as indicated by one or more of the following occurring at least 3 months after date of original cancer diagnosis but before or on 31st Dec 2011 (in these 4 anatomical sites):

West Midlands legacy data (GRACE): recurrence code = 'M', metastatic disease

- CAS (NCRS): typeofrecurrence = '05', metastatic disease

- CWT: mets_site is not null or CTE_type (cancer treatment event type) =

'05' Treatment for a distant recurrence of cancer (metastatic disease)

'07' First treatment for metastatic disease following an unknown primary

'08' Second or subsequent treatment for metastatic disease following an unknown primary

- HES: All women regarded as stage IV from the HES diagnosis codes (ICD-10 codes C77-C79) using criteria listed in appendix B.

C03a(ii) (Distant metastases in other locations after 3m):

There is a registration of a new secondary malignancy in an anatomical site not covered by C03a that occurs at least 93 days after the date of the original diagnosis

C3a (iii): Metastatic disease receiving palliative therapy at least 3 months after date of original cancer diagnosis but before or on 31st Dec 2011:

- Palliative care is only given to cancer patients with distant metastases, so anyone receiving it after 3 months from diagnosis can be classified as C3a.

- Any of the following three event codes are recorded: 'HP' – hospice stay (GRACE), '07' – specialist palliative care (CWT), '09' – non-specialist palliative care (CWT). (Although these last two codes exist in CAS – none have been allocated from CAS to this cohort within the timeframe).

C3b. women who do not have any other invasive cancer recorded (ICD-10 codes: C00-C97, including C50, but excluding C44; ICD-9 codes: 140-209, including 174, but excluding 173) but have one or more of the following OPCS4 codes in CAS/HES after 3 months of diagnosis ('T122', 'T121', 'T013', 'A025', 'W085', 'W097', 'A021', 'A028', 'B343', 'W091', 'E543', 'A023', 'E544', 'V432', 'W098', 'A029', 'W088', 'A024', 'S049', 'W093', 'A022', 'J023', 'W089', 'E545', 'T018', 'E552', 'V431', 'J024', 'J031', 'J124', 'T019', 'V257', 'J029', 'Q241', 'V051', 'W096', 'E558', 'A026', 'A178', 'B401', 'E541', 'E542', 'E548', 'J027', 'J032', 'T078', 'T079', 'T101', 'W099', 'T102', 'T103', 'T131', 'T095', 'T094', 'L912', 'L943', 'Z941', 'ABLATION', 'T124', 'T882', 'L915') irrespective of the laterality of the procedure.

- Women in C3b may either have (1) one or more of the cancers in C08a that was not registered, (2) an occult breast cancer recurrence or (3) no second cancer. The classification algorithm will be re-

run on these women after excluding the 'first event' to see which category they subsequently fall into.

C4a. Recurrence of cancer in the ipsilateral regional lymph nodes

Indicated by one or more of the following occurring at least 3 months after date of original cancer diagnosis but before or on 31st Dec 2011:

- HES/CAS: {OPCS4 treatment type = 'MT' (Surgery for nodes and metastases: i.e. OPCS4 codes 'T873', 'T878', 'T879', 'T911', 'T918', 'T919', 'T928', 'Z613', 'T862', 'T869', 'T85', 'T852', 'T859') OR OPCS4 = 'O142'} AND the laterality relating to the event is the same as recorded for the tumour
- CAS/CWT: Treatment event type '04': Treatment for a regional recurrence of cancer
- GRACE: recurrence event of 'B' Loco-nodal recurrence OR 'R' Nodal recurrence

C4b. Recurrence of cancer in the regional lymph nodes with no laterality recorded (assumed to be ipsilateral).

Indicated by one or more of the events listed in category 'C4 (i)', occurring at least 3 months after date of original cancer diagnosis but before or on 31st Dec 2011, but with no laterality recorded in any routinely collected dataset. In this setting it is assumed that the recurrence of cancer is in the ipsilateral regional lymph nodes, in the absence of evidence to the contrary.

C5. Recurrence of cancer in the ipsilateral breast

Indicated by one or more of the following occurring at least 3 months after date of original cancer diagnosis but before or on 31st Dec 2011:

- CAS: Another, later breast tumour with same laterality
- CAS/CWT: event type of '03', i.e. Treatment for a local recurrence of a primary cancer.
- GRACE: recurrence type of 'L' Local recurrence or progression of primary disease
- HES/CAS: (OPCS4 codes = 'B28', 'B281', 'B282', 'B283', 'B285', 'B286', 'B287', 'B288', 'B289', 'B323', 'B34', 'B341', 'B342', 'B343', 'B345', 'B348', 'B349', 'B35', 'B27', 'B271', 'B272', 'B273', 'B274', 'B275', 'B276', 'B278', 'B279') AND the laterality relating to the event is the same as recorded for the original tumour

C6. Breast cancer recurrence of unknown location

Indicated by one or more of the following before or on 31st Dec 2011:

- CWT: patient status in ('15', '16', '17', '18', '19', '20') all indicating recurrent cancer suspected or confirmed.
- CAS/CWT: event type in '06' Treatment for multiple recurrence of cancer (local/regional/distant)
 - '09' 2nd or subsequent treatment for relapse of primary cancer
 - '10' 2nd or subsequent treatment for progression of primary cancer
 - '99'/'U' Unknown recurrence treatment
- CAS: recurrence flag = 'Y'
- CAS/HES: one or more of the following OPCS4 codes in CAS/HES:
 - Initiation of chemotherapy ≥ 1 year after diagnosis ('X352', '02')
 - Initiation of radiotherapy ≥ 1 year after diagnosis (RTX)

Initiation of hormone therapy at any time after diagnosis following an absence of hormone therapy of ≥ 1 year ('03')

Two or more of the following types of scan, within a 1 year time period, at least one year following diagnosis of breast cancer: 'CT scan with contrast', 'CT scan with unspecified contrast', 'MRI scan with contrast', 'MRI scan with unspecified contrast', 'PET scan', 'Radio-isotope bone scan'.

-DID: Two or more of the following types of scan, within a 1 year time period, at least one year following diagnosis of breast cancer:'

Computed tomography of abdomen (procedure) (169070004) [CABDO], Computed tomography and biopsy of abdomen (procedure) (419940006) [CABDOB], Computed tomography of abdomen and pelvis (procedure) (419394006) [CABPE], Computed tomography of chest and abdomen (procedure) (418891003) [CCABD], Computed tomography of chest, abdomen and pelvis (procedure) (418023006) [CCHAP], Computed tomography of chest (procedure) (169069000) [CCHES], Computed tomography of cervical spine (procedure) (241578008) [CCSPN], Computed tomography of liver (procedure) (241549007) [CLIVE], Computed tomography and biopsy of liver (procedure) (418749009) [CLIVEB], Computed tomography of lumbar spine (procedure) (241580002) [CLSPN], Computed tomography of mediastinum (procedure) (241543008) [CMEDM], Computed tomography of pituitary fossa (procedure) (241519005) [CPITF], Computed tomography of posterior fossa (procedure) (241518002) [CPOSF], Computed tomography of sacral spine (procedure) (241581003) [CSACM], Computed tomography of thoracic spine (procedure) (241579000) [CTSPN], Virtual computed tomography bronchoscopy (procedure) (418419008) [CVBRY], Magnetic resonance imaging of abdomen (procedure) (241621009) [MABDO], Ablation of lesion of liver using computed tomography guidance (procedure) (432226009) [CLMAAA], Radiofrequency ablation of lesion of liver using computed tomography guidance (procedure) (431475009) [CLMAFA], Laser ablation of lesion of liver using computed tomography guidance (procedure) (433866007) [CLMASA], Computed tomography of neck and thorax (procedure) (430448007) [CNECH], Computed tomography of neck and thorax with contrast (procedure) (429927002) [CNECHC], Computed tomography of neck with contrast (procedure) (431326009) [CNECKC], Computed tomography of sacral spine with contrast (procedure) (430452007) [CSACMC], Computed tomography of sternum (procedure) (429874005) [CSTRM], Magnetic resonance imaging of chest with contrast (procedure) (432815006) [MCHESC], Magnetic resonance imaging of liver with contrast (procedure) (431839003) [MLIVEC].

Or one or more of the following (SNOMED) or [NICIP] codes in DID after 1 year of diagnosis:

Computed tomography of breast for radiotherapy planning (procedure) (429866009) [CRTMAR], Computed tomography of head for radiotherapy planning (procedure) (429867000) [CRTSKR], Computed tomography of spine for radiotherapy planning (procedure) (430438005) [CRTSPR], Magnetic resonance imaging of chest for radiotherapy planning (procedure) (431766006) [MRTCHR], Magnetic resonance imaging of breast for radiotherapy planning (procedure) (430243000) [MRTMAR], Radiofrequency ablation using ultrasound guidance (procedure) (433058002) [URAFGA], Wire guided localisation of lesion using ultrasonography guidance (procedure) (915381000000106) [UWLLE], Wire guided localisation of breast lesion using magnetic resonance imaging guidance (procedure) (911831000000104) [MWLBL], Wire

guided localisation of breast lesion using magnetic resonance imaging guidance (procedure) (911831000000104) [MWLBR], Wire guided localisation of breast lesion using magnetic resonance imaging guidance (procedure) (911831000000104) [MWLBB], Biopsy of femur using fluoroscopic guidance (procedure) (431455000) [FTHIBB], Biopsy of lung using computed tomography guidance (procedure) (429932001) [CLUNGB], Biopsy of brain using computed tomography guidance (procedure) (432666003) [CSKUHB].

C7. Cancer in the contralateral breast or the contralateral lymph nodes

This is also referred to as the woman having had bilateral breast cancer. Indicated by one or more of the following before or on 31st Dec 2011:

- HES: {OPCS4 treatment type = 'MT' (Surgery for nodes and metastases, covering OPCS4 codes 'T873', 'T878', 'T879', 'T911', 'T918', 'T919', 'T928', 'Z613', 'T862', 'T869', 'T85', 'T852', 'T859') OR OPCS4 = 'O142'} AND the laterality relating to the event is not the same as recorded for the tumour.
- HES/CAS: {(OPCS4 treatment (surgery for breast), covering OPCS4 codes 'B28', 'B281', 'B282', 'B283', 'B285', 'B286', 'B287', 'B288', 'B289', 'B323', 'B34', 'B341', 'B342', 'B343', 'B345', 'B348', 'B349', 'B35') OR OPCS4 = 'B27', 'B271', 'B272', 'B273', 'B274', 'B275', 'B276', 'B278', 'B279'} AND the laterality relating to the event is not the same as recorded for the tumour
- GRACE: Recurrence type = 'C' Contralateral disease
- CAS: Another, later breast tumour with different laterality

C8. Occurrence of any other invasive cancer (ICD-10 codes: C00-C97, including C50, but excluding C44; ICD-9 codes: 140-209, including 174, but excluding 173)

CAS: another, later, tumour of any type other than non-melanoma skin more before or on 31st Dec 2011:

where ICD10 codes are between C00-C97 exc. C44; and ICD9 codes are between 140 and 209, exc.173, and cross referenced with behaviour being in (3,6,9) to indicate malignancy/invasive tumour.

This group is subdivided into C8a, C8b and C8c as follows:

C8a. women falling into category C8 who have cancer (s) of liver (ICD 9: 155, ICD 10: D13.4), lung (ICD 9: 162, ICD 10: C34), bone (ICD 9: 170, ICD 10: C40-41) or brain (ICD 9: 191-192, ICD 10: C70-C72) and the 'basis of diagnosis' are one or more of 'specific tumour markers', 'cytology', 'histology of metastasis' and 'histology of primary'.

C8b. women falling into category C8 who have cancer (s) of liver (ICD 9: 155, ICD 10: D13.4), lung (ICD 9: 162, ICD 10: C34), bone (ICD 9: 170, ICD 10: C40-41) or brain (ICD 9: 191-192, ICD 10: C70-C72) and the 'basis of diagnosis' as one or more of 'death certificate', 'clinical', 'clinical investigation', 'x not known'

C8c. women falling into category C8 who have invasive cancers other than those listed in C08a irrespective of the 'basis of diagnosis'.

C9. Death from a cancer other than breast cancer, i.e. with one of the relevant codes entered in any cause of death field 1a, 1b, 1c or 'underlying'. (ICD-10 codes: C00-C97, excluding C50; ICD-9 codes: 140-209, excluding 174). Codes in cause of death field 2 are not taken into account.

C10. Deaths where no cause has been attributed.

C11. Women who died of a cause other than cancer

Women who are not in any of the above categories but who died before or on 31st Dec 2011 and for whom the cause of death is known but for whom there is no mention of cancer entered in any cause of death fields 1a, 1b, 1c or 'underlying' (ie no ICD-10 codes: C00-C97; ICD-9 codes: 140-209). Codes in cause of death field 2 are not taken into account.

D. Women not in A, B, or C and with no further cancer-related events recorded during the time-period under consideration

D1. Women for whom there was no indication that any further event related to breast cancer had occurred by the end of the time period under consideration (ie 31 December 2011), nor had they died by the end of the time period under consideration.

E. Women who cannot be classified into any of the above categories.

Document history

17 Dec 2015 Draft by JC and JB
18 Dec 2015 Revisions by SD and GM
8 Jan 2016 Revisions by SD and GM
11 Jan 2016 Revisions by SD
18 Jan 2016 Revisions by GM
19 Jan 2016 Revisions by GM and SD
20 Jan 2016 Revisions by GM
2 Feb 2016 Revisions by GM
15 May 2016 Revisions by GM
30 June 2016 Revisions by GM
4 July 2016 Revisions by GM
19 July 2016 Revisions by GM
3 October 2016 Revisions by GM
19 Dec 2016 Revisions by GM
8 Jan 2017 Revisions by GM
26 Jan 2017 Revisions by GM
31 Jan 2017 Revisions by GM

Supplementary material A. Women to be excluded from the initial dataset
(taken directly from Project Description version of 3 November 2015)

- A. Anyone not recorded as female.
- B. Anyone not living in the West Midlands area at time of first registration of breast cancer.
- C. Anyone with a diagnosis of invasive breast cancer prior to 1 Jan 1997, even if they have another breast cancer registered after 1 Jan 1997.
- D. Anyone who had an invasive cancer of a site other than invasive breast cancer prior to their initial diagnosis of invasive breast cancer, with the exception of non-melanoma skin cancer which is to be ignored.

The ICD codes for non-melanoma skin cancer are: (ICD-10 code: C44; ICD-9 code 173)

NB: Previously diagnosed cancers that are NOT invasive, including in situ breast cancers, benign breast tumours, and breast tumours of uncertain or unknown behaviour are NOT grounds for exclusion. The ICD codes for these cancers that are not classified as invasive cancers are: (ICD-10: D05, D25, D48.6; ICD-9 217, 233.0, 238.3, 239.3).

- E. Women who, at the time of their initial diagnosis of breast cancer, or within three months of it, were diagnosed with an invasive cancer in the other breast (ie had bilateral invasive breast cancer). NB Women who, within a period of three months, have several diagnoses of invasive breast cancer reported in the same breast will be included in the dataset. Their date of diagnosis will be taken to be the earliest of these reported dates and the stage of breast cancer will be taken to be the highest recorded.
- F. Women who were registered with a primary invasive cancer of another site as well as breast cancer at the same time as their initial diagnosis of invasive breast cancer, or within three months of it (with the exception of non-melanoma skin cancer, which is to be ignored, see above).

Supplementary material B. For women registered as stage VI (unknown stage), ICD-10 codes C77-C79 from HES data (within 3 months of diagnosis of breast cancer) will be used to determine cancer stage

ICD 10 code	Sub-classification	Cancer stage to be allocated
C77 Secondary and unspecified malignant neoplasm of lymph nodes Excl.: malignant neoplasm of lymph nodes, specified as primary (C81-C86, C96.-)	C77.0 Lymph nodes of head, face and neck Supraclavicular lymph nodes	II
	C77.1 Intrathoracic lymph nodes	II
	C77.2 Intra-abdominal lymph nodes	VI (Not staged)
	C77.3 Axillary and upper limb lymph nodes Pectoral lymph nodes	II
	C77.4 Inguinal and lower limb lymph nodes	VI (Not staged)
	C77.5 Intrapelvic lymph nodes	VI (Not staged)
	C77.8 Lymph nodes of multiple regions	II
	C77.9 Lymph node, unspecified	II
C78 Secondary malignant neoplasm of respiratory and digestive organs	C78.0 Secondary malignant neoplasm of lung	IV
	C78.1 Secondary malignant neoplasm of mediastinum	IV
	C78.2 Secondary malignant neoplasm of pleura Malignant pleural effusion NOS	IV
	C78.3 Secondary malignant neoplasm of other and unspecified respiratory organs	IV
	C78.4 Secondary malignant neoplasm of small intestine	VI (Not staged)
	C78.5 Secondary malignant neoplasm of large intestine and rectum	VI (Not staged)
	C78.6 Secondary malignant neoplasm of retroperitoneum and peritoneum, Malignant ascites NOS	IV
	C78.7 Secondary malignant neoplasm of liver and intrahepatic bile duct	IV
	C78.8 Secondary malignant neoplasm of other and unspecified digestive organs	IV
C79 Secondary malignant neoplasm of other and unspecified sites	C79.0 Secondary malignant neoplasm of kidney and renal pelvis	VI (Not staged)
	C79.1 Secondary malignant neoplasm of bladder and other and unspecified urinary organs	VI (Not staged)
	C79.2 Secondary malignant neoplasm of skin	IV
	C79.3 Secondary malignant neoplasm of brain and cerebral meninges	IV
	C79.4 Secondary malignant neoplasm of other and unspecified parts of nervous system	IV
	C79.5 Secondary malignant neoplasm of bone and bone marrow	IV
	C79.6 Secondary malignant neoplasm of ovary	VI (Not staged)
	C79.7 Secondary malignant neoplasm of adrenal gland	VI (Not staged)
	C79.8 Secondary malignant neoplasm of other specified sites	IV
	C79.9 Secondary malignant neoplasm, unspecified site, Carcinomatosis (secondary), Disseminated (secondary): cancer NOS, malignancy NOS, Generalized (secondary): cancer NOS, Malignancy NOS, Multiple secondary cancer NOS, Sarcomatosis (secondary) NOS	IV

Supplementary material C. Priority order and rules defining the categorisation

The priority order is applied from top to bottom of the following list at the time of each potential classifying event, in order of time from diagnosis. Each tumour is assigned one classification and inclusion in a particular classification is only possible if that tumour cannot be classified by any of the preceding classifications. (i.e. for B1 onwards, the criteria for B2 can be considered as being implicitly stated in the inverse as part of the criteria for B1 and so on throughout the priority order.)

So, for instance, if a patient (as their first potential classifying events occurring on the same date) is recorded as developing both contralateral recurrence and being diagnosed with a secondary cancer of the liver, then they will be classified to C07, not C08a. This patient's treatment pathway can also be said not to have fulfilled any of the criteria before C07 in the priority order.

B2 (DCO or death & diagnosis on the same day):

The tumour recorded as a DCO or the date of death is the same as the date of diagnosis.

B1 (Stage 4 or metastases within 3 months of diagnosis):

The stage at diagnosis for the tumour is 4 or there is an indication of metastases.

Metastases are indicated by:

- the registration of a secondary malignancy that is distant to the breast in HES
- the notification of a metastatic recurrence to NCRAS by a Trust
- the treatment for a distant recurrence of cancer in CWT

C01 (Death from breast cancer):

The tumour is not a DCO and any part of the death certificate includes an invasive breast cancer code

C02 (Palliative treatment):

An event notified to NCRAS as palliative care by a Trust

C03a(i) (Distant metastases of liver/lung/bone/brain after 3m):

There is a registration of a new secondary malignancy in the above locations that occurs at least 93 days after the date of diagnosis

C03a(ii) (Distant metastases in other locations after 3m):

There is a registration of a new secondary malignancy not covered by C03a that occurs at least 93 days after the date of the original diagnosis

C03b (Distant metastases of liver/lung/bone/brain (unconfirmed) after 3m):

There is a procedure on the above locations that occurs at least 93 days after the date of diagnosis, which may represent treatment for a distant breast cancer metastasis in the absence of registration of a new primary cancer of that location.

C07 (Contralateral breast/lymph node cancer)/bilateral breast cancer:

One of the following conditions is met

- Sentinel lymph node biopsy on the other side from the original tumour
- Any surgery to the other breast
- Diagnosis of a breast tumour in the other breast

C04a (Recurrence in ipsilateral regional lymph nodes):

One of the following conditions is met

- Sentinel lymph node biopsy on the same side as the original tumour at least 93 days after the date of diagnosis
- CWT records treatment for a regional recurrence of cancer

C4b (Recurrence of cancer in the regional lymph nodes (assumed to be ipsilateral)):

- One of the conditions listed in C04 (i) is met but no information on laterality is recorded.
- In this setting it is assumed that the recurrence of cancer is in the ipsilateral regional lymph nodes (in the absence of evidence to the contrary).

C05 (Recurrence in ipsilateral breast):

One of the following conditions is met, occurring at least 93 days after the date of the original diagnosis

- A diagnosis of another tumour in the same breast
- CWT records treatment for a local recurrence of a primary cancer
- There is a mastectomy or further BCS

C06 (Recurrence – unknown location):

One of the following conditions is met

- CWT records the patient status as having suspected or diagnosed recurrent cancer
- CWT records treatment for one of
 - multiple recurrence of cancer (local and/or regional and/or distant)
 - relapse of primary cancer (second or subsequent)
 - progression of primary cancer (second or subsequent)
- Notification to NCRAS by a Trust of recurrence

C08a (New microscopically confirmed primary cancer of liver/lung/bone/brain after 3 months):

There is a registration of a new malignancy in the above locations that occurs at least 93 days after the date of diagnosis with microscopic confirmation of a primary cancer

C08b (New microscopically unconfirmed primary cancer of liver/lung/bone/brain after 3 months):

There is a registration of a new malignancy in the above locations that occurs at least 93 days after the date of diagnosis without microscopic confirmation of a primary cancer

C08c (New primary cancer in other locations after 3m):

There is a registration of a new malignancy not covered by C08a that occurs at least 93 days after the date of the original diagnosis

Note that C5, C7 prevent another breast tumour diagnoses from being categorised here although C50s are not explicitly excluded.

C09 (Death from another cancer):

The tumour is not a DCO and any part of the death certificate includes an invasive cancer code but no invasive breast cancer codes

C10 (Death with no cause recorded):

The tumour is not a DCO and there is no recorded cause of death information

C11 (Death from cause other than cancer):

The tumour is not a DCO and there is no record of cancer in the cause of death information

D01 (Embarkation):

Where NCRAS have been notified by ONS that the patient has emigrated

D02 (No further events within the time period):

No events (diagnostic or treatment) were found between the date of diagnosis and the end of the period of study.

E (Unclassified):

Tumours and patients that do not meet any of the above criteria.

Appendix M: Contributions to and Planned Extensions of Studies in the Thesis

The introductory and discussion chapters (1, 2 and 7) were the work of GM alone. The four studies presented in chapters 3 -6 were all highly collaborative with contributions described below. SD and FH provided appropriate supervision throughout.

Chapter 3: Differences in Risk Factors for Ductal Carcinoma in Situ and Invasive Breast Cancer: Prospective Cohort Study

-GM: Application for funding to the Royal College of Surgeons of Edinburgh for UK Biobank application costs, chief investigator on data requests to the UK Biobank scientific committee, design of the study, interpretation of data and composition of chapter.

-PM and ZW: Assistance with (and programming for) statistical analyses

Planned extensions:

-Analyses of risk factors for cases of DCIS and IBC diagnosed before entry into UK Biobank

-Comparison of genetic risk factors for DCIS and IBC

Chapter 4: The accuracy of preoperative biopsy in Ductal Carcinoma in Situ

-GM: Design of the study, analysis and programming, interpretation of data and composition of chapter.

-ZW: Assistance with (and programming for) statistical analyses

-MC and IJ: data extraction from clinical records

-FL: statistical and methodological input

-MS: database design and methodological input

-EG: data extraction from pathology slides

-EL: pathology input

-JW: methodological input and data extraction from pathology slides

-GW/JT: data extraction from mammograms

Planned extensions:

-Examination of the association between genetic biomarkers and preoperative biopsy underestimation rate for IBC or DCIS grade

-Detailed examination of characteristics of IBC in upstaged biopsies.

Chapter 5: Population-based Study of Invasive Breast Cancer in Women Diagnosed with Ductal Carcinoma in Situ

-GM: Design of the study, principal investigator for Research Ethics Committee (REC) application and Office of Data Release (ODR) application, interpretation of data and composition of chapter.

-ZW: Assistance with (and programming for) statistical analyses

-SC and OK: Extraction and transfer of data from the NHS Breast Screening Programme (NHSBSP) audit.

-JAC and JB: Extraction and transfer of data from the Cancer Analysis System (CAS).

Planned extensions:

-Examination of the association between DCIS morphology and geographical region of DCIS diagnosis and IBC rate.

-Comparison of IBC rate and breast cancer mortality in women diagnosed with screen-detected DCIS with those diagnosed with non-screen-detected DCIS

Chapter 6: The Use of Routine Data Sources to Identify Recurrence in Women Registered with Invasive Breast Cancer in England

-GM: Design of the study, data cleaning, development of the algorithm, principal investigator for REC application and ODR application, interpretation of data and composition of chapter.

-DK: Assistance with (and programming for) statistical analyses

-JAC and JB: Assistance with (and programming for) algorithm and statistical analyses

Planned extensions:

-External validation of findings using data from a UK randomised controlled trial and/or from a regional cohort study.

-Extension of the algorithm to the identify recurrences for women diagnosed in the whole of England.

Contributors' initials (alphabetical order):

DK: David Kennedy, Statistician, Nuffield Department of Population Health, University of Oxford

EG: Emilie Groen, Pathologist, Netherlands Cancer Institute Antoni van Leeuwenhoek

EL: Esther Lips, Staff scientist, Netherlands Cancer Institute Antoni van Leeuwenhoek

FL: Flora van Leeuwen, Professor in Epidemiology, Netherlands Cancer Institute Antoni van Leeuwenhoek

GM: Gurdeep S. Mannu (thesis author), Nuffield Department of Population Health, University of Oxford

IJ: Ires Joore, Researcher, Netherlands Cancer Institute Antoni van Leeuwenhoek

JB: John Broggio, Principal Analyst, National Cancer Registration and Analysis Service, Public Health England

JC: Jackie Charman, Senior Cancer Analyst, National Cancer Registration and Analysis Service, Public Health England

JT = Jelle Teerstra, Radiologist, Netherlands Cancer Institute Antoni van Leeuwenhoek

JW: Jelle Wesseling, Professor in Molecular Pathology, Netherlands Cancer Institute Antoni van Leeuwenhoek

MC: Monica Chung, Researcher, Netherlands Cancer Institute Antoni van Leeuwenhoek

MS: Michael Schaapveld, Associate staff scientist, Netherlands Cancer Institute Antoni van Leeuwenhoek

OK: Olive Kearins, Head of QA, Midlands and East, National Lead Breast Screening QA, Public Health England

SC: Shan Cheung, National Audit Project QA Facilitator, Screening QA Service (Midlands and East), Public Health England

SD: Sarah Darby, Professor of Medical Statistics, Nuffield Department of Population Health, University of Oxford

ZW: Zhe Wang, Statistician, Nuffield Department of Population Health, University of Oxford