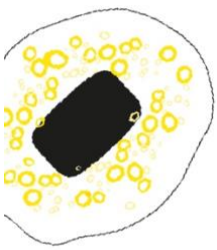
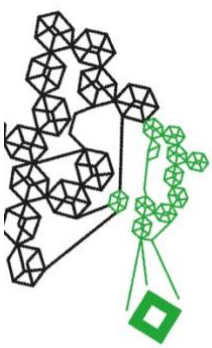
Master Thesis



The risk of developing lung cancer after breast cancer

Sophie Nagtegaal



1st supervisor: prof. dr. S. Siesling
2nd supervisor: dr. J. Mikhal
Daily supervision: dr. A. Eijkelboom

June 2025

Faculty of S&T
Health Sciences
Optimization of Healthcare Processes
University of Twente

Abstract

Background

Second primary lung cancer (SPLC) is a serious long-term complication among breast cancer survivors, with limited knowledge on how treatment characteristics influence this risk. This study aimed to assess associations between breast cancer treatments and SPLC risk, and to compare tumor characteristics of SPLCs in breast cancer survivors with those of primary lung cancer in women without a history of cancer.

Methods

All women newly diagnosed with breast cancer between 2005 and 2020 were selected from the Netherlands Cancer Registry. Breast cancer patients were grouped into four cohorts based on year of diagnosis and cancer being either invasive or in situ. Multivariate logistic regression, Cox proportional hazards models, and a competing risk Cox model were used to assess associations between patient, tumor, and treatment characteristics and the risk of SPLC. In addition, all women diagnosed with primary lung cancer or SPLC between 2010 and 2023 were selected to compare tumor characteristics between SPLCs and primary lung cancers. Lung tumor characteristics were compared between breast cancer survivors with SPLC and women diagnosed with a primary lung tumor without a history of cancer.

Results

Increased risk of SPLC was observed among women aged 50-74 years compared to those under 50 years, although this association was not consistent for patients aged 75+. Lower socioeconomic status and HR-/HER2- breast tumor receptor status were linked to higher risk of SPLC, while stage III breast tumors and multifocal breast tumors were associated with decreased risk. In the 2010-2020 invasive cohort, radiotherapy and chemotherapy were associated with reduced risk of SPLC. These findings were largely confirmed in Cox regression. The competing risk Cox model confirmed that older patients had higher mortality before SPLC, highlighting the importance of accounting for competing events. Compared to women with primary lung cancer, SPLCs in breast cancer survivors were more frequently located in the upper lobe and diagnosed at an earlier stage.

Conclusion

This study shows that certain patient characteristics, breast tumor characteristics, and treatment characteristics are associated with the risk of developing SPLC in breast cancer survivors. Differences were found in tumor location and stage at diagnosis between SPLCs and primary lung cancers.

Table of Contents

1 Introduction

Breast cancer is the most frequently diagnosed malignancy in women worldwide, with an estimated 2.3 million new cases reported in 2022 (1). Advancements in screening programs, surgical techniques, radiotherapy, and systemic therapies have led to significant improvements in survival rates in recent decades (2). As a result, the number of long-term breast cancer survivors continues to grow. This shift in survivorship has led to an increasing focus on late effects of treatment. This includes the development of second primary malignancies (SPMs). Among these, second primary lung cancer (SPLC), is particularly concerning due to its poor prognosis, with a 5-year survival rate of only 19% (3, 4).

Several studies have found that breast cancer survivors have an increased risk of developing SPLC compared to individuals without a history of cancer (5, 6). This risk may be influenced by a combination of patient characteristics, characteristics of the primary breast tumor, and treatment-related factors.

Age and socioeconomic status (SES) are important patient-related factors that may contribute to the development of SPLC according to previous research. The incidence of breast cancer increases with age, and older breast cancer survivors may be more at risk of developing secondary malignancies due to cumulative carcinogenic exposure, immune senescence and impaired DNA repair mechanisms (7-10). Furthermore, low SES is associated with behavioral and environmental risk factors, including higher smoking prevalence and increased exposure to air pollution, as well as inequalities in healthcare access and quality. These factors may all contribute to lung cancer risk (11-13). Although these associations are supported by existing evidence, further research is needed to improve our understanding of how patient-related factors influence the development of SPLC in breast cancer survivors.

Several characteristics of the initial breast tumor may influence the likelihood of developing SPLC. Higher TNM-stage appears to have a protective effect against the development of SPLC, likely due to competing risks. Patients with advanced breast cancer may not survive long enough to develop SPLC (14, 15). Furthermore, poorly or undifferentiated breast tumors have been associated with a higher risk of SPLC, potentially due to their aggressive biological behavior and underlying genomic instability (16, 17). Receptor status of the primary breast tumor is also linked to SPLC by previous research. Hormone receptor-negative breast cancers, particularly triple-negative, have been linked to an increased SPLC risk, possibly because of shared etiological factors or tumor aggressiveness (16, 18). EGFR mutations in SPLC seem to be more frequently observed in survivors of HER2-negative breast cancer, suggesting a possible interaction between receptor pathways (19). Invasive ductal carcinoma, the most prevalent subtype, has been linked to an increased risk of SPLC compared to other histologies. This may reflect its unique biological features or differences in treatment patterns (16). Although several breast tumor characteristics have been linked to SPLC risk, further research is needed to clarify their individual and combined contributions. Histological subtype also appears to play a role.

Radiotherapy is widely recognized as a potential contributor to this increased risk, possibly due to the exposure of lung tissue to ionizing radiation. The risk is particularly high following locoregional radiotherapy, which targets the breast and regional lymph nodes, such as those in the axilla (5). Exposure to radiation can induce DNA damage, oxidative stress, inflammation and fibrosis, known to contribute to carcinogenesis over time (20-22). Also, SPLCs are more frequently observed in the ipsilateral lung compared to the contralateral lung, further supporting the hypothesis of radiation-induced carcinogenesis (23).

While the focus has been on radiotherapy, the role of other breast cancer treatments remains underexplored. Chemotherapeutic agents may cause changes in genetic material that carry on after treatment, but their contribution to the development of SPLC risk is unclear (24). Similarly, endocrine and targeted therapies may have long-term effects on lung tissue (25). As the use of systemic therapies alongside radiotherapy and surgery has become more prevalent, a more comprehensive understanding of the individual effects of these treatments on SPLC risk is needed.

In addition to risk factors, gaps remain in understanding the biological and clinical characteristics of SPLCs in breast cancer survivors. Although adenocarcinoma is the most common subtype in both primary lung cancer and SPLC (26-28), it remains unclear whether SPLCs differ in terms of morphology, localization, differentiation, or molecular profiles compared to primary lung cancers in individuals without a history of cancer. Further research is needed to determine whether SPLCs represent distinct tumor types, which could have implications for the individualization of diagnostic and therapeutic strategies.

For this study, data from the Netherlands Cancer Registry (NCR) is used. The aim is to investigate how specific breast cancer treatment characteristics, such as radiotherapy, chemotherapy, endocrine therapy and targeted therapy, as well as patient and tumor characteristics, are associated with the risk of developing SPLC. It will also examine whether the tumor characteristics of SPLCs in breast cancer survivors differ from those of primary lung cancers in individuals without a history of cancer.

2 Methods

2.1 Study design

This study is a retrospective, population-based cohort study using data from the NCR, which systematically collects information on all newly diagnosed malignancies in the Netherlands. The registry includes detailed data on patient demographics, tumor characteristics, treatments. Data for this study were de-identified and provided following the guidelines of the General Data Protection Regulation (29).

2.2 Study population

Two cohorts were constructed for this study. The first cohort, i.e. the breast cancer cohort, aimed to investigate the association between specific breast cancer treatment characteristics and the risk of developing SPLC. This cohort included all female patients aged 18 years or older, diagnosed with primary invasive breast cancer or ductal carcinoma in situ (DCIS) between 2005 and 2020.

Due to the implementation of a new clinical radiotherapy protocol, using a 15x 2.67 Gy regimen from 2010 onwards, the breast cancer dataset was stratified into two periods: before 2010 and from 2010 onwards. Data containing SES was only available for diagnoses from 2010 onwards, justifying separate analyses for this subgroup. Within each period, the breast cancer cohort was further subdivided based on tumor behavior into DCIS and invasive cancer. Molecular subtyping based on estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) status is only performed in the Netherlands for invasive tumors. Therefore, subtype information was only available for patients with invasive disease. Subtype-based analyses and imputations were restricted to the invasive subgroups.

Patients in the breast cancer dataset who did not undergo surgery were excluded. To avoid the potential bias of including synchronous breast tumors with different treatment strategies, we excluded all patients diagnosed with a second breast tumor within 91 days. This threshold is consistent with common definitions of synchronous malignancies and ensures that only primary, independently treated tumors were included in the analysis.

The second cohort, i.e. the lung cancer cohort, aimed to compare tumor characteristics of SPLC with those of primary lung cancer in individuals without a history of cancer. This cohort included all female patients aged 18 years or older, diagnosed with invasive lung cancer between 2010 and 2023. Women with a previous malignancy other than breast cancer were excluded. Unlike the breast cancer cohort, we did not exclude patients with synchronous tumors in the lung cancer cohort, as we were specifically interested in the tumor characteristics following breast cancer, rather than patient characteristics.

2.3 Variables and definitions

For the breast cancer cohort, patient-related variables included age at diagnosis (categorized as <50, 50-59, 60-74, and 75+ years) and socioeconomic status (SES: low income [<24,300 EUR], medium income [24,300-31,000 EUR], and high income [>31,000 EUR]). Tumor-related variables included morphology (ductal, lobular + mixed lobular/ductal, and other carcinomas), behavior (invasive and in situ), receptor status (HR+/HER2-, HR+/HER2+, HR-/HER2+, HR-/HER2-, and unknown), laterality (left and right), differentiation grade (grade 1 [well differentiated], grade 2 [moderately differentiated], and grade 3-4 [poorly/undifferentiated]), TNM stage (stage 0, stage I, stage II, and stage III), and multifocality (yes and no). Hormone receptor (HR) status was defined as HR-positive (HR+) when either estrogen receptor (ER) or progesterone receptor was (PR) positive, and HR-negative (HR-) when both ER and PR were negative, based on ER, PR, and human epidermal growth factor receptor 2 (HER2) status. Treatment-related variables included radiotherapy (breast/chest wall, breast/chest wall + regional, partial breast irradiation, and other), chemotherapy (yes and no), endocrine therapy (yes and no), targeted therapy (yes and no), and type of surgery (breast-conserving [including lumpectomy, resection, and node dissection], non-breast-conserving [including mastectomy], other surgery, and not applicable). A combined variable was created to streamline the baseline table, in which morphology and tumor behavior were merged into one variable

(ductal [malignant], ductal [in situ], lobular + mixed ductal/lobular [malignant], lobular + mixed ductal/lobular [in situ], other carcinoma [malignant], and other carcinoma [in situ].

For the lung cancer cohort, patient-related variables included age at diagnosis (categorized as <59, 60-74, and 75+ years) and SES (low income [<24,300 EUR], middle income [24,300-31,000 EUR], and high income [>31,000 EUR]). Tumor-related variables included morphology (squamous cell carcinoma, adenocarcinoma [including adenosquamous carcinoma], large cell carcinoma [not otherwise specified], no pathological confirmation, and other), tumor sublocalization (main bronchus, upper lobe of lung, middle lobe of lung, lower lobe of lung, overlapping lesions of lung, and lung NOS), laterality (left, right, bilateral, and missing), TNM stage (stage 0, stage I, stage II, stage III, and stage IV), multifocality (yes and no), and differentiation grade (grade 1, grade 2, grade 3, and grade 4). Time between previous malignancy and lung cancer diagnosis was included as both a continuous variable (in days) and categorical variable (no previous cancer, <1, 2-5 years, 5-9 years, 10-14 years, 15-19 years, 20-24 years, 25-29 years, 30-34 years, 35-39 years, 40-44 years, 45-49 years, and 50-54 years). For patients with a previous malignancy, additional variables included the topography of the previous tumor (breast cancer or not applicable), behavior of the previous tumor (in situ, malignant, metastatic, or not applicable), and tumor laterality in relation to the breast tumor (same side, opposite side, missing, or not applicable). Molecular data included mutation status (EGFR mutation, KRAS mutation, ALK mutation, other mutation(s), and unknown/not tested/no oncogenic mutation), which was available for patients with stage IV non-squamous tumors from 2017 onwards. PD-L1 expression was available as a continuous variable (percentage) for patients with stage IV tumors from 2018 onwards.

2.4 Statistical analysis

All statistical analyses were performed with Stata Statistical Software: Release 17 (StataCorp, College Station, TX). Descriptive statistics were used to summarize the baseline characteristics of the cohorts. Categorical variables were presented as frequencies and percentages, and continuous variables were summarized as medians with interquartile ranges (IQRs). Normality of continuous variables was assessed using the Shapiro-Wilk test. Due to non-normal distributions, Mann-Whitney U tests were used to compare continuous variables between groups, and chi-squared tests were used for categorical variables.

2.4.1 Handling of missing data

As several variables contained missing data, multiple imputation by chained equations (MICE) was used to create 25 imputed datasets with 10 burn-in iterations. Imputation models were chosen based on variable type, where multinomial logistic regression was used for nominal variables, ordered logistic regression for ordinal variables, and binary logistic regression for dichotomous variables, all with augmentation to handle potential perfect prediction.

The variables included in the imputation model depended on tumor behavior and year of diagnosis. For invasive breast cancer patients diagnosed between 2005 and 2009, six variables with missing values were imputed: receptor status, tumor stage, differentiation grade, laterality, multifocality and surgery type. For DCIS patients diagnosed in this period, receptor status was excluded, as receptor status is not routinely tested for this tumor type. Laterality and tumor stage had no missing values in this cohort, and thus did not need imputation. Only differentiation grade, multifocality, and surgery type were imputed.

For patients diagnosed from 2010 onwards, the imputation model for the invasive breast cancer additionally included type of radiotherapy and SES. In the DCIS cohort from 2010 onwards, the same variables were imputed as for the 2005-2009 cohort, but SES was included.

All imputation models used a common set of predictors, including age group, tumor behavior, morphology, type of later malignancy, chemotherapy, endocrine therapy, radiotherapy, targeted therapy, and SPLC status. For DCIS cohorts, chemotherapy, endocrine therapy, and targeted therapy were excluded as predictor variables, as DCIS is not typically treated with these types of therapies. Also, tumor behavior was not included in the imputation models for the DCIS cohorts, as it is the same value for all patients. Imputed variables were also used as predictors for the imputation of other missing values.

For analysis purposes, local tumor resection and lymph node dissection were grouped with breast-conserving surgery, but the original categories were retained for descriptive summaries. Rubin's rules were used to combine outcomes, assuming missing at random (MAR).

2.4.2 Breast cancer cohort

Univariate logistic regression was used to estimate odds ratios (ORs) and 95% confidence intervals (CIs), to evaluate the association between each individual patient-, tumor-, and treatment-related variable and the risk of developing SPLC. A multivariable logistic regression model was used to assess the adjusted association between patient-, tumor-, and treatment-related variables and SPLC risk. Variables included in the model were selected based on clinical relevance and univariate analyses, and included age group, SES, tumor grade, multifocality, laterality, morphology, tumor stage, receptor status, radiotherapy, chemotherapy, endocrine therapy, and targeted therapy.

A Cox proportional hazards model was used to estimate the time-dependent risk of SPLC. For this analysis, a new variable was created representing the time in years between the first breast cancer diagnosis and SPLC diagnosis. This interval was used as the time variable in the Cox model, with patients who did not develop SPLC censored at the end of follow-up.

A separate cause-specific Cox proportional hazards model was used to evaluate factors associated with death before SPLC occurred. In this model, patients who developed SPLC or were still alive at the end of the follow-up period were censored. This approach took into account the competing risk of death, enabling a more accurate interpretation of SPLC risk in the context of overall survival.

For these analyses, SPLC was only considered if it was the first subsequent primary malignancy after the first breast cancer diagnosis of a patient. Patients who developed other primary cancers before a lung tumor were not classified as having SPLC and were excluded from the comparison groups. This approach ensured that lung tumors classified as SPLC reflected the earliest second malignancy following a breast cancer diagnosis.

2.4.3 Lung cancer cohort

Among patients with a history of breast cancer, the laterality of the lung tumor was compared to the laterality of the previous breast cancer. Patients were classified as having lung tumors on the same side (ipsilateral) or the opposite side (contralateral) as their breast cancer.

To examine the time interval between breast cancer and SPLC diagnoses, the same variable (in years) was used as in the Fine and Gray analysis. This variable was categorized into one-year intervals to analyze the latency distribution, and the number of cases in each interval was calculated. The distribution of SPLC cases was visualized using a line graph.

3 Results

3.1 Study population and baseline characteristics

The breast cancer cohort was analyzed separately for women diagnosed with invasive breast cancer and those diagnosed with ductal carcinoma in situ (DCIS). Tables 1 and 2 present the baseline characteristics for each group, stratified by year of diagnosis (2005-2009 vs. 2010-2020) and SPLC development. In total, 61,153 women diagnosed between 2005 and 2009 and 150,787 women diagnosed between 2010 and 2020. Within these cohorts, 2.2% (n = 1,390) and 1.2% (n = 1,925) of patients, respectively, developed an SPLC.

The lung cancer cohort consisted of 71,955 women diagnosed with invasive lung cancer between 2010 and 2023, 11.01% (n = 7,924) had a prior diagnosis of breast cancer.

3.1.1 Breast cancer cohort

The cohort of women with invasive breast cancer consisted of 55,160 women diagnosed between 2005 and 2009 and 131,381 women diagnosed between 2010 and 2020. In both time periods, women who developed SPLC were significantly more likely to be aged 60-74 years at breast cancer diagnosis than those who did not (42.7% vs. 32.3% in 2005-2009; 56.6% vs. 36.7% in 2010-2020). Also, SPLC patients were also more likely to have the lowest SES (35.6% vs. 27.6% in 2010-2020) (11).

Regarding tumor characteristics, a larger proportion of women with SPLC had stage I breast tumors (49.5% vs. 42.0% in 2005-2009; 57.9% vs. 48.2% in 2010-2020), and were more likely to have well-differentiated grade 1 tumors (25.0% vs. 20.3% in 2005-2009; 26.5% vs. 21.6% in 2010-2020), compared to women who did not develop SPLC. They also more often had HR+/HER2- tumors (70.0% vs. 63.6% in 2005-2009; 75.2% vs. 71.9 % in 2010-2020).

Looking at treatment, women with SPLC were more likely to receive breast-conserving surgery (63.4% vs. 52.2% in 2005-2009; 65.1% vs. 60.2% in 2010-2020), compared to women who did not develop SPLC. Although radiotherapy was more frequently administered in 2005-2009 (72.4% vs. 64.2%), this difference was not present in 2010-2020 (71.2% vs. 70.8%). In contrast, SPLC patients were less likely to receive chemotherapy (35.4% vs. 40.2% in 2005-2009; 30.8% vs. 40.6% in 2010-2020) or target therapy (5.1% vs. 8.4% in 2005-2009; 7.7% vs. 10.4% in 2010-2020).

Similar differences were observed in the DCIS cohort, but systemic treatments were rarely provided. Women who developed SPLC were more likely to be aged 60-74 years (46.4% vs. 38.1% in 2005-2009; 57.8% vs. 42.2% in 2010-2020) and more likely to be in the lowest socioeconomic group (39.0% vs. 24.3%) in 2010-2020. Grade 1 tumors occurred slightly less frequent in SPLC patients in 2005-2009 (11.3% vs. 15.3%), but slightly more frequent in 2010-2020 (18.5% vs. 16.9%).

Women who developed SPLC were slightly more likely to undergo breast-conserving surgery (63.6% vs. 58.2% in 2005-2009; 64.8% vs. 65.5% in 2010-2020). Similarly, radiotherapy was also slightly more prevalent among women that developed SPLC (53.6% vs. 51.0% in 2005-2009; 55.7% vs. 53.6% in 2010-2020). However, chemotherapy and targeted therapy were not used for any patients that developed SPLC, and endocrine therapy was rarely used.

Table 1. Baseline characteristics of the invasive breast cancer cohort, stratified by year of diagnosis (2005-2009 vs. 2010-2020) and development of SPLC. Values are presented as N (%).

	2005-2009				**2010-2020**			
	Total	**No SPLC**	**SPLC**	***p* Value**	**Total**	**No SPLC**	**SPLC**	***p* Value**
	N=55,160	N=53,921	N=1,239		N=131,381	N=129,743	N=1,638	
Age at breast cancer diagnosis				<0.001				<0.001
<50 years	14,011 (25.4%)	13,821 (25.6%)	190 (15.3%)		29,255 (22.3%)	29,134 (22.5%)	121 (7.4%)	
50-59 years	14,391 (26.1%)	13,939 (25.9%)	452 (36.5%)		33,372 (25.4%)	32,944 (25.4%)	428 (26.1%)	
60-74 years	17,963 (32.6%)	17,434 (32.3%)	529 (42.7%)		48,542 (36.9%)	47,615 (36.7%)	927 (56.6%)	
75+ years	8,795 (15.9%)	8,727 (16.2%)	68 (5.5%)		20,212 (15.4%)	20,050 (15.5%)	162 (9.9%)	
Socioeconomic status (SES)								<0.001
Low income (< €24,300)	-	-	-		36,372 (27.7%)	35,789 (27.6%)	583 (35.6%)	
Middle income (€24,300 - €31,000)	-	-	-		45,518 (34.6%)	44,916 (34.6%)	602 (36.8%)	
High income (> €31,000)	-	-	-		48,131 (36.6%)	47,690 (36.8%)	441 (26.9%)	
Missing					1,360 (1.0%)	1,348 (1.0%)	12 (0.7%)	
TNM-stage				<0.001				<0.001
Stage I	23,282 (42.2%)	22,669 (42.0%)	613 (49.5%)		63,463 (48.3%)	62,515 (48.2%)	948 (57.9%)	
Stage II	23,321 (42.3%)	22,816 (42.3%)	505 (40.8%)		52,068 (39.6%)	51,505 (39.7%)	563 (34.4%)	
Stage III	8,270 (15.0%)	8,153 (15.1%)	117 (9.4%)		15,594 (11.9%)	15,468 (11.9%)	126 (7.7%)	
Missing	287 (0.5%)	283 (0.5%)	4 (0.3%)		256 (0.2%)	255 (0.2%)	1 (0.1%)	
Differentiation grade				<0.001				<0.001
Grade 1 (Well differentiated)	11,233 (20.4%)	10,923 (20.3%)	310 (25.0%)		28,439 (21.6%)	28,005 (21.6%)	434 (26.5%)	
Grade 2 (Moderately differentiated)	21,741 (39.4%)	21,236 (39.4%)	505 (40.8%)		57,067 (43.4%)	56,350 (43.4%)	717 (43.8%)	
Grade 3-4 (Poorly/ Undifferentiated)	15,662 (28.4%)	15,337 (28.4%)	325 (26.2%)		31,829 (24.2%)	31,474 (24.3%)	355 (21.7%)	
Missing	6,524 (11.8%)	6,425 (11.9%)	99 (8.0%)		14,046 (10.7%)	13,914 (10.7%)	132 (8.1%)	
Lateralization				0.260				0.058
Left side	28,427 (51.5%)	27,817 (51.6%)	610 (49.2%)		67,080 (51.1%)	66,198 (51.0%)	882 (53.8%)	
Right side	26,695 (48.4%)	26,067 (48.3%)	628 (50.7%)		64,255 (48.9%)	63,499 (48.9%)	756 (46.2%)	
Missing	38 (0.1%)	37 (0.1%)	1 (0.1%)		46 (0.0%)	46 (0.0%)	0 (0.0%)	

Receptor status				<0.001				0.001
HR+/HER2-	35,144 (63.7%)	34,277 (63.6%)	867 (70.0%)		94,564 (72.0%)	93,333 (71.9%)	1,231 (75.2%)	
HR+/HER2+	4,384 (7.9%)	4,293 (8.0%)	91 (7.3%)		11,357 (8.6%)	11,254 (8.7%)	103 (6.3%)	
HR-/HER2+	2,722 (4.9%)	2,689 (5.0%)	33 (2.7%)		5,484 (4.2%)	5,421 (4.2%)	63 (3.8%)	
HR-/HER2-	5,748 (10.4%)	5,634 (10.4%)	114 (9.2%)		14,050 (10.7%)	13,866 (10.7%)	184 (11.2%)	
Missing	7,162 (13.0%)	7,028 (13.0%)	134 (10.8%)		5,926 (4.5%)	5,869 (4.5%)	57 (3.5%)	
Tumor multifocality				0.002				0.001
No	44,851 (81.3%)	43,796 (81.2%)	1,055 (85.1%)		107,709 (82.0%)	106,309 (81.9%)	1,400 (85.5%)	
Yes	8,778 (15.9%)	8,623 (16.0%)	155 (12.5%)		22,098 (16.8%)	21,877 (16.9%)	221 (13.5%)	
Missing	1,531 (2.8%)	1,502 (2.8%)	29 (2.3%)		1,574 (1.2%)	1,557 (1.2%)	17 (1.0%)	
Morphology				0.510				0.970
Ductal	42,917 (77.8%)	41,938 (77.8%)	979 (79.0%)		105,250 (80.1%)	103,941 (80.1%)	1,309 (79.9%)	
Lobular + mixed ductal/lobular	8,093 (14.7%)	7,917 (14.7%)	176 (14.2%)		18,821 (14.3%)	18,585 (14.3%)	236 (14.4%)	
Other carcinomas	4,150 (7.5%)	4,066 (7.5%)	84 (6.8%)		7,310 (5.6%)	7,217 (5.6%)	93 (5.7%)	
Surgery type				<0.001				<0.001
Breast-conserving surgery (incl. lumpectomy & local resection)	28,914 (52.4%)	28,129 (52.2%)	785 (63.4%)		79,216 (60.3%)	78,150 (60.2%)	1,066 (65.1%)	
Non-breast-conserving surgery (incl. mastectomy)	23,249 (42.1%)	22,807 (42.3%)	442 (35.7%)		44,078 (33.5%)	43,541 (33.6%)	537 (32.8%)	
Unknown	20 (0.0%)	19 (0.0%)	1 (0.1%)		71 (0.1%)	68 (0.1%)	3 (0.2%)	
Other surgery (incl. incidental finding & lymph node dissection)	73 (0.1%)	71 (0.1%)	2 (0.2%)		160 (0.1%)	155 (0.1%)	5 (0.3%)	
No surgery	2,904 (5.3%)	2,895 (5.4%)	9 (0.7%)		7,856 (6.0%)	7,829 (6.0%)	27 (1.6%)	
Chemotherapy				<0.001				<0.001
No	33,052 (59.9%)	32,252 (59.8%)	800 (64.6%)		78,222 (59.5%)	77,089 (59.4%)	1,133 (69.2%)	
Yes	22,108 (40.1%)	21,669 (40.2%)	439 (35.4%)		53,159 (40.5%)	52,654 (40.6%)	505 (30.8%)	
Target therapy				<0.001				<0.001
No	50,593 (91.7%)	49,417 (91.6%)	1,176 (94.9%)		117,751 (89.6%)	116,239 (89.6%)	1,512 (92.3%)	
Yes	4,567 (8.3%)	4,504 (8.4%)	63 (5.1%)		13,630 (10.4%)	13,504 (10.4%)	126 (7.7%)	
Radiotherapy				<0.001				0.74

No	19,621 (35.6%)	19,279 (35.8%)	342 (27.6%)		38,338 (29.2%)	37,866 (29.2%)	472 (28.8%)	
Yes	35,539 (64.4%)	34,642 (64.2%)	897 (72.4%)		93,043 (70.8%)	91,877 (70.8%)	1,166 (71.2%)	
Endocrine therapy				0.037				<0.001
No	27,320 (49.5%)	26,670 (49.5%)	650 (52.5%)		56,834 (43.3%)	56,014 (43.2%)	820 (50.1%)	
Yes	27,840 (50.5%)	27,251 (50.5%)	589 (47.5%)		74,547 (56.7%)	73,729 (56.8%)	818 (49.9%)	
Radiotherapy type				<0.001				<0.001
Breast/Chest wall irradiation	4,497 (8.2%)	4,390 (8.1%)	107 (8.6%)		57,581 (43.8%)	56,822 (43.8%)	759 (46.3%)	
Breast/Chest wall + regional irradiation	1,828 (3.3%)	1,795 (3.3%)	33 (2.7%)		23,395 (17.8%)	23,188 (17.9%)	207 (12.6%)	
Partial breast irradiation	3 (0.0%)	3 (0.0%)	0 (0.0%)		2,062 (1.6%)	2,032 (1.6%)	30 (1.8%)	
Other (e.g. only regional lymph nodes)	73 (0.1%)	72 (0.1%)	1 (0.1%)		1,049 (0.8%)	1,035 (0.8%)	14 (0.9%)	
Not applicable	19,621 (35.6%)	19,279 (35.8%)	342 (27.6%)		38,338 (29.2%)	37,866 (29.2%)	472 (28.8%)	
Missing	29,138 (52.8%)	28,382 (52.6%)	756 (61.0%)		8,956 (6.8%)	8,800 (6.8%)	156 (9.5%)	
Later malignancies				<0.001				<0.001
No	43,701 (79.2%)	43,701 (81.0%)	0 (0.0%)		116,401 (88.6%)	116,401 (89.7%)	0 (0.0%)	
Yes	11,459 (20.8%)	10,220 (19.0%)	1,239 (100.0%)		14,980 (11.4%)	13,342 (10.3%)	1,638 (100.0%)	

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; HR, hormone receptor.
HR status was defined as HR-positive (HR+) when either estrogen receptor (ER) or progesterone receptor was (PR) positive, and HR-negative (HR-) when both ER and PR were negative, based on ER, PR, and human epidermal growth factor receptor 2 (HER2) status.

Table 2. Baseline characteristics of the DCIS breast cancer cohort, stratified by year of diagnosis (2005-2009 vs. 2010-2020) and development of SPLC. Values are presented as N (%).

	2005-2009				**2010-2020**			
	Total	**No SPLC**	**SPLC**	***p* Value**	**Total**	**No SPLC**	**SPLC**	***p* Value**
	N=5,993	N=5,842	N=151		N=19,406	N=19,119	N=287	
Age at breast cancer diagnosis				0.030				<0.001
<50 years	1,120 (18.7%)	1,102 (18.9%)	18 (11.9%)		2,916 (15.0%)	2,894 (15.1%)	22 (7.7%)	
50-59 years	2,186 (36.5%)	2,128 (36.4%)	58 (38.4%)		7,214 (37.2%)	7,127 (37.3%)	87 (30.3%)	
60-74 years	2,298 (38.3%)	2,228 (38.1%)	70 (46.4%)		8,233 (42.2%)	8,067 (42.2%)	166 (57.8%)	
75+ years	389 (6.5%)	384 (6.6%)	5 (3.3%)		1,043 (5.4%)	1,031 (5.4%)	12 (4.2%)	
Socioeconomic status (SES)								<0.001
Low income (< €24,300)	-	-	-		4,755 (24.5%)	4,643 (24.3%)	112 (39.0%)	
Middle income (€24,300 - €31,000)	-	-	-		6,741 (34.7%)	6,650 (34.8%)	91 (31.7%)	
High income (> €31,000)	-	-	-		7,701 (39.7%)	7,619 (39.9%)	82 (28.6%)	
Missing					209 (1.1%)	207 (1.1%)	2 (0.7%)	
Differentiation grade				0.150				0.170
Grade 1 (Well differentiated)	909 (15.2%)	892 (15.3%)	17 (11.3%)		3,290 (17.0%)	3,237 (16.9%)	53 (18.5%)	
Grade 2 (Moderately differentiated)	1,687 (28.1%)	1,650 (28.2%)	37 (24.5%)		6,738 (34.7%)	6,650 (34.8%)	88 (30.7%)	
Grade 3-4 (Poorly/ Undifferentiated)	2,932 (48.9%)	2,852 (48.8%)	80 (53.0%)		8,489 (43.7%)	8,351 (43.7%)	138 (48.1%)	
Missing	465 (7.8%)	448 (7.7%)	17 (11.3%)		889 (4.6%)	881 (4.6%)	8 (2.8%)	
Lateralization				0.071				0.890
Left side	3,178 (53.0%)	3,087 (52.8%)	91 (60.3%)		10,034 (51.7%)	9,889 (51.7%)	145 (50.5%)	
Right side	2,815 (47.0%)	2,755 (47.2%)	60 (39.7%)		9,368 (48.3%)	9,226 (48.3%)	142 (49.5%)	
Missing	0 (0.0%)	0 (0.0%)	0 (0.0%)		4 (0.0%)	4 (0.0%)	0 (0.0%)	
Tumor multifocality				0.990				0.720
No	5,020 (83.8%)	4,894 (83.8%)	126 (83.4%)		17,803 (91.7%)	17,538 (91.7%)	265 (92.3%)	
Yes	635 (10.6%)	619 (10.6%)	16 (10.6%)		1,269 (6.5%)	1,253 (6.6%)	16 (5.6%)	
Missing	338 (5.6%)	329 (5.6%)	9 (6.0%)		334 (1.7%)	328 (1.7%)	6 (2.1%)	
Morphology				0.015				0.100

Ductal	5,802 (96.8%)	5,662 (96.9%)	140 (92.7%)		18,735 (96.5%)	18,457 (96.5%)	278 (96.9%)	
Lobular + mixed ductal/lobular	107 (1.8%)	101 (1.7%)	6 (4.0%)		454 (2.3%)	451 (2.4%)	3 (1.0%)	
Other carcinomas	84 (1.4%)	79 (1.4%)	5 (3.3%)		217 (1.1%)	211 (1.1%)	6 (2.1%)	
Surgery type				0.037				0.006
Breast-conserving surgery (incl. lumpectomy & local resection)	3,498 (58.4%)	3,402 (58.2%)	96 (63.6%)		12,714 (65.5%)	12,528 (65.5%)	186 (64.8%)	
Non-breast-conserving surgery (incl. mastectomy)	2,343 (39.1%)	2,290 (39.2%)	53 (35.1%)		5,646 (29.1%)	5,556 (29.1%)	90 (31.4%)	
Unknown	18 (0.3%)	16 (0.3%)	2 (1.3%)		5 (0.0%)	4 (0.0%)	1 (0.3%)	
Other surgery (incl. incidental finding & lymph node dissection)	4 (0.1%)	4 (0.1%)	0 (0.0%)		59 (0.3%)	58 (0.3%)	1 (0.3%)	
No surgery	130 (2.2%)	130 (2.2%)	0 (0.0%)		982 (5.1%)	973 (5.1%)	9 (3.1%)	
Chemotherapy				0.780				0.830
No	5,990 (99.9%)	5,839 (99.9%)	151 (100%)		19,403 (100.0%)	19,116 (100.0%)	287 (100.0%)	
Yes	3 (0.1%)	3 (0.1%)	0 (0%)		3 (0.0%)	3 (0.0%)	0 (0.0%)	
Target therapy				0.820				0.860
No	5,991 (100.0%)	5,840 (100.0%)	151 (100.0%)		19,404 (100.0%)	19,117 (100.0%)	287 (100.0%)	
Yes	2 (0.0%)	2 (0.0%)	0 (0.0%)		2 (0.0%)	2 (0.0%)	0 (0.0%)	
Radiotherapy				0.520				0.460
No	2,933 (48.9%)	2,863 (49.0%)	70 (46.4%)		9,004 (46.4%)	8,877 (46.4%)	127 (44.3%)	
Yes	3,060 (51.1%)	2,979 (51.0%)	81 (53.6%)		10,402 (53.6%)	10,242 (53.6%)	160 (55.7%)	
Endocrine therapy				0.340				0.640
No	5,958 (99.4%)	5,807 (99.4%)	151 (100.0%)		19,308 (99.5%)	19,023 (99.5%)	285 (99.3%)	
Yes	35 (0.6%)	35 (0.6%)	0 (0.0%)		98 (0.5%)	96 (0.5%)	2 (0.7%)	
Radiotherapy type				0.880				0.130
Breast/Chest wall irradiation	354 (5.9%)	347 (5.9%)	7 (4.6%)		9,105 (46.9%)	8,976 (46.9%)	129 (44.9%)	
Breast/Chest wall + regional irradiation	1 (0.0%)	1 (0.0%)	0 (0.0%)		27 (0.1%)	27 (0.1%)	0 (0.0%)	
Partial breast irradiation	1 (0.0%)	1 (0.0%)	0 (0.0%)		157 (0.8%)	154 (0.8%)	3 (1.0%)	
Other (e.g. only regional lymph nodes)	0 (0.0%)	0 (0.0%)	0 (0.0%)		6 (0.0%)	6 (0.0%)	0 (0.0%)	

Not applicable	2,933 (48.9%)	2,863 (49.0%)	70 (46.4%)		9,004 (46.6%)	8,877 (46.4%)	127 (44.3%)	
Missing	2,705 (45.1%)	2,631 (45.0%)	74 (49.0%)		1,107 (5.7%)	1,079 (5.6%)	28 (9.8%)	
Later malignancies				<0.001				<0.001
No	4,207 (70.2%)	4,207 (72.0%)	0 (0.0%)		16,176 (83.4%)	16,176 (84.6%)	0 (0.0%)	
Yes	1,786 (29.8%)	1,635 (28.0%)	151 (100.0%)		3,230 (16.6%)	2,943 (15.4%)	287 (100.0%)	

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; HR, hormone receptor.
HR status was defined as HR-positive (HR+) when either estrogen receptor (ER) or progesterone receptor was (PR) positive, and HR-negative (HR-) when both ER and PR were negative, based on ER, PR, and human epidermal growth factor receptor 2 (HER2) status.

1

3.1.2 Lung cancer cohort

Among the lung cancer cohort, women with a history of breast cancer were significantly older at lung cancer diagnosis, compared to women without a history of breast cancer, with women who had a previous breast cancer diagnosis being more often 60-74 years old (56.9% vs. 52.2%), or 75+ years old (29.7% vs. 25.5%). In addition, a higher proportion of women with breast cancer history were diagnosed with stage I lung cancer (29.7% vs. 25.5%), and stage IV lung cancer was less common in this group (44.3% vs. 51.5%). Tumor laterality analysis showed 53.2% of patients with SPLC having tumors on the same side as their breast cancer (ipsilateral), compared to 43.5% with tumors on the opposite side (contralateral), though no significant difference was observed in laterality between the two groups ($p = 0.059$).

Table 3. Baseline characteristics of the lung cancer cohort, stratified by lung cancer being either primary lung cancer or SPLC after breast cancer. Data are presented as median (IQR) for continuous measures, and No. (%) for categorical measures.

	Total	**Primary lung cancer**	**Second primary lung cancer**	***p* Value**
	N=71,955	N=64,031	N=7,924	
Age at lung cancer diagnosis				<0.001
<59 years	16,028 (22.3%)	14,971 (23.4%)	1,057 (13.3%)	
60-74 years	37,567 (52.2%)	33,057 (51.6%)	4,510 (56.9%)	
75+ years	18,360 (25.5%)	16,003 (25.0%)	2,357 (29.7%)	
Socioeconomic status (SES)				<0.001
Low income (< €24,300)	30,219 (42.0%)	27,080 (42.3%)	3,139 (39.6%)	
Middle income (€24,300 - €31,000)	24,247 (33.7%)	21,635 (33.8%)	2,612 (33.0%)	
High income (> €31,000)	16,864 (23.4%)	14,762 (23.1%)	2,102 (26.5%)	
Missing	625 (0.9%)	554 (0.9%)	71 (0.9%)	
Interval between primary and current malignancy (years)	10.9 (4.9-18.6)	NA*	10.9 (4.9-18.6)	
Categorical interval between primary to current malignancy				<0.001
No previous cancer	64,031 (89.0%)	64,031 (100.0%)	0 (0.0%)	
<1 years	905 (1.3%)	0 (0.0%)	905 (11.4%)	
2-4 years	890 (1.2%)	0 (0.0%)	890 (11.2%)	
5-9 years	1,648 (2.3%)	0 (0.0%)	1,648 (20.8%)	
10-14 years	1,424 (2.0%)	0 (0.0%)	1,424 (18.0%)	
15-19 years	1,129 (1.6%)	0 (0.0%)	1,129 (14.2%)	
20-24 years	805 (1.1%)	0 (0.0%)	805 (10.2%)	
25-29 years	503 (0.7%)	0 (0.0%)	503 (6.3%)	
30-34 years	242 (0.3%)	0 (0.0%)	242 (3.1%)	
35-39 years	86 (0.1%)	0 (0.0%)	86 (1.1%)	
40-44 years	39 (0.1%)	0 (0.0%)	39 (0.5%)	
45-49 years	12 (0.0%)	0 (0.0%)	12 (0.2%)	
50-54 years	5 (0.0%)	0 (0.0%)	5 (0.1%)	
Behavior of previous malignancy				<0.001
In situ	791 (1.1%)	0 (0.0%)	791 (10.0%)	
Malignant	7,131 (9.9%)	0 (0.0%)	7,131 (90.0%)	

Metastatic	2 (0.0%)	0 (0.0%)	2 (0.0%)	
Not applicable	64,031 (89.0%)	64,031 (100.0%)	0 (0.0%)	
cTNM-stage				<0.001
Stage 0	82 (0.1%)	66 (0.1%)	16 (0.2%)	
Stage I	13,923 (19.3%)	11,871 (18.5%)	2,052 (25.9%)	
Stage II	4,772 (6.6%)	4,179 (6.5%)	593 (7.5%)	
Stage III	14,625 (20.3%)	13,057 (20.4%)	1,568 (19.8%)	
Stage IV	37,066 (51.5%)	33,553 (52.4%)	3,513 (44.3%)	
Missing	1,487 (2.1%)	1,305 (2.0%)	182 (2.3%)	
Laterality lung tumor				0.059
Left side	28,883 (40.1%)	25,606 (40.0%)	3,277 (41.4%)	
Right side	41,253 (57.3%)	36,811 (57.5%)	4,442 (56.1%)	
Bilateral	73 (0.1%)	68 (0.1%)	5 (0.1%)	
Missing	1,746 (2.4%)	1,546 (2.4%)	200 (2.5%)	
Laterality lung and breast tumor				****
Same side	4213 (5.9%)	0 (0.0%)	4213 (53.2%)	
Opposite side	3445 (4.8%)	0 (0.0%)	3445 (43.5%)	
Missing	266 (0.0%)	0 (0.0%)	266 (3.4%)	
Not Applicable	64,031 (89.0%)	64,031 (100.0%)	0 (0.0%)	
Morphology category				<0.001
Squamous Cell Carcinoma	9,224 (12.8%)	8,057 (12.6%)	1,167 (14.7%)	
Adenocarcinoma (incl. adenosquamous)	32,886 (45.7%)	29,198 (45.6%)	3,688 (46.5%)	
Large Cell Carcinoma, Not Otherwise Specified	6,366 (8.8%)	5,734 (9.0%)	632 (8.0%)	
No Pathological Confirmation	10,125 (14.1%)	9,050 (14.1%)	1,075 (13.6%)	
Other	13,345 (18.5%)	11,984 (18.7%)	1,361 (17.2%)	
Missing	9 (0.0%)	8 (0.0%)	1 (0.0%)	
Sublocalization				<0.001
Main Bronchus	5,640 (7.8%)	5,136 (8.0%)	504 (6.4%)	
Upper Lobe of Lung	37,134 (51.6%)	32,951 (51.5%)	4,183 (52.8%)	
Middle Lobe of Lung	3,166 (4.4%)	2,725 (4.3%)	441 (5.6%)	
Lower Lobe of Lung	19,452 (27.0%)	17,362 (27.1%)	2,090 (26.4%)	

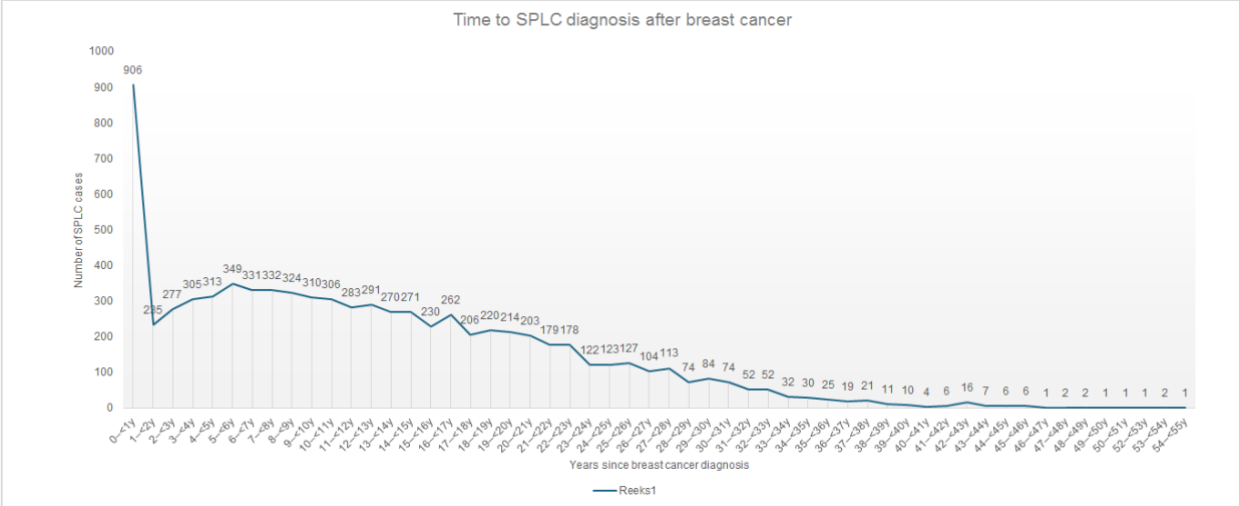
Overlapping Lesions of Lung	2,090 (2.9%)	1,874 (2.9%)	216 (2.7%)	
Lung, Not Otherwise Specified	4,473 (6.2%)	3,983 (6.2%)	490 (6.2%)	
Tumor multifocality				<0.001
No	47,442 (65.9%)	41,999 (65.6%)	5,443 (68.7%)	
Yes	13,069 (18.2%)	11,755 (18.4%)	1,314 (16.6%)	
Missing	11,444 (15.9%)	10,277 (16.1%)	1,167 (14.7%)	
Mutationtype **				
	N=7,580	N=6,939	N=641	
				<0.001
EGFR Mutation	3,108 (41.0%)	2,851 (41.1%)	257 (40.1%)	
KRAS Mutation	3,217 (42.4%)	2,929 (42.2%)	288 (44.9%)	
ALK Mutation	176 (2.3%)	169 (2.4%)	7 (1.1%)	
Other Mutation(s)	1,079 (14.2%)	990 (14.3%)	89 (13.9%)	
PDL-1 Expression ***				
	N=17,620	N=15,824	N=1,796	
PDL-1 (%)	33.5 (37.9)	33.8 (38.0)	31.3 (37.7)	0.048

Abbreviations: SPLC, second primary lung cancer.

*: All values indicating unknown or missing were excluded from the analysis. Continuous variables are presented as medians with interquartile ranges (IQR).
**: Cohort consists of all patients with stage IV and non-squamous tumors from 2017 onwards with a documented oncogenic mutation. Patients without a detected mutation, unknown status, or untested were excluded.
***: Cohort consists of all patients with stage IV tumors from 2018 onwards.
****: p-value not applicable, variable only relevant for patients with SPLC.

The time between breast cancer diagnosis and SPLC diagnosis ranged from 0 to 55 years (Figure 1). A total of 7,924 SPLC cases were included in this analysis. Most SPLC cases occurred within the first 10-15 years after breast cancer, with a peak in the first year (n = 906), followed by a decline in incidence in the second year and then a rise reaching a second peak around year 6. After year 6, the number of SPLC diagnoses gradually declines over time, with a small increase again around year 17.

Figure 1. Distribution of SPLC diagnoses by one-year intervals since initial breast cancer diagnosis.



3.2 Predictors of second primary lung cancer

3.2.1 Univariate logistic regression

Univariate logistic regression analyses were performed to examine the associations between patient, tumor and treatment characteristics, and the risk of second primary lung cancer (SPLC). Significant associations are described below. Full results are provided in Tables 4 and 5.

In the 2005-2009 DCIS cohort (Table 4), patients aged 60-74 had increased odds of SPLC (OR: 1.92, 95% CI: 1.14-3.25) compared to those under 50 years. For patients aged 50-59 and 75+ no significant association showed. Patients with lobular/mixed (OR: 2.45, 95% CI: 1.05-5.67) or other carcinomas (OR: 2.50, 95% CI: 1.00-6.28) had higher odds of SPLC compared to those with ductal carcinoma.

In the 2010-2020 DCIS cohort (Table 4), patients aged 50-59 (OR: 1.68, 95% CI: 1.04-2.71) and 60-74 years (OR: 2.78, 95% CI: 1.76-4.39) had higher odds of SPLC compared to those under 50. Patients with middle (OR: 0.59, 95% CI: 0.44-0.78) and high (OR: 0.46, 95% CI: 0.35-0.62) SES had a lower odds compared to those with low SES.

In the 2005-2009 invasive cohort (Table 5), patients aged 50-59 and 60-74 years had increased odds of SPLC compared to those under 50 (OR: 2.35, 95% CI: 1.98-2.79; OR: 2.21, 95% CI: 1.87-2.62 respectively). For patients aged 75+ no significant association showed. Patients with stage II (stage II: OR: 0.84, 95% CI: 0.75-0.95) and stage III tumors (OR: 0.56, 95% CI: 0.45-0.68) had lower odds of SPLC compared to those with stage I tumors. Similarly, patients with grade 2 (OR: 0.84, 95% CI: 0.73-0.95) and grade 3-4 tumors (OR: 0.74, 95% CI: 0.63-0.86) had lower odds compared to those with grade 1 tumors. Patients with multifocal tumors had decreased risk of SPLC (OR: 0.73, 95% CI: 0.62-0.87), as did patients with a tumor with a HR+/HER2+ receptor status compared to HR+/HER2- (OR: 0.82, 95% CI: 0.67-0.99). Patients who underwent non-breast conserving surgery had decreased odds of SPLC compared to patients who underwent breast conserving surgery (OR: 0.70, 95% CI: 0.62-0.78). Patients who received chemotherapy (OR: 0.76, 95% CI: 0.67-0.85) also had decreased odds of SPLC, while patients who received radiotherapy were linked to an increased risk of SPLC (OR: 1.29, 95% CI: 1.13-1.46).

In the 2010-2020 invasive cohort (Table 5), patients aged 50-59 (OR: 3.13, 95% CI: 2.56-3.84), 60-74 years (OR: 4.69, 95% CI: 3.88-5.68), and over 75 years (OR: 1.95, 95% CI: 2.04-3.32) had increased odds of SPLC compared to patients under 50. Patients with a middle (OR: 0.82, 95% CI: 0.69-0.87) and high

SES (OR: 0.57, 95% CI: 0.47-0.60) had lower odds of SPLC compared to those with a low SES. Patients with stage II (OR: 0.72, 95% CI: 0.67-0.83) and stage III tumors (OR: 0.54, 95% CI: 0.46-0.68) were associated with lower odds of SPLC, compared to those with stage I tumors. Patients with a differentiation grade 2 (OR: 0.82, 95% CI: 0.74-0.94) and grade 3-4 tumor (OR: 0.73, 95% CI: 0.63-0.83) had reduced odds compared to those with grade 1 tumors. Patients with right-sided tumors (OR: 0.89, 95% CI: 0.81-0.98) had slightly lower odds of SPLC compared to those with left-sided tumors. Multifocality was also associated with decreased odds (OR: 0.76, 95% CI: 0.65-0.86). Furthermore, patients who received chemotherapy (OR: 0.65, 95% CI: 0.54-0.67), targeted therapy (OR: 0.72, 95% CI: 0.56-0.82) or endocrine therapy (OR: 0.76, 95% CI: 0.71-0.87) were associated with lower odds of SPLC.

Table 4. Association between patient, tumor, and treatment characteristics and risk of SPLC in DCIS breast cancer patients diagnosed in 2005-2009 and 2010-2020.

	Univariate regression				Multivariate regression			
	2005-2009		2010-2020		2005-2009		2010-2020	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Age								
<50 years	1.00	reference	1.00	reference	1.00	reference	1.00	reference
50-59 years	1.66	0.98-2.84	1.68	1.04-2.71	1.62	0.95-2.78	1.75	1.08-2.83
60-74 years	1.92	1.14-3.25	2.78	1.76, 4.39	1.89	1.12-3.21	2.80	1.76-4.44
75+ years	0.92	0.34-2.50	1.28	0.57, 2.90	0.93	0.34-2.56	1.19	0.52-2.71
Socioeconomic status								
Low income (< €24,300)	-	-	1.00	reference	-	-	1.00	reference
Middle income (€24,300 - €31,000)	-	-	0.59	0.44-0.78	-	-	0.60	0.45-0.80
High income (> €31,000)	-	-	0.46	0.35-0.62	-	-	0.49	0.36-0.65
Differentiation grade								
Grade 1 (Well differentiated)	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Grade 2 (Moderately differentiated)	1.15	0.64-2.04	0.76	0.53-1.07	1.20	0.67-2.17	0.66	0.45-0.96
Grade 3-4 (Poorly differentiated)	1.37	0.82-2.29	0.94	0.68-1.30	1.52	0.54-1.05	0.77	0.54-1.12
Laterality								
Left	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Right	0.74	0.54-1.04	1.07	0.85-1.36	0.76	0.54-1.05	1.07	0.85-1.36
Morphology								
Ductal carcinoma	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Lobular + Mixed	2.45	1.05-5.67	0.30	0.07-1.20	2.58	1.10-6.06	0.33	0.08-1.32
Other carcinomas	2.50	1.00-6.28	2.06	0.91-4.69	2.52	1.00-6.36	1.90	0.82-4.37
Tumor multifocality								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	1.00	0.20-1.69	0.89	0.52-1.45	1.01	0.58-1.73	0.88	0.53-1.48
Surgery type								
Breast-conserving surgery	1.00	reference	1.00	reference	1.00	reference	1.00	reference

Non-breast-conserving surgery	0.82	0.58-1.15	1.09	0.84-1.40	0.62	0.35-1.09	1.61	1.04-2.50
Radiotherapy								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	1.07	0.78-1.49	1.05	0.83-1.33	0.70	0.41-1.20	1.35	0.90-2.03

Abbreviations: SPLC, second primary lung cancer; CI, confidence interval; OR, odds ratio.

Table 5. Association between patient, tumor, and treatment characteristics and risk of SPLC in invasive breast cancer patients diagnosed in 2005-2009 and 2010-2020.

	Univariate regression				Multivariate regression			
	2005-2009		2010-2020		2005-2009		2010-2020	
	OR	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Age								
<50 years	1.00	reference	1.00	reference	1.00	reference	1.00	reference
50-59 years	2.35	1.98-2.79	3.13	2.56-3.84	2.22	1.86-2.64	2.99	2.43-3.67
60-74 years	2.21	1.87-2.62	4.69	3.88-5.68	1.94	1.62-2.34	4.11	3.37-5.02
75+ years	0.76	0.58-1.01	1.95	2.04-3.32	0.69	0.51-0.94	1.98	1.52-2.58
Socioeconomic status								
Low income (< €24,300)	-	-	1.00	reference	-	-	1.00	reference
Middle income (€24,300 - €31,000)	-	-	0.82	0.69-0.87	-	-	0.78	0.70-0.88
High income (> €31,000)	-	-	0.57	0.47-0.60	-	-	0.55	0.49-0.63
TNM-Stage								
Stage I	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Stage II	0.84	0.75-0.95	0.72	0.67-0.83	1.02	0.88-1.20	0.91	0.80-1.03
Stage III	0.56	0.45-0.68	0.54	0.46-0.68	0.71	0.54-0.93	0.76	0.60-0.97
Differentiation grade								
Grade 1 (Well differentiated)	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Grade 2 (Moderately differentiated)	0.84	0.73-0.97	0.82	0.74-0.94	0.92	0.80-1.07	0.92	0.81-1.06
Grade 3-4 (Poorly differentiated)	0.74	0.63-0.86	0.73	0.63-0.83	0.98	0.80-1.19	0.89	0.74-1.06
Laterality								
Left	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Right	1.09	0.98-1.23	0.89	0.81-0.98	1.10	0.98-1.23	0.89	0.81-0.99
Morphology								
Ductal carcinoma	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Lobular + Mixed	0.97	0.82-1.14	1.01	0.90-1.19	0.96	0.81-1.13	1.01	0.87-1.17
Other carcinomas	1.06	0.84-1.34	1.02	0.87-1.36	1.07	0.84-1.35	0.99	0.79-1.24

Tumor multifocality								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	0.73	0.62-0.87	0.76	0.65-0.86	0.83	0.69-0.99	0.86	0.73-1.00
Receptor status								
HR+/HER2-	1.00	reference	1.00	reference	1.00	reference	1.00	reference
HR+/HER2+	0.84	0.68-1.04	0.69	0.54-0.82	1.09	0.82-1.44	0.82	0.59-1.13
HR-/HER2+	0.49	0.35-0.69	0.88	0.67-1.12	0.71	0.46-1.08	1.07	0.72-1.57
HR-/HER2-	0.82	0.67-0.99	1.03	0.86-1.17	0.95	0.73-1.23	1.33	1.07-1.64
Surgery type								
Breast-conserving surgery	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Non-breast-conserving surgery	0.70	0.62-0.78	0.90	0.81-1.00	1.03	0.82-1.29	1.05	0.88-1.26
Chemotherapy								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	0.76	0.67-0.85	0.65	0.54-0.67	0.88	0.74-1.05	0.86	0.73-1.00
Target therapy								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	0.57	0.44-0.73	0.72	0.56-0.82	0.71	0.49-1.03	1.12	0.79-1.60
Radiotherapy								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	1.29	1.13-1.46	1.02	0.77-0.97	1.18	0.94-1.49	0.82	0.69-0.98
Endocrine therapy								
No	1.00	reference	1.00	reference	1.00	reference	1.00	reference
Yes	0.94	0.84-1.06	0.76	0.71-0.87	1.09	0.92-1.30	1.01	0.88-1.16

Abbreviations: SPLC, second primary lung cancer; CI, confidence interval; OR, odds ratio; TNM, cancer, lymph node, metastasis; HR, hormone receptor.
HR status was defined as HR-positive (HR+) when either estrogen receptor (ER) or progesterone receptor was (PR) positive, and HR-negative (HR-) when both ER and PR were negative, based on ER, PR, and human epidermal growth factor receptor 2 (HER2) status.

3.2.2 Multivariate logistic regression

To assess the adjusted associations between patient, tumor, and treatment characteristics and the risk of SPLC, multivariable logistic regression analyses were performed for each cohort. Significant results on associations are reported below. Full results are provided in Tables 4 and 5.

In the 2005-2009 DCIS cohort (Table 4), patients aged 60-74 years had significantly higher odds of developing SPLC compared to those under 50 years (OR: 1.89, 95% CI: 1.12-3.21). Patients with lobular or mixed tumor morphology (OR: 2.58, 95% CI: 1.10-6.06) and those with other carcinoma types (OR: 2.52, 95% CI: 1.00-6.36) also had significantly higher odds of SPLC, compared to patients with ductal carcinoma.

In the 2010-2020 DCIS cohort (Table 4), patients aged 50-59 (OR: 1.75, 95% CI: 1.08-2.83) and 60-74 (OR: 2.80, 95% CI: 1.76-4.44) had increased odds of SPLC compared to those under 50. Patients with middle and high SES had significantly lower odds of SPLC (middle: OR: 0.60, 95% CI: 0.45-0.80; high: OR: 0.49, 95% CI: 0.36-0.65 respectively) compared to low SES. Furthermore, patients with a tumor with differentiation grade 2 was associated with decreased odds (OR: 0.66, 95% CI: 0.45-0.96) compared to those with a differentiation grade of 1. Patients who received non-breast-conserving surgery had an increased risk of SPLC (OR: 1.61, 95% CI: 1.04-2.50).

In the 2005-2009 invasive cohort (Table 5), patients aged 50-59 (OR: 2.22, 95% CI: 1.86-2.64) and 60-74 (OR: 1.94, 95% CI: 1.62-2.34) were significantly associated with increased risk of SPLC compared to those under 50, whereas patients aged 75+ had a lower risk of SPLC (OR: 0.69, 95% CI: 0.51-0.94). Patients with stage III tumors (OR: 0.71, 95% CI: 0.54-0.93) had lower odds of SPLC compared to those with stage I tumor. Similarly, patients with multifocal tumors (OR: 0.83, 95% CI: 0.69-0.99) had lower odds of SPLC.

In the 2010-2020 invasive cohort (Table 5), patients aged 60-74 had higher odds of developing SPLC (OR: 4.11, 95% CI: 3.37-5.02) compared to those under 50. Patients with a middle (OR: 0.78, 95% CI: 0.70-0.88) and high SES (OR: 0.55, 95% CI: 0.49-0.63) had reduced odds of SPLC compared to patients with a low SES. Patients with HR-/HER2- receptor status had higher odds of SPLC (OR: 1.33, 95% CI: 1.07-1.64) compared to patients with a HR+/HER2- receptor status. Patients with stage III tumors had a decreased risk of SPLC (OR: 0.76, 95% CI: 0.60-0.97) compared to those with stage I tumors, as well as patients with multifocal tumors (OR: 0.86, 95% CI: 0.73-1.00). For treatments, radiotherapy and chemotherapy were associated with reduced odds of SPLC (OR: 0.82, 95% CI: 0.69-0.98; OR: 0.86, 95% CI: 0.73-1.00 respectively).

3.3 Cox Regression: risk of SPLC

Cox proportional hazard models were used to examine the time to SPLC diagnosis after breast cancer. Below, we summarize whether the results from the Cox regression are consistent with the multivariate logistic regression, and newly identified or no longer significant associations will be highlighted. Adjusted hazard ratios (HRs) with 95% confidence intervals (CIs) are provided in Tables 5 and 6.

In the 2005-2009 DCIS cohort, results were largely consistent with those from the multivariate logistic regression (Table 5). Only the association between other carcinoma types and the risk of SPLC was no longer statistically significant.

In the 2010-2020 DCIS cohort, findings were consistent with those from the multivariate logistic regression (Table 5).

In the 2005-2009 invasive cohort, findings were consistent with those from the multivariate logistic regression (Table 6).

In the 2010-2020 invasive cohort, results were largely consistent with those from the multivariate logistic regression (Table 6). A new found association is on age, where patients aged 60-74 (HR: 4.24, 95% CI:

3.48-5.18) and those over 75 years (HR: 2.07, 95% CI: 1.59-2.70) have a higher hazard of SPLC compared to those younger than 50.

Table 5. Hazard ratios and 95% confidence intervals for the association between patient, tumor, and treatment characteristics and risk of SPLC in the 2005-2009 and 2010-2020 DCIS breast cancer cohort.

	2005-2009			**2010-2020**		
	HR	**95% CI**	***p* Value**	**HR**	**95% CI**	***p* Value**
Age						
<50 years	1.00	reference	-	1.00	reference	-
50-59 years	1.63	0.96-2.78	0.070	1.74	1.07-2.80	0.024
60-74 years	2.02	1.20-3.41	0.008	2.85	1.80-4.51	0.000
75+ years	1.02	0.38-2.78	0.962	1.28	0.56-2.89	0.559
Socioeconomic status						
Low income (< €24,300)	-	-	-	1.00	reference	-
Middle income (€24,300 - €31,000)	-	-	-	0.60	0.45-0.80	0.000
High income (> €31,000)	-	-	-	0.49	0.37-0.66	0.000
Differentiation Grade						
Grade 1 (Well differentiated)	1.00	reference	-	1.00	reference	-
Grade 2 (Moderately differentiated)	1.19	0.67-2.13	0.555	0.66	0.45-0.95	0.027
Grade 3-4 (Poorly differentiated)	1.49	0.87-2.54	0.143	0.78	0.54-1.12	0.175
Laterality						
Left	1.00	reference	-	1.00	reference	-
Right	0.77	0.56-1.07	0.125	1.07	0.85-1.35	0.571
Morphology						
Ductal carcinoma	1.00	reference	-	1.00	reference	-
Lobular + Mixed	2.62	1.14-6.00	0.023	0.34	0.08-1.35	0.124
Other carcinomas	2.35	0.96-5.76	0.062	1.86	0.82-4.23	0.137
Tumor multifocality						
No	1.00	reference	-	1.00	reference	-
Yes	1.03	0.60-1.75	0.924	0.89	0.53-1.48	0.643
Surgery type						
Breast-conserving surgery	1.00	reference	-	1.00	reference	-
Non-breast-conserving surgery	0.60	0.34-1.05	0.074	1.58	1.02-2.44	0.040
Radiotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.70	0.41-1.18	0.183	1.33	-	0.169

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; CI, confidence interval; HR, hazard ratio.

Table 6. Hazard ratios and 95% confidence intervals for the association between patient, tumor, and treatment characteristics and risk of SPLC in the 2005-2009 and 2010-2020 invasive breast cancer cohort.

	2005-2009			**2010-2020**		
	HR	**95% CI**	***p* Value**	**HR**	**95% CI**	***p* Value**
Age						
<50 years	1.00	reference	-	1.00	reference	-
50-59 years	2.26	1.90-2.69	0.000	3.01	2.45-3.69	0.000
60-74 years	2.07	1.73-2.49	0.000	4.24	3.48-5.18	0.000

75+ years	0.72	0.53-0.98	0.037	2.07	1.59-2.70	0.000
Socioeconomic status						
Low income (< €24,300)	-	-	-	-	-	-
Middle income (€24,300 - €31,000)	-	-	-	0.78	0.70-0.88	0.000
High income (> €31,000)	-	-	-	0.55	0.49-0.63	0.000
TNM-Stage						
Stage I	1.00	reference	-	1.00	reference	-
Stage II	1.01	0.87-1.18	0.895	0.93	0.79-1.03	0.116
Stage III	0.70	0.54-0.92	0.010	0.78	0.61-0.98	0.033
Differentiation grade						
Grade 1 (Well differentiated)	1.00	reference	-	1.00	reference	-
Grade 2 (Moderately differentiated)	0.92	0.79-1.07	0.27	0.92	0.81-1.06	0.241
Grade 3-4 (Poorly differentiated)	0.97	0.80-1.17	0.735	0.88	0.74-1.06	0.174
Laterality						
Left	1.00	reference	-	1.00	reference	-
Right	1.10	0.99-1.23	0.088	0.89	0.81-0.99	0.025
Morphology						
Ductal carcinoma	1.00	reference	-	1.00	reference	-
Lobular + Mixed	0.96	0.81-1.13	0.628	1.01	0.87-1.16	0.941
Other carcinomas	1.07	0.85-1.35	0.576	0.99	0.79-1.24	0.959
Tumor multifocality						
No	1.00	reference	-	1.00	reference	-
Yes	0.84	0.70-1.00	0.049	0.86	0.74-1.00	0.044
Receptor status						
HR+/HER2-	1.00	reference	-	1.00	reference	-
HR+/HER2+	1.06	0.81-1.39	0.682	0.81	0.59-1.12	0.203
HR-/HER2+	0.69	0.45-1.06	0.088	1.05	0.72-1.54	0.799
HR-/HER2-	0.93	0.72-1.21	0.604	1.31	1.06-1.62	0.011
Surgery type						
Breast-conserving surgery	1.00	reference	-	1.00	reference	-
Non-breast-conserving surgery	1.02	0.82-1.27	0.856	1.06	0.89-1.26	0.523
Chemotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.87	0.73-1.04	0.130	0.85	0.73-0.99	0.032
Target therapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.72	0.50-1.04	0.077	1.13	0.79-1.60	0.507
Radiotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	1.18	0.94-1.48	0.153	0.82	0.69-0.98	0.029
Endocrine therapy						
No	1.00	reference	-	1.00	reference	-
Yes	1.07	0.90-1.26	0.456	1.00	0.87-1.14	0.981

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; CI, confidence interval; HR, hazard ratio.

3.4 Cox regression: risk of death before SPLC

Cox proportional hazard models were used to examine the time to SPLC diagnosis after breast cancer. Below, we summarize whether the results from the Cox regression are consistent with the multivariate logistic regression, and newly identified or no longer significant associations will be highlighted. Adjusted hazard ratios (HRs) with 95% confidence intervals (CIs) are provided in Tables 7 and 8.

In the 2005-2009 DCIS cohort, patients aged 60-74 years old and 75+ had a higher risk of mortality before SPLC (HR: 6.04, 95% CI: 4.66-7.82; HR: 15.44, 95% CI: 11.71-20.35) compared to those under 50. Patients with lobular or mixed tumors were also associated with a higher risk of mortality before SPLC (HR: 1.54, 95% CI: 1.00-2.38) compared to patients with ductal carcinoma.

In the 2010-2020 DCIS cohort, patients aged 60-74 years old and 75+ had a higher hazard risk of mortality before SPLC (HR: 6.07, 95% CI: 4.51-8.16; HR: 22.86, 95% CI: 16.79-31.12). Higher SES status was had lower hazard ratios for both middle SES patients (HR: 0.78, 95% CI: 0.68-0.88) and high SES patients (HR: 0.56, 95% CI: 0.49-0.64), compared to patients with low SES.

In the 2005-2009 invasive cohort, patients aged 50-59, 60-74, and 75+ had a higher risk of mortality before SPLC (HR: 3.01, 95% CI: 2.45-3.69; HR: 4.24, 95% CI: 3.48-5.18; HR: 2.07, 95% CI: 1.59-2.70) compared to those under 50. Patients with higher tumor stage were also at higher risk of mortality before SPLC (stage III: HR: 1.79, 95% CI: 1.69-1.89) compared to patients with stage I tumors. Similarly, patients with a tumor differentiation grade of 2 and 3/4 were at higher risk of mortality before SPLC (HR: 1.10, 95% CI: 1.05-1.14; HR: 1.20, 95% CI: 1.14-1.26). Patients who received chemotherapy (HR: 0.89, 95% CI: 0.85-0.93), endocrine therapy (HR: 0.88, 95% CI: 0.84-0.91), and targeted therapy (HR: 0.77, 95% CI: 0.71-0.83) were associated with a lower risk of mortality before SPLC.

In the 2010-2020 invasive cohort, patients aged 50-59, 60-74, and 75+ again showed an increased risk (HR: 1.23, 95% CI: 1.17-1.30; HR: 2.36, 95% CI: 2.25-2.48; HR: 4.99, 95% CI: 4.72-5.28) compared to those under 50. Middle and high socioeconomic status were associated with a lower risk (HR: 0.88, 95% CI: 0.85-0.91 and HR: 0.75, 95% CI: 0.72-0.78, respectively). Patients with a higher tumor stage and differentiation grade were again associated with higher mortality risk before SPLC (stage III: HR: 2.69, 95% CI: 2.56-2.83; grade 2: HR: 1.14, 95% CI: 1.09-1.19; grade 3-4: HR: 1.46, 95% CI: 1.39-1.53) compared to lower stage and grade. Patients who received chemotherapy (HR: 0.95, 95% CI: 0.91-0.99), endocrine therapy (HR: 0.89, 95% CI: 0.86-0.93), and targeted therapy (HR: 0.53, 95% CI: 0.48-0.57) were again associated with reduced risk of mortality before SPLC. Patients with a lobular or mixed tumor morphology (HR: 0.93, 95% CI: 0.89-0.96) and multifocality (HR: 0.95, 95% CI: 0.92-0.99) were associated with a lower risk of death before SPLC compared to those with ductal carcinoma.

Table 7. Hazard ratios and 95% confidence intervals for the association between patient, tumor, and treatment characteristics and the risk of death before SPLC in the 2005-2009 and 2010-2020 DCIS breast cancer cohort.

	2005-2009			**2010-2020**		
	HR	**95% CI**	***p* Value**	**HR**	**95% CI**	***p* Value**
Age						
<50 years	1.00	reference	-	1.00	reference	-
50-59 years	1.39	1.04-1.86	0.025	1.55	1.12-2.14	0.008
60-74 years	6.04	4.66-7.82	0.000	6.07	4.51-8.16	0.000
75+ years	15.44	11.71-20.35	0.000	22.86	16.79-31.12	0.000
Socioeconomic status						
Low income (< €24,300)	-	-	-	1.00	reference	-
Middle income (€24,300 - €31,000)	-	-	-	0.78	0.68-0.88	0.000
High income (> €31,000)	-	-	-	0.56	0.49-0.64	0.000
Differentiation Grade						

Grade 1 (Well differentiated)	1.00	reference	-	1.00	reference	-
Grade 2 (Moderately differentiated)	0.99	0.83-1.19	0.953	0.84	0.71-1.00	0.054
Grade 3-4 (Poorly differentiated)	1.06	0.90-1.25	0.487	0.92	0.77-1.10	0.345
Laterality						
Left	1.00	reference	-	1.00	reference	-
Right	1.03	0.92-1.15	0.660	0.90	0.81-1.01	0.070
Morphology						
Ductal carcinoma	1.00	reference	-	1.00	reference	-
Lobular + Mixed	1.54	1.00-2.38	0.050	0.85	0.54-1.34	0.487
Other carcinomas	0.83	0.50-1.39	0.484	0.90	0.52-1.56	0.715
Tumor multifocality						
No	1.00	reference	-	1.00	reference	-
Yes	0.96	0.79-1.17	0.698	0.91	0.72-1.15	0.437
Surgery type						
Breast-conserving surgery	1.00	reference	-	1.00	reference	-
Non-breast-conserving surgery	0.87	0.72-1.06	0.650	1.21	1.01-1.45	0.370
Radiotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.89	0.74-1.07	0.222	1.33	0.89-1.99	0.169

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; CI, confidence interval; HR, hazard ratio.

Table 8. Hazard ratios and 95% confidence intervals for the association between patient, tumor, and treatment characteristics and the risk of death before SPLC in the 2005-2009 and 2010-2020 invasive breast cancer cohort.

	2005-2009			**2010-2020**		
	HR	**95% CI**	***p* Value**	**HR**	**95% CI**	***p* Value**
Age						
<50 years	1.00	reference	-	1.00	reference	-
50-59 years	3.01	2.45-3.69	0.000	1.23	1.17-1.30	0.000
60-74 years	4.24	3.48-5.18	0.000	2.36	2.25-2.48	0.000
75+ years	2.07	1.59-2.70	0.000	4.99	4.72-5.28	0.000
Socioeconomic status						
Low income (< €24,300)	-	-	-	-	-	-
Middle income (€24,300 - €31,000)	-	-	-	0.88	0.85-0.91	0.000
High income (> €31,000)	-	-	-	0.75	0.72-0.78	0.000
TNM-Stage						
Stage I	1.00	reference	-	1.00	reference	-
Stage II	1.23	1.18-1.28	0.000	1.47	1.42-1.53	0.000
Stage III	1.79	1.69-1.89	0.000	2.69	2.56-2.83	0.000
Differentiation grade						
Grade 1 (Well differentiated)	1.00	reference	-	1.00	reference	-
Grade 2 (Moderately differentiated)	1.10	1.05-1.14	0.000	1.14	1.09-1.19	0.000
Grade 3-4 (Poorly differentiated)	1.20	1.14-1.26	0.000	1.46	1.39-1.53	0.000
Laterality						
Left	1.00	reference	-	1.00	reference	-
Right	1.01	0.98-1.04	0.626	0.97	0.94-0.99	0.016

Morphology						
Ductal carcinoma	1.00	reference	-	1.00	reference	-
Lobular + Mixed	1.04	1.00-1.08	0.056	0.93	0.89-0.96	0.000
Other carcinomas	1.00	0.95-1.06	0.900	1.02	0.96-1.08	0.500
Tumor multifocality						
No	1.00	reference	-	1.00	reference	-
Yes	1.00	0.96-1.04	0.834	0.95	0.92-0.99	0.012
Receptor status						
HR+/HER2-	1.00	reference	-	1.00	reference	-
HR+/HER2+	0.92	0.92-1.04	0.558	1.16	1.08-1.24	0.000
HR-/HER2+	0.95	0.87-1.03	0.193	1.10	1.01-1.20	0.038
HR-/HER2-	0.97	0.91-1.02	0.255	1.12	1.06-1.18	0.000
Surgery type						
Breast-conserving surgery	1.00	reference	-	1.00	reference	-
Non-breast-conserving surgery	1.14	1.08-1.19	0.000	1.27	1.22-1.32	0.000
Chemotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.89	0.85-0.93	0.000	0.95	0.91-0.99	0.015
Target therapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.77	0.71-0.83	0.000	0.53	0.48-0.57	0.000
Radiotherapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.99	0.95-1.04	0.738	0.84	0.80-0.87	0.000
Endocrine therapy						
No	1.00	reference	-	1.00	reference	-
Yes	0.88	0.84-0.91	0.000	0.89	0.86-0.93	0.000

Abbreviations: SPLC, second primary lung cancer; TNM, cancer, lymph node, metastasis; CI, confidence interval; HR, hazard ratio.

4 Discussion

This study investigated the association between breast cancer treatment characteristics and the development of SPLC in breast cancer survivors. Additionally, it compared tumor characteristics of primary lung cancer and SPLC. The study's key findings revealed a consistent association between an age of 60-74 years and lowest socioeconomic status, and a higher risk of SPLC. Patients with stage I and grade 1 breast tumors were also at increased risk. SPLCs were more frequently detected in the upper lobe and on the same side as the previous breast tumor.

In all breast cancer cohorts, patients aged 60-74 years consistently showed an increased risk of SPLC, across all analyses. This finding is consistent with previous studies linking older age as a risk factor for SPMs, possibly due to longer latency periods and cumulative exposure to risk factors (7, 30). However, the competing risk Cox model showed that older patients, particularly those 60 and older, were at higher risk of dying before SPLC. In patients aged 75+, SPLC incidence was comparatively lower, which may not reflect lower actual risk, but is more likely due to a higher probability of dying before SPLC develops. These findings show the importance of accounting for competing risks when evaluating SPLC incidence across age groups. Earlier research also suggests that higher smoking and environmental exposure rates among individuals with lower SES may contribute to increased SPLC (11-13). This aligns with our results that women with a low SES were at higher risk of SPLC. Additionally, the competing risk Cox model showed that a higher SES was linked to a lower risk of death before SPLC, particularly among women in the 2010-2020 cohorts. This suggests that SES influences the risk of developing SPLC, with lower SES being associated with both a higher SPLC risk and higher early mortality. Early mortality could have potentially led to underdiagnoses of SPLC.

Patients with stage I and grade 1 tumors were more likely to develop SPLC. This pattern was consistent across periods and the different statistical model approaches. An explanation for this may be due to survivor effect, meaning patients with less aggressive breast cancer live long enough to develop SPLC (16). This is supported by the competing risk analysis, which showed that patients in the invasive cohorts, with higher stage and grade tumors had an increased hazard of death before SPLC. This suggests that patients with more advanced breast cancer may have died before SPLC could be developed, which could have led to an underestimation of SPLC risk in these groups.

Radiotherapy was not associated with an increased risk of SPLC in our adjusted models. This does not align with previous findings, as previous studies suggest an increased SPLC risk from radiotherapy, especially when using older techniques or larger radiation fields (17, 20-23, 31, 32). In the 2010-2020 invasive cohort, a protective effect was observed, suggesting patients who underwent radiotherapy had lower odds of SPLC. A possible explanation for the observed protective effect is confounding by indication (33). In that case radiotherapy may have been used more frequently in healthier patients with early-stage disease, fewer comorbidities, or better overall health. These patients may have had a lower risk of developing SPLC regardless of treatment, which could make radiotherapy appear protective if differences are not fully adjusted for. This interpretation is supported by the results of the competing risk Cox model, which found a reduced risk of death among patients who received radiotherapy in the 2010-2020 cohort. However, we did not have detailed radiotherapy data, such as dose, field size and technique. This limited the possibility to evaluate whether specific treatment characteristics influence SPLC risk. Such information could have helped differentiate between older methods involving broader lung exposure and newer techniques designed to minimize radiation to healthy tissue. Most existing studies on the effect of radiotherapy on SPLC included patients treated before 2005, whereas patients in our study were more likely to be treated with more modern radiotherapy protocols. For example, Taylor et al. (2017) reported significantly lower lung doses in modern breast cancer radiotherapy (21). This may explain the protective association observed in our study, suggesting that improvements in radiotherapy techniques have reduced the long-term risk of secondary lung malignancies.

The roles of chemotherapy, endocrine therapy, and targeted therapy differed across cohorts. Chemotherapy showed a protective effect in the 2010-2020 invasive cohort, and also showed reduced risk of death before

SPLC. This may indicate that chemotherapy was more often administered to patients with better overall health or less advanced disease. It is in conflict with some previous studies, as these suggest that chemotherapeutic agents may contribute to the development of SPLC (24). This difference may reflect changes in treatment protocols, but could also be due to chance. Endocrine and targeted therapies were not significantly associated with SPLC risk. These findings suggest that these therapies do not appear to influence the long-term risk of developing SPLC (25-27, 34).

Looking at lung cancer characteristics, SPLC was more likely to be diagnosed at an earlier stage and more frequently located in the upper lobe compared to primary lung cancer. The earlier stage at diagnosis may reflect more intensive follow-up among breast cancer survivors, increasing the likelihood of detecting SPLCs at an asymptomatic and earlier stage. In line with this, an increased incidence of SPLC was observed during the first year after breast cancer diagnosis. This early peak is likely due to increased diagnostic procedures following breast cancer treatment and surveillance, which may lead to incidental detection of otherwise undiagnosed lung tumors (16). A small number of patients had two SPLC diagnoses within the first year, which contributed to this early peak. These findings are likely to reflect detection bias rather than a real increase in the biological risk shortly after breast cancer diagnosis. To account for this, time since breast cancer diagnosis was included as a covariate in the Cox regression analyses.

This study has several strengths. It uses large, nationwide, population-based datasets from the NCR, with complete follow-up information, which eliminates selection bias and enhances the generalizability of the findings to similar Western populations with comparable breast cancer management practices (35). In addition, the use of multiple imputation by chained equations reduced the potential bias of results due to missing data (36).

This study has several limitations. First, no data were available on smoking history, which is a major risk factors for lung cancer and may have resulted in residual confounding (37). Second, environmental exposures such as air pollution and occupational hazards were not recorded. These factors are known to contribute to lung cancer risk. Their inclusion would have allowed for a more accurate risk estimate and better differentiation between treatment-related and environmental risks (38). Third, data on SES were only available from 2010 onward, limiting our ability to adjust for SES in the earlier cohorts. As a result, differences in SPLC risk between cohorts may be partly caused by differences in socioeconomic status (38). Fourth, the length of follow-up for patients in our cohorts differs. Patients diagnosed earlier (2005-2009) had a longer potential follow-up period for SPLC detection compared to those diagnosed in later years (2010-2020). This variation has contributed to the higher SPLC incidence in the earlier cohort. Finally, although multiple imputation was used to address missing data (36), this approach assumes that data are missing at random (MAR), which cannot be definitively verified. In particular, SES data were missing completely at random (MCAR), which does not meet the MAR assumption required for valid imputation. However, as the proportion of missing SES values was only 1.0%, the potential impact on the results is likely minimal.

Based on these findings, two recommendations can be made. First, breast cancer survivors aged between 60 and 74 years, those with lower SES, and those with stage I or grade 1 tumors appear to be at higher risk for SPLC, and may benefit from individualized follow-up protocols. Second, it is recommended that clinicians remain aware of the potential long-term risks of second primary lung cancer (SPLC) when making treatment decisions for breast cancer patients. This includes considering our findings, such as the observed protective associations with radiotherapy and chemotherapy in the 2010–2020 cohort. While these associations may partly reflect patient selection, they underline the importance of evaluating both the immediate benefits and possible long-term consequences of breast cancer treatment (24, 39).

Future research could focus on developing predictive models for SPLC to identify breast cancer survivors at increased risk and inform screening or follow-up protocols (34, 40). This study already contributes to understanding the clinical presentation of SPLC by comparing tumor characteristics with those of primary lung cancer. Furthermore, a protective association for radiotherapy was observed in the invasive, 2010-2020 cohort. A longer follow-up of breast cancer survivors for at least 10 to 15 years is recommended, to assess whether changes in radiotherapy protocols in 2010 have contributed to a reduced SPLC risk. In

addition, future studies in countries with similar healthcare systems are needed to assess the generalizability of these findings. Repeating similar analysis in the coming years, as longer follow-up becomes available and treatment practices continue to evolve, may provide further insight into the long-term risk of SPLC.

5 Conclusions

This study shows that among breast cancer survivors, patients aged 60-74 years and those with lower SES are at increased risk of SPLC. Radiotherapy was not associated with an increased risk of SPLC in any cohort after adjustment, a protective association was observed in the earlier cohorts. These findings reflect statistical associations and should be interpreted as potential indicators of SPLC risk. The lack of data on smoking history, environmental exposures, and detailed radiotherapy characteristics, limits the ability to account for residual confounding. Differences in lung tumor characteristics between SPLC and primary lung cancer, such as stage and location, suggest differences in tumor biology. Finally, most SPLCs were diagnosed within the first 15 years after breast cancer diagnosis, with a peak in the first year.

6 Ethical Approval

The study was conducted following a formal application submitted to and approved by the privacy committee of the Netherlands Cancer Registry and the NABON BOOG Scientific Committee (24-00580). Ethical approval was granted in accordance with the relevant regulations.

7 References

1. WHO WHO. Breast cancer 2024 [updated March 13, 2024. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer.
2. Łukasiewicz S, Czeczelewski M, Forma A, Baj J, Sitarz R, Stanisławek A. Breast Cancer—Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies—An Updated Review. Cancers. 2021;13(17).
3. Schaapveld M, Visser O, Louwman MJ, de Vries EGE, Willemse PHB, Otter R, et al. Risk of New Primary Nonbreast Cancers After Breast Cancer Treatment: A Dutch Population-Based Study. Journal of Clinical Oncology. 2008;26(8):1239-46.
4. Hendriks LEL, Dingemans A-MC, De Ruysscher DKM, Aarts MJ, Barberio L, Cornelissen R, et al. Lung Cancer in the Netherlands. Journal of Thoracic Oncology. 2021;16(3):355-65.
5. Grantzau T, Overgaard J. Risk of second non-breast cancer after radiotherapy for breast cancer: A systematic review and meta-analysis of 762,468 patients. Radiotherapy and Oncology. 2015;114(1):56-65.
6. Berrington de Gonzalez A, Gilbert E, Curtis R, Inskip P, Kleinerman R, Morton L, et al. Second Solid Cancers After Radiation Therapy: A Systematic Review of the Epidemiologic Studies of the Radiation Dose-Response Relationship. International Journal of Radiation Oncology*Biology*Physics. 2013;86(2):224-33.
7. Prathap R, Kirubha S, Rajan AT, Manoharan S, Elumalai K. The increasing prevalence of cancer in the elderly: An investigation of epidemiological trends. Aging Medicine. 2024;7(4):516-27.
8. Schneider JL, Rowe JH, Garcia-de-Alba C, Kim CF, Sharpe AH, Haigis MC. The aging lung: Physiology, disease, and immunity. Cell. 2021;184(8):1990-2019.
9. Campisi J. Aging, Cellular Senescence, and Cancer. Annual Review of Physiology. 2013;75(1):685-705.
10. Huang Y-J, Huang T-W, Lin F-H, Chung C-H, Tsao C-H, Chien W-C. Radiation Therapy for Invasive Breast Cancer Increases the Risk of Second Primary Lung Cancer: A Nationwide Population-Based Cohort Analysis. Journal of Thoracic Oncology. 2017;12(5):782-90.
11. Sandström N, Johansson M, Jekunen A, Andersén H. Socioeconomic status and lifestyle patterns in the most common cancer types-community-based research. BMC Public Health. 2023;23(1).
12. Hitchman SC, Fong GT, Zanna MP, Thrasher JF, Chung-Hall J, Siahpush M. Socioeconomic status and smokers' number of smoking friends: Findings from the International Tobacco Control (ITC) Four Country Survey. Drug and Alcohol Dependence. 2014;143:158-66.
13. Hajat A, Hsia C, O'Neill MS. Socioeconomic Disparities and Air Pollution Exposure: a Global Review. Current Environmental Health Reports. 2015;2(4):440-50.
14. Liu J, Hu Z, Feng Y, Zeng S, Zhong M. Problems to affect long-term survival for breast cancer patients. Medicine. 2018;97(39).
15. Ford MB, Sigurdson AJ, Petrulis ES, Ng CS, Kemp B, Cooksley C, et al. Effects of smoking and radiotherapy on lung carcinoma in breast carcinoma survivors. Cancer. 2003;98(7):1457-64.
16. Wang R, Yin Z, Liu L, Gao W, Li W, Shu Y, et al. Second Primary Lung Cancer After Breast Cancer: A Population-Based Study of 6,269 Women. Frontiers in Oncology. 2018;8.
17. Lu J, Shen Z, Wang X, Lin Y, Han Z, Kang M. Impact of radiotherapy on secondary lung cancer risk and survival in elderly female breast cancer survivors. Translational Oncology. 2025;52.

18. Schonfeld SJ, Curtis RE, Anderson WF, Berrington de González A. The risk of a second primary lung cancer after a first invasive breast cancer according to estrogen receptor status. Cancer Causes & Control. 2012;23(10):1721-8.
19. Jang YJ, Kim SY, Jung HK, Kim H-R, Kim CH, Lee H-R, et al. Association of EGFR mutations in second primary lung cancer and HER2 expression in breast cancer survivors. Translational Cancer Research. 2021;10(12):5204-11.
20. Dracham CB, Shankar A, Madan R. Radiation induced secondary malignancies: a review article. Radiation Oncology Journal. 2018;36(2):85-94.
21. Taylor C, Correa C, Duane FK, Aznar MC, Anderson SJ, Bergh J, et al. Estimating the Risks of Breast Cancer Radiotherapy: Evidence From Modern Radiation Doses to the Lungs and Heart and From Previous Randomized Trials. Journal of Clinical Oncology. 2017;35(15):1641-9.
22. Wei J, Wang B, Wang H, Meng L, Zhao Q, Li X, et al. Radiation-Induced Normal Tissue Damage: Oxidative Stress and Epigenetic Mechanisms. Oxidative Medicine and Cellular Longevity. 2019;2019:1-11.
23. Wong L-Y, Kapula N, He H, Guenthart BA, Vitzthum LK, Horst K, et al. Risk of developing subsequent primary lung cancer after receiving radiation for breast cancer. JTCVS Open. 2023;16:919-28.
24. Reyna VF, Nelson WL, Han PK, Pignone MP. Decision making and cancer. American Psychologist. 2015;70(2):105-18.
25. Liu K-L, Chen J-S, Chen S-C, Chu P-H. Cardiovascular Toxicity of Molecular Targeted Therapy in Cancer Patients- A Double-Edged Sword. Acta Cardiologica Sinica. 2013;29(4):295-303.
26. Duma N, Santana-Davila R, Molina JR. Non–Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. Mayo Clinic Proceedings. 2019;94(8):1623-40.
27. Chevallier M, Borgeaud M, Addeo A, Friedlaender A. Oncogenic driver mutations in non-small cell lung cancer: Past, present and future. World Journal of Clinical Oncology. 2021;12(4):217-37.
28. Liu M, Fu Y, Zhu S, Hong Q, Huang W, Chen C, et al. Survival Analysis of Secondary Primary Lung Cancer After Breast Cancer Patients: Insights From a Retrospective Single-Center Study of Clinical Outcomes and Prognostic Indicators. Thoracic Cancer. 2025;16(7).
29. Regulation GDP. What is GDPR, the EU's new data protection law? [Available from: https://gdprinfo.eu/what-is-gdpr-the-eus-new-data-protection-law.
30. Grantzau T, Thomsen MS, Væth M, Overgaard J. Risk of second primary lung cancer in women after radiotherapy for breast cancer. Radiotherapy and Oncology. 2014;111(3):366-73.
31. O'Dwyer E, Halpenny DF, Ginsberg MS. Lung cancer screening in patients with previous malignancy: Is this cohort at increased risk for malignancy? European Radiology. 2020;31(1):458-67.
32. Mukesh M, Harris E, Jena R, Evans P, Coles C. Relationship between irradiated breast volume and late normal tissue complications: A systematic review. Radiotherapy and Oncology. 2012;104(1):1-10.
33. Joseph KS, Mehrabadi A, Lisonkova S. Confounding by Indication and Related Concepts. Current Epidemiology Reports. 2014;1(1):1-8.
34. Liu B, Zhou H, Tan L, Siu KTH, Guan X-Y. Exploring treatment options in cancer: tumor treatment strategies. Signal Transduction and Targeted Therapy. 2024;9(1).
35. Nederland IK. Netherlands Cancer Registry (NCR) [updated October 23, 2023. Available from: https://iknl.nl/en/NCR.
36. Heymans MW, Twisk JWR. Handling missing data in clinical research. Journal of Clinical Epidemiology. 2022;151:185-8.
37. Phua ZJ, MacInnis RJ, Jayasekara H. Cigarette smoking and risk of second primary cancer: a systematic review and meta-analysis. Cancer Epidemiology. 2022;78.

38. WHO WHO. Cancer 2025 [Available from: https://www.who.int/news-room/fact-sheets/detail/cancer#:~:text=Tobacco%20use%2C%20alcohol%20consumption%2C%20unhealthy,%2D%20and%20middle%2Dincome%20countries.
39. Workman P, Clarke Paul A, Al-Lazikani B. Personalized Medicine: Patient-Predictive Panel Power. Cancer Cell. 2012;21(4):455-8.
40. Klaassen-Dekker A, Drossaert CHC, Van Maaren MC, Van Leeuwen-Stok AE, Retel VP, Korevaar JC, et al. Personalized surveillance and aftercare for non-metastasized breast cancer: the NABOR study protocol of a multiple interrupted time series design. BMC Cancer. 2023;23(1).