



ELSEVIER

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Data describing the poor outcome associated with a breast cancer diagnosis in the post-weaning period



Hanne Lefrère^{a,b}, Giuseppe Floris^{c,d,e}, Marjanka K. Schmidt^{f,g},
 Patrick Neven^{a,e,h}, Ellen Warnerⁱ, Elyce Cardonick^j,
 Fedro Alessandro Peccatori^k, Sibylle Loibl^l, Charlotte Maggen^{a,h},
 Hanne De Mulder^h, Katarzyna J. Jerzakⁱ, Diether Lambrechts^{m,n},
 Liesbeth Lenaerts^a, Frédéric Amant^{a,b,h,o,*}

^a Department of Oncology, Laboratory of Gynecological Oncology, KU Leuven, Leuven, Belgium

^b Department of Gynecology, The Netherlands Cancer Institute–Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands

^c Department of Imaging and Pathology, Unit of Translational Cell & Tissue Research, KU Leuven, Leuven, Belgium

^d Department of Pathology, Unit of Translational Cell & Tissue Research, University Hospitals Leuven, Leuven, Belgium

^e Multidisciplinary Breast Centre, UZ-KU Leuven Cancer Institute (LKI), University Hospitals Leuven, Leuven, Belgium

^f Division of Molecular Pathology, The Netherlands Cancer Institute–Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands

^g Division of Physiological Research and Epidemiology, The Netherlands Cancer Institute–Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands

^h Department of Gynecology and Obstetrics, University Hospitals Leuven, Leuven, Belgium

ⁱ Division of Medical Oncology, Department of Medicine, Sunnybrook Odette Cancer Centre, University of Toronto, Toronto, Ontario, Canada

^j Department of Obstetrics and Gynecology, Cooper University Health Care, Camden, New Jersey, USA

^k Division of Gynecological Oncology, Department of Gynecology, IEO European Institute of Oncology IRCCS, Milan, Italy

^l Centre for Haematology and Oncology Bethanien, German Breast Group, Neu-Isenburg, Frankfurt, Hessen, Germany

^m Center for Cancer Biology, VIB, Leuven, Belgium

ⁿ Laboratory of Translational Genetics, Department of Human Genetics, KU Leuven, Herestraat 49, Leuven 3000, Belgium

^o Department of Gynecological Oncology, Amsterdam University Medical Centers, Amsterdam, the Netherlands

* Corresponding author at: Department of Oncology, Laboratory of Gynecological Oncology, KU Leuven, Herestraat 49, 3000 Leuven, Belgium.

E-mail address: frederic.amant@uzleuven.be (F. Amant).

<https://doi.org/10.1016/j.dib.2021.107354>

2352-3409/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

ARTICLE INFO

Article history:

Received 26 July 2021

Revised 27 August 2021

Accepted 31 August 2021

Available online 10 September 2021

Keywords:

Postpartum breast cancer

Involution

Post-weaning

Lactation

Metastasis

Prognosis

ABSTRACT

Postpartum breast cancer (PPBC) - which according to new data, can extend to 5–10 years after the birth - are estimated to represent 35–55% of all cases of breast cancer in women younger than 45 years. Increasing clinical evidence indicates that PPBC represents a high-risk form of breast cancer in young women with an approximately 2-fold increased risk for metastasis and death. Yet, the exact mechanisms that underlay this poor prognosis are incompletely understood and, hence, it is unknown why postpartum breast cancer has an enhanced risk for metastasis or how it should be effectively targeted for improved survival. This article is an accompanying resource of the original article entitled “Breast cancer diagnosed in the post-weaning period is indicative for a poor outcome” and present epidemiological data that compare standard prognostic parameters, first site of metastatic disease and survival and metastatic rates in young women with primary invasive breast cancer diagnosed within two years postpartum (PP-BC), in young women diagnosed during pregnancy (Pr-BC) and nulliparous women (NP-BC). Via an international collaboration of 13 centres participating in the International Network on Cancer, Infertility and Pregnancy (INCIP), retrospective data of 1180 patients with primary invasive breast cancer, aged 25–40 years and diagnosed between January 1995 and December 2017 were collected. In particular, tumour-, patient, and therapy-related characteristics were collected. Furthermore, patient files were reviewed thoroughly to assess, for each parity, if and for how long breastfeeding was given. For PP-BC patients, breastfeeding history was used to differentiate breast cancers identified during lactation (PP-BC_{DL}) from those diagnosed post-weaning (PP-BC_{PW}). Primary exposures were prior childbirth or no childbirth, time between most recent childbirth and breast cancer diagnosis, time between cessation of lactation and breast cancer diagnosis and time between breast cancer diagnosis and metastasis or death. Distribution of standard prognostic parameters and first site of distant metastasis among study groups was determined applying fisher's exact, chi-squared, One-Way ANOVA or Kruskal-Wallis tests or logistic regression models, where applicable. The risks for metastasis and death were assessed using Cox proportional hazards models. A subgroup analysis was performed in PP-BC_{PW} patients that never lactated (PP-BC_{PW/NL}), lactated ≤ 3 months (PP-BC_{PW/Lshort}) or lactated > 3 months (PP-BC_{PW/Llong}).

© 2021 The Author(s). Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Specifications Table

Subject	Health and medical sciences, Oncology
Specific subject area	Prognostic parameters, survival and metastasis in women with a primary invasive breast cancer diagnosed in a 2-year post-weaning period.
Type of data	Table Chart Graph
How data were acquired	Patient data were collected via retrospective clinical chart review of patients in databases of the University Hospitals Leuven and via the registry of the International Network on Cancer, Infertility and Pregnancy (INCIP). INCIP members systematically enter extensive obstetric, perinatal, oncological and general medical information, including info on breastfeeding, as well as follow-up data about any new patient with cancer during or shortly after pregnancy.
Data format	Raw Analyzed
Parameters for data collection	Pre-menopausal women with primary invasive breast cancer at age 25 to 40 years and diagnosed between January 1995 and December 2017 were included in this retrospective cohort study. Exclusion criteria were a diagnosis >2 years post-weaning, postmenopausal status, invasive cancer history, pregnancy lasting < 24 weeks and insufficient data on 2 or more parameters or a lack of follow-up (< 2 years).
Description of data collection	Patient data and medical records were obtained from two prospectively maintained databases: (1) the database from the University Hospitals Leuven, and (2) the registry from the International Network on Cancer, Infertility and Pregnancy (INCIP). Tumour-, patient- and therapy-related characteristics, as well as detailed information regarding lactation history, were reviewed and compiled in a new Excel dataset.
Data source location	Institution: University Hospitals Leuven (UZ Leuven), Department of Gynaecological Oncology City/Town/Region: Leuven Country: Belgium
Data accessibility	Access to raw clinical data can be requested from the host institution UZ Leuven, via e-mail frederic.amant@uzleuven.be . Requests will be assessed by UZ Leuven Ethics Committee Research UZ/KU Leuven in order to evaluate concordance with information in the patient consent forms.
Related research article	H. Lefrère, G. Floris, M.K. Schmidt, P. Neven, E. Warner, E. Cardonick, F.A. Peccatori, S. Loibl, C. Maggen, H. De Mulder, K.J. Jerzak, D. Lambrechts, L. Lenaerts, F. Amant, Breast cancer diagnosed in the post-weaning period is indicative for a poor outcome. <i>European Journal of Cancer</i> . https://doi.org/10.1016/j.ejca.2021.06.009 .

Value of the Data

- This is a unique and large retrospective dataset compiling extensive tumour-, patient-, and therapy-related characteristics, as well as detailed breastfeeding information, from an international cohort of postpartum, pregnant, and nulliparous breast cancer patients.
- Cancer researchers, epidemiologists, oncologists, and gynaecologists can benefit from these data.
- Our data would allow researchers to conduct further in-depth studies on the influence of a (recent) pregnancy or lactation behaviour on breast cancer prognosis.

1. Data Description

1.1. Patient data collection

We present supporting data to the research article “Breast cancer diagnosed in the post-weaning period is indicative for a poor outcome” by Lefrère et al. Using patient data from the Multidisciplinary Breast Centre (MBC) of the University Hospitals Leuven (UZ Leuven) and from

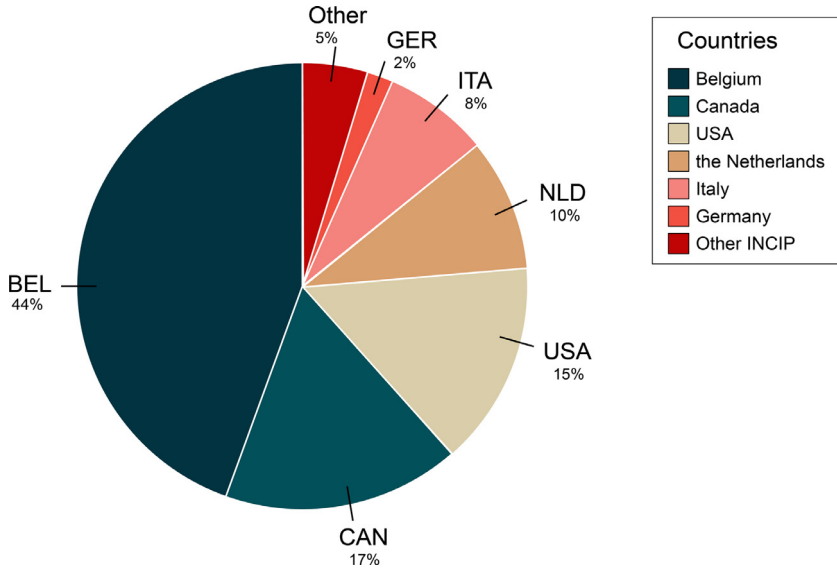


Fig. 1. Pie chart of all participating countries from which patient data was derived for this study. The largest groups of breast cancer patients were selected from the Multidisciplinary Breast Centre Database from the University Hospitals Leuven (Belgium, BEL). Additional patient data were retrieved from the registry coupled to the International Network on Cancer, Infertility and Pregnancy (INCIP, <http://www.cancerinpregnancy.org>). In particular, the division of cases from INCIP centres was as follows: 17% from the Sunnybrook Health Sciences Centre in Toronto (Canada, CAN), 15% from the Cooper University Hospital in New York (USA), 10% from the Netherlands Cancer Institute-Antoni Van Leeuwenhoek Hospital - NKI-AVL (the Netherlands, NLD), 8% from the European Institute of Oncology in Milan (Italy, ITA) and 2% from the German Breast Group in Berlin (Germany, GER), the remaining 5% of patient data were selected from other countries from the INCIP network, i.e. Austria, Czech Republic, Denmark, Russia, Spain, Swiss, Poland and Greece.

13 other international centres participating to the International Network on Cancer, Infertility and Pregnancy network (INCIP), a database was created consisting of clinicopathological data from 1180 eligible patients, aged ≤ 40 years and being diagnosed between 1995 and 2017 with primary invasive breast cancer. For each patient the following raw data were retrieved from the previously described databases: (1) tumour-related characteristics, including clinical stage, tumour size, lymph node (LN) infiltration, pN subtype (N0, N1, N2, N3), grade, histological type and surrogate molecular subtype; (2) patient-related characteristics (if applicable), including age at diagnosis, year of diagnosis, date of most recent delivery, gravidity, parity, number of miscarriages, lactation history, date of distant recurrence, site of distant recurrence and clinical outcome; (3) therapy-related characteristics, including surgery, radiotherapy (RT), chemotherapy (CT), hormonal therapy (HT) and/or anti-HER-2 treatment. In addition, patient files were thoroughly reviewed to retrieve data on whether or not breastfeeding was given and the duration of lactation. Fig. 1 shows a pie chart of all 1180 breast cancer patients enrolled in this study that were retrospectively collected from University Hospitals Leuven and 13 other countries participating in INCIP.

1.2. Clinicopathological characterisation

Based on the timing of their breast cancer diagnosis relative to their pregnancy history and lactation status at diagnosis, 189 women in the cohort of breast cancer patients were assigned to postpartum breast cancer diagnosed in the 2-year post-weaning period (PP-BC_{PW}), 53 to postpartum breast diagnosed during lactation (PP-BC_{DL}), 492 to pregnant breast cancer (Pr-BC) and 446 to nulliparous breast cancer (NP-BC). Table 2 assesses the differences in categorical

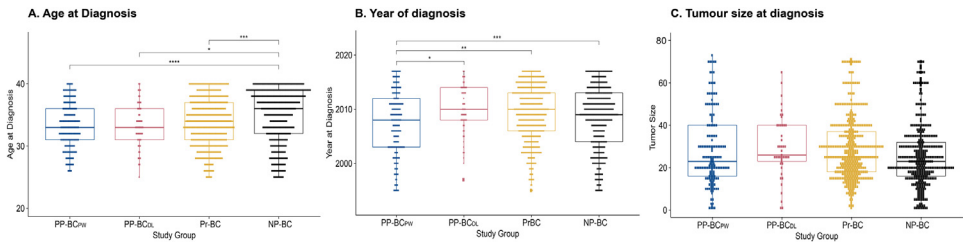


Fig. 2. Kruskal-Wallis and pairwise Tukey HSD testing in continuous parameters between study groups. A. Boxplot of age at diagnosis for all study groups. Overall Kruskal-Wallis analysis is $P < 0.001$. Tukey-HSD tests for pairwise comparisons indicate a significant result between PP-BC_{PW} versus NP-BC ($P < 0.001$). No significant difference is found between PP-BC_{PW} vs PP-BC_{DL} ($P = 0.949$) and PP-BC_{PW} vs Pr-BC ($P = 0.217$). B. Boxplot of year of diagnosis for all study groups. Overall Kruskal-Wallis analysis is $P = 0.012$. Tukey-HSD tests for pairwise comparisons indicate a significant result between PP-BC_{PW} vs PP-BC_{DL} ($P = 0.025$) and PP-BC_{PW} vs Pr-BC ($P = 0.026$) and between PP-BC_{PW} vs NP-BC ($P = 0.061$). C. Boxplot of tumour size at diagnosis for all study groups. Overall Kruskal-Wallis analysis is $P = 0.804$. Tukey-HSD tests for pairwise comparisons indicate no significant results between PP-BC_{PW} vs PP-BC_{DL} ($P = 0.761$), PP-BC_{PW} vs Pr-BC ($P = 0.761$) and PP-BC_{PW} vs NP-BC ($P = 0.187$).

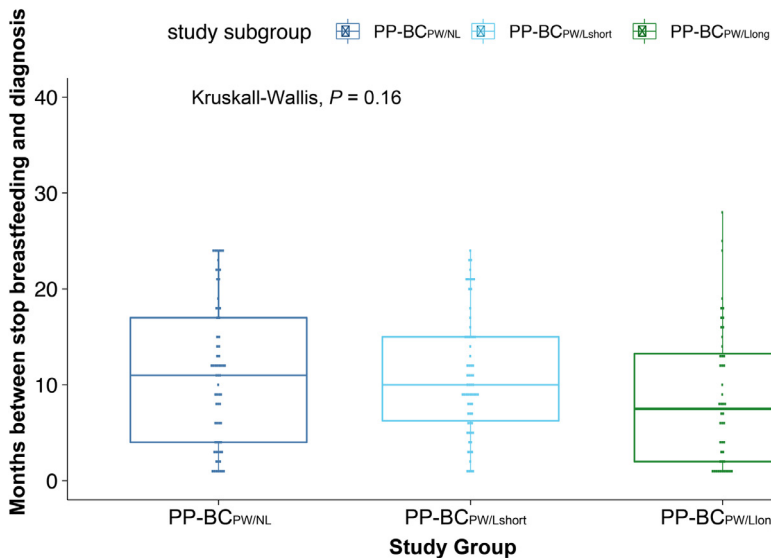


Fig. 3. Boxplot of time between cessation of lactation and cancer diagnosis (months). Overall Kruskal-Wallis analysis is $P = 0.160$. Tukey-HSD tests for pairwise comparisons indicate no significant results between PP-BC_{PW} vs PP-BC_{PW/Lshort} ($P = 0.993$) and PP-BC_{PW/Llong} ($P = 0.184$). Also, no significant differences could be observed between PP-BC_{PW/Lshort} and PP-BC_{PW/Llong} ($P = 0.222$).

tumour-, patient- and therapy-related characteristics between the study groups. Two-tailed Wald tests were used to determine the odds ratio (OR) and 95% confidence interval (CI) in a multinomial logistic regression model including all parameters listed in this table. Kruskal-Wallis and pairwise Tukey HSD analyses were performed to identify differences in continuous parameters, including age at diagnosis (Fig. 2A), year of diagnosis (Fig. 2B) and tumour size (Fig. 2C), between the study groups. In addition, Fig. 3 evaluates differences in time between cessation of lactation and cancer diagnosis in postpartum post-weaning patients that never lactated (PP-BC_{PW/NL}), lactated ≤ 3 months (PP-BC_{PW/Lshort}) or lactated > 3 months (PP-BC_{PW/Llong}) using Kruskal-Wallis and Tukey HSD tests.

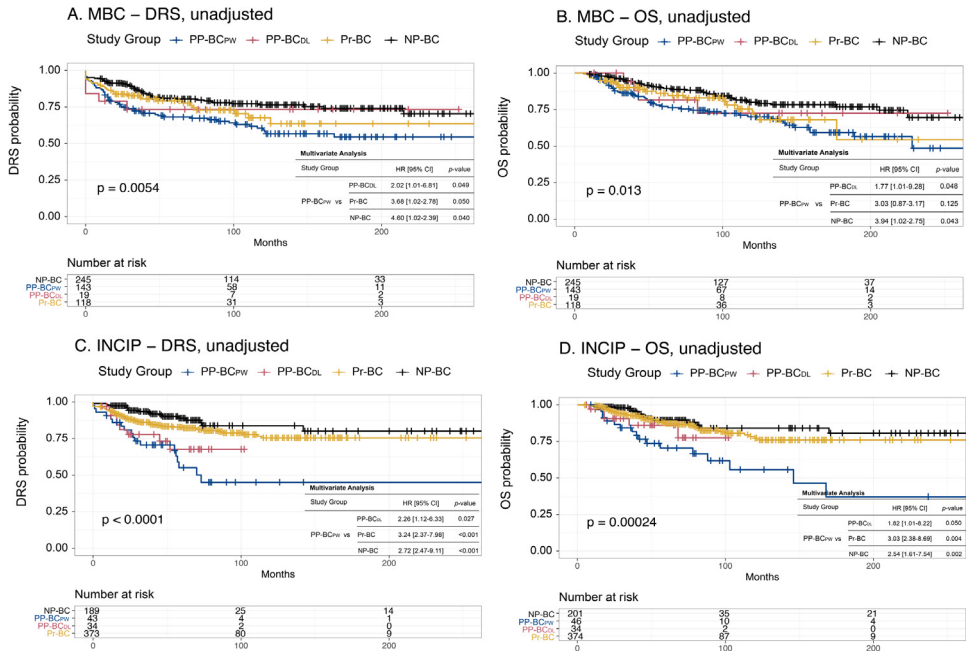


Fig. 4. Unadjusted and adjusted Cox regression analyses in University Hospitals Leuven patients (Belgium, Multidisciplinary Breast Centre, MBC) and INCIP patients. A and B. Risk for metastasis (A) and death (B) in MBC patients from University Hospitals Leuven ($n = 525$). C and D. Risk for metastasis (A) and death (B) in patients diagnosed in other INCIP centres ($n = 655$). Unadjusted analyses are shown in the figures whereas the results of the adjusted analyses are displayed in the inserted tables. Both in the cohort of patients diagnosed in University Hospitals Leuven (Belgium) as well in the cohort of patients diagnosed in one of the participating INCIP centres, OS and DRS probabilities in PP-BC_{PW} patients were statistically significantly differing from those observed in PP-BC_{DL}, Pr-BC and NP-BC cases. Only the observed difference in OS in PP-BC_{PW} versus Pr-BC patients in the cohort of University Hospitals Leuven patients did not reach statistical significance. Given the same trends, pointing to the poorest outcome parameters in PP-BC_{PW} patients, independent of the centre of diagnosis (i.e. when comparing A and C, and B and D), any influence of centre of diagnosis on prognostic differences can be ruled.

1.3. Cox proportional hazard models

Differences in overall survival (OS) and distant recurrence-free survival (DRS) among the study groups were assessed using both univariate and multivariate Cox proportional hazard models. The multivariate model was adjusted for prognostic host- and tumour-related variables that (i) significantly differed between patients groups (Table 1) and/or (ii) were significantly correlated to OS and/or DRS (Table 2).

Univariate and multivariate Cox proportional hazard models were drawn to assess the influence of study group or study subgroup on prognosis (results in original article). To evaluate the influence of centre of diagnosis on prognosis in our population, Kaplan-Meier and Cox regression analyses were performed on patients subdivided based on whether they were registered via the University Hospitals Leuven ($n = 525$) or via INCIP ($n = 655$) (Fig. 4). Unadjusted and adjusted Cox regression analyses were also performed in all patients stratified for study group (Fig. 5) or study subgroup (Fig. 6), excluding patients diagnosed with stage IV disease. Stage IV patients were excluded to investigate true population outcomes, as patients with metastatic disease at diagnosis are biased towards a poorer prognosis. Finally, the influence of a diagnosis post-weaning for patients that lactated ≤ 3 months on prognosis was specifically investigated in Table 3.

Table 1Comparison of the frequencies of host- and tumour-related prognostic parameters in PP-BC_{DL} versus the other patient groups.

Study group	PP-BC _{PW} (n = 189)		PP-BC _{DL} (n = 53)		Pr-BC (n = 492)		NP-BC (n = 446)		PP-BC _{DL} VS		
	n	%	n	%	n	%	n	%	PP-BC _{PW} OR [95% CI]*	Pr-BC OR [95% CI]*	NP-BC OR [95% CI]*
Mean age of diagnosis (SD)	33.3	SD 3.6	33.2	SD 3.6	33.7	SD 3.9	34.8	SD 4.3	1.0 [0.9–1.1]	1.0 [0.9–1.0]	0.9 [0.8–1.0]
Mean year at diagnosis (SD)	2007	SD 6.0	2009	SD 4.6	2009	SD 5.1	2008	SD 5.7	1.0 [1.0–1.1]	1.0 [1.0–1.1]	1.0 [1.0–1.1]
Stage											
Stage IA	39	21.0	4	7.5	70	14.3	103	23.3	–	–	–
Stage IB	6	3.2	1	1.9	6	1.2	28	6.3	1.6 [0.2–17.1]	2.9 [0.3–30.4]	0.9 [0.1–8.6]
Stage IIA	56	30.1	18	34.0	169	34.6	143	32.3	3.1 [0.9–9.9]	1.9 [0.6–5.7]	3.2 [1.1–9.9]
Stage IIB	28	15.1	12	22.6	111	22.7	75	16.9	4.2 [1.2 – 14.3]	1.9 [0.6 – 6.1]	4.1 [1.3 – 13.3]
Stage IIIA	26	14.0	4	7.5	70	14.3	48	10.8	1.5 [0.3 – 6.5]	1.0 [0.2 – 4.2]	2.2 [0.5 – 8.9]
Stage IIIB	11	5.9	6	11.3	17	3.5	18	4.1	5.3 [1.3 – 22.2]	6.2 [1.6 – 24.3]	8.6 [2.2 – 33.5]
Stage IIIC	13	7.0	4	7.5	28	5.7	14	3.2	3.0 [0.7 – 13.7]	2.5 [0.6 – 10.7]	7.4 [1.7 – 32.8]
Stage IV	7	3.8	4	7.5	18	3.7	14	3.2	5.6 [1.1 – 27.7]	3.9 [0.9 – 17.1]	7.4 [1.7 – 32.8]
Missing	3		0		3		3				
LN involvement											
Negative	80	43.5	21	39.6	217	44.4	226	51.1	–	–	–
Positive	104	56.5	32	60.4	273	55.6	215	48.9	1.1 [0.6–2.0]	1.2 [0.7–2.2]	1.6 [0.9–2.9]
Missing	5		0		3		4				
pN status											
N0	80	43.5	21	39.6	217	44.4	226	51.1	–	–	–
N1	79	42.9	25	47.2	192	39.3	154	34.8	1.2 [0.6–2.3]	1.4 [0.7–2.5]	1.8 [0.9–3.2]
N2	15	8.2	2	3.8	46	9.4	35	7.9	0.5 [0.1–2.4]	0.5 [0.1–2.0]	0.6 [0.1–2.7]
N3	10	5.4	5	9.4	34	7.0	27	6.1	1.9 [0.6–6.2]	1.5 [0.5–4.3]	2.0 [0.7–5.7]
Missing	5		0		3		4				
Mean tumour size (mm) (SD)	33.9	SD 26.6	32.8	SD 21.4	32.6	SD 24.0	30.1	SD 23.4	1.0 [0.9–1.0]	1.0 [0.9–1.0]	1.0 [0.9–1.0]
Grade											
Grade I	7	3.8	5	9.4	15	3.1	60	13.9	–	–	–
Grade II	49	26.5	9	17.0	111	23.1	161	37.3	0.3 [0.1–1.0]	0.4 [0.2–1.0]	0.7 [0.2–2.1]
Grade III	129	69.7	39	73.6	354	73.8	211	48.8	0.4 [0.1–1.4]	0.5 [0.1–1.2]	2.2 [0.8–5.9]
Missing	4		0		12		14				
Surrogate Molecular Subtype											
Luminal A-like	36	19.6	3	5.7	35	8.1	127	30.5	–	–	–
Luminal B-like	47	25.5	15	28.3	115	26.6	106	25.5	3.8 [1.0–14.2]	1.5 [0.4–5.6]	6.0 [1.7–21.3]
Luminal HER-2	30	16.3	10	18.9	77	17.8	70	16.8	4.0 [1.0–15.9]	1.5 [0.4–5.9]	6.1 [1.6–22.7]

(continued on next page)

Table 1 (continued)

Study group	PP-BC _{PW} (n = 189)		PP-BC _{DL} (n = 53)		Pr-BC (n = 492)		NP-BC (n = 446)		PP-BC _{DL} VS		
	n	%	n	%	n	%	n	%	PP-BC _{PW} OR [95% CI]*	Pr-BC OR [95% CI]*	NP-BC OR [95% CI]*
HER-2-like	19	10.3	10	18.9	52	12.0	25	6.0	6.3 [1.6–25.7]	2.2 [0.6–8.7]	16.9 [4.4–65.9]
Triple Negative	52	28.3	15	28.3	154	35.6	88	21.2	3.5 [0.9–12.8]	1.1 [0.3–4.1]	7.2 [2.0–25.7]
Missing	5		0		59		30				
Histological Subtype											
IDC	153	81.8	41	77.4	404	83.1	351	79.1	–	–	–
ILC	23	12.3	5	9.4	50	10.3	67	15.1	0.8 [0.3–2.3]	1.0 [0.4–2.6]	0.6 [0.2–1.7]
Other (Special)	11	5.9	7	13.2	32	6.6	26	5.9	2.4 [0.9–6.5]	2.2 [0.9–5.2]	2.3 [0.9–5.6]
Missing	2		0		6		2				
Chemotherapy											
No	29	15.4	5	9.4	27	5.5	92	20.8	–	–	–
Yes	159	84.6	48	90.6	463	94.4	350	79.2	1.8 [0.6–4.8]	0.6 [0.2–1.5]	1.8 [0.9–6.5]
Adj.	103	64.8	25	52.1	311	67.2	252	72.0	1.0 [0.3–3.1]	1.3 [0.4–4.1]	3.7 [1.1–12.4]
Neoadj.	39	24.5	19	39.6	115	24.8	87	24.9	2.0 [0.9–4.0]	2.0 [0.9–3.9]	2.2 [1.56–4.19]
(Neo)-Adj.	17	10.7	4	8.3	37	8.0	11	3.1	0.7 [0.3–2.0]	2.3 [0.8–6.5]	0.6 [0.2–1.5]
Missing	1		0		2		4				
Surgery											
No	10	5.3	4	7.5	19	3.9	12	2.7	–	–	–
Yes	178	94.7	49	92.5	470	96.1	430	97.3	0.7 [0.2–2.3]	0.0 [0.2–1.5]	0.3 [0.1–1.1]
Missing	1		0		3		4				
Radiotherapy											
No	34	18.2	17	32.1	170	36.0	119	26.9	–	–	–
Yes	153	81.8	36	67.9	302	64.0	323	73.1	0.5 [0.2–0.9]	1.2 [0.7–2.2]	0.8 [0.4–1.4]
Missing	2		0		20		4				
Hormone therapy											
No	97	51.6	35	66.0	275	56.9	172	39.0	–	–	–
Yes	91	48.4	18	34.0	208	43.1	269	61.0	0.6 [0.3–1.0]	0.7 [0.4–1.2]	0.3 [0.2–0.6]
Missing	1		0		9		5				
Anti HER-2 therapy											
No	158	83.6	38	71.7	379	78.5	369	83.7	–	–	–
Yes	31	16.4	15	28.3	104	21.5	72	16.3	2.0 [0.9–4.1]	1.4 [0.8–2.7]	2.0 [1.1–3.9]
Missing	0		0		9		5				

* Multinomial logistic regression model – 2-tailed Wald tests are used to determine Odds Ratio (OR) and 95% Confidence Intervals (95% CI). OR larger (lower) than 1 indicates increased (decreased) occurrence of that parameter in PP-BCDL compared to the reference type (either PP-BCPW, Pr-BC or NP-BC). Significant values are indicated in bold.

Table 2

Covariate risk estimate of host- and tumor-related prognostic variables and treatment modalities for risk for death (OS) and metastases (DRS) for all 1180 breast cancer patients.

Covariates	#	OS HR [95% CI] †	DRS HR [95% CI] †
Age of diagnosis (continuous)			
≥ 25 years and ≤ 40 years	1180	1.0 [1.0–1.1]	1.0 [1.0–1.1]
Year at diagnosis (continuous)			
1995 – 2017	1180	1.0 [1.0–1.1]	1.0 [1.0–1.1]
Stage			
Stage IA	216	–	–
Stage IB	41	2.6 [0.9–8.0]	1.9 [0.7–5.1]
Stage IIA	386	1.3 [0.7–2.4]	1.4 [0.8–2.3]
Stage IIB	226	3.0 [1.7–5.3]	2.8 [1.7–4.8]
Stage IIIA	148	4.2 [2.3–7.7]	3.8 [2.3–6.5]
Stage IIIB	52	8.4 [4.4–15.9]	7.2 [4.0–13.1]
Stage IIIC	59	7.9[4.2–14.9]	7.6 [4.3–13.5]
Stage IV	43	27.3 [14.7–50.6]	x
LN Involvement			
No	544	–	–
Yes	624	3.7 [2.7–5.2]	3.0 [2.2–4.0]
pN subtype			
N0	544	–	–
N1	450	2.7 [1.9–3.8]	2.2 [1.6–3.0]
N2	98	4.0 [2.5–3.4]	3.6 [2.4–5.3]
N3	76	7.0 [4.4–10.9]	5.8 [3.9–8.7]
Tumor size at diagnosis (continuous)			
mm	1054	1.0 [1.0–1.1]	1.0 [1.0–1.1]
Grade			
Grade I	87	–	–
Grade II	330	6.3 [1.5–25.7]	3.5 [1.4–8.6]
Grade III	733	7.4 [2.0–29.7]	4.0 [1.6–9.7]
Surrogate molecular subtype			
Luminal A-like	201	–	–
Luminal B-like	283	1.9 [1.2–3.0]	1.6 [1.0–2.3]
Luminal HER-2	187	1.2 [0.7–2.1]	1.2 [0.7–1.9]
HER-2-like	106	1.5 [0.8–2.7]	1.5 [0.9–2.5]
Triple Negative	309	2.0 [1.3–3.2]	1.5 [1.0–2.2]
Histological subtype			
IDC	949	–	–
ILC	145	1.1 [0.7–1.7]	1.1 [0.8–1.6]
Other	76	1.3 [0.8–2.0]	1.2 [0.7–1.9]
Chemotherapy			
No	153	–	–
Yes	1020	1.3 [0.9–2.0]	1.0 [0.7–1.4]
Adjuvant	691	1.0 [0.6–1.6]	0.8 [0.5–1.1]
Neoadjuvant	260	2.1 [1.1–3.8]	1.4 [0.8–2.4]
Adjuvant + Neoadjuvant	69	2.5 [1.6–4.1]	1.5 [0.9–2.2]
Surgery			
No	45	–	–
Yes	1127	0.1 [0.1–0.2]	0.1 [0.1–0.2]
Radiotherapy			
No	340	–	–
Yes	814	1.1 [0.8–1.5]	1.1 [0.9–1.5]
Hormone Therapy			
No	579	–	–
Yes	586	0.7 [0.5–0.9]	0.8 [0.6–0.9]
Anti HER-2 therapy			
No	944	–	–
Yes	222	0.5 [0.3–0.9]	0.6 [0.4–0.9]

† Univariate models to determine the influence of each parameter on OS and DRS – Hazard Ratio (HR) and 95% Confidence Interval (CI) are determined for each parameter. HR of more (less) than 1 indicates higher (lower) risk of death or metastasis. Significant values are in bold. As expected, more LN infiltration and/or positive LNs, larger tumours, higher grade and more hormone receptor-negativity were all associated with an increased risk for death and metastasis. Except for RT, all other therapy-related characteristics also seemed to influence either OS or DRS. More (neo-)adjuvant CT, less surgery, less HT and less anti-HER-2 therapy could be related to an increased risk for metastasis and/or death.

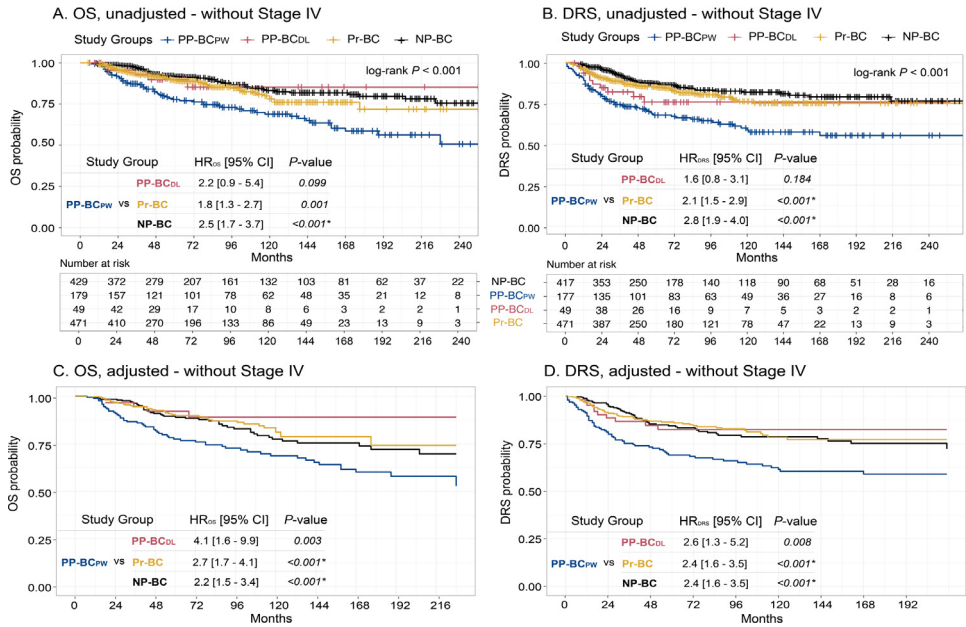


Fig. 5. Unadjusted and adjusted Cox regression analyses excluding stage IV cases in PP-BC_{pw} (n = 179), PP-BC_{DL} (n = 49), Pr-BC (n = 471) and NP-BC (n = 429) patients demonstrated an increased risk of death (A) and metastasis (B) in PP-BC_{pw}. Adjusted probability in PP-BC_{pw} (n = 173), PP-BC_{DL} (n = 49), Pr-BC (n = 394) and NP-BC (n = 394) patients also indicates an increased risk of death (C) and metastasis (D) in PP-BC_{pw}. Multivariate Proportional Hazards models were adjusted for age at diagnosis, year of diagnosis, stage, grade, surrogate molecular subtype, surgery, and (neo-)adjuvant CT, RT, HT and anti-HER-2 therapy. Grade, HT and surgery were stratified to comply with the proportional hazard's assumption. HR of more (less) than 1 indicates higher (lower) risk of death or metastasis. Significant values are indicated with an asterisk (*).

Table 3

Unadjusted and adjusted models of OS and DRS in relation to study group. The PP-BC_{pw} group was further subdivided in this analysis based on duration of breastfeeding prior to breast cancer diagnosis.

Study Group	OS		DRS			
	HR [95% CI]*	p-value	HR [95% CI]*	p-value		
Part I. Univariate Model						
PP-BC _{pw} /Lshort	VS	PP-BC _{pw} /NL	0.7 [0.4–1.2]	0.178	0.7 [0.4–1.2]	0.164
		PP-BC _{pw} /Llong	1.5 [0.7–3.1]	0.271	1.4 [0.7–2.8]	0.290
		PP-BC _{DL}	1.4 [0.6–2.9]	0.425	1.2 [1.6–2.3]	0.549
		Pr-BC	1.7 [1.1–2.7]	0.022	1.8 [1.2–2.8]	0.006
		NP-BC	2.3 [1.4–3.6]	0.001	2.3 [1.5–3.6]	<0.001
Part II. Multivariate Model						
PP-BC _{pw} /Lshort	VS	PP-BC _{pw} /NL	0.5 [0.3–0.8]	0.008	0.5 [0.3–1.8]	0.006
		PP-BC _{pw} /Llong	1.2 [0.6–2.6]	0.621	1.4 [0.7–2.7]	0.395
		PP-BC _{DL}	2.4 [1.1–5.4]	0.035	1.7 [1.0–3.3]	0.050
		Pr-BC	1.8 [1.0–2.9]	0.036	1.7 [1.0–2.7]	0.038
		NP-BC	1.7 [1.0–2.9]	0.044	1.7 [1.0–2.7]	0.035

Multivariate Proportional Hazards model for OS and DRS were adjusted for age at diagnosis, year of diagnosis, grade, stage (accounting for tumour size and LN infiltration), surrogate molecular subtype, surgery, and (neo-)adjuvant CT, RT, HT and anti-HER-2 therapy. Grade, hormone therapy and surgery were stratified to comply with the proportional hazard's assumption.

* Hazard Ratio (HR) and 95% Confidence Interval (CI) for the OS and DRS proportional hazards models were determined using Cox regression analyses and Kaplan-Meier curves. HR of more (less) than 1 indicates higher (lower) risk of death or metastasis.

* P-values of less than 0.05 were considered significant.

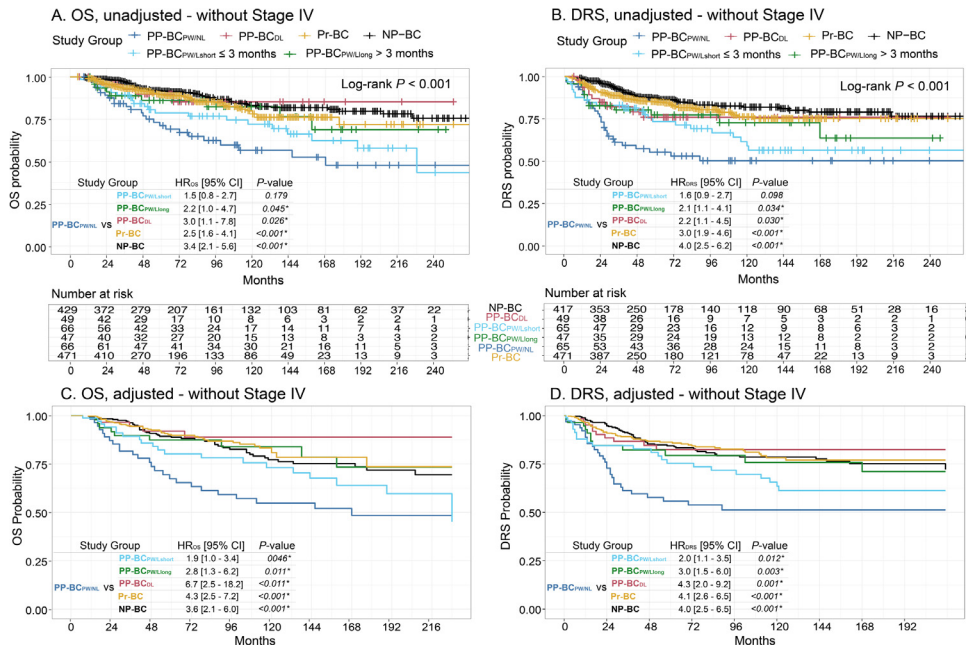


Fig. 6. Unadjusted and adjusted Cox regression analyses excluding stage IV cases in PP-BC_{PW/NL} (n = 66), PP-BC_{PW/Lshort} (n = 66), PP-BC_{PW/Llong} (n = 47), PP-BC_{DL} (n = 49), Pr-BC (n = 471) and NP-BC patients (n = 429) demonstrated an increased risk of death (A) and metastasis (B) in PP-BC_{PW/NL} and PP-BC_{PW/Lshort} patients. Adjusted OS and DRS probability in PP-BC_{PW/NL} (n = 65), PP-BC_{PW/Lshort} (n = 61), PP-BC_{PW/Llong} (n = 47), PP-BC_{DL} (n = 49), Pr-BC (n = 394) and NP-BC (n = 394) patients indicated an increased risk of death (C) and metastasis (D) in both PP-BC_{PW/NL} and PP-BC_{PW/Lshort} patients. Multivariate Proportional Hazards models were adjusted for age at diagnosis, year of diagnosis, stage, grade, surrogate molecular subtype, surgery, and (neo-)adjuvant CT, RT, HT and anti-HER-2 therapy. Grade, HT and surgery were stratified to comply with the proportional hazard's assumption. HR of more (less) than 1 indicates higher (lower) risk of death or metastasis. Significant values are indicated with an asterisk (*).

1.4. Site of metastasis analyze

Differences in preferred site of metastasis were analysed among the different study groups by binary logistic regression models (Table 4). To assess the association of site of metastatic disease on breast cancer prognosis, unadjusted (Fig. 7A) and adjusted (Fig. 7B) OS models were drawn for our patient population stratified based on first site of metastasis (liver, n=67; bone, n=65; brain, n=25; lungs, n=45; other, n=9). These analyses excluded patients with multi-site metastatic disease or unknown site of first metastatic recurrence to avoid potential confounding due to multi-site metastasis. Similar analyses were performed in the same patient cohort, separated for liver (Fig. 8A), bone (Fig. 8B), brain (Fig. 8C) and lung (Fig. 8D) metastatic patients and stratified for study group to evaluate differences in prognosis for the metastatic site in a specific study group.

2. Experimental Design, Materials and Methods

2.1. Patient data collection

Patient data were collected via the University Hospitals Leuven and the International Network on Cancer, Infertility and Pregnancy (INCP) network. At UZ Leuven, patient data are readily available via the Multidisciplinary Breast Centre (MBC) patient database. This database contains

Table 4
Binary logistic regression models for site of distant metastasis.

Organ	Study Group	Metastases		Multivariate OR [95% CI] *	p-value
		Yes (%)	No (%)		
Liver	PP-BC _{PW} (n = 183)	25.1%	74.9%	–	–
	vs PP-BC _{DL} (n = 53)	3.8%	96.2%	8.6 [2.0–36.6]	0.004
	vs Pr-BC (n = 486)	8.4%	91.6%	3.8 [2.2–6.7]	<0.001
	vs NP-BC (n = 419)	6.4%	93.6%	4.9 [2.7–9.1]	<0.001
Bone	PP-BC _{PW} (n = 183)	10.3%	89.7%	–	–
	vs PP-BC _{DL} (n = 53)	14.3%	85.7%	0.7 [0.3–1.7]	0.460
	vs Pr-BC (n = 486)	11.7%	88.3%	0.7 [0.3–1.5]	0.364
	vs NP-BC (n = 419)	16.4%	83.6%	0.6 [0.3–1.2]	0.172
Brain	PP-BC _{PW} (n = 183)	6.6%	93.4%	–	–
	vs PP-BC _{DL} (n = 53)	6.1%	93.9%	0.6 [0.3–3.7]	0.530
	vs Pr-BC (n = 486)	6.6%	93.4%	2.2 [0.7–6.3]	0.161
	vs NP-BC (n = 419)	6.0%	94.0%	1.2 [0.5–3.2]	0.690
Lung	PP-BC _{PW} (n = 183)	9.0%	91.0%	–	–
	vs PP-BC _{DL} (n = 53)	10.2%	89.8%	1.1 [0.2–4.9]	0.906
	vs Pr-BC (n = 486)	8.8%	91.2%	2.0 [0.8–4.7]	0.128
	vs NP-BC (n = 419)	5.4%	94.6%	1.7 [0.8–3.8]	0.192

Binary Logistic regression model for liver, bone, brain and lung metastasis. Models adjusted for age at diagnosis, year of diagnosis, stage, grade, surrogate molecular subtype, surgery, CT, RT, HT and anti-HER-2 therapy. As histological subtype did not significantly differ between study groups (Tables 1 and 2), this factor was not further taken into account.

* Odds Ratio (OR) and 95% Confidence Intervals (95% CI). OR larger (lower) than 1 indicates increased (decreased) occurrence of that parameter in PP-BCPW compared to the reference type (either PP-BCDL, Pr-BC or NP-BC). Significant values are indicated in bold.

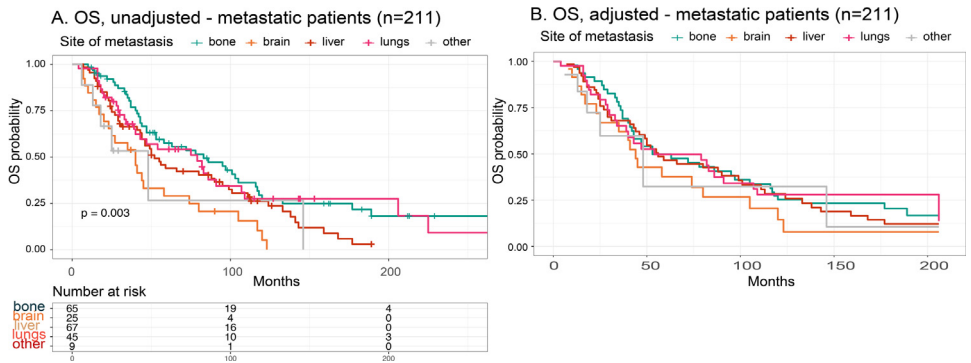


Fig. 7. Unadjusted and adjusted OS model of metastatic patients with known primary site of metastasis (n = 211) stratified based on site of metastasis. A. Unadjusted OS model, where patients were stratified based on site of metastasis (liver, n = 67; bone, n = 65; brain, n = 25; lungs, n = 45; other, n = 9). Patients with brain metastasis had an overall, more than 2-fold increased risk of death in comparison to patients that metastasized to other sites (log-rank, P = 0.003). B. OS model adjusted for age at diagnosis, year of diagnosis, stage, grade, surrogate molecular subtype and surgery, (neo)adjuvant CT-, RT-, HT- and/or anti-HER-2 therapy. Patients with brain metastasis remained to have a significant, almost 2-fold increased risk compared to patients that metastasized to other metastatic sites. No other significant differences could be observed.

information on tumour and therapy characteristics as well as patient information of all breast cancer patients diagnosed in UZ Leuven since 1960. Another important arm of data recruitment occurred through the INCIP network. The INCIP patient registry was founded in 2009 by Prof. Amant within the European Society of Gynaecological Oncology with the primary objective to facilitate large-scale studies on cancer in pregnancy. Today, INCIP has 97 members from 30 countries and 75 different centres. In the online registration program (www.cancerinpregnancy.org)

Details on pregnancy and breastfeeding were retrieved through extensive investigation of the patient records, and subsequently used to stratify patients in one of the study groups.

Next, breast cancer patients were divided into groups based on their pregnancy and lactation status. NP-BC are nulliparous patients with no history of pregnancy; Pr-BC are women diagnosed during pregnancy; PP-BC_{DL} are patients diagnosed during lactation; PP-BC_{PW} are women diagnosed post-weaning. PP-BC_{PW} cases were further subdivided based on breastfeeding duration prior to the pregnancy-associated cancer diagnosis. PP-BC_{PW/NL} are women diagnosed post-weaning who never lactated; PP-BC_{PW/Lshort} are women diagnosed post-weaning who breastfed for ≤ 3 months; PP-BC_{PW/Llong} are women diagnosed post-weaning who breastfed for ≥ 3 months prior to diagnosis.

2.2. Statistical analyze

A priori power calculations (STATA) indicated that, based on 200 patients per group with an event rate of 20% (and an average follow-up of 8 years), we have sufficient power ($> 80\%$) to detect differences in metastasis and survival (HR 2.0) and small intergroup differences of 10 to 15% between PP-BC_{PW} subgroups and Pr-BC and NP-BC groups. Larger clinicopathological differences ($> 20\%$) and variations in prognosis (HR ≥ 3.0) within the smaller PP-BC subgroups are identifiable at $> 70\%$ power. Frequencies of prognostic categorical variables were evaluated using fisher's exact tests or chi-square testing. Continuous variables, *i.e.* age at diagnosis, year of diagnosis and tumour size, were compared via One-Way ANOVA (parametric) or kruskal-wallis (non-parametric) analyses, and in addition, Tukey HSD testing was used for performing multiple pairwise-comparison between the means of groups. Odds ratios (OR) with 95% confidence interval (CI) were determined using (multinomial) logistic regression.

The risk of distant recurrence and death of any cause was determined using Kaplan-Meier analyses. Log-rank tests were applied to assess differences between distant recurrence and survival probabilities across groups. Distant recurrence-free survival (DRS) was defined as the time of diagnosis until the time of metastasis, or, if no metastasis was recorded, until the time of the last follow-up visit. Overall survival (OS) was calculated from the date of diagnosis until the date of death or last follow-up visit among those alive at the end of follow-up period. Using R (packages survival, survminer and KMsurv), Cox proportional multivariate logistic regression models were performed, and Kaplan-Meier curves were drawn to investigate differences in DRS and OS hazard ratios across study groups. To assess the association between patient group and prognosis, multivariate proportional hazards models were adjusted for age at diagnosis, year of diagnosis, stage (accounting for lymph node infiltration and tumour size), grade, surrogate molecular subtype, surgery, and (neo-)adjuvant CT, RT, HT and anti-HER-2 therapy. Grade, HT and surgery were stratified to comply with the proportional hazard's assumption [4]. HR of more (less) than 1 indicated higher (lower) risk of death or metastasis. The adequacy of the Cox proportional hazards assumptions for the included variables was checked by $\log(-\log(\text{survival}))$ curves. To rule out potential bias introduced due to centre of diagnosis on prognosis, OS and DRS calculations were stratified based on inclusion via UZ Leuven or INCIP. Proportional hazard models for study groups were performed both in- and excluding stage IV cases, with distant metastasis at diagnosis, to evaluate the effect on outcome.

Binary logistic regression models were used to assess the effect of study group on the frequency of metastatic disease to either the liver, bone, brain or lungs, while adjusting for the same confounding variables as described before. Subsequent analyses were performed on the subset of patients with metastatic disease, either at diagnosis or during follow-up. To avoid potential confounding due to multi-site metastasis, we performed these additional analyses only taking into account the primary site of metastatic disease. Patients were excluded if site of first metastatic recurrence was unknown, or if diagnosed with multi-site metastatic disease. The association between liver, bone, brain or lung metastases and our study groups was assessed using two-sided fisher's exact tests. In addition, we performed Cox regression models to determine the association between the distinct metastatic sites and OS probability.

All statistical analyses were performed using R version 3.4.4. A p-value less than 0.05 was considered statistically significant.

Ethics Statement

The study was approved by the Ethics Committee Research UZ/KU Leuven, Belgium (study number: S25470). Written informed consent was available for every patient included in the study. The study protocol conforms to the ethical guidelines of the 2000 Declaration of Helsinki.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Hanne Lefrère: Conceptualization, Methodology, Formal analysis, Investigation, Writing – original draft, Visualization; **Giuseppe Floris:** Conceptualization, Validation, Resources, Writing – review & editing; **Marjanka K. Schmidt:** Methodology, Formal analysis, Writing – review & editing; **Patrick Neven:** Resources, Writing – review & editing; **Ellen Warner:** Resources, Writing – review & editing; **Elyce Cardonick:** Resources, Writing – review & editing; **Fedro Alessandro Peccatori:** Resources, Writing – review & editing; **Sibylle Loibl:** Resources, Writing – review & editing; **Charlotte Maggen:** Resources, Writing – review & editing; **Hanne De Mulder:** Resources, Writing – review & editing; **Katarzyna J. Jerzak:** Resources, Writing – review & editing; **Diether Lambrechts:** Conceptualization, Methodology, Writing – review & editing, Supervision; **Liesbeth Lenaerts:** Conceptualization, Methodology, Writing – original draft, Supervision, Project administration, Funding acquisition; **Frédéric Amant:** Conceptualization, Methodology, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition.

Acknowledgments

We thank all participating patients and all (para-)medical staff involved in registering cases in the INCIP database (see www.cancerinpregnancy.org). We also thank Prof. Flora van der Leeuwen for providing additional Dutch patient data.

References

- [1] E.K. Hartman, G.D. Eslick, The prognosis of women diagnosed with breast cancer before, during and after pregnancy: a meta-analysis, *Breast Cancer Res. Treat.* 160 (2016) 347–360.
- [2] M.E. Hammond, D.F. Hayes, A.C. Wolff, P.B. Mangu, S. Temin, American society of clinical oncology/college of american pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer, *J. Oncol. Pract.* 6 (2010) 195–197.
- [3] A.C. Wolff, M.E. Hammond, D.G. Hicks, M. Dowsett, L.M. McShane, K.H. Allison, et al., Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American society of clinical oncology/college of American pathologists clinical practice guideline update, *J. Clin. Oncol.* 31 (2013) 3997–4013.
- [4] F.Y. Hsieh, P.W. Lavori, Sample-size calculations for the Cox proportional hazards regression model with nonbinary covariates, *Control Clin. Trials* 21 (2000) 552–560.