



Universiteit
Leiden
The Netherlands

Ductal carcinoma in situ of the breast : cancer precursor or not?

Visser, L.L.

Citation

Visser, L. L. (2020, March 10). *Ductal carcinoma in situ of the breast : cancer precursor or not?*. Retrieved from <https://hdl.handle.net/1887/86290>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/86290>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/86290> holds various files of this Leiden University dissertation.

Author: Visser, L.L.

Title: Ductal carcinoma in situ of the breast : cancer precursor or not?

Issue Date: 2020-03-10

Ductal carcinoma in situ of the breast: cancer precursor or not?

Lindy Larissa Visser

The work presented in this thesis was performed at the Netherlands Cancer Institute - Antoni van Leeuwenhoek, Amsterdam, The Netherlands. In cooperation with the Netherlands Comprehensive Cancer Organization (IKNL) and the nationwide network and registry of cytopathology and histopathology in the Netherlands (PALGA foundation).

The research was funded by grants from Pink Ribbon, A Sister's Hope, the Dutch Cancer Society (KWF kankerbestrijding), and Cancer Research UK.

Financial support for publication of this thesis was kindly provided by the Netherlands Cancer Institute.

Cover design: Lindy Visser; Layout: Lindy Visser; Printed by: Gildeprint
ISBN: 978-94-6323-981-3

Copyright © 2020 L.L. Visser, Amsterdam, the Netherlands.

All rights preserved. No part of this publication may be reproduced, stored in an archival system, or transmitted in any form or by any means, without prior permission of the copyright owner.

Ductal carcinoma in situ of the breast: cancer precursor or not?

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 10 maart 2020
klokke 13.45 uur

door
Lindy Larissa Visser

geboren te Alkmaar
in 1989

Promotor: prof. dr. J. Wesseling

Co-promotor: dr. E.H. Lips *NKI-AVL*

Leden

promotiecommissie: prof.dr. M.K. Schmidt
prof. dr. L.F.A. Wessels *Technische Universiteit Delft*
prof. dr. E. Sawyer *King's College London*
prof. dr. P. Devilée
prof. dr. V.T.H.B.M. Smit

*Strive not to be a success,
but rather to be of value*

Albert Einstein
(1879 - 1955)

Table of contents

Chapter 1	Scope of the thesis	9
Chapter 2	Predictors of an invasive breast cancer recurrence after DCIS: A systematic review and meta-analyses	17
Chapter 3	Clinicopathological risk factors for an invasive breast cancer recurrence after ductal carcinoma in situ - A nested case-control study	43
Chapter 4	Deciphering the genomic and transcriptomic landscapes of ductal carcinoma in situ associated with and without subsequent invasive breast cancer	73
Chapter 5	Discordant marker expression between invasive breast carcinoma and corresponding synchronous and preceding DCIS	103
Chapter 6	Approximately 30% of invasive ipsilateral breast cancers after treatment of ductal carcinoma in situ may be new independent primary cancers	127
Chapter 7	Summarizing discussion	157
Appendices		
	Nederlandse samenvatting	171
	List of publications	175
	About the author	177

CHAPTER 1

Scope of the thesis

Introduction

Primum non nocere, “first do no harm”, is one of the promises within the Hippocratic Oath, that medical students must take as an important step in becoming a doctor. In order to honor this promise, doctors must be able to reliably estimate the risks and benefits of treatment for a patient. However, in daily practice these estimates are often uncertain and prone to error. This is a reminder that we need high-quality research to help us make a more balanced treatment choice for every patient. One disease that is in high need of reliable risk estimation is ductal carcinoma in situ (DCIS).

Carcinoma in situ

DCIS belongs to the group of non-invasive breast lesions, called carcinoma in situ (CIS). It is known that CIS lesions have the potential to progress to invasive breast cancer within a time period varying from a few years to several decades.¹ Only a subset of CIS lesions will undergo the transition from in situ to invasive cancer, and not all invasive cancers arise from CIS lesions.^{2,3} Most CIS lesions are asymptomatic. Before the introduction of population-based breast cancer screening in 1989, CIS was rarely diagnosed, but nowadays they represent about 20-25% of all newly diagnosed breast cancers.⁴⁻⁶

DCIS is the most common form of CIS. It is characterized by the neoplastic proliferation of epithelial cells inside the milk duct of the breast. DCIS is considered a different entity than lobular carcinoma in situ (LCIS), as the latter is characterized by the proliferation of epithelial cells in the lobules of the breast. Whereas DCIS is generally accepted to be a potential precursor of invasive breast cancer, LCIS is considered to be a risk factor of future development of invasive disease.

DCIS as breast cancer precursor

DCIS is considered a precursor of invasive breast cancer, because DCIS and IBC are frequently found next to each other sharing genetic alterations as well as risk factors (e.g. age, family history of breast cancer, etc).^{2,7-10} However, other findings reject DCIS being a precursor. First, in contrast to other in situ carcinoma, it has never been demonstrated that surgical removal of DCIS reduces the incidence of subsequent invasive breast cancer. A possible explanation for this is that most DCIS is nowadays detected by mammography screening, which next to an increased detection of DCIS also results in overdiagnosis of breast cancer, thus masking a possible reduction in incidence caused by removal of DCIS.² Second, not all DCIS progresses to IBC and definite proof of DCIS progression to IBC is still lacking. The fact that most DCIS is treated, hinders studies on the natural course of DCIS.

Dilemma of overdiagnosis and overtreatment

Although some DCIS will develop into IBC, the majority of DCIS, if left untreated, is not destined to progress and thus will never become life-threatening.¹¹ This implies that many women are “overdiagnosed”, as they are diagnosed with a disease that would not have caused any symptoms or death.^{12,13} Nonetheless, we are lacking prognostic markers that can predict which DCIS patients will subsequently develop invasive disease. As a result, current treatment guidelines for DCIS dictate that all women diagnosed with DCIS should undergo treatment to prevent the development of invasive breast cancer. This treatment consists of breast conserving surgery (BCS), often followed by radiotherapy and/or endocrine therapy, or a mastectomy. This makes that many women, who have a low-risk to develop subsequent IBC, are being harmed by this intensive treatment without any benefit.¹⁴ Prognostic markers are urgently needed to avoid such “overtreatment”. A better risk prediction will safely spare many women intensive treatment and so preserve their quality of life.

The thesis

Objectives

The objectives of this thesis were to identify prognostic markers predictive of the development of subsequent ipsilateral IBC after DCIS and to explore the clonal relatedness of patient-matched DCIS and subsequent ipsilateral IBC. This will help to stratify a woman’s individual risk of subsequent invasive breast cancer development and will help to avoid overtreatment.

Thesis outline

Numerous prognostic markers have been reported by previous studies, but none of them have shown to be of sufficient value for implementation into the clinic. So, what do we really know? To answer this question, in **chapter 2** we performed a systematic literature review to identify the factors with the strongest predictive value of subsequent IBC after DCIS that should be considered for validation. In addition, we provide insight into frequently introduced biases in prognostic factor studies for DCIS.

In **chapter 3**, we performed our own prognostic factors study for DCIS, by comparing clinicopathological characteristics of DCIS lesions associated with or without subsequent IBC. Next to this, information on the future disease course of DCIS might be encoded in the genomic and transcriptomic profile present at diagnosis. In **chapter 4**, we aimed to identify molecular alterations differentiating lesions that will remain indolent from those that will likely progress to IBC.

To shed light upon the mechanism of DCIS progression, most studies performed so far compared IBC with an adjacent DCIS component, referred to as synchronous DCIS. However, to our knowledge, it has never been investigated whether the synchronous DCIS

and IBC comparison is a good surrogate of primary DCIS and subsequent IBC. In **chapter 5**, we analyzed immunohistochemical marker expression between IBC and corresponding synchronous and preceding DCIS. Furthermore, in **chapter 6** we evaluated the clonal relatedness of 78 patient-matched DCIS and subsequent ipsilateral IBC, to provide evidence of direct progression from DCIS to IBC,.

Lastly, in **chapter 7** the main findings of this thesis are described and put in perspective. Furthermore, strengths, limitations, and methodological challenges are discussed, concluding with recommendations for future research and clinical implications.

References

1. Barnes NLP, Ooi JL, Yarnold JR, Bundred NJ. Ductal carcinoma in situ of the breast. *BMJ*. 2012;344:e797.
2. To T, Wall C, Baines CJ, Miller AB. Is Carcinoma in Situ a Precursor Lesion of Invasive Breast Cancer? *Int J Cancer*. March 2014:1-7.
3. Levi F, Randimbison L, Te V-C, La Vecchia C. Invasive breast cancer following ductal and lobular carcinoma in situ of the breast. *Int J cancer*. 2005;116(5):820-823.
4. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst*. 2002;94(20):1546-1554.
5. Netherlands Comprehensive Cancer Organization (IKNL). DCIS incidence trends over time, the Netherlands.
6. Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2014*. National Cancer Institute. Bethesda, MD
7. O'Connell P, Pekkel V, Fuqua SAW, Osborne CK, Clark GM, Allred DC. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst*. 1998;90(9):697-703.
8. Burkhardt L, Grob TJ, Hermann I, et al. Gene amplification in ductal carcinoma in situ of the breast. *Breast Cancer Res Treat*. 2010;123(3):757-765.
9. Buerger H, Otterbach F, Simon R, et al. Comparative genomic hybridization of ductal carcinoma in situ of the breast-evidence of multiple genetic pathways. *J Pathol*. 1999;187(4):396-402.
10. Buerger H, Otterbach F, Simon R, et al. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol*. 1999;189(4):521-526.
11. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchió C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*. 2010;57(2):171-192.
12. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102(9):605-613.
13. Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res*. 2015;17(1):63.
14. Groen EJ, Elshof LE, Visser LL, et al. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). *Breast*. September 2016:1-10.

Lindy L. Visser
Emma J. Groen
Flora E. van Leeuwen
Esther H. Lips
Marjanka K. Schmidt*
Jelle Wesseling*

Can Epidem Biol Prev. 2019 May;28(5):835-845

*: Shared last author

CHAPTER 2

Predictors of an invasive breast cancer recurrence after DCIS: A systematic review and meta-analyses

Abstract

We performed a systematic review with meta-analyses to summarize current knowledge on prognostic factors for invasive disease after a diagnosis of ductal carcinoma in situ (DCIS). Eligible studies assessed risk of invasive recurrence in women primarily diagnosed and treated for DCIS, and included at least ten ipsilateral invasive breast cancer events and one year of follow-up. Quality In Prognosis Studies (QUIPS) tool was used for risk of bias assessment. Meta-analyses were performed to estimate the average effect size of the prognostic factors. Of 1,781 articles reviewed, 40 articles met the inclusion criteria. Highest risk of bias was attributable to insufficient handling of confounders and poorly described study groups. Six prognostic factors were statistically significant in the meta-analyses: African-American race [pooled estimate (ES), 1.43; 95% confidence interval (CI), 1.15-1.79], premenopausal status (ES, 1.59; 95% CI, 1.20-2.11), detection by palpation (ES, 1.84; 95% CI, 1.47-2.29), involved margins (ES, 1.63; 95% CI, 1.14-2.32), high histologic grade (ES, 1.36; 95% CI, 1.04-1.77), and high p16 expression (ES, 1.51; 95% CI, 1.04-2.19). Six prognostic factors associated with invasive recurrence were identified, whereas many other factors need confirmation in well-designed studies on large patients numbers. Furthermore, we identified frequently occurring biases in studies on invasive recurrence after DCIS. Avoiding these common methodological pitfalls can improve future study designs.

Introduction

With the introduction of the population-based breast cancer screening program in the wealthy world, the incidence of ductal carcinoma *in situ* (DCIS) has increased almost 6-fold.¹⁻⁶ Although some DCIS will develop into invasive breast cancer, the majority of DCIS, if left untreated, is not destined to progress and thus will never become life-threatening.⁷ This implies that many women are overtreated, as they are diagnosed with a disease that would not have caused symptoms or death.⁸ However, we are currently unable to predict which DCIS patients will subsequently develop invasive disease. As a result, almost all women diagnosed with DCIS are nowadays intensively treated with surgical treatment, adjuvant treatment, or both. Many women, who have a low risk to develop subsequent invasive disease, do not benefit from this treatment and thus suffer from overtreatment. Until breast cancer screening programs will include strategies to only detect hazardous disease, we will continue to be faced with large numbers of women diagnosed with low-risk DCIS annually worldwide.

Despite repeated calls for development of prognostic factors for predicting invasive recurrences following DCIS, progress in this field has been slow.⁹ Numerous prognostic factors have been reported, but none have shown to be of sufficient value for implementation into the clinic.¹⁰ This is due to a variety of reasons. For example, sufficiently large, unbiased patient cohorts are lacking to set-up validation studies. Current guidelines dictate surgical excision of DCIS when such a lesion is detected. This makes that almost all DCIS is treated and the natural course of DCIS is poorly understood. On top of this, many previous prognostic factor studies have only limited power, given the low event rate in treated patients and the fact that it can take a decade before the presentation of an invasive recurrence. In all this, in-dept molecular analysis of DCIS is challenging due to the minimal quantity and often limited quality of the DNA and RNA extracted from DCIS. As a result, a multitude of factors are now lost in transition.

In this systematic review, we (1) give an overview of previously published studies on prognostic factors for subsequent invasive recurrence after DCIS, (2) assess these studies for potential bias using a standardized risk assessment tool, and (3) identify the factors with the strongest prognostic value that should be considered for validation. With these results, we want to make recommendations for future studies.

Materials and methods

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement to guide the conduct and reporting of this review.¹¹

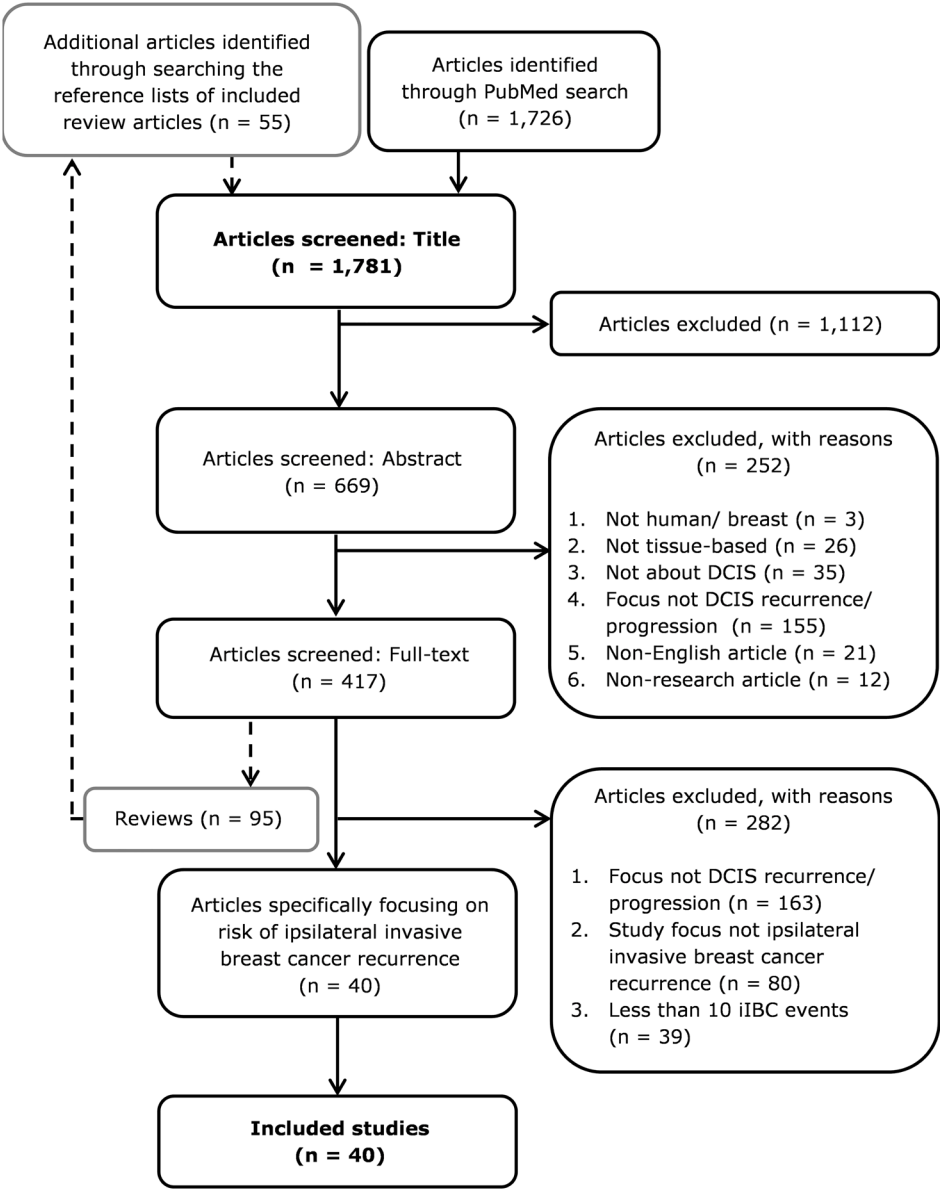


Figure 1. Flow chart of the identification of eligible articles in the systematic literature review. Note that 1,781 papers were identified in the Pubmed database, of which 40 met our inclusion criteria. Reference lists of review articles were searched, and any reference with an ambiguous title was included for screening (arrows with dotted lines).

Eligibility criteria and search strategy

Studies were identified through a systematic search in Pubmed until June 1, 2018 with no language restrictions using the search strategy that can be found in Supplementary Table S1; no limits were set. One reviewer (L.L. Visser) screened titles and abstracts of all papers and assessed their eligibility for the research topic: factors associated with the risk of subsequent ipsilateral invasive breast cancer (iIBC) in women that were primarily diagnosed and treated for DCIS. Studies not reporting original data, letters to the editor, and commentaries were excluded from the review (non-research articles), as were non-English articles (Figure 1). In addition, we selected for studies including at least 1 year of follow-up. Next, full-text articles were screened for inclusion by two reviewers (L.L. Visser and E.J. Groen) independently. Studies including less than ten subsequent invasive breast cancer events after DCIS treatment were excluded, as were studies that did not focus on subsequent *invasive* recurrences as primary end-point. Discrepancies were resolved by group discussion with team members. Reference lists of review articles were searched and any reference with an ambiguous title was included for screening. When multiple studies using the same study population had been published, the study with the largest number of subjects and longest follow-up time was included. If studies used the same study population but reported different prognostic factors, each factor was included separately.

Following the definition of our search strategy, only tumour-related factors and age, race and/or ethnicity, detection method, or menopausal status were included in this systematic review. Incidentally, factors such as treatment, family history of breast cancer, BMI, or lifestyle factors were described in the included studies, but these factors were not included in the analyses.

Data extraction and definitions

During the full-text screening phase, the following data were extracted: source of the study population, single- or multi-centre study, study design, number of DCIS patients, number of iIBC events, period of recruitment, median follow-up time in years, received treatment for DCIS, the identified prognostic factors, the risk estimates – i.e., hazard ratio (HR) or odds ratio (OR) with 95% confidence interval (CI), adjustments, and the statistical method used.

Quality assessment

Next, each study was evaluated independently by two reviewers (L.L. Visser and E.J. Groen) using the Quality in Prognosis Study (QUIPS) tool developed by Hayden and colleagues.^{12,13} Details on the tool used for the assessment are shown in Supplementary Table S2. In brief, domains assessed for bias were study participation, study attrition, prognostic factor measurement, end-point definition, confounding measurement and handling, and statistical analysis and reporting. Each domain was assessed with the help of three to six

Table 1. Study and patient characteristics of included papers

Ref	First author, year	Source population	RCT or Obs	Single- or multi-center	Study design	No. DCIS patients	No. iBC events	Period	Mean follow-up (years)	Adjacent invasion included?	Treatment
17	Rakovitch, 2018	HQ EOC E5194, Ontario DCIS Cohort	RCT*, Obs	multi	cohort	773	65	1994-2003	9.4	n/a	BCS alone
18	Visser, 2018	HQ Netherlands Cancer Registry	Obs	multi	case-control	674	200	1989-2004	12.0	no	BCS alone
19	Pruneri, 2017	European Institute of Oncology	Obs	single	cohort	1488	136	1997-2008	8.2	n/a	all modalities
20	Molinaro, 2016	HQ SEER Northern California	Obs	multi	case-control	1492	167	1983-1996	12.6	no	BCS alone
21	Borgquist, 2015	Uppland, Västmanland (Sweden)	Obs	multi	cohort	324	46	1986-2004	15.3	n/a	all modalities
22	Williams, 2015	DCIS I, IBIS II, IRESSA trial, ERISSAC, lapatinib DCIS	RCT	multi	cohort	314	22	1990-2010	5.0	yes	all modalities
23	Curigliano, 2015	HQ European Institute of Oncology	Obs	single	cohort	1667	201	1996-2008	7.6	n/a	all modalities
24	Cheung, 2014	HQ West Midlands Cancer Registry	Obs	multi	cohort	3930	297	1988-2008	n/a	n/a	all modalities
25	Generali, 2014	John Radcliffe Hospital, Royal Brisbane and Women's Hospital	Obs	multi	cohort	174	25	n/a	12.1	n/a	all modalities
26	Kong, 2014	HQ Ontario Cancer Registry	Obs	multi	cohort	1607	148	1994-2003	10	no	BCS+RT
27	v Bockstal, 2013	Ghent University Hospital	Obs	single	cohort	64	12	1991-2003	9.3	yes	all modalities
28	Solin, 2013	ECOG E5194	RCT*	multi	cohort	327	20	1997-2002	8.8	n/a	BCS +/- HT
29	Donker, 2013	EORTC	RCT	multi	cohort	1010	123	1986-1996	15.8	no	BCS +/- RT
30	Collins, 2013	HQ KPNC, KPSC, HPHC	Obs	multi	case-control	619	225	1990-2001	4.8	n/a	BCS +/- RT
31	Holmberg, 2013	HQ SweDCIS	RCT	multi	case-control	1046	n/a	n/a	n/a	yes	BCS +/- RT
32	Knudsen, 2012	Thomas Jefferson University Hospital	Obs	single	cohort	236	27	1978-2008	9.0	no	BCS alone
33	Alvarado, 2012	University of Texas MD Anderson Cancer Center	Obs	single	cohort	2037	16	1996-2009	5.2	n/a	all modalities
34	Rakovitch, 2012	HQ Sunnybrook Health Sciences Centre	Obs	single	cohort	213	21	1982-2000	7.7	n/a	BCS +/- RT
35	Han, 2012	HQ Sunnybrook Health Sciences Centre	Obs	single	cohort	180	22	1987-2000	7.8	yes	BCS +/- RT

Ref	First author, year	Source population	RCT or Obs	Single- or multi-center	Study design	No. DCIS patients	No. iBC events	Period	Mean follow-up (years)	Adjacent invasion included?	Treatment
36	Witkiewicz, 2011	Thomas Jefferson University Hospital	Obs	single	cohort **	126	16	n/a	n/a	n/a	BCS alone
37	Wapnir, 2011	NSABP B-17, NSABP B-24	RCT	multi	cohort	2612	263	1985-1994	14.7	n/a	BCS+/-RT +/-HT
38	Falk, 2011	Norwegian Breast Cancer Screening Programme	Obs	multi	cohort	3046	96	1993-2007	5.2	no	all modalities
39	Tunon-de-Lara, 2010	Institut Bergonie	Obs	single	cohort	812	47	1971-2001	9.8	no	all modalities
40	Zhou, 2010	Uppland, Västmanland (Sweden)	Obs	multi	cohort	392	34	1986-2004	10.2	n/a	all modalities
41	Pinder, 2010	UKCCR/ANZ	RCT	multi	cohort	1224	55	1990-1998	n/a	yes	BCS+/-RT +/-HT
42	Kerlikowske, 2010	SEER Northern California	Obs	multi	case-control	329	72	1983-1994	8.2	no	BCS alone
43	Witkiewicz, 2009	Thomas Jefferson University Hospital	Obs	single	cohort **	78	11	n/a	12.2	n/a	BCS alone
44	Nofech-Mozes, 2008	Sunnybrook Health Sciences Centre	Obs	single	cohort	133	21	1982-2000	8.9	yes	BCS alone
45	Rakovitch, 2007	Sunnybrook Health Sciences Centre	Obs	single	cohort	615	36	1982-2000	5.9	yes	BCS +/- RT
46	Ringberg, 2007	SweDCIS	RCT	multi	case-control	1046	155	1987-1999	5.2	n/a	BCS +/- RT
47	Hwang, 2007	Breast Cancer Surveillance Consortium	Obs	multi	cohort	3274	83	1993-2005	3.2	no	all modalities
48	Smith, 2006	SEER	Obs	multi	cohort	3409	107	1992-1999	5.0	n/a	BCS +/- RT
49	Li, 2006	SEER	Obs	multi	cohort	37692	1504	1988-2002	n/a	n/a	BCS +/- RT
50	Bijker, 2006	EORTC	RCT	multi	cohort	1010	108	1986-1996	10.5	no	BCS +/- RT
51	Warren, 2005	SEER	Obs	multi	cohort	1103	62	1991-1992	7.5	no	BCS +/- RT
52	Kerlikowske, 2003	SEER Northern California	Obs	multi	cohort	1036	71	1983-1994	6.5	no	BCS alone
53	Teo, 2003	Merseyside (UK)	Obs	multi	case-control	52	12	1989-1999	5.3	no	all modalities
54	Bijker, 2001	EORTC	RCT	multi	cohort	863	66	1986-1996	5.4	no	BCS +/- RT
55	Warberg, 2001	Swedish Cancer Registry	Obs	multi	case-control	570	70	1960-1996	10	no	all modalities
56	Habel, 1998	SEER Washington	Obs	multi	cohort	709	55	1980-1992	5.2	n/a	BCS +/- RT

Abbreviations: "All modalities" consist of treatment with BCS alone, BCS+RT/HT, and mastectomy; HPHC, Harvard Pilgrim Health Care; KPSC, Kaiser Permanente Southern California; N/a = data not available; RCT, randomized clinical trial; Obs, observational study.

*: Non-randomized clinical trial; **: Selection of cohort used in analysis.

prompting questions of which several were modified for the purpose of this study. The assessment for each study was completed by assigning a grade of low, moderate, or high risk of bias to each domain. Any discrepancy in grading was discussed, and if no consensus was reached, a third reviewer (M.K. Schmidt) was consulted. For consistence of assessment, we tested the QUIPS instrument between the two reviewers (L.L. Visser and E.J. Groen) before rating the included studies. The kappa for interobserver agreement was 0.9 (SE of 0.2). In addition, because we found DCIS treatment to be the most strongly confounding variable in previous studies, we explicitly specified this confounder in the QUIPS tool. We classified studies as “high quality (HQ) studies” if they were properly designed and well conducted: these studies were not allowed to have high risk of bias in any of the QUIPS domains and should account for the confounding effect of treatment. Prognostic factors identified in these HQ studies were considered as factors with the strongest predictive value.

Statistical analysis

To estimate the average effect size of the prognostic factors, meta-analyses were performed using the univariate effect sizes reported by the different studies; this was done for all factors reported by more than 1 HQ study. For the absolute effect size difference between studies, pooled estimates were calculated using weighting based on the number of included iIBC events per study [weight per study (%) = (n of DCIS patients with subsequent iIBC in that specific study / total number of DCIS patients with subsequent iIBC of all studies which were used to form the pooled estimate)*100]. In a few articles, effect sizes were not reported or used categories were not comparable with the other studies assessing that specific prognostic factor; hence, these were excluded from the analysis. For the reported effect sizes, pooled estimates were visualized and summarized using a forest plot, and statistical heterogeneity was assessed using Random effect analysis.¹⁴

A funnel plot was used to assess possible publication bias.^{15,16} Because there were only a few estimates/studies for each of the factors, it was only possible to do this for all factors combined. Chi-square tests were performed to compare year of publication and risk of bias per QUIPS domain. For this, studies were divided at the median into publication years 1998-2011 and 2012-2018 and compared the risk of bias per domain. *P* values ≤ 0.05 (2-sided test) were considered statistically significant. All statistical analyses were done using Stata/SE (version 13.1, Statacorp).

Results

Until June 2018, 1,781 papers were identified in the Pubmed database, of which 40 met our inclusion criteria (Figure 1).¹⁷⁻⁵⁶ This low number of included studies was because only a few studies specifically focused on iIBC recurrence after DCIS. Many studies did not specify for the type of recurrence, *in situ* or invasive, and thus were excluded (n = 80).

Table 2. Risk of bias per QUIPS domain for each individual study

First author	Year	HQ	Level of risk of bias due to:				
			Study participation	Study attrition	Prognostic factor measurement	End-point definition	Confounding measurement and handling
Rakovitch	2018	HQ					
Visser	2018	HQ					
Pruneri	2017						
Molinaro	2016	HQ					
Borgquist	2015						
Curigliano	2015	HQ					
Williams	2015						
Cheung	2014	HQ					
Generali	2014						
Kong	2014	HQ					
Holmberg	2013	HQ					
Donker	2013						
Collins	2013	HQ					
Solin	2013						
Van Bockstal	2013						
Alvarado	2012						
Han	2012	HQ					
Knudsen	2012						
Rakovitch	2012	HQ					
Tunon-de-Lara	2011						
Witkiewicz	2011						
Wapnir	2011						
Falk	2011						
Kerlikowske	2010	HQ					
Pinder	2010						
Zhou	2010						
Witkiewicz	2009						
Nofech-Mozes	2008						
Rakovitch	2007						
Ringberg	2007	HQ					
Hwang	2007	HQ					
Smith	2006	HQ					
Li	2006	HQ					
Bijker	2006						
Warren	2005						
Kerlikowske	2003	HQ					
Teo	2003						
Bijker	2001	HQ					
Warnberg	2001						
Habel	1998						

Risk of bias
 : Low
 : Moderate
 : High

Study and patient characteristics

Study and patient characteristics of the included studies can be found in Table 1. The sample size of the included studies ranged from 52 to 37,692 patients, and mean follow-up time ranged from 3.2 to 15.8 years. Seven studies included DCIS patients who also had an adjacent invasive component or microinvasion and seven other studies explicitly excluded these patients. Furthermore, 14 studies included patients from all treatment modalities, breast conserving surgery (BCS) alone, BCS + radiotherapy (RT)/ hormonal therapy (HT),

Chapter 2

and mastectomy, whereas 16 other studies included only BCS treated patients (+/- RT). Ten studies included patients who underwent one treatment modality: BCS+RT (n=1) and BCS alone (n=9).

For all studies, data was collected retrospectively, regarding patients diagnosed with DCIS between 1960 and 2010. For this, hospital registries, national registries, or data from clinical trials were used. Both cohort (80%) and case-control designs (20%) were used. Seventy percent were multi-center studies, and 30% involved only a single center.

Table 3. List of factors that were assessed in the high quality studies

Factor	Number of HQ studies:	
	Assessed factor	Statistical significant finding
Age at DCIS diagnosis	6	4
Calcification	2	1
Calgranulin status	1	0
COX-2 status	2	1
Cyclin D1 status	1	0
DCIS architecture	6	2
Detection method	4	2
ER status	3	0
Focality	1	0
Grade, histologic	7	1
HER2 status	3	0
Ki67 status	2	0
Lesion size	7	0
Margin status	4	0
Menopausal status	2	2
Necrosis	4	0
p16 status	2	1
p21 status	1	0
p53 status	3	0
Periductal fibrosis	1	1
Periductal lymphocytes	1	0
PR status	3	0
Psoriasin status	1	0
Race and/or ethnicity	2	1
Subtypes, intrinsic	2	0
Year of DCIS diagnosis	1	0

Assessment of quality of prognosis studies (QUIPS)

We assessed six QUIPS domains: study participation, study attrition, end-point definition, prognostic factor measurement, confounding measurement and handling, and statistical analysis and reporting (Supplementary Table S2). A high or moderate risk of bias was identified in at least one domain in 39 of the 40 studies, with 22 studies having a high risk of bias in at least one domain (Table 2). The domains with the highest risk of bias were confounding measurement and handling and study participation, which had a high risk of bias in 16 and 8 of the 40 studies, respectively.

In total, 11 of the 40 studies (27.5%) used the study design to account for potential confounding through either matching, stratification, or initial assembly of comparable groups. Eighteen of the 40 studies (45.0%) accounted for confounding effect in the analysis stage. The remaining 11 studies (27.5%) did not perform adjustments for confounding. The reasons for the high-risk-of-bias ratings in the study participation domain were incomplete description of inclusion and exclusion criteria and/or poorly described baseline characteristics of the study group. Cox proportional hazard analysis was performed in all studies but two. One of these two studies was assessed as having a high risk of bias in the statistical analysis domain, since the analysis used was not appropriate for the design of the study.

None of the studies had a high risk of bias in the domains end-point definition and prognostic factor measurement.

Finally, we assessed the effect of time period of publication on risk of bias. We divided the studies at the median into publication years 1998-2011 and 2012-2018 and compared the risk of bias per domain. There was no significant difference in any of the study domains.

Exploring publication bias

Supplementary Figure S1 shows the funnel plot that was used to assess publication bias by including all prognostic factors together in one plot. The funnel plot shows that both significant and non-significant factors related to outcome were published. As such, we conclude that there was no evidence for publication bias.

Identification of the high quality (HQ) studies and their reported prognostic factors

We filtered for the studies without a high risk of bias in any of the QUIPS domains and selected only those studies that accounted for the confounding effect of treatment (HQ studies). Only 17 studies met these criteria (Table 1 and 2). All together, these 17 HQ studies assessed 26 different factors and identified 10 different potential prognostic factors, which were assessed in Y HQ studies and reported to have statistically significant association with subsequent invasive breast cancer in X of these studies (X/Y): high histologic grade (1/7), young age at DCIS diagnosis (4/6), solid DCIS architecture (2/6), detection by palpation (2/4), pre-menopausal status (2/2), African-American race (1/2), presence of calcification (1/2), high p16 expression (1/2), high COX-2 expression (1/2), and presence of periductal

fibrosis (1/1; Table 3 and Supplementary Table S3). None of the studies assessed all prognostic factors. Notably, studies examining the same prognostic factor often showed inconsistent results (Supplementary Table S3).

Meta-analyses

Meta-analyses were performed to estimate the average effect size of the prognostic factors; this was done for all factors reported by more than 1 HQ study, regardless of their statistical significance (Figure 2 and Supplementary Figure S2). Most of the factors seemed to point to a higher relative risk of subsequent iIBC for DCIS patients, although effects were generally small. Six prognostic factors had a statistically significant pooled estimate: African-American race [pooled estimate (ES), 1.43; 95% CI, 1.15-1.79], premenopausal status (ES, 1.59; 95% CI, 1.20-2.11), detection by palpation (ES, 1.84; 95% CI, 1.47-2.29), involved margins (ES, 1.63; 95% CI 1.14-2.32), high histologic grade (poorly differentiated; ES, 1.36; 95% CI, 1.04-1.77), and high p16 expression (ES, 1.51; 95% CI, 1.04-2.19). For these six prognostic factors, the heterogeneity test demonstrated consistency of the estimates reported in the included studies. Although, histologic grade showed a trend towards heterogeneity ($P = 0.09$). None of the studies reported all these six prognostic factors. Meta-analyses could not be performed for the factors age at diagnosis, DCIS architecture, lesion size, and year of DCIS diagnosis because the categories used in the studies were not comparable.

Discussion

The purpose of this review was two-fold. First, we aimed to identify prognostic factors with statistically significant association with subsequent ipsilateral invasive breast cancer (iIBC) that deserve validation. We identified 17 high quality (HQ) studies, assessing 26 factors, of which six prognostic factors were statistically significantly associated with subsequent iIBC risk in the meta-analyses: African-American race, premenopausal status, detection by palpation, involved margins, high histologic grade (poorly differentiated), and high p16 expression. Second, we aimed to give insight into bias that was frequently introduced in previously published prognostic factor studies for subsequent iIBC after preceding DCIS. Highest risk of bias in the studies was attributable to insufficient measurement and handling of confounders and poorly described study groups.

The association between the six unfavorable prognostic factors and subsequent iIBC risk can be biologically explained. When DCIS has involved margins this indicates that residual tumor cells are left behind at the resection site. These cells can subsequently grow out and form a recurrence, which could be invasive disease. Premenopausal status and African-American race are known independent predictors of a worse breast cancer outcome.^{57,58} Furthermore, literature has shown that DCIS detected by palpation would be more

aggressive than screening-detected DCIS, as these DCIS lesions are more often ER negative and HER2 positive.⁵⁹ The same holds true for DCIS lesions of high histologic grade.⁶⁰ Lastly, p16 mediates cell-cycle arrest through the p16/Rb signaling pathway. Disruption of the p16/Rb signaling pathway is an oncogenic event and results in sustained cellular proliferation, which can lead to DCIS progression to iIBC.⁶¹

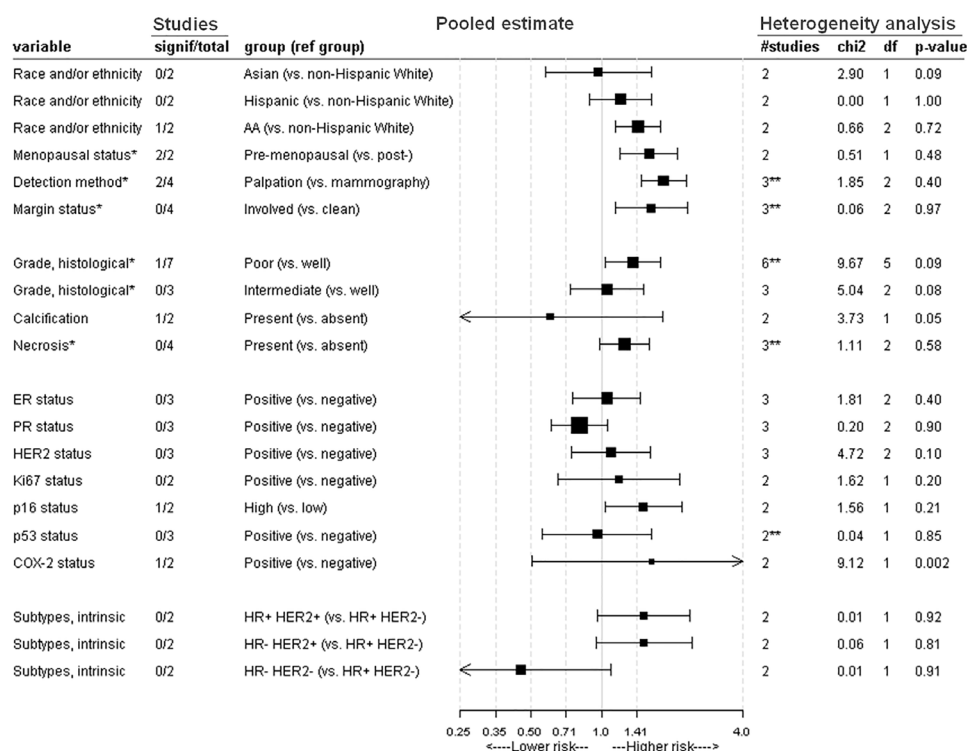


Figure 2. Pooled estimates and heterogeneity analysis of prognostic factors reported in more than one HQ study. Pooled estimates were calculated using weighting based on the number of included iIBC events per study [weight per study (%) = (n of DCIS patients with subsequent iIBC in that specific study / total number of DCIS patients with subsequent iIBC of all studies which are used to form the pooled estimate)*100]. Heterogeneity was assessed using Random effect (DerSimonian and Laird) analyses. The column “studies signif/total” represents the number of high quality studies that reported a statistical significant association for the prognostic factor and subsequent iIBC risk and the total number of HQ studies that assessed the prognostic factor. *, factors used in routine clinical practice for DCIS; **, Number of studies included in the analysis: A few studies did not report effect-sizes or used categories were not comparable with the other studies assessing that specific prognostic factor, hence these studies were excluded from the analysis; AA, African-American.

Whether or not to use histologic grade as a prognostic marker for invasive recurrence after DCIS is a matter of debate. In our meta-analysis, histologic grade showed a trend towards heterogeneity, which is likely caused by differences in histologic classification methods.^{41,62–64} Moreover, all methods suffer from reproducibility problems causing high inter-observer variability.^{65,66}

Zhang and colleagues carried out the first meta-analysis specifically focusing on ipsilateral *invasive* recurrence after DCIS.⁶⁷ In line with our study, they found that positive margins and non-screening detected lesions were associated with a higher risk of iIBC after DCIS. However, they included only 18 studies. Although Zhang and colleagues performed bias assessment of the included articles, using a different method than we did, they did not report on the results from the bias assessment, making it likely that also studies with a high risk of bias were included in their meta-analyses. In addition to the study by Zhang and colleagues, two other meta-analyses have been published, although focusing on ipsilateral tumor recurrence (both *in situ* and invasive) preceding DCIS. Boyages and colleagues found that the presence of necrosis, involved margin, high histologic grade, and large tumor size were predictive of ipsilateral recurrence for DCIS.⁶⁸ In addition to these factors, Wang and colleagues reported that multifocality and symptomatic DCIS were also associated with high risk of ipsilateral breast recurrence.⁶⁹ In our study, multifocality was not assessed, and meta-analyses of necrosis, histologic grade, and tumor size yielded non-statistically significant results. However, previous literature has indicated that risk factors for subsequent invasive disease and recurrence of DCIS may not be identical; thus, combining *in situ* recurrence and invasive recurrence into a single group may obscure the real risk factors for invasive disease after DCIS.⁵² This could explain the inconsistent meta-analysis results of our study and the studies mentioned above and highlights the need to specify for the type of recurrence when performing a prognostic factor study for DCIS.

Next to the prognostic factors we found to be statistically significant in the meta-analyses, many other factors were identified in the included studies. This variability could firstly be explained by underreporting of the prognostic factors, because none of the studies assessed all prognostic factors. Secondly, the presence of unadjusted confounding could also play a role in this, because this makes that any risk estimate could be misleading. The most important confounder in the studies was DCIS treatment: This variable was a risk factors for subsequent iIBC among DCIS patients while at the same time associated with the prognostic factors of interest.⁷⁰ Confounding can be accounted for at the design stage of the study (e.g., by matching or randomization) and/or at the analysis stage, given the confounders have been measured properly. Twenty-nine included studies properly adjusted for confounding effect. Remarkably, 11 studies did not include any adjustments at all.

All patients included in prognostic factor studies for DCIS are treated. As most studies did not include genomic characterization, we could not confirm whether the invasive recurrences studied were indeed all clonally related to the primary DCIS lesion. As they might also be

second primary tumors, the prognostic factors identified could also be risk factors for any second invasive breast event after DCIS. In addition, some DCIS cases developed early recurrences (within 4 months), questioning if these were not missed invasive cancers. As we know that the rate of missed invasive disease at DCIS diagnosis is 11-25%, it is unlikely that this will be a major percentage of the recurrences reported.⁷¹⁻⁷⁴

High risk of bias attributed to selective study participation was mostly because the source of patient (clinical and histopathologic) information was often not mentioned or not properly described. The same holds true for details on inclusion and exclusion criteria. Incomplete description of these criteria can bias the estimates in an uncertain direction. In addition, baseline characteristics were often not adequately described and should at least comprise the factors that are reported during routine diagnosis and treatment, such as: age at diagnosis, histologic grade, clinical presentation, received treatment for DCIS, lesion size, and margin status. Furthermore, some studies included DCIS patients with an adjacent invasive component or microinvasion. Prognostic factor studies for DCIS are aiming to find predictors of subsequent invasive diseases. Yet, DCIS lesion with a (micro) invasive component are already invasive disease. Including these lesions in the analysis is not appropriate, because this may obscure the risk factors for subsequent iIBC after DCIS. Thus, DCIS with an adjacent invasive component or microinvasion should be excluded from such a study. The same holds true for the inclusion of patients treated by mastectomy. Because the recurrence risk after mastectomy is negligible, inclusion of these patients into a study assessing risk of invasive recurrence after DCIS is likely to be less adequate. Of note, two HQ studies, Cheung and colleagues and Curigliano and colleagues, included a substantial proportion of patients treated with mastectomy. Despite this, these studies were still considered as HQ studies following our predefined criteria. Exclusion of these two studies from the meta-analyses did not substantially alter the results (data not shown).

Although study attrition did not introduce a high risk of bias, it was a recurrent problem. Next to the proportion of the initial patient group available for analysis at the end of the study, it is also important to report the reasons why certain patients were not included in the analysis. If the reason for exclusion was related to the study's end-point (missingness not at random), this can substantially affect risk estimates, either towards unity or away from it. Only a few studies included in this review explored differences between drop-outs and non-drop-outs. This could contribute to the wide variations in prognostic factors identified and non-reproducibility of prognostic factors between studies. Most of the factors identified were associated with small effect sizes, and the clinical relevance of these factors therefore is questionable.

Many studies included in this systematic review are retrospective studies that used hospital registries or national registries as a data source and working with these data is a challenge. Registry-based studies often depend on the size, quality, completeness of relevant variables, and features of the registry on which the study is based.⁷⁵ Furthermore,

there are worries about data quality related to end-point measures in registries, end-point information such as migration abroad or death from other causes is not always included.^{76,77} This is a general concern regarding registry-based studies which can only be solved by improving source data. The remainder of the studies used clinical trial data as data source. Clinical trials have the advantage in finding prognostic factors as patient groups are often randomized and thus the analysis does not suffer from confounding. Yet, as clinical trials may focus on highly selected patient groups (e.g., specific age range, lesion size range, etc.), generalizability of trial results might be limited.

This systematic review has several strengths. First, to our knowledge we are the first to perform bias assessment on prognostic factor studies for DCIS. Second, using the QUIPS tool, we were able to provide insight into the most frequently occurring biases in prognostic factor studies in a standardized way. This enabled us to subsequently identify the studies including the least bias, in order to identify factors with the strongest predictive value regarding subsequent iIBC risk after DCIS.

Our study also has some limitations. First, use of the QUIPS tool still involved subjective judgment in assigning a score for each of the 6 domains, although we minimized this by assessing the included studies in a consistent manner using specific criteria for each domain and by assigning two independent assessors. The interobserver kappa value showed an excellent consistency between the two assessors. Second, because the prognostic factors examined differed widely among the studies, the prognostic evidence of the factors obviously only relied on a few publications available: but all studies included in the analysis were HQ studies. Third, studies that we classified as HQ were not allowed to have high risk of bias in any of the QUIPS domains. However, studies with high risk of bias in at least one QUIPS domain might be as good (or bad) as studies with moderate risk of bias in three domains.

In conclusion, measurement and evaluation of prognostic factors have the potential to improve the clinical management of women diagnosed with DCIS. Nonetheless, studies assessing these factors should be of sufficient rigor to reach a high level of specificity and sensitivity. We highly recommend the six prognostic factors for independent validation, although with a critical note added to the use of histologic grade as a prognostic factor. Next to this, we encourage researchers to remain searching for other factors. Also, we could not assess all reported prognostic factors in our meta-analyses, as some were only assessed by a single study. Thus, the potential of these factors remains unproven, but could be confirmed in future studies. In addition, we showed that not accounting for the confounding effect of DCIS treatment is the main cause of study bias, indicating that is of utmost importance to correct for this. Furthermore, we encourage researchers to describe their used patient groups in high detail. Lastly, in the analysis stage, the type of recurrence should be specified: *in situ* or invasive. This, because invasive recurrences increase a patient's risk of dying from breast cancer, and thus should be an (additional) important end-point of interest in prognostic factor studies of DCIS. These insights and the use of for example the STROBE guidelines⁷⁸

can help researchers improve their study designs and avoid common methodological pitfalls.

This systematic review underlines the high need of well-designed studies with large patient numbers that undergo independent validation.⁷⁹ Currently, initiatives have been established to make this happen and translate promising prognostic factors to clinical practice. One of these initiatives is the PRECISION (PREvent ductal Carcinoma In Situ Invasive Overtreatment Now) initiative, funded by Cancer Research UK and the Dutch Cancer Society (<https://www.cancerresearchuk.org/funding-for-researchers/how-we-deliver-research/grand-challenge-award/funded-teams-wesseling>).⁸⁰ In addition, non-inferiority trials, like LORD, LORIS, and COMET, have been initiated and will be important in prospective validation of prognostic factors.^{81–83} We hope our review will ultimately contribute to the identification of reliable and clinically meaningful prognostic factors for DCIS in the near future. This may help us to distinguish indolent from potentially hazardous DCIS, thereby putting an end to the current overtreatment dilemma.

Article information

Funding: This work was jointly funded by Cancer Research UK and the Dutch Cancer Society (grant number C38317/A24043, to J. Wesseling).

Conflict of Interest: The authors declare no conflict of interest.

Author Contributions: LLV participated in the design of the study, data collection, bias assessment, performed the statistical analysis with input from MKS, and wrote the manuscript. EJG assisted in the data collection and bias assessment. EHL assisted in writing the manuscript. JW, FvL, and MSK participated in the design of the study, assisted in the interpretation of the findings, and writing the manuscript.

Supplementary data: Supplementary Table S1–3 are available with the published article.

References

1. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst.* 2002;94(20):1546-1554.
2. Netherlands Comperhensive Cancer Organization (IKNL). DCIS incidence in the Netherlands.
3. Information Services Division (ISD) Scotland. DCIS incidence trends over time, Scotland.
4. Welsh Cancer Intelligence and Surveillance Unit. DCIS incidence trends over time, Wales.
5. Office for National Statistics. DCIS incidence trends over time, UK.
6. Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2014*. National Cancer Institute. Bethesda, MD
7. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchió C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology.* 2010;57(2):171-192.
8. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst.* 2010;102(9):605-613.
9. Bartlett JMS, Nofech-Moses S, Rakovitch E. Ductal carcinoma in situ of the breast: can biomarkers improve current management? *Clin Chem.* 2014;60(1):60-67.
10. Lari S a, Kuerer HM. Biological Markers in DCIS and Risk of Breast Recurrence: A Systematic Review. *J Cancer.* 2011;2:232-261.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ.* 2009;339.
12. Hayden J a, Côté P, Bombardier C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann Intern Med.* 2006;144(6):427-437.
13. Hayden J a, Windt D a Van Der, Cartwright JL, Co P. Research and Reporting Methods Annals of Internal Medicine Assessing Bias in Studies of Prognostic Factors. *Ann Intern Med.* 2013;144:427-437.
14. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials.* 1986;7(3):177-188.
15. Sterne JAC, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ.* 2001;323(7304):101-105.
16. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *J Clin Epidemiol.* 2001;54(10):1046-1055.
17. Rakovitch E, Gray R, Baehner FL, et al. Refined estimates of local recurrence risks by DCIS score adjusting for clinicopathological features: a combined analysis of ECOG-ACRIN E5194 and Ontario DCIS cohort studies. *Breast Cancer Res Treat.* 2018;169(2):359-369.
18. Visser LL, Elshof LE, Schaapveld M, et al. Clinicopathological Risk Factors for an Invasive Breast Cancer Recurrence after Ductal Carcinoma In Situ-A Nested Case-Control Study. *Clin Cancer Res.* April 2018:clincanres.0201.2018.
19. Pruneri G, Lazzeroni M, Bagnardi V, et al. The prevalence and clinical relevance of tumor-infiltrating lymphocytes (TILs) in ductal carcinoma in situ of the breast. *Ann Oncol Off J Eur Soc Med Oncol.* 2017;28(2):321-328.
20. Molinaro AM, Sison JD, Ljung B-M, Tlsty TD, Kerlikowske K. Risk prediction for local versus regional/metastatic tumors after initial ductal carcinoma in situ diagnosis treated by lumpectomy. *Breast Cancer Res Treat.* 2016;157(2):351-361.
21. Borgquist S, Zhou W, Jirström K, et al. The prognostic role of HER2 expression in ductal breast carcinoma in situ (DCIS); a population-based cohort study. *BMC Cancer.* 2015;15(1):468.
22. Williams K, Barnes N, Cramer A, et al. Molecular Phenotypes of DCIS predict overall and invasive recurrence. *Ann Oncol.* 2015.
23. Curigliano G, Disalvatore D, Esposito A, et al. Risk of subsequent in situ and invasive breast cancer in human epidermal growth factor receptor 2-positive ductal carcinoma in situ. *Ann Oncol.* 2015;26(4):682-687.
24. Cheung S, Booth ME, Kearins O, Dodwell D. Risk of subsequent invasive breast cancer after a diagnosis of ductal carcinoma in situ (DCIS). *Breast.* September 2014.

25. Generali D, Buffa FM, Deb S, et al. COX-2 expression is predictive for early relapse and aromatase inhibitor resistance in patients with ductal carcinoma in situ of the breast, and is a target for treatment. *Br J Cancer*. 2014;111(1):46-54.
26. Kong I, Narod SA, Taylor C, et al. Age at diagnosis predicts local recurrence in women treated with breast-conserving surgery and postoperative radiation therapy for ductal carcinoma in situ: a population-based outcomes analysis. *Curr Oncol*. 2014;21(1):e96-e104.
27. Van Bockstal M, Lambein K, Gevaert O, et al. Stromal architecture and periductal decorin are potential prognostic markers for ipsilateral locoregional recurrence in ductal carcinoma in situ of the breast. *Histopathology*. 2013;63(4):520-533.
28. Solin LJ, Gray R, Baehner FL, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. *J Natl Cancer Inst*. 2013;105(10):701-710.
29. Donker M, Liti??re S, Werutsky G, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma In Situ: 15-year recurrence rates and outcome after a recurrence, from the EORTC 10853 randomized phase III trial. *J Clin Oncol*. 2013;31(32):4054-4059.
30. Collins LC, Achacoso N, Haque R, et al. Risk factors for non-invasive and invasive local recurrence in patients with ductal carcinoma in situ. *Breast Cancer Res Treat*. 2013;139(2):453-460.
31. Holmberg L, Wong YNS, Tabár L, et al. Mammography casting-type calcification and risk of local recurrence in DCIS: analyses from a randomised study. *Br J Cancer*. 2013;108(4):812-819.
32. Knudsen ES, Pajak TF, Queenan M, et al. Retinoblastoma and phosphate and tensin homolog tumor suppressors: impact on ductal carcinoma in situ progression. *J Natl Cancer Inst*. 2012;104(23):1825-1836.
33. Alvarado R, Lari S a., Roses RE, et al. Biology, Treatment, and Outcome in Very Young and Older Women with DCIS. *Ann Surg Oncol*. 2012;19(12):3777-3784.
34. Rakovitch E, Nofech-Mozes S, Hanna W, et al. HER2/neu and Ki-67 expression predict non-invasive recurrence following breast-conserving therapy for ductal carcinoma in situ. *Br J Cancer*. 2012;106(6):1160-1165.
35. Han K, Nofech-Mozes S, Narod S, et al. Expression of HER2neu in ductal carcinoma in situ is associated with local recurrence. *Clin Oncol (R Coll Radiol)*. 2012;24(3):183-189.
36. Witkiewicz AK, Rivadeneira DB, Ertel A, et al. Association of RB/p16-pathway perturbations with DCIS recurrence: dependence on tumor versus tissue microenvironment. *Am J Pathol*. 2011;179(3):1171-1178.
37. Wapnir IL, Dignam JJ, Fisher B, 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. *J Natl Cancer Inst*. 2011;103(6):478-488.
38. Falk RS, Hofvind S, Skaane P, Haldorsen T. Second events following ductal carcinoma in situ of the breast: a register-based cohort study. *Breast Cancer Res Treat*. 2011;129(3):929-938.
39. Tunon-de-Lara C, André G, Macgrogan G, et al. Ductal carcinoma in situ of the breast: influence of age on diagnostic, therapeutic, and prognostic features. Retrospective study of 812 patients. *Ann Surg Oncol*. 2011;18(5):1372-1379.
40. Zhou W, Jirström K, Johansson C, et al. Long-term survival of women with basal-like ductal carcinoma in situ of the breast: a population-based cohort study. *BMC Cancer*. 2010;10:653.
41. Pinder SE, Duggan C, Ellis IO, et al. A new pathological system for grading DCIS with improved prediction of local recurrence: results from the UKCCCR/ANZ DCIS trial. *Br J Cancer*. 2010;103(1):94-100.
42. Kerlikowske K, Molinaro AM, Gauthier ML, et al. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma in situ diagnosis. *J Natl Cancer Inst*. 2010;102(9):627-637.
43. Witkiewicz AK, Dasgupta A, Nguyen KH, et al. Stromal caveolin-1 levels predict early DCIS progression to invasive breast cancer. *Cancer Biol Ther*. 2009;8(11):1071-1079.
44. Nofech-Mozes S, Spayne J, Rakovitch E, et al. Biological Markers Predictive of Invasive Recurrence in DCIS. *Clin Med Oncol*. 2008;2(416):7-18.

45. Rakovitch E, Pignol J-PJ-P, Hanna W, et al. Significance of Multifocality in Ductal Carcinoma In Situ: Outcomes of Women Treated With Breast-Conserving Therapy. *J Clin Oncol*. 2007;25(35):5591-5596.
46. Ringberg A, Nordgren H, Thorstensson S, 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. *Eur J Cancer*. 2007;43(2):291-298.
47. Hwang ES, Miglioretti DL, Ballard-Barbash R, Weaver DL, Kerlikowske K. Association between breast density and subsequent breast cancer following treatment for ductal carcinoma in situ. *Cancer Epidemiol Biomarkers Prev*. 2007;16(12):2587-2593.
48. Smith BD, Haffty BG, Buchholz TA, et al. Effectiveness of radiation therapy in older women with ductal carcinoma in situ. *J Natl Cancer Inst*. 2006;98(18):1302-1310.
49. Li CI, Malone KE, Saltzman BS, Daling JR. Risk of invasive breast carcinoma among women diagnosed with ductal carcinoma in situ and lobular carcinoma in situ, 1988-2001. *Cancer*. 2006;106(10):2104-2112.
50. EORTC Breast Cancer Cooperative Group, EORTC Radiotherapy Group, Bijker N, 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. *J Clin Oncol*. 2006;24(21):3381-3387.
51. Warren JL, Weaver DL, Bocklage T, et al. The frequency of ipsilateral second tumors after breast-conserving surgery for DCIS: a population based analysis. *Cancer*. 2005;104(9):1840-1848.
52. Kerlikowske K, Molinaro A, Cha I, et al. Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy. *J Natl Cancer Inst*. 2003;95(22):1692-1702.
53. Teo NB, Shoker BS, Jarvis C, Martin L, Sloane JP, Holcombe C. Angiogenesis and invasive recurrence in ductal carcinoma in situ of the breast. *Eur J Cancer*. 2003;39(1):38-44.
54. Bijker N, Peterse JL, Duchateau L, 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. *J Clin Oncol*. 2001;19(8):2263-2271.
55. Wärnberg F, Bergh J, Zack M, Holmberg L. Risk factors for subsequent invasive breast cancer and breast cancer death after ductal carcinoma in situ: a population-based case-control study in Sweden. *Cancer Epidemiol Biomarkers Prev*. 2001;10(5):495-499.
56. Habel LA, Daling JR, Newcomb PA, et al. Risk of recurrence after ductal carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev*. 1998;7(8):689-696.
57. Newman LA, Mason J, Cote D, et al. African-American ethnicity, socioeconomic status, and breast cancer survival: a meta-analysis of 14 studies involving over 10,000 African-American and 40,000 White American patients with carcinoma of the breast. *Cancer*. 2002;94(11):2844-2854.
58. Partridge AH, Hughes ME, Warner ET, et al. Subtype-Dependent Relationship Between Young Age at Diagnosis and Breast Cancer Survival. *J Clin Oncol*. 2016;34(27):3308-3314.
59. 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. *Eur J Surg Oncol*. 2014;40(3):249-254.
60. Bijker N, Peterse JL, Duchateau L, 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. *Br J Cancer*. 2001;84(4):539-544.
61. Gauthier ML, Berman HK, Miller C, et al. Abrogated response to cellular stress identifies DCIS associated with subsequent tumor events and defines basal-like breast tumors. *Cancer Cell*. 2007;12(5):479-491.
62. Holland R, Peterse JL, Millis RR, et al. Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol*. 1994;11(3):167-180.
63. Poller DN, Silverstein MJ, Galea M, et al. Ideas in pathology. Ductal carcinoma in situ of the breast: a proposal for a new simplified histological classification association between cellular proliferation and c-erbB-2 protein expression. *Mod Pathol*. 1994;7(2):257-262.

64. Scott MA, Lagios MD, Axelsson K, Rogers LW, Anderson TJ, Page DL. Ductal carcinoma in situ of the breast: reproducibility of histological subtype analysis. *Hum Pathol*. 1997;28(8):967-973.
65. Sloane JP, Amendoeira I, Apostolikas N, 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. *Hum Pathol*. 1998;29(10):1056-1062.
66. Van Bockstal M, Baldewijns M, Colpaert C, et al. Dichotomous histopathological assessment of ductal carcinoma in situ of the breast results in substantial inter-observer concordance. *Histopathology*. August 2018:0-3.
67. Zhang X, Dai H, Liu B, Song F, Chen K. Predictors for local invasive recurrence of ductal carcinoma in situ of the breast: a meta-analysis. *Eur J Cancer Prev*. 2016;25(1):19-28.
68. Boyages J, Delaney G, Taylor R. Predictors of local recurrence after treatment of ductal carcinoma in situ: a meta-analysis. *Cancer*. 1999;85(3):616-628.
69. Wang S-Y, Shamliyan T, Virnig BA, Kane R. Tumor characteristics as predictors of local recurrence after treatment of ductal carcinoma in situ: a meta-analysis. *Breast Cancer Res Treat*. 2011;127(1):1-14.
70. Delgado-Rodriguez M. Bias. *J Epidemiol Community Heal*. 2004;58(8):635-641.
71. Burak WE, Owens KE, Tighe MB, et al. Vacuum-assisted stereotactic breast biopsy: histologic underestimation of malignant lesions. *Arch Surg*. 2000;135(6):700-703.
72. Brennan ME, Turner RM, Ciatto S, et al. Ductal carcinoma in situ at core-needle biopsy: meta-analysis of underestimation and predictors of invasive breast cancer. *Radiology*. 2011;260(1):119-128.
73. Soumian S, Verghese ET, Booth M, 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):1337-1340.
74. Grimm LJ, Ryser MD, Partridge AH, et al. Surgical Upstaging Rates for Vacuum Assisted Biopsy Proven DCIS: Implications for Active Surveillance Trials. *Ann Surg Oncol*. 2017;24(12):3534-3540.
75. Li G, Sajobi TT, Menon BK, et al. Registry-based randomized controlled trials- what are the advantages, challenges, and areas for future research? *J Clin Epidemiol*. 2016;80:16-24.
76. Becher H, Winkler V. Estimating the standardized incidence ratio (SIR) with incomplete follow-up data. *BMC Med Res Methodol*. 2017;17(1):1-10.
77. Lauer MS, D'Agostino RB. The randomized registry trial--the next disruptive technology in clinical research? *N Engl J Med*. 2013;369(17):1579-1581.
78. Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007;Volume1(10):e296.
79. Gierisch JM, Myers ER, Schmit KM, Crowley MJ. Prioritization of Research Addressing Management Strategies for Ductal Carcinoma In Situ. *Ann Intern Med*. 2014;(January).
80. Shelley Hwang E, Thompson A. Editorial: What Can Molecular Diagnostics Add to Locoregional Treatment Recommendations for DCIS? *JNCI J Natl Cancer Inst*. 2017;109(4).
81. Elshof LE, Tryfonidis K, Slaets L, 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. *Eur J Cancer*. 2015;51(12):1497-1510.
82. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen detected DCIS; The LORIS trial. *Eur J Cancer*. 2014;51(16):2296-2303.
83. Youngwirth LM, Boughey JC, Hwang ES. Surgery versus monitoring and endocrine therapy for low-risk DCIS: The COMET trial. *Bull Am Coll Surg*. 2017;102(1):62-63.

Supplementary figures

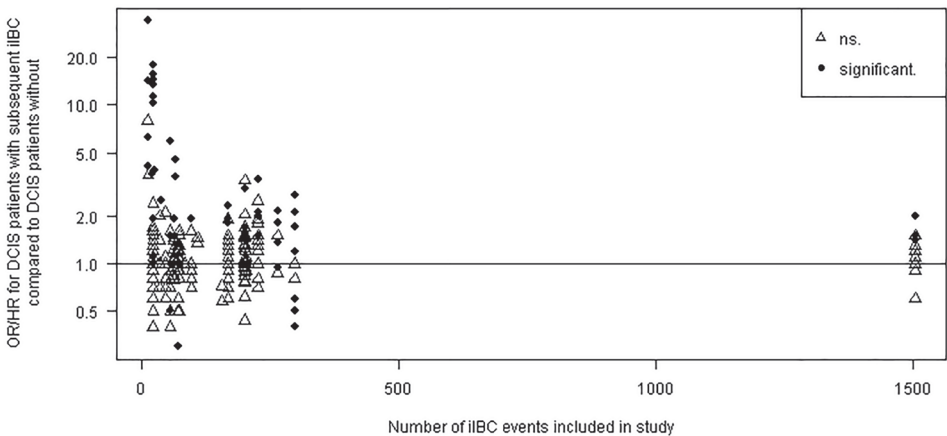


Figure S1. Funnel plot showing the number of DCIS patients with a subsequent ipsilateral invasive breast cancer (iIBC) event included in the study related to the effect of a specific prognostic factor for DCIS patient with subsequent iIBC compared to DCIS patients without a recurrence. This funnel plot includes all factors that were assessed for their possible association with iIBC in all studies.

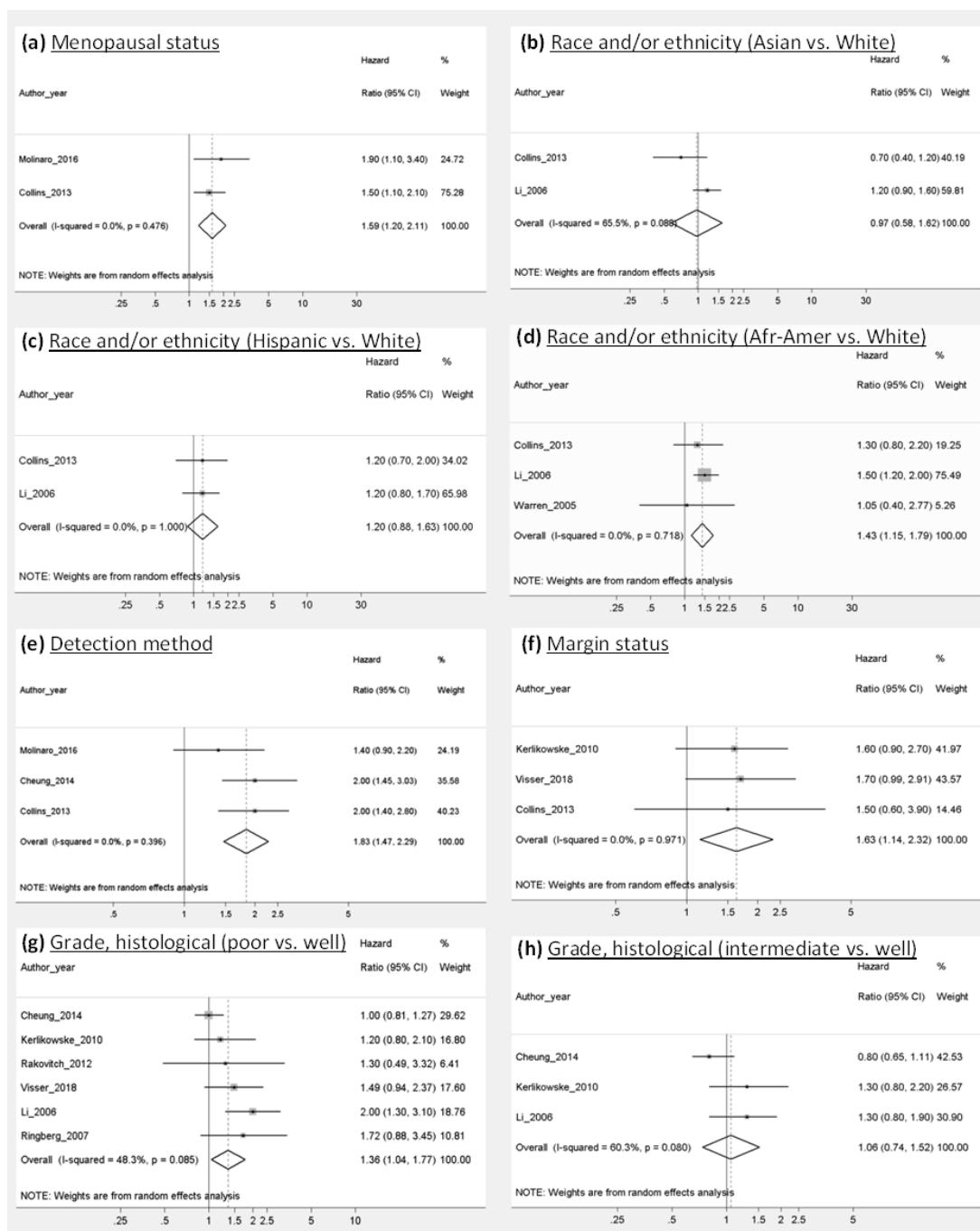


Figure S2. (continues on next page)

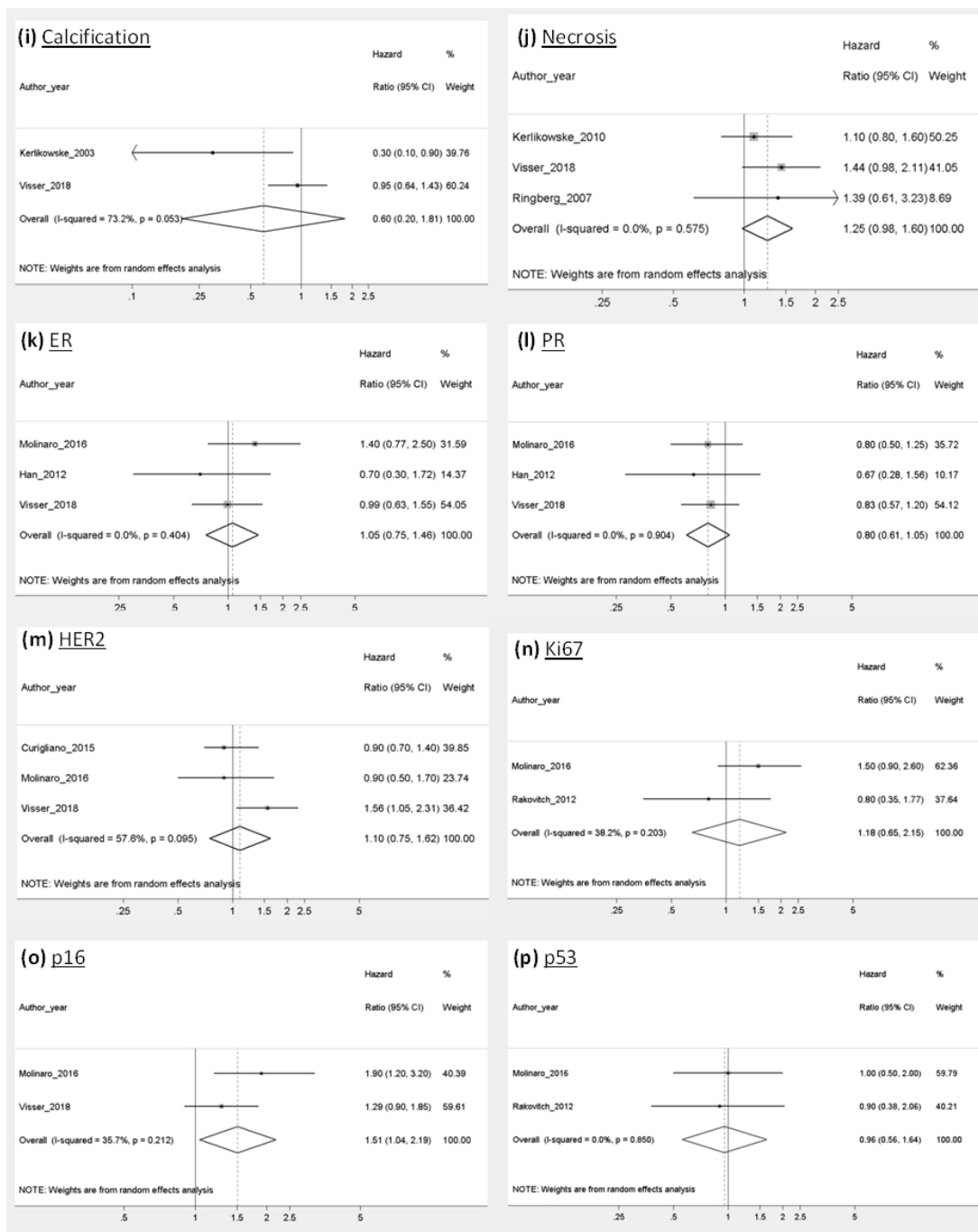


Figure S2. (continues on next page)

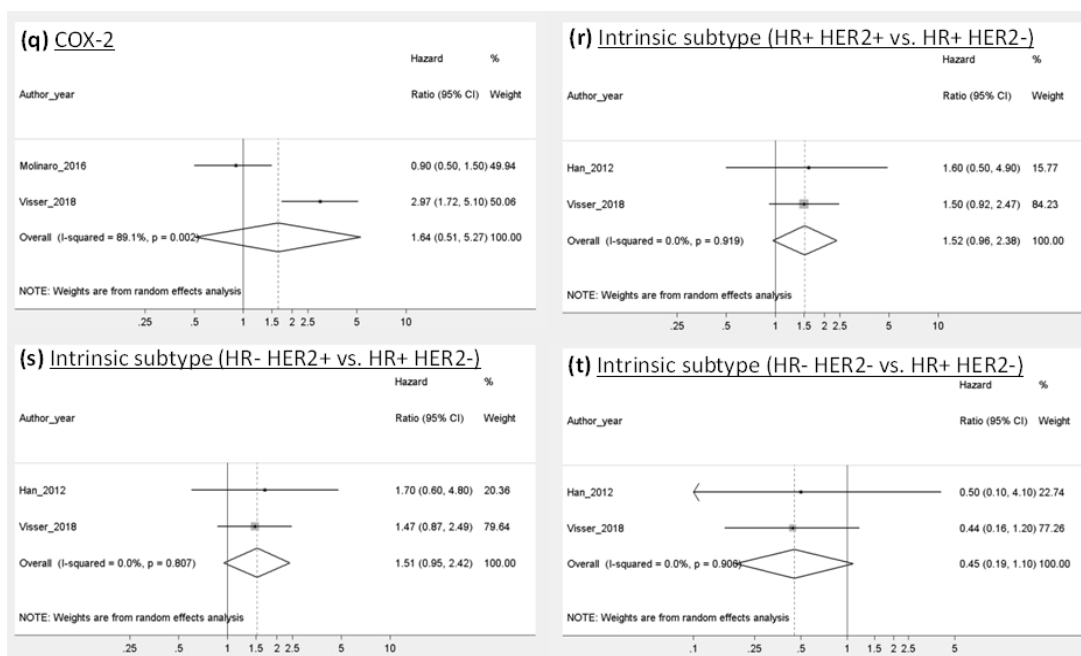


Figure S2. Forest plots of the meta-analyses performed for the prognostic factors and their association with subsequent invasive breast cancer that were reported by more than 1 high quality study: **(a)** menopausal status (pre- vs. post-); **(b)** Race and/or ethnicity (Asian vs. non-Hispanic White); **(c)** Race and/or ethnicity (Hispanic vs. non-Hispanic White); **(d)** Race and/or ethnicity (African-American vs. non-Hispanic White); **(e)** detection method (palpation vs. mammography); **(f)** margin status (involved vs. clear); **(g)** histological grade (poor vs. well differentiated); **(h)** histological grade (intermediate vs. well differentiated); **(i)** calcification (present vs. absent); **(j)** necrosis (present vs. absent); **(k)** ER (positive vs. negative); **(l)** PR (positive vs. negative); **(m)** HER2 (positive vs. negative); **(n)** Ki67 (positive vs. negative); **(o)** p16 (high vs. low); **(p)** p53 (positive vs. negative); **(q)** COX-2 (positive vs. negative); **(r)** intrinsic subtype (HR+ HER2+ vs. HR+ HER2-); **(s)** intrinsic subtype (HR- HER2+ vs. HR+ HER2-); **(t)** intrinsic subtype (HR- HER2- vs. HR+ HER2-); CI, confidence interval.

Translational relevance

There is increasing concern about the current overtreatment of ductal carcinoma in situ (DCIS). Although numerous prognostic markers for DCIS have been reported, none have shown to be of value for clinical implementation. One of the main reasons for this is the frequently introduced bias caused by the lack of sufficiently large patient cohorts. In the context of a nation-wide cohort, we performed a nested case-control study including patients treated by breast conserving surgery alone with long-term follow-up. We found a 4-fold higher prevalence of subsequent ipsilateral invasive breast cancer (iIBC) for women diagnosed with HER2⁺/COX-2^{high} DCIS as compared to women with HER2⁻/COX-2^{low} DCIS lesions. Furthermore, patients with COX-2^{low} DCIS were at lowest risk of iIBC as their risk was comparable to the general population. These prognostic markers are excellent candidates for validation and, ultimately, use in personalized patient risk stratification.

Lindy L. Visser

Lotte E. Elshof

Michael Schaapveld

Koen K. van de Vijver

Emma J. Groen

Mathilde M. Almekinders

Carolien Bierman

Flora E. van Leeuwen

Emiel J. Rutgers

Marjanka K. Schmidt

Esther H. Lips

Jelle Wesseling

CHAPTER 3

Clinicopathological risk factors for an invasive breast cancer recurrence after ductal carcinoma in situ - A nested case-control study

Abstract

Purpose: Ductal carcinoma in situ (DCIS) is treated to prevent progression to invasive breast cancer. Yet, most lesions will never progress, implying that overtreatment exists. Therefore, we aimed to identify factors distinguishing harmless from potentially hazardous DCIS using a nested case-control study.

Experimental Design: We conducted a case-control study nested in a population-based cohort of DCIS patients treated with breast conserving surgery (BCS) alone ($n = 2,658$) between 1989-2005. We compared clinical, pathological, and IHC DCIS characteristics of 200 women who subsequently developed ipsilateral invasive breast cancer (iIBC; cases) and 474 women who did not (controls), in a matched setting. Median follow-up time was 12.0 years (interquartile range 9.0-15.3). Conditional logistic regression models, were used to assess associations of various factors with subsequent iIBC risk after primary DCIS.

Results: High COX-2 protein expression showed the strongest association with subsequent iIBC (OR = 2.97, 95% confidence interval [95% CI] 1.72-5.10). In addition, HER2 overexpression (OR = 1.56, 95% CI 1.05-2.31) and presence of periductal fibrosis (OR = 1.44, 95% CI 1.01-2.06) were associated with subsequent iIBC risk. Patients with HER2⁺/COX-2^{high} DCIS had a 4-fold higher risk of subsequent iIBC (vs. HER2⁻/COX-2^{low} DCIS), and an estimated 22.8% cumulative risk of developing subsequent iIBC at 15 years.

Conclusions: With this unbiased study design and representative group of DCIS patients treated by BCS alone, COX-2, HER2, and periductal fibrosis were revealed as promising markers predicting progression of DCIS into iIBC. Validation will be done in independent data sets. Ultimately, this will aid individual risk stratification of women with primary DCIS.

Introduction

Ductal carcinoma *in situ* (DCIS) is a potential precursor of invasive breast cancer (IBC). It is characterized by proliferation of ductal epithelial cells confined within the ductal-lobular system. Most women (80-85%) are diagnosed with DCIS by screening mammography in which breast abnormalities are found without the women having clinical symptoms.¹ In the western world, the incidence of DCIS has increased almost 6-fold with the introduction of population-based breast cancer screening, and accounts for about 20-30% of all newly diagnosed breast neoplasms.²⁻⁷

Although DCIS is not life threatening, it does increase a woman's risk of developing IBC later in life, which subsequently could lead to a breast cancer-specific death.⁸ However, we are currently unable to distinguish DCIS lesions that will progress to IBC from those that will not, since there is only limited information on the long-term natural history of DCIS.⁹ As a consequence, almost all DCIS is treated by mastectomy or breast conserving surgery (BCS) with or without radiotherapy. This is done under the assumption that this will prevent IBCs and subsequently breast cancer-specific deaths, despite the fact that breast cancer-specific mortality after DCIS is uncommon: <2%.^{10,11} On top of that, the long-term benefit of treatment of asymptomatic DCIS that may or may not progress to IBC is difficult to quantify.¹² As a result, screening programs are nowadays criticized for being associated with overdiagnosis and overtreatment.^{13,14}

Distinguishing, at diagnosis, DCIS that might cause life-threatening disease from indolent DCIS is therefore of great importance. A multitude of studies have tried to find markers that could predict local recurrence or progression of DCIS.¹⁵ In a few studies, investigators showed that various histopathologic characteristics of DCIS, such as lesion size, marginal status, histologic grade, architectural patterns, and presence of necrosis were associated with recurrence.^{16,17} However, these studies did not discriminate between invasive and *in situ* recurrences. Furthermore, due to limited patient numbers and lack of validation studies, none of the markers studied to date show sufficiently strong evidence for an association with subsequent ipsilateral IBC (iIBC).

The primary objective of this study was to identify clinical and histologic characteristics of the initial DCIS lesion that are associated with the development of subsequent iIBC. Here, we report the results of our case-control study nested within a large nation-wide population-based cohort of Dutch women with DCIS treated by BCS alone between 1989 and 2005.

Patients and Methods

Study population and design

The study population has been previously described.¹⁸ In brief, we used a nation-wide population-based patient cohort derived from the Netherlands Cancer Registry (NCR), in

which we included all women diagnosed with primary DCIS, and treated with BCS alone within the Netherlands from January 1, 1989 to December 31, 2004. Patients did not receive tamoxifen or other anti-hormonal adjuvant treatment. According to the Dutch guidelines, patients diagnosed and treated for DCIS were followed by undergoing annual mammograms for at least 5 years. If no recurrence occurred, women could participate in population-based screening again, if applicable regarding age group. Patients with adjacent invasive disease or a prior cancer diagnosis except for nonmelanoma skin cancer were not included. This resulted in 2,658 eligible female participants.

Data provided by the NCR included information on age at and date of diagnosis, histology and treatment for DCIS, and any subsequent IBCs. Follow-up for subsequent iIBC and vital status were complete until at least January 1, 2011. The median follow-up was 12.0 years (interquartile range, 9.0-15.3). 374 of the 2,658 women developed subsequent iIBC, as first invasive cancer, after a primary diagnosis of DCIS.¹⁸ At the start of this present study, the first 316 women with a subsequent iIBC were identified and the remaining 58 cases were identified when data collection of this study was completed. These 316 women were included in this study and were considered “cases”. Controls were matched to cases based on age in years at DCIS diagnosis (exact) using a variable matching ratio. Controls had to have remained free from ipsilateral and contralateral IBC for at least as long as the initial DCIS diagnosis to iIBC development of the case they were matched to. Controls were selected with replacement, so some individuals were a control for more than one case.

Cases and controls originated from 58 hospitals within the Netherlands. We could not obtain FFPE tissue blocks from some hospitals, either because tissue blocks were unavailable or because the hospital refused to provide tissue for research (61 cases and 388 controls; Supplementary Table S1). Furthermore, some patients were excluded during pathology review because: no DCIS component was found, a (micro)invasive component or LCIS was present, or because the specimen was not assessable (53 cases and 173 controls). Finally, some patients were excluded because no matched case or control was available for case-control sets they belonged to (2 cases and 268 controls).

Together with the tissue blocks, pathology reports were retrieved from the participating hospitals. Pathology reports were reviewed for measurements of lesion size and margin status. Notable, data on lesion size was often not routinely described for DCIS in these old retrospective series.

We categorized year of DCIS diagnosis into two time periods: 1989-1998 (screening implementation phase) and 1999-2004 (full nationwide coverage phase). Clinical presentation of DCIS was subdivided into screen-detected (mean age 59 years; range 50-74) and non-screening-related (mean age 54 years; range 30-89). This information was available for 91% of the included patient group.

The study was approved by the review boards of the NCR (request K12.281; 03-01-2013) and PALGA (LZV990; 16-04-2013). The secondary use of tissue and data under

an opt-out regime in this study is conform Dutch regulations and the Code of Conduct of Federa-COREON.¹⁹

Pathology review

New haematoxylin and eosin-stained whole slides were prepared for all tissue specimens and subsequently histopathologic re-examined by three consultant breast pathologists (J. Wesseling, E.J. Groen, and K. van de Vijver). Slides were assessed on morphological characteristics, including DCIS architecture and nuclear grade, the presence of calcifications and necrosis, and microenvironmental characteristics like stromal features and the presence of lymphocytes. This was done blinded of case or control status. In addition, for every patient a representative part of the lesion was selected for further evaluation. Clinical characteristics of patients in- and excluded in this study are presented in Supplementary Table S2. Pathology review data was available for 200 case-control sets, including at least one control per case, resulting in a case-control series of 200 cases and 474 controls, which was representative of the original case-control selection (Supplementary Table S3).

IHC assessment

IHC staining was used to identify DCIS phenotypes using slides from FFPE tissue blocks of 185 DCIS cases and 420 DCIS controls. Some tissue specimens were excluded because insufficient tissue material was available for IHC (12 cases and 26 controls; Supplementary Table 1). DCIS lesions were scored for estrogen receptor (ER) and progesterone receptor (PR) status, overexpression of human epidermal growth factor 2 (HER2), Ki67, and expression of the tumor suppressor proteins p16 and p53, and cyclooxygenase 2 (COX-2). These markers were selected because the antigens have been associated with subsequent IBC after DCIS, based on multivariable analyses, and these results had been reported previously in at least two papers.^{15,20} Moreover, the ability to perform good quality IHC on FFPE material was decisive. Details about the used antibodies, IHC staining procedure, and scoring criteria can be found in Supplementary Materials and Methods. All antibodies used in this study were previously tested in our laboratory using normal tissue and tumor samples known to contain the antigens.

Statistical Analysis

Logistic regression models, conditional on matched sets, were used to assess associations of various clinical and histopathologic characteristics with subsequent iIBC risk after primary DCIS. Wald-based 95% confidence intervals (95% CI) and *P*-values are reported for overall effect for factors with more than two categories.

Variables were selected for inclusion in multivariable models based on a *p*-value ≤ 0.1 in univariate analyses. Due to the amount of missing data margin status and lesion size were excluded from the multivariable models. In addition, since histologic grade was correlated (*r*

> 0.4) with necrosis and periductal fibrosis, this variable was excluded from the multivariable models. Subsequently, the likelihood ratio (LR) chi-square was used to identify the models with the strongest association with subsequent iIBC development.

Approximate cumulative incidence of subsequent iIBC by HER2 and COX-2 status was estimated using the iIBC ORs for HER2 and COX-2 status and cumulative risk of iIBC in the entire cohort. Death due to causes other than breast cancer was considered as a competing risk in this analysis. The expected cumulative incidence of breast cancer for our study population was derived from age-specific breast cancer incidence and all-cause mortality rates in the Dutch female population using the Hakulinen method.²¹

All statistical analyses were performed using Stata/SE (version 13.1, statacorps, Texas). *P* values ≤ 0.05 were considered statistically significant.

Results

Patient characteristics

Clinical characteristics of cases and controls were comparable (Table 1): For both cases and controls median age was 57 years, main period of DCIS diagnosis was 1989-1998, and around 50% of DCIS was screen-detected. For cases, the median time to iIBC was 6.2 years (range 0.5-19.2). Ninety-five percent of all subsequent iIBC lesions recurred at or near the side of the DCIS excision (data not shown). Furthermore, of 141 cases (71%), we were able to assess the ER status of the matched subsequent IBC, which showed an agreement of 89% between the primary DCIS and matched IBC (data not shown).

Univariate results of characteristics associated with subsequent iIBC

The presence of periductal fibrosis was associated with increased risk of subsequent iIBC (OR = 1.44, 95% CI 1.01-2.06) compared to women who did not develop iIBC (Table 2). Furthermore, there was a trend towards a larger lesion size ($P = 0.08$) and more frequent positive resection margins ($P = 0.06$) among cases as compared to controls. However, it should be stressed that around 65% of data on lesion size and around 15% of margin data is missing within our case-control series. There was a trend for necrosis ($P = 0.06$), periductal lymphocytes ($P = 0.13$), and high histologic grade ($P = 0.08$) to be more often present in DCIS lesions of women who subsequently developed iIBC. DCIS architecture, calcifications, and periductal lymphocytes were not associated with subsequent invasive disease. The IHC markers HER2 (OR = 1.56, 95% CI 1.05-2.31) and COX-2 (OR = 2.97, 95% CI 1.72-5.10) were associated with subsequent iIBC risk, but ER, PR, p16, and p53 expression and immunohistochemical subtype were not associated with subsequent iIBC risk (Table 3). Ki67 was excluded from the analysis because of unreliable staining results.

Table 1. Clinical characteristics of female primary DCIS patients treated with BCS alone, who subsequently did (DCIS cases) or did not (DCIS controls) develop subsequent iIBC

Characteristics	DCIS cases (n = 200)		DCIS controls (n = 474)	
	n	(%)	n	(%)
Age at DCIS diagnosis (years)				
<40	14	(7.0)	30	(6.3)
40-49	27	(13.5)	60	(12.7)
50-59	79	(39.5)	204	(43.0)
60-69	55	(27.5)	125	(26.4)
70-79	19	(9.5)	41	(8.6)
≥80	6	(3.0)	14	(3.0)
Year of DCIS diagnosis, mean (range)	1996	(1989 - 2004)	1997	(1989 - 2004)
Period of DCIS diagnosis *				
1989-1998 (screening implementation phase)	147	(73.5)	335	(70.7)
1999-2004 (full nationwide coverage)	53	(26.5)	139	(29.3)
Clinical presentation of DCIS				
Screen-detected	96	(48.0)	245	(51.7)
Non-screening-related	89	(44.5)	184	(38.8)
Unknown	15	(7.5)	45	(9.5)
Time to iIBC, mean in years (range)	6.2	(0.5 - 19.2)	-	-

NOTE: Controls were matched to cases on the basis of age at diagnosis (exact), using a variable matching ratio, and followed at least as long as the case they were matched to, by conditional logistic regression analysis.

iIBC: Ipsilateral invasive breast cancer.

*: Based on the gradual implementation of the national breast cancer screening program in the Netherlands for women >50 years of age, we divided year of DCIS diagnosis into two time periods: 1989-1998, which was the period of implementation of the national mammographic screening program within the Netherlands; and within the period of 1999-2004 the screening program was fully implemented.

Multivariable results of characteristics independently associated with subsequent iIBC

COX-2 was also significantly associated with the risk of subsequent iIBC in multivariable analyses (Table 4 and Supplementary Table S4). Subsequent invasive disease was significantly associated with high COX-2 expression in combination with: (1) overexpression of HER2 (LR $\chi^2 = 6.47$; OR = 3.98); (2) the presence of periductal fibrosis (LR $\chi^2 = 6.34$; OR = 4.87); or (3) the presence of necrosis (LR $\chi^2 = 5.18$; OR = 5.76). Combination of periductal fibrosis, HER2 and COX-2 was also significantly associated with subsequent iIBC, although achieving a lower LR chi-square ratio (LR $\chi^2 = 4.98$; OR = 4.63; Table 4). When combining COX-2, HER2, periductal fibrosis, and necrosis, the addition of necrosis deteriorated the performance of the model, since necrosis is positively correlated with HER2 (data not shown). Eighty-seven percent of DCIS lesions associated with subsequent iIBC (DCIS cases) showed high expression of COX-2. Of this subset of COX-2^{high} DCIS lesions, 34% was HER2 positive, and in 37% periductal fibrosis was present (Table 2 and 3). HER2 overexpression was most frequently accompanied with high COX-2 expression. In contrast, high COX-2 expression was independent of HER2 overexpression. The cumulative risk of subsequent iIBC of HER2⁺/COX-2^{high} DCIS was almost 4 times higher as compared to the

Table 2. Univariate results of histopathologic characteristics associated with subsequent iIBC

Characteristics	DCIS cases (n = 200)		DCIS controls (n = 474)		OR (95% CI) ^a	P ^b
	n	(%)	n	(%)		
Lesion size, millimeter, mean (range)	13	(2-30)	10	(1-70)	1.03 (0.98-1.07)	0.25
Lesion size	200		474			
≤10 mm	29	(14.5)	122	(25.7)	1.00 (reference)	
11-20 mm	27	(13.5)	41	(8.6)	2.04 (0.95-4.34)	
>20 mm	6	(3.0)	8	(1.7)	3.35 (0.78-14.33)	0.08
Unknown	138	(69.0)	303	(63.9)		
Margin status						
Free	136	(68.0)	376	(79.3)	1.00 (reference)	
Not free	28	(14.0)	47	(9.9)	1.70 (0.99-2.91)	0.06
Unknown	36	(18.0)	51	(10.8)		
Growth pattern						
Solid	76	(38.0)	150	(31.6)	1.00 (reference)	
Cribriform	14	(7.0)	33	(7.0)	0.89 (0.45-1.77)	
(Micro)papillary	9	(4.5)	13	(2.7)	1.24 (0.45-3.42)	
Clinging	2	(1.0)	8	(1.7)	0.62 (0.13-3.04)	
Mixed	99	(49.5)	270	(57.0)	0.76 (0.53-1.07)	0.52
Dominant growth pattern						
Solid	119	(59.5)	275	(58.0)	1.00 (reference)	
Cribriform	45	(22.5)	110	(23.2)	0.93 (0.61-1.41)	
(Micro)papillary	24	(12.0)	54	(11.4)	1.02 (0.59-1.78)	
Clinging	12	(6.0)	35	(7.4)	0.83 (0.42-1.66)	0.95
Histologic grade						
Low (grade 1)	29	(14.5)	96	(20.3)	1.00 (reference)	
High (grade 2 and 3)	171	(85.5)	378	(79.7)	1.49 (0.94-2.37)	0.08
Necrosis						
Absent	45	(22.5)	141	(29.7)	1.00 (reference)	
Present	155	(77.5)	333	(70.3)	1.44 (0.98-2.11)	0.06
Microcalcification						
Absent	51	(25.5)	115	(24.3)	1.00 (reference)	
Present	149	(74.5)	358	(75.5)	0.95 (0.64-1.43)	0.82
N/A	0	(0.0)	1	(0.2)		
Periductal fibrosis						
Absent	124	(62.0)	336	(70.9)	1.00 (reference)	
Present	76	(38.0)	137	(28.9)	1.44 (1.01-2.06)	<0.05
N/A	0	(0.0)	1	(0.2)		
Periductal lymphocytes						
Absent	136	(68.0)	353	(74.5)	1.00 (reference)	
Present	64	(32.0)	120	(25.3)	1.33 (0.92-1.92)	0.13
N/A	0	(0.0)	1	(0.2)		

Abbreviations: HR: hormone receptors, ER and PR, where HR+ are ER+ and/or PR+ lesions, and HR- are ER- and PR- lesions. N/A: Not assessable (N/As and unknowns were not included in the analysis).

^a: Comparisons between DCIS cases and DCIS controls were made by univariate conditional logistic regression in which matching was taken into account; ^b: For variables with >2 categories the P-value for overall effect was calculated by Wald-test, for variables with only 2 categories the prob >chi2 was used.

Table 3. Univariate results of IHC markers associated with subsequent iIBC

Characteristics	DCIS cases (n = 185)		DCIS controls (n = 420)		OR (95% CI) ^a	P ^b
	n	(%)	n	(%)		
ER						
Negative	35	(18.9)	79	(18.8)	1.00 (reference)	0.95
Positive	149	(80.5)	341	(81.2)	0.99 (0.63-1.55)	
N/A	1	(0.5)	0	(0.0)		
PR						
Negative	74	(40.0)	149	(35.5)	1.00 (reference)	0.31
Positive	108	(58.4)	264	(62.9)	0.83 (0.57-1.20)	
N/A	3	(1.6)	7	(1.7)		
HER2						
Negative	120	(64.9)	310	(73.8)	1.00 (reference)	0.03
Positive	62	(33.5)	104	(24.8)	1.56 (1.05-2.31)	
N/A	3	(1.6)	6	(1.4)		
Subtypes						
HR+ HER2-	115	(62.2)	284	(67.6)	1.00 (reference)	0.06
HR+ HER2+	33	(17.8)	54	(12.9)	1.50 (0.92-2.47)	
HR- HER2+	29	(15.7)	50	(11.9)	1.47 (0.87-2.49)	
HR- HER2-	5	(2.7)	26	(6.2)	0.44 (0.16-1.20)	
N/A	3	(1.6)	6	(1.4)		
p16						
Low	90	(48.6)	228	(54.3)	1.00 (reference)	0.16
High	93	(50.3)	187	(44.5)	1.29 (0.90-1.85)	
N/A	2	(1.1)	5	(1.2)		
p53						
<30% positive cells	100	(54.1)	240	(57.1)	1.00 (reference)	0.13
30-70% positive cells	40	(21.6)	61	(14.5)	1.67 (1.01-2.77)	
>70% positive cells	25	(13.5)	58	(13.8)	1.08 (0.62-1.87)	
Negative	17	(9.2)	58	(13.8)	0.78 (0.43-1.40)	
N/A	3	(1.6)	3	(0.7)		
COX-2						
Low	19	(10.3)	106	(25.2)	1.00 (reference)	<0.001
High	161	(87.0)	306	(72.9)	2.97 (1.72-5.10)	
N/A	5	(2.7)	8	(1.9)		

Abbreviations: N/A: Not assessable (N/As were not included in the analysis).

^a: Comparisons between DCIS cases and DCIS controls were made by univariate conditional logistic regression; ^b: For variables with >2 categories the P-value for overall effect was calculated by Wald-test, for variables with only 2 categories the prob >chi2 was used.

risk for HER2⁻/COX-2^{low} DCIS lesions (Table 4). Furthermore, analysis of IHC data of 141 DCIS and matched IBC pairs showed that of the 43 patients with HER2⁺/COX-2^{high} DCIS that subsequently developed an iIBC, 35% developed an ER negative invasive recurrence (Table 5).

In our study group, the estimated overall 10-year and 15-year cumulative incidence of iIBC were 10.9% and 13.8%, respectively. For patients with HER2⁻/COX-2^{low} DCIS, the estimated 10-year and 15-year cumulative incidence of iIBC were 4.8% and 6.0%, respectively; for patients with HER2⁺/COX-2^{low} DCIS, 4.5% and 5.6%; for patient with HER2⁻/COX-2^{high} DCIS, 11.3% and 14.3%; and for patients with HER2⁺/COX-2^{high} DCIS, 18.1% and 22.8%, respectively (Figure 1). Within our study group, 29.7% of cases and 18.3% of controls had this unfavorable DCIS subtype of HER2⁺/COX-2^{high} DCIS.

The positive predictive value of the HER2⁺/COX-2^{high} DCIS subtype is 42% (sensitivity: 31%; specificity: 81%), indicating that more than half of the HER2⁺/COX-2^{high} DCIS lesions are not associated with invasive recurrence and thus are false positives. The strength of this marker combination is most likely in its negative predictive value (NPV = 73%), indicating that the risk of subsequent iIBC after DCIS is low in the non-HER2⁺/COX-2^{high} subgroup.

Table 4. Multivariable results of histopathologic characteristics and IHC markers independently associated with subsequent invasive disease

	DCIS cases		DCIS controls		OR (95% CI) ^a	P ^b
	n	(%)	n	(%)		
Periductal fibrosis/Necrosis						
Absent/Absent	34	(17.0)	127	(26.8)	1.00 (reference)	
Present/Absent	11	(5.5)	14	(3.0)	2.75 (1.12-6.75)	
Absent/Present	90	(45.0)	209	(44.1)	1.59 (1.01-2.49)	
Present/Present	65	(32.5)	123	(25.9)	1.88 (1.16-3.07)	0.04
N/A	0	(0.0)	1	(0.2)		
HER2/COX-2						
Negative/Low	14	(7.6)	77	(18.3)	1.00 (reference)	
Positive/Low	5	(2.7)	26	(6.2)	0.94 (0.30-2.95)	
Negative/High	105	(56.8)	227	(54.0)	2.44 (1.30-4.59)	
Positive/High	55	(29.7)	77	(18.3)	3.98 (2.01-7.91)	<0.001
N/A	6	(3.2)	13	(3.1)		
Periductal fibrosis/HER2/ COX-2						
All other groupings	19	(10.3)	103	(24.5)	1.00 (reference)	
Negative/Negative/High	79	(42.7)	174	(41.4)	2.54 (1.42-4.54)	
Positive/Negative/High	26	(14.1)	52	(12.4)	2.45 (1.22-4.92)	
Negative/Positive/High	23	(12.4)	37	(8.8)	3.51 (1.69-7.29)	
Positive/Positive/High	32	(17.3)	40	(9.5)	4.63 (2.26-9.50)	<0.001
N/A	6	(3.2)	14	(3.3)		

Abbreviations: N/A: Not assessable (N/As were not included in the analysis).

^a: Comparisons between DCIS cases and DCIS controls were made by multivariable conditional logistic regression; ^b: P-values for overall effect were calculated by Wald-test.

LR chi2 corrected for degrees of freedom were 2.85 for periductal fibrosis/necrosis, 6.47 for HER2/COX-2, and 4.98 for periductal fibrosis/HER2/COX-2.

Discussion

In this study we identified promising risk factors for progression of DCIS into iIBC, by conducting a nested case-control study comparing women who did and did not develop invasive disease after primary DCIS. To avoid confounding by radiation effects, we analyzed all women treated by BCS alone, within our well-characterized nation-wide population-based cohort of women diagnosed with primary DCIS between 1989 and 2005 in the Netherlands with a median follow-up time of 12.0 years. The large size of our series, the design applied, the comprehensive data, and the long term-follow-up are essential to overcome limitations due to bias and lack of power of small sample series with a relatively short follow-up.

We found that initial DCIS lesions with HER2⁺/COX-2^{high} expression were associated with increased risk of subsequent iIBC, which are often ER negative (35%). This is an important finding, as ER negative tumors have in general a worse prognosis than ER positive breast cancers. Women diagnosed with this unfavorable DCIS subtype had a 4-fold higher risk and an estimated 22.8% cumulative 15-year risk of developing subsequent iIBC. This was higher than the overall cumulative incidence of iIBC in this patient cohort, i.e. 13.8% at 15 years. The estimated cumulative risk for patients with HER2⁻/COX-2^{low} and HER2⁺/COX-2^{low} DCIS was comparable to the risk for the general population, i.e. 4-6% at 15 years. The positive and negative predictive value of the HER2/COX-2 marker combination is 42% and 73%, respectively. For clinical purposes the predictive values should definitely be improved.

Extensive granular cytoplasmic expression of COX-2 was the marker which had the strongest association with subsequent iIBC both in univariate and multivariable analysis. This is in line with one previous study that showed an association between COX-2 and subsequent invasive breast cancer in univariate analysis.²² We also found an association of subsequent iIBC with HER2 positive primary DCIS, and presence of periductal fibrosis. HER2 overexpression was associated with subsequent iIBC in univariate analysis, which was also found in two previous studies.^{23,24} Interestingly, results from our multivariable analysis showed that HER2 overexpression is not predictive for subsequent iIBC in the absence of high COX-2 expression. This indicates that the prognostic value of HER2 overexpression in the risk of subsequent iIBC is probably limited and the increase in risk is primarily driven by COX-2 overexpression. Another prognostic factor we found was periductal fibrosis. This microenvironmental factor was associated with subsequent iIBC in both univariate and multivariable analysis and is supported by a previous study.²⁴ So far, diagnosis of breast disease has been limited to the morphological interpretation of epithelial cells and the assessment of epithelial tissue architecture, in which the stromal compartment is largely ignored. The role of periductal fibrosis described here underline the importance of assessment of the DCIS stromal compartment.

There is biological support for a role of COX-2 in invasive breast cancer recurrence. COX-2 is a cytoplasmic enzyme involved in prostaglandin synthesis. It is induced rapidly in

response to growth factors, tumor promoters, hormones, and cytokines.²⁵ It has been shown that overexpression of COX-2 leads to an increased prostaglandin E2 (PGE2) level which can potentially affect most of the key processes in cancer development, including proliferation, resistance to apoptosis, angiogenesis, immune suppression and invasion.²⁶ Furthermore, overexpression of COX-2 has been shown to result in p16-mediated cell cycle arrest through the p16/Rb-signaling pathway. When the p16/Rb-signaling pathway is disrupted, cellular proliferation continues, resulting in additional high Ki67 expression in the presence of high p16 and high COX-2 expression.²⁷ It has previously been shown that p16⁺/COX-2⁺/Ki67⁺ DCIS is associated with subsequent iIBC.²⁸ Unfortunately, in our study we were unable to assess p16/COX-2/Ki67 protein expression, since Ki67 suffers from loss of antigenicity over time and is very sensitive to improper formalin fixation.^{29,30}

Table 5. HER2/COX-2 status DCIS related to ER status of subsequent iIBC

	Invasive breast cancer		Total	p
	ER_pos n (%)	ER_neg n (%)		
DCIS				
HER2+/COX-2high	28 (65.1)	15 (34.9)	43	<0.001
All other groupings	91 (92.9)	7 (7.1)	98	
Total	119 (84.4)	22 (15.6)	141	

NOTE: Comparison between DCIS and invasive breast cancer were made by marginal homogeneity test.

Next to the factors described above, a wide range of other prognostic factors have been reported in literature related to recurrent disease, albeit with small effect sizes.^{1,15} In contrast to our study, most studies did not discriminate between invasive and *in situ* recurrences as a primary endpoint.^{31–33} As invasive recurrences may lead to breast cancer mortality, it is of utmost importance to make a distinction between *in situ* and invasive recurrences in risk prediction.

Our study group comprised of DCIS patients diagnosed between 1989 and 2005 and treated by BCS alone. Regarding this time period, we have to keep in mind that treatment strategies and screening techniques for DCIS have evolved over the years, which may have impacted treatment or other care for these patients. The main period of DCIS diagnosis of our patient group was 1989–1998 and about 50% of DCIS was screen-detected. Within the time period 1989–1998, guidelines for DCIS treatment in the Netherlands recommended mastectomy or BCS alone. From 1999, the addition of radiotherapy after BCS was included. Clinical trials have shown that adjuvant radiotherapy reduces the risk of both *in situ* and invasive recurrence with about 50%.^{34,35} Moreover, our group has shown that women diagnosed with DCIS between 1999 and 2004 were less likely to develop iIBC than women diagnosed between 1989 and 1998, regardless of treatment and age.¹⁸ With the introduction of digital mammography, the coverage and sensitivity of screening improved significantly

and led to an increase in the percentage of screen-detected DCIS.³⁶

In our study group, the estimated overall 10-year and 15-year cumulative incidence of iIBC were 10.9% and 13.8%, respectively. This is comparable to the 10-year cumulative incidence reported in two non-randomized prospective studies of women with DCIS treated by BCS alone.^{37,38}

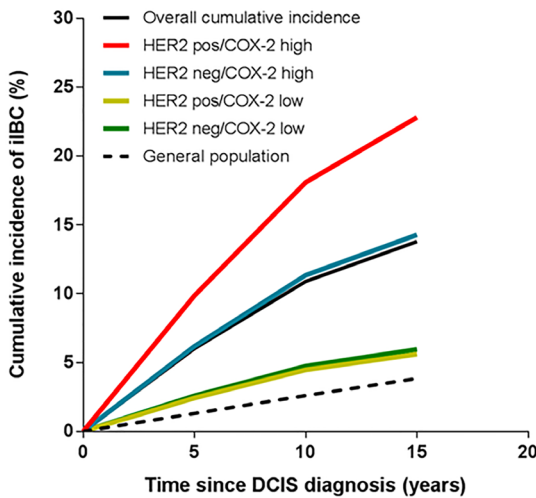


Figure 1. Cumulative incidence per category of HER2/COX-2 status.

Cumulative incidence of iIBC among women with an initial diagnosis of DCIS treated by BCS alone. Approximate cumulative incidence of subsequent iIBC by HER2 and COX-2 status was estimated using the iIBC ORs for HER2 and COX-2 status from the current study and cumulative risk of iIBC and death due to other causes derived from the entire cohort.

HER2/COX-2	10-year	15-year
Overall cumulative incidence (-)	10.9%	13.8%
Positive/High (-)	18.1%	22.8%
Negative/High (-)	11.3%	14.3%
Positive/Low (-)	4.5%	5.6%
Negative/Low (-)	4.8%	6.0%
General population (-)	2.6%	3.9%

Our study has several strengths. First, our study is nested in a large, unique population-based study of women with DCIS treated by BCS alone that provides information on clinical, histopathological and immunohistochemical characteristics, and focusses specifically on subsequent ipsilateral *invasive* breast cancer with a median follow-up time of 12.0 years. This well-annotated nature of our patient series enabled us to prevent a great amount of bias often seen in previous studies. Second, we were able to collect 85% of the requested tissue blocks, from which new whole slides were developed for re-assessment by specialized breast pathologists. Third, our large sample allowed us to assess the combinations of clinical, histopathological, and immunohistochemical characteristics, that were independently associated with subsequent invasive disease, by using multivariable models.

Our study also has some limitations. First, we were only able to assess prognostic factors for a subset of all women included in the case-control study since this depended on the

availability of FFPE tissue blocks in the participating hospitals. Fortunately, this did not cause significant bias, as these samples were missing randomly (hospital participation) and the group for which we could successfully collect the tissue blocks was an excellent representation of the complete case-control selection based on the patients' clinical characteristics. Second, the interpretation of immunohistochemical markers can be challenging because of the heterogeneous expression of certain proteins. Nevertheless, we succeeded in minimizing the inter-observer variability by our well designed scoring method. Third, margin status and size of the DCIS lesion were not known in 15% and 65% of the cases, respectively, as these data were not always routinely described in these old retrospective series. As far as margin status and DCIS lesions size were known, no statistically significant differences were present between the two groups, i.e. cases and controls. Missing data were also equally randomly distributed among these groups (Table 2), and thus was not related to the outcome of interest. Therefore, resulting over- or underrepresentation of some factors is unlikely. Fourth, we did not have data on BRCA status or family history of breast cancer of our patient group. Yet, it has to be taken into account that BRCA status could be a confounding factor in subsequent iIBC development.

Conclusions

In summary, identification of prognostic factors has the potential to improve the clinical management of women diagnosed with DCIS. We found a prognostic role for COX-2, HER2, and periductal fibrosis. In addition, women diagnosed with HER2⁺/COX-2^{high} DCIS and treated by BCS alone had a 4-fold higher prevalence of subsequent iIBC than women with HER2⁻/COX-2^{low} DCIS lesions. Furthermore, HER2⁺/COX-2^{high} DCIS was associated with ER negative invasive recurrences. Our results underline the importance of assessment of the DCIS stromal compartment and protein expression of HER2 and COX-2 to estimate the risk of subsequent invasive disease after a diagnosis of DCIS. Patients with COX-2^{low} DCIS are at lowest risk of iIBC as their risk is comparable to the general population. Our study design and unique retrospective patient series provided us with excellent candidate prognostic markers for use in personalized patient risk stratification. As a next step, these prognostic markers will need to be validated in independent data sets, as a major step to distinguish harmless from potentially hazardous DCIS.

Article information

Funding: This work was supported by Pink Ribbon (grand number 2013.WO29); and A Sisters Hope (grand number 2011.WO19.C88).

Acknowledgements: We thank all patients whose stored residual material was used in this study and their treating physicians, all the collaborating hospitals and especially the pathology departments, and all persons who helped in the data collection and analysis. Specifically, we acknowledge Otto Visser, Annemarie Eeltink and the registration teams of the Netherlands Comprehensive Cancer Organization for the collection of data for the Netherlands Cancer Registry; Lucy Overbeek and PALGA, the nationwide histopathology and cytopathology data network and archive, for providing pathology data; the pathologists of the PALGA-group: Dr. P. Bult, Dr. C.H.M. van Deurzen, Dr. E. Geuken, and Dr. C.M. van Dijk for their cooperation; Frank Nieboer, Danah Kamphuis, and Chaima Loukili for their assistance; and the NKI-AVL Core Facility Molecular Pathology & Biobanking (CFMPB) for their lab support.

References

1. Bartlett JMS, Nofech-Moses S, Rakovitch E. Ductal carcinoma in situ of the breast: can biomarkers improve current management? *Clin Chem*. 2014;60(1):60-67.
2. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst*. 2002;94(20):1546-1554.
3. Netherlands Comprehensive Cancer Organization (IKNL). DCIS incidence trends over time, the Netherlands.
4. Information Services Division (ISD) Scotland. DCIS incidence trends over time, Scotland.
5. Welsh Cancer Intelligence and Surveillance Unit. DCIS incidence trends over time, Wales.
6. Office for National Statistics. DCIS incidence trends over time, UK.
7. Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2014*. National Cancer Institute. Bethesda, MD
8. Lopez-Garcia MA, Geyer FC, Lacroix-Triki M, Marchi  C, Reis-Filho JS. Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*. 2010;57(2):171-192.
9. Erbas B, Provenzano E, Armes J, Gertig D. The natural history of ductal carcinoma in situ of the breast: a review. *Breast Cancer Res Treat*. 2006;97(2):135-144.
10. Narod S a, Iqbal J, Giannakeas V, Sopik V, Sun P. Breast Cancer Mortality After a Diagnosis of Ductal Carcinoma In Situ. *JAMA Oncol*. 2015;1(7):888-896.
11. Elshof LE, Schmidt MK, Rutgers EJT, van Leeuwen FE, Wesseling J, Schaapveld M. Cause-specific Mortality in a Population-based Cohort of 9799 Women Treated for Ductal Carcinoma In Situ. *Ann Surg*. 2018;267(5):952-958.
12. Gierisch JM, Myers ER, Schmit KM, et al. Prioritization of research addressing management strategies for ductal carcinoma in situ. *Ann Intern Med*. 2014;160(7):484-491.
13. L berg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res*. 2015;17(1):63.
14. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102(9):605-613.
15. Lari S a, Kuerer HM. Biological Markers in DCIS and Risk of Breast Recurrence: A Systematic Review. *J Cancer*. 2011;2:232-261.
16. Ringberg A, Nordgren H, Thorstensson S, 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. *Eur J Cancer*. 2007;43(2):291-298.
17. Collins LC, Achacoso N, Haque R, et al. Risk factors for non-invasive and invasive local recurrence in patients with ductal carcinoma in situ. *Breast Cancer Res Treat*. 2013;139(2):453-460.
18. 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 Res Treat*. 2016;159:553-63.
19. Federa-COREON. Dutch regulations and the Code of Conduct.
20. Sarode VR, Han JS, Morris DH, Peng Y, Rao R. A Comparative Analysis of Biomarker Expression and Molecular Subtypes of Pure Ductal Carcinoma In Situ and Invasive Breast Carcinoma by Image Analysis: Relationship of the Subtypes with Histologic Grade, Ki67, p53 Overexpression, and DNA Ploidy. *Int J Breast Cancer*. 2011;2011:217060.
21. Hakulinen T. Cancer survival corrected for heterogeneity in patient withdrawal. *Biometrics*. 1982;38(4):933-942.
22. Generali D, Buffa FM, Deb S, et al. COX-2 expression is predictive for early relapse and aromatase inhibitor resistance in patients with ductal carcinoma in situ of the breast, and is a target for treatment. *Br J Cancer*. 2014;111(1):46-54.
23. Nofech-Mozes S, Spayne J, Rakovitch E, et al. Biological Markers Predictive of Invasive Recurrence in DCIS. *Clin Med Oncol*. 2008;2(416):7-18.
24. Van Bockstal M, Lambein K, Gevaert O, et al. Stromal architecture and periductal

- decorin are potential prognostic markers for ipsilateral locoregional recurrence in ductal carcinoma in situ of the breast. *Histopathology*. 2013;63(4):520-533.
25. Singh-Ranger G, Salhab M, Mokbel K. The role of cyclooxygenase-2 in breast cancer: review. *Breast Cancer Res Treat*. 2008;109(2):189-198.
26. Bocca C, Ievolella M, Autelli R, et al. Expression of Cox-2 in human breast cancer cells as a critical determinant of epithelial-to-mesenchymal transition and invasiveness. *Expert Opin Ther Targets*. 2014;18(2):121-135.
27. Gauthier ML, Berman HK, Miller C, et al. Abrogated response to cellular stress identifies DCIS associated with subsequent tumor events and defines basal-like breast tumors. *Cancer Cell*. 2007;12(5):479-491.
28. Kerlikowske K, Molinaro AM, Gauthier ML, et al. Biomarker expression and risk of subsequent tumors after initial ductal carcinoma in situ diagnosis. *J Natl Cancer Inst*. 2010;102(9):627-637.
29. Combs SE, Han G, Mani N, Beruti S, Nerenberg M, Rimm DL. Loss of antigenicity with tissue age in breast cancer. *Lab Invest*. 2016;96(3):264-269.
30. Arima N, Nishimura R, Osako T, et al. The importance of tissue handling of surgically removed breast cancer for an accurate assessment of the Ki-67 index. *J Clin Pathol*. 2016;69(3):255-259.
31. Rudloff U, Jacks LM, Goldberg JI, et al. Nomogram for predicting the risk of local recurrence after breast-conserving surgery for ductal carcinoma in situ. *J Clin Oncol*. 2010;28(23):3762-3769.
32. Benson JR, Jatoi I, Toi M, et al. Treatment of low-risk ductal carcinoma in situ: is nothing better than something? *Lancet Oncol*. 2016;17(10):e442-e451.
33. Wang F, Li H, Tan PH, et al. Validation of a nomogram in the prediction of local recurrence risks after conserving surgery for Asian women with ductal carcinoma in situ of the breast. *Clin Oncol (R Coll Radiol)*. 2014;26(11):684-691.
34. Wapnir IL, Dignam JJ, Fisher B, 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. *J Natl Cancer Inst*. 2011;103(6):478-488.
35. Donker M, Liti?re S, Werutsky G, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma In Situ: 15-year recurrence rates and outcome after a recurrence, from the EORTC 10853 randomized phase III trial. *J Clin Oncol*. 2013;31(32):4054-4059.
36. Karssemeijer N, Bluekens AM, Beijerinck D, et al. Breast Cancer Screening Results 5 Years after Introduction of Digital Mammography in a Population-based Screening Program. *Radiology*. 2009;253(2):353-358.
37. Wong JS, Chen YH, Gadd MA, et al. Eight-year update of a prospective study of wide excision alone for small low- or intermediate-grade ductal carcinoma in situ (DCIS). *Breast Cancer Res Treat*. 2014;143(2):343-350.
38. Solin LJ, Gray R, Hughes LL, et al. Surgical excision without radiation for ductal carcinoma in situ of the breast: 12-year results from the ECOG-ACRIN E5194 study. *J Clin Oncol*. 2015;33(33):3938-3944.

Supplementary materials and methods

Pathology review and assessment of inter-observer variability

A random sample of 50 DCIS specimens were reviewed by all pathologists to assess the inter-observer agreement. The concordance between the pathologists was around 85% with an interclass correlation coefficient of >0.4 for every characteristic assessed. Subsequently, slides from each included patient were reviewed by one of these pathologists, unaware of case or control status.

Immunohistochemical staining procedure

Immunohistochemical staining of ER, PR, HER2, Ki67, p16, p53, and COX-2 was performed using a Benchmark ULTRA autostainer (Ventana Medical Systems, AZ, USA). For each lesion and for each staining, a 3 μ m thick whole slide paraffin section was heated at 75 °C for 28 minutes and deparaffinized in the instrument with EZ prep solution (Ventana Medical Systems). Heat-induced antigen retrieval was carried out using Cell Conditioning 1 (CC1; Ventana Medical Systems) for 64 min at 95 °C. Next, the primary antibody, indicated in Table M1, was applied to the tissue section. Additionally, for the p16 staining signal amplification was performed using the Optiview Amplification kit (Ventana Medical Systems). For the PR staining the slides were additionally incubated with normal antibody diluent (ABB999, Immunologic) to reduce the background signal. Reactions were detected using the UltraView Universal DAB Detection kit (#760-500; Roche) for visualization of ER, PR, and HER2, or the OptiView DAB Detection kit (#760-700; Roche) for visualization of Ki67, p16, p53 and COX-2. Finally, the slides were counterstained with Hematoxylin II and Bluing Reagent (Ventana Medical Systems).

Table M1. Primary antibody sources and dilutions used in this study

Antigen	Clone	Dilution	Manufacturer
ER	SP1	Ready-to-use	Ventana Medical Systems
PR	1E2	Ready-to-use	Ventana Medical Systems
HER2	4B5	Ready-to-use	Ventana Medical Systems
Ki67	MIB1	1/250	DAKO
p16	JC8	1/800	Santa Cruz
p53	DO-7	1/7000	DAKO
COX-2	CX294	1/100	DAKO

ER, estrogen receptor; PR, progesteron receptor; HER2, human epidermal growth factor receptor 2; p16, cyclin-dependent kinase inhibitor 2A; COX-2, cyclooxygenase-2.

Assessment of immunohistochemistry and inter-observer variability

IHC stained slides were scanned using an Aperio AT2 Slide Scanner (Leica Biosystems) and subsequently the digital images were scored by a team of 7 observers, including 5 pathologists; all were blinded of clinical outcome. All 7 observers scored ER, PR, HER2, Ki67, p16 and p53 stains. A random sample of 20 DCIS specimens were scored by all observers to assess the inter-observer agreement. The concordance between the observers was around 97% with an interclass correlation coefficient (ICC) of >0.8 for ER, PR, HER2, and p53, and an ICC of 0.7 for p16. Ki67 was excluded for further scoring because of unreliable staining results. COX-2 stains were scored by two observers (LLV, CB). To assess the inter-observer agreement, a random sample of 20 DCIS specimens were scored by both observers. With this, COX-2 reached a concordance between the two observers of 94% with a k statistic of 0.7. Consequently, the remaining ER, PR, HER2, p16, and p53 stains were distributed over all observers, and the remaining COX-2 stains were distributed over two observers (LLV, CB), in which each patient was scored by one observer.

Immunohistochemical scoring criteria

Representative examples of each immunohistochemical marker and their corresponding staining categories can be found in Figure M1. For ER and PR, the percentage of luminal epithelial cells that showed staining of any intensity was assessed, and were considered positive when at least 10% of the cells nuclei showed staining. Similarly, p53 positive staining was assessed based on the percentage of nuclei that showed moderate to strong staining. p53 protein accumulation was considered when 70% or more nuclei showed moderate to strong staining.

HER2 overexpression was analyzed according to the American Society of Clinical Oncology and College of American Pathologists (ASCO-CAP) 2013 recommendations: HER2-positive staining was considered positive if circumferential membrane staining was complete, intense, and within more than 10% of tumor cells (HER2 3+), or if circumferential membrane staining was incomplete and/or weak/moderate and within more than 10% of tumor cells (HER2 2+) for which the overexpression could be confirmed by HER2 SISH. HER2 was considered negative when incomplete membrane staining was faint/barely visible and within >10% of the tumor cells (HER2 1+) or when no staining was observed (HER2 0).¹³

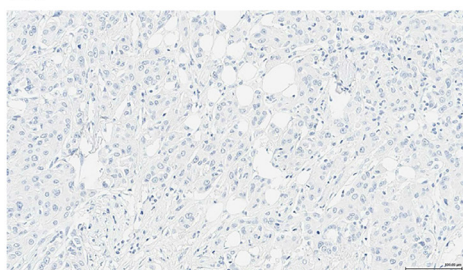
For the assessment of p16 a semi-quantitative approach was used in which H-scores were generated by multiplying the staining intensity (0 = no staining, 1 = weak, 2 = moderate, 3 = intense) by the percentage of positive cells (0-100%). H-scores below or equal to 100 were categorized as score 1, H-scores between 101-200 were categorized as score 2, H-scores between 201-300 were categorized as score 3. Score 1 was considered low p16 expression and score 2 and 3 were considered high p16 expression.

COX-2 expression was evaluated according to the following criteria (adapted from

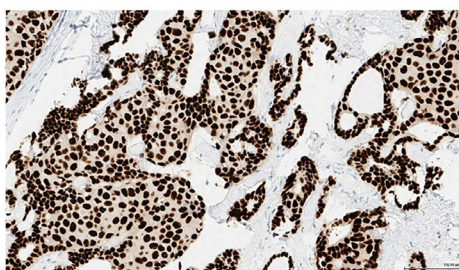
Ristimäki et al.¹⁴): 1 = weak diffuse cytoplasmic staining that may contain moderate to strong granular cytoplasmic staining in less than 10% of tumor cells; 2 = moderate to strong granular cytoplasmic staining in 10-90% of the tumor cells; 3 = moderate to strong granular cytoplasmic staining in over 90% of the tumor cells. Score 2 and 3 are considered high expression of COX-2 and score 1 low COX-2 expression.

Figure M1. Representative examples the immunohistochemical markers and their corresponding staining categories

ER

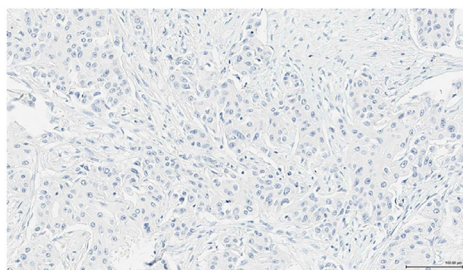


Negative

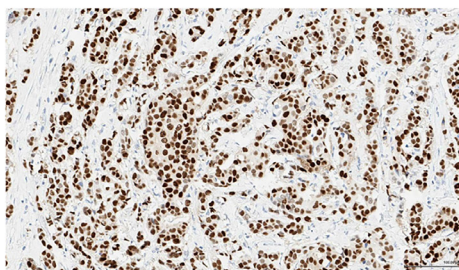


Positive

PR

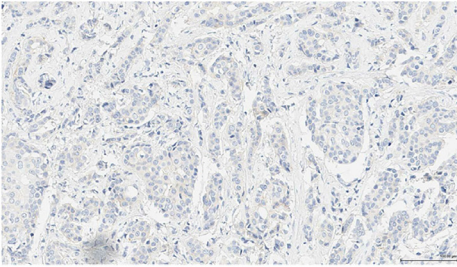


Negative

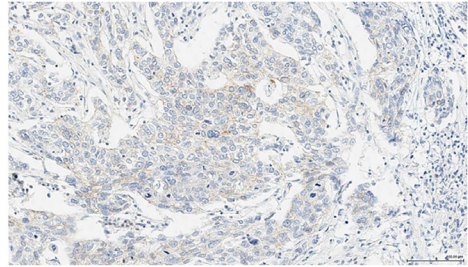


Positive

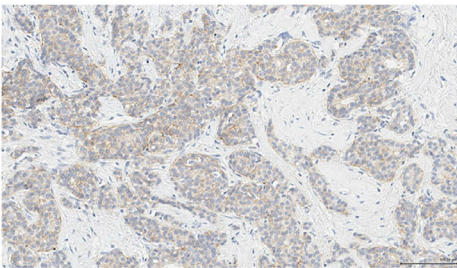
HER2



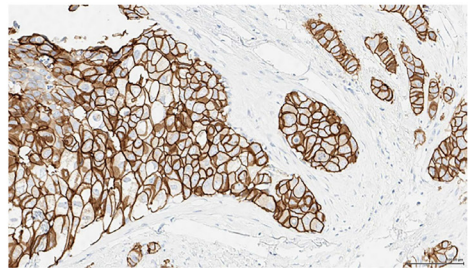
No staining = 0



Weak staining = 1

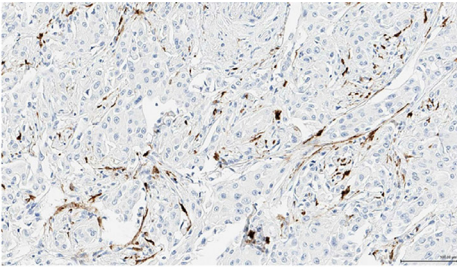


Moderate/ equivocal = 2

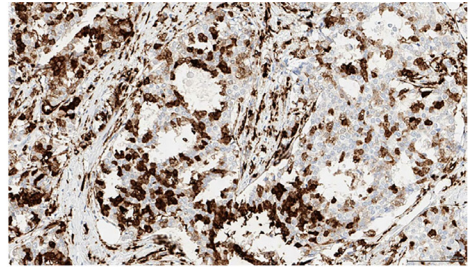


Strong staining = 3

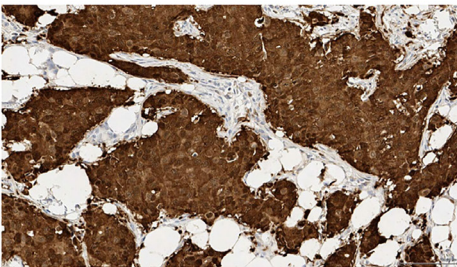
p16



H-score = 1

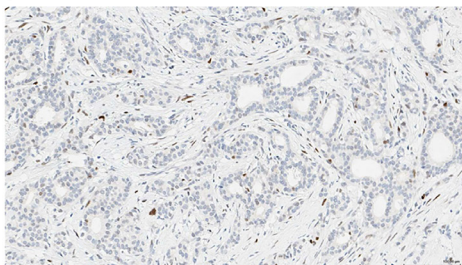


H-score = 2

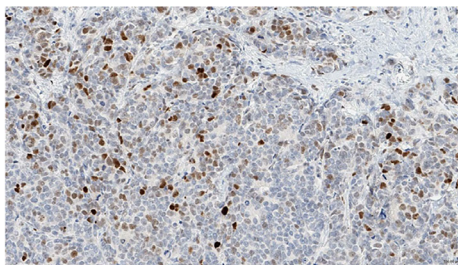


H-score = 3

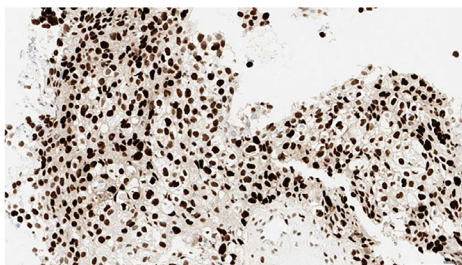
p53



Percentage cells with moderate/strong stained nuclei;
1-30%

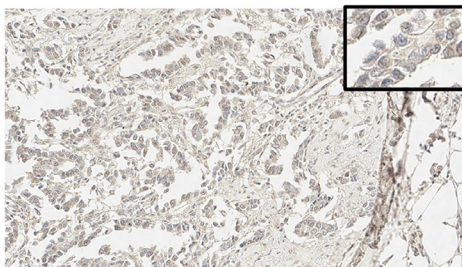


Percentage cells with moderate/strong stained nuclei;
30-70%

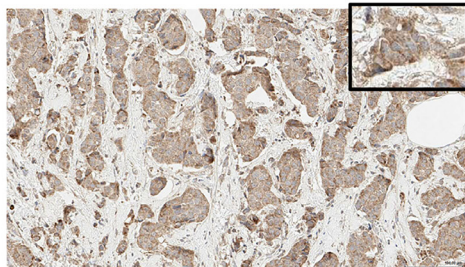


Percentage cells with moderate/strong stained nuclei;
>70%

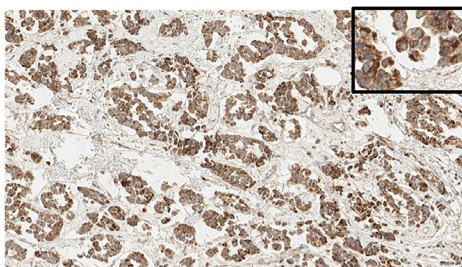
COX-2



1: Weak diffuse cytoplasmic staining that may contain moderate to strong granular cytoplasmic staining in less than 10% of tumor cells



2: Moderate to strong granular cytoplasmic staining in 10-90% of the tumor cells



3: Moderate to strong granular cytoplasmic staining in over 90% of tumor cells

Supplementary tables

Table S1. Distribution of cases and controls who were initially selected for the case-control study by final study status

A. Collected tissue specimens and pathology review

	Number of Cases		Number of Controls	
	No.	%	No.	%
Initially selected	316	100.0	1303	100.0
Excluded	114	36.1	561	43.1
- No tissue blocks available	61	19.3	388	29.8
- No DCIS component found	18	5.7	74	5.7
- (Micro)invasive component present	6	1.9	11	0.8
- LCIS present	9	2.8	31	2.4
- Papillary lesion other than DCIS	14	4.4	34	2.6
- Specimen not assessable	5	1.6	12	0.9
- Other reason	1	0.3	11	0.8
Administrative reason for exclusion	2	0.6	268	20.6
- Selected as case, controls excluded	2	0.6	-	-
- Selected as control, case excluded	-	-	268	20.6
Total excluded	116	36.7	829	63.7
Total included in study	200	63.3	474	36.4
Duplicate controls	-	-	113	23.8
Case as control	-	-	33	7.0
- Once	-	-	28	5.9
- Twice	-	-	5	1.1
Control	-	-	72	15.2
- Once	-	-	65	13.7
- Twice	-	-	6	1.3
- Three times	-	-	1	0.2
Total number of unique individuals in study	200	100.0	361	76.2

Table S1. Distribution of cases and controls who were initially selected for the case-control study by final study status

B. Immunohistochemistry (IHC) assessment				
	Number of Cases		Number of Controls	
	No.	%	No.	%
Initially selected for IHC	200	100.0	474	100.0
Excluded	12	6.0	26	5.5
- Insufficient tissue available	12	6.0	26	5.5
Administrative reason for exclusion	3	1.5	28	5.9
- Selected as case, controls excluded	3	1.5	-	-
- Selected as control, case excluded	-	-	28	5.9
Total excluded	15	7.5	54	11.4
Total included in IHC analysis	185	92.5	420	88.6
Duplicate controls	-	-	97	23.1
Case as control	-	-	27	6.4
- Once	-	-	23	5.5
- Twice	-	-	4	1.0
Control	-	-	61	14.5
- Once	-	-	54	12.9
- Twice	-	-	6	1.4
- Three times	-	-	1	0.2
Total number of unique individuals in IHC analysis	185	100	323	76.9

Table S2. Clinical characteristics of 561 unique patients included in this study and of 669 patients not included in this case-control study.

Characteristics	Included, No. (%)		Not included, No. (%)		P ^a
Overall	561	(100.0)	669	(100.0)	
Patient group					
Cases	200	(35.7)	116	(17.3)	
Controls	361	(64.3)	553	(82.7)	<0.001
Age at DCIS diagnosis, mean (range)	58	(30 - 89)	59	(30 - 89)	0.008
Age at DCIS diagnosis (years)					
<40	33	(5.9)	29	(4.33)	
40-49	74	(13.2)	108	(16.1)	
50-59	225	(40.1)	210	(31.4)	
60-69	160	(28.5)	178	(26.6)	
70-79	52	(9.3)	105	(15.7)	
≥80	17	(3.0)	39	(5.8)	<0.001
Year of DCIS diagnosis, mean (range)	1996	(1989 - 2004)	1997	(1989 - 2004)	0.17
Period of DCIS diagnosis					
1989-1998 (screening implementation phase)	393	(70.1)	441	(65.9)	
1999-2004 (full nationwide coverage)	168	(29.9)	228	(34.1)	0.12
Clinical presentation of DCIS					
Screen-detected	282	(50.3)	273	(40.8)	
Non-screening-related	230	(41.0)	332	(49.6)	0.001
Unknown	49	(8.73)	64	(9.57)	

^a: For continuous variables, the p-value was calculated by unpaired T-test, and for categorical variables the p-value was calculated by chi-square test; Unknown clinical presentation was not included in the analysis.

Table S3. Comparison of the 316 cases and 914 controls (unique) initially selected for the case-control study with the unique cases and controls included in this study

Cases					
Characteristics	Initial selection		Included in study		P ^a
	(n = 316)		(n = 200)		
	n	(%)	n	(%)	
Age at DCIS diagnosis (years)					
<40	16	(5.1)	14	(7.0)	
40-49	49	(15.5)	27	(13.5)	
50-59	109	(34.5)	79	(39.5)	
60-69	85	(26.9)	55	(27.5)	
70-79	41	(13.0)	19	(9.5)	
≥80	16	(5.1)	6	(3.0)	0.48
Year of DCIS diagnosis, mean (range)	1996	(1989 - 2004)	1996	(1989 - 2004)	0.88
Period of DCIS diagnosis					
1989-1998 (implementation phase)	234	(74.1)	147	(73.5)	
1999-2004 (full nationwide coverage)	82	(25.9)	53	(26.5)	0.92
Clinical presentation of DCIS					
Screen-detected	135	(42.7)	96	(48.0)	
Non-screening-related	155	(49.1)	89	(44.5)	0.23
Unknown	26	(8.2)	15	(7.5)	
Controls					
Characteristics	Initial selection		Included in study		P ^a
	(n = 914)		(n = 361)		
	n	(%)	n	(%)	
Age at DCIS diagnosis (years)					
<40	46	(5.0)	19	(5.3)	
40-49	133	(14.6)	47	(13.0)	
50-59	326	(35.7)	146	(40.4)	
60-69	253	(27.7)	105	(29.1)	
70-79	116	(12.7)	33	(9.1)	
≥80	40	(4.4)	11	(3.0)	0.30
Year of DCIS diagnosis, mean (range)	1997	(1989 - 2004)	1997	(1989 - 2004)	0.53
Period of DCIS diagnosis					
1989-1998 (implementation phase)	600	(65.6)	246	(68.1)	
1999-2004 (full nationwide coverage)	314	(34.4)	115	(31.9)	0.41
Clinical presentation of DCIS					
Screen-detected	420	(46.0)	186	(51.5)	
Non-screening-related	407	(44.5)	141	(39.1)	0.07
Unknown	87	(9.5)	34	(9.4)	

^a: For continuous variables, the p-value was calculated by paired T-test, and for categorical variables the p-value was calculated by chi-square test; Unknown clinical presentation was not included in the analysis.

Table S4. Multivariable analysis of histopathological characteristics and immunohistochemical markers independently associated with subsequent invasive disease

	DCIS cases		DCIS controls		OR (95% CI) ^a	P ^b
	n	(%)	n	(%)		
Periductal fibrosis/COX-2						
Absent/Low	11	(5.9)	86	(20.5)	1.00 (reference)	
Present/Low	8	(4.3)	20	(4.8)	3.26 (1.13-9.42)	
Absent/High	102	(55.1)	213	(50.7)	3.82 (1.94-7.51)	
Present/High	59	(31.9)	92	(21.9)	4.87 (2.37-9.99)	<0.001
N/A	5	(2.7)	9	(2.1)		
Necrosis/COX-2						
Absent/Low	3	(1.6)	39	(9.3)	1.00 (reference)	
Present/Low	16	(8.6)	67	(16.0)	2.34 (0.63-8.66)	
Absent/High	37	(20.0)	76	(18.1)	5.07 (1.48-17.33)	
Present/High	124	(67.0)	230	(54.8)	5.76 (1.75-18.95)	0.001
N/A	5	(2.7)	8	(1.9)		
Periductal fibrosis/HER2						
Absent/Negative	89	(48.1)	244	(58.1)	1.00 (reference)	
Present/Negative	31	(16.8)	65	(15.5)	1.25 (0.75-2.09)	
Absent/Positive	24	(13.0)	54	(12.9)	1.26 (0.74-2.17)	
Present/Positive	38	(20.5)	50	(11.9)	2.04 (1.23-3.38)	0.050
N/A	3	(1.6)	7	(1.7)		
Necrosis/HER2						
Absent/Negative	35	(18.9)	107	(25.5)	1.00 (reference)	
Present/Negative	85	(45.9)	203	(48.3)	1.26 (0.80-1.99)	
Absent/Positive	5	(2.7)	9	(2.1)	1.72 (0.52-5.63)	
Present/Positive	57	(30.8)	95	(22.6)	1.84 (1.10-3.09)	0.12
N/A	3	(1.6)	6	(1.4)		
Periductal fibrosis/Necrosis/COX-2						
All other groupings	19	(10.3)	106	(25.2)	1.00 (reference)	
Negative/Negative/High	28	(15.1)	66	(15.7)	2.39 (1.21-4.71)	
Positive/Negative/High	9	(4.9)	10	(2.4)	4.60 (1.62-13.03)	
Negative/Positive/High	74	(40.0)	147	(35.0)	2.90 (1.62-5.21)	
Positive/Positive/High	50	(27.0)	82	(19.5)	3.39 (1.81-6.34)	0.002
N/A	5	(2.7)	9	(2.1)		
Fibrosis/Necrosis/HER2						
All other groupings	40	(21.6)	116	(27.6)	1.00 (reference)	
Negative/Positive/Negative	62	(33.5)	149	(35.5)	1.19 (0.74-1.91)	
Positive/Positive/Negative	23	(12.4)	53	(12.6)	1.15 (0.62-2.12)	
Negative/Positive/Positive	21	(11.4)	45	(10.7)	1.39 (0.73-2.66)	
Positive/Positive/Positive	36	(19.5)	50	(11.9)	2.02 (1.14-3.56)	0.18
N/A	3	(1.6)	7	(1.7)		

(Continues on next page)

Table S4. Multivariable analysis of histopathological characteristics and immunohistochemical markers independently associated with subsequent invasive disease

	DCIS cases		DCIS controls		OR (95% CI) ^a	P ^b
	n	(%)	n	(%)		
Necrosis/HER2/COX-2						
All other groupings	13	(7.0)	65	(15.5)	1.00 (reference)	
Positive/Negative/Low	11	(5.9)	43	(10.2)	1.14 (0.45-2.89)	
Negative/Negative/High	32	(17.3)	71	(16.9)	2.12 (1.00-4.50)	
Positive/Negative/High	73	(39.5)	156	(37.1)	2.23 (1.13-4.41)	
Positive/Positive/High	50	(27.0)	72	(17.1)	3.43 (1.66-7.08)	0.004
N/A	6	(3.2)	13	(3.1)		

^a: Comparisons between DCIS cases and DCIS controls were made by multivariate conditional logistic regression; ^b: P-values for overall effect were calculated by Wald-test.

N/A: Not assessable. N/As were not included in the analysis.

LR chi2 corrected for degrees of freedom were 6.34 for periductal fibrosis/COX-2, 5.18 for necrosis/COX-2, 2.60 for periductal fibrosis/HER2, 1.95 for necrosis/HER2, 4.39 for periductal fibrosis/necrosis/COX-2, 1.57 for periductal fibrosis/necrosis/HER2, and 3.83 for necrosis/HER2/COX-2.

Lindy L. Visser
Marlous Hoogstraat
Tycho Bismeijer
Frank Nieboer
Michiel de Maaker
Petra Kristel
Lennart Mulder
Marielle Kreté
Wim Brugman
Michael Schaapveld
Marjanka K. Schmidt
Lodewyk F. A. Wessels
Jelle Wesseling*
Esther H. Lips*
On behalf of the PRECISION team

Manuscript in preparation

*: Shared last author

CHAPTER 4

Deciphering the genomic and transcriptomic landscape of ductal carcinoma in situ associated with and without subsequent invasive breast cancer

Abstract

Introduction. Ductal carcinoma in situ (DCIS) is a heterogeneous disease with variable risk of developing into invasive breast cancer (IBC). Treatment decisions are complicated by the lack of clinicopathological factors that can reliably predict the risk of future invasive disease. Identification of molecular alterations differentiating lesions that will remain indolent from those that will likely progress to IBC is needed to provide more accurate risk assessment.

Methods. We performed extensive molecular profiling, including whole transcriptome RNA sequencing, low coverage whole genome sequencing, and panel sequencing of DCIS lesions associated with (n=101) and without (n=82) subsequent IBC, to find marker associated with subsequent IBC after DCIS.

Results. Gene expression analysis comparing all DCIS with and without subsequent IBC did not identify markers associated with subsequent IBC. After stratifying DCIS by PAM50 subtype, we found that genes belonging to immune-related pathways were positively enriched in luminal A DCIS with subsequent IBC, whereas the opposite was seen in luminal B DCIS. Copy number analysis restricted to 53 breast cancer driver genes showed that DCIS cases with subsequent IBC had more copy number events than DCIS without subsequent IBC ($p=0.022$). Comparison of CNA and mutations at the gene level resulted in subtle differences between cases and controls, of which non were statistically significant.

Conclusion. Although this study is one of the largest in its sort, we only identified subtle differences between DCIS with and without a subsequent IBC. Our analysis revealed that immune-related pathways play a role in the progression of DCIS. In addition, we found more copy number alterations in DCIS with a subsequent recurrence.

Introduction

Ductal carcinoma in situ (DCIS) is a neoplastic, non-invasive breast lesion with a variety of biological behaviors.¹ Some lesions have the potential to progress to IBC within a time period varying from a few years to several decades, whereas others will remain DCIS within a patient's lifetime. As such, DCIS is considered to be a non-obligate precursor of invasive breast cancer.

DCIS at first was a rarely diagnosed disease, but with the introduction of population based breast cancer screening it nowadays represents 20-25% of all newly diagnosed so called "breast cancers".²⁻⁴ After detection, the optimal clinical management of DCIS remains controversial, as we are unable to identify those DCIS lesions with invasive potential. As a result, almost all women with DCIS undergo treatment preventive of invasive breast cancer development. Treatment for DCIS consists of breast conserving surgery (BCS), often followed by radiotherapy and/or endocrine therapy, or mastectomy. Many women, who have a low risk to develop subsequent invasive disease, do not benefit from this treatment and thus suffer from overdiagnosis and overtreatment.^{5,6}

Thus, there clearly is a high need for robust prognostic markers that can reliably predict the disease course of DCIS. Whereas genomic and transcriptomic characterization have significantly improved prognostication of invasive breast cancer (e.g. Mammaprint™, PAM50™), DCIS prognostication is still largely based on clinical and histopathological parameters. Factors such as African-American race, premenopausal status, detection by palpation, involved margins, high histological grade, and high p16 expression have been shown to be predictive of subsequent IBC after DCIS.⁷ However, these clinicopathological features are insufficient in current practice to distinguish lesions that will ultimately progress to invasive breast cancer from those that will not.

We hypothesize that information on the future disease course of DCIS might be encoded in the genomic and transcriptomic profile present at diagnosis. Our hypothesis is supported by the fact that genomic and transcriptomic profiles of DCIS are highly similar to invasive disease.^{8,9} Yet, molecular profiling of DCIS poses challenges. First, the study cohort should include long term clinical follow up to obtain information on potential recurrences. Second, technologies need to be selected that are compatible with low inputs of low-quality FFPE-derived DNA and RNA, as DCIS lesions are often small and thus yield small amounts of DNA and RNA, which is often highly degraded.

For this study we made use of a well-annotated DCIS cohort with long-term follow-up, including women treated for DCIS with BCS alone, that subsequently did or did not develop invasive breast cancer. The primary objective was to identify molecular markers predictive of subsequent invasive breast cancer development after DCIS.

Patients and Methods

Study population

The subjects in this study were derived from a case-control study nested in a Dutch population-based cohort of women with primary DCIS treated with breast conserving surgery (BCS) alone (n=2,658) between 1989 and 2005, which is described previously.¹⁰ In brief, this cohort was derived from the Netherlands Cancer Registry (NCR). Clinicopathological data was obtained from the NCR and the nation-wide histopathology and cytopathology data network (PALGA). Patients included in this study did not receive tamoxifen or other anti-hormonal adjuvant treatment, did not present with any evidence of invasive disease and did not have a prior cancer diagnosis except for nonmelanoma skin cancer. Follow-up for subsequent ipsilateral IBC and vital status were complete until at least January 1, 2011 and the mean follow-up was 12.0 years. A nested case-control study was set up including 316 women with subsequent ipsilateral IBC, which were considered “cases”. Controls, representing women without subsequent IBC, were matched to cases based on age at DCIS diagnosis and follow-up time.

Formalin-fixed paraffin-embedded (FFPE) tissue blocks were successfully retrieved and pathologically re-examined to confirm DCIS diagnosis, of 200 cases and 474 controls.¹⁰ DNA and RNA, of sufficient yield and quality, were extracted from 101 cases and 82 controls (Figure 1). Data on histological grade was available for all DCIS lesions and ER status (IHC) was available for 98 cases and 80 controls, and HER2 status (IHC; CISH) was available for 98 cases and 77 controls (Supplementary File 1).¹⁰

The study was approved by the review board of the NKR (request K12.281; January 3, 2013) and PALGA (LZV990; April 16, 2013). Secondary use of tissue and data for this study was done under an opt-out regime which is conform Dutch regulations and the Code of Conduct of Federa-COREON.

Laser microdissection and simultaneous DNA and RNA extraction

To reassure a more cellular-based rather than tissue-based resolution for our molecular studies, laser-microdissection was performed. Tissue processing prior to laser-microdissection has been described previously (Visser, et al. Chapter 6). Microdissection of 8- μ m-thick, toluidine blue-stained, FFPE tissue sections was performed using a laser microdissection (LMD) microscope (Leica Microsystems). The number of sections per case used for microdissection was determined by the size and cellularity of the DCIS lesion, with an average of 8 sections (range 2-15) per lesion.

Laser microdissected tissue fragments were subjected to simultaneous DNA and RNA extraction using the Qiagen AllPrep DNA/RNA FFPE kit (QIAGEN), following manufacturer's instructions. DNA samples were RNase treated. Using a Bioanalyzer 2100 (Agilent), quality and quantity of total RNA was assessed and the percentage RNA fragments >200bp (DV200)

was determined. DNA concentrations were measured using the Qubit dsDNA High-Sensitivity assay and Qubit 2.0 Fluorometer (Life Technologies).

RNA samples were submitted for RNA sequencing and DNA samples were submitted for low coverage whole genome sequencing and panel sequencing (PanelSeq). In cases with sufficient DNA (>50 ng), we performed both assays. In cases with insufficient DNA for both assays, only low coverage whole genome sequencing was performed.

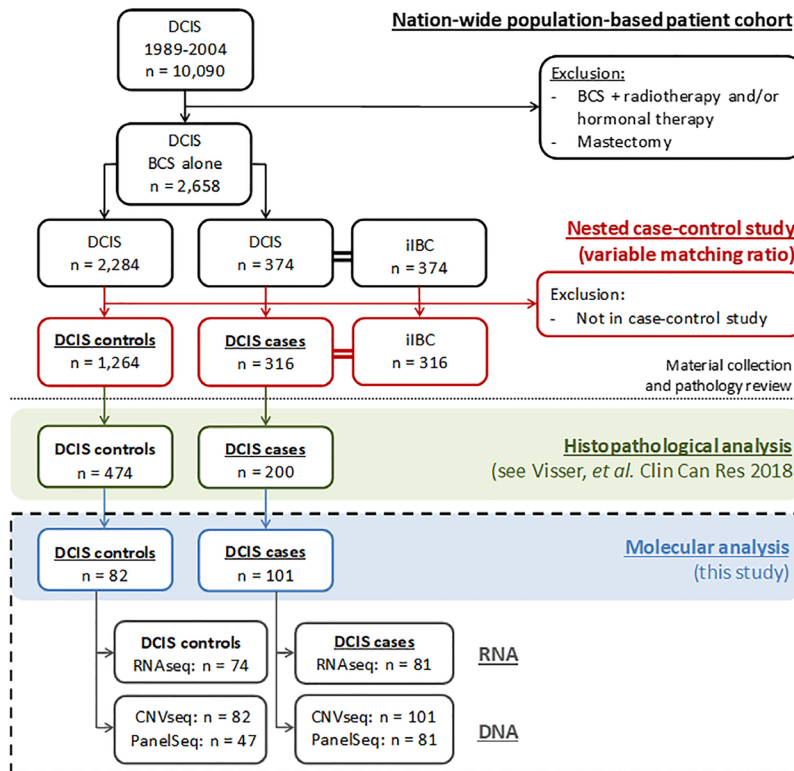


Figure 1. Design of the study. In this study we make use of a nation-wide population-based patient cohort. Within this group of patients, we selected for the women treated with breast conserving surgery (BCS) only, of which a substantial part developed an ipsilateral invasive breast cancer (iIBC; black). A nested case-control study was set up where patients with a subsequent iIBC were defined as cases, and patients without an iIBC in the same follow-up period were defined as controls. Each DCIS case was matched with controls based on age and follow-up time, using a variable matching ratio (red). After material collection and pathology re-evaluation, 200 cases and 474 controls were available (green marking), and 101 cases and 82 controls were available for the molecular studies presented here (dashed line with blue marking).

Whole transcriptome RNA sequencing (RNAseq)

Total RNA samples were converted to strand-specific libraries using TruSeq RNA Access Library Prep kit (Illumina, RS-301-2001/2), according manufacturer's instructions, protocol part #15049525 Rev. B. The libraries were sequenced with 65-bp single-reads using a HiSeq 2500 instrument (Illumina) to an average of 20 million reads per sample. Reads were aligned to the reference genome GRCh38 (hg38), using Tophat single end.¹¹ Gene expression was calculated using Icount.

Normalization was performed using DESeq2.¹² Read counts were log transformed. We excluded outlier samples based on clustering and principal component analysis (PCA; Supplementary Figure S1). Subtyping of DCIS lesions based on the PAM50 gene signature was performed using the geneFu package.¹³

Analysis of *ESR1* and *ERBB2* by gene expression showed good correlation with IHC protein expression (Supplementary Figure S2). Likewise, *MKI67* gene expression showed a good correlation with histological grade.

Low coverage whole genome sequencing (CNVseq)

Sequencing methodologies and bioinformatics procedures for low coverage whole genome sequencing (CNVseq) are described elsewhere (Visser, et al. Chapter 6). In short, DNA samples were library prepped with the KAPA hyper prep kit (KAPA Biosystems). Sequencing was performed on a HiSeq 2500 instrument (Illumina) with single-end 65-bp reads, obtaining a coverage of 0.2x. Reads were aligned to the reference genome GRCh38 (hg38) using BWA-MEM.¹⁴ Copy number segmentation was performed using DNACopy.¹⁵ Segments were identified as lost or gained at log2 ratios of <-1 and >1.5, respectively. Recurrent copy number events were identified using RUBIC.¹⁶ To clean up the data, we removed losses in noisy samples (>1000 segments).

Panel sequencing (PanelSeq)

Panel sequencing (PanelSeq) was performed using a custom-made amplicon panel covering 53 breast cancer driver genes. More details on this panel can be found elsewhere (Visser, et al. Chapter 6). Sequencing was performed on the Ion Torrent Personal Genome Machine (PGM; Life Technologies). Reads were aligned to the reference genome GRCh37 (hg19) using the Torrent Mapping Alignment Program, and variant calling was performed using Torrent Variant Caller (TVC) version 5.6.

Variant data in VCF format was first translated to GRCh38 and annotated using bedtools, picard (<https://broadinstitute.github.io/picard/command-line-overview.html>), samtools, bcftools and the Ensembl VEP, and further analyzed in R, employing vcfr, and tidyverse.¹⁷⁻²⁰ Variants were subjected to a filtering pipeline during which low quality variants (variant allele frequency [VAF] <10%, coverage <100x, and a quality [QUAL] of <1000), artifacts (found in >90% of samples), and germline variants (>5 cases in GNOMAD and GoNL; except

for *BRCA1*, *BRCA2*, and *TP53*) were removed, resulting in a list of high confidence somatic variants.

Statistical analysis

Differential gene expression (DE) analysis was performed using DESeq2.¹² A threshold of $p < 0.1$ was used to select differentially expressed genes. Cluster analysis and visualization was performed using the NMF package (aheatmap function).²¹ Gene set enrichment analysis (GSEA) with the Hallmark gene sets, consisting of 50 gene sets, was performed to identify pathways associated with genes differentially expressed in our analysis of subsequent IBC.²²

Differential copy number analysis was performed using the Mann-Whitney test. The Chi² test and Fisher's exact test were used to assess differences in mutation frequencies between DCIS cases and controls. *P* values ≤ 0.05 were considered statistically significant.

Results

Patient characteristics

Clinicopathological characteristics of the study group are shown in Supplementary Table S1. It includes women diagnosed with DCIS and treated with BCS alone, who subsequently did (cases) or did not (controls) develop subsequent ipsilateral IBC. There were no clinicopathological differences between the DCIS cases and controls. This study group is a representative sample of a previously published case-control study (Supplementary Table S2).¹⁰

Characterization of DCIS transcriptomic profiles

Differential gene expression (DE) analysis between DCIS associated with and without subsequent IBC resulted in 24 differentially expressed genes ($p < 0.1$) of which 4 were statistically significant ($p < 0.05$). Unsupervised clustering did not show a good separation of the two groups (Supplementary Figure S3).

Next, we categorized the DCIS lesions based on PAM50 subtype (Supplementary Figure S4) to repeat the DE analysis per subtype. Luminal DCIS (both A and B) were all ER-positive based on IHC (Supplementary Figure S4; Supplementary File 1)). Luminal A DCIS consisted of grade 1 and 2 lesions and basal-like DCIS only of grade 2 and 3. Seventy percent of HER2-positive DCIS (30 of 43; IHC, CISH) were classified as Her2-enriched, 14.0% as basal-like (6 of 43), and 11.6% (5 of 43) as luminal B. The majority of HER2-negative DCIS was classified as luminal (83%; 35 of 106 luminal A; 53 of 106 luminal B). PAM50 subtype was not predictive of subsequent IBC after DCIS ($p = 0.12$; Figure 2A).

Per PAM50 subtype, *Y* differentially expressed genes were identified of which *X* were statistically significant ($p < 0.05$; *X/Y*): luminal A (64/91), luminal B (1/34), Her2 (13/33), basal-like (11/13), and normal-like (56/80; Figure 2B-F). Based on the differentially

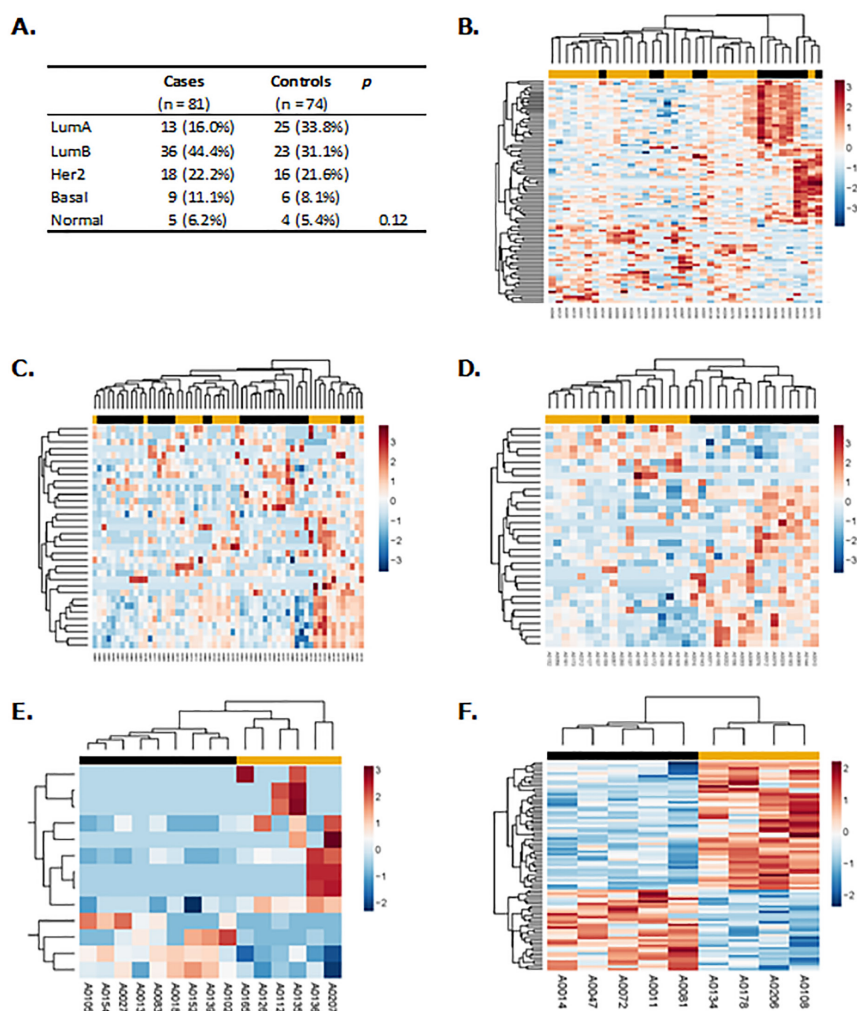


Figure 2. Heatmaps visualizing results of differential gene expression between DCIS associated with (cases; black) and without (controls; yellow) subsequent iIBC per PAM50 subtype. **A.** Comparison of the number of cases and controls per PAM50 subtype by Fisher' exact test; **B.** Luminal A DCIS; **C.** Luminal B DCIS; **D.** Her2-enriched DCIS; **E.** Basal-like DCIS; **F.** Normal-like DCIS. Heatmaps (B-F) include differentially expressed genes $p < 0.1$.

expressed genes, we were able to discriminate between DCIS with and without subsequent IBC.

Gene set enrichment analysis (GSEA) showed that the genes associated with subsequent IBC belonged to immune-related pathways, indicating that these pathways play a role in DCIS progression. Genes belonging to immune related pathways were enriched in luminal A DCIS cases (Figure 3; Supplementary Figure S5), but also enriched in luminal B DCIS controls.

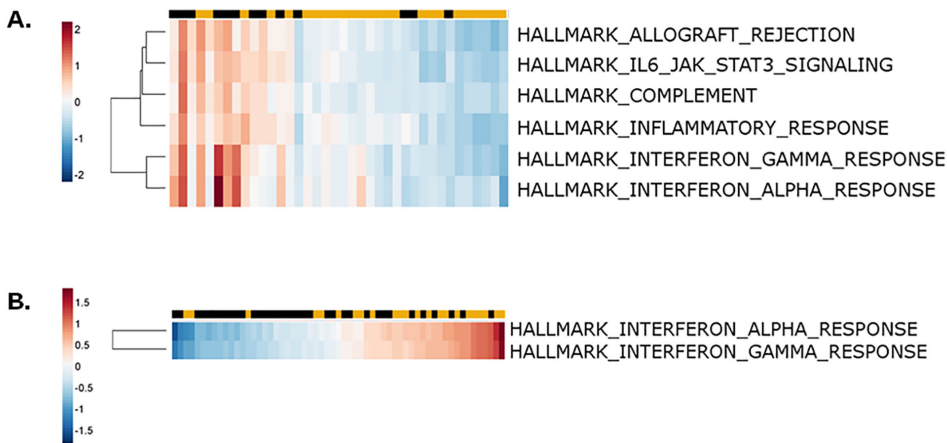


Figure 3. Results of the gene set enrichment analysis visualized in a heatmap. **A.** Results of analysis of Luminal A DCIS. **B.** Results of analysis of Luminal B DCIS. Annotation bar on top: black, DCIS associated with subsequent iIBC; yellow, DCIS without subsequent iIBC. Enriched immune-related gene sets with NORM p-value <0.05 and FDR q-value <0.1 are shown.

Characterization of DCIS genomic profiles

Next, we assessed the burden and spectrum of CNA and mutations in DCIS. We observed that marked aneuploidy is already present in the DCIS stage, with copy number alterations (CNAs) present across the genome. The most frequent copy number gains were observed on chromosome 1q, 8q, 16p, 17q, and 20q, while copy number loss was detected on 1p, 3p, 8p, 11q, and 16q, also known to be the most common genomic aberrations in IBC (Figure 4; Supplementary Figure S6). Frequent gene amplifications were observed in *ERBB2*, *ZNF217*, *CCND1*, *FGFR1* and *MYC*, and frequent gene deletions were observed in *CDKN2A/B*, *MAP2K4*, *TP53*, *CBFB*, *KMT2A*, *RB1*, and *ESR1* (Figure 5).

Differential copy number analysis between all cases and all controls showed no differences when looking at mean copy number profiles (Supplementary Figure S6; Supplementary File 2). However, visual inspection of mean copy number plots of DCIS cases and controls,

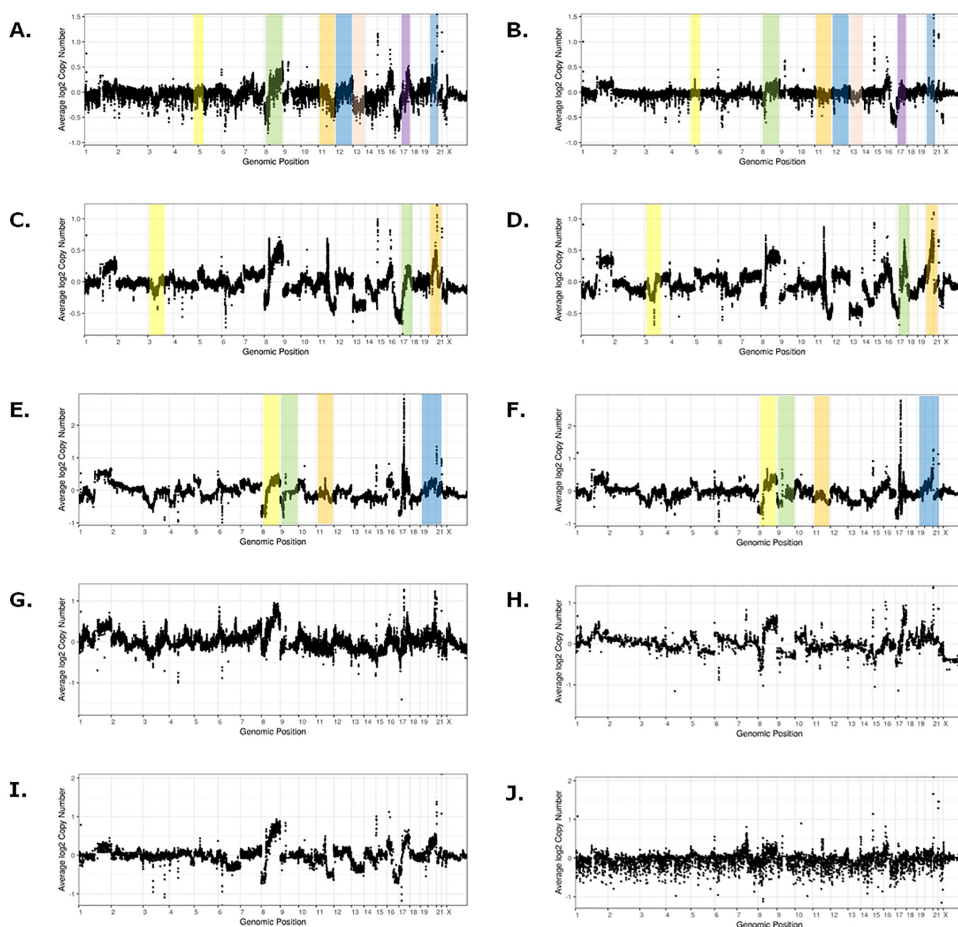


Figure 4. Mean copy number plots of DCIS with (cases, left column) and without (controls; right column) subsequent iIBC **A.** Luminal A DCIS cases, $n=13$; **B.** Luminal A DCIS controls, $n=25$; **C.** Luminal B DCIS cases, $n=36$; **D.** Luminal B DCIS controls, $n=23$; **E.** Her2 enriched DCIS cases, $n=18$; **F.** Her2 enriched DCIS controls, $n=16$; **G.** Basal-like DCIS cases, $n=9$; **H.** Basal-like DCIS controls, $n=6$; **I.** Normal-like DCIS cases, $n=5$; **J.** Normal-like DCIS controls, $n=4$.

stratified per PAM50, did result in differences, although not statistically significant (Figure 4; Supplementary File 2). Restricted to the 53 breast cancer driver genes, 61.4% of cases (62/101) and 57.3% (47/82) controls harbored at least one copy number alteration. DCIS associated with subsequent IBC harbored more CNA than DCIS without subsequent IBC, as on average 1.6 (range 0-11) alterations were found in cases and 1.0 (range 0-5) alterations in controls ($p=0.022$; Supplementary File 1).

Mutation analysis included 81 cases and 47 controls. 76.5% of DCIS cases (62/81) and 72.3% of DCIS controls harbored at least one somatic mutation (34/47; Figure 6; Supplementary File 1 and 3). Frequent somatic mutations were observed in *PIK3CA* (36%) and *TP53* (30%) followed by *BRCA2* (10%), *NF1* (10%), *GATA3* (9%), *KMT2C* (7%), and *AKT1* (7%; Figure 6).

There was no difference in the number of somatic variants harbored by DCIS cases and controls ($p=0.52$; Supplementary File 1 and 3). Also, no difference was found when assessing the number of somatic variants per DCIS PAM50 subtype. Lastly, we compared the most frequently found mutations (*PIK3CA*, *TP53*, *NF1*, and *BRCA2*; found in at least 10 samples) between DCIS cases and controls (Supplementary Table S3). Mutations in *PIK3CA* were present more often in DCIS controls (36.2%; 17/47) as compared to DCIS cases (22.2%; 18/81; $p=0.041$), but this did not remain statistically significant after multiple testing correction.

Discussion

We performed extensive molecular profiling of DCIS lesions associated with and without subsequent invasive breast cancer, involving transcriptomic and genomic analyses to identify molecular markers predictive of subsequent invasive breast cancer development after DCIS. To date, this is the largest molecular profiling study of DCIS. Based on transcriptomic data, we found a role for immune-related pathways in the progression of DCIS to IBC. DCIS with subsequent IBC had more CNAs as compared to DCIS without an invasive recurrence. Furthermore, comparison of CNA and mutations at the gene level resulted in only subtle differences, of which none were statistically significant.

Differential gene expression analysis comparing all DCIS with and without subsequent IBC, resulted in only a few differentially expressed genes. Interestingly, more differentially expressed genes and a clearer difference was found when DCIS lesion were stratified based on their PAM50 subtype. This is in line with a previous finding of Lesurf and colleagues, where they showed that comparison between DCIS and IBC was compromised when unmatched for intrinsic subtype, as expression profiles were strongly driven by the intrinsic subtypes.²⁵ Thus, although the subtypes have limited prognostic value for DCIS, it is still of value to stratify lesions based on intrinsic subtype because of distinct evolutionary disease paths.

Gene set enrichment analysis (GSEA) showed that genes belonging to immune related pathways were enriched in Luminal A DCIS associated with subsequent IBC. This finding is especially interesting as the samples were free of stromal contamination and only contained few immune cells, since we used laser microdissection to reassess enrichment of neoplastic cells for DNA and RNA extraction. In Luminal B DCIS we observed the opposite: genes belonging to the immune related pathways were enriched in DCIS lesions without subsequent IBC. Thus, in low proliferative luminal A DCIS, involvement of the immune system led to the

Copy number alterations found in 109 (60%) of 183 DCIS lesions

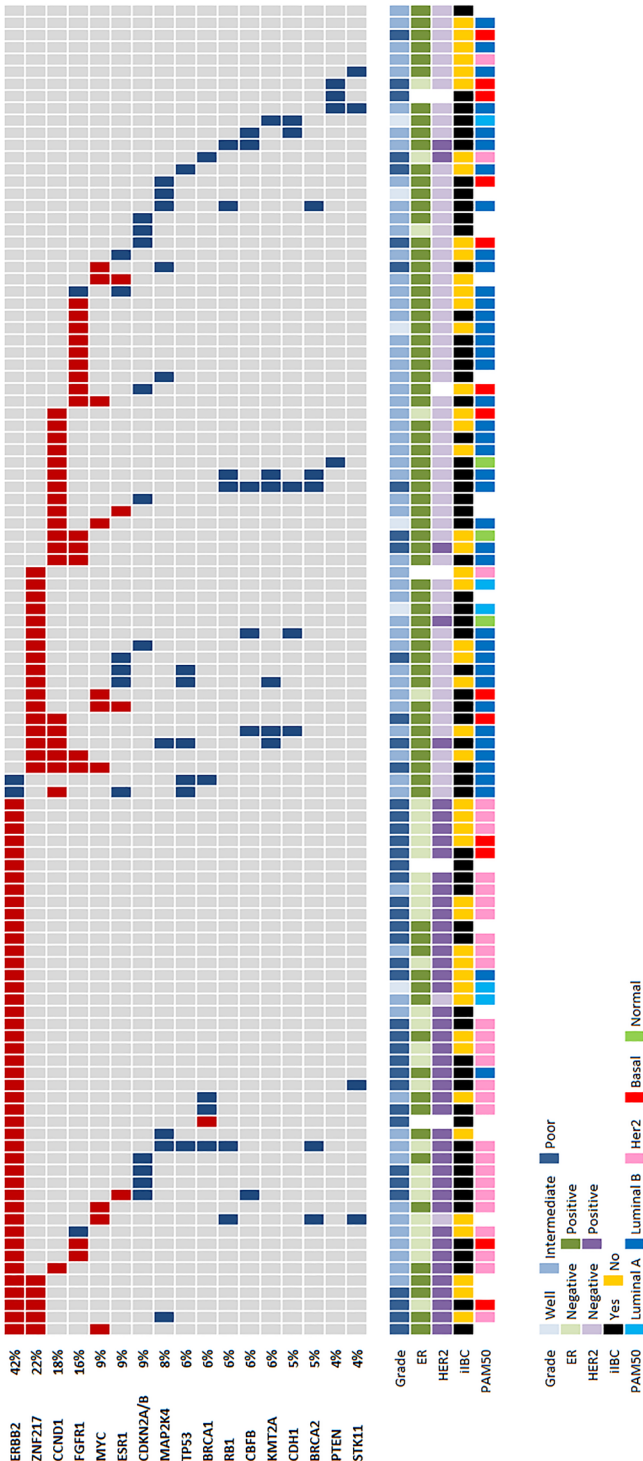


Figure 5. Overview figure of copy number alterations (red, amplifications; blue, losses) restricted to 53 driver breast cancer driver genes. Genes affected in at least 4 DCIS samples are shown. Copy number alterations were found in 109 (columns) of 183 DCIS samples. Samples are annotated with information on grade, ER and HER2 status (IHC), whether the DCIS was associated with subsequent iIBC, and PAM50 subtype.

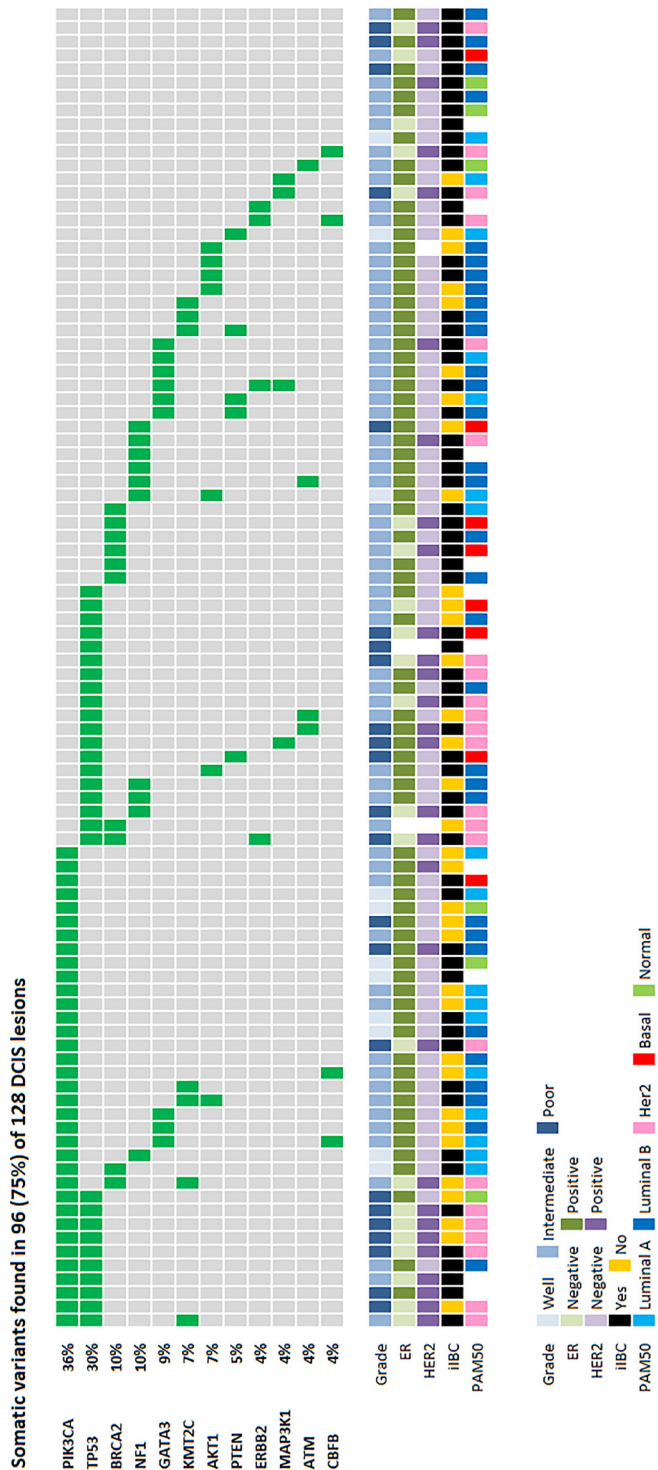


Figure 6. Overview figure of somatic variants detected by panel sequencing using custom 53-gene panel. Genes affected in at least 4 DCIS samples are shown. Somatic variants were found in samples of 96 (columns) of 128 DCIS samples. Samples are annotated with information on grade, ER and HER2 status (IHC), whether the DCIS was associated with subsequent iIBC, and PAM50 subtype.

development of invasive disease (pro-tumor), whereas among high proliferative luminal B DCIS involvement of the immune system seemed to inhibit invasive outgrowth (anti-tumor). This interplay between immunity and proliferation has been previously described by Nagalla and colleagues.²⁶ They investigated the relationship between proliferation and immune gene signatures within breast cancer and suggested that a productive interplay among immune cell types at the tumor site promotes long-term anti-metastatic immunity in a proliferation-dependent manner.

We hypothesized that information on the future disease course of DCIS might be encoded in the genomic and transcriptomic profile present at diagnosis. In this study, we only found subtle differences between DCIS associated with and without subsequent IBC, based on CNA and mutation data. A reason for this could be DCIS heterogeneity at the genetic level, as this complicates identification of differences between DCIS with and without subsequent IBC. We identified many different CNAs and mutations, but these were only present in a small number of samples. Outside of transcriptomic and (somatic) genomic factors, there are also suggestions that the tumor micro-environment plays an important role in the disease course of DCIS.²⁷ In addition, a DCIS GWAS study identified genetic polymorphisms that predispose to DCIS.²⁸ Genetic predisposition could also play a role in DCIS progression. Furthermore, lifestyle factors could make a woman prone to develop subsequent IBC after DCIS.^{29,30}

Previous molecular profiling studies on DCIS mainly focused on analyzing DCIS adjacent to an invasive component ("synchronous" or "mixed" DCIS) whereby the two components are compared to find differences that may be related to invasion. Furthermore, studies have been conducted examining DCIS in the absence of invasive disease ("pure" DCIS) in comparison to unmatched IBC lesions (from different individuals). Copy number changes and promotor methylation have mainly been associated with poor prognostic factors such as high nuclear grade.³¹⁻³³ To date, only a single study evaluated the association of copy number changes with recurrence.³⁴ In line with our findings, Gorringer and colleagues reported that DCIS cases with recurrence (*in situ* or invasive) had more CN events than DCIS without recurrence. In addition, they found several genomic regions showing an increase in frequency in DCIS associated with recurrent disease, including 20q gain, *ERBB2* amplification and 15q loss. In our study, these associations were not found, which could be explained by the fact that we specifically focused on invasive recurrences after DCIS. Furthermore, one whole-genome methylation study has been undertaken focused on finding methylation makers associated with recurrence. Johnson and colleagues found that differentially methylated loci in DCIS associated with subsequent IBC were enriched for homeobox-containing genes and genes involved in limb morphogenesis (e.g. *HOXB13* and *EN1*).⁸ Somatic mutations are relatively infrequently assessed in DCIS, and when assessed it involved only small patient cohorts and/or small gene panels.³⁵⁻⁴¹ In line with our study, *PIK3CA* and *TP53* were found to be the most frequently mutated genes in DCIS.^{9,35-42} They are believed to be early events in the development of DCIS rather than progression.

Compared with genetic alteration-based prognostic markers, gene expression-based prognostic markers have had more success in discriminating DCIS with and without subsequent IBC. Several gene expression studies have been performed for both pure and synchronous DCIS.^{43–49} Yet, individual gene lists differ between studies. The first and only commercially available multigene expression panel is the Oncotype DX DCIS scoreTM.⁵⁰ This 12-gene quantitative PCR-based assay has been adapted for DCIS and predicts the risk of local recurrence (*in situ* or invasive) after BCS for DCIS, estimates the benefit from RT, and gives a numeric score that places patients into low- intermediate- or high risk categories. The assay has been tested in two cohorts in which the test had prognostic value, but they were unable to demonstrate any difference in outcome between the risk groups (low, intermediate, high).^{51,52} The two studies also had limitations in that only about 50% of patients could be tested, a result was not possible for about 15% of cases on top of this 50%, and confidence intervals were very wide. Furthermore, prospective validations of the DCIS score is still lacking. Taken together, results from previous studies are insufficient for finding an association with invasive recurrence.

Our study has several strengths. First, we made use of a unique, well-annotated series of women diagnosed with and treated for DCIS with BCS alone and subsequently did or did not experience subsequent invasive breast cancer during a mean follow-up of 12.0 years. Second, use of laser microdissection enabled us to reassure enrichment of DCIS epithelial cells for our molecular study. Third, for the majority of DCIS samples we collected both whole transcriptome RNA sequencing and low coverage whole-genome sequencing data, with the addition of PanelSeq data for a subset of these samples.

Our study also has some limitations. First, we compared CNA and mutations based on a gene panel. This method covered only 53 genes and thus represent a part of the full picture. Second, due to the lack of normal DNA, we used germline databases (GNOMAD and GoNL) to remove common germline variants from our PanelSeq data. We cannot fully exclude the possibility that germline mutations that are not recognized as such are called as somatic mutations. This is probably also why we detected such a high rate of *BRCA2* mutations. Third, as all women in our cohort were treated for DCIS, we have to keep in mind that treatment in this study could have introduced potential bias. Where there is no recurrence, the patient could have been treated adequately. In other words, the absence of a recurrence does not mean that the preceding DCIS lesion did not have invasive potential. Also, the invasive recurrences could either be a true recurrence of the preceding DCIS or a new primary tumor.

In summary, our study accounts for the high need of well-designed studies with large patient numbers, with uniformly treated patients, and long-term follow-up. Despite our efforts in minimizing potential bias as much as possible, we found only limited differences between DCIS with and without subsequent IBC. We confirmed the need for stratification based on intrinsic subtype. This way, we found a role for immune-related pathways in the

progression of DCIS, involving a proliferation-based anti-tumor immunity. These results support the findings of Lesurf et al, describing the existence of subtype-specific evolutionary disease paths in DCIS progression.²⁵

As a next step, we will validate our findings in an independent cohort. Currently, we take an integrative, comprehensive approach that involves analysis of the DCIS methylome, the tumor micro-environment, and tumor associated stroma. In addition, studies are initiated to explore the role of pathogenic germline variants and lifestyle factors in determining the risk of developing subsequent IBC after DCIS.

Article information

Author's contributions: LLV, MH, MS, MKS, LFAW, EHL, and JW designed the study. LLV, FN, MdM, PK, and LM set-up and performed DNA and RNA extractions. MK and WB set-up and performed library preparations and sequencing. LLV, MH, and TB analysed the data. LLV, MH, TB, MS, MKS, LFAW, EHL, and JW interpreted the data. LLV wrote the report. All authors provided comments and approved the final version.

Funding: This work was jointly funded by Cancer Research UK and the Dutch Cancer Society (grant number C38317/A24043, to J. Wesseling).

Acknowledgements: We thank all patients of which stored residual material was used in this study, all collaborating hospitals and in particular the pathology departments, and all persons who helped in the process of data collection and analysis. We would like to specifically acknowledge Otto Visser, Annemarie Eeltink, and the registration teams of the Netherlands Comprehensive Cancer Organization for collection of the data for the Netherlands Cancer Registry. We thank Lucy Overbeek and PALGA, the nation-wide histopathology and cytopathology data network and archive, for providing pathology data and for their help in the collection of the residual patient material. We acknowledge the staff of the NKI-AVL Core Facility Molecular Pathology & Biobanking for their technical support, the staff of the NKI-AVL Genomics Core Facility and Ronald van Marion of the Erasmus University Medical Center for their sequencing support, and Asli Kucukosmanoglu for her assistance during data analysis.

Supplementary data: Supplementary File 1-3 will be available after publication of the article.

References

- To T, Wall C, Baines CJ, Miller AB. Is Carcinoma in Situ a Precursor Lesion of Invasive Breast Cancer? *Int J Cancer*. March 2014;1-7.
- Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst*. 2002;94(20):1546-1554.
- Netherlands Comprehensive Cancer Organization (IKNL). DCIS incidence trends over time, the Netherlands.
- Howlander N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2014*. National Cancer Institute. Bethesda, MD
- Løberg M, Lousdal ML, Bretthauer M, Kalager M. Benefits and harms of mammography screening. *Breast Cancer Res*. 2015;17(1):63.
- Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102(9):605-613.
- Visser LL, Groen EJ, Leeuwen FE Van, Lips EH. Predictors of an Invasive Breast Cancer Recurrence after DCIS : A Systematic Review and Meta-analyses. *Cancer Epidemiol Biomarkers Prev*. 2019;(11):1-12.
- Johnson KC, Koestler DC, Fleischer T, et al. DNA methylation in ductal carcinoma in situ related with future development of invasive breast cancer. *Clin Epigenetics*. 2015;7(1):75.
- Kim SY, Jung S, Kim MS, et al. Genomic differences between pure ductal carcinoma in situ and synchronous ductal carcinoma in situ with invasive breast cancer. *Oncotarget*. 2015;6(10):7597-7607.
- Visser LL, Elshof LE, Schaapveld M, et al. Clinicopathological Risk Factors for an Invasive Breast Cancer Recurrence after Ductal Carcinoma In Situ-A Nested Case-Control Study. *Clin Cancer Res*. April 2018;clincanres.0201.2018.
- Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- Gendoo DMA, Ratanasirigulchai N, Schröder MS, et al. Genefu: an R/Bioconductor package for computation of gene expression-based signatures in breast cancer. *Bioinformatics*. 2016;32(7):1097-1099.
- Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. March 2013.
- Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*. 2007;23(6):657-663.
- van Dyk E, Hoogstraat M, Ten Hoeve J, Reinders MJT, Wessels LFA. RUBIC identifies driver genes by detecting recurrent DNA copy number breaks. *Nat Commun*. 2016;7:12159.
- Quinlan AR. BEDTools : the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinforma*. 2014.
- Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011;27(21):2987-2993.
- McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
- Knaus BJ, Grünwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour*. 2017;17(1):44-53.
- Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics*. 2010;11:367.
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425.
- Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61-70.
- Michailidou K, Lindström S, Dennis J, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature*. 2017;551(7678):92-94.
- Lesurf R, Aure MR, Mørk HH, et al. Molecular Features of Subtype-Specific Progression

- from Ductal Carcinoma In Situ to Invasive Breast Cancer. *Cell Rep.* 2016;16(4):1166-1179.
26. Nagalla S, Chou JW, Willingham MC, et al. Interactions between immunity, proliferation and molecular subtype in breast cancer prognosis. *Genome Biol.* 2013;14(4):R34.
27. Nelson AC, Machado HL, Schwertfeger KL. Breaking through to the Other Side: Microenvironment Contributions to DCIS Initiation and Progression. *J Mammary Gland Biol Neoplasia.* August 2018.
28. Petridis C, Brook MN, Shah V, et al. Genetic predisposition to ductal carcinoma in situ of the breast. *Breast Cancer Res.* 2016;18(1):22.
29. Habel LA, Daling JR, Newcomb PA, et al. Risk of recurrence after ductal carcinoma in situ of the breast. *Cancer Epidemiol Biomarkers Prev.* 1998;7(8):689-696.
30. McLaughlin VH, Trentham-Dietz A, Hampton JM, Newcomb PA, Sprague BL. Lifestyle factors and the risk of a second breast cancer after ductal carcinoma in situ. *Cancer Epidemiol Biomarkers Prev.* 2014;23(3):450-460.
31. Pang JB, Dobrovic A, Fox SB. DNA methylation in ductal carcinoma in situ of the breast. *Breast Cancer Res.* 2013;15(3):206.
32. Hwang ES, DeVries S, Chew KL, et al. Patterns of chromosomal alterations in breast ductal carcinoma in situ. *Clin Cancer Res.* 2004;10(15):5160-5167.
33. Reis-Filho JS, Simpson PT, Gale T, Lakhani SR. The molecular genetics of breast cancer: the contribution of comparative genomic hybridization. *Pathol Res Pract.* 2005;201(11):713-725.
34. Gorringe KL, Hunter SM, Pang J-M, et al. Copy number analysis of ductal carcinoma in situ with and without recurrence. *Mod Pathol.* 2015;28(9):1174-1184.
35. Pang J-MB, Savas P, Fellowes AP, et al. Breast ductal carcinoma in situ carry mutational driver events representative of invasive breast cancer. *Mod Pathol.* 2017;30(7):952-963.
36. Johnson CE, Gorringe KL, Thompson ER, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast Cancer Res Treat.* 2012;133(3):889-898.
37. Sakr R a, Weigelt B, Chandarlapaty S, et al. PI3K Pathway Activation in High-Grade Ductal Carcinoma In Situ--Implications for Progression to Invasive Breast Carcinoma. *Clin Cancer Res.* April 2014:2326-2337.
38. Miron A, Varadi M, Carrasco D, et al. PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res.* 2010;70(14):5674-5678.
39. Done SJ, Eskandarian S, Bull S, Redston M, Andrulis IL. p53 missense mutations in microdissected high-grade ductal carcinoma in situ of the breast. *J Natl Cancer Inst.* 2001;93(9):700-704.
40. Zhou W, Muggerud AA, Vu P, et al. Full sequencing of TP53 identifies identical mutations within in situ and invasive components in breast cancer suggesting clonal evolution. *Mol Oncol.* 2009;3(3):214-219.
41. Vincent-Salomon A, Lucchesi C, Gruel N, et al. Integrated genomic and transcriptomic analysis of ductal carcinoma in situ of the breast. *Clin Cancer Res.* 2008;14(7):1956-1965.
42. Abba MC, Gong T, Lu Y, et al. A Molecular Portrait of High-Grade Ductal Carcinoma In Situ. *Cancer Res.* 2015;75(18):3980-3990.
43. Knudsen ES, Ertel A, Davicioni E, Kline J, Schwartz GF, Witkiewicz AK. Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. *Breast Cancer Res Treat.* 2012;133(3):1009-1024.
44. Lee S, Stewart S, Nagtegaal I, et al. Differentially expressed genes regulating the progression of ductal carcinoma in situ to invasive breast cancer. *Cancer Res.* 2012;72(17):4574-4586.
45. Hannemann J, Velds A, Halfwerk JBG, Kreike B, Peterse JL, van de Vijver MJ. Classification of ductal carcinoma in situ by gene expression profiling. *Breast Cancer Res.* 2006;8(5):R61.
46. Muggerud AA, Hallett M, Johnsen H, et al. Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Mol Oncol.* 2010;4(4):357-368.
47. Elias EV, de Castro NP, Pineda PHB, et al. Epithelial cells captured from ductal carcinoma in situ reveal a gene expression signature associated with progression to invasive breast cancer. *Oncotarget.*

- September 2016.
48. Doebar SC, Sieuwerts AM, de Weerd V, Stoop H, Martens JWM, van Deurzen CHM. Gene Expression Differences between Ductal Carcinoma in Situ with and without Progression to Invasive Breast Cancer. *Am J Pathol.* 2017;187(7):1648-1655.
 49. Schultz S, Bartsch H, Sotlar K, et al. Progression-specific genes identified in microdissected formalin-fixed and paraffin-embedded tissue containing matched ductal carcinoma in situ and invasive ductal breast cancers. *BMC Med Genomics.* 2018;11(1):80.
 50. Solin LJ, Gray R, Baehner FL, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. *J Natl Cancer Inst.* 2013;105(10):701-710.
 51. Solin LJ, Gray R, Hughes LL, et al. Surgical excision without radiation for ductal carcinoma in situ of the breast: 12-year results from the ECOG-ACRIN E5194 study. *J Clin Oncol.* 2015;33(33):3938-3944.
 52. Rakovitch E, Nofech-Mozes S, Hanna W, et al. A population-based validation study of the DCIS Score predicting recurrence risk in individuals treated by breast-conserving surgery alone. *Breast Cancer Res Treat.* 2015;152(2):389-398.

Supplementary tables and figures

Table S1. Clinicopathological characteristics of women diagnosed with DCIS and treated by BCS alone, who subsequently did (DCIS cases) or did not (DCIS controls) develop subsequent iIBC

Characteristics	DCIS cases (n = 101)		DCIS controls (n = 82)	
	n	(%)	n	(%)
Age at DCIS diagnosis (years)				
<40	8	(7.9)	5	(6.1)
40-49	12	(11.9)	10	(12.2)
50-59	42	(41.6)	36	(43.9)
60-69	27	(26.7)	22	(26.8)
70-79	9	(8.9)	7	(8.5)
≥80	3	(3.0)	2	(2.4)
Clinical presentation of DCIS				
Detected via screening	48	(47.5)	39	(47.6)
Nonscreening-related	44	(43.6)	34	(41.5)
Unknown	9	(8.9)	9	(11.0)
Histological grade				
Well differentiated	14	(13.9)	11	(13.4)
Intermediately differentiated	63	(62.4)	51	(62.2)
Poorly differentiated	24	(23.8)	20	(24.4)
ER status				
Negative	19	(18.8)	14	(17.1)
Positive	79	(78.2)	66	(80.5)
N/A	3	(3.0)	2	(2.4)
HER2 status				
Negative	69	(68.3)	57	(69.5)
Positive	29	(28.7)	20	(24.4)
N/A	3	(3.0)	5	(6.1)
Time to iIBC, mean in years (range)	6.4	(0.5-17.0)	-	

N/A, not assessable

Table S2. Comparison of the 200 cases and 361 controls included in the CCR 2018 case-control study with the cases and controls included in this study

Cases					
Characteristics	Included in CCR 2018		Included in study		P ^a
	(n = 200)		(n = 101)		
	n	(%)	n	(%)	
Age at DCIS diagnosis (years)					
<40	14	(7.0)	8	(7.9)	
40-49	27	(13.5)	12	(11.9)	
50-59	79	(39.5)	42	(41.6)	
60-69	55	(27.5)	27	(26.7)	
70-79	19	(9.5)	9	(8.9)	
≥80	6	(3.0)	3	(3.0)	1.00
Year of DCIS diagnosis, mean (range)	1996	(1989 - 2004)	1996	(1989 - 2004)	0.38
Period of DCIS diagnosis					
1989-1998 (implementation phase)	147	(73.5)	75	(74.3)	
1999-2004 (full nationwide coverage)	53	(26.5)	26	(25.7)	0.91
Clinical presentation of DCIS					
Screen-detected	96	(48.0)	48	(47.5)	
Symptomatic	89	(44.5)	44	(43.6)	1.00
Unknown	15	(7.5)	9	(8.9)	

Controls					
Characteristics	Included in CCR 2018		Included in study		P ^a
	(n = 361)		(n = 82)		
	n	(%)	n	(%)	
Age at DCIS diagnosis (years)					
<40	19	(5.3)	5	(6.1)	
40-49	47	(13.0)	10	(12.2)	
50-59	146	(40.4)	36	(43.9)	
60-69	105	(29.1)	22	(26.8)	
70-79	33	(9.1)	7	(8.5)	
≥80	11	(3.0)	2	(2.4)	0.99
Year of DCIS diagnosis, mean (range)	1997	(1989 - 2004)	1996	(1989 - 2004)	0.14
Period of DCIS diagnosis					
1989-1998 (implementation phase)	246	(68.1)	62	(75.6)	
1999-2004 (full nationwide coverage)	115	(31.9)	20	(24.4)	0.18
Clinical presentation of DCIS					
Screen-detected	186	(51.5)	39	(47.6)	
Symptomatic	141	(39.1)	34	(41.5)	0.61
Unknown	34	(9.4)	9	(11.0)	

^a: For continuous variables, the p-value was calculated by paired T-test, and for categorical variables the p-value was calculated by chi-square test; Unknown clinical presentation was not included in the analysis; CCR 2018, Visser, et al. Clin Can Res 2018.

Table S3. Comparison of mutations found in DCIS associated with (cases) and without (controls) subsequent iIBC

	Cases (n = 81)		Controls (n = 47)		P	
	n	(%)	n	(%)		
<i>PIK3CA</i>	18	(22.2)	17	(36.2)	0.041	^a
<i>TP53</i>	17	(21)	12	(25.5)	0.39	^a
<i>NF1</i>	7	(8.6)	3	(6.4)	1.00	^b
<i>BRCA2</i>	8	(9.9)	2	(4.3)	0.49	^b

Comparisons between DCIS cases and controls were made by chi2 test (a) and Fisher's exact test (b).

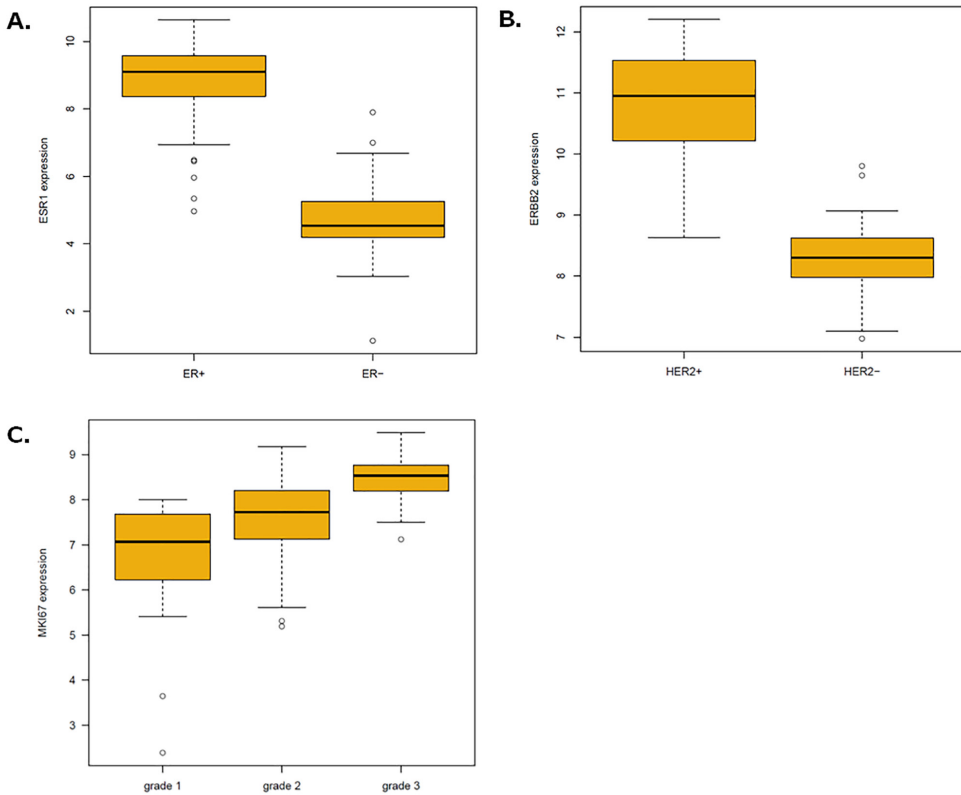


Figure S1. Identification of samples for exclusion using clustering and principle component analysis (PCA). **A.** Clustering of all DCIS samples, showing outliers at right and left side of the dendrogram. **B.** PCA plot before exclusion of the outlier samples, Red dotted line represents cutoff for exclusion. **C.** PCA plot after exclusion of outlier samples. Outliers: A0006, A0025, A0040, A0049, A0061, A0062, A0069, A0073, A0087, A0095, A0096, A0147, A0149, A0153, A0202.

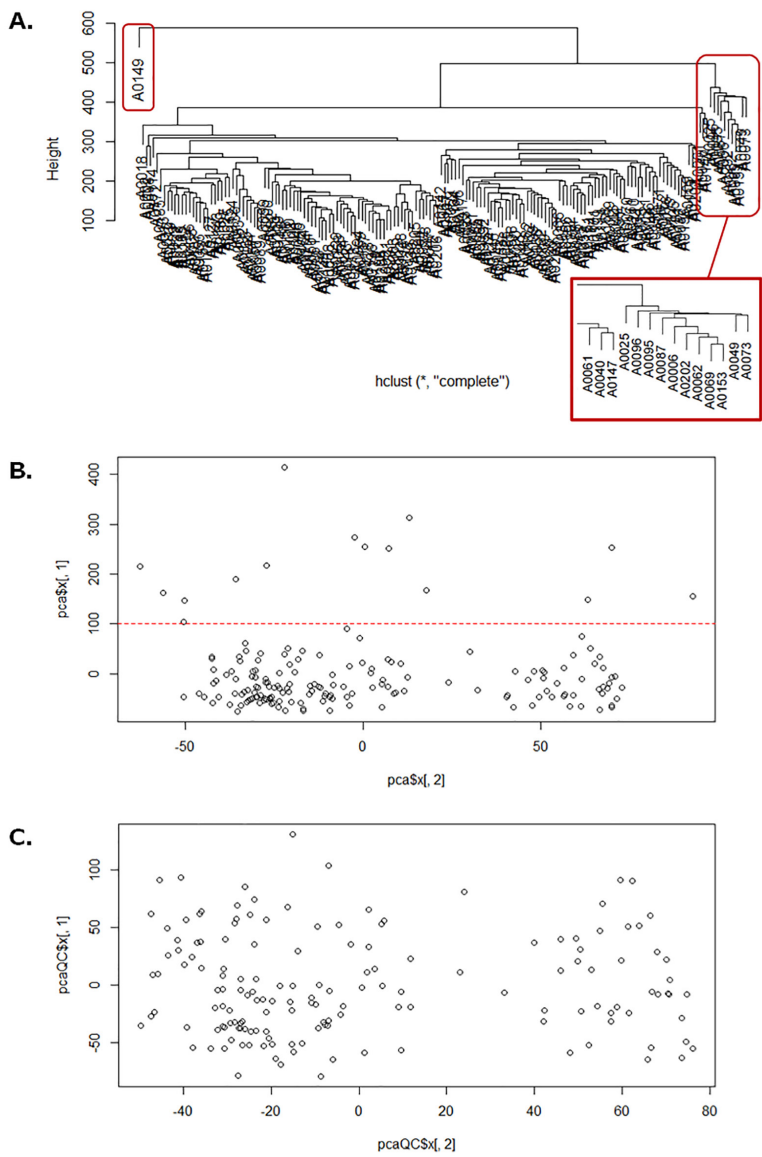


Figure S2. Protein expression versus gene expression. **A.** ER protein vs. ESR1 gene; **B.** HER2 protein vs. ERBB2 gene; **C.** DCIS histologic grade versus MKI67 gene. Abbreviations: ER, Estrogen receptor.

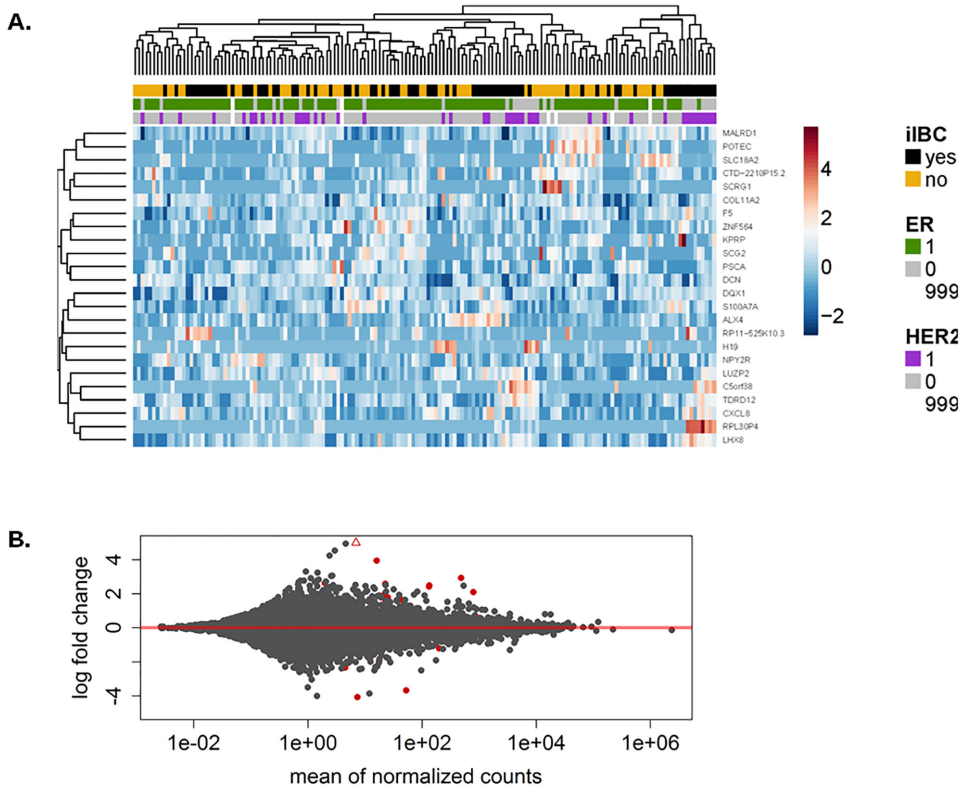


Figure S3. Visualization of differentially expressed genes between DCIS with (cases) and without subsequent ipsilateral invasive breast cancer (iIBC; controls). **A.** Unsupervised clustering of differentially expressed genes; **B.** MA plot with in red the genes with a differential expression. $P < 0.1$

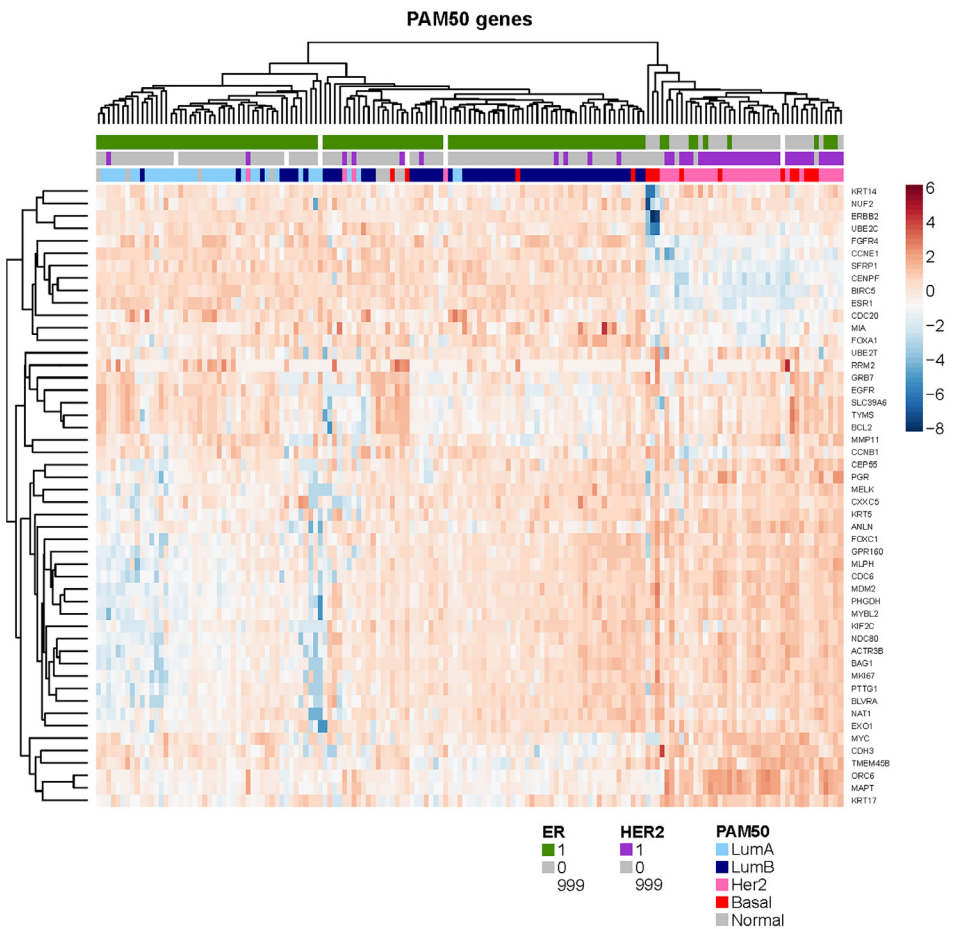
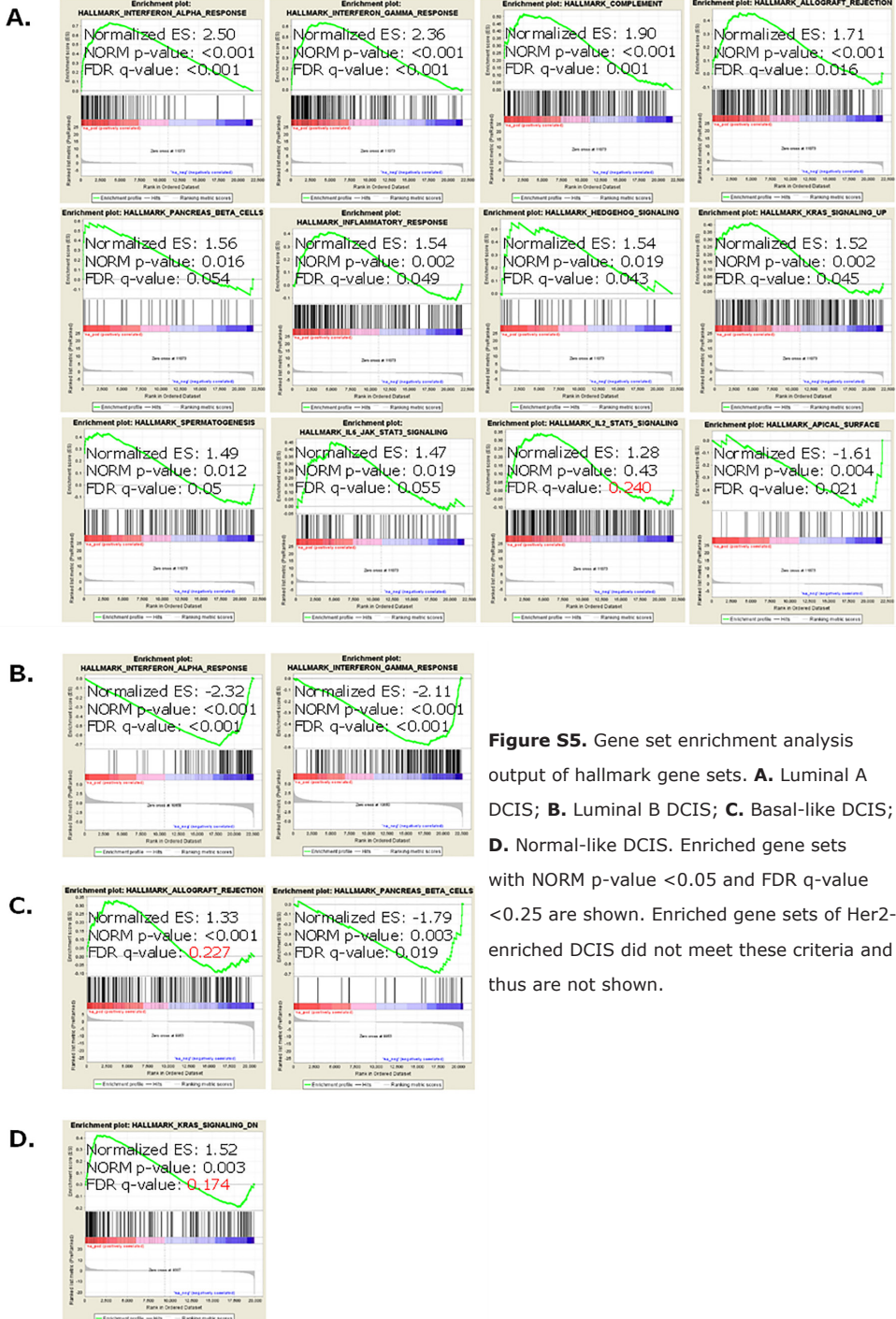


Figure S4. Classification of DCIS lesions into the PAM50 subtypes. The heatmap shows the expression of the 50 genes of the PAM50 classifier. Green and purple annotation bars give information on the ER and HER2 status of the DCIS lesions.



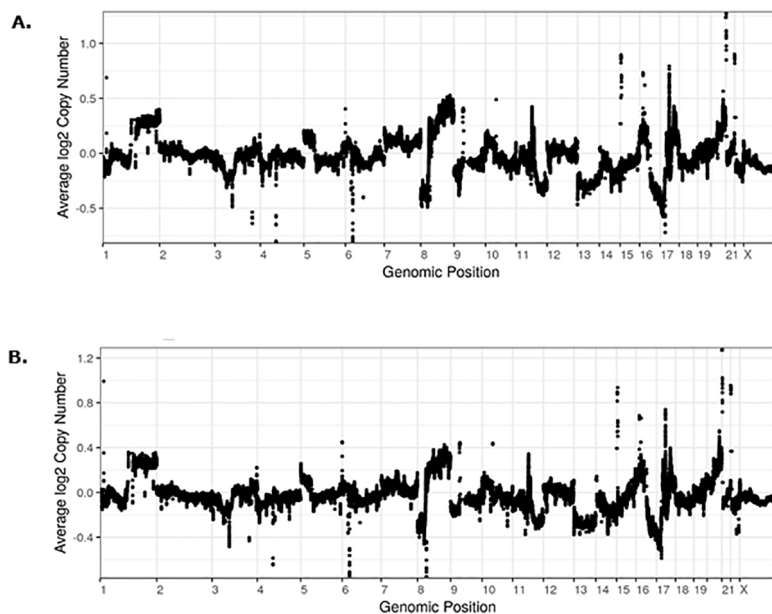


Figure S6. Mean copy number profiles of DCIS associates with subsequent iIBC (**A**) and DCIS without subsequent iIBC (**B**).

Lindy L. Visser
Lotte E. Elshof
Koen K. van de Vijver
Emma J. Groen
Mathilde M. Almekinders
Joyce Sanders
Carolien Bierman
Dennis Peters
Ingrid Hofland
Annegien Broeks
Flora E. van Leeuwen
Emiel J. Th. Rutgers
Marjanka K. Schmidt
Michael Schaapveld
Esther H. Lips
Jelle Wesseling

CHAPTER 5

Discordant marker expression between invasive breast carcinoma and corresponding synchronous and preceding DCIS

Abstract

Ductal carcinoma in situ (DCIS) is considered a potential precursor of invasive breast carcinoma (IBC). Studies aiming to find markers involved in DCIS progression generally have compared characteristics of IBC lesions with those of adjacent *synchronous* DCIS lesions. The question remains whether *synchronous* DCIS and IBC comparisons are a good surrogate for *primary* DCIS and *subsequent* IBC. In this study, we compared both primary DCIS and synchronous DCIS with the associated IBC lesion, based on immunohistochemical marker expression. Immunohistochemical analysis of ER, PR, HER2, p53, and cyclo-oxygenase 2 (COX-2) was performed for 143 primary DCIS and subsequent IBC lesions, including 81 IBC lesions with synchronous DCIS. Agreement between DCIS and IBC were assessed using kappa, and symmetry tests were performed to assess the pattern in marker conversion. The primary DCIS and subsequent IBC more often showed discordant marker expression than synchronous DCIS and IBC. Strikingly, 18 out of 49 (36%) women with HER2-positive primary DCIS developed a HER2-negative IBC. Such a difference in HER2 expression was not observed when comparing synchronous DCIS and IBC. The frequency of discordant marker expression did not increase with longer time between primary DCIS and IBC. In conclusion, comparison of primary DCIS and subsequent IBC yields different results than comparison of synchronous DCIS and IBC, in particular regarding HER2 status. To gain more insight into the progression of DCIS to IBC, it is essential to focus on the relationship between *primary* DCIS and *subsequent* IBC, rather than comparing IBC with *synchronous* DCIS.

Keywords: Ductal carcinoma in situ; invasive breast carcinoma; local recurrence; synchronous lesions; immunohistochemistry

Introduction

Ductal carcinoma in situ (DCIS) is generally accepted as a non-obligate precursor of invasive breast carcinoma (IBC).¹ This because they are frequently found next to each other sharing the genetic alterations as well as risk factors (e.g. age, family history of breast carcinoma, etc).²⁻⁶ While DCIS itself is not life-threatening, it does increase a woman's risk of developing IBC later in life, which subsequently could lead to breast cancer-specific death.⁷ To prevent progression to invasive disease, almost all DCIS lesions are treated by mastectomy or breast conserving surgery with or without adjuvant radiotherapy and/or endocrine therapy.

If it holds true that DCIS directly progresses to IBC, one would expect that primary DCIS and subsequent ipsilateral IBC share multiple features, for example hormone receptor and HER2 status. It has been shown that the histological grade of the DCIS component adjacent to invasive disease (*synchronous* DCIS) and the grade of the IBC lesion are significantly correlated, i.e. well differentiated DCIS relates to grade I IBC and poorly differentiated DCIS to grade III IBC.^{8,9} Therefore, it is thought that, if progression occurs, well differentiated DCIS will give rise to grade I IBC and poorly differentiated DCIS to grade III IBC. Nonetheless, comparison of marker expression shows conflicting results. Allred, et al. found that HER2 overexpression is more frequently observed in DCIS than in IBC.¹⁰ Multiple studies have tried to address this, mostly by comparing synchronous DCIS and IBC.^{3,11-13} Many of these studies showed that synchronous DCIS and IBC components were, however, very similar on the quantitative level and discordant HER2 status was rarely observed.

Several studies aimed to find markers involved in DCIS progression to IBC, mostly by comparing IBC lesions and an adjacent DCIS component, referred to as *synchronous* DCIS.¹⁴ However, to our knowledge, it has never been investigated whether the *synchronous* DCIS and IBC comparisons are a good surrogate for primary DCIS and *subsequent* IBC. Therefore, we performed a comparative analysis between primary DCIS and subsequent ipsilateral IBC, and between this IBC and the adjacent synchronous DCIS component based on immunohistochemical marker expression (Figure 1). With this, we aimed to (1) assess the concordance in marker expression between primary DCIS and subsequent ipsilateral IBC, and IBC and synchronous DCIS and (2) to identify factors that may explain the potential discordance in marker expression.

Patients and Methods

Study population and design

The source population was derived from the Netherlands Cancer Registry (NCR), linked to the nation-wide network and registry of histology and cytopathology in the Netherlands (PALGA). It included all women diagnosed with primary DCIS within the Netherlands and

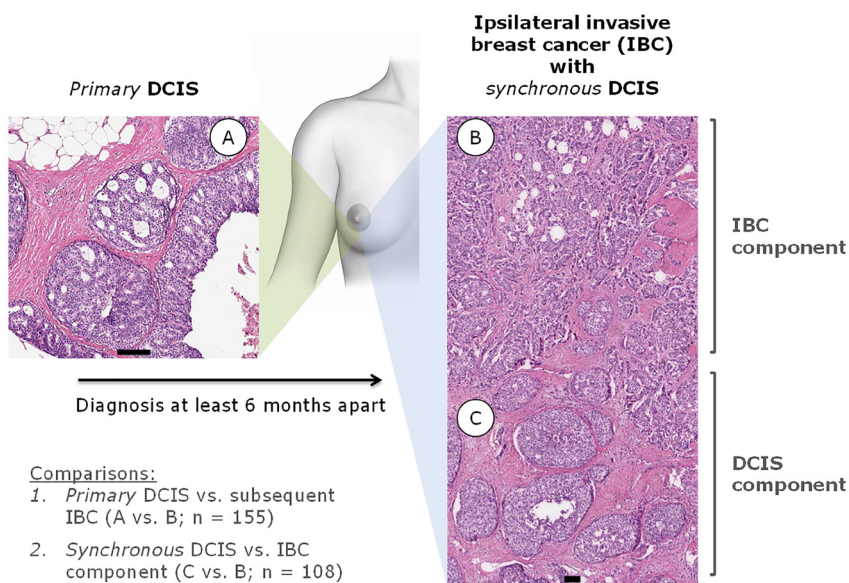


Figure 1. Set-up of the study and the number of included lesions. Out of 155, there were 108 IBC lesions that harbored a synchronous DCIS component.

treated with breast conserving surgery alone between January 1, 1989 to December 31, 2004 (n=2,658).⁷ The NCR provided information on age at diagnosis, date of diagnosis, treatment of DCIS, and the development of subsequent IBC. Follow-up and vital status were complete until at least January 1, 2011. The median follow-up was 12.0 years (interquartile range, 9.0-15.3). In total, 374 of 2,658 women developed ipsilateral IBC as first invasive carcinoma after a primary diagnosis of DCIS.⁷ At the tissue block collection for this current study, the first 328 women with a subsequent iIBC had been identified.

Formalin-fixed paraffin-embedded (FFPE) tissue samples from both the primary DCIS and the corresponding subsequent IBC were requested from 58 pathology laboratories within the Netherlands. All specimens were histopathologic re-examined by a team of consultant breast pathologists (JW, EJG, KvdV), using new hematoxylin and eosin-stained whole slides. Slides were assessed for histological grade: DCIS was graded according to the World Health Organization criteria (WHO; 2012).¹⁵ IBC was graded according to Elston/Nottingham modification of the Bloom-Richardson system, based on tumor tubule formation, number of mitotic figures, and nuclear polymorphism.¹⁶ We were unable to collect tissue blocks from some hospitals, either because the hospital refused to provide tissue for research or because tissue blocks were unavailable (n=61 pairs excluded). Women for whom DCIS and/or IBC diagnoses could not be confirmed by pathology were excluded from this study (n=59), as

were DCIS including a (micro)invasive component or LCIS (lobular carcinoma in situ; n=53). This resulted in 155 patient-matched primary DCIS and subsequent IBC pairs.

The study was approved by the review board on the NKR (request K12.281; January 3, 2013) and PALGA (LZV990; April 16, 2013). Secondary use of tissue and data for this study was done under an opt-out regime which is conform Dutch regulations and the Code of Conduct of Federa-COREON.

Immunohistochemistry

Both primary DCIS and the corresponding invasive breast carcinoma were immunohistochemically stained for estrogen receptor (ER; SP1, ready-to-use, Ventana Medical Systems) and progesterone receptor (PR; 1E2, ready-to-use, Ventana Medical Systems) status, overexpression of HER2 (4B5, ready-to-use, Ventana Medical Systems), Ki67 (MIB1, 1/250, DAKO), and expression of tumor suppressor protein p53 (DO-7, 1/7000, DAKO), and COX-2 (CX294, 1/100, DAKO), using a Benchmark ULTRA autostainer (Ventana Meducan Systems, AZ, USA), using 3- μ m thick whole slides. Some DCIS tissue specimens were excluded because insufficient tissue was available for immunohistochemistry (IHC; n=13). Details on the IHC staining procedure can be found elsewhere.¹⁷ Positive and negative controls were included in all staining runs.

IHC assessment was performed by a team of 7 observers, including 5 pathologists. Assessment of inter-observer agreement is described elsewhere.¹⁷ For ER, PR, HER2, Ki67 and p53 a concordance was reached of 97% with an interclass correlation coefficient (ICC) of >0.8. Ki67 was excluded for further scoring because of unreliable staining results. COX-2 reached a concordance of 94% with a *k* statistic of 0.7. ER and PR status were considered positive when $\geq 10\%$ of the luminal epithelial cells showed nuclear staining, based on scoring guidelines by FCCC pathologists from 2007-2011. Similarly, p53-positive staining was assessed based on the percentage of cells that showed moderate to strong nuclear staining. The presence of >70% positive cells or complete lack of p53 expression was considered as mutant p53 expression, and 1-70% positive cells was considered wild-type (WT) p53 expression.^{18,19} HER2 overexpression was assessed according to the American Society of Clinical Oncology and College of American Pathologist (ASCO-CAP) 2013 recommendations.²⁰ A membrane score of 3+ was considered as HER2 positive, as was a membrane score of 2+ for which the overexpression could be confirmed by HER2 CISH (chromogenic in situ hybridization). A membrane score of 0 and 1+ was considered HER2 negative. COX-2 expression was evaluated according to criteria adapted from Ristimäki et al: 1= weak diffuse cytoplasmic staining that may contain moderate to strong granular cytoplasmic staining in less than 10% of tumor cells; 2= moderate to strong granular cytoplasmic staining in 10-90% of the tumor cells; 3= moderate to strong granular cytoplasmic staining in over 90% of the tumor cells. Score 1 was considered low COX-2 expression and score 2 and 3 were considered as high expression of COX-2.²¹

DCIS and IBC lesions were categorized into the following IHC subtypes: hormone receptor (HR)+ HER2-, HR+ HER2+, HR- HER2+, and HR- HER2-. Lesions were classified as HR+ when ER and/or PR was scored as positive. Lesions were classified as HR- when ER and PR both were negative.

To assess intra-lesional heterogeneity within our study cohort, we randomly selected 10 DCIS ducts within one primary DCIS lesion and independently scored the IHC markers for these ducts. Subsequently, we used the ER, PR, and HER2 status to classify the DCIS duct into the IHC subtypes. When there were different IHC subtypes present within a single DCIS lesions, we considered these DCIS lesions heterogeneous, and the DCIS ducts with different IHC subtypes were considered separate DCIS subclones.

Statistical Analysis

Concordance of marker expression was determined by calculating the agreement between DCIS and IBC samples using kappa. In addition, symmetry tests were performed to assess whether there was a pattern in the marker conversion in case of discordant expression. For variables with two categories an asymptotic symmetry test was performed. When a variable was composed of more than two categories then a marginal homogeneity test was used. Chi-square tests were performed to compare time to event and the frequency of discordant marker expression. Time to event was defined as a categorical variable i.e. the study group was divided into two groups at the median time to ipsilateral IBC diagnosis: ≤ 6.1 vs. > 6.1 years. These two groups were compared with regard to the frequency of marker expression concordance. All statistical analyses were performed using Stata/SE (version 13.1, StataCorp). *P* values ≤ 0.05 were considered statistically significant.

Results

Baseline characteristics

155 DCIS patients with a subsequent ipsilateral IBC were included in this study, which is a good representation of all 328 patients that developed an invasive recurrence after primary DCIS within our cohort (Supplementary Table S1). 108 out of these 155 lesions (69.7%) had a DCIS component adjacent to the invasive disease (synchronous DCIS). The mean time to invasive recurrence was 6.3 years (range 0.5 to 17.0 years). 79.4% of the invasive recurrences occurred in the same quadrant as the initial DCIS (77 of 97; 58 pairs location not specified; Supplementary Table S2).

The primary DCIS was well differentiated in 28 lesions (18.1%), intermediately differentiated in 94 lesions (60.6%) and poorly differentiated in 33 lesions (21.3%). Of the subsequent invasive breast carcinomas, 131 (84.5%) were invasive carcinoma of no special type, 10 (6.5%) were invasive lobular carcinoma and 14 (9.0%) were mixed.

Immunohistochemical staining was performed for 142 out of 155 primary DCIS and

subsequent IBC pairs and 81 out of 108 IBC and synchronous DCIS pairs. The frequency of ER, PR, and COX-2 positivity was similar in primary DCIS, IBC, and synchronous DCIS (Table 1). HER2 was overexpressed in 34.5% of primary DCIS, 26.1% of IBC, and 24.7% of synchronous DCIS. In line with previous literature, low-grade lesions (both DCIS and IBC) were associated with ER and PR positivity and low expression of HER2, p53, and COX-2.^{22,23} High-grade lesions were often ER and PR negative, frequently overexpressed HER2 and p53, and had high expression of COX-2 (Supplementary Table S3).

Table 1. Overview of immunohistochemical marker expression of primary DCIS, invasive breast cancer, and synchronous DCIS

Characteristics	Primary DCIS (n = 142)		Invasive BC (n = 142)		Synchronous DCIS (n = 81)	
	n	(%)	n	(%)	n	(%)
IHC subtype						
HR+ HER2-	90	(63.4)	96	(67.6)	56	(69.1)
HR+ HER2+	28	(19.7)	23	(16.2)	8	(9.9)
HR- HER2+	21	(14.8)	14	(9.9)	12	(14.8)
HR- HER2-	3	(2.1)	9	(6.3)	5	(6.2)
ER						
Negative	24	(16.9)	23	(16.2)	17	(21.0)
Positive	118	(83.1)	119	(83.8)	63	(77.8)
N/A	0	(0.0)	0	(0.0)	1	(1.2)
PR						
Negative	51	(35.9)	58	(40.8)	31	(38.3)
Positive	90	(63.4)	84	(59.2)	50	(61.7)
N/A	1	(0.7)	0	(0.0)	0	(0.0)
HER2						
Negative	93	(65.5)	105	(73.9)	61	(75.3)
Positive	49	(34.5)	37	(26.1)	20	(24.7)
p53						
0% positive cells (mutant)	12	(8.5)	18	(12.7)	6	(7.4)
1-70% positive cells (WT)	108	(76.1)	101	(71.2)	65	(80.2)
>70% positive cells (mutant)	21	(14.8)	23	(16.2)	10	(12.4)
N/A	1	(0.7)	0	(0.0)	0	(0.0)
COX-2						
Low	16	(11.3)	17	(12.0)	13	(16.0)
High	125	(88.0)	124	(87.3)	67	(82.7)
N/A	1	(0.7)	1	(0.7)	1	(1.2)

*: HER2 positive group includes all synchronous DCIS lesions with membrane score 3+; WT: wild-type

Discordant marker expression is more frequently observed between primary DCIS and subsequent IBC as compared to synchronous DCIS and IBC

To assess the frequency of discordant marker expression, we compared IHC marker expression between primary DCIS and subsequent IBC pairs and synchronous IBC and DCIS pairs (Table 2 and 3). IBC and synchronous DCIS were discordant for ER, PR, HER2, p53, and COX-2 marker expression in 5.0%, 19.7%, 4.9%, 12.3%, and 12.5% of the pairs, respectively (Table 2 and 4). For all five markers, conversion was observed both from positive/high to negative/low as well as from negative/low to positive/high. Four IBC and synchronous DCIS pairs consisted of ER positive IBC and an ER negative synchronous DCIS component ($p = 0.046$; Table 2), although this number is too small to draw any conclusions from. 9.9% of the synchronous DCIS and IBC pairs did not share the same IHC subtype (Supplementary Table S4).

Marker expression of primary DCIS and the subsequent ipsilateral IBC were discordant for ER, PR, HER2, p53, and COX-2 expression in 12.0%, 27.7%, 16.9%, 22.7%, and 19.3% of the pairs, respectively (Table 3 and 4). Notably, 18 out of 49 patients (36%) had a HER2-positive primary DCIS which was followed by HER2-negative IBC (symmetry $p=0.014$; Table 3). This also caused a conversion of the IHC subtype in that 33 of the 142 primary DCIS and IBC pairs (23.2%) did not share the same IHC subtype (symmetry $p=0.040$; Supplementary Table S5), which was irrespective of the presence of synchronous DCIS adjacent to the IBC. The discordant marker expression rates between primary DCIS and synchronous DCIS were comparable to those of primary DCIS and subsequent IBC (Table 4; Supplementary Table S6-7).

Table 2. Marker expression of synchronous DCIS related to invasive component

		Invasive component (n)		Agreement	Symmetry
		Negative/ Low/WT	Positive/ High/Mutant	(%)	P
Synchronous DCIS (n)					
ER	Negative	13	4	95.0	0.046
	Positive	0	63		
PR	Negative	25	6	80.3	0.32
	Positive	10	40		
HER2	Negative	58	3	95.1	0.32
	Positive	1	19		
P53	WT	57	8	87.7	0.06
	Mutant	2	14		
COX-2	Low	7	6	87.5	0.53
	High	4	63		

Agreement was calculated by non-weighted kappa; p -values were calculated by asymptotic symmetry test. Total number of pairs included: ER n=80; PR n=81; HER2 n=142; p53 n=140; COX-2 n=140. ER/PR positive: >10% positive cells; HER2 positive: membrane score 3 or membrane score 2 confirmed by CISH; p53 wild-type (WT): 1-70% positive cells; p53 mutant: >70% positive cells or complete lack of p53 expression; COX-2 high: score 2-3.

Table 3. Immunohistochemical marker expression of primary DCIS related to subsequent invasive breast cancer

		Invasive breast cancer (n)		Agreement	Symmetry
		Negative/ Low/WT	Positive/ High/Mutant	(%)	P
Primary DCIS (n)					
ER	Negative	15	9	88.0	0.81
	Positive	8	110		
PR	Negative	35	16	72.3	0.26
	Positive	23	67		
HER2	Negative	87	6	83.1	0.014
	Positive	18	31		
p53	WT	88	29	77.3	0.16
	Mutant	12	21		
COX-2	Low	3	13	80.7	0.85
	High	14	110		

Agreement was calculated by non-weighted kappa; *p*-values were calculated by asymptotic symmetry test
 Total number of pairs included: ER n=142; PR n=141; HER2 n=142; p53 n=140; COX-2 n=140. ER/
 PR positive: >10% positive cells; HER2 positive: membrane score 3 or membrane score 2 confirmed by
 CISH; p53 wild-type (WT): 1-70% positive cells; p53 mutant: >70% positive cells or complete lack of p53
 expression; COX-2 high: score 2-3.

Table 4. Percentage discordance in grade and marker expression between pairs of DCIS and IBC

	sDCIS vs. IBC		pDCIS vs. IBC		pDCIS vs. sDCIS	
	Discordance (%)	Symmetry <i>P</i> -value	Discordance (%)	Symmetry <i>P</i>	Discordance (%)	Symmetry <i>P</i>
IHC						
ER	5.0	0.046	12.0	0.81	13.7	0.76
PR	19.7	0.32	27.7	0.26	29.6	0.68
HER2	4.9	0.32	16.9	0.014	13.6	0.007
p53	12.3	0.06	22.7	0.16	22.5	0.64
COX-2	12.5	0.53	19.3	0.85	25.0	0.37
Subtype	9.9	0.09	23.2	0.040	16.0	0.17

pDCIS: primary DCIS; sDCIS: synchronous DCIS.

Discordance of marker expression is not associated with time to event

When comparing primary DCIS with a subsequent IBC, a time factor is obviously present, which is absent when comparing synchronous DCIS and adjacent IBC. Thus, we questioned whether time to event could play a role in the higher frequency of discordant marker expression between primary DCIS and subsequent IBC. For this, the study group was divided by the median time to IBC. While 56.9% of the women that developed IBC within 6.1 years after their DCIS diagnosis showed discordant marker expression between primary DCIS and subsequent IBC involving at least one IHC marker, this was 64.3% in the

group of women that developed IBC after more than 6.1 years after their DCIS diagnosis (p=0.37).

These data suggest that the probability of discordant marker expression between the primary DCIS and subsequent IBC does not increase with longer time to IBC.

Specific subclones might be responsible for the invasive outgrowth

From previous studies we know that this intra-lesional heterogeneity already exists at the DCIS stage.^{24,25} Discordant marker expression could be caused by heterogeneity within the DCIS lesion. Therefore, we assessed IHC staining in ten 10 individual ducts per DCIS lesions. In 10 out of 94 DCIS lesions (10.6%) we observed heterogeneity, defined by the presence of multiple IHC subtypes, or subclones, within one DCIS lesion. We compared the IHC subtypes of the DCIS subclones, to the IHC subtype of the subsequent IBC lesion (Figure 2). In nine of the 10 pairs, the subtype of the IBC lesion was shared with a subclone of the DCIS lesion. In two of these DCIS lesions, four different IHC subtypes were present and seven DCIS lesions consisted of two different IHC subtypes. In one pair, the subtype of the IBC lesion was not shared with any of the DCIS subclones.

These results show that intra-lesional heterogeneity exists within DCIS lesions. This may be causative for the discordant marker expression between DCIS and IBC.

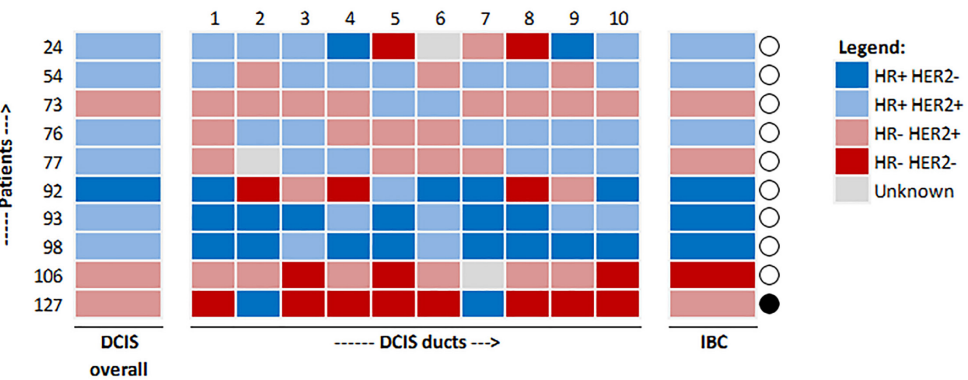


Figure 2. Matrix table of the IHC subtypes of 10 DCIS and subsequent IBC pairs in which intra-lesional heterogeneity was found in the DCIS lesion. For every DCIS lesion 10 individual DCIS ducts were assessed for IHC subtype using hormone receptor (HR) and HER2 status. On the left, columns are individual DCIS ducts (n=10); rows are individual patients. On the left, the overall IHC subtype of the DCIS lesion is shown. On the right, the IHC subtype of the corresponding IBC lesion is shown. Circle without fill indicates pairs of which the subtype of the IBC lesion was shared with a subclone of the DCIS lesion. Circle with black fill indicates a pair of which the subtype of the IBC lesion was not shared with any of the DCIS subclones.

Discussion

In this study, we demonstrated that comparative analysis between primary DCIS and subsequent ipsilateral invasive breast carcinoma (IBC) versus IBC and adjacent synchronous DCIS yields different results. This was most prominently illustrated for HER2, as we found that 36% of HER2-positive primary DCIS lesions were followed by HER2-negative IBC. Such a difference was not observed in our comparison between synchronous DCIS and IBC.

Our finding that HER2-negative IBC is preceded by HER2-positive DCIS is challenging our current understanding of the role of HER2 in the progression of DCIS to invasive breast carcinoma, as HER2 has been described as a predictor of recurrence after DCIS.^{17,26,27} If the overexpression of HER2 plays a major role in DCIS progression, the overexpression of HER2 in IBC might be expected to be equal or exceeding the level of the preceding DCIS. This emphasizes that the role of HER2 in progression of DCIS to IBC remains to be elucidated. It could be hypothesized that HER2-overexpression promotes a higher proliferative rate, but does not lead a higher invasive potential of DCIS. Bijker, et al. and Karlsson, et al. performed the same comparison, but the number of matched DCIS and IBC pairs in these two studies were too small to notice the major findings presented here.^{22,28}

Our group recently reported that high COX-2 expression was strongly associated with development of subsequent IBC.¹⁷ In the current study, we showed that the level of COX-2 expression is almost similar when comparing primary DCIS and subsequent IBC. This may suggest that COX-2 could play a role in the invasive outgrowth of DCIS.

Overall, marker expression between *primary* DCIS and *subsequent* IBC was less concordant than *synchronous* DCIS and IBC. Yet, the frequency of discordant marker expression between primary DCIS and subsequent IBC did not increase with longer time to IBC. However, we found that DCIS intra-lesional heterogeneity exists in DCIS lesions, suggesting that discordant marker expression may be caused by heterogeneity. This makes it plausible that subsequent invasive disease arises from only one or just a few of these subclones. We were unable to assess multiple ducts of synchronous DCIS, as our intra-lesional heterogeneity analysis was based on lesions with at least 10 ducts. Yet, there is no reason to assume that our findings on heterogeneity do not apply to synchronous DCIS. Obviously, it is highly likely that synchronous DCIS has the potential to become invasive, as the DCIS is present in or in the close proximity of the IBC. This might imply that the DCIS subclone responsible for the invasive outgrowth is just a small minority subclone within this synchronous DCIS lesions or that it is even not existing anymore after the invasive outgrowth. As a consequence, one could argue that the comparative analyses of *synchronous* DCIS and IBC is biased and thus the preferred analysis is the comparison of *primary* DCIS and *subsequent* IBC.

In the current study, we used a 10% cut-off for ER and PR positivity, as established in the guidelines by FCCC pathologists. Using a 1% cut-off would have minimal impact on the

results of ER (concordance 10% cut-off, 88.1%; 1% cutoff, 86.1%; Supplementary Figure S1) and moderate impact on the results of PR (concordance 10% cut-off, 72.3%; 1% cutoff, 83.0%), but without changing our conclusion.

We intentionally refrained from the comparative analysis of grade between paired DCIS and IBC lesions. Previous histopathological studies of synchronous DCIS and IBC have shown a close link between the grade of the in situ and invasive component.^{8,9} However, histopathological studies of primary DCIS and subsequent IBC have shown only moderate correlation between the grades of the subsequent invasive tumor and the original DCIS.^{22,29} A likely explanation is that the criteria used for grading DCIS and IBC are different. For DCIS grading multiple classification methods are in use.^{30–33} These grading systems are predominantly based on nuclear polymorphism, cell polarization, and also, in some, on the presence or absence of necrosis. None of the classification systems for DCIS is evidently the golden standard as they are all based on subjective criteria. In contrast, invasive breast carcinoma is graded according to a standardized classification system, i.e. the Elston and Ellis modification of Bloom-Richardson system, which is based on more objective features, i.e. tumor tubule formation, number of mitotic figures, and nuclear polymorphism.¹⁵ Although in the current study we re-examined all DCIS lesions using a single DCIS grading system, the problem of the comparability of grading systems for DCIS and IBC still stands. Indeed, we and others, found a closer association was found between the nuclear polymorphism score for the subsequent invasive carcinoma and the original DCIS grade (Supplementary Table S8).²⁹

Previous studies comparing receptor status between primary invasive breast carcinoma and the corresponding recurrence demonstrated discordances for ER, PR, and HER2 in a range of 10–37%, 24–48%, and 3–30%, respectively.^{34–41} However, as almost all of these primary invasive breast carcinomas were treated by radiotherapy and/or chemotherapy, part of these discordances could be caused by treatment.⁴² Furthermore, the difference in frequency of HER2 overexpression in DCIS and IBC was investigated by previous studies using synchronous DCIS and IBC.^{3,11–13} In line with our study, none of these studies found a significant difference of HER2 overexpression between synchronous DCIS and IBC. Another explanation for this change in HER2 status is that the invasive component arose from a DCIS subclone that did not harbor HER2 amplification in the first place. The results from our IHC heterogeneity analysis at least suggest that this might be a valid option.

Our study has some limitations. First, our study group consisted of women that were all treated for DCIS by BCS alone. DCIS treated by BCS carries a risk of recurrent disease, but the origin of the subsequent IBC after primary DCIS could be: (1) from residual DCIS that was left behind after BCS, or (2) unrelated to the preceding DCIS, and thus be a second primary tumor. In the current study, the contribution of second primary tumors after treatment of DCIS is unknown and should be further elucidated by molecular analysis of the primary DCIS and subsequent IBC. Second, we cannot exclude the possibility of receptor

measurement error as the source of discordance in marker expression. Third, for the intra-lesional heterogeneity analysis, inclusion of more heterogeneously expressed IHC markers would be more informative when assessing heterogeneity within DCIS lesions, as now we only found 10 cases of heterogenous DCIS based on IHC subtype.

Our study has several strengths. First, we made use of a unique, large series of 155 patient-matched primary DCIS and subsequent IBC pairs, derived from a well-defined cohort of patients treated for DCIS with BCS alone. Within this patient group we were able to compare both primary DCIS versus subsequent IBC and synchronous DCIS versus the IBC component. Second, for all tissue specimens included in this study new H&E whole slides were developed, which were reassessed by specialized breast pathologists. Third, new IHC stained whole slides were used for IHC assessment, to reassure that all stains of the different samples were performed using the same protocol. Fourth, a good agreement was achieved in the interobserver analysis of IHC assessment.

In summary, this study demonstrated that marker expression between *primary* DCIS and *subsequent* IBC is less concordant than *synchronous* DCIS and IBC. This indicates that synchronous and subsequent lesions are not that similar after all. HER2 marker expression showed the largest discrepancy: 36% of HER2-positive *primary* DCIS lesions were followed by HER2-negative IBC. Surprisingly, the frequency of discordant marker expression between primary DCIS and subsequent IBC did not increase with longer time to IBC. Intra-lesional heterogeneity was identified as a possible cause of the observed discordant marker expression. We suggest that future studies investigating the progression of DCIS to IBC, should study primary DCIS and subsequent IBC, instead of synchronous DCIS and IBC lesions. Only this comparison could result in the identification of solid markers for DCIS progression.

More research is needed to assess the contribution of second primary tumors after treatment of DCIS. Molecular analysis of a large, well-annotated patient series including patient matched primary DCIS and subsequent IBC, is of high importance. Currently, there are multiple initiatives that are setting up study cohorts including primary DCIS and subsequent IBC. These include the Sloane project, the PRECISION (PREvent ductal Carcinoma In Situ Invasive Overtreatment Now) initiative, and also noninferiority trials (i.e. LORD, LORIS, and COMET; <http://www.cancerresearchuk.org/funding-for-researchers/how-we-deliver-research/grand-challenge-award/funded-teams-wesseling>).^{43–46} These initiatives are highly relevant in finding markers for DCIS progression. Additionally, as copy number alterations are acquired at early stages of tumorigenesis, comparative analysis of these alterations and mutations between these matched pairs would be the preferable choice to assess the clonal-relationship between DCIS and IBC.^{47,48}

Article information

Acknowledgements: We thank all patients of which stored residual material was used in this study, all collaborating hospitals and in particular the pathology departments, and all persons who helped in the process of data collection and analysis. We would like to specifically acknowledge Otto Visser, Annemarie Eeltink, and the registration teams of the Netherlands Comprehensive Cancer Organization for collection of the data for the Netherlands Cancer Registry. We thank Lucy Overbeek and PALGA, the nation-wide histopathology and cytopathology data network and archive, for providing pathology data and for their help in the collection of the residual patient material; the pathologists of the PALGA group: i.a. P.Bult, C.H.M. van Deurzen, E. Geuken, and C.M van Dijk for their cooperation; and Danah Kamphuis, Chaima Loukili, and Frank Nieboer for their assistance.

This work was funded by Pink Ribbon (grant number 2013.WO29, to J. Wesseling), A Sisters Hope (grant number 2011.WO19.C88, to J. Wesseling), and was jointly funded by Cancer Research UK and the Dutch Cancer Society (grant number C38317/A24043, to J. Wesseling).

Author's contributions: LLV, FvL, EJR, MS, MKS, EHL, and JW designed the study. LLV and LEE collected data. DP, IH, and AB set up and performed IHC stains. LLV, KvdV, EJG, MMA, JW, JS, and CB did the pathology review and IHC assessment. LLV, FvL, EJR, MS, MKS, EHL, and JW analyzed and interpreted the data. LLV wrote the report. All authors provided comments and approved the final version.

References

1. Cowell CF, Weigelt B, Sakr R, et al. Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Mol Oncol*. 2013;7(5):859-869.
2. O'Connell P, Pekkel V, Fuqua SAW, Osborne CK, Clark GM, Allred DC. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst*. 1998;90(9):697-703.
3. Burkhardt L, Grob TJ, Hermann I, et al. Gene amplification in ductal carcinoma in situ of the breast. *Breast Cancer Res Treat*. 2010;123(3):757-765.
4. Buerger H, Otterbach F, Simon R, et al. Comparative genomic hybridization of ductal carcinoma in situ of the breast-evidence of multiple genetic pathways. *J Pathol*. 1999;187(4):396-402.
5. Buerger H, Otterbach F, Simon R, et al. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol*. 1999;189(4):521-526.
6. To T, Wall C, Baines CJ, Miller AB. Is carcinoma in situ a precursor lesion of invasive breast cancer? *Int J Cancer*. 2014;135(7):1646-1652.
7. 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 Res Treat*. 2016;159(3):553-563.
8. Lampejo OT, Barnes DM, Smith P, Millis RR. Evaluation of infiltrating ductal carcinomas with a DCIS component: correlation of the histologic type of the in situ component with grade of the infiltrating component. *Semin Diagn Pathol*. 1994;11(3):215-222.
9. Goldstein NS, Murphy T. Intraductal carcinoma associated with invasive carcinoma of the breast: A comparison of the two lesions with implications for intraductal carcinoma classification systems. *Am J Clin Pathol*. 1996;106(3):312-318.
10. Allred DC, Clark GM, Tandon AK, et al. HER-2/neu in node-negative breast cancer: prognostic significance of overexpression influenced by the presence of in situ carcinoma. *J Clin Oncol*. 1992;10(4):599-605.
11. Latta EK, Tjan S, Parkes RK, O'Malley FP. The role of HER2/neu overexpression/amplification in the progression of ductal carcinoma in situ to invasive carcinoma of the breast. *Mod Pathol*. 2002;15(12):1318-1325.
12. Park K, Han S, Kim HJ, Kim J, Shin E. HER2 status in pure ductal carcinoma in situ and in the intraductal and invasive components of invasive ductal carcinoma determined by fluorescence in situ hybridization and immunohistochemistry. *Histopathology*. 2006;48(6):702-707.
13. Hui Y, Lu S, Wang H, Resnick MB, Wang Y. Discordant HER2 Immunohistochemical Expression and Gene Amplification in Ductal Carcinoma In Situ - Evaluating HER2 in Synchronous In Situ and Invasive Carcinoma. *Histopathology*. August 2019;74(2):358-362.
14. Cowell CF, Weigelt B, Sakr RA, et al. Progression from ductal carcinoma in situ to invasive breast cancer: Revisited. *Mol Oncol*. 2013;7(5):859-869.
15. BLOOM HJ, RICHARDSON WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer*. 1957;11(3):359-377.
16. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 2002;41(3A):154-161.
17. Visser LL, Elshof LE, Schaapveld M, et al. Clinicopathological Risk Factors for an Invasive Breast Cancer Recurrence after Ductal Carcinoma In Situ-A Nested Case-Control Study. *Clin Cancer Res*. April 2018;clincanres.0201.2018.
18. Yemelyanova A, Vang R, Kshirsagar M, et al. Immunohistochemical staining patterns of p53 can serve as a surrogate marker for TP53 mutations in ovarian carcinoma: an immunohistochemical and nucleotide sequencing analysis. *Mod Pathol*. 2011;24(9):1248-1253.
19. Boyle DP, McCart DG, Irwin G, et al. The

- prognostic significance of the aberrant extremes of p53 immunophenotypes in breast cancer. *Histopathology*. 2014;65(3):340-352.
20. Wolff AC, Hammond MEH, Hicks DG, et al. Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol*. 2013;31(31):3997-4013.
 21. Ristimäki A, Sivula A, Lundin J, et al. Prognostic significance of elevated cyclooxygenase-2 expression in breast cancer. *Cancer Res*. 2002;62(3):632-635.
 22. Bijker N, Peterse JL, Duchateau L, 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. *Br J Cancer*. 2001;84(4):539-544.
 23. Boland GP, Butt IS, Prasad R, Knox WF, Bundred NJ. COX-2 expression is associated with an aggressive phenotype in ductal carcinoma in situ. *Br J Cancer*. 2004;90(2):423-429.
 24. Pape-Zambito D, Jiang Z, Wu H, 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.
 25. Casasent AK, Schalck A, Gao R, et al. Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell*. 2018;172(1-2):205-217.
 26. Van Bockstal M, Lambein K, Gevaert O, et al. Stromal architecture and periductal decorin are potential prognostic markers for ipsilateral locoregional recurrence in ductal carcinoma in situ of the breast. *Histopathology*. 2013;63(4):520-533.
 27. Nofech-Mozes S, Spayne J, Rakovitch E, et al. Biological Markers Predictive of Invasive Recurrence in DCIS. *Clin Med Oncol*. 2008;2(416):7-18.
 28. Karlsson E, Sandelin K, Appelgren J, et al. Clonal alteration of breast cancer receptors between primary ductal carcinoma in situ (DCIS) and corresponding local events. *Eur J Cancer*. 2014;50(3):517-524.
 29. Millis RR, Pinder SE, Ryder K, Howitt R, Lakhani SR. Grade of recurrent in situ and invasive carcinoma following treatment of pure ductal carcinoma in situ of the breast. *Br J Cancer*. 2004;90(8):1538-1542.
 30. Holland R, Peterse JL, Millis RR, et al. Ductal carcinoma in situ: a proposal for a new classification. *Semin Diagn Pathol*. 1994;11(3):167-180.
 31. Poller DN, Silverstein MJ, Galea M, et al. Ideas in pathology. Ductal carcinoma in situ of the breast: a proposal for a new simplified histological classification association between cellular proliferation and c-erbB-2 protein expression. *Mod Pathol*. 1994;7(2):257-262.
 32. Scott MA, Lagios MD, Axelsson K, Rogers LW, Anderson TJ, Page DL. Ductal carcinoma in situ of the breast: reproducibility of histological subtype analysis. *Hum Pathol*. 1997;28(8):967-973.
 33. Pinder SE, Duggan C, Ellis IO, et al. A new pathological system for grading DCIS with improved prediction of local recurrence: results from the UKCCCR/ANZ DCIS trial. *Br J Cancer*. 2010;103(1):94-100.
 34. Liedtke C, Broglio K, Moulder S, et al. Prognostic impact of discordance between triple-receptor measurements in primary and recurrent breast cancer. *Ann Oncol Off J Eur Soc Med Oncol*. 2009;20(12):1953-1958.
 35. Lindström LS, Karlsson E, Wilking UM, et al. Clinically used breast cancer markers such as estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 are unstable throughout tumor progression. *J Clin Oncol*. 2012;30(21):2601-2608.
 36. Niikura N, Liu J, Hayashi N, et al. Loss of human epidermal growth factor receptor 2 (HER2) expression in metastatic sites of HER2-overexpressing primary breast tumors. *J Clin Oncol*. 2012;30(6):593-599.
 37. Kamby C, Rasmussen BB, Kristensen B. Oestrogen receptor status of primary breast carcinomas and their metastases. Relation to pattern of spread and survival after recurrence. *Br J Cancer*. 1989;60(2):252-257.
 38. Curigliano G, Bagnardi V, Viale G, et al. Should liver metastases of breast cancer be biopsied to improve treatment choice? *Ann Oncol Off J Eur Soc Med Oncol*. 2011;22(10):2227-2233.
 39. Thompson AM, Jordan LB, Quinlan P, et al. Prospective comparison of switches

- in biomarker status between primary and recurrent breast cancer: the Breast Recurrence In Tissues Study (BRITS). *Breast Cancer Res.* 2010;12(6):R92.
40. Amir E, Miller N, Geddie W, et al. Prospective study evaluating the impact of tissue confirmation of metastatic disease in patients with breast cancer. *J Clin Oncol.* 2012;30(6):587-592.
 41. Wilking U, Karlsson E, Skoog L, et al. HER2 status in a population-derived breast cancer cohort: discordances during tumor progression. *Breast Cancer Res Treat.* 2011;125(2):553-561.
 42. Kramer I, Schaapveld M, Oldenburg HSA, et al. The influence of adjuvant systemic regimens on contralateral breast cancer risk and receptor subtype. *JNCI J Natl Cancer Inst.* January 2019.
 43. Thompson AM, Clements K, Cheung S, et al. Management and 5-year outcomes in 9938 women with screen-detected ductal carcinoma in situ: the UK Sloane Project. *Eur J Cancer.* 2018;101:210-219.
 44. Elshof LE, Tryfonidis K, Slaets L, 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. *Eur J Cancer.* 2015;51(12):1497-1510.
 45. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen detected DCIS; The LORIS trial. *Eur J Cancer.* 2014;51(16):2296-2303.
 46. Youngwirth LM, Boughey JC, Hwang ES. Surgery versus monitoring and endocrine therapy for low-risk DCIS: The COMET Trial. *Bull Am Coll Surg.* 2017;102(1):62-63.
 47. Wang Y, Waters J, Leung ML, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature.* 2014;512(7513):155-160.
 48. Gao R, Davis A, McDonald TO, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet.* 2016;48(10):1119-1130.

Supplementary tables and figures

Table S1. Clinical variables of all patients and of the selection used in this paper

Characteristics	Selection <i>n</i> = 155		All patients <i>n</i> = 328		<i>P</i>
Age	58	(32 - 87)	59	(30 - 89)	0.32 ^b
Year of DCIS diagnosis	1996	(1989 - 2004)	1996	(1989 - 2004)	0.58 ^b
Time to event					
0-3 years	37	(23.9%)	96	(29.3%)	
4-6 years	40	(25.8%)	81	(24.7%)	
7-9 years	44	(28.4%)	78	(23.8%)	
≥10 years	34	(21.9%)	73	(22.3%)	0.56 ^c
Screen-detection ^a					
Yes	73	(47.1%)	139	(42.4%)	
No	67	(43.2%)	161	(49.1%)	
Unknown	15	(9.7%)	28	(8.5%)	0.48 ^c

^a: Group of screen-detected lesions include women >50 years. DCIS of women <50 years were not detected via screening; ^b: *p*-value calculated by unpaired T-test; ^c: *p*-value calculated by Fisher's exact test

Table S2.

Available with the published article.

Table S3. Immunohistochemical marker expression by DCIS histological grade and IBC grade

	Primary DCIS			Invasive breast cancer					
	Histological grade			Grade					
	Well (n = 26) n (%)	Interm. (n = 86) n (%)	Poor (n = 30) n (%)	I (n = 27) n (%)		II (n = 65) n (%)		III (n = 48) n (%)	
ER									
Negative	0 (0)	11 (12.8)	13 (43.3)	1 (3.7)	10 (15.4)	11 (22.9)			
Positive	26 (100.0)	75 (87.2)	17 (56.7)	26 (96.3)	55 (84.6)	37 (77.1)			
PR									
Negative	5 (19.2)	25 (29.1)	21 (70.0)	4 (14.8)	24 (36.9)	28 (58.3)			
Positive	21 (80.8)	60 (69.8)	9 (30.0)	23 (85.2)	41 (63.1)	20 (41.7)			
N/A	0 (0)	1 (1.2)	0 (0)	0 (0)	0 (0)	0 (0)			
HER2									
Negative	24 (92.3)	64 (74.4)	5 (16.7)	24 (88.9)	55 (84.6)	25 (52.1)			
Positive	2 (7.7)	22 (25.6)	25 (83.3)	3 (11.1)	10 (15.4)	23 (47.9)			
p16									
Low	17 (65.4)	41 (47.7)	14 (46.7)	14 (51.9)	37 (56.9)	27 (56.3)			
High	9 (34.6)	45 (52.3)	16 (53.3)	13 (48.1)	27 (41.5)	21 (43.8)			
N/A	0 (0)	0 (0)	0 (0)	0 (0)	1 (1.5)	0 (0)			
p53									
0% positive cells (mutant)	2 (7.7)	4 (4.7)	6 (20.0)	1 (3.7)	7 (10.8)	9 (18.8)			
1-70% positive cells (WT)	22 (84.6)	72 (83.7)	14 (46.7)	25 (92.6)	50 (76.9)	26 (54.2)			
>70% positive cells (mutant)	2 (7.7)	9 (10.5)	10 (33.3)	1 (3.7)	8 (12.3)	13 (27.1)			
N/A	0 (0)	1 (1.2)	0 (0)	0 (0)	0 (0)	0 (0)			
COX-2									
Low	2 (7.7)	10 (11.6)	4 (13.3)	6 (22.2)	7 (10.8)	3 (6.3)			
High	24 (92.3)	76 (88.4)	25 (83.3)	21 (77.8)	57 (87.7)	45 (93.8)			
N/A	0 (0)	0 (0)	1 (3.3)	0 (0)	0 (0)	0 (0)			

P-values were calculated by Fisher's exact test. N/As were not included in the statistical analysis.

Note: Grade could not be assessed for two invasive lesions because of poor fixation.

Table S4. Immunohistochemical subtype synchronous DCIS related to invasive component

	Invasive component				
	HR+ HER2-	HR+ HER2+	HR- HER2+	HR- HER2-	Total
Synchronous DCIS					
HR+ HER2-	53	3	0	0	56
HR+ HER2+	0	8	0	0	8
HR- HER2+	0	3	8	1	12
HR- HER2-	1	0	0	4	5
Total	54	14	8	5	81

Agreement of 90.1%; marginal homogeneity (symmetry) test $p = 0.09$

Table S5. Immunohistochemical subtype primary DCIS related to invasive recurrence

Invasive breast cancer					
	HR+ HER2-	HR+ HER2+	HR- HER2+	HR- HER2-	Total
Primary DCIS					
HR+ HER2-	81	4	2	3	90
HR+ HER2+	9	16	2	1	28
HR- HER2+	5	3	10	3	21
HR- HER2-	1	0	0	2	3
Total	96	23	14	9	142

Agreement of 76.8%; marginal homogeneity test $p = 0.040$

Table S6. Marker expression primary DCIS related to synchronous DCIS

		Synchronous DCIS		% Agreement	<i>P</i>
		Negative/ Low/WT	Positive/ High/Mutant		
Primary DCIS					
ER	Negative	11	5	86.3	0.76
	Positive	6	58		
PR	Negative	20	13	70.4	0.68
	Positive	11	37		
HER2	Negative	51	1	86.4	0.007
	Positive	10	19		
p53	WT	54	8	77.5	0.64
	Mutant	10	8		
COX-2	Low	1	8	75.0	0.37
	High	12	59		

Comparison between DCIS and DCIS-IBC were made by marginal homogeneity test.

Table S7. IHC subtype primary DCIS related to synchronous DCIS

	Synchronous DCIS				Total
	HR+ HER2-	HR+ HER2+	HR- HER2+	HR- HER2-	
Primary DCIS					
HR+ HER2-	48	0	1	1	50
HR+ HER2+	5	6	3	1	15
HR- HER2+	3	2	8	1	14
HR- HER2-	0	0	0	2	2
Total	56	8	12	5	81

Agreement of 79.0%; marginal homogeneity $p=0.034$

Table S8. Percentage discordance in grade between pairs of DCIS and IBC

	sDCIS vs. IBC		pDCIS vs. IBC		pDCIS vs. sDCIS	
	%	Symmetry P	%	Symmetry P	%	Symmetry P
PA						
Grade	37.7	0.06	56.2	0.015	45.4	0.17
Tumor tubule formation*	58.9	<0.001	66.9	<0.001	-	-
Number of mitotic figures *	55.7	<0.001	60.2	<0.001	-	-
Nuclear polym. *	29.9	0.001	52.6	<0.001	-	-

*: Histological grade DCIS vs. the different components of the grade of invasive disease (tumor tubule formation, number of mitotic figures, and nuclear polymorphism); discordance (%) = 100-(% Agreement calculated by kappa).

pDCIS: Primary DCIS; sDCIS: synchronous DCIS.

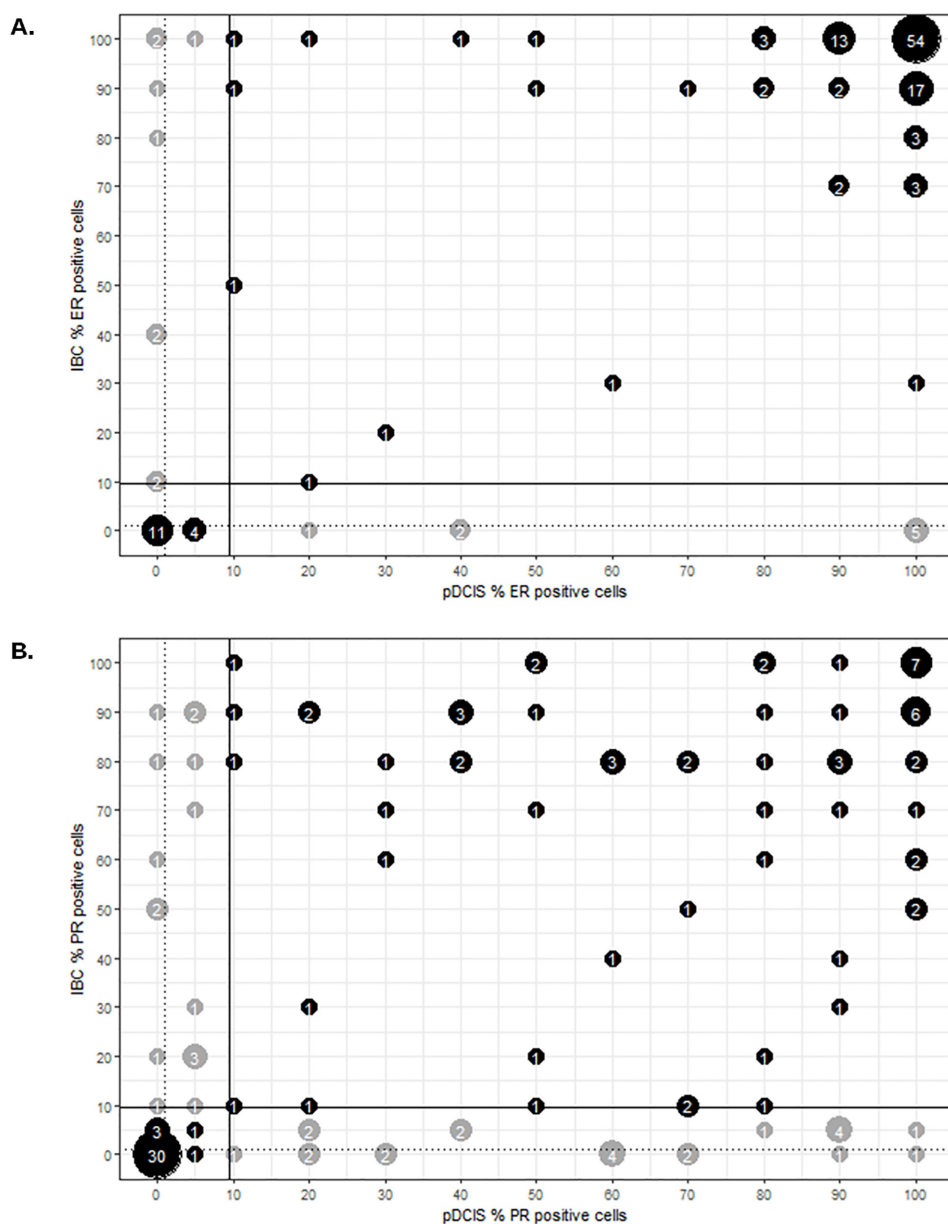


Figure S1. Continuous marker expression primary DCIS related to invasive recurrence. Solid lines represents 10% cut-off and dotted lines represent 1% cut-off for ER and PR positivity. Gray filled circles represent pairs with discordant marker expression and black filled circles represent pairs with concordant marker expression. Numbers represent number of pairs. **A.** ER expression: 10% cut-off, discordant n=17; 1% cutoff, discordant n=20. **B.** PR expression: 10% cut-off, discordant n=39; 1% cutoff, discordant n=24.

Lindy L. Visser
Marlous Hoogstraat
Tycho Bismeijer
Frank Nieboer
Michiel de Maaker
Petra Kristel
Lennart Mulder
Marielle Kreté
Wim Brugman
Michael Schaapveld
Marjanka K. Schmidt
Elinor Sawyer
Alastair Thompson
Andrew Futreal
Helen Davies
Serena Nik-Zainal
Lodewyk F. A. Wessels
Esther H. Lips*
Jelle Wesseling*
On behalf of the PRECISION team

Manuscript in preparation

*: Shared last author

Approximately 30% of invasive ipsilateral breast cancers after treatment of ductal carcinoma in situ may be new independent primary cancers

Abstract

Ductal carcinoma *in situ* (DCIS) is regarded as a precursor of invasive breast cancer (IBC). Yet, not all DCIS progresses into IBC and proof of direct progression of DCIS to IBC is lacking. Therefore, we assessed the clonal relatedness between the initial DCIS and subsequent ipsilateral invasive breast cancer (IBC) by combined analysis of somatic copy number changes and mutations. This revealed that a substantial percentage (29.5%; i.e. 23 out of 78) of subsequent ipsilateral IBC are likely not clonally related to the preceding DCIS and therefore could be new primary tumors. This challenges the dogma that almost all subsequent ipsilateral IBC represent outgrowth of the initial DCIS lesion. If true, DCIS might not just be an IBC precursor, but could be an "independent" risk factor for developing IBC as well. This will substantially impact current clinical DCIS risk estimates, if validated by ongoing confirmatory analyses.

Main

Ductal carcinoma *in situ* (DCIS) is a non-invasive breast lesion with the potential to progress to invasive breast cancer (IBC) within a time period varying from a few years to decades.¹ Only a subset of DCIS will undergo the transition from *in situ* to invasive cancer, and not all invasive cancers arise from DCIS.^{2,3} Factors mediating the progression of DCIS to IBC are still largely unknown. Before the introduction of population breast cancer screening, DCIS was rarely diagnosed, but nowadays it represents about 20-25% of all newly diagnosed breast cancers.⁴⁻⁶

Although DCIS is considered a precursor of IBC, definite proof of its progression to IBC is still lacking. The fact that most DCIS is treated, hinders studies on the natural course of DCIS. Treatment of DCIS consists of breast conserving surgery (BCS), often followed by radiotherapy and/or endocrine therapy, or even mastectomy. After breast conserving treatment, recurrent disease may occur. The subsequent ipsilateral IBC could be progression of residual DCIS that was left behind after BCS, or may be a new primary tumor that is unrelated to the preceding DCIS.

Traditionally, assessment whether a subsequent cancer is related to the preceding one, is primarily based on comparison of clinical and histopathologic information, i.e time interval, location, morphology, tumor grade, and immunohistochemical marker expression.⁷⁻¹¹ In recent years, genetic markers have started to play an increasingly important role in assessing the clonal relatedness of two lesions.¹²⁻¹⁶

In the present study, we evaluated the clonal relatedness of 78 patient-matched DCIS and subsequent ipsilateral IBC. This will help us to provide evidence of direct progression from DCIS to IBC. The most ideal patient group for this study would have been a group of women with untreated DCIS that subsequently developed ipsilateral IBC. However, this is unethical and almost all women with DCIS are treated. Therefore, we investigated women with DCIS and subsequent ipsilateral IBC, who were treated for DCIS with BCS alone. This patient group was part of a nation-wide population-based cohort including all women diagnosed with and treated for DCIS with BCS alone in the Netherlands between 1989 and 2005, with a median follow-up time of 12.0 years.¹⁷ Formalin-fixed paraffin-embedded (FFPE) tissue specimens of patient-matched DCIS and subsequent ipsilateral IBC were retrieved from across the Netherlands, and pathology review and molecular profiling was performed. Patterns of shared and unique somatic copy number alterations and mutations were evaluated between the patient-matched DCIS and subsequent IBC. Pairs classified as having a shared clonal origin, showing one or more shared somatic events and only few to none sample-specific somatic events, were considered true recurrences. Pairs classified as having independent origin, showing only few to none shared somatic events and one or more sample-specific events, were considered new primary tumors.

Table 1. Overview of clinical and histopathologic characteristics of the primary DCIS and matched subsequent invasive breast cancer

Patient ID	Age (years)	Lesion			Primary DCIS				Time to event (years)							Invasive breast cancer					
		size (mm)	Marginal status	Histologic grade	ER	PR	HER2	Same quadrant?	Histologic subtype	Stage	Grade	ER	PR	HER2	Discordance						
9698	61.9	20	free	poor	+	+	+	0.5	unknown	NST	1	3	+	+	+						
6325	60.2	NA	unknown	intermediate	+	+	-	0.5	yes	mixed	2B	2	+	+	-						
3630	48.4	NA	free	intermediate	+	+	-	0.7	no	NST	3A	1	+	+	-	location, ER					
8417	37.7	NA	involved	intermediate	+	+	+	1.0	no	NST	2A	2	+	-	-	location, PR, HER2					
1569	68.9	NA	free	poor	NA	NA	NA	1.0	no	NST	2A	3	-	-	-	location					
4863	58.4	NA	free	poor	-	-	+	1.1	no	NST	1	3	-	-	+	location					
10048	52	14	free	intermediate	-	-	-	1.1	unknown	NST	4	2	-	-	-						
4590	72.4	20	unknown	intermediate	+	+	-	1.4	yes	NST	1	2	+	+	-						
6280	55.9	NA	free	poor	+	+	-	1.9	no	NST	1	2	+	+	-	location, grade					
9663	51.5	NA	free	poor	+	-	-	1.9	unknown	NST	3A	3	+	+	-	PR					
1800	66.2	NA	free	poor	-	-	+	1.9	unknown	NST	1	2	-	-	-	grade, HER2					
7114	73.7	NA	free	intermediate	+	+	-	1.9	yes	NST	1	3	+	+	-	grade					
971	56	11	free	intermediate	+	-	-	2.1	yes	NST	1	1	+	+	-	grade, PR					
9371	50.4	NA	free	intermediate	-	-	+	2.1	unknown	NST	1	2	-	-	+						
8088	60.1	NA	unknown	intermediate	-	-	+	2.3	unknown	NST	1	2	-	-	-	HER2					
6929	43.8	7	free	intermediate	+	+	-	2.3	unknown	NST	2B	3	+	+	-	grade					
783	66.4	NA	free	intermediate	+	+	-	2.4	yes	NST	2A	2	+	+	-						
384	39.3	8	unknown	intermediate	+	+	-	2.8	yes	NST	2B	2	+	-	-	PR					
6921	49.6	15	free	poor	+	-	+	2.9	yes	NST	1	3	+	-	+						
2010	55.5	NA	free	well	+	+	-	3.0	no	NST	2A	3	+	-	-	location, grade, PR					
6890	53.8	NA	unknown	intermediate	+	+	-	3.0	unknown	NST	2A	2	+	+	-						
2002	66.9	NA	free	poor	-	-	+	3.0	yes	NST	1	3	-	-	+						
293	58.8	20	free	poor	+	+	-	3.1	yes	NST	2A	1	+	+	-	grade					
88	74.7	NA	free	intermediate	+	+	-	3.5	yes	NST	2A	2	+	+	-						
3507	62.6	30	involved	intermediate	-	-	+	3.7	yes	NST	1	1	-	-	+	grade					
5736	68.3	25	free	intermediate	+	+	-	4.2	yes	NST	1	2	+	+	-						
7323	56.6	24	free	well	+	+	-	4.2	yes	NST	2A	2	+	+	-	grade					
9355	67.5	NA	unknown	well	+	-	-	4.2	yes	NST	1	2	+	+	-	grade, PR					
7199	57.9	NA	free	intermediate	+	+	-	4.3	unknown	NST	3B	1	+	+	-	grade					

Table 1. Overview of clinical and histopathologic characteristics of the primary DCIS and matched subsequent invasive breast cancer

Patient ID	Primary DCIS			Invasive breast cancer												
	Age (years)	Lesion size (mm)	Margin status	Histologic grade	ER	PR	HER2	Time to event (years)	Same quadrant?	Histologic subtype	Stage	Grade	ER	PR	HER2	Discordance
6919	46.4	NA	unknown	poor	+	+	+	4.6	yes	NST	2A	3	+	-	+	PR
8126	58.5	NA	free	poor	-	-	+	4.6	yes	NST	1	3	+	-	-	ER, HER2
7097	47.8	NA	involved	intermediate	+	+	-	5.0	unknown	mixed	2B	2	+	+	-	
3956	47.1	5	free	well	+	-	-	5.0	yes	mixed	2A	2	+	+	-	grade, PR
5230	46.4	NA	free	intermediate	+	+	-	5.1	unknown	NST	NA	2	+	+	-	
9413	51.7	NA	free	intermediate	+	-	+	5.1	yes	NST	1	3	+	-	+	grade
9253	56	NA	unknown	well	+	+	-	5.3	yes	mixed	2A	3	+	+	-	grade
3275	53	15	free	well	+	+	-	5.3	yes	NST	1	1	+	+	-	
455	86.7	NA	free	intermediate	+	+	-	5.4	yes	NST	2B	2	+	+	-	
4879	59.7	NA	free	intermediate	+	+	+	5.5	yes	NST	1	3	-	-	-	grade, ER, PR, HER2
7831	66.6	NA	free	poor	-	-	+	5.5	no	NST	4	3	-	-	+	location
8759	33.9	NA	free	intermediate	-	-	+	6.1	unknown	NST	1	3	+	+	+	grade, ER, PR
4965	58	NA	involved	poor	+	+	+	6.3	yes	lobular	4	3	+	-	+	PR
3800	63.9	NA	unknown	intermediate	+	+	-	6.3	yes	NST	2B	2	+	-	-	PR
3386	86.5	NA	unknown	intermediate	-	-	-	6.4	yes	NST	2A	3	+	-	-	grade, ER
1678	69	8	involved	intermediate	+	+	-	6.6	yes	NST	1	2	+	+	-	
9076	65.5	NA	involved	intermediate	+	+	-	6.7	yes	mixed	1	1	+	+	-	grade
2364	53	NA	free	poor	+	-	-	6.9	yes	NST	1	3	+	+	+	PR, HER2
5613	34.5	22	free	well	+	+	+	7.0	unknown	NST	3A	2	+	-	+	grade, PR
567	53.8	NA	involved	well	+	+	-	7.1	yes	NST	4	2	-	-	+	grade, ER, PR, HER2
8300	49.7	10	free	intermediate	+	+	-	7.1	yes	NST	3A	2	+	+	-	
7993	58.1	NA	free	intermediate	+	-	+	7.3	unknown	NST	1	3	+	+	+	grade, PR
809	51.6	10	free	intermediate	+	-	+	7.3	yes	NST	3A	3	+	-	+	grade
471	59.6	17	free	poor	-	-	+	7.4	yes	NST	3A	2	+	+	-	grade, ER, PR, HER2
8206	58.5	15	free	intermediate	+	+	-	7.5	yes	NST	1	3	+	+	-	grade
4901	56.6	10	free	intermediate	+	NA	-	8.0	yes	mixed	1	3	+	+	-	grade
216	51.2	NA	involved	intermediate	+	-	-	8.0	unknown	NST	1	2	+	-	-	
7829	65.3	NA	free	intermediate	+	-	+	8.1	unknown	NST	1	2	+	-	+	

Table 1. Overview of clinical and histopathologic characteristics of the primary DCIS and matched subsequent invasive breast cancer

Patient ID	Age (years)	Primary DCIS					Invasive breast cancer									
		Lesion size (mm)	Margin status	Histologic grade	ER	PR	HER2	Time to event (years)	Same quadrant?	Histologic subtype	Stage	Grade	ER	PR	HER2	Discordance
7784	80.7	NA	involved	intermediate	+	-	-	8.4	no	NST	1	3	+	-	-	location, grade
9196	57.1	NA	free	intermediate	+	-	-	8.4	yes	NST	2B	2	+	-	-	
820	70.3	15	free	well	+	+	-	8.7	yes	NST	1	1	+	+	-	
9337	50.7	NA	free	intermediate	+	+	-	8.7	unknown	NST	1	3	-	-	-	grade, ER, PR
264	58	15	free	intermediate	+	-	-	8.8	no	NST	1	3	-	-	+	location, grade, ER, HER2
5753	69.5	NA	free	intermediate	+	+	-	9.1	unknown	NST	1	3	+	+	-	grade
6737	43.8	12	involved	well	+	+	-	9.7	unknown	NST	2B	1	NA	NA	NA	
7815	73.3	8	free	intermediate	+	+	-	9.9	no	NST	2A	3	+	+	-	location, grade
9255	63.2	NA	unknown	intermediate	+	-	-	10.4	unknown	NST	2B	2	+	+	-	PR
9530	54.1	NA	free	intermediate	+	-	-	10.4	yes	NST	2A	2	+	-	-	
8168	55	NA	unknown	intermediate	+	+	-	10.8	unknown	NST	1	3	+	+	-	grade
277	50.1	NA	free	intermediate	+	+	-	10.9	unknown	NST	1	2	+	+	-	
3543	76.9	4	free	well	+	+	-	11.7	yes	NST	2A	2	+	+	-	grade
8210	45.7	2	free	intermediate	+	+	-	13.7	no	NST	2A	3	+	+	-	location, grade
6287	54.5	12	involved	intermediate	+	+	-	13.9	unknown	NST	1	2	+	+	-	
9117	38.2	NA	unknown	poor	NA	NA	NA	13.9	yes	NST	2A	3	-	-	-	
9202	71.1	NA	free	intermediate	-	-	+	14.4	no	NST	1	2	+	-	+	location, ER
2553	42.9	NA	unknown	well	+	+	-	15.3	unknown	mixed	4	2	+	+	-	grade
2802	37.5	NA	unknown	well	+	+	-	15.5	yes	NST	3A	2	+	-	-	grade, PR
9127	60.6	NA	unknown	poor	NA	NA	NA	15.7	unknown	NST	2A	3	-	-	+	
10018	44.7	NA	unknown	intermediate	+	+	-	17.0	unknown	NST	2A	3	+	-	-	grade, PR

NA: not assessable; NST: invasive carcinoma of no special type.

Results

Patient characteristics

We made use of a nation-wide population-based cohort including all women diagnosed and treated for DCIS in the Netherlands between 1989 and 2005. Details on this cohort are described before.¹⁷ In brief, clinicopathological information was obtained from the Netherlands Cancer Registry (NCR) and the nation-wide histopathology and cytopathology data network (PALGA). The cohort comprises 10,090 women of which 2,658 were treated with BCS alone. Patients did not receive tamoxifen or other endocrine adjuvant treatment.

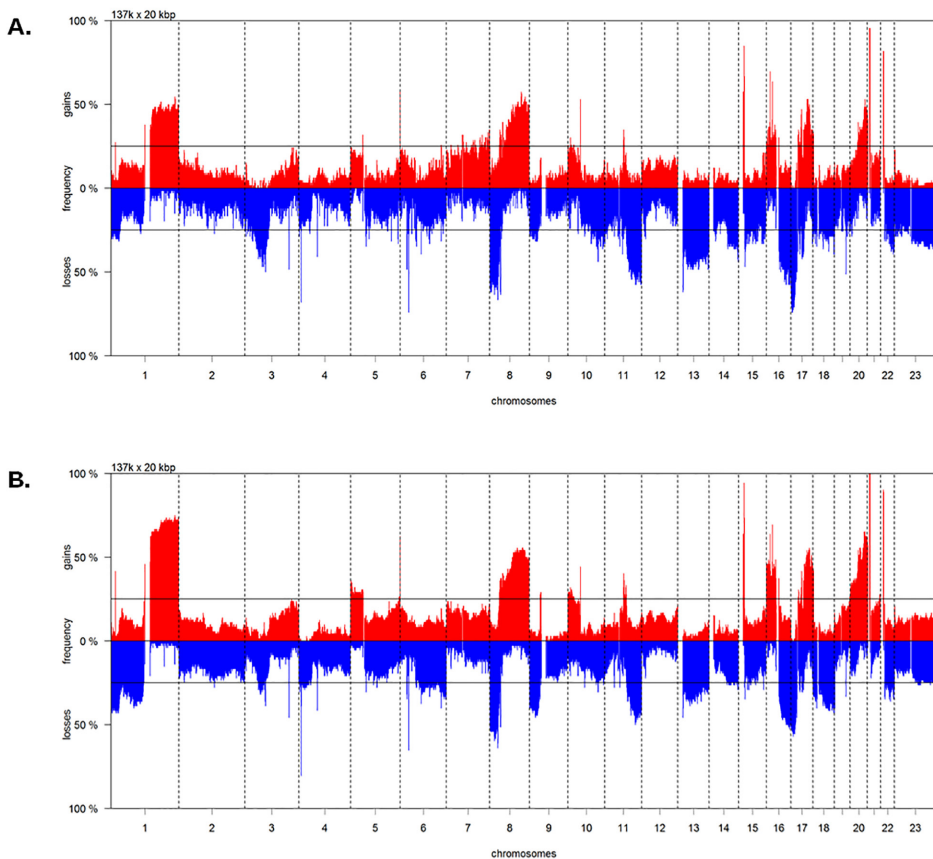


Figure 1. Genome-wide frequency plots of DNA copy number gains (red) and losses (blue) of DCIS and IBC samples separately. Both homozygous and heterozygous copy number events are present ranging from very focal to whole arm events. **(A)** Frequency plot of 66 DCIS samples and **(B)** frequency plot of 72 IBC samples. Noisy copy number profiles were excluded.

Fifteen years after DCIS diagnosis, 374 ipsilateral IBC events were observed after DCIS treatment with BCS alone (cumulative incidence of ipsilateral IBC of 15.4%). FFPE tissue specimens of patient-matched DCIS and subsequent IBC were successfully retrieved and pathologically re-examined of 155 women with an ipsilateral IBC event.¹⁸ DNA and RNA of sufficient yield and quality was successfully extracted from both DCIS and IBC of 78 women.

Clinical characteristics of the 78 patients are shown in Table 1. Mean time to IBC event was 6.3 years (range 0.5-17.0). Mean age at DCIS diagnosis was 57.5 years (range 33.9-86-7). Histological grade of DCIS was well differentiated in 16.7% (13/78), intermediately differentiated in 61.5% (48/78), and poorly differentiated in 21.8% (17/78) of the cases. 79.5% of DCIS was ER positive and 29.2% was HER2 positive.

Discordances in histopathological characteristics

Histopathological features were compared between patient-matched DCIS and IBC, including tumor location, histological grade, and the expression of ER, PR, and HER2 (Table 1). 65.4% of the patients (51/78) showed discordances between the DCIS and IBC lesion, with 25 of the 78 patients (32.1%) showing two or more discordant changes. Most changes were found when comparing grade (47.4%; 37/78 patients), while tumor location differed in 23.1% (12/52 patients; for 27 patients tumor location of DCIS and/or IBC was unknown), ER status in 12.2% (9/74 patients), PR status in 28.8% (21/73 patients), and HER2 status in 12.2% (9/74 patients).

Copy number alterations and mutations

DNA copy number analysis using low coverage whole genome sequencing (CNVseq) was performed for all DCIS and IBC pairs to identify regions of DNA copy number gain (red) and loss (blue) that occurred in at least 25% of the DCIS and IBC lesions in our study group. In both DCIS and IBC lesions, frequent DNA gains were identified on chromosomes 1q, 8q, 16p, 17q, and 20q, while DNA losses were detected on 1p, 3p, 8p, 11q, and 16q (Figure 1). These results are in line with DNA gains and losses frequently identified in breast cancer.¹⁹ Overall, although the frequencies of allelic changes are slightly lower in the *in situ* lesions, the general pattern of broad genetic instability is similar.

Ultra-deep sequencing (average >2000x) was performed to assess mutations in breast cancer driver genes, using a custom 53-gene panel. PIK3CA, TP53, NF1, BRCA2, and GATA3 were most frequently mutated in DCIS in 33.3%, 28.7%, 9.3%, 9.3% and 8.3% of the lesion, respectively. In the IBCs, we found TP53, PIK3CA, BRCA2, GATA3, and KMT2C to be the most frequently mutated genes, as they were affected in 31.7%, 27.0%, 11.1%, 9.5%, and 9.5% of the lesions, respectively. and events unique to the DCIS or IBC lesion.

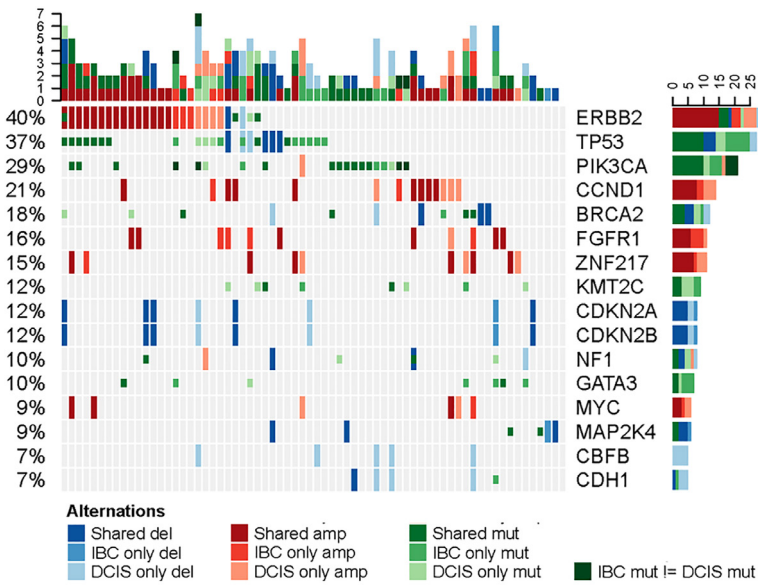


Figure 2. OncoPrint of 68 patient-matched DCIS and subsequent iIBC. Ten patients did not harbor any somatic events in the 53 assessed genes. IBC mut != DCIS mut, gene mutation present in both lesions but different mutation.

In Figure 2, an overview is given of the most frequent somatic events found in patient-matched DCIS and IBC (for full list of genes see Supplementary Figure S1 and Supplementary File 1). A distinction was made between somatic events shared by the DCIS and IBC lesion

Assessment of clonal relatedness

We determined if the patient-matched DCIS and IBC pairs had a shared clonal origin or represented two independent lesions, by evaluating patterns of shared and unique somatic copy number alterations and mutations (representative example shown in Figure 3). For this, we made use of two independent methods. First, we used the Clonality R package which is based on genome-wide copy number data. A cutoff for clonal origin and independence was set on $p < 0.1$ as suggested by the developers of the method. Second, we developed a panel-based clonality score making use of CNA and mutation data, restricted to our panel of 53 driver genes. The score consists of the sum of shared aberrations divided by the sum of and unique aberrations for every patient, corrected for the probability of observing these aberrations in breast cancer. Both true pairs (test set) and artificial pairs (independent pairs; reference set) were analyzed using the two methods. The artificial pairs are obtained by constructing pairs of lesion from different patients. Since these pairs are by definition

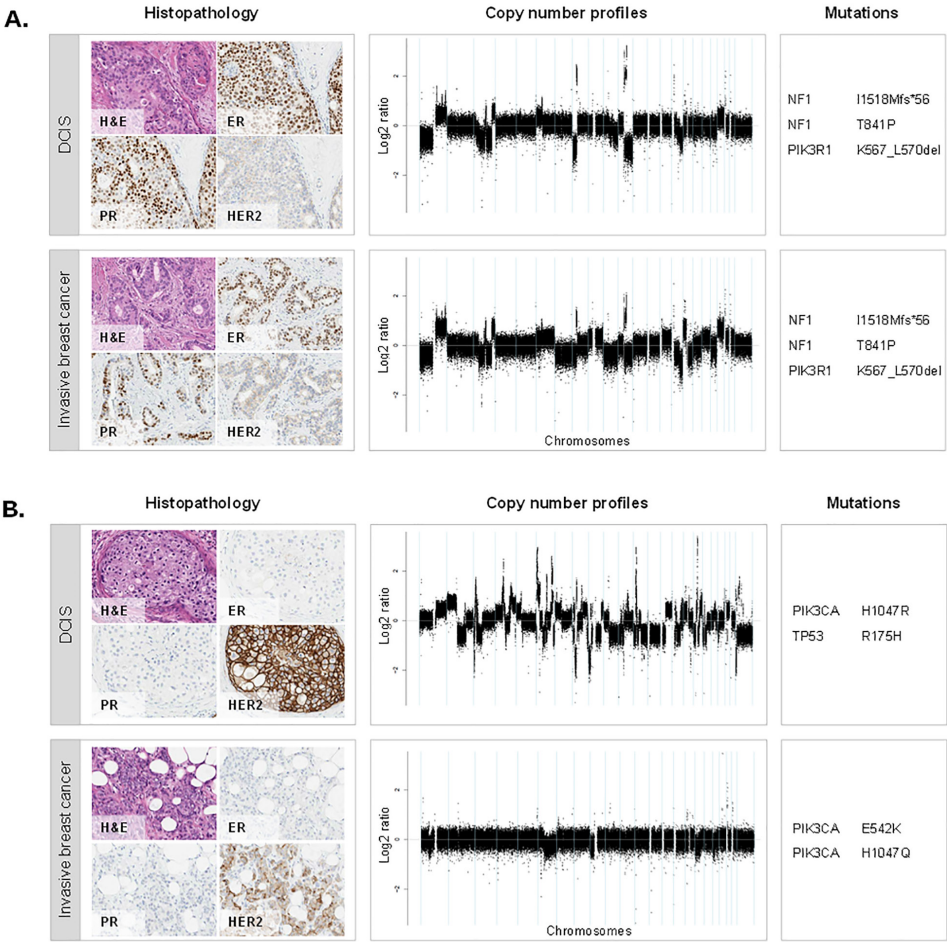


Figure 3. Example of a clonal (A) and independent (B) DCIS IBC pair

independent they provide the reference distribution, which is used to permit calculation of the p values of the real paring. A cutoff for clonal origin and independence was set on the 99th percentile of the artificial pairs' scores. Pairs classified as "independent" based on the Clonality package and with a low panel-based clonality score were considered new primary tumors. Pairs classified as "clonal" and with a high panel-based clonality score were considered true recurrences.

Based solely on copy number data, 45 of 78 pairs were classified as clonally related and 33 of 78 suspected of having an independent origin (Clonality package; Figure 4; Supplementary Figure S2; Supplementary File 2). Using our panel-based clonality score, 43 pairs were classified as clonal and 25 were likely to be independent (n=10 no somatic events found;

Supplementary Figure S3). The agreement between the two methods was 79% (54 of 68, with 10 unknowns; Supplementary Figure S4). 14 of 68 pairs showed discordant classification based on clonality package and panel-based clonality score. With this, 41 IBCs (52.6%) were classified as “true recurrences” and 23 IBCs (29.5%) were likely to be “new primary tumors”. For 14 pairs (17.9%) our method was inconclusive and were classified as “equivocal” (Figure 4).

Next, we assessed the clinicopathological characteristics of true recurrences and new primary tumors. Discordant grade was more often observed among new primary tumors (low (1-2) and high (3) grade; $p=0.049$; Figure 4; Supplementary Table S1). Discordant ER status ($p=0.021$) was higher among new primary tumors as compared to true recurrences. Furthermore, new primary IBCs trended to originate more often from HER2-positive DCIS index lesions ($p=0.054$) and were more often of HER2-negative type ($p=0.009$). Interestingly, involved margins ($p=0.70$), size of the DCIS index lesion ($p=0.32$), and discordant tumor location ($p=0.72$) are not associated with a higher rate of new primary tumors. No difference was found when comparing IBC stage between true recurrences and new primary IBCs ($p=0.52$).

Calculating the proportion of new primary tumors after DCIS based on epidemiologic data

Lastly, we estimated the chance for a woman with DCIS to develop a new primary tumor, using epidemiologic data. From a previous study by our group, making use of the same patient cohort, we know that: 1) fifteen years after DCIS diagnosis, cumulative incidence of ipsilateral IBC was 15.4% after BCS alone; and 2) that the cumulative incidence of contralateral invasive breast cancer (cIBC) at 15 years was 6.4% and did not differ by treatment, period of diagnosis, or age group.¹⁷ We assumed that the incidence of new primary tumors is similar to the incidence of cIBC. This led us to the following proportions of true recurrences and new primary tumor: 1) after DCIS treatment with BCS alone, 41.6% of the subsequent ipsilateral IBCs (6.4% of 15.4%) are new primary tumors and the remaining 58.4% are true recurrences, clonally related to the initial DCIS index lesion. These numbers support our estimate that 29.5% of subsequent ipsilateral IBCs might very well be new primary tumors based on the molecular characterization.

Discussion

This study included a unique series of 78 patient-matched DCIS and subsequent ipsilateral IBC, that occurred during a median follow-up time of 12 years. Clonality analysis based on genomic aberrations showed that 52.6% of the pairs (41/78) are clonally related. This provides evidence for direct progression from *in situ* to invasive breast disease. Strikingly, our results also showed that 29.5% of the subsequent IBCs (23/78) may be new primary tumors. This high rate of new primary tumors after DCIS treated by BCS alone indicates that: 1) *true* recurrence rates after DCIS are likely to

be overestimated; and, if true, 2) DCIS might not just be a precursor, but could also be risk lesion of IBC, as has been estimated

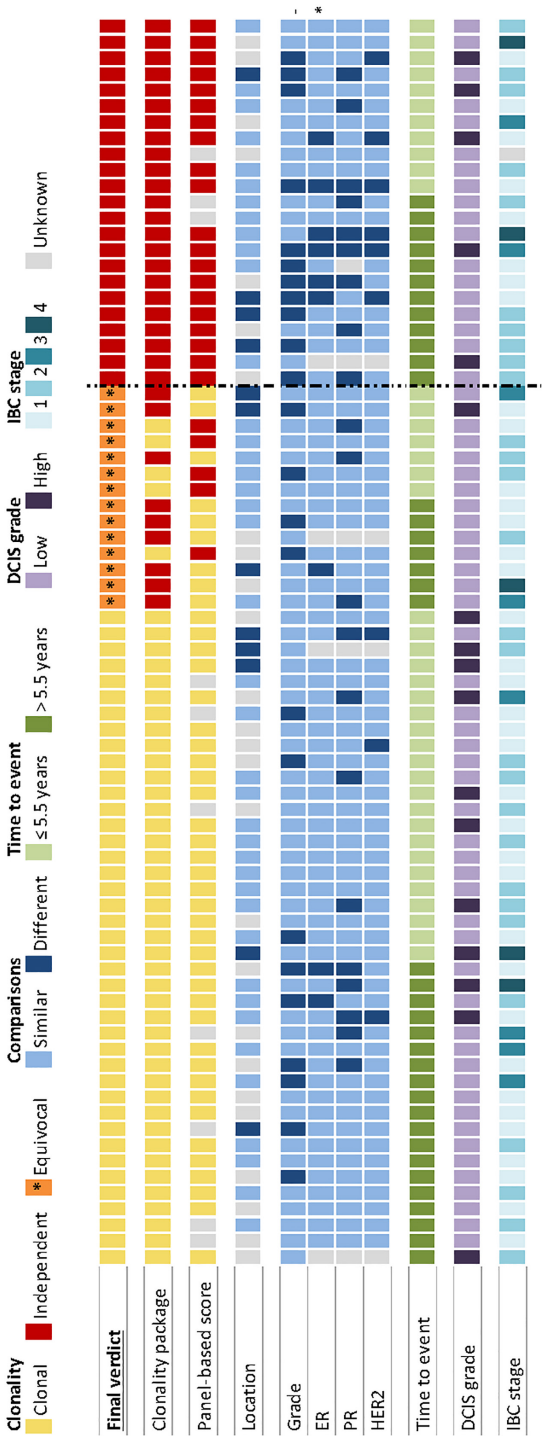


Figure 4. Overview figure of all datatypes available in this study. Discordant ER status (*) was higher among independent pairs compared to clonal pairs (p=0.021). Discordant grade (-) trended to be more often observed among independent pairs (low (1-2) vs. high (3) grade; p=0.054).

for lobular carcinoma in situ (LCIS).

These findings may impact personalized risk stratification for women with DCIS substantially, as two different risk issues have to be addressed. First, the risk of the DCIS lesion progressing to IBC. Second, the risk to develop a non-clonally related, most likely independent new IBC. As a consequence, prognostic factors studies so far might have an intrinsic limitation, as they focus on DCIS features predicting the risk of developing subsequent IBC. However, it is far less likely that the risk of a subsequent independent, new primary IBC is dictated by the initial DCIS lesion.

Currently, an ipsilateral tumor after BCS is diagnosed as a local recurrence, as it is considered outgrowth of residual cancer cells outside the surgical margins. However, studies on tumor adjacent histologically normal tissue raise the possibility that non-malignant, pre-cancerous cells within these tissues can also contribute to subsequent ipsilateral IBC.^{20–24} This hypothesis is further acknowledged by two concepts which have been recognized to underlie the development of breast cancer: the sick lobe theory and the field cancerization concept.^{25–29} The sick lobe theory postulates that breast carcinoma, both *in situ* and invasive, are lobular diseases, meaning that synchronous and metachronous lesions appear in a single lobe of the breast. Both concepts imply that the epithelial lining within a lobe can consist of cells which have undergone early genetic events (first “hit”), either during mammary development (sick lobe theory) or at unspecified time points (field cancerization concept), but have not acquired all the changes (second “hit”) necessary for tumorigenesis. This breast lobe constitutes the field cancerized tissue or “sick lobe”, and is hypothesized to be the mediator of disease progression.

An alternative explanation for finding such a high rate of new primary tumors is tumor heterogeneity. It could be that not all cells in a DCIS lesion harbor the same genetic aberrations. When analyzing bulk samples, this intra-lesional heterogeneity could be missed. The issue of intra-DCIS heterogeneity complicates clonality analysis and makes that we cannot rule out that new primary tumors originate from the index lesion.

Previous studies have attempted to distinguish NP from TR after IBC treated with BCS, with or without RT, and have reported different proportions of patients with NP, ranging from 38–52%.^{7–9,11,30,31} One other study reported a NP rate of 15%.¹⁰ All the studies reported that TRs occur sooner than NPs. Furthermore, some studies showed that TRs have worse survival rates,^{7,9,10,30,31} and were associated with involved margins.^{10,30} Patients with NPs more often developed cIBC.^{7,9,10} Our results on DCIS do not confirm these associations.

There is no consensus about which type of data (e.g. copy number, mutation, histology) and especially which statistical method is most suitable to distinguish clonal recurrences from independent primary tumors.¹⁶ Many studies have been published on tumor clonality using different data types and statistical methods,^{12,13,16,32–34} whereas other studies defined their own methods and cohort-specific cut-offs.^{14,15} Yet, studies comparing clonality assessment based on clinical and molecular data have shown that the clinical assessment

of clonality is inaccurate for most ipsilateral breast tumor recurrences.^{16,35,36} Furthermore, Bierman and colleagues have shown that the type of molecular data analysis had a stronger impact on clonality determination than the analytical methods used.¹⁶ Data suggests that mutations evolve gradually over long periods of time, generating extensive clonal diversity.³⁷ In contrast, CNAs are acquired at early stages of tumorigenesis, which makes them the most stable type of biological data you can use for clonality assessment, but depending on the methodology, it is difficult to identify CNAs present in a (small) subset of cells in a sample.^{37,38}

For the DCIS and IBC pairs in our cohort, both copy number and mutation data were available to evaluate clonal relatedness of the lesions. The results from these two methods were mostly concordant. However, in several cases, our panel-based clonality score demonstrated evidence of clonality, while the clonality package did not. This may be due to higher noise levels in the data of these pairs. In case of discordance, pairs were classified as equivocal.

Our study group comprised of patients with DCIS diagnosed between 1989 and 2004, which were treated with BCS alone. Regarding this time period, we have to keep in mind that treatment strategies and screening techniques for DCIS have evolved over the years. This may have impacted treatment or other care for these patients and with this possibly also the rates of new primary tumors. Also, adjuvant radiotherapy is nowadays standard of care after BCS for DCIS. Radiotherapy significantly reduces the risk of recurrence (both *in situ* and invasive) after DCIS.^{39,40} Fifteen years after DCIS diagnosis, cumulative incidence of ipsilateral IBC was 15.4% after BCS alone and 8.8% after BCS followed by RT.¹⁷ Based on our calculations using epidemiological data, we estimated that also the proportion of true recurrences is lower when radiotherapy is added to the treatment of women with DCIS as compared to women treated with BCS alone: 72.7% of the subsequent ipsilateral IBCs (6.4% of 8.8%) are new primary tumors and the remaining 27.3% are true recurrences. Lastly, it is important to note that we do not know what the recurrence rate is if the women in our cohort were left untreated.

Our study has several strengths. First, we made use of a unique, well-annotated set of 78 patient-matched DCIS and subsequent IBC pairs derived from a nation-wide population-based cohort of DCIS patients treated by BCS alone. Differences between DCIS and IBC could not be caused by treatment. Second, using microdissection we achieved a high neoplastic cell purity for our molecular studies. This improved the intensity of detected CNAs and small cell populations. With this we ruled out that a lack of clonal relatedness could be due to a lack of tumor cells. Third, we reached an average coverage >2000x for variant calling in the DCIS sample using PanelSeq. With this high coverage, mutations in potential DCIS subclones should have been picked up. In addition, the positions of variants solely called in the IBC lesions were checked in the matching DCIS lesion and vice versa, to reassure the absence of this variant in the paired sample. Fourth, we took along loci-specific probability

of CNA and mutation to correct for the frequency in which an aberration is found in breast cancer. Fifth, our epidemiological calculation based on same source population supports our experimental results.

Our study also has some limitations. First, with our panel-based clonality score we compared CNA and mutations based on a gene panel. This method covered only 53 genes and thus represent a part of the full picture. Whole exome sequencing would have given us more information, although with the cost of lower sequencing depth, but is not feasible for this cohort of old-FFPE derived DNA. Yet, the analysis using the clonality package did gave us a whole genome-based assessment. Second, due to the lack of normal DNA, we used germline databases (GNOMAD and GoNL) to remove common germline variants from our PanelSeq data. Because of this, we cannot fully exclude the possibility that germline mutations that are not recognized as such are called as somatic mutations. This is probably also why we detected such a high rate of BRCA2 mutations. Third, intra-lesional heterogeneity could have influence the outcome of our clonality analysis due to biological differences in different parts of a tumor and subclone evolution.

In conclusion, we provided proof of the direct progression of DCIS to IBC. Additionally, we found that many ipsilateral IBCs after BCS alone treated-DCIS are new primary tumors, challenging the current dogma that almost all subsequent ipsilateral IBC are due to progression of the initial DCIS. Furthermore, our findings will have major impact on assessing the risk of subsequent ipsilateral IBC after diagnosis and treatment of DCIS.

More research is needed to assess the role of DCIS heterogeneity in the outcome of our clonality analysis. Single cell or multi-regional sequencing could be used to explore the existence of minority DCIS subclones. Furthermore, better understanding of mammary field cancerization ("sick lobe theory") could improve informed decision-making in patient risk stratification and treatment. Ultimately, this may reveal markers that could be applied to identify and create a "map" of the affected lobe(s) within the breast and could be used complemented to margin assessment.⁴¹ Yet, one has to bear in mind that choosing optimal resection margins required balancing the competing risks of morbidities associated with more radical surgery with the increased rapid recurrence from a cancerized field left *in situ*.²⁸ Lastly, we recommend validation of our results using an independent cohort of women treated for DCIS, or a group of women that did not receive any treatment for DCIS.

Article information

Author's contributions: LLV, MH, MS, MKS, LFAW, EHL, and JW designed the study. LLV, FN, MdM, PK, and LM set-up and performed DNA and RNA extractions. MK and WB set-up and performed library preparations and sequencing. LLV, MH, and TB analysed that data. LLV, MH, TB, MS, MKS, ES, AT, AF, HD, SNZ, LFAW, EHL, and JW interpreted the data. LLV wrote the report. All authors provided comments and approved the final version.

Funding: This work was jointly funded by Cancer Research UK and the Dutch Cancer Society (grant number C38317/A24043, to J. Wesseling).

Acknowledgements: We thank all patients of which stored residual material was used in this study, all collaborating hospitals and in particular the pathology departments, and all persons who helped in the process of data collection and analysis. We would like to specifically acknowledge Otto Visser, Annemarie Eeltink, and the registration teams of the Netherlands Comprehensive Cancer Organization for collection of the data for the Netherlands Cancer Registry. We thank Lucy Overbeek and PALGA, the nation-wide histopathology and cytopathology data network and archive, for providing pathology data and for their help in the collection of the residual patient material. We acknowledge the staff of the NKI-AVL Core Facility Molecular Pathology & Biobanking for their technical support and the staff of the NKI-AVL Genomics Core Facility and Ronald van Marion of the Erasmus University Medical Center for their sequencing support.

Supplementary data: Supplementary File 1-5 will be available after publication of the article.

Methods

Study population

We made use of a nation-wide population-based cohort including all women diagnosed and treated for DCIS in the Netherlands between 1989 and 2005, with a median follow-up time of 12.0 years. Details on this cohort are published elsewhere.¹⁷ In brief, clinicopathological information was obtained from the Netherlands Cancer Registry (NCR) and the nation-wide histopathology and cytopathology data network (PALGA). The cohort included 10,090 women of which 2,658 were treated with BCS alone. Patients did not receive tamoxifen or other anti-hormonal adjuvant treatment. Fifteen years after DCIS diagnosis, 374 iIBC events were observed after DCIS treatment with BCS alone (cumulative incidence of iIBC of 15.4%). Formalin-fixed paraffin-embedded (FFPE) tissue specimens of patient-matched DCIS and subsequent iIBC were successfully retrieved and pathologically re-examined of 155 women.¹⁸ DNA and RNA, of sufficient

yield and quality, was successfully extracted from both DCIS and IBC of 78 women. Data on tumor location (ICD-10) was available for n=52 pairs. DCIS lesion size (n=21) and margin status (n=61) was extracted from pathology reports. Immunohistochemical data on ER, PR, and HER2 was available for n=74, n=73, and n=74 pairs, respectively.⁴²

The study was approved by the review board on the NKR (request K12.281; January 3, 2013) and PALGA (LZV990; April 16, 2013). Secondary use of tissue and data for this study was done under an opt-out regime which is conform Dutch regulations and the Code of Conduct of Federa-COREON.

Tissue processing and microdissection

For laser microdissection, 8- μ m-thick FFPE tissue section were mounted on PolyEthylene Naphthalate (PEN) slides. Prior to sectioning, the slides were UV-treated for 30 minutes at 254 nm, to overcome the membrane's hydrophobic nature and to allow better section adherence. Sections were incubated overnight at 56°C. After deparaffinization and rehydration, FFPE sections were stained with 1% toluidin blue for 5 seconds to allow identification of epithelial areas of DCIS, followed by a washing step in distilled water and dehydration in serial ethanol dilutions (70, 90, 100%).

Microdissection of in situ lesions was performed using a laser microdissection (LMD) microscope (Leica Microsystems). Invasive lesions were microdissected from 10- μ m-thick sections either using the LMD system or manually using a scalpel. The number of sections per case used for microdissection varied by lesions size and cellularity, with an average of 8 sections per in situ lesion (range 2-15) and an average of 9 sections per invasive lesion (range 5-10). The microdissected tissue fragments were collected in the cap of a 0.5ml tube containing 30 μ l PKD buffer, the first reagent of the extraction method. Tissue fragments were stored at 4°C for a maximum of two days until DNA and RNA extraction or were processed immediately.

DNA and RNA extraction

DNA and RNA were simultaneously extracted from (laser) microdissected tissue fragments using the Qiagen AllPrep DNA/RNA FFPE kit (QIAGEN), following manufacturer's instructions. DNA samples were RNase treated. DNA concentrations were measured using the Qubit dsDNA High-Sensitivity assay and Qubit 2.0 Fluorometer (Life Technologies). Quality and quantity of total RNA from FFPE tissue was assessed using a Bioanalyzer 2100 (Agilent) and the percentage RNA fragments >200bp (DV200) were determined.

RNA samples were submitted for RNA sequencing and DNA samples were submitted for low coverage whole genome sequencing and panel sequencing (PanelSeq). In cases with sufficient DNA, we performed both assays. In cases with insufficient DNA for both assays, only low coverage whole genome sequencing was performed. An overview of the available datasets per sample is presented in Supplementary File 3.

Whole transcriptome RNA sequencing (RNAseq)

Total RNA samples were converted to strand-specific libraries using TruSeq RNA Access Library Prep kit (Illumina, RS-301-2001/2), according manufacturer's instructions, protocol part #15049525 Rev. B. The libraries were sequenced with 65-bp single-reads using a HiSeq 2500 instrument (Illumina). With

this, an average sequence depth was reached of 20 million reads per sample. Reads were aligned to the reference genome GRCh38 (hg38), using BWA-MEM.⁴³

Low coverage whole genome sequencing (CNVseq)

DNA was sheared to 160-200bp using the Covaris S220 Focused-ultrasonicator, cleaned with SPRI beads, analyzed for proper length on Perkin Elmer GX and library prepped with the KAPA hyper prep kit (KAPA Biosystems), protocol KR0961-v5.16. Agilent S5XT-2 (1-96) adapters with Illumina P5 and P7 sequences were used, containing 8bp Agilent indices. All samples received 9 PCR cycles, a SPRI beads cleanup and Perkin Elmer GX quality control.

Each sample was sequenced single-end 65bp on a HiSeq 2500 instrument (Illumina), one pool per lane, to obtain low coverage (0.2x) whole genome sequencing data. Reads were aligned to the reference genome GRCh38 (hg38) using BWA-MEM.⁴³ For each sample, per bin of 20kb, reads of a mapping quality of >37 were counted and read counts were rated against a similar mapping of all attainable 65bp sequences on the reference genome. Sample counts were corrected per bin for local GC effects using a non-linear loess fit of mappability over 0.8 on autosomes. Reference values were scaled according to the slope of a linear fit, forced to intercept at the origin, of reference mappability after GC correction. Bins with a mappability below 0.2 or overlapping ENCODE blacklisted regions were removed from further analysis.⁴⁴

Copy number segmentation was performed using DNACopy;⁴⁵ genome-wide copy number frequency plots were created using CGHcall.⁴⁶ Segments were identified as lost or gained at log2 ratios of <-1 and >1.5 , respectively. To clean up the data, we removed losses in noisy samples (>1000 segments). Lost and gained segments were subsequently annotated with overlapping genes using BEDTools intersect.⁴⁷

Panel sequencing

Samples were screened for mutations with the Ion Torrent Personal Genome Machine (PGM; Life Technologies), using a custom-made amplicon panel encompassing 2778 amplicons covering 53 genes including hotspot mutation regions and single nucleotide polymorphisms (SNPs). Genes included in this panel were selected based on a recent publication describing frequently found breast cancer drivers and by consulting the METABRIC and ICGC/TCGA breast cancer datasets.^{19,48,49} Amplicons for the multiplex PCR assay were designed using the IonTorrent AmpliSeq Designer tool, aiming for 150-bp amplicons and allowing efficient amplification of fragmented DNA isolated from FFPE tissue specimens.

Samples analyzed using the IonTorrent AmpliSeq custom 53-gene panel were processed according to the Ion AmpliSeq Library Kit Plus protocol (ThermoFisher Scientific). The multiplexed PCR was split into two reactions of 10 ng of DNA, allowing the amplification and sequencing of overlapping amplicons, required to obtain full sequence coverage of large exons. Each sample was barcoded using IonTorrent Xpress bar-coded adapters, allowing multiplexed sequencing. A total of 19 PCR cycles were performed on the FFPE samples. Ten samples were multiplexed on an Ion 540 Chip and sequenced on the Ion S5XL Semiconductor sequencer. Reads were aligned to the reference genome GRCh37 (hg19) using the Torrent Mapping Alignment Program, and variant calling was performed using Torrent Variant Caller (TVC) version 5.6.

Variant data in VCF format was first translated to GRCh38 and annotated using bedtools, picard (<https://broadinstitute.github.io/picard/command-line-overview.html>), samtools, bcftools and VEP, and further analyzed in R, employing vcfr, and tidy verse.^{47,50–52} True somatic variants, identified via filtering during which low quality variants (variant allele frequency [VAF] <10%, coverage <100x, and a quality [QUAL] of <1000), artifacts (found in >90% of samples), and germline variants (>5 cases in GNOMAD and GoNL; except for BRCA1, BRCA2, and TP53) were removed, and validated in RNAseq data using samtools mpileup.^{50,53} Details regarding the amplicon panel design, performance, and filtering strategy are provided in Supplementary File 4-5.

Based on germline mutation data, we confirmed that patient-matched DCIS and IBC pairs indeed originated from the one patient.

Statistical analysis

We determined if the patient-matched DCIS and subsequent iIBC pairs had a shared clonal origin or represented two independent lesions, by evaluating patterns of shared and unique somatic copy number alterations and mutations. For this, two independent methods were used. Clonality package was used to define the likelihood ratio with individual comparisons (LR2) and LR2 *p*-value.⁵⁴ Using this package, pairs were classified as either clonal or independent, based on copy number data. Both true pairs (test set) and artificial pairs (independent pairs; reference set) were analyzed using the clonality package. The artificial pairs are obtained by constructing pairs of lesions from different patients. Since these pairs are definitely independent they provide the reference distribution, which is used to permit calculation of the *p* values of the real paring. A cutoff for clonal origin and independence was set on *p*<0.1. In addition, we developed a panel-based clonality score making used of CNA and mutation data. This score is restricted to our panel of 53 driver genes and was computed as follows: 1) loci-specific probabilities of observing a somatic mutation/CNA were obtained from the METABRIC breast cancer dataset;^{19,49} 2) a score was computed based on the sum of shared and unique aberrations for every patient, corrected for the probability and squared to give less weight to highly common, and more weight to highly rare aberrations (e.g. PIK3CA mut 17% = $(1-0.17)^2$); 3) the following formula was applied: $\text{score} = \log(\text{"shared-score"} + 1) / (\text{"unique-score"} + 1)$. The panel-based clonality score was computed for both true pairs (test set) and artificial pairs (reference set). The reference set included 80 extra DCIS samples without a matching IBC lesion. A cutoff for clonal origin and independence was set on the 99th percentile of the artificial pairs' scores.

Pairs classified as "independent" based on the Clonality package and with a low panel-based clonality score were considered new primary tumors (NP). Pairs classified as "clonal" and with a high panel-based clonality score were considered true recurrences (TR).

We used t-test and Fisher's exact test to ascertain differences in clinicopathological characteristics between TR and NP. *P* values ≤0.05 were considered statistically significant.

References

1. Barnes NLP, Ooi JL, Yarnold JR, Bundred NJ. Ductal carcinoma in situ of the breast. *BMJ*. 2012;344:e797.
2. To T, Wall C, Baines CJ, Miller AB. Is Carcinoma in Situ a Precursor Lesion of Invasive Breast Cancer? *Int J Cancer*. March 2014;1-7.
3. Levi F, Randimbison L, Te V-C, La Vecchia C. Invasive breast cancer following ductal and lobular carcinoma in situ of the breast. *Int J cancer*. 2005;116(5):820-823.
4. Ernster VL, Ballard-Barbash R, Barlow WE, et al. Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst*. 2002;94(20):1546-1554.
5. Netherlands Comprehensive Cancer Organization (IKNL). DCIS incidence trends over time, the Netherlands.
6. Howlader N, Noone AM, Krapcho M, et al. *SEER Cancer Statistics Review, 1975-2014*. National Cancer Institute. Bethesda, MD
7. Huang E, Buchholz TA, Meric F, et al. Classifying local disease recurrences after breast conservation therapy based on location and histology: new primary tumors have more favorable outcomes than true local disease recurrences. *Cancer*. 2002;95(10):2059-2067.
8. Freedman GM, Anderson PR, Hanlon AL, Eisenberg DF, Nicolaou N. Pattern of local recurrence after conservative surgery and whole-breast irradiation. *Int J Radiat Oncol Biol Phys*. 2005;61(5):1328-1336.
9. Nishimura S, Takahashi K, Akiyama F, et al. Classification of ipsilateral breast tumor recurrence after breast-conserving therapy: new primary cancer allows a good prognosis. *Breast Cancer*. 2005;12(2):112-117.
10. Komoike Y, Akiyama F, Iino Y, et al. Analysis of ipsilateral breast tumor recurrences after breast-conserving treatment based on the classification of true recurrences and new primary tumors. *Breast Cancer*. 2005;12(2):104-111.
11. Panet-Raymond V, Truong PT, McDonald RE, et al. True recurrence versus new primary: an analysis of ipsilateral breast tumor recurrences after breast-conserving therapy. *Int J Radiat Oncol Biol Phys*. 2011;81(2):409-417.
12. Banelli B, Casciano I, di Vinci A, et al. Pathological and molecular characteristics distinguishing contralateral metastatic from new primary breast cancer. *Ann Oncol*. 2009;21(6):1237-1242.
13. Begg CB, Ostrovnaya I, Carniello JVS, et al. Clonal relationships between lobular carcinoma in situ and other breast malignancies. *Breast Cancer Res*. 2016;18(1):1-11.
14. Bollet MA, Servant N, Neuvial P, et al. High-Resolution Mapping of DNA Breakpoints to Define True Recurrences Among Ipsilateral Breast Cancers. *JNCI J Natl Cancer Inst*. 2008;100(1):48-58.
15. Hwang ES, Nyante SJ, Chen YY, et al. Clonality of lobular carcinoma in situ and synchronous invasive lobular carcinoma. *Cancer*. 2004;100(12):2562-2572.
16. Biermann J, Parris TZ, Nemes S, et al. Clonal relatedness in tumour pairs of breast cancer patients. *Breast Cancer Res*. 2018;20(1):96.
17. 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 Res Treat*. 2016;159(3):553-563.
18. Visser LL, Elshof LE, van de Vijver K, et al. Discordant marker expression between invasive breast carcinoma and corresponding synchronous and preceding DCIS. *Am J Surg Pathol*. 2019.
19. Curtis C, Shah SP, Chin S-F, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403).
20. Li Z, Moore DH, Meng ZH, Ljung B-M, Gray JW, Dairkee SH. Increased risk of local recurrence is associated with allelic loss in normal lobules of breast cancer patients. *Cancer Res*. 2002;62(4):1000-1003.
21. Larson PS, De las Morenas A, Bennett SR, Cupples LA, Rosenberg CL. Loss of heterozygosity or allele imbalance in histologically normal breast epithelium is distinct from loss of heterozygosity or allele imbalance in co-existing carcinomas. *Am J Pathol*. 2002;161(1):283-290.

22. Heaphy CM, Bisoffi M, Fordyce CA, et al. Telomere DNA content and allelic imbalance demonstrate field cancerization in histologically normal tissue adjacent to breast tumors. *Int J cancer*. 2006;119(1):108-116.
23. Forsberg LA, Rasi C, Pekar G, et al. Signatures of post-zygotic structural genetic aberrations in the cells of histologically normal breast tissue that can predispose to sporadic breast cancer. *Genome Res*. 2015;25(10):1521-1535.
24. Takeshima H, Ushijima T. Accumulation of genetic and epigenetic alterations in normal cells and cancer risk. *npj Precis Oncol*. 2019;3(1):7.
25. Tot T. DCIS, cytokeratins, and the theory of the sick lobe. *Virchows Arch*. 2005;447(1):1-8.
26. Heaphy CM, Griffith JK, Bisoffi M. Mammary field cancerization: molecular evidence and clinical importance. *Breast Cancer Res Treat*. 2009;118(2):229-239.
27. Lochhead P, Chan AT, Nishihara R, et al. Etiologic field effect: reappraisal of the field effect concept in cancer predisposition and progression. *Mod Pathol*. 2015;28(1):14-29.
28. Curtius K, Wright NA, Graham TA. An evolutionary perspective on field cancerization. *Nat Rev Cancer*. 2018;18(1):19-32.
29. Tan MP, Tot T. The sick lobe hypothesis, field cancerisation and the new era of precision breast surgery. *Gland Surg*. 2018;7(6):611-618.
30. Haffty BG, Carter D, Flynn SD, et al. Local recurrence versus new primary: clinical analysis of 82 breast relapses and potential applications for genetic fingerprinting. *Int J Radiat Oncol Biol Phys*. 1993;27(3):575-583.
31. Smith TE, Lee D, Turner BC, Carter D, Haffty BG. True recurrence vs. new primary ipsilateral breast tumor relapse: an analysis of clinical and pathologic differences and their implications in natural history, prognoses, and therapeutic management. *Int J Radiat Oncol Biol Phys*. 2000;48(5):1281-1289.
32. Nemes S, Danielsson A, Parris TZ, et al. A diagnostic algorithm to identify paired tumors with clonal origin. *Genes Chromosomes Cancer*. 2013;52(11):1007-1016.
33. Andrade VP, Ostrovnaya I, Seshan VE, et al. Clonal relatedness between lobular carcinoma in situ and synchronous malignant lesions. *Breast Cancer Res*. 2012;14(4):R103.
34. Mauguen A, Seshan VE, Ostrovnaya I, Begg CB. Estimating the probability of clonal relatedness of pairs of tumors in cancer patients. *Biometrics*. 2018;74(1):321-330.
35. Goldstein NS, Vicini FA, Hunter S, et al. Molecular clonality determination of ipsilateral recurrence of invasive breast carcinomas after breast-conserving therapy: comparison with clinical and biologic factors. *Am J Clin Pathol*. 2005;123(5):679-689.
36. McGrath S, Antonucci J, Goldstein N, et al. Long-term patterns of in-breast failure in patients with early stage breast cancer treated with breast-conserving therapy: a molecular based clonality evaluation. *Am J Clin Oncol*. 2010;33(1):17-22.
37. Wang Y, Waters J, Leung ML, et al. Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature*. 2014;512(7513):155-160.
38. Gao R, Davis A, McDonald TO, et al. Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat Genet*. 2016;(August):1-15.
39. Wapnir IL, Dignam JJ, Fisher B, 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. *J Natl Cancer Inst*. 2011;103(6):478-488.
40. Donker M, Liti??re S, Werutsky G, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma In Situ: 15-year recurrence rates and outcome after a recurrence, from the EORTC 10853 randomized phase III trial. *J Clin Oncol*. 2013;31(32):4054-4059.
41. Lebya K, Garcia-Smith R, Swaminathan R, et al. Towards a personalized surgical margin for breast conserving surgery-Implications of field cancerization in local recurrence. *J Surg Oncol*. 2017;115(2):109-115.
42. Visser LL, Elshof LE, Schaapveld M, et al. Clinicopathological Risk Factors for an Invasive Breast Cancer Recurrence after Ductal Carcinoma In Situ-A Nested Case-Control Study. *Clin Cancer Res*. April 2018:clincanres.0201.2018.
43. Li H. Aligning sequence reads, clone

- sequences and assembly contigs with BWA-MEM. March 2013.
44. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489(7414):57-74.
 45. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. *Bioinformatics*. 2007;23(6):657-663.
 46. Van de Wiel M, Vosse S. CGHcall: Calling aberrations for array CGH tumor profiles. 2018.
 47. Quinlan AR. BEDTools : the Swiss-army tool for genome feature analysis. *Curr Protoc Bioinforma*. 2014.
 48. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605):47-54.
 49. Pereira B, Chin S-F, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. *Nat Commun*. 2016;7(May):11479.
 50. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*. 2011;27(21):2987-2993.
 51. McLaren W, Gil L, Hunt SE, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 2016;17(1):122.
 52. Knaus BJ, Grünwald NJ. vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour*. 2017;17(1):44-53.
 53. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res*. 2002;12(6):996-1006.
 54. Ostrovskaya I, Seshan VE, Olshen AB, Begg CB. Clonality: an R package for testing clonal relatedness of two tumors from the same patient based on their genomic profiles. *Bioinformatics*. 2011;27(12):1698-1699.

Supplementary tables and figures

Table S1. Comparison of clinipathological characteristics between clonal pairs and independent pairs

	Clonal pairs n = 41		Independent pairs n = 23		P
Age at DCIS diagnosis in years, mean (range)	58.5	(33.9-86.5)	58.0	(38.2-86.7)	0.87 ^a
DCIS lesion size in mm, mean (range)	15.9	(4-30)	11.75	(2-20)	0.21 ^a
DCIS lesion size					
1-10 mm	5	(12.2%)	4	(17.4%)	0.32 ^b
11-20 mm	5	(12.2%)	4	(17.4%)	
>20 mm	4	(9.8%)	0	(0.0%)	
Unknown	27	(65.9%)	15	(65.2%)	
Margin status					
Free	27	(65.9%)	15	(65.2%)	0.70 ^b
Involved	7	(17.1%)	2	(8.7%)	
Unknown	7	(17.1%)	6	(26.1%)	
Time to iIBC event in years, mean (range)	5.6	(0.5-15.7)	6.8	(0.5-17.0)	0.24 ^a
Same breast quadrant					
Yes	20	(44.4%)	12	(52.2%)	0.72 ^b
No	5	(11.1%)	4	(17.4%)	
Unknown	16	(35.6%)	7	(30.4%)	
Grade *					
Concordant	32	(78.0%)	12	(52.2%)	0.049 ^b
Discordant	9	(22.0%)	11	(47.8%)	
ER					
Concordant	37	(90.2%)	16	(69.6%)	0.021 ^b
Discordant	2	(4.9%)	6	(26.1%)	
NA	2	(4.9%)	1	(4.3%)	
PR					
Concordant	30	(73.2%)	12	(52.2%)	0.14 ^b
Discordant	9	(22%)	9	(39.1%)	
NA	2	(4.9%)	2	(8.7%)	
HER2					
Concordant	36	(87.8%)	16	(69.6%)	0.06 ^b
Discordant	3	(7.3%)	6	(26.1%)	
NA	2	(4.9%)	1	(4.3%)	

Table S1. Comparison of clinipathological characteristics between clonal pairs and independent pairs

	Clonal pairs		Independent pairs		P
	n = 41		n = 23		
DCIS					
Histologic grade *					
Low	30	(73.2%)	18	(78.3%)	0.77 ^b
High	11	(26.8%)	5	(21.7%)	
ER status					
Positive	31	(75.6%)	18	(78.3%)	1.00 ^b
Negative	8	(19.5%)	4	(17.4%)	
NA	2	(4.9%)	1	(4.3%)	
PR status					
Positive	20	(48.8%)	14	(60.9%)	0.29 ^b
Negative	19	(46.3%)	7	(30.4%)	
NA	2	(4.9%)	2	(8.7%)	
HER2 status					
Positive	17	(41.5%)	4	(17.4%)	0.054 ^b
Negative	22	(53.7%)	18	(78.3%)	
NA	2	(4.9%)	1	(4.3%)	
IHC subtype					
HR+ HER2-	21	(51.2%)	17	(73.9%)	0.12 ^b
HR+ HER2+	10	(24.4%)	1	(4.3%)	
HR- HER2+	7	(17.1%)	3	(13%)	
HR- HER2-	1	(2.4%)	1	(4.3%)	
NA	2	(4.9%)	1	(4.3%)	
IBC					
Histologic type					
NST	39	(95.1%)	21	(91.3%)	0.71 ^b
Lobular	1	(2.4%)	0	(0.0%)	
Mixed	1	(2.4%)	2	(8.7%)	
Status					
1	21	(51.2%)	8	(34.8%)	0.52 ^b
2A	10	(24.4%)	6	(26.1%)	
2B	4	(9.8%)	4	(17.4%)	
3A	4	(9.8%)	1	(4.3%)	
3B	0	(0.0%)	1	(4.3%)	
4	2	(4.9%)	2	(8.7%)	
NA	0	(0.0%)	1	(4.3%)	
Status (3)					
1	21	(51.2%)	8	(34.8%)	0.30 ^b
Other	20	(48.8%)	14	(60.9%)	
NA	1	(2.4%)	1	(4.3%)	
Grade *					
Low	21	(51.2%)	13	(56.5%)	0.80 ^b
High	20	(48.8%)	10	(43.5%)	

Table S1. Comparison of clinipathological characteristics between clonal pairs and independent pairs

	Clonal pairs		Independent pairs		<i>P</i>
	n = 41		n = 23		
ER status					
Positive	33	(80.5%)	16	(69.6%)	0.366 ^b
Negative	8	(19.5%)	7	(30.4%)	
PR status					
Positive	19	(46.3%)	12	(52.2%)	0.795 ^b
Negative	22	(53.7%)	11	(47.8%)	
HER2 status					
Positive	17	(41.5%)	2	(8.7%)	0.009 ^b
Negative	24	(58.5%)	21	(91.3%)	
IHC subtype					
HR+ HER2-	22	(53.7%)	16	(69.6%)	0.006 ^b
HR+ HER2+	11	(26.8%)	0	(0.0%)	
HR- HER2+	6	(14.6%)	2	(8.7%)	
HR- HER2-	2	(4.9%)	5	(21.7%)	

Abbreviations: mm, millimeter; IIBC, ipsilateral invasive breast cancer; ER, estrogen receptor; PR, progesterone receptor; NA, not assessable (NAs were not included in the analysis); HR, hormone receptor (ER and/or PR); NST, no special type.

*: Low grade group consists of grade 1 and 2, and high grade group consists of grade 3.

Comparisons between clonal and independent pairs were made by t-test (^a) or Fisher's exact test (^b).

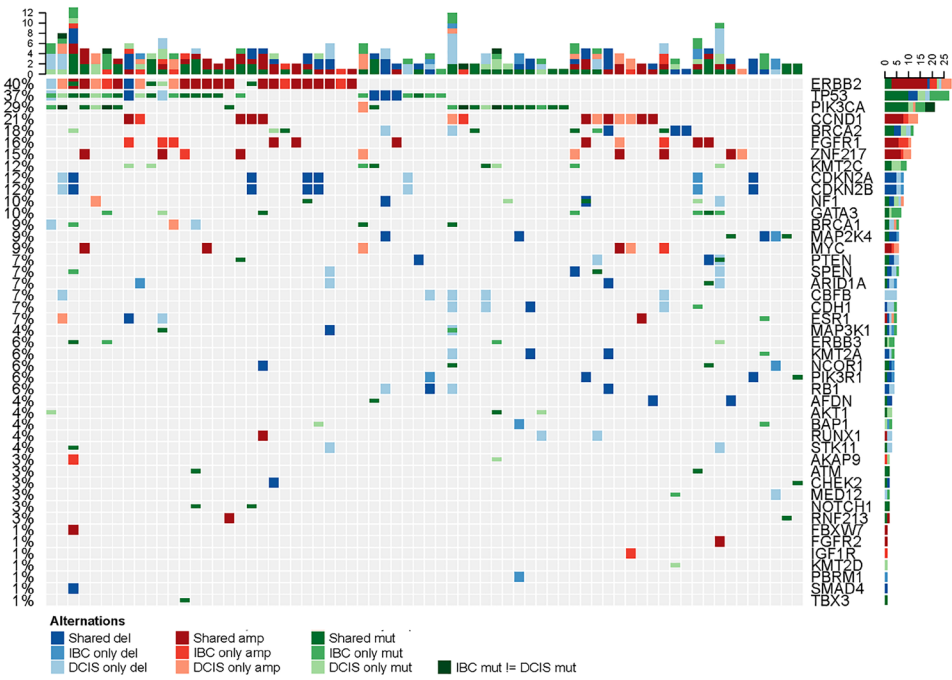


Figure S1. OncoPrint of 68 patient-matched DCIS and subsequent iIBC. Ten patients did not harbor any somatic events in the 53 assessed genes. IBC mut != DCIS mut, gene mutation in both lesions but different mutation.

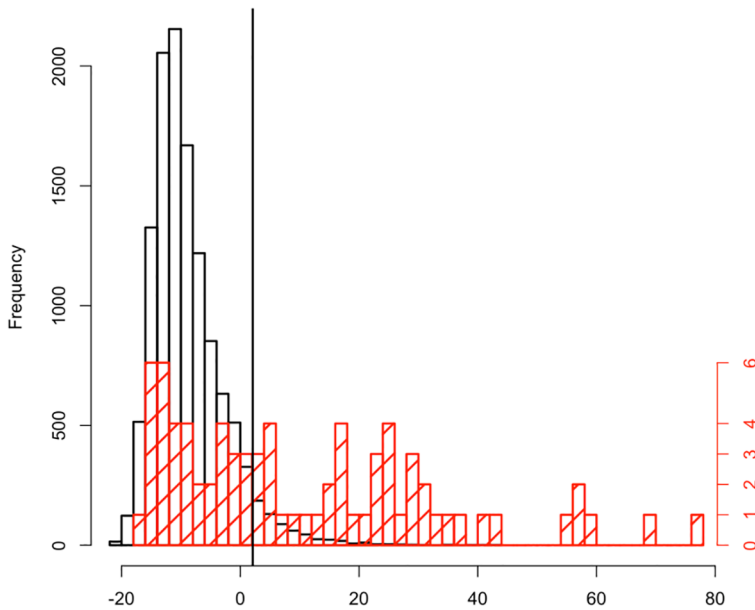


Figure S2. Distribution of the measure of clonality outputted by the Clonality package. Clonality measure of true pairs (test set) is depicted in red and artificial pairs (reference set) in black. The scale of the horizontal axis is the log-likelihood (LogLR), the measure of evidence in favor of clonality. The vertical axes represent the frequencies of pairs: left axis, black (reference set) histogram; right axis, red (test set) histogram. A cutoff for clonal origin or independence was set on $p < 0.1$ (black vertical line).

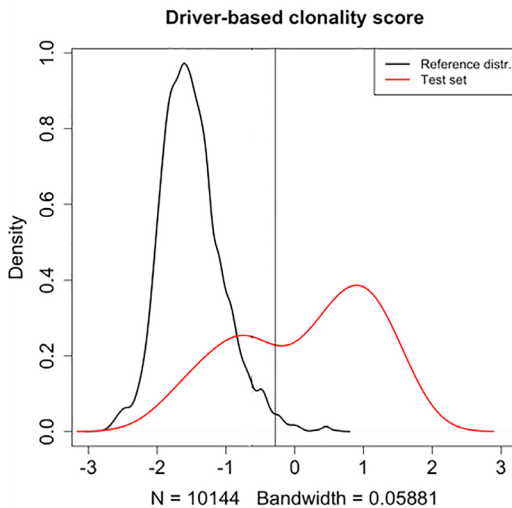


Figure S3. Distribution of the measure of clonality outputted by the panel-based clonality score. Clonality measure of true pairs (test set) is depicted in red and artificial pairs (reference set) in black. A cutoff for clonal origin and independence was set on the 99th percentile of the artificial pairs' scores (black vertical line). $N = 10144$ is number of pairs in the reference distribution.

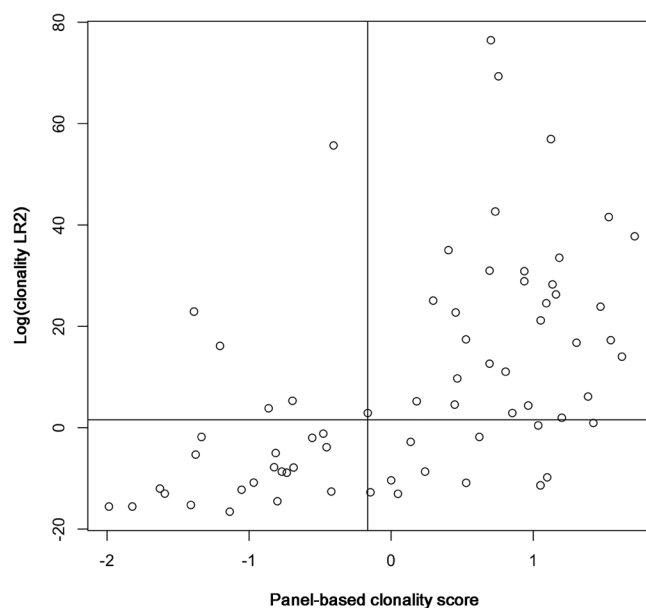


Figure S4. Correlation plot of clonality measure outputted by the clonality package (LR2 value; y-axis), based on whole genome copy number data, and the panel-based clonality score score (x-axis) which is based on copy number and mutation data, restricted to 53 breast cancer driver genes. Lines represent the used cutoffs. Pairs with a value above the cutoff are clonal and pairs with a value below the cutoff are independent. Agreement between the two methods is 79%.

CHAPTER 7

Summarizing discussion

The general aims of this thesis were to identify prognostic markers predictive of the development of subsequent ipsilateral invasive breast cancer (iIBC) after DCIS and to explore the clonal relatedness of patient-matched DCIS and subsequent iIBC. In this concluding chapter we will discuss the main findings and interpret them in a broader context. The methodological challenges of performing observational research and the strengths and limitations of our studies are discussed. Finally, recommendations for future research and clinical implications are given.

Main findings in context of other literature

Risk factors for an invasive breast cancer after DCIS

There is a high need for robust prognostic markers that can reliably predict the disease course of DCIS. Numerous prognostic factors have been reported by previous studies, but none of them have shown to be of sufficient value for implementation into the clinic. So, what do we really know? To answer this question, we performed a systematic review and meta-analyses (**chapter 2**). Previously, three meta-analyses have been published on prognostic factors for recurrence after DCIS, although only one specifically focused on ipsilateral *invasive* recurrence after DCIS.¹⁻³ Yet, we are the first to perform bias assessment on prognostic factors studies for DCIS and incorporate the results into meta-analyses. We concluded that the following six factors had the strongest predictive value for subsequent IBC after DCIS: African-American race, premenopausal status, detection by palpation, involved margins, high histologic grade, and high p16 protein expression. Common biases were insufficient measurement and handling of confounders and poorly described study groups.

In **chapter 3** the results of our own prognostic factor study for DCIS are described. The large size of our series, the case-control design, the comprehensive data, and the long-term follow-up were essential to overcome biases, as often seen in previous studies. Our results showed a prognostic role for COX-2, HER2, and periductal fibrosis. Previous studies support these findings.⁴⁻⁶ In addition, we found a 4-fold higher prevalence of subsequent iIBC for women diagnosed with HER2⁺/COX-2^{high} DCIS as compared to women with HER2⁻/COX-2^{low} DCIS lesions. Furthermore, patients with COX-2^{low} DCIS were at lower risk of iIBC as their risk was comparable to the general population.

More information on the future disease course of DCIS may be found in the genomic and epigenomic profile present at diagnosis. This hypothesis is supported by the fact that DCIS and IBC are highly similar at the molecular level.^{7,8} Previous molecular profiling studies on DCIS mainly focused on synchronous DCIS and IBC or pure DCIS and unmatched IBC lesions (from different individuals), whereby the two components were compared to find differences that may be related to invasion.⁸⁻¹³ To date, only two studies evaluated the association of molecular alterations and recurrence (both *in situ* and invasive).^{7,14} Furthermore, the

Oncotype DX DCIS score is the first and only commercially available multigene expression panel. This score enables risk prediction of local recurrence (in situ and invasive) after BCS for DCIS and estimates the benefit from RT.¹⁵ Yet, prospective validation of the score is still lacking.

Despite the large patient numbers, our gene expression analysis comparing DCIS associated with and without subsequent IBC did not result in the identification of markers statistically significantly associated with subsequent IBC (**chapter 4**). This is in line with a previous study, showing that expression analysis is compromised when unmatched for intrinsic subtype as profiles are strongly driven by the PAM50 intrinsic subtypes.¹⁶ Gene expression analysis, stratified by PAM50 subtype, showed that immune-related pathways play a role in the progression of DCIS. With this, we showed that although the subtypes have limited prognostic value for DCIS, it is still of value to stratify lesions based on intrinsic subtype because of distinct evolutionary disease paths. Additionally, we found that DCIS associated with subsequent IBC harbored more CNAs as compared to DCIS without invasive recurrence, which is in line with literature.¹⁴ Alternatively, we might have to focus on the tumor microenvironment, epigenetics, genetic predispositions, and lifestyle factors in combination with integrative analysis to distinguish DCIS that will potentially progress to IBC from the ones that will not.

IBC and corresponding synchronous and preceding DCIS

As described above, studies aiming to find markers related to DCIS progression and invasion generally have compared characteristics of IBC lesions with those of adjacent synchronous DCIS. The question remains whether synchronous DCIS and IBC comparison are a good surrogate of primary DCIS and subsequent IBC. In **chapter 5**, we demonstrated that marker expression between primary DCIS and subsequent IBC is less concordant than synchronous DCIS and IBC. HER2 marker expression showed the largest discrepancy: 36% of HER2-positive primary DCIS lesions were followed by HER2-negative IBC. In previous studies, numbers of matched DCIS and IBC pairs were too small to notice the major finding presented here.^{17,18} Intra-lesional heterogeneity was identified as a possible cause of the observed discordant marker expression.

Clonal relatedness of patient-matched DCIS and subsequent IBC

DCIS is regarded as a precursor of invasive breast cancer. Yet, not all DCIS progresses into IBC and proof of direct progression is lacking. To provide evidence of the direct progression of DCIS to IBC, we assessed the clonal-relatedness of patient-matched DCIS and subsequent IBC. Traditionally, whether a subsequent cancer is related to the preceding one, is primarily based on comparison of clinical and histopathologic information.^{19–23} More recently, genetic markers have started to play an increasingly important role in assessing the clonal relatedness of two lesions.^{24–28}

Strikingly, we found that approximately 30% of the subsequent IBC after DCIS treated with breast conserving surgery alone may be independent new primary cancers (**chapter 6**). This high rate of new primary tumors indicates that: 1) true recurrence rates after DCIS are likely to be overestimated; and, if true, 2) DCIS might not just be a precursor, but could also be risk lesion of IBC.

Strengths, limitations and methodological challenges

Access to large, well-annotated patient cohort

A strength of the studies presented in this thesis is that we made use of a case-case control study nested in a nation-wide population-based cohort of all women diagnosed with and treated for DCIS between 1989-2005 within the Netherlands.²⁹ Women with subsequent iIBC (cases) were matched to women without subsequent iIBC (controls) by age and follow-up time (**chapter 3 and 4**). The women in this cohort have been considered to represent the Dutch population of women with DCIS.²⁹ We were fortunate to have access to data of the Netherlands Cancer Registry (NCR), collected by the Netherlands Comprehensive Cancer Organization, and pathology data from PALGA, the nationwide histopathology and cytopathology data network and archive. Data provided by the NCR included information on age at and data of DCIS diagnosis, histology and treatment for DCIS, and any subsequent IBCs as well as any other history of cancer. The well-annotated nature of our patient series enabled us to prevent a great amount of bias often seen in previous studies. Yet, our cohort did have missing data on margin status and DCIS lesions size (15% and 65%) and information on and family history of breast cancer was unavailable. Still, the work described in **chapter 3, 4, 5, and 6** account for the high need of well-designed studies with large patient numbers.

Use of registry-based data

A challenge of the studies described in **chapter, 3, 4, 5, and 6** is that they are retrospective studies. A general concern when working with retrospective data is that you are dependent on the size, quality, and completeness of relevant variables of the registry on which the study is based.³⁰ Data quality related to end-point measures in registries often is incomplete, as end-point information such as migration abroad or death from other causes is not always included.^{31,32} This challenge can only be solved by improving source data.

Availability of tumor tissue and high quality data due to central pathology review

With the help of PALGA, we succeeded in retrieving 85% of the requested material from 58 participating hospitals within the Netherlands. Next to the DCIS specimens, we also collected the invasive recurrences of women that experience subsequent IBC. This enables

us to compare both DCIS associated with and without a subsequent invasive breast cancer event, but also to study the clonal relatedness of patient-matched DCIS and subsequent IBC. New whole slides were developed for reassessment by specialized breast pathologists, yielding in high quality pathology data. In addition, all IHC stainings within our studies were performed in one center and scored by a panel of observers with good inter-observer agreement. With this, inter-laboratory and inter-observer variability was prevented.^{33,34}

Use of a patient group that was treated for DCIS

Our study group comprised of women diagnosed with DCIS between 1989 and 2005 and treated by BCS alone. We refrained from including women who also received adjuvant radiotherapy to avoid confounding by radiation effects. Nonetheless, we have to keep in mind that treatment strategies, and also screening techniques, for DCIS have evolved over the years, which may have impacted treatment or other care for these patients. Between 1989 and 1998, mastectomy or BCS alone were standard of care, according to Dutch treatment guidelines. From 1999 onwards, the addition of radiotherapy after BCS was included, which significantly reduces ipsilateral breast recurrences (both *in situ* and invasive).^{35,36} Also, our group has shown that women diagnosed with DCIS between 1999 and 2004 were less likely to develop iIBC than women diagnosed between 1989 and 1998, regardless of treatment and age.²⁹ In addition, the introduction of digital mammography significantly improved the coverage and sensitivity of screening and led to an increase in the percentage of screening-detected DCIS.³⁷

The fact that all patients included in our study group were treated by BCS alone, also brings a limitation of the studies presented in **chapter 3 and 4**. As analysis for clonal-relatedness of patient-matched DCIS and IBC (**chapter 6**) was not yet performed at the time of selecting patients for our case-control study, we did not take into account whether subsequent IBCs were indeed all clonally related to the primary DCIS lesion. As they might be second primary tumors, the prognostic factors identified could have attenuated prognostic significance or could be risk factors for any second invasive breast event after DCIS. The most ideal patient group for these studies would have been a group of women with untreated DCIS that subsequently developed ipsilateral IBC. However, such study groups are not yet available. Therefore, using a group of women treated for DCIS with BCS alone was the best possible alternative.

Use of FFPE tissue for molecular studies

Another challenge was the use of archival FFPE tissue for our molecular studies. Tissue blocks were collected from several pathology laboratories within the Netherlands, and as a result the embedding and storage of the tissue blocks may have been different between laboratories. This could have affected the results presented in this thesis. However, this would only have confounded the analyses if case-control status would have been differently

distributed between laboratories and this was not the case.

Use of archival FFPE tissue also has a limitation, since formalin-fixation induces cross-links between the nucleic acids that cause fragmentation of DNA and RNA with prolonged storage and sequence artefacts after amplification by PCR.^{38,39} However, the use of FFPE tissue for molecular studies has been proven to be successful by others.^{40–45} On top of this, we were challenged by the fact that DCIS lesions are often small and thus yield small amounts of DNA and RNA. Thus, we selected methods and kits that were compatible with low inputs of low-quality FFPE-derived DNA and RNA extracted.

A strength of the molecular studies described in **chapter 4 and 6** is the use of laser-microdissection. Morphologically, breast tumors are generally highly heterogeneous. Hence, to reassure a more cellular-based rather than a tissue-based resolution for our molecular studies, laser-microdissection is essential.^{46,47} In addition, we succeeded in collecting both gene expression, copy number, and mutation data for a large number of DCIS samples included in our molecular studies. Mutation analysis was restricted to a 53-gene panel and thus represents a part of the full picture. Whole exome sequencing would have given us more information, although with the cost of lower sequencing depth, but is not feasible for this cohort of old-FFPE derived DNA.

Clinical implications and suggestions for further research

With the work described in this thesis, we have identified several prognostic markers for subsequent IBC after DCIS (e.g. HER2 and COX-2), but before implementation into the clinic validation in independent cohorts is needed. We encourage researchers to set up new unbiased patient cohorts for validation purposes and to remain searching for new prognostic markers. Furthermore, the type of recurrence should be specified, *in situ* or invasive, as it are the invasive recurrences that increase a woman's risk of dying from breast cancer and are the most clinically relevant end-point in prognostic factor studies for DCIS. Currently, there are multiple initiatives that are setting up study cohorts including primary DCIS and subsequent IBC (PRECISION, <https://www.cancerresearchuk.org/funding-for-researchers/how-we-deliver-research/grand-challenge-award/funded-teams-wesseling>).^{48–51} In addition, non-inferiority trials, such as LORD, LORIS and COMET, will be important in prospective validation of prognostic factors.^{49–51}

The results from our clonality analysis may substantially impact personalized risk stratification for women with DCIS, as two different risk issues have to be addressed. First, the risk of the DCIS lesion progressing to IBC. Second, the risk to develop a non-clonally related, most likely independent new IBC. Future risk stratification for women with DCIS, therefor should include the risk of the DCIS lesion progressing to IBC as well as the risk to develop a new independent breast cancer.

Lastly, we advise future researchers to take an integrated, comprehensive approach that involves new technologies and areas of research (e.g., detailed genomics and epigenomics analysis, the role of the microenvironment, and creation of a risk stratification tool) to assess the likelihood of DCIS progression.

Conclusions

Based on the results presented in this thesis, we can conclude that we have identified excellent candidate prognostic markers for use in personalized risk stratification. From previous high quality prognostic factor studies we identified six factors that predict invasive breast cancer risk after DCIS and are highly recommended for validation. The results of our case-control study underline the importance of assessment of DCIS stromal compartment and protein expression of HER2 and COX-2 to estimate the risk of subsequent invasive disease after a diagnosis of DCIS and are excellent candidate prognostic markers for use in personalized patient risk stratification. Lastly, although our molecular study is one of the largest in its sort, we only identified subtle differences between DCIS with and without a subsequent IBC.

Furthermore, we have provided evidence of the direct progression of DCIS into IBC. The results from our clonality analysis have changed our view on DCIS as *true* recurrence rates after DCIS may be overestimated and it seems that DCIS might not just be a precursor, but could also be risk lesion of IBC, as has been estimated for LCIS.

We hope our efforts will ultimately contribute to the identification of reliable and clinically relevant prognostic factors for DCIS in the near future. This will subsequently help to distinguish indolent from potential hazardous DCIS, and thereby putting an end to the current overtreatment of an often indolent disease.

References

1. Zhang X, Dai H, Liu B, Song F, Chen K. Predictors for local invasive recurrence of ductal carcinoma in situ of the breast: a meta-analysis. *Eur J Cancer Prev.* 2016;25(1):19-28.
2. Boyages J, Delaney G, Taylor R. Predictors of local recurrence after treatment of ductal carcinoma in situ: a meta-analysis. *Cancer.* 1999;85(3):616-628.
3. Wang S-Y, Shamliyan T, Virnig BA, Kane R. Tumor characteristics as predictors of local recurrence after treatment of ductal carcinoma in situ: a meta-analysis. *Breast Cancer Res Treat.* 2011;127(1):1-14.
4. Nofech-Mozes S, Spayne J, Rakovitch E, et al. Biological Markers Predictive of Invasive Recurrence in DCIS. *Clin Med Oncol.* 2008;2(416):7-18.
5. Van Bockstal M, Lambein K, Gevaert O, et al. Stromal architecture and periductal decorin are potential prognostic markers for ipsilateral locoregional recurrence in ductal carcinoma in situ of the breast. *Histopathology.* 2013;63(4):520-533.
6. Generali D, Buffa FM, Deb S, et al. COX-2 expression is predictive for early relapse and aromatase inhibitor resistance in patients with ductal carcinoma in situ of the breast, and is a target for treatment. *Br J Cancer.* 2014;111(1):46-54.
7. Johnson KC, Koestler DC, Fleischer T, et al. DNA methylation in ductal carcinoma in situ related with future development of invasive breast cancer. *Clin Epigenetics.* 2015;7(1):75.
8. Kim SY, Jung S, Kim MS, et al. Genomic differences between pure ductal carcinoma in situ and synchronous ductal carcinoma in situ with invasive breast cancer. *Oncotarget.* 2015;6(10):7597-7607.
9. Muggerud AA, Hallett M, Johnsen H, et al. Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer. *Mol Oncol.* 2010;4(4):357-368.
10. Knudsen ES, Ertel A, Davicioni E, Kline J, Schwartz GF, Witkiewicz AK. Progression of ductal carcinoma in situ to invasive breast cancer is associated with gene expression programs of EMT and myoepithelia. *Breast Cancer Res Treat.* 2012;133(3):1009-1024.
11. Fleischer T, Frigessi A, Johnson KC, et al. Genome-wide DNA methylation profiles in progression to in situ and invasive carcinoma of the breast with impact on gene transcription and prognosis. *Genome Biol.* 2014;15(8):435.
12. Johnson CE, Gorringer KL, Thompson ER, et al. Identification of copy number alterations associated with the progression of DCIS to invasive ductal carcinoma. *Breast Cancer Res Treat.* 2012;133(3):889-898.
13. Doebar SC, Sieuwerts AM, de Weerd V, Stoop H, Martens JWM, van Deurzen CHM. Gene Expression Differences between Ductal Carcinoma in Situ with and without Progression to Invasive Breast Cancer. *Am J Pathol.* 2017;187(7):1648-1655.
14. Gorringer KL, Hunter SM, Pang J-M, et al. Copy number analysis of ductal carcinoma in situ with and without recurrence. *Mod Pathol.* 2015;28(9):1174-1184.
15. Solin LJ, Gray R, Baehner FL, et al. A multigene expression assay to predict local recurrence risk for ductal carcinoma in situ of the breast. *J Natl Cancer Inst.* 2013;105(10):701-710.
16. Lesurf R, Aure MR, Mørk HH, et al. Molecular Features of Subtype-Specific Progression from Ductal Carcinoma In Situ to Invasive Breast Cancer. *Cell Rep.* 2016;16(4):1166-1179.
17. Bijker N, Peterse JL, Duchateau L, 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. *Br J Cancer.* 2001;84(4):539-544.
18. Karlsson E, Sandelin K, Appelgren J, et al. Clonal alteration of breast cancer receptors between primary ductal carcinoma in situ (DCIS) and corresponding local events. *Eur J Cancer.* 2014;50(3):517-524.
19. Huang E, Buchholz TA, Meric F, et al. Classifying local disease recurrences after breast conservation therapy based on location and histology: new primary tumors have more favorable outcomes than true local disease recurrences. *Cancer.* 2002;95(10):2059-2067.
20. Freedman GM, Anderson PR, Hanlon AL, Eisenberg DF, Nicolaou N. Pattern of local

- recurrence after conservative surgery and whole-breast irradiation. *Int J Radiat Oncol Biol Phys.* 2005;61(5):1328-1336.
21. Nishimura S, Takahashi K, Akiyama F, et al. Classification of ipsilateral breast tumor recurrence after breast-conserving therapy: new primary cancer allows a good prognosis. *Breast Cancer.* 2005;12(2):112-117.
22. Komoike Y, Akiyama F, Iino Y, et al. Analysis of ipsilateral breast tumor recurrences after breast-conserving treatment based on the classification of true recurrences and new primary tumors. *Breast Cancer.* 2005;12(2):104-111.
23. Panet-Raymond V, Truong PT, McDonald RE, et al. True recurrence versus new primary: an analysis of ipsilateral breast tumor recurrences after breast-conserving therapy. *Int J Radiat Oncol Biol Phys.* 2011;81(2):409-417.
24. Banelli B, Casciano I, di Vinci A, et al. Pathological and molecular characteristics distinguishing contralateral metastatic from new primary breast cancer. *Ann Oncol.* 2009;21(6):1237-1242.
25. Begg CB, Ostrovnaya I, Carniello JVS, et al. Clonal relationships between lobular carcinoma in situ and other breast malignancies. *Breast Cancer Res.* 2016;18(1):1-11.
26. Bollet MA, Servant N, Neuvial P, et al. High-Resolution Mapping of DNA Breakpoints to Define True Recurrences Among Ipsilateral Breast Cancers. *JNCI J Natl Cancer Inst.* 2008;100(1):48-58.
27. Hwang ES, Nyante SJ, Chen YY, et al. Clonality of lobular carcinoma in situ and synchronous invasive lobular carcinoma. *Cancer.* 2004;100(12):2562-2572.
28. Biermann J, Parris TZ, Nemes S, et al. Clonal relatedness in tumour pairs of breast cancer patients. *Breast Cancer Res.* 2018;20(1):96.
29. 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 Res Treat.* 2016;159(3):553-563.
30. Li G, Sajobi TT, Menon BK, et al. Registry-based randomized controlled trials- what are the advantages, challenges, and areas for future research? *J Clin Epidemiol.* 2016;80:16-24.
31. Becher H, Winkler V. Estimating the standardized incidence ratio (SIR) with incomplete follow-up data. *BMC Med Res Methodol.* 2017;17(1):1-10.
32. Lauer MS, D'Agostino RB. The randomized registry trial--the next disruptive technology in clinical research? *N Engl J Med.* 2013;369(17):1579-1581.
33. Mengel M, von Wasielewski R, Wiese B, Rüdiger T, Müller-Hermelink HK, Kreipe H. Inter-laboratory and inter-observer reproducibility of immunohistochemical assessment of the Ki-67 labelling index in a large multi-centre trial. *J Pathol.* 2002;198(3):292-299.
34. van Doijeweert C, van Diest PJ, Willems SM, Kuijpers CCHJ, Overbeek LIH, Deckers IAG. Significant inter- and intra-laboratory variation in grading of ductal carcinoma in situ of the breast: a nationwide study of 4901 patients in the Netherlands. *Breast Cancer Res Treat.* 2018;0(0):0.
35. Wapnir IL, Dignam JJ, Fisher B, 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. *J Natl Cancer Inst.* 2011;103(6):478-488.
36. Donker M, Liti?re S, Werutsky G, et al. Breast-conserving treatment with or without radiotherapy in ductal carcinoma In Situ: 15-year recurrence rates and outcome after a recurrence, from the EORTC 10853 randomized phase III trial. *J Clin Oncol.* 2013;31(32):4054-4059.
37. Karssemeijer N, Bluekens AM, Beijerinck D, et al. Breast Cancer Screening Results 5 Years after Introduction of Digital Mammography in a Population-based Screening Program. *Radiology.* 2009;253(2):353-358.
38. Williams C, Pontén F, Moberg C, et al. A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am J Pathol.* 1999;155(5):1467-1471.
39. Quach N, Goodman MF, Shibata D. In vitro mutation artifacts after formalin fixation and error prone translesion synthesis during PCR. *BMC Clin Pathol.* 2004;4(1):1.

40. Cieslik M, Chugh R, Wu Y-M, et al. The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res.* 2015;25(9):1372-1381.
41. Eikrem O, Beisland C, Hjelle K, et al. Transcriptome sequencing (RNAseq) enables utilization of formalin-fixed, paraffin-embedded biopsies with clear cell renal cell carcinoma for exploration of disease biology and biomarker development. *PLoS One.* 2016;11(2):1-19.
42. Hoogstraat M, Hinrichs JWJ, Besselink NJM, et al. Simultaneous detection of clinically relevant mutations and amplifications for routine cancer pathology. *J Mol Diagn.* 2015;17(1):10-18.
43. Ellison G, Huang S, Carr H, et al. A reliable method for the detection of BRCA1 and BRCA2 mutations in fixed tumour tissue utilising multiplex PCR-based targeted next generation sequencing. *BMC Clin Pathol.* 2015;15:5.
44. Chung J, Son D-S, Jeon H-J, et al. The minimal amount of starting DNA for Agilent's hybrid capture-based targeted massively parallel sequencing. *Sci Rep.* 2016;6(May):26732.
45. Kader T, Goode DL, Wong SQ, et al. Copy number analysis by low coverage whole genome sequencing using ultra low-input DNA from formalin-fixed paraffin embedded tumor tissue. *Genome Med.* 2016;8(1):121.
46. Hernández S, Lloreta J. Manual versus laser micro-dissection in molecular biology. *Ultrastruct Pathol.* 2006;30(3):221-228.
47. Espina V, Wulfkühle JD, Calvert VS, et al. Laser-capture microdissection. *Nat Protoc.* 2006;1(2):586-603.
48. Thompson AM, Clements K, Cheung S, et al. Management and 5-year outcomes in 9938 women with screen-detected ductal carcinoma in situ: the UK Sloane Project. *Eur J Cancer.* 2018;101:210-219.
49. Elshof LE, Tryfonidis K, Slaets L, 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. *Eur J Cancer.* 2015;51(12):1497-1510.
50. Francis A, Thomas J, Fallowfield L, et al. Addressing overtreatment of screen detected DCIS; The LORIS trial. *Eur J Cancer.* 2014;51(16):2296-2303.
51. Youngwirth LM, Boughey JC, Hwang ES. Surgery versus monitoring and endocrine therapy for low-risk DCIS: The COMET Trial. *Bull Am Coll Surg.* 2017;102(1):62-63.

APPENDICES

Nederlandse samenvatting

List of publications

About the author

Nederlandse samenvatting

Ductaal carcinoom in situ is een voorstadium van borstkanker, of toch niet?

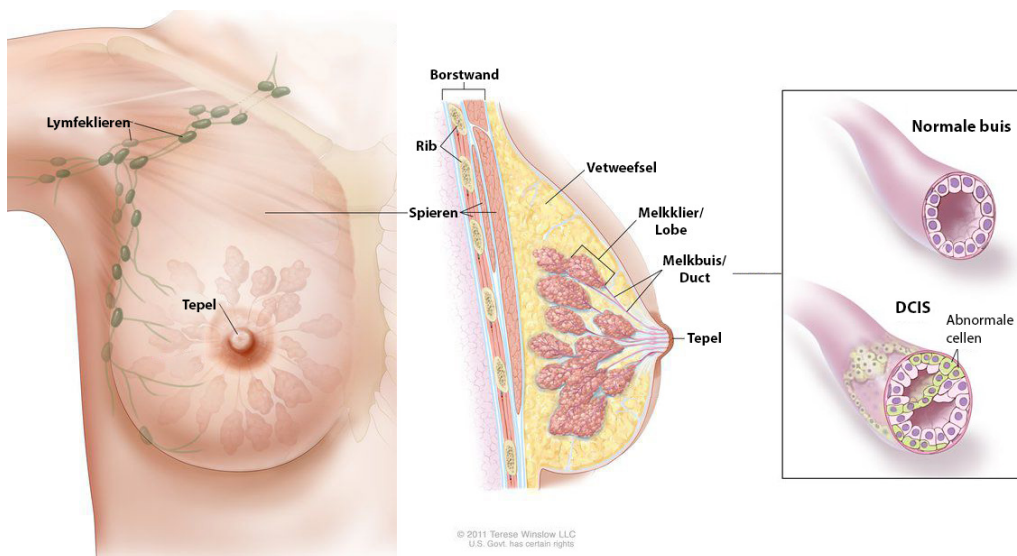
Ten eerste geen kwaad doen

Primum non nocere, "ten eerste geen kwaad doen", is een advies uit de eed van Hippocrates die artsen steeds in gedachten moeten houden. Vooral als ze de keus hebben tussen enerzijds handelen met onzeker resultaat of afwachten. Om de juiste behandeling te bepalen voor een patiënt, moeten artsen een goede, gebalanceerde inschatting kunnen maken van de risico's en voordelen. Hiervoor is wetenschappelijk onderzoek erg belangrijk. Een ziekte waarbij betrouwbare risico-inschatting hard nodig is, is ductaal carcinoom in situ, ook wel bekend als "DCIS".

Carcinoom in situ

De vrouwelijke borst heeft een complexe biologie en anatomie. Onder invloed van hormonen gaat de borst meer veranderingen door dan andere delen van het lichaam. Deze veranderingen vinden plaats tijdens geboorte, pubertijd, zwangerschap en borstvoeding, en gaan door tot de menopauze. Iedere borst bevat 15-20 buizenstelsels bestaande uit melkklieren (lobes) en melkbuizen (ducten; Figuur 1). De melkklieren zijn verantwoordelijk voor de melkproductie wanneer een vrouw borstvoeding geeft. De melkbuizen transporteren de melk van de melkklieren naar de tepel. De rest van de borst bestaat uit vetweefsel.

Net als alle andere delen van het lichaam, bestaat de vrouwelijke borst uit miljarden microscopisch kleine bouwsteentjes, genaamd cellen. Normaal gesproken vermenigvuldigen deze cellen zich via een strikt georganiseerd programma, waarbij nieuwe gezonde cellen doorgaan met delen en de plaats innemen van oude cellen die afsterven. Maar soms ontstaan er tijdens de ontwikkeling van een cel bepaalde foutjes (mutaties). Deze mutaties ontstaan wanneer genen, die normaal gesproken controleren of een cel zich op de juiste manier vermenigvuldigt, falen in het oppikken van een fout. In de meeste gevallen schiet het afweersysteem hier te hulp en ruimt de afwijkende cel op, maar soms is het afweersysteem niet in staat de afwijkende cel als zodanig te herkennen. De afwijkende cel zal dan doorgaan met vermenigvuldigen en volgt daarbij niet het strikt georganiseerde programma. Hierdoor kan deze afwijkende cel ongeremd en sneller delen dan normaal en op die manier snel in aantal toenemen. In dit stadium hebben we nog niet te maken met "kanker". De afwijkende cellen zijn namelijk alleen aanwezig in de melkbuis (ductaal carcinoom *in situ* [DCIS]; *in situ* = in de melkbuis) of in de melkklier (lobulair carcinoom *in situ*; LCIS) en kunnen daardoor geen schade aanrichten aan het omliggende borstweefsel (Figuur 1). Deze "*in situ* carcinomen" veroorzaken vaak geen klachten, zijn vaak niet voelbaar (geen knobbeltje in



Figuur 1. Anatomie van de vrouwelijk borst. Links: vooraanzicht; midden: zijaanzicht; rechts: doornede van een normale melkbuis en een melkbuis met afwijkende cellen, welke DCIS wordt genoemd.

de borst) en zijn niet levensbedreigend. De meest voorkomende vorm van carcinoom in situ is DCIS.

DCIS is lang een aandoening geweest die maar sporadisch werd vastgesteld (gediagnostiseerd). De introductie van het bevolkingsonderzoek borstkanker heeft ervoor gezorgd dat DCIS vaker wordt gevonden. Tijdens het bevolkingsonderzoek wordt een borstfoto (mammogram) gemaakt. Aanwezigheid van kalkspatjes op het mammogram zijn een teken dat er DCIS aanwezig zou kunnen zijn. Dit moet dan nog worden bevestigd door een bezoek aan het ziekenhuis. Hier wordt met een holle naald een klein beetje van het afwijkend weefsel afgenomen (biopt) voor onderzoek onder de microscoop. DCIS bedraagt tegenwoordig 20-25% van alle borstafwijkingen die met het bevolkingsonderzoek worden gevonden. In Nederlands krijgen jaarlijks 2.500 vrouwen de diagnose DCIS.

Dilemma van overdiagnose en overbehandeling

We spreken pas van borstkanker wanneer de afwijkende cellen de eigenschap ontwikkelen om het gezonde, omliggende weefsel binnen te dringen (invasie). Borstkanker moet behandeld worden omdat het levensbedreigende problemen kan veroorzaken voor de patiënt.

DCIS wordt beschouwd als een mogelijk voorstadium van borstkanker. Onderzoek heeft

uitgewezen dat een deel van de DCIS-en zich kan ontwikkelen tot borstkanker, maar dat het overgrote deel van de DCIS-en, zonder behandeling, nooit borstkanker zal worden en dus ook nooit levensbedreigende problemen zal veroorzaken voor de patiënt. Dit wijst erop dat veel vrouwen met DCIS een “overdiagnose” ontvangen, omdat ze worden gediagnostiseerd met een ziekte die niet zou hebben geleid tot klachten of een risico op overlijden.

Op dit moment is het nog niet mogelijk om betrouwbaar onderscheid te maken tussen DCIS wat zich zal ontwikkelen tot borstkanker en DCIS die dit waarschijnlijk nooit zal worden. Als gevolg adviseren de huidige richtlijnen voor de behandeling van DCIS om alle vrouwen met DCIS te behandelen, om zo de ontwikkeling van borstkanker te voorkomen. Dit gebeurt in Nederland door middel van een operatie waarbij een deel van het borstweefsel met de DCIS wordt verwijderd (borst-sparende chirurgie), waarna de borst vaak ook nog worden behandeld met bestraling (radiotherapie). Soms wordt met een operatie zelfs de hele borst verwijderd (mastectomie). Vrouwen met DCIS met een laag risico op het ontwikkelen van borstkanker wordt met deze intensive behandeling eerder kwaad aangedaan dan goed, omdat zij geen baat hebben bij deze behandeling. Factoren die ons kunnen helpen voorspellen welke DCIS uiteindelijke borstkanker zal worden, zijn hard nodig om deze “overbehandeling” te voorkomen. Gebruik van deze voorspellende factoren zal er uiteindelijk toe leiden dat vrouwen met een “laag risico” DCIS deze intensieve behandeling bespaard kan blijven.

Bevindingen beschreven in dit proefschrift

Het onderzoek beschreven in dit proefschrift had als doel het vinden van voorspellende factoren voor de ontwikkeling van borstkanker na DCIS. Daarnaast wilden we bewijs vinden dat borstkanker inderdaad kan voortkomen uit DCIS.

In **hoofdstuk 1** wordt het probleem beschreven wat we met dit onderzoek willen oplossen. Dit probleem is ook hierboven beschreven. **Hoofdstuk 2** geeft een samenvatting van de bevindingen van voorgaande onderzoeken op het gebied van voorspellende factoren voor de ontwikkeling van borstkanker na DCIS. Aan de hand van deze samenvatting hebben we de factoren met de beste voorspellende waarde kunnen identificeren. Ook hebben we zwakheden van voorgaande studies uitgelicht, zodat toekomstige onderzoekers hiervan op de hoogte zijn en deze zwakke punten in het vervolg zo veel mogelijk kunnen voorkomen. **Hoofdstuk 3 en 4** beschrijven ons eigen onderzoek naar voorspellende factoren voor de ontwikkeling van borstkanker na DCIS. We hebben gevonden dat een hoge hoeveelheid van de eiwitten HER2 en COX-2 en aanwezigheid van periductale fibrose (een bindweefsel patroon) in verband staan met de ontwikkeling van borstkanker na DCIS. Vrouwen met DCIS met een hoge hoeveelheid van deze twee eiwitten hebben een duidelijk hogere kans om borstkanker te ontwikkelen dan vrouwen met DCIS met een kleine hoeveelheid van deze eiwitten. Daarnaast laat dit onderzoek zien dat het risico op borstkanker van vrouwen met DCIS met lage hoeveelheid COX-2 eiwit vergelijkbaar is met het risico op borstkanker van vrouwen zonder DCIS. Ook speelt de reactie van het afweersysteem op de DCIS een rol in

het wel of niet ontwikkelen van borstkanker.

Welke eigenschappen heeft DCIS nodig om het omliggende weefsel te kunnen binnendringen en hoe kunnen we dit het beste onderzoeken? Hierover gaat **hoofdstuk 5**. Voorgaande onderzoeken hebben mogelijk de verkeerde aanpak gebruikt om belangrijke factoren voor invasie te vinden en dat proberen we in dit hoofdstuk aan het licht te brengen.

Hoofdstuk 6 levert het bewijs dat borstkanker inderdaad kan voortkomen uit DCIS, maar laat ook zien dat dit in 29% van de gevallen niet zo is. In dit laatste geval is de borstkanker een nieuwe ziekte, onafhankelijk van de eerdere DCIS. De bevinding dat borstkanker vaak een nieuwe ziekte is, is erg belangrijk omdat hieruit blijkt dat: (1) het aandeel in borstkanker wat voortkomt uit een eerdere DCIS wordt overschat; en (2) dat DCIS niet alleen een voorstadium van borstkanker kan zijn maar ook een waarschuwing (risico-indicator) is voor een hogere kans op borstkanker in het algemeen.

Als laatste, in **hoofdstuk 7**, worden de resultaten uit dit proefschrift samengevat en bediscussieerd.

Impact van het onderzoek

Dit onderzoek heeft geleid tot de identificatie van verschillende voorspellende factoren (zoals de eiwitten HER2 en COX-2) die grote potentie hebben om onderscheid te maken tussen DCIS die uiteindelijk borstkanker zal worden en DCIS die niet deze eigenschap heeft. Deze factoren kunnen bijdragen aan betere risico-inschatting voor vrouwen met DCIS en kunnen helpen om overdiagnose en overbehandeling te voorkomen. Daarnaast hebben we zwakke punten van voorgaande studies uitgelicht, zodat toekomstige onderzoekers hiervan op de hoogte zijn en er hopelijk alles aan doen om deze in het vervolg te voorkomen.

De bevinding dat borstkanker na DCIS ook een nieuwe ziekte kan zijn moet in ogenschouw worden genomen wanneer het risico op borstkanker voor een vrouw met DCIS wordt berekend. In deze risicoberekening moet namelijk niet alleen het risico opgenomen worden dat de DCIS uitgroeit tot borstkanker, maar ook de kans op borstkanker als een nieuwe, onafhankelijke ziekte.

We hopen dat de bevindingen beschreven in dit proefschrift zullen bijdragen aan de toepassing van betrouwbare voorspellende factoren voor DCIS in de nabije toekomst. Daarmee zal er dan een einde komen aan de overdiagnose en overbehandeling van vrouwen met DCIS.

List of publications

L.L. Visser, M. Hoogstraat, T. Bismeyer, F. Nieboer, M. de Maaker, P. Kristel, L. Mulder, M. Kreté, W. Brugman, M. Schaapveld, M.K. Schmidt, L.F.A. Wessels, J. Wesseling*, E.H. Lips*, on behalf of the PRECISION team. Deciphering the genomic and epigenomic landscapes of ductal carcinoma in situ associated with and without subsequent invasive breast cancer. *Manuscript in preparation*.

*: Shared last author

L.L. Visser, M. Hoogstraat, T. Bismeyer, F. Nieboer, M. de Maaker, P. Kristel, L. Mulder, M. Kreté, W. Brugman, M. Schaapveld, M.K. Schmidt, E. Sawyer, A. Thompson, A. Futreal, H. Davies, S. Nik-Zainal, L.F.A. Wessels, E.H. Lips*, J. Wesseling*, on behalf of the PRECISION team. Approximately 30% of invasive ipsilateral breast cancers after treatment of ductal carcinoma in situ may be new independent primary cancers. *Manuscript in preparation*.

L.L. Visser, L.E. Elshof, K. van de Vijver, E.J. Groen, M.M. Almekinders, J. Sanders, C. Bierman, D. Peters, I. Hofland, A. Broeks, F.E. van Leeuwen, E.J. Rutgers, M.K. Schmidt, M. Schaapveld, E.H. Lips, J. Wesseling. Discordant marker expression between invasive breast carcinoma and corresponding synchronous and preceding DCIS. *Am J Surg Pathol*. 2019 Nov;43(11):1574-1582. doi: 10.1097/PAS.0000000000001306.

L.L. Visser, E.J. Groen, F. E. van Leeuwen, E.H. Lips, M.K. Schmidt*, and J. Wesseling*. Predictors of an invasive breast cancer recurrence after DCIS: a systematic review and meta-analyses. *Cancer Epidemiol Biomarkers Prev*. 2019 May;28(5):835-845. doi: 10.1158/1055-9965.

L.L. Visser, L.E. Elshof, M. Schaapveld, K. Van de Vijver, E.J. Groen, M.M. Almekinders, C. Bierman, F.E. van Leeuwen, E.J. Rutgers, M.K. Schmidt, E.H. Lips, J. Wesseling. Clinicopathological risk factors for an invasive breast cancer recurrence after ductal carcinoma in situ. *Clin Cancer Res*. 2018 Aug 1;24(15):3593-3601. doi: 10.1158/1078-0432.

Not in this thesis:

E.J. Groen, L.E. Elshof, L.L. Visser, E.J. Rutgers, G.H.A. Winter-Warnars, E.H. Lips, J. Wesseling. Finding the balance between over- and under-treatment of ductal carcinoma in situ (DCIS). *Breast*. 2017 Feb;31:274-283. doi: 10.1016/j.breast.2016.09.001.

Appendices

J.M. Ruben, L.L. Visser, K.M. Heinhuis, T. O'Toole, H.J. Bontkes, T.M. Westers, G.J. Ossenkoppele, T.D. de Gruijl, A.A. van de Loosdrecht. A human cell line model for interferon- α driven dendritic cell differentiation. *PLoS One*. 2015 Aug 7;10(8):e0135219. doi: 10.1371/journal.pone.0135219.

S. Ariotti, M.A. Hogenbirk, F.E. Dijkgraaf, L.L. Visser, M.E. Hoekstra, J. Song, H. Jacobs, J.B. Haanen, T.N. Schumacher. Induction of tissue-wide pathogen alert upon skin-resident memory CD8⁺ T cell triggering. *Science*. 2014 Oct 3;346(6205):101-5. doi: 10.1126/science.1254803.

J.M. Ruben, L.L. Visser, H.J. Bontkes, T.M. Westers, G.J. Ossenkoppele, T.D. de Gruijl, A.A. van de Loosdrecht. Targeting the leukemic stem cell compartment by enhancing tumor cell-based vaccines. *Immunotherapy*. 2013 Aug;5(8):859-68. doi: 10.2217/imt.13.76.

About the author

Lindy Larissa Visser was born on October 23th, 1989 in Alkmaar. She graduated from secondary education (HAVO) in 2007 at Christelijke Scholengemeenschap Jan Arentsz. Subsequently, she enrolled the Biology and Medical Laboratory Research program at the University of Applied Sciences in Leiden, with a major in cytology and histopathology. Triggered by her interest in oncology, she performed her research internship in the field of oncology. After obtaining her bachelor degree in 2011, she continued with a Master's program in Oncology, during which she participated in several courses on medical oncology, tumor immunology, and research ethics. She performed her first internship in the lab of prof. dr. Arjan A. van de Loosdrecht at the VU University Medical Center – Cancer Center Amsterdam (CCA). Her second internship was carried out at the Netherlands Cancer Institute in the lab of prof. dr. Ton N.M. Schumacher. Both internships were focused on tumor immunology. She obtained her master's degree in 2013. On May 15th, 2014 she joined the group of prof. dr. Jelle Wesseling as a graduate student, where she conducted the research described in the current thesis. In 2018, Lindy received the AACR-Pezcollor Scholar-In-Training Award 2018 for her work on molecular changes associated with progression of DCIS to subsequent invasive breast cancer. That same year, she also was awarded with the PALGA-prijs, a prize for the best research project making use of data from the Dutch cytopathology and histopathology registry (PALGA). July 1st, 2019 Lindy continued her work in the direction of molecular and computational biology within the single cell genomics team of prof. dr. Frank C.P. Holstege at the Prinsess Máxima Center for pediatric oncology.

