Potential risk factors for local breast cancer recurrence after mastectomy in the Netherlands:
a retrospective nationwide cohort study

T.D.A. Stindt

Master thesis for Health Sciences

Faculty of Science & Technology

University of Twente

Supervised by

Prof. Dr. S. Siesling

Dr. M.C. Van Maaren

22th of July 2024

Abstract

Objective: Despite the supposed complete removal of glandular breast tissue in mastectomy, breast cancer patients still experience local recurrence after the surgery. Moreover, limited research currently exists on isolated local recurrence as opposed to locoregional recurrence of breast cancer after mastectomy. This study aims to identify potential risk factors associated with the development of local breast cancer recurrence after mastectomy.

Methods: Based on data from the Netherlands Cancer Registry, a retrospective nationwide cohort study was conducted on women diagnosed with primary invasive nonmetastatic breast cancer between 2012 and 2016 and treated with mastectomy. Potentially relevant variables were selected based on univariable Cox regression before inclusion in multivariable cause-specific Cox regression to estimate the hazard ratios for local recurrence. Separate models were constructed for patients with and without neoadjuvant systemic therapy. Collinearity was assessed using variance inflation factors and interaction variables for potentially collinear variables. The goodness of fit was visually assessed using the Cox-Snell residuals. Multiple imputation was used to account for missing data.

Results: In total, 22,304 patients were selected, of which 17,250 patients were included in the non-NST group and 5,054 patients in the NST group. In the non-NST group, statistically significant associations with local recurrence were found for menopausal status, socioeconomic status, screening detection, sublocalisation, differentiation grade, pathologic tumour and nodal stages, immediate breast reconstruction, hormone receptor status and treatment, HER2 status and treatment, radiotherapy type, and chemotherapy. For the NST group, statistically significant associations with local recurrence were found for clinical tumour stage, pathologic complete response, presence of a DCIS component, HER2 status, hormone receptor status and treatment, and radiotherapy type. No collinearity between variables was detected and all models showed adequate goodness of fit.

Conclusion: The study results indicate potential risk factors for local recurrence after mastectomy. Further investigation of these variables is essential for improving the outcomes of breast cancer patients by optimizing post-surgical care based on individual patient, tumour and treatment characteristics.

Keywords: Breast cancer, local recurrence, mastectomy, risk factors

Acknowledgements

I would like to express my gratitude to my direct supervisors, Dr. M.C. van Maaren and Prof. Dr. S. Siesling, for allowing me to contribute to this research and for their continuous guidance and feedback. I also want to thank Prof. Dr. L.J. Boersma and Dr. O. Kaidar-Person for conceptualising this research topic. My thanks do not only go out to them, but also to the other clinicians for their contributions to my understanding of the clinical practices in breast cancer care and to the interpretation of the study results. Lastly, I want to thank all the coworkers at the Netherlands Comprehensive Cancer Organisation for their warm welcome and for allowing me to use their workspaces.

Table of contents

Introduction

In recent decades, the incidence of female breast cancer in The Netherlands has steadily risen. In 2023, over 15,000 women were diagnosed with invasive breast cancer. Contrary to the incidence, breast cancer mortality has declined [1]. This can be attributed to early diagnosis, for example through screening, and improved treatment [2]. The primary surgical treatment of nonmetastatic invasive breast cancer consists of either breast-conserving therapy (BCT) or mastectomy [2,3]. In mastectomy, the glandular breast tissue is removed, though various types that preserve anatomical structures intact exist, such as radical, modified radical, simple, skin-sparing, and nipple-areolar-sparing mastectomy [3,4]. In 2022, about 26% of all breast cancer patients were treated by mastectomy, though this percentage is decreasing in favour of BCT [5]. After treatment, breast cancer survivors have a risk of recurrence, which can occur locally, regionally or distantly [6]. Together with the rising incidence of invasive breast cancer, the population at risk of recurrence is also rising.

Considering that mastectomy involves the complete removal of the glandular breast tissue, it is puzzling how local recurrence occurs in these patients. This raises questions about underlying factors that contribute to the occurrence of local recurrence. However, in breast cancer research, there is a lack of studies that report on local recurrence specifically and its risk factors are poorly understood. Researchers often group local recurrence with regional recurrence without distinguishing between the two [6–8]. This limits further advancement in the personalised treatment of breast cancer patients. For example, post-mastectomy radiotherapy (PMRT) has been shown to reduce the risk of both locoregional and overall recurrence and mortality [9,10]. In 2022, 10% of breast cancer patients received PMRT after mastectomy [5]. Despite the proven benefits of PMRT, the current guidelines need improvement in terms of when PMRT is indicated based on individual risk of recurrence and how patients should be irradiated to improve oncological outcomes [4,11–13]. To better understand when PMRT is indicated, a crucial first step is to know which factors influence a patient's predisposition to specifically local recurrence after mastectomy. This is one example of the relevance of a better understanding of the risk factors associated with local recurrence after mastectomy. To address the current gap in knowledge, this study aims to develop an explanatory model that evaluates the presence and magnitude of potential risk factors for local recurrence in breast cancer patients treated with mastectomy based on observational data in The Netherlands. This patient group is selected because of their eligibility for post-mastectomy radiotherapy. Thus, this study aims to answer the main research question: *'What variables are associated with local recurrence in patients diagnosed between 2012 and 2016 with primary invasive nonmetastatic breast cancer who were treated with mastectomy?'*

Methods

Study population

The data used in this retrospective cohort study were obtained from the Netherlands Cancer Registry (NCR). This registry consists of population-based data of all cancer patients in the Netherlands diagnosed as of 1989. The NCR is maintained by the Netherlands Comprehensive Cancer Organisation (IKNL). The study population consisted of women diagnosed with primary invasive nonmetastatic breast cancer between 2012 and 2016 who underwent mastectomy. Patients treated in a hospital outside of the Netherlands or with synchronous breast cancer were excluded. Additionally, male patients, patients whose diagnosis was an incidental finding, patients with irradical surgery, patients with unknown surgery dates, and patients with incomplete follow-up were excluded.

The study population was divided into two subgroups based on whether patients had received neoadjuvant systemic therapy (NST) to account for the differences in prognostic interpretation between these groups for variables such as clinical and pathologic tumour stage and nodal involvement. The subgroup of patients who did not receive NST, *i.e.* non-NST, consisted of patients who received adjuvant systemic therapy, *i.e.* post-surgical chemotherapy, endocrine therapy or targeted therapy, or patients with no systemic therapy at all. The subgroup of patients who did receive NST consisted of patients who only received pre-surgical chemotherapy or endocrine therapy or patients who received these treatments both pre- and post-surgically. Because we only consider patients with invasive breast cancer, patients from the non-NST group with pT0 or with unknown pT or pN were excluded. In the NST group, patients with cT0, cTIS, or unknown cT or cN were similarly excluded. Lastly, NST patients who only received neoadjuvant radiotherapy or targeted therapy without chemotherapy were excluded. The exclusion and division of these patients to achieve the described population and subgroups were conducted outside of this study in preparation for a separate study by Van Maaren *et al.* [14].To achieve a population of only patients treated with mastectomy for this study, patients treated with breast-conserving therapy were subsequently excluded.

Data collection

To avoid extensive manual recurrence data collection from individual patient files, this study's data collection was performed through linkage with the Dutch Nationwide Pathology Databank (PALGA). Patient data was collected using an algorithm that selected patients suspected of having had an LRR. Once these patients were identified, NCR data managers collected the registered data in these patient files.

Outcome and definitions

The primary outcome was a local recurrence, defined as any invasive component or DCIS in ipsilateral breast tissue following mastectomy [6]. This was regardless of whether these events were marked in the dataset as a recurrence or a second primary tumour. The occurrence of the outcome was monitored during the follow-up period of 10 years, defined as the time between the date of the primary surgery and the occurrence of an event or the last observation. An event was defined as either a local recurrence or any of the competing events. These risks consisted of regional recurrence, distant metastases, contralateral breast event or death. In case of multiple events, the first event was considered. The only exception was if a local recurrence occurred within 30 days after a competing event. In such cases, the local recurrence was assumed to pre-exist at the time of diagnosis of the competing event. Socioeconomic status was determined for each patient based on the postal code at diagnosis. A positive hormone receptor (HR) status was defined as either positive estrogen receptor (ER) or progesterone receptor (PR) status. Moreover, HR status and endocrine therapy were combined into a single variable to account for the potential collinearity between these variables. For the same reason, human epidermal growth factor receptor 2 (HER2) status and targeted treatment were also combined into a single variable in the non-NST group. This was not viable in the NST group because the number of events was too low. For the NST group, a pathologic complete response (pCR) was defined as ypT0N0 or ypTISN0 to indicate the absence of residual malignancy.

Statistical analysis

Various patient-, tumour, and treatment-related characteristics were considered as the exposure variables. These potential exposure variables were summarised and analysed using univariable Cox proportional hazard regression. While univariable Cox regression is suitable for identifying the association between a single variable and the outcome, it neglects to account for possible interactions between multiple variables that are associated with the outcome [15]. However, its simplicity of construction and interpretation makes it a suitable method for identifying variables that are likely to be associated with the outcome in multivariable analysis and assessing each variable's adherence to the proportional hazard assumption. This assumption states that patients have a constant baseline hazard for the outcome that is modulated by a time-independent, variable-specific hazard ratio [16,17]. The adherence of each variable to the proportional hazard assumption was assessed by visually inspecting the time independence of the Schoenfeld residuals, which represent the discrepancy between the observed values of each variable in the model and the average values of those variables across the entire cohort for a patient when they experience a local recurrence, and parallelism of the log-log survival curves [18]. Viable variables that were differently distributed between outcomes status and showed relatively significant association with the outcome in

univariable Cox regression, defined by a p-value of less than 0.2, were included in a multivariable Cox regression model. Selecting variables likely to be associated with the outcome using univariable regression is a method to limit the risk of overfitting, *i.e.* inflating the association between the variables and the outcome by including too many variables [19]. The multivariable model was used to determine statistically significant exposure variables with estimates of their hazard ratios (HRs) and their 95% confidence intervals. Non-significant variables that did not significantly (p-value < 0.05) contribute to the model fit according to the likelihood ratio test were manually removed using backward selection to obtain a reduced model that contains the most impactful variables [20]. In both the univariable and the multivariable Cox regression, competing events were corrected for by censoring the patient when one of the competing events occurred. This approach follows the principle of cause-specific Cox regression [21]. This method allows researchers to estimate the cause-specific hazards to evaluate the association between the included variables and only the event of interest. By censoring the patients with a competing event, they are no longer considered at risk of the event of interest [22,23]. Not all items in the dataset were complete. Missing data were considered missing at random based on the pattern of missingness. The missing data was imputed using futuremice, which is a wrapper function to run for the Multivariate Imputation by Chained Equations (MICE) function in R in parallel using multiple processor cores. MICE fills missing data using various statistical methods, such as a regression model between the missing variables and the observer variables in the dataset. It does so iteratively for an imposed number of imputations and iterations. MICE relies on the assumption that the missing values are missing at random. This means that the missingness is not completely random but associated with the observed data and not the unobserved data [24]. Von Hippel's method was used to determine the minimum number of imputations required to produce replicable standard errors in the multivariable Cox regression model [25]. For the non-NST group, the data was imputed 33 times, and for the NST group, the data was imputed 55 times. For the non-NST group, each imputation consisted of 10 iterations to achieve convergence. For the NST group, the number of iterations was 40. Four processor cores were used to run MICE in parallel to speed up the computation. Only the missing data was imputed and the other exposure variables were used as imputer variables. The imputation methods were logistic regression for binomial variables, polytomous logistic regression for nominal variables, and proportional odds logistic regression for ordinal variables. To avoid contradictory imputation results, the composite variables, *i.e.* hormonal status with endocrine therapy, HER2 status and targeted therapy, and pCR, were imputed using the 'impute, then transform' method, which involves imputing the separate component variables and combining them afterwards [26]. To avoid missing values after imputation, unclear HER2 status was recoded as missing to ensure that this category was imputed and there were no missing values in HER2 status and targeted therapy. The validity of the imputed data was evaluated using the convergence plots of the mean and standard

deviation and by comparing the distributions of the observed data and imputed data. The imputed data was used and the estimates of the imputed datasets were pooled and compared with estimates obtained from the complete cases. When considering multiple variables in multivariable regression, variables that describe the same or similar phenomena exhibit collinearity. In other words, these variables are determined by each other in the analysis. Collinearity between variables in the multivariable model was assessed using three steps. Firstly, a variance inflation factor (VIF) threshold of 4 was maintained [27]. Secondly, potential variable interactions were evaluated for age groups with menopausal status and immediate breast reconstruction with radiotherapy type by observing the effects of including interaction variables in the models. Lastly, for these same potentially interacting variables, the effect of removing one of both variables and retaining the other on the estimated HRs was observed. Finally, the model's goodness of fit was evaluated by visually inspecting the Cox-Snell residuals. This method involves graphically inspecting whether the estimated cumulative hazard by a model matches the expected cumulative hazard, represented by a line with a slope of 1. The agreement between the estimated and expected cumulative hazard gives an indication of how well the model fits the data. [28]. All statistical analyses were performed in R, version 4.3.1.

Results

Patient inclusion

A total of 23,474 patients were considered for inclusion. Of this total, 418 male patients were excluded. Furthermore, 29 incidental findings, 208 patients with irradical surgeries, 5 patients with unknown surgery dates, and 5 patients with incomplete follow-up were excluded. The remaining 22,809 patients were divided into the non-NST group and the NST group. 17,646 patients were assigned to the non-NST group, of which 187 patients with pT0 or unknown pT and 209 patients with unknown pN were excluded. 5,163 patients were assigned to the NST group, of which 53 patients with cT0, cTIS or unknown cT, 42 patients with unknown cN and 9 patients with either only neoadjuvant radiotherapy or only targeted therapy without chemotherapy were excluded. After exclusion, 17,250 patients remained in the non-NST group and 5,054 patients remained in the NST group (Figure 1).

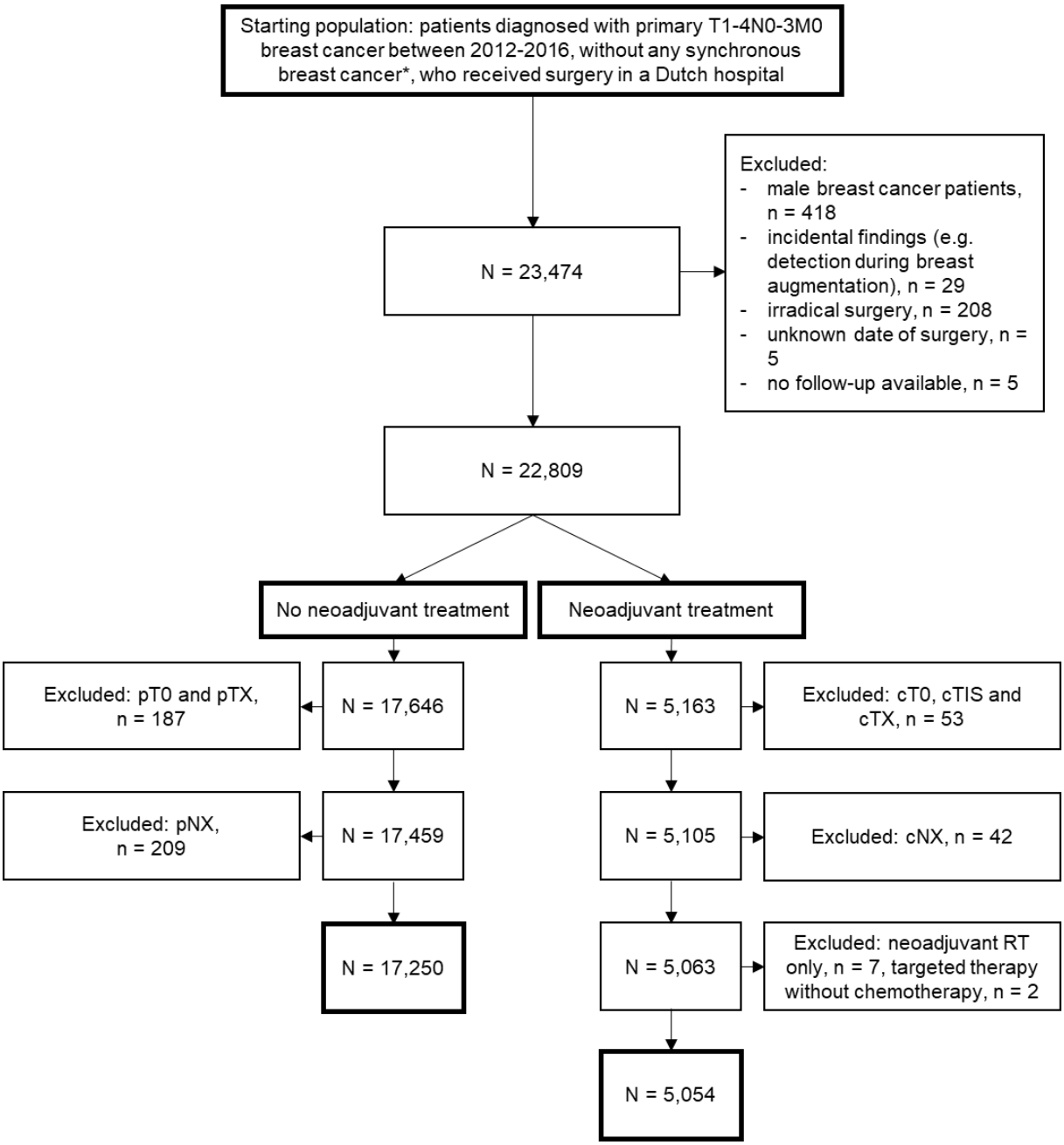

Figure 1: Flowchart of patient inclusion results

**a second primary breast tumour diagnosed within 90 days of the first*

Both subgroups show different baseline characteristics in terms of demographics, tumour characteristics, and treatment characteristics (Table 1). Patients in the non-NST group were generally older than patients in the NST group. Furthermore, the non-NST group had a larger percentage of postmenopausal women. Differentiation grade was significantly more missing in the NST group. Patients in the non-NST group had relatively favourable tumour characteristics, such as lower tumour and nodal stages and positive HR status. They also receive less extensive treatment, evidenced by the lower percentages of patients with radiotherapy, targeted therapy when HER2 positive or

chemotherapy and the larger percentage of HR-positive patients without endocrine therapy. Most patients in the NST subgroup who were HER2 positive also received targeted therapy. In the non-NST group, 354 patients (2.1%) experienced a local recurrence as the first event after a median follow-up of 7.8 years after surgery. In the NST group, this was the case for 82 patients (1.6%) after a median follow-up of 6.7 years after surgery.

Table 1: Baseline characteristics of the study population (n = 22,304).

	No neoadjuvant systemic treatment (N=17250)	**Neoadjuvant systemic treatment (N=5054)**
Year of diagnosis		
2012	3969 (23.0%)	790 (15.6%)
2013	3754 (21.8%)	894 (17.7%)
2014	3486 (20.2%)	1026 (20.3%)
2015	3230 (18.7%)	1186 (23.5%)
2016	2811 (16.3%)	1158 (22.9%)
Age group		
<40	938 (5.4%)	872 (17.3%)
40-49	2879 (16.7%)	1647 (32.6%)
50-59	3852 (22.3%)	1217 (24.1%)
60-69	3992 (23.1%)	862 (17.1%)
70-79	3298 (19.1%)	285 (5.6%)
>79	2291 (13.3%)	171 (3.4%)
Menopausal status at diagnosis		
Premenopausal	3263 (18.9%)	2115 (41.8%)
Perimenopausal[i]	796 (4.6%)	309 (6.1%)

	No neoadjuvant systemic treatment (N=17250)	Neoadjuvant systemic treatment (N=5054)
Postmenopausal	11467 (67.3%)	2038 (40.3%)
Unknown	1587 (9.2%)	592 (11.7%)
Socioeconomic status[ii]		
Low	5487 (31.8%)	1378 (27.3%)
Medium	6708 (38.9%)	1939 (38.4%)
High	5055 (29.3%)	1737 (34.4%)
Detected after population screening		
No / unknown	12132 (70.3%)	4488 (88.8%)
Yes	5118 (29.7%)	566 (11.2%)
Sublocalisation		
Outer quadrants	7083 (41.1%)	1872 (37.0%)
Inner quadrants	2892 (16.8%)	640 (12.7%)
Central parts	1615 (9.4%)	396 (7.8%)
Overlapping lesions	5398 (31.3%)	2087 (41.3%)
Unknown	262 (1.5%)	59 (1.2%)
Morphology		
Ductal	13031 (75.5%)	4068 (80.5%)
Lobular	2759 (16.0%)	706 (14.0%)
Mixed ductal lobular	736 (4.3%)	140 (2.8%)
Other	724 (4.2%)	140 (2.8%)
Differentiation grade		
Grade 1	3068 (17.8%)	367 (7.3%)

	No neoadjuvant systemic treatment (N=17250)	Neoadjuvant systemic treatment (N=5054)
Grade 2	8325 (48.3%)	1507 (29.8%)
Grade 3/4	5253 (30.5%)	1107 (21.9%)
Unknown	604 (3.5%)	2073 (41.0%)
Multifocality		
No	12110 (70.2%)	3103 (61.4%)
Yes	5058 (29.3%)	1886 (37.3%)
Unknown	82 (0.5%)	65 (1.3%)
Clinical tumour stage (cT)		
cTIS	607 (3.5%)	-
cT1	7418 (43.0%)	582 (11.5%)
cT2	7303 (42.3%)	2304 (45.6%)
cT3	1148 (6.7%)	1479 (29.3%)
cT4	216 (1.3%)	689 (13.6%)
Unknown	558 (3.2%)	-
Clinical nodal stage (cN)		
cN0	13992 (81.1%)	1727 (34.2%)
cN1	2984 (17.3%)	2763 (54.7%)
cN2	44 (0.3%)	118 (2.3%)
cN3	50 (0.3%)	446 (8.8%)
Unknown	180 (1.0%)	-
Pathological tumour stage (pT)		
pT0	-	1013 (20.0%)

	No neoadjuvant systemic treatment (N=17250)	Neoadjuvant systemic treatment (N=5054)
pTIS	-	212 (4.12%)
pT1	8391 (48.6%)	1576 (31.2%)
pT2	7501 (43.5%)	1425 (28.2%)
pT3	1153 (6.7%)	573 (11.3%)
pT4	205 (1.2%)	106 (2.1%)
Unknown	-	149 (2.9%)
Pathological nodal stage (pN)		
pN0	9704 (56.3%)	2299 (45.5%)
pN1	5573 (32.3%)	1689 (33.4%)
pN2	1224 (7.1%)	598 (11.8%)
pN3	749 (4.3%)	297 (5.9%)
Unknown	-	171 (3.4%)
Pathologic complete response		
No	NA	3941 (78.0%)
Yes	NA	954 (18.9%)
Unknown	NA	159 (3.1%)
Presence of DCIS component next to invasive component		
No	7987 (46.3%)	2998 (59.3%)
Yes	9086 (52.7%)	1916 (37.9%)
Unknown	177 (1.0%)	140 (2.8%)
Hormonal receptor status & treatment		

	No neoadjuvant systemic treatment (N=17250)	Neoadjuvant systemic treatment (N=5054)
Positive with endocrine therapy	10725 (62.2%)	3489 (69.0%)
Positive without endocrine therapy	3618 (21.0%)	189 (3.7%)
Negative	2728 (15.8%)	1332 (26.4%)
Unknown	179 (1.0%)	44 (0.9%)
HER2 status and treatment[iii]		
Negative	14196 (82.3%)	3759 (74.4%)
Positive with targeted therapy	1675 (9.7%)	1156 (22.9%)
Positive without targeted therapy	834 (4.8%)	41 (0.8%)
Unknown	545 (3.2%)	98 (1.9%)
Immediate breast reconstruction		
No	13428 (77.8%)	3594 (71.1%)
Yes	3822 (22.2%)	1460 (28.9%)
Axillary node dissection		
No	11958 (69.3%)	2582 (51.1%)
Yes	5292 (30.7%)	2472 (48.9%)
Sentinel node procedure		
No	3154 (18.3%)	2432 (48.1%)
Yes	14096 (81.7%)	2622 (51.9%)
Molecular diagnosis[iv]		
No	16012 (92.8%)	4953 (98.0%)
Yes, low risk	701 (4.1%)	39 (0.8%)
Yes, high risk	545 (2.6%)	58 (1.1%)

	No neoadjuvant systemic treatment (N=17250)	Neoadjuvant systemic treatment (N=5054)
Yes, results unknown	83 (0.5%)	4 (0.1%)
Radiotherapy type		
No radiotherapy	12560 (72.8%)	1669 (33.0%)
Breast/chestwall with boost	202 (1.2%)	61 (1.2%)
Breast/chestwall without boost	1159 (6.7%)	633 (12.5%)
Breast/chestwall with regional and boost	208 (1.2%)	279 (5.5%)
Breast/chestwall with regional without boost	2734 (15.8%)	2196 (43.5%)
Partial breast or other	267 (1.5%)	54 (1.0%)
Unknown[v]	120 (0.7%)	162 (3.2%)
Chemotherapy		
No	10213 (59.2%)	380 (7.5%)
Pre-surgical	-	4386 (86.8%)
Post-surgical	7037 (40.8%)	38 (0.8%)
Pre- and post-surgical	-	250 (4.9%)
Local recurrence as first event		
No	16896 (97.9%)	4972 (98.4%)
Yes	354 (2.1%)	82 (1.6%)

[i] Period between the last regular menstrual cycle and a year after the last menstruation

[ii] Socioeconomic status is based on postal code at diagnosis

[iii] For the non-NST group, HER2 status & treatment was reduced to HER2 status positive and negative because positive patients without targeted therapy were rare

[iv] Either Oncotype DX, MammaPrint or both. Intermediate oncotype risk was classified as low risk

[v] Patients who received radiotherapy but for whom the radiotherapy type is unknown

Regression models based on multiple imputation

Multiple imputation successfully replaced missing values in both the non-NST and the NST groups. For the non-NST group, the imputed variables were differentiation grade, multifocality, presence of DCIS component, ER status, (PR) status, HER2 status, menopausal status, radiotherapy type, and sublocalisation. The variables for HR status and treatment and HER2 status and treatment were constructed after imputation. For the NST group, the imputed variables were pathologic tumour and nodal stages, differentiation degree, multifocality, presence of DCIS component, ER status, PR status, HER2 status, menopausal status, radiotherapy type, and sublocalisation. The variables for HR status and treatment and pathologic complete response were constructed after imputation. The same variables that were used for the complete case analysis were included in the multivariable Cox regression models. The analyses based on the complete cases yielded similar results as those based on the imputed dataset, but with broader confidence intervals due to missingness (Supplementary A).

In the non-NST dataset, age group, menopausal status, socioeconomic status, screening detection, tumour sublocalisation, tumour morphology, tumour differentiation grade, multifocality, pathologic tumour stage, pathologic nodal stage, presence of DCIS, immediate reconstruction, HR status and treatment, HER2 status and treatment, radiotherapy type, and chemotherapy had a p-value of less than 0.2 in the univariable Cox regression. No impactful breaches of the proportional hazard assumption were found. These variables were included in the multivariable complete model (Table 2).

Table 2: Results from the complete non-NST multivariable Cox regression model based on the imputed dataset. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Age group	*60-69 (reference)*			
	<40	1.46	0.77 - 2.8	0.246
	40-49	0.95	0.56 - 1.6	0.839
	50-59	1.42	0.99 - 2.04	0.056
	70-79	1.14	0.8 - 1.61	0.463
	>79	0.87	0.59 - 1.28	0.485
Menopausal status	*Post (reference)*			
	Pre	1.01	0.66 - 1.53	0.979
	Peri	**0.53**	**0.29 - 0.97**	**0.039**
	Medium (reference)			

Factor	Level	HR	95% CI	p-value
Socioeconomic status	**Low**	**0.77**	**0.6 - 1**	**0.05**
	High	0.98	0.76 - 1.25	0.859
Screening	*No (reference)*			
	Yes	**0.63**	**0.47 - 0.85**	**0.002**
Sublocalisation	*Outer quadrants (reference)*			
	Inner quadrants	**1.42**	**1.07 - 1.87**	**0.015**
	Central parts	0.96	0.65 - 1.42	0.843
	Overlapping lesions	1.06	0.81 - 1.38	0.685
Morphology	*Ductal (reference)*			
	Lobular	1.12	0.81 - 1.56	0.482
	Mixed ductal lobular	0.72	0.38 - 1.37	0.32
	Other	0.65	0.37 - 1.15	0.138
Differentiation grade	*Grade 2 (reference)*			
	Grade 1	**0.72**	**0.52 - 1**	**0.048**
	Gade 3/4	1.04	0.78 - 1.38	0.793
Multifocality	*No (reference)*			
	Yes	1.1	0.85 - 1.42	0.479
Pathologic tumour stage	*pT1 (reference)*			
	pT2	**1.44**	**1.11 - 1.88**	**0.007**
	pT3	**1.89**	**1.16 - 3.07**	**0.011**
	pT4	2.08	0.97 - 4.49	0.061
Pathologic nodal stage	*pN0 (reference)*			
	pN1	**1.43**	**1.09 - 1.86**	**0.009**
	pN2	**2.45**	**1.48 - 4.04**	**0.001**
	pN3	**3.38**	**1.92 - 5.94**	**<0.001**
Presence of DCIS component	*No (reference)*			
	Yes	0.95	0.75 - 1.21	0.693
Immediate reconstruction	*No (reference)*			
	Yes	**1.45**	**1.08 - 1.95**	**0.012**
	Positive with endocrine therapy (reference)			

Factor	Level	HR	95% CI	p-value
Hormone receptor status and treatment	**Positive without endocrine therapy**	**2.1**	**1.58 - 2.8**	**<0.001**
	Negative	**2.78**	**2.05 - 3.78**	**<0.001**
HER2 status and treatment	*Negative (reference)*			
	Positive with targeted therapy	**0.29**	**0.14 - 0.58**	**0.001**
	Positive without targeted therapy	0.83	0.54 - 1.29	0.41
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.21**	**0.11 - 0.44**	**<0.001**
	Breast/chest wall and regional	**0.24**	**0.15 - 0.38**	**<0.001**
	Partial breast or other	0.83	0.37 - 1.91	0.668
Chemotherapy	*No (reference)*			
	Yes, post-surgical	**0.53**	**0.38 - 0.74**	**<0.001**

To produce the reduced model, morphology, multifocality, and the presence of a DCIS component were removed during backward selection (Table 3).

Table 3: Results from the reduced non-NST multivariable Cox regression model based on the imputed dataset. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Age group	*60-69 (reference)*			
	<40	1.46	0.76 - 2.78	0.253
	40-49	0.95	0.56 - 1.6	0.838
	50-59	1.42	0.99 - 2.04	0.056
	70-79	1.13	0.8 - 1.6	0.486
	>79	0.86	0.58 - 1.26	0.44
Menopausal status	*Post (reference)*			
	Pre	1	0.66 - 1.52	0.987

Factor	Level	HR	95% CI	p-value
	Peri	**0.53**	**0.3 - 0.97**	**0.039**
Socioeconomic status	*Medium (reference)*			
	Low	0.77	0.6 - 1	0.051
	High	0.98	0.76 - 1.25	0.861
Screening	*No (reference)*			
	Yes	**0.64**	**0.48 - 0.85**	**0.002**
Sublocalisation	*Outer quadrants (reference)*			
	Inner quadrants	**1.42**	**1.07 - 1.87**	**0.014**
	Central parts	0.96	0.65 - 1.42	0.829
	Overlapping lesions	1.07	0.83 - 1.39	0.594
Differentiation grade	*Grade 2 (reference)*			
	Grade 1	**0.71**	**0.51 - 0.97**	**0.034**
	Gade 3/4	1.03	0.78 - 1.35	0.857
Pathologic tumour stage	*pT1 (reference)*			
	pT2	**1.45**	**1.12 - 1.88**	**0.006**
	pT3	**1.91**	**1.18 - 3.07**	**0.008**
	pT4	2.11	0.98 - 4.55	0.055
Pathologic nodal stage	*pN0 (reference)*			
	pN1	**1.44**	**1.1 - 1.88**	**0.007**
	pN2	**2.51**	**1.52 - 4.13**	**<0.001**
	pN3	**3.5**	**1.99 - 6.14**	**<0.001**
Immediate reconstruction	*No (reference)*			
	Yes	**1.45**	**1.08 - 1.94**	**0.013**
Hormone receptor status and treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	**2.06**	**1.55 - 2.74**	**<0.001**
	Negative	**2.67**	**1.97 - 3.6**	**<0.001**
HER2 status and treatment	*Negative (reference)*			
	Positive with targeted therapy	**0.29**	**0.15 - 0.58**	**0.001**

Factor	Level	HR	95% CI	p-value
	Positive without targeted therapy	0.84	0.54 - 1.29	0.415
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.21**	**0.11 - 0.44**	**<0.001**
	Breast/chest wall with regional	**0.23**	**0.15 - 0.38**	**<0.001**
	Partial breast or other	0.84	0.37 - 1.91	0.67
Chemotherapy	*No (reference)*			
	Yes, post-surgical	**0.53**	**0.38 - 0.74**	**<0.001**

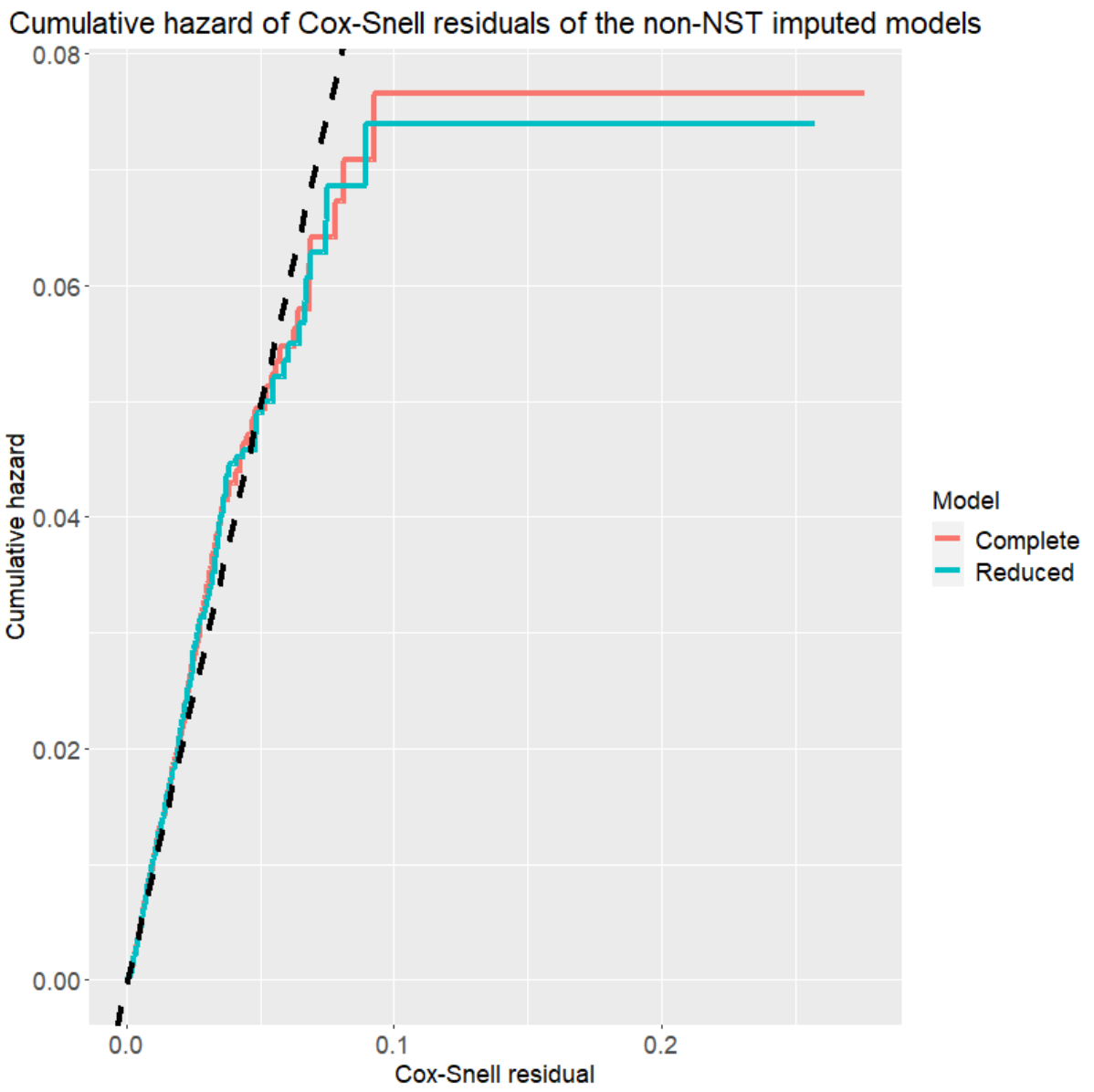

Figure 2: Cumulative hazard of the Cox-Snell residuals of the complete (red) and reduced (blue) non-NST models based on the imputed data compared to the expected cumulative hazard (black dashed).

Variables that were included in the multivariable model were menopausal status, socioeconomic status, screening detection, sublocalisation, morphology, tumour differentiation grade, multifocality, clinical tumour stage, clinical nodal stage, pathologic complete response, presence of DCIS component, immediate breast reconstruction, HR status and treatment, HER2 status and treatment, radiotherapy type and chemotherapy (Table 4). No impactful breaches of the proportional hazard assumption were found. With the exception of multifocality, these variables had p-values below 0.2 in the univariable analysis. Therefore, they were included in the multivariable model. Multifocality was included because of its contribution to the insightfulness of the model.

Table 4: Results from the complete NST multivariable Cox regression model based on the imputed dataset. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Menopausal status	*Post (reference)*			
	Pre	0.83	0.47 - 1.46	0.513
	Peri	1.38	0.64 - 2.99	0.405
Socioeconomic status	*Medium (reference)*			
	Low	0.72	0.39 - 1.31	0.272
	High	1.16	0.7 - 1.92	0.554
Screening	*No (reference)*			
	Yes	0.46	0.16 - 1.33	0.146
Sublocalisation	*Outer quadrants (reference)*			
	Inner quadrants	0.97	0.43 - 2.22	0.946
	Central parts	1.36	0.57 - 3.25	0.478
	Overlapping lesions	1.6	0.94 - 2.73	0.083
Morphology	*Ductal (reference)*			
	Lobular	0.68	0.27 - 1.72	0.41
	Mixed ductal lobular	0.94	0.22 - 4.1	0.937
	Other	0.66	0.15 - 2.85	0.574
Differentiation grade	*Grade 2 (reference)*			
	Grade 1	0.87	0.32 - 2.39	0.781
	Gade 3/4	1.63	0.87 - 3.04	0.12
Multifocality	*No (reference)*			
	Yes	0.88	0.53 - 1.45	0.615
Clinical tumour stage	*cT2 (reference)*			
	cT1	0.94	0.4 - 2.19	0.88
	cT3	1.35	0.76 - 2.42	0.303
	cT4	**2.26**	**1.15 - 4.41**	**0.019**
Clinical nodal stage	*cN1 (reference)*			
	cN0	1.17	0.65 - 2.09	0.6

Factor	Level	HR	95% CI	p-value
	>cN1	0.47	0.2 - 1.09	0.077
Pathologic complete response	*No (reference)*			
	Yes	**0.34**	**0.13 - 0.91**	**0.032**
Presence of DCIS component	*No (reference)*			
	Yes	**1.62**	**1.01 - 2.59**	**0.045**
HER2 status	*Negative (reference)*			
	Positive	**0.55**	**0.31 - 1**	**0.048**
Immediate reconstruction	*No (reference)*			
	Yes	0.97	0.55 - 1.73	0.922
Hormone receptor status & treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	2.19	0.75 - 6.44	0.149
	Negative	**2.31**	**1.3 - 4.1**	**0.005**
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.32**	**0.11 - 0.94**	**0.039**
	Breast/chest wall and regional	0.96	0.51 - 1.84	0.912
	Partial breast or other	1.11	0.18 - 6.91	0.91
Chemotherapy	*Yes, pre-surgical (reference)*			
	No	0.72	0.23 - 2.22	0.561
	Yes, post-surgical	2.39	0.3 - 18.75	0.401
	Yes, pre- and post-surgical	1.21	0.5 - 2.94	0.66

After backward selection, clinical tumour stage, pathologic complete response, presence of DCIS component, HER2 status, HR status and treatment, and radiotherapy type remained variables in the reduced model (Table 5).

Table 5: Results from the reduced NST multivariable Cox regression model based on the imputed dataset. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Clinical tumour stage	*cT2 (reference)*			
	cT1	0.91	0.4 - 2.1	0.824
	cT3	1.34	0.76 - 2.35	0.305
	cT4	**2.2**	**1.19 - 4.1**	**0.013**
Pathologic complete response	*No (reference)*			
	Yes	**0.35**	**0.14 - 0.92**	**0.033**
Presence of DCIS component	*No (reference)*			
	Yes	**1.7**	**1.09 - 2.65**	**0.021**
HER2 status	*Negative (reference)*			
	Positive	0.58	0.33 - 1.04	0.068
Hormone receptor status & treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	2.14	0.75 - 6.13	0.152
	Negative	**2.99**	**1.86 - 4.81**	**<0.001**
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.34**	**0.12 - 0.96**	**0.041**
	Breast/chest wall and regional	0.89	0.52 - 1.52	0.675
	Partial breast or other	1.11	0.19 - 6.61	0.904

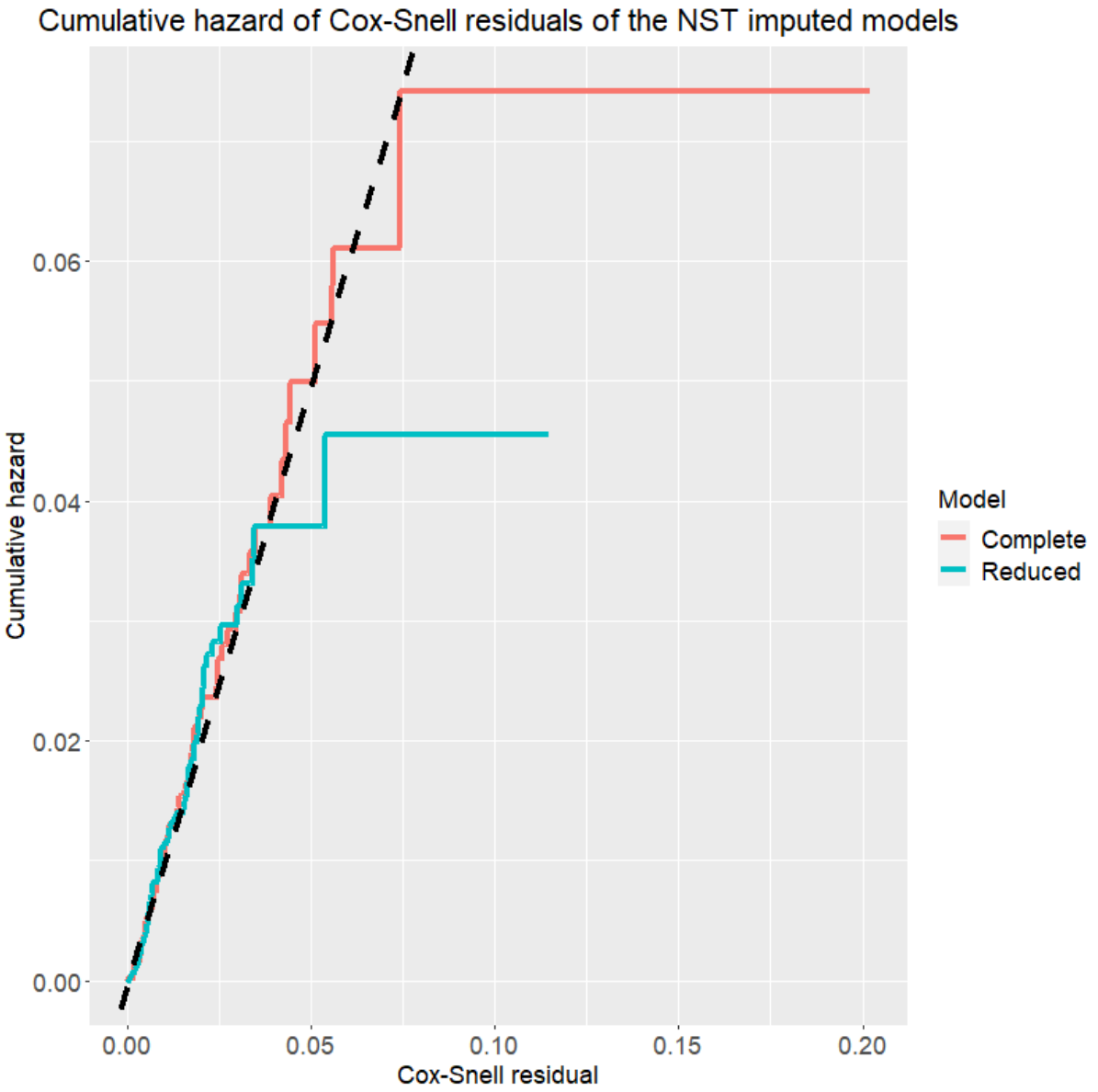

Figure 3: Cumulative hazard of the Cox-Snell residuals of the complete (red) and reduced (blue) NST models based on imputed data compared to the expected cumulative hazard (black dashed).

For both the non-NST and the NST models, the VIFs of the included variables were below the threshold value of 4. Furthermore, the interaction variables between age group and menopausal status and between immediate breast reconstruction and radiotherapy type were not statistically significant. Removing either of these variables and retaining the other did not significantly impact the estimates of the models.

Discussion

Interpretation of results

The aim of this study was to identify potential risk factors for local breast cancer recurrence after mastectomy. The Cox-Snell residuals for both patient groups show that the estimated cumulative hazard of both the complete and reduced models roughly match the expected cumulative hazard. This indicates that the models adequately fit the data. The reduced model does not show a significantly worse fit than the complete model, which indicates that no impactful variables were omitted during backwards selection. The estimated cumulative hazard for the reduced model is lower than for the complete model. This is in line with expectations, as the reduced model uses fewer variables to estimate the hazard compared to the full model, which leads to a lower estimate. At the ends of the graphs, large outliers are visible with a constant cumulative hazard. These could be attributed to patients who are censored or do not experience local recurrence during the follow-up period and thus do not contribute to the overall hazard function. As such, these outliers would not indicate a bad model fit.

While few studies have been conducted on risk factors for local recurrences after mastectomy, the results presented in this study show similarities with established literature regarding locoregional and overall recurrence. In the non-NST group, age was not significantly associated with local recurrence. However, the estimated HRs still somewhat indicated that younger ages increase the risk of local recurrence. These effects coincide with findings from Van Werkhoven *et al.,* who identified an increased risk of ipsilateral relapse for patients treated with either BCT or mastectomy below the age of 50. Above 50 years, the risk remained relatively stable [29]. However, Van Werkhoven *et al.* also included patients treated with BCT, which could affect the study outcomes if treatment allocation is different depending on the age of the patient. Furthermore, this turning point at 50 years could correspond to the perimenopausal status described in this study, which increased the risk of local recurrence compared to postmenopausal status. This effect could be attributed to hormonal changes during perimenopause, such as declining estrogen and progesterone levels, that influence the recurrence of breast cancer. Low socioeconomic status showed a protective effect for local recurrence compared to medium status. This is challenging to interpret on an individual level because the socioeconomic status was determined based solely on the postal code at diagnosis and no other individual circumstances. These results also contrast the literature, where high socioeconomic status demonstrated protective effects for any recurrence [30]. This difference in results can be explained by the fact that, contrary to what has been described in the literature, the characteristics of low, medium

and high socioeconomic status were almost identical [31]. Moreover, the current study considers the effect of socioeconomic status on exclusively local recurrence.

Detection through screening instead of clinical diagnosis is often associated with early detection and more favourable tumour characteristics [32,33]. This could explain the findings of this study, where screening was associated with reduced risk of local recurrence. Sublocalisation in the inner quadrants compared to the outer quadrants showed protective effects for local recurrence. A possible explanation could be the unfavourable surgical position of the inner quadrants, causing positive or close margins. However, patients with irradical surgery were excluded from this study. Therefore, the effect of the surgery would be less severe. Alternatively, clinicians might underdose these inner quadrants in an attempt to avoid cardiotoxicity. However, there was no significant variance in tumour laterality between the different sublocalisations. Tumour characteristics, such as differentiation grade, tumour stage and nodal stage were also associated with local recurrence, with higher stages increasing the risk. This coincides with other studies that relate higher tumour stage and nodal involvement to poor recurrence prognosis [10,12,34]. For the non-NST group, tumour stage 4 was the only stage that was not statistically significant, but this can be attributed to the significantly lower number of patients with this stage compared to the other stages. For the NST group, a lower clinical nodal stage at diagnosis was associated with a higher risk of recurrence, which contradicts expectations and results from the non-NST group. Unlike the non-NST group, the nodal stage could not be pathologically confirmed due to prior neoadjuvant treatment. Therefore, a possible explanation could be that the clinical nodal stage does not accurately describe the true nodal stage of the NST patients. High differentiation grade, indicative of poor differentiation, was found to increase the risk of locoregional recurrence for breast cancer patients after surgery and neoadjuvant treatment in a univariable analysis, but other results in the literature are scarce [35]. Good differentiation tends to be associated with a better prognosis due to slower growth and reduced aggressiveness, which can also help avoid local recurrence.

The presence of a DCIS component, in addition to the invasive tumour, was associated with an increased risk of recurrence after mastectomy for patients in the NST group. Similar observations have been reported for ipsilateral breast relapse after BCT [29]. Unlike in BCT, however, mastectomy should excise the DCIS that is confined to the ducts of the breast tissue. Nevertheless, if the DCIS extends over a larger area in the breast, this could cause residual DCIS in the surgical margins, especially if the foci are located further from the primary tumour. Additionally, if the DCIS components have different sensitivity to (neo)adjuvant treatment than then primary tumour, they might persist and lead to recurrence.

Lack of endocrine treatment increased the risk of recurrence for HR-positive patients. Similarly, the protective effects of targeted treatment for patients with positive HER2 status were demonstrated for

the non-NST group. This agrees with findings in the literature that identify the importance of HR status and HER2 as prognostic factors for breast cancer outcomes and recurrence [36]. For the NST group, explicitly combining HER2 status with targeted therapy was not viable because there were too few patients and insufficient events for HER2-positive patients who did not receive targeted therapy. The analysis does include HER2 status for NST patients, which reduced the risk of local recurrence. This inherently accounts for the impact of targeted therapy because, in practice, HER2-positive patients are almost always treated with targeted therapy, and HER2-negative patients are not.

In the literature, pCR, defined as ypT0N0 or ypT0/IS, has been demonstrated as a favourable prognostic factor to improve disease-free and overall survival after neoadjuvant chemotherapy [37–39]. The results of this study add to this, as pCR was associated with a reduced risk of local recurrence.

In this analysis, immediate post-mastectomy reconstruction was also associated with an elevated risk of local recurrence. Conversely, two small-scale retrospective studies found no significant association between immediate reconstruction and local recurrence rates [40,41]. However, these studies were based on small, localized cohorts. In 2022, a systematic review and meta-analysis examined the oncological outcomes after immediate and delayed breast reconstruction. The review also found no significant differences in the rates of local recurrence between immediate and delayed reconstruction after mastectomy. However, the authors highlight the lack of quality surveillance to detect recurrence after reconstruction, which could lead to underestimation of recurrence rates. In addition, they cite the low risk of local recurrence as an important obstacle to generating robust evidence on the oncological effects of different techniques and timing of reconstruction. They also suggest that immediate autologous breast reconstruction could form a risk of local recurrence due to the preservation of the skin envelope [42]. Indeed, less radical mastectomy types, such as skin-sparing and nipple-sparing mastectomies, have been shown to increase the likelihood of residual glandular breast tissue, mainly in the lower outer quadrant and the middle circle of the superficial dissection plane [43–45]. This corresponds with the hypothesis that residual glandular breast tissue after less radical mastectomies, whether or not to facilitate subsequent reconstruction, increases the risk of local recurrence. Moreover, local recurrence predominately occurs near the original tumour site [7]. While this study does not provide insight into the type of mastectomy or the type of reconstruction performed, the increased risk of local recurrence demonstrated after an immediate reconstruction could indicate a relation between the execution of the mastectomy procedure and local recurrence. However, patients with irradical surgery were excluded from the study population, so this should not significantly impact the risk of local recurrence. Alternatively, the observed increased risk after immediate reconstruction could be attributed to reluctance to irradiate after reconstruction or to reconstruct after radiotherapy [46,47]. Patients in both groups of this cohort who had immediate

reconstruction were less often irradiated than patients who did not have immediate reconstruction. However, no collinearity was detected between immediate reconstruction and radiotherapy.

The causal effect of treatment on the risk of recurrence should be carefully interpreted. Because the study is based on observational data, the treatments were not randomized between patients. Therefore, there is a risk of confounding by indication, and the estimated HRs should not be interpreted as indications of the treatment effect [48]. However, the results do concur with the described benefits of breast or chest wall radiotherapy, as both of these treatments demonstrate protective effects for recurrence [9,10]. Chemotherapy has been shown as an effective treatment of isolated local recurrence after breast cancer [49]. The results from this study suggest that post-surgical chemotherapy for non-NST can also reduce the risk of developing local recurrence.

Study limitations

The data collection for this cohort study was performed using the PALGA linkage. While this method saves on manual labour, some potential limitations should be addressed. The first limitation is that the algorithm flagged only pathologically confirmed diagnoses of LRR. In a separate validation performed by manually comparing the data collection using the PALGA linkage with a cohort from 2012 when this method was not yet implemented, it was determined that this method resulted in about 90% completeness regarding LRR. This implies that about 10% of LRR diagnoses were missed. In this cohort, 53% of the identified locoregional recurrences actually concerned isolated local recurrence as first event. Assuming a similar distribution between local and regional recurrence in the missed LRR would result in 5.3% missed local recurrence diagnoses. Therefore, the risk of local recurrence could have been underestimated slightly in the analysis of this study. Another limitation is that patients with only distant metastases were missed because the algorithm only flagged patients based on LRR. This could have affected the correction for competing events in the analysis. In a comparison of the effects of this missingness for distant metastases on the predicted risks and model performance in the influence 2.0 models, this limitation did not have a clinically relevant impact (data not yet published).

Another important note is that the results depend on proper documentation and completeness of the data. As previously discussed, some results can be explained by the limited number of patients or events in certain categories. When considering the complete dataset that was at our disposal, some variables had to be omitted in the multivariable analysis due to the large number of values, low patient numbers for each category or due to insufficient cases of local recurrence. For example, the variables radicality of the DCIS and the invasive components and molecular diagnosis could not be included in the analysis. For the same reasons, clinical nodal stages 2 and 3 had to be combined for the NST group and the distinction between radiotherapy boost was omitted. While this does somewhat dilute the

effects of these categories, the practical implications are limited as they concerned categories with few patients or no events of local recurrence. The effect of low event rates or limited numbers of patients was most prevalent in the complete case analysis of the NST group due to the overall lower number of patients included in this subgroup. This is reflected in the larger confidence intervals in the multivariable regression. While the use of multiple imputations did narrow these intervals, variables like HER2 status, hormone receptor status and treatment, radiotherapy type, and chemotherapy still exhibit wide confidence intervals.

The likelihood ratio test was used for the complete case analysis to assess the impact of the removed variables during backwards selection on the model fit. However, the likelihood ratio test was not used for the imputed dataset because it requires alternative methods for approximating or pooling the likelihood ratio statistic [50]. Instead, the reduced models for the imputed datasets were constructed using the reduced models of the complete case analyses as templates and including variables with a p-value below 0.07. With this approach, no additional variables were removed from the complete models of the imputed data that were not initially removed from the complete models of the complete cases. Moreover, visual inspection of the Cox-Snell residual plots did not indicate a worse model fit of the reduced models of the imputed data compared to the reduced models of the complete cases. Nevertheless, a more formal method for evaluating the effect of the variable removal on the model fit of the imputed data models would increase the methodological robustness of the approach.

While this study uncovered several potential risk factors for local recurrence after mastectomy, future research is advised. A similar study design based on a larger dataset with more cases of surgical treatment through mastectomy and more events can particularly achieve stronger scientific evidence for these identified factors. This could be achieved by including patients from more recent years of diagnosis, though this would require more time for a proper follow-up period and would increase the risk of introducing bias due to changing protocols in diagnosis and treatment throughout those years. Alternatively, international collaboration could provide larger patient databases, but beforehand, the generalizability of the study results to other populations outside of the female breast cancer population in the Netherlands should be assessed. With the results presented in this study, future research should additionally focus on personalizing patient treatment and follow-up to reduce the risk of local recurrence for patients presenting with these risk factors.

Conclusion

This study successfully identified potential risk factors for local breast cancer recurrence after mastectomy. These findings shed more light on the development of local recurrence after mastectomy

and help fill the research gap in this topic. With continued efforts to better understand the patient, tumour and treatment characteristics associated with local recurrence and by incorporating them into personalised treatment before and after mastectomy, breast cancer care can be further improved.

References

[1] IKNL. NKR Viewer 2023. https://nkr-cijfers.iknl.nl/ (accessed February 6, 2024).

[2] Liefers GJ, Scholten AN, Elkhuizen PHM, Linn SC, Schröder P, van Asperen CJ, et al. Mammatumoren. In: van Krieken JHJM, Beets-Tan RGH, Gelderblom AJ, Olofsen MJJ, Rutten HJT, editors. Leerboek oncologie. 10th ed., Houten: Bohn Stafleu van Loghum; 2020, p. 433 – 458. https://doi.org/10.1007/978-90-368-2449-1.

[3] Moo TA, Sanford R, Dang C, Morrow M. Overview of Breast Cancer Therapy. PET Clin 2018;13:339–54. https://doi.org/10.1016/j.cpet.2018.02.006.

[4] Kaidar-Person O, Offersen B V., Boersma LJ, de Ruysscher D, Tramm T, Kühn T, et al. A multidisciplinary view of mastectomy and breast reconstruction: Understanding the challenges. Breast 2021;56:42–52. https://doi.org/10.1016/j.breast.2021.02.004.

[5] IKNL. Behandeling borstkanker 2023. https://iknl.nl/kankersoorten/borstkanker/registratie/behandeling (accessed February 18, 2024).

[6] Moossdorff M, Van Roozendaal LM, Strobbe LJA, Aebi S, Cameron DA, Dixon JM, et al. Maastricht Delphi consensus on event definitions for classification of recurrence in breast cancer research. J Natl Cancer Inst 2014;106. https://doi.org/10.1093/jnci/dju288.

[7] Kaidar-Person O, Poortmans P, Offersen BV, Siesling S, Sklair-Levy M, Meattini I, et al. Spatial location of local recurrences after mastectomy: a systematic review. Breast Cancer Res Treat 2020;183:263–73. https://doi.org/10.1007/s10549-020-05774-4.

[8] Spronk I, Schellevis FG, Burgers JS, de Bock GH, Korevaar JC. Incidence of isolated local breast cancer recurrence and contralateral breast cancer: A systematic review. The Breast 2018;39:70–9. https://doi.org/10.1016/j.breast.2018.03.011.

[9] EBCTCG (Early Breast Cancer Trialists' Collaborative Group). Effect of radiotherapy after mastectomy and axillary surgery on 10-year recurrence and 20-year breast cancer mortality: meta-analysis of individual patient data for 8135 women in 22 randomised trials. The Lancet 2014;383:2127–35. https://doi.org/10.1016/S0140-6736(14)60488-8.

[10] Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of radiotherapy and of differences in the extent of surgery for early breast cancer on local recurrence and 15-year survival: an overview of the randomised trials. The Lancet 2005;366:2087–106. https://doi.org/10.1016/S0140-6736(05)67887-7.

[11] Kaidar-Person O, Vrou Offersen B, Hol S, Arenas M, Aristei C, Bourgier C, et al. ESTRO consensus guideline for target volume delineation in the setting of postmastectomy radiation

therapy after implant-based immediate reconstruction for early stage breast cancer. Radiotherapy and Oncology 2019;137:159–66. https://doi.org/10.1016/j.radonc.2019.04.010.

[12] Salvestrini V, Valzano M, Meattini I, Becherini C, Visani L, Francolini G, et al. Anatomical assessment of local recurrence site in breast cancer patients after breast reconstruction and post-mastectomy radiotherapy: implications for radiation volumes and techniques. Radiol Med 2024;129:845–54. https://doi.org/10.1007/s11547-024-01812-z.

[13] Offersen B V., Boersma LJ, Kirkove C, Hol S, Aznar MC, Sola AB, et al. ESTRO consensus guideline on target volume delineation for elective radiation therapy of early stage breast cancer, version 1.1. Radiotherapy and Oncology 2016;118:205–8. https://doi.org/10.1016/J.RADONC.2015.12.027.

[14] Van Maaren M, Hueting T, Van Uden D, Van Hezewijk M, De Munck L, Mureau M, et al. The INFLUENCE 3.0 model: updated predictions of locoregional recurrence and contralateral breast cancer, now also suitable for patients treated with neoadjuvant systemic therapy. Research Square (Research Square) 2024. https://doi.org/10.21203/rs.3.rs-4304136/v1.

[15] Bradburn MJ, Clark TG, Love SB, Altman DG. Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. Br J Cancer 2003;89:431–6. https://doi.org/10.1038/sj.bjc.6601119.

[16] Bewick V, Cheek L, Ball J. Statistics review 12: Survival analysis. Crit Care 2004;8:389–94. https://doi.org/10.1186/cc2955.

[17] Singh R, Mukhopadhyay K. Survival analysis in clinical trials: Basics and must know areas. Perspect Clin Res 2011;2:145. https://doi.org/10.4103/2229-3485.86872.

[18] In J, Lee DK. Survival analysis: Part II – applied clinical data analysis. Korean J Anesthesiol 2019;72:441–57. https://doi.org/10.4097/kja.19183.

[19] Ogundimu EO, Altman DG, Collins GS. Adequate sample size for developing prediction models is not simply related to events per variable. J Clin Epidemiol 2016;76:175–82. https://doi.org/10.1016/j.jclinepi.2016.02.031.

[20] Sanchez-Pinto LN, Venable LR, Fahrenbach J, Churpek MM. Comparison of variable selection methods for clinical predictive modeling. Int J Med Inform 2018;116:10–7. https://doi.org/10.1016/j.ijmedinf.2018.05.006.

[21] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. Stat Med 2007;26:2389–430. https://doi.org/10.1002/sim.2712.

[22] Zhang Z. Survival analysis in the presence of competing risks. Ann Transl Med 2017;5:47–47. https://doi.org/10.21037/atm.2016.08.62.

[23] Van Der Pas S, Nelissen R, Fiocco M. Different competing risks models for different questions may give similar results in arthroplasty registers in the presence of few events. Acta Orthop 2018;89:145–51. https://doi.org/10.1080/17453674.2018.1427314.

[24] Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res 2011;20:40–9. https://doi.org/10.1002/mpr.329.

[25] von Hippel PT. How Many Imputations Do You Need? A Two-stage Calculation Using a Quadratic Rule. Sociol Methods Res 2020;49:699–718. https://doi.org/10.1177/0049124117747303.

[26] von Hippel PT. 8. How to Impute Interactions, Squares, and other Transformed Variables. Sociol Methodol 2009;39:265–91. https://doi.org/10.1111/j.1467-9531.2009.01215.x.

[27] Kim JH. Multicollinearity and misleading statistical results. Korean J Anesthesiol 2019;72:558–69. https://doi.org/10.4097/kja.19087.

[28] Collett D. Model checking in the Cox Regression model. 4th ed. CRC Press; 2023.

[29] Werkhoven E van, Hart G, Tinteren H van, Elkhuizen P, Collette L, Poortmans P, et al. Nomogram to predict ipsilateral breast relapse based on pathology review from the EORTC 22881-10882 boost versus no boost trial. Radiotherapy and Oncology 2011;100:101–7. https://doi.org/10.1016/j.radonc.2011.07.004.

[30] van Maaren MC, Rachet B, Sonke GS, Mauguen A, Rondeau V, Siesling S, et al. Socioeconomic status and its relation with breast cancer recurrence and survival in young women in the Netherlands. Cancer Epidemiol 2022;77. https://doi.org/10.1016/j.canep.2022.102118.

[31] Klassen AC, Smith KC. The enduring and evolving relationship between social class and breast cancer burden: A review of the literature. Cancer Epidemiol 2011;35:217–34. https://doi.org/10.1016/j.canep.2011.02.009.

[32] Starikov A, Askin G, Blackburn A, Lu CM, Reznik E, Kim J, et al. Mode of detection matters: Differences in screen-detected versus symptomatic breast cancers. Clin Imaging 2021;80:11–5. https://doi.org/10.1016/j.clinimag.2021.06.032.

[33] Linda de Munck. Breast cancer: screening, stage, and outcome: Studies based on the Netherlands Cancer Registry. University of Groningen, 2020. https://doi.org/10.33612/diss.144606375.

[34] Hui C, Dirbas FM, Horst KC. Management of Local–Regional Recurrence of Breast Cancer. Curr Breast Cancer Rep 2023;15:185–95. https://doi.org/10.1007/s12609-023-00498-y.

[35] Chou H-H, Chung W-S, Ding R-Y, Kuo W-L, Yu C-C, Tsai H-P, et al. Factors affecting locoregional recurrence in breast cancer patients undergoing surgery following neoadjuvant treatment. BMC Surg 2021;21:160. https://doi.org/10.1186/s12893-021-01158-7.

[36] Kyndi M, Sørensen FB, Knudsen H, Overgaard M, Nielsen HM, Overgaard J. Estrogen receptor, progesterone receptor, HER-2, and response to postmastectomy radiotherapy in high-risk breast cancer: The Danish Breast Cancer Cooperative Group. Journal of Clinical Oncology 2008;26:1419–26. https://doi.org/10.1200/JCO.2007.14.5565.

[37] ORSARIA P, GRASSO A, IPPOLITO E, PANTANO F, SAMMARRA M, ALTOMARE C, et al. Clinical Outcomes Among Major Breast Cancer Subtypes After Neoadjuvant Chemotherapy: Impact on Breast Cancer Recurrence and Survival. Anticancer Res 2021;41:2697–709. https://doi.org/10.21873/anticanres.15051.

[38] Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. The Lancet 2014;384:164–72. https://doi.org/10.1016/S0140-6736(13)62422-8.

[39] Choi JY, Woen D, Jang SY, Lee H, Shin DS, Kwak Y, et al. Risk factors of breast cancer recurrence in pathologic complete response achieved by patients following neoadjuvant chemotherapy: a single-center retrospective study. Front Oncol 2023;13. https://doi.org/10.3389/fonc.2023.1230310.

[40] Kroll SS, Schusterman MA, Tadjalli HE, Singletary SE, Ames FC. Risk of recurrence after treatment of early breast cancer with skin-sparing mastectomy. Ann Surg Oncol 1997;4:193–7. https://doi.org/10.1007/BF02306609.

[41] Murphy RX, Wahhab S, Rovito PF, Harper G, Kimmel SR, Kleinman LC, et al. Impact of Immediate Reconstruction on the Local Recurrence of Breast Cancer After Mastectomy. Ann Plast Surg 2003;50:333–8. https://doi.org/10.1097/01.SAP.0000041488.88950.A2.

[42] Bargon CA, Young-Afat DA, Ikinci M, Braakenburg A, Rakhorst HA, Mureau MAM, et al. Breast cancer recurrence after immediate and delayed postmastectomy breast reconstruction—A systematic review and meta-analysis. Cancer 2022;128:3449–69. https://doi.org/10.1002/cncr.34393.

[43] Kaidar-Person O, Boersma LJ, Poortmans P, Sklair-Levy M, Offersen BV, Cardoso MJ, et al. Residual Glandular Breast Tissue After Mastectomy: A Systematic Review. Ann Surg Oncol 2020;27:2288–96. https://doi.org/10.1245/s10434-020-08516-4.

[44] Deutschmann C, Singer CF, Gschwantler-Kaulich D, Pfeiler G, Leser C, Baltzer PAT, et al. Residual fibroglandular breast tissue after mastectomy is associated with an increased risk of a local recurrence or a new primary breast cancer". BMC Cancer 2023;23. https://doi.org/10.1186/s12885-023-10764-y.

[45] Griepsma M, de Roy van Zuidewijn DBW, Grond AJK, Siesling S, Groen H, de Bock GH. Residual Breast Tissue after Mastectomy: How Often and Where Is It Located? Ann Surg Oncol 2014;21:1260–6. https://doi.org/10.1245/s10434-013-3383-x.

[46] See MS-F, Farhadi J. Radiation Therapy and Immediate Breast Reconstruction. Clin Plast Surg 2018;45:13–24. https://doi.org/10.1016/j.cps.2017.08.001.

[47] van Bommel ACM, Schreuder K, Veenstra RK, de Ligt KM, Vrancken Peeters M-JTFD, Maduro JH, et al. Discrepancies Between Surgical Oncologists and Plastic Surgeons in Patient Information Provision and Personal Opinions Towards Immediate Breast Reconstruction. Ann Plast Surg 2018;81:383–8. https://doi.org/10.1097/SAP.0000000000001572.

[48] Sendor R, Stürmer T. Core concepts in pharmacoepidemiology: Confounding by indication and the role of active comparators. Pharmacoepidemiol Drug Saf 2022;31:261–9. https://doi.org/10.1002/pds.5407.

[49] Aebi S, Gelber S, Anderson SJ, Láng I, Robidoux A, Martín M, et al. Chemotherapy for isolated locoregional recurrence of breast cancer (CALOR): a randomised trial. Lancet Oncol 2014;15:156–63. https://doi.org/10.1016/S1470-2045(13)70589-8.

[50] Grund S, Lüdtke O, Robitzsch A. Pooling methods for likelihood ratio tests in multiply imputed data sets. Psychol Methods 2023;28:1207–21. https://doi.org/10.1037/met0000556.

Supplementary A – Complete case analyses

Non-NST multivariable Cox regression

Table A1: Results from the complete non-NST multivariable Cox regression model based on the complete cases. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Age group	*60-69 (reference)*			
	<40	1.18	0.56 - 2.51	0.664
	40-49	0.71	0.37 - 1.34	0.29
	50-59	1.34	0.91 - 1.98	0.141
	70-79	1.01	0.71 - 1.44	0.953
	>79	0.69	0.46 - 1.04	0.074
Menopausal status	*Post (reference)*			
	Pre	1.28	0.77 - 2.11	0.339
	Peri	**0.35**	**0.14 - 0.88**	**0.025**
Socioeconomic status	*Medium (reference)*			
	Low	0.77	0.58 - 1.01	0.06
	High	0.95	0.72 - 1.25	0.718
Screening	*No (reference)*			
	Yes	**0.65**	**0.48 - 0.89**	**0.007**
Sublocalisation	*Outer quadrants (reference)*			
	Inner quadrants	1.31	0.97 - 1.78	0.082
	Central parts	0.99	0.65 - 1.49	0.946
	Overlapping lesions	0.99	0.74 - 1.32	0.945
Morphology	*Ductal (reference)*			
	Lobular	1.17	0.82 - 1.66	0.394
	Mixed ductal lobular	0.79	0.4 - 1.55	0.491
	Other	0.56	0.29 - 1.1	0.093
Differentiation grade	*Grade 2 (reference)*			
	Grade 1	0.73	0.52 - 1.02	0.068
	Gade 3/4	0.98	0.72 - 1.32	0.876

Factor	Level	HR	95% CI	p-value
Multifocality	*No (reference)*			
	Yes	1.09	0.82 - 1.43	0.556
Pathologic tumour stage	*pT1 (reference)*			
	pT2	1.27	0.96 - 1.69	0.095
	pT3	**1.82**	**1.09 - 3.03**	**0.022**
	pT4	1.97	0.87 - 4.47	0.104
Pathologic nodal stage	*pN0 (reference)*			
	pN1	**1.61**	**1.21 - 2.14**	**0.001**
	pN2	**3.07**	**1.83 - 5.13**	**<0.001**
	pN3	**4.56**	**2.51 - 8.27**	**<0.001**
Presence of DCIS component	*No (reference)*			
	Yes	1.07	0.83 - 1.37	0.626
Immediate reconstruction	*No (reference)*			
	Yes	**1.41**	**1.02 - 1.95**	**0.036**
Hormone receptor status & treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	**1.95**	**1.43 - 2.66**	**<0.001**
	Negative	**3.13**	**2.25 - 4.35**	**<0.001**
HER2 status & treatment	*Negative (reference)*			
	Positive with targeted therapy	**0.27**	**0.12 - 0.6**	**0.001**
	Positive without targeted therapy	0.75	0.46 - 1.21	0.231
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.2**	**0.09 - 0.43**	**<0.001**
	Breast/chest wall and regional	**0.21**	**0.13 - 0.35**	**<0.001**
	Partial breast or other	0.91	0.4 - 2.09	0.825
Chemotherapy	*No (reference)*			
	Yes, post-surgical	**0.45**	**0.32 - 0.65**	**<0.001**

Table A2: Results from the reduced non-NST multivariable Cox regression model based on the complete cases. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Age group	*60-69 (reference)*			
	<40	1.16	0.55 - 2.45	0.705
	40-49	0.69	0.37 - 1.31	0.261
	50-59	1.33	0.9 - 1.96	0.151
	70-79	1	0.7 - 1.43	0.992
	>79	0.68	0.45 - 1.03	0.067
Menopausal status	*Post (reference)*			
	Pre	1.29	0.78 - 2.12	0.323
	Peri	**0.35**	**0.14 - 0.87**	**0.025**
Screening	*No (reference)*			
	Yes	**0.65**	**0.48 - 0.88**	**0.006**
Pathologic tumour stage	*pT1 (reference)*			
	pT2	1.28	0.97 - 1.69	0.085
	pT3	**1.81**	**1.1 - 2.99**	**0.02**
	pT4	1.87	0.83 - 4.21	0.128
Pathologic nodal stage	*pN0 (reference)*			
	pN1	**1.59**	**1.2 - 2.11**	**1.59**
	pN2	**3.12**	**1.86 - 5.21**	**3.12**
	pN3	**4.65**	**2.56 - 8.43**	**4.65**
Immediate reconstruction	*No (reference)*			
	Yes	**1.41**	**1.03 - 1.95**	**0.035**
Hormone receptor status & treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	**1.78**	**1.32 - 2.42**	**<0.001**
	Negative	**2.97**	**2.22 - 3.97**	**<0.001**
HER2 status & treatment	*Negative (reference)*			
	Positive with targeted therapy	**0.28**	**0.13 - 0.62**	**0.002**

Factor	Level	HR	95% CI	p-value
	Positive without targeted therapy	0.78	0.49 - 1.25	0.303
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.2**	**0.09 - 0.43**	**<0.001**
	Breast/chest wall and regional	**0.21**	**0.13 - 0.35**	**<0.001**
	Partial breast or other	0.93	0.41 - 2.14	0.873
Chemotherapy	*No (reference)*			
	Yes, post-surgical	**0.47**	**0.33 - 0.67**	**<0.001**

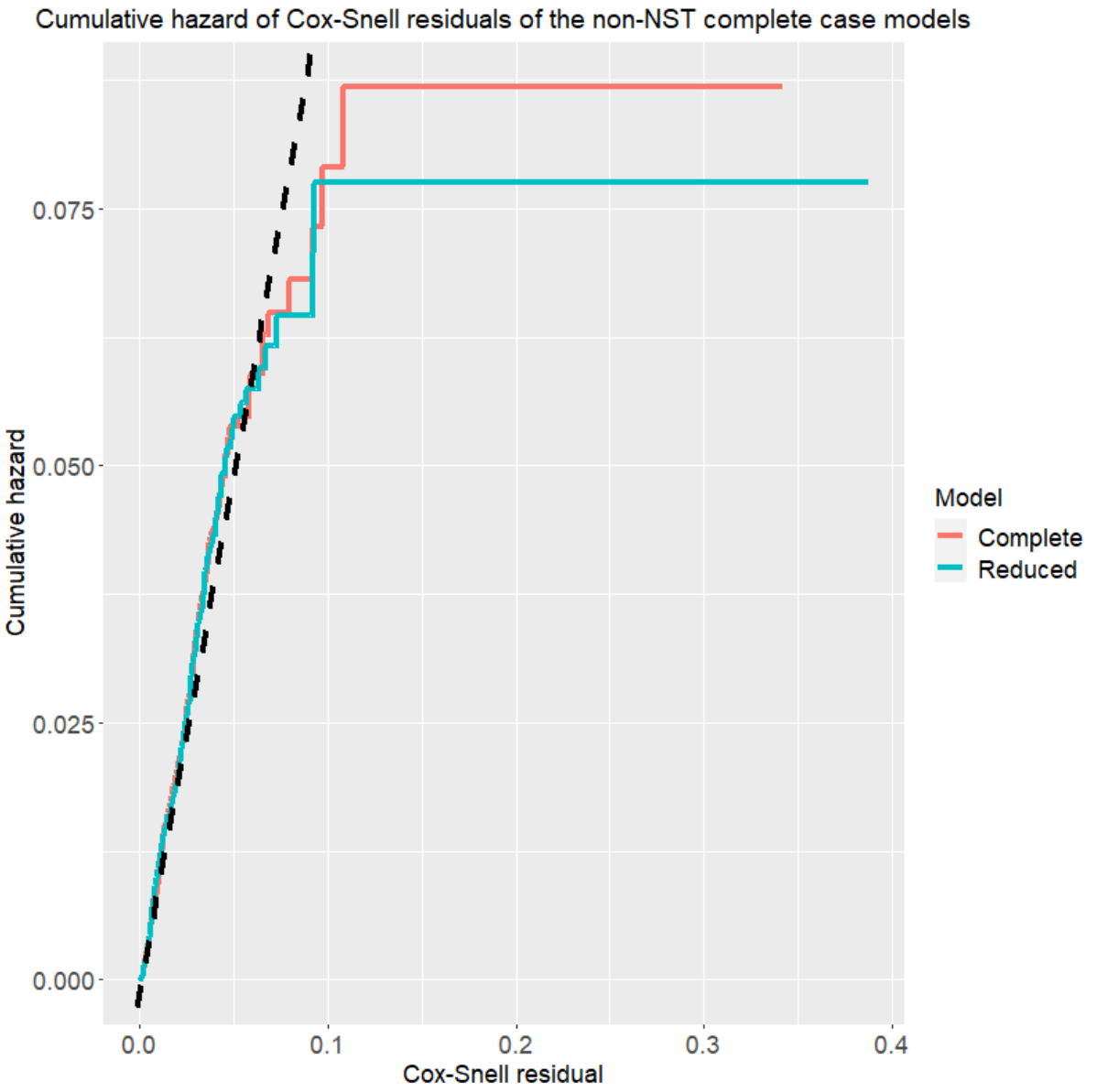

Figure A1: Cumulative hazard of the Cox-Snell residuals of the complete (red) and reduced (blue) non-NST models based on the complete cases compared to the expected cumulative hazard (black dashed).

NST multivariable Cox regression

Table A3: Results from the complete NST multivariable Cox regression model based on the complete cases. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Menopausal status	*Post (reference)*			
	Pre	1.11	0.54 - 2.31	0.771
	Peri	2.3	0.84 - 6.28	0.103
	Medium (reference)			

Factor	Level	HR	95% CI	p-value
Socioeconomic status	Low	0.91	0.43 - 1.9	0.797
	High	0.94	0.47 - 1.87	0.857
Screening	*No (reference)*			
	Yes	0.45	0.1 - 1.97	0.289
Sublocalisation	*Outer quadrants (reference)*			
	Inner quadrants	0.7	0.23 - 2.13	0.526
	Central parts	1.64	0.58 - 4.63	0.35
	Overlapping lesions	1.51	0.75 - 3.02	0.248
Morphology	*Ductal (reference)*			
	Lobular	0.71	0.2 - 2.52	0.597
	Mixed ductal lobular	1	0.13 - 7.7	0.999
	Other[i]	0	0 - Inf	0.995
Differentiation grade	*Grade 2 (reference)*			
	Grade 1	1.08	0.39 - 3.02	0.884
	Gade 3/4	1.91	0.94 - 3.91	0.076
Multifocality	*No (reference)*			
	Yes	0.64	0.32 - 1.25	0.188
Clinical tumour stage	*cT2 (reference)*			
	cT1	0.67	0.22 - 2.01	0.474
	cT3	0.98	0.45 - 2.13	0.952
	cT4	2.01	0.8 - 5.05	0.136
Clinical nodal stage	*cN1 (reference)*			
	cN0	1.44	0.67 - 3.12	0.352
	>cN1	0.35	0.1 - 1.21	0.096
Pathologic complete response	*No (reference)*			
	Yes	**0.31**	**0.11 - 0.89**	**0.029**
Presence of DCIS component	*No (reference)*			
	Yes	1.68	0.9 - 3.14	0.103
HER2 status	*Negative (reference)*			

Factor	Level	HR	95% CI	p-value
	Positive	0.82	0.37 - 1.79	0.614
Immediate reconstruction	*No (reference)*			
	Yes	1.09	0.52 - 2.29	0.828
Hormone receptor status & treatment	*Positive with endocrine therapy (reference)*			
	Positive without endocrine therapy	0.89	0.12 - 6.73	0.906
	Negative	**2.63**	**1.26 - 5.48**	**0.01**
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.1**	**0.01 - 0.81**	**0.031**
	Breast/chest wall and regional	1.01	0.45 - 2.26	0.983
	Partial breast or other	1.89	0.24 - 14.93	0.547
Chemotherapy	*Yes, pre-surgical (reference)*			
	No	1.28	0.38 - 4.32	0.688
	Yes, post-surgical	2.88	0.36 - 23.27	0.322
	Yes, pre- and post-surgical	1.96	0.73 - 5.28	0.183

[i] Insufficient events due to missingness in the data

Table A4: Results from the reduced NST multivariable Cox regression model based on the complete cases. Statistically significant estimates (p-value < 0.05) are shown in bold, protective effects are marked in green, and risk-increasing effects are marked in red.

Factor	Level	HR	95% CI	p-value
Pathologic complete response	*No (reference)*			
	Yes	**0.32**	**0.12 - 0.85**	**0.022**
Presence of DCIS component	*No (reference)*			
	Yes	1.79	1 - 3.2	0.052
	Positive with endocrine therapy (reference)			

Factor	Level	HR	95% CI	p-value
Hormone receptor status & treatment	Positive without endocrine therapy	0.83	0.11 - 6.12	0.851
	Negative	**3.27**	**1.77 - 6.04**	**<0.001**
Radiotherapy type	*No radiotherapy (reference)*			
	Breast/chest wall	**0.11**	**0.01 - 0.84**	**0.033**
	Breast/chest wall and regional	0.86	0.47 - 1.57	0.617
	Partial breast or other	1.46	0.19 - 10.96	0.712

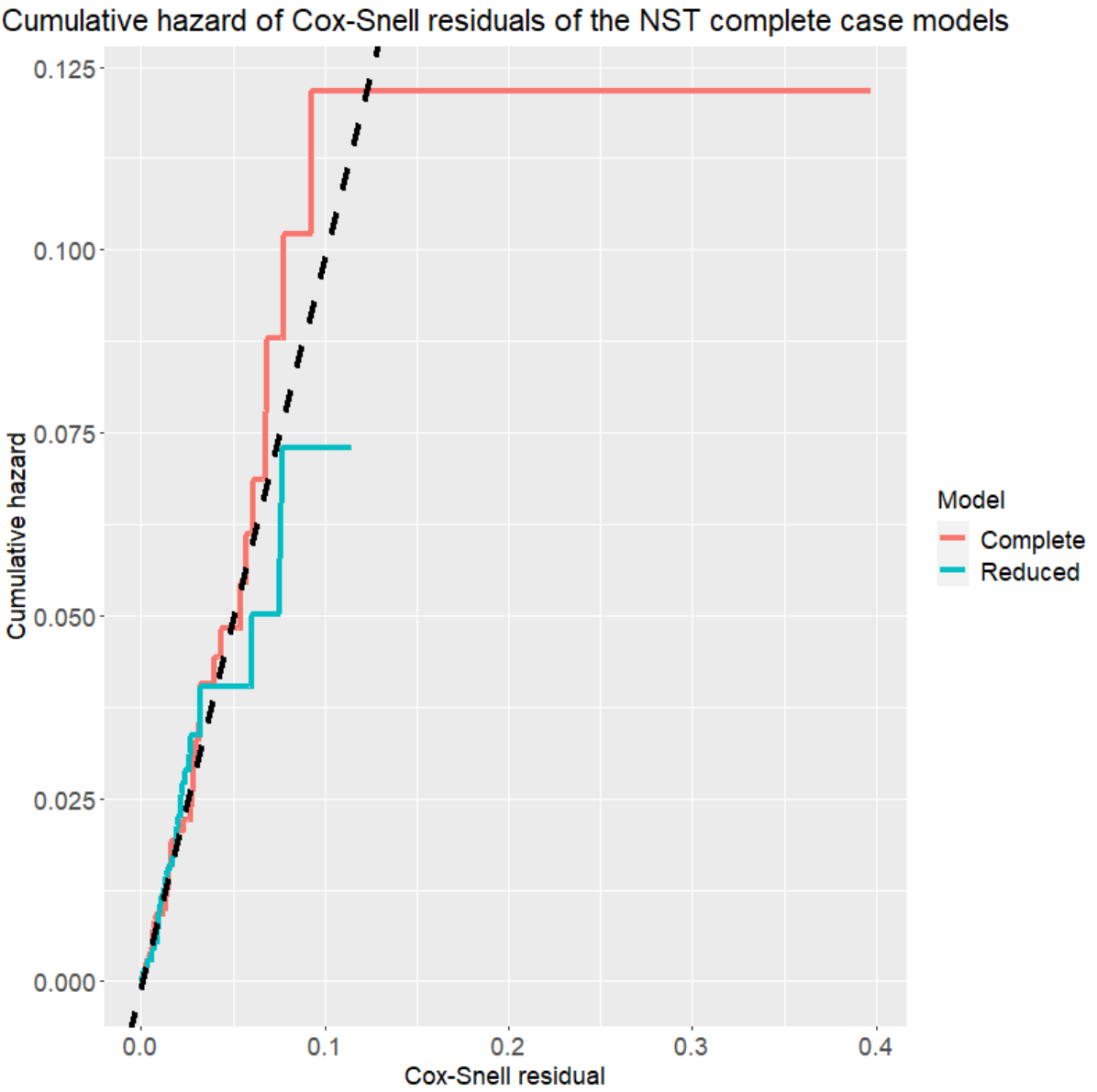

Figure A2: Cumulative hazard of the Cox-Snell residuals of the complete (red) and reduced (blue) NST models based on the complete cases compared to the expected cumulative hazard (black dashed).