Incidence of breast cancer subtypes in Belgium: a population-based study

L. van Walle, MD[1], J. Vandeven[1], C. Colpaert, MD, PhD[1,2], FP. Duhoux, MD, PhD[3], P. Neven, MD, PhD[4], L. Van Eycken, MD[1], N. Van Damme, PhD[1]

SUMMARY

The aim of this study is to provide a reference for the Belgian breast cancer population, offering detailed information on various patient and tumour characteristics for the breast cancer population as a whole, as well as for the different molecular subtypes. Incidence data for primary invasive breast cancer in females diagnosed in 2014 were selected in the Belgian cancer registration database and underwent individual manual reviewing of the pathology protocols. Subsequently, in 95% of the study population a surrogate molecular subtype was successfully derived, using the combined expression of oestrogen receptor, progesterone receptor, human epidermal growth factor receptor-2, and tumour differentiation grade as surrogate for the proliferation marker Ki67, in conformity with the 2011 St Gallen surrogate classification. Ultimately, differences between the molecular subtypes regarding initial presentation and histopathological features were evaluated by means of a Pearson Chi-squared test for independence. Furthermore, relative survival was calculated for the different molecular subtypes. Histologically, the large majority of the Belgian breast cancer population presents with invasive breast carcinoma of no special type (NST), formerly called invasive ductal carcinoma (75.2%), 14.5% with invasive lobular carcinoma and 5.8% with mixed ductal/lobular invasive carcinoma. Less than five percent of the population harbours less frequently occurring histological subtypes. The Belgian breast cancers are predominantly of the luminal A-like subtype (54.4%), followed by the luminal B-like HER2 negative (14.7%) and the luminal B-like HER2 positive subtype (12.2%). The mean age at diagnosis is 62 years, with almost a third of the patients being 70 years or older. One out of five patients is younger than 50 years, and in the triple negative population this group counts for 31.9%, compared to 16.6% in the luminal A-like breast carcinomas. Most patients (69.4%) are diagnosed with early stage breast cancer (clinical stage 0-II); six percent of the breast cancers are clinically metastasised at the time of diagnosis. For 19% of the patients, information on clinical stage was lacking or staging was not applicable. The unadjusted five-year relative survival proportion for the Belgian cohort is 91.4%. Luminal A-like breast cancer opposed to triple negative breast cancer have the best and worst relative survival, with respectively 96.8% and 77.4% five-year relative survival proportions.

(BELG J MED ONCOL 2020;14(6):263-73)

[1]Belgian Cancer Registry, Brussels, Belgium, [2]General Hospital Turnhout and University Hospitals Leuven, Leuven, Belgium, [3]King Albert II Cancer Institute, Department of Medical Oncology, University Hospitals Saint-Luc and Institute of Experimental and Clinical Research (pole MIRO), Leuven Catholic University Centre (UCL), Brussels, Belgium, [4]Department of Obstetrics and Gynecology, Multidisciplinary Breast Centre and Leuven University Centre for Cancer Prevention (LUCK), University Hospitals Leuven, Leuven, Belgium

Please send all correspondence to: L. van Walle, MD, Stichting Kankerregister, Koningsstraat 215 B7, 1210 Brussels, Belgium, tel: +32 474/736614, e-mail: lien.vanwalle@kankerregister.org.

Conflict of interest: The authors have nothing to disclose and indicate no potential conflict of interest.

Keywords: breast cancer, demographics, molecular subtypes, population-based, prognosis.

Financial support: The Fonds Pink Ribbon provided financial support for the manual retrieval of information from the pathology protocols.

Acknowledgements: The authors would like to thank the team of data managers at the Belgian Cancer Registry for their work in retrieving information from the pathology reports, Dr. An Poppe, previous collaborator at the Belgian Cancer Registry, for her contribution to this study and the oncological care programs and services for pathological anatomy in the Belgian hospitals for their cooperation with cancer registration.

INTRODUCTION

For the past decades, breast cancer has been the most frequently diagnosed malignancy and the most important cause of cancer death in females worldwide.[1-3] In 2017, Belgium counted 10,627 new diagnoses of breast cancer, consistent with an age-standardised incidence rate of 104.4 using the World Standard Population (WSR; N/100,000 person years).[4] Although breast cancer mortality rates are declining in the Belgian population, breast cancer currently remains the leading cause of cancer death in women, albeit closely followed by lung cancer.[5,6]

The clinical, cellular and molecular heterogeneity of breast cancer is generally accepted and various initiatives to identify prognostic subtypes have been undertaken.[7-11] Accurate characterisation of breast cancer subtypes is necessary for selecting the initial therapeutic approach, guiding (neo-)adjuvant treatment, predicting clinical behaviour and determining prognosis.[12,13] Ultimately, breast cancer subtyping enables a more refined definition of quality of care indicators, enhancing the evaluation of cancer care.

This study describes the female Belgian breast cancer population anno 2014. Therefore, this study aims to serve as a reference for the female Belgian breast cancer population, offering several descriptive parameters including the distribution and characterisation of the different molecular subtypes.

MATERIAL AND METHODS

DATA SOURCES

The Belgian Cancer Registry (BCR) is a population-based cancer registry collecting national data since the incidence year 2004. The available data consist of clinical information registered by the oncological care programs (*e.g.* patient and tumour characteristics), supplemented with histopathological information provided by the laboratories of pathology (*i.e.* encoded information on every (pre-)malignant specimen, supplemented by the pathology report in full text).[5,6] All data that enters the Cancer Registry is submitted to an extended set of automated and manual validation procedures to ensure validity and quality of the data. Information on the patient's vital status and official residence are obtained through the Belgian Crossroads Bank for Social Security. If necessary, the BCR registry data can be linked with health insurance claims data provided by the Intermutualistic Agency (IMA) to obtain more detailed information on diagnostic and therapeutic procedures reimbursed by the mandatory Belgian health insurance. Linkage is done via the unique Social Security Identification Number (SSIN).

STUDY POPULATION

Incidence data for unilateral primary invasive breast cancers in females, diagnosed between January 1st and December 31st, 2014 and with an official residence in Belgium at the moment of diagnosis, were selected. Bilateral breast cancers - either synchronically or sequentially - (N=1,125 according to follow-up status until 31/12/2014) and sarcoma (N=4) were excluded from the analysis. A total of 10,082 cases were finally included in this study.

DESCRIPTIVE PARAMETERS

Patient and tumour characteristics

For this study an extra effort was done through review of the pathology reports to enhance completeness of specific tumour characteristics in the registry database such as focality, tumour size, nodal involvement and differentiation grade (using the Elston-Ellis grading system).[14] If unavailable in the pathology reports, they were completed with the information derived from the oncological care programs. As a proxy for tumour size and number of metastatically involved lymph nodes, pathological T (pT) and N (pN) categories were deducted according to the 7th edition of the tumour, node and metastasis (TNM) classification.[15] For the patients who received neo-adjuvant treatment, a ypT and ypN was defined. Information regarding the histological subtype as well as the clinical TNM classification (*i.e.* clinical T-, N-, M- category and clinical stage) was obtained from the registry database. Histological subtypes were defined based on the WHO International Classification of Diseases for Oncology (ICD-O) (3rd edition) of morphology codes.[16]

Biomarkers

Information regarding hormone receptor expression and human epidermal growth factor receptor 2 (HER2) overexpression and/or HER2/neu gene amplification was derived from the pathology reports by manual review. In general, the result of the resection specimen was preferred over the biopsy, with the exception of the neo-adjuvant setting in which the result of the biopsy was prioritised. Oestrogen receptor (ER) and progesterone receptor (PR) status were categorised as positive, negative or unknown based on the results of immunohistochemistry (IHC) assays. Different scoring systems were used in the original pathology reports (*e.g.* Allred Quick score, H score, *etc.*) according to the preference of the pathologist.

HER2 status was labelled as positive, negative or unknown. HER2 was evaluated through IHC and/or *in situ* hybridisation (ISH).[17,18] In general, the result of the ISH test was decisive for HER2 categorisation. In case there was no HER2 ISH test result available or if the HER2 ISH result was equivocal, the health insurance claims data were consulted for trastuzumab administration within twelve months after incidence date. If trastuzumab administration was confirmed, HER2 status was categorised as positive. If absent, HER2 sta-

TABLE 1. Surrogate molecular subtypes according to tumour differentiation grade, hormonal receptor and HER2 status.

Tumour Differentiation Grade	ER+ PR+ HER2-	ER+ PR- HER2-	ER- PR+ HER2-	ER+ PR+ HER2+	ER+ PR- HER2+	ER- PR+ HER2+	ER- PR- HER2+	ER- PR- HER2-	ER and PR unk, HER2 equ or unk
1	Lum A	Lum A	Lum A	Lum B HER2+	Lum B HER2+	Lum B HER2+	HER2+ non-lum	TN	Unk
2	Lum A	Lum A	Lum A	Lum B HER2+	Lum B HER2+	Lum B HER2+	HER2+ non-lum	TN	Unk
3	Lum B HER2-	Lum B HER2-	Lum B HER2-	Lum B HER2+	Lum B HER2+	Lum B HER2+	HER2+ non-lum	TN	Unk
Unk	Unk	Unk	Unk	Lum B HER2+	Lum B HER2+	Lum B HER2+	HER2+ non-lum	TN	Unk

Lum A= luminal A-like; Lum B HER2-= luminal B-like HER2 negative; Lum B HER2+= luminal B-like HER2 positive; HER2+ non-lum= HER2 positive non-luminal; TN= triple negative; Unk = unknown; equ = equivocal.

tus was further completed with the IHC test result if the latter was more informative. Remaining equivocal results were set to negative.

MOLECULAR SUBTYPE DEFINITION

Based on different combinations of ER-, PR-, HER2 status and tumour differentiation grade, five molecular subtypes were distinguished according to the 2011 St Gallen surrogate classification.[19] Luminal A-like, luminal B-like HER2 negative, luminal B-like HER2 positive, HER2 positive non-luminal and triple negative subtypes were determined, as shown in *Table 1*. Tumour differentiation grade was used as surrogate for the proliferation marker Ki67.[19]

DESCRIPTIVE AND STATISTICAL ANALYSIS

The association between the molecular subtype and patient and tumour baseline characteristics was evaluated with the Pearson Chi-squared test for independence in contingency tables. When independence is rejected by this test, Pearson residuals allow to identify cells which show dependence. Cells with large positive (negative) Pearson residual values have a higher (lower) proportion than expected under independence. An absolute residual value of two or more is indicative of lack of fit in that cell under the assumption of independence.[20]

Survival time was calculated from the incidence date to the date of death or until the last known date alive. Follow-up was complete up to December 31st, 2019. Relative survival was calculated as the ratio of the observed survival and the expected survival for a similar group of persons from the general Belgian population (stratified on gender, age, calendar year and region) using an actuarial approach in one year wide time intervals. Observed survival (OS) proportions were estimated using the Kaplan-Meier method.[21] The Ederer II method was applied to estimate the expected survival using the Belgian national lifetables.[22] Fourteen out of 10,082 patients were lost-to-follow-up on the date of diagnosis and were therefore not considered in survival analysis, resulting in N=10,068 numbers at risk for the Belgian cohort.

All analyses were performed with SAS version 9.4 (SAS Institute Inc., Cary, NC, USA).

RESULTS

GENERAL BREAST CANCER POPULATION

In total, 10,082 incident cases of unilateral invasive breast carcinoma were diagnosed in Belgium in 2014, hereafter referred to as the general Belgian breast cancer population.

Patient and tumour characteristics such as age at diagnosis, tumour focality and differentiation grade, histological subtype, clinical stage (cStage), clinical T-, N- and M-categories, pathological stage (pStage) and pathological (y)T-, (y)N- and (y)M-categories are provided in *Tables 1S to 4S (Supplementary File)*.

The mean age at diagnosis of the Belgian breast cancer population is 62 years. Almost half of the patients (48.1%; N=4,846) are between 50 and 69 years old at the time of diagnosis, 20.2% (N=2,307) of the patients is younger than 50 years and about a third is aged 70 years or older (31.7%; N=3,199).

Of the available information on tumour focality (71.2%; N=7,176), 79.4% (N=5,700) of the patients display an unifocal breast tumour compared to 20.6% (N=1,476) having a multifocal tumour. The majority of the breast carcinomas is moderately differentiated (*i.e.* grade 2) (46.6%; N=4,698),

TABLE 2. Unadjusted five-year relative survival (RS) by surrogate molecular subtype, for female breast cancer in Belgium 2014, bilateral breast cancers and sarcoma excluded.

	N Total	%	N at Risk (*)	Estimate	95% CI
Overall	10,082		10,068	91.4	[90.6, 92.2]
Molecular Subtype					
LumA	5,485	54.4	5,482	96.8	[95.8, 97.8]
LumB, HER2-	1,484	14.7	1,483	86.8	[84.4, 89.1]
LumB, HER2+	1,235	12.2	1,234	92.2	[89.9, 94.2]
HER2+, non-lum	468	4.6	467	86.9	[82.7, 90.5]
TN	866	8.6	864	77.4	[73.9, 80.5]
Unknown	544	5.4	538	73.1	[68.1, 77.7]

Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative.
(*) N=14 patients were lost-to-follow-up on the date of diagnosis and were not considered for survival analysis.

the second major proportion being poorly differentiated (*i.e.* grade 3) (31.2%; N=3,148). In 3.4% of the patients (N=343) information on tumour differentiation grade is lacking.

Histologically, 75.2% (N=7,580) of all unilateral invasive breast carcinomas are invasive ductal carcinomas NST (IDC), 14.5% (N=1,462) invasive lobular carcinomas (ILC) and 5.8% (N=581) mixed ductal/lobular invasive carcinomas. The remaining 4.6% (N=459) of the cases represent various less frequently occurring histological subtypes, with 1.0% (N=98) mucinous carcinomas, 0.6% (N=58) papillary carcinomas, 0.2% (N=24) micropapillary carcinomas, 0.3% (N=32) medullary carcinomas, 0.4% (N=42) metaplastic carcinomas, 0.2% (N=19) tubular carcinomas, 0.1% (N=14) malignant Phyllodes tumours and 0.4% (N=39) Paget's disease (*Table 1S of Supplementary File*).

Most patients are diagnosed with early stage breast cancer, 85.4% (N=6,993) of the patients with a known clinical stage at the time of diagnosis having a cStage 0-II. This is confirmed by the pathological stage, as 89.5% (N=6,692) of the operated patients who are not treated with neo-adjuvant therapy and for whom the pStage is known, have a pStage I-II. For the operated patients who received neo-adjuvant therapy and have a known ypStage, 14.6% (N=145) have a complete pathological response. 62.0% of this population (N=615) has a ypStage I-II. More detailed results for the c/pTNM categories and the c/pStage can be found in *Tables 2S to 4S (Supplementary File).*

POPULATION WITH KNOWN MOLECULAR SUBTYPE

In 94.6% of the study population, corresponding to a total of 9,538 cases, a surrogate molecular subtype was successfully derived whereas for 5.4% (N=544) there was insufficient information available on either tumour grade, hormone receptor or HER2 status to define a molecular subtype (*Table 2*). The Belgian breast cancer population predominantly presents with the luminal A-like subtype (54.4%; N=5,485) followed by the luminal B-like HER2 negative (14.7%; N=1,484) and luminal B-like HER2 positive subtypes (12.2%; N=1,235). In total, 8.6% of the patients have triple negative breast cancer (N=866), and 4.6% have HER2 positive non-luminal breast cancers (N=468). Patient and tumour characteristics and relative survival proportion for the total population with known surrogate molecular subtype (N=9,538) and for the different surrogate molecular subtypes separately are described hereafter according to the results presented in *Tables 2 to 4.*

Age at diagnosis

Table 3 shows that a significantly larger proportion of the luminal A-like breast cancers are diagnosed between 50 and 69 years of age (51.2%; N=2,806), in contrast with the triple negative population that is significantly less represented in this age group (40.3%; N=349). Patients younger than 50 years are significantly less common in the luminal A-like population (16.6%; N=911) and meanwhile more frequent in the triple negative (31.9%; N=276) and in the luminal B-like HER2 positive population (26.9%; N=332). Patients aged 70 years or older at diagnosis are significantly less frequent in the luminal B-like HER2 positive population (25.5%; N=315), contrasting with the luminal B-like HER2 negative population of which 35.1% (N=521) falls in this age category.

TABLE 3. Tumour and patient characteristics for the population with known molecular subtype (N=9,538).

Characteristic	Total N (%)	LumA N (%)	[PR]	LumB, HER2- N (%)	[PR]	LumB, HER2+ N (%)	[PR]	HER2+, non-lum N (%)	[PR]	TN N (%)	[PR]	P
All Patients	9,538 (100.0)	5,485 (100.0)		1,484 (100.0)		1,235 (100.0)		468 (100.0)		866 (100.0)		
Age at Diagnosis												
Mean (SD)	62 (14.0)	63 (13.3)		63 (14.7)		60 (14.2)		60 (13.6)		59 (16.2)		
Median (IQR)	62 (52-73)	63 (53-73)		62 (52-75)		59 (49-70)		59 (50-71)		59 (46-71)		
Age Categories												**<.0001**
<50 years	1,938 (20.3)	911 (16.6)	**[-10.47]**	307 (20.7)	[0.38]	332 (26.9)	**[6.14]**	112 (23.9)	[1.99]	276 (31.9)	**[8.86]**	
50-69 years	4,628 (48.5)	2,806 (51.2)	**[5.99]**	656 (44.2)	**[-3.62]**	588 (47.6)	[-0.69]	229 (48.9)	[0.18]	349 (40.3)	**[-5.08]**	
≥70 years	2,972 (31.2)	1,768 (32.2)	**[2.63]**	521 (35.1)	**[3.57]**	315 (25.5)	**[-4.60]**	127 (27.1)	[-1.93]	241 (27.8)	**[-2.22]**	
Tumour Focality												**<.0001**
Unifocal	5,535 (58.0)	3,405 (62.1)	**[9.32]**	781 (52.6)	**[-4.59]**	619 (50.1)	**[-6.04]**	231 (49.4)	**[-3.90]**	499 (57.6)	[-0.26]	
Multifocal	1,417 (14.9)	870 (15.9)	**[3.21]**	200 (13.5)	[-1.63]	192 (15.5)	[0.73]	68 (14.5)	[-0.20]	87 (10.0)	**[-4.17]**	
Unknown	2,586 (27.1)	1,210 (22.1)	**[-12.91]**	503 (33.9)	**[6.40]**	424 (34.3)	**[6.12]**	169 (36.1)	**[4.49]**	280 (32.3)	**[3.62]**	
Histological Subtype												**<.0001**
IDC	7,261 (76.1)	3,862 (70.4)	**[-15.24]**	1,205 (81.2)	**[4.99]**	1,043 (84.5)	**[7.36]**	407 (87.0)	**[5.64]**	744 (85.9)	**[7.08]**	
ILC	1,356 (14.2)	1,046 (19.1)	**[15.79]**	164 (11.1)	**[-3.80]**	107 (8.7)	**[-5.99]**	13 (2.8)	**[-7.27]**	26 (3.0)	**[-9.91]**	
Mixed ductal/lobular	556 (5.8)	391 (7.1)	**[6.30]**	74 (5.0)	[-1.51]	61 (4.9)	[-1.43]	14 (3.0)	**[-2.69]**	16 (1.8)	**[-5.24]**	
Other carcinoma types	365 (3.8)	186 (3.4)	**[-2.58]**	41 (2.8)	**[-2.33]**	24 (1.9)	**[-3.70]**	34 (7.3)	**[3.98]**	80 (9.2)	**[8.70]**	
Clinical Stage (cStage)												**<.0001**
0/I (*)	3,709 (38.9)	2,597 (47.3)	**[19.72]**	405 (27.3)	**[-9.97]**	402 (32.6)	**[-4.90]**	95 (20.3)	**[-8.46]**	210 (24.2)	**[-9.27]**	
II	3,055 (32.0)	1,403 (25.6)	**[-15.71]**	604 (40.7)	**[7.79]**	492 (39.8)	**[6.30]**	170 (36.3)	**[2.04]**	386 (44.6)	**[8.30]**	
III	576 (6.0)	215 (3.9)	**[-10.11]**	124 (8.4)	**[4.08]**	83 (6.7)	[1.08]	66 (14.1)	**[7.51]**	88 (10.2)	**[5.34]**	
IV	477 (5.0)	208 (3.8)	**[-6.30]**	94 (6.3)	**[2.56]**	92 (7.4)	**[4.23]**	47 (10.0)	**[5.13]**	36 (4.2)	[-1.20]	
X/NA	1,721 (18.0)	1,062 (19.4)	**[3.89]**	257 (17.3)	[-0.79]	166 (13.4)	**[-4.51]**	90 (19.2)	[0.68]	146 (16.9)	[-0.95]	

TABLE 3. Continued.

Characteristic	Total N (%)	LumA N (%)	LumA [PR]	LumB, HER2- N (%)	LumB, HER2- [PR]	LumB, HER2+ N (%)	LumB, HER2+ [PR]	HER2+, non-lum N (%)	HER2+, non-lum [PR]	TN N (%)	TN [PR]	P
cT Category												**<.0001**
is	86 (0.9)	47 (0.9)	[-0.54]	6 (0.4)	**[-2.21]**	18 (1.5)	**[2.21]**	11 (2.4)	**[3.40]**	4 (0.5)	[-1.44]	
1	3,905 (40.9)	2,693 (49.1)	**[18.84]**	440 (29.6)	**[-9.63]**	427 (34.6)	**[-4.88]**	102 (21.8)	**[-8.64]**	243 (28.1)	**[-8.08]**	
2	2,831 (29.7)	1,278 (23.3)	**[-15.87]**	567 (38.2)	**[7.82]**	465 (37.7)	**[6.57]**	174 (37.2)	**[3.64]**	347 (40.1)	**[7.02]**	
3	449 (4.7)	183 (3.3)	**[-7.35]**	94 (6.3)	**[3.22]**	78 (6.3)	**[2.86]**	39 (8.3)	**[3.80]**	55 (6.4)	**[2.39]**	
4	464 (4.9)	191 (3.5)	**[-7.30]**	106 (7.1)	**[4.44]**	62 (5.0)	[0.27]	46 (9.8)	**[5.12]**	59 (6.8)	**[2.79]**	
x	1,803 (18.9)	1,093 (19.9)	**[2.97]**	271 (18.3)	[-0.69]	185 (15.0)	**[-3.77]**	96 (20.5)	[0.91]	158 (18.2)	[-0.52]	
cN Category												**<.0001**
0	5,614 (58.9)	3,526 (64.3)	**[12.53]**	788 (53.1)	**[-4.91]**	690 (55.9)	**[-2.29]**	193 (41.2)	**[-7.94]**	417 (48.2)	**[-6.72]**	
1	1,419 (14.9)	535 (9.8)	**[-16.36]**	274 (18.5)	**[4.22]**	263 (21.3)	**[6.79]**	130 (27.8)	**[8.04]**	217 (25.1)	**[8.83]**	
2	142 (1.5)	45 (0.8)	**[-6.27]**	31 (2.1)	**[2.08]**	21 (1.7)	[0.66]	20 (4.3)	**[5.10]**	25 (2.9)	**[3.56]**	
3	116 (1.2)	37 (0.7)	**[-5.61]**	30 (2.0)	**[3.08]**	20 (1.6)	[1.39]	16 (3.4)	**[4.46]**	13 (1.5)	[0.80]	
x	2,247 (23.6)	1,342 (24.5)	**[2.43]**	361 (24.3)	[0.76]	241 (19.5)	**[-3.59]**	109 (23.3)	[-0.14]	194 (22.4)	[-0.84]	
cM Category												**<.0001**
0	6,719 (70.4)	3,853 (70.2)	[-0.49]	1,060 (71.4)	[0.90]	876 (70.9)	[0.40]	310 (66.2)	**[-2.04]**	620 (71.6)	[0.78]	
1	477 (5.0)	208 (3.8)	**[-6.30]**	94 (6.3)	**[2.56]**	92 (7.4)	**[4.23]**	47 (10.0)	**[5.13]**	36 (4.2)	[-1.20]	
x	2,342 (24.6)	1,424 (26.0)	**[3.71]**	330 (22.2)	**[-2.26]**	267 (21.6)	**[-2.57]**	111 (23.7)	[-0.43]	210 (24.2)	[-0.22]	
Treatment												**<.0001**
Operated, no NAT	7,294 (76.5)	4,576 (83.4)	**[18.63]**	1,102 (74.3)	**[-2.19]**	854 (69.1)	**[-6.50]**	265 (56.6)	**[-10.38]**	497 (57.4)	**[-13.88]**	
Operated, with NAT	1,071 (11.2)	263 (4.8)	**[-23.15]**	183 (12.3)	[1.46]	226 (18.3)	**[8.44]**	133 (28.4)	**[12.08]**	266 (30.7)	**[19.05]**	
Not operated	1,173 (12.3)	646 (11.8)	[-1.80]	199 (13.4)	[1.42]	155 (12.6)	[0.29]	70 (15.0)	[1.80]	103 (11.9)	[-0.38]	

Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative, PR = Pearson residual value (cells with large positive (negative) Pearson residual values have a higher (lower) proportion than expected under independence. An absolute residual value of two or more (reported in bold) is indicative of lack of fit in that cell under the assumption of independence), SD= Standard Deviation, IQR= Inter-quartile Range, IDC= Invasive Carcinoma of no special type (NST), ILC= Invasive Lobular Carcinoma, X= Unknown, NA= Not Applicable (TNM stage is not applicable, e.g. Phyllodes tumours), NAT= Neo-adjuvant Treatment. (*) In correspondence with TNM 7th edition, cTis cN0 cM0 tumours are categorised as cStage 0, i.e. these tumours were clinically assessed as in situ but appeared to be invasive after resection.

TABLE 4. Distribution of pathological stage and pT, pN and pM category for operated patients with known molecular subtypes who were not treated with neo-adjuvant therapy.

pTNM	Total N (%)	LumA N (%)	[PR]	LumB, HER2- N (%)	[PR]	LumB, HER2+ N (%)	[PR]	HER2+, non-lum N (%)	[PR]	TN N (%)	[PR]
Operated patients, no NAT	7,294 (100.0)	4,576 (100.0)		1,102 (100.0)		854 (100.0)		265 (100.0)		497 (100.0)	
Pathological Stage (pStage)											
I	3,869 (53.0)	2,752 (60.1)	**[15.65]**	391 (35.5)	**[-12.70]**	412 (48.2)	**[-2.87]**	120 (45.3)	**[-2.63]**	194 (39.0)	**[-6.38]**
II	2,627 (36.0)	1,452 (31.7)	**[-10.03]**	530 (48.1)	**[9.11]**	324 (37.9)	[1.35]	102 (38.5)	[0.82]	219 (44.1)	**[3.99]**
III	751 (10.3)	349 (7.6)	**[-9.80]**	173 (15.7)	**[6.42]**	109 (12.8)	**[2.57]**	42 (15.8)	**[3.01]**	78 (15.7)	**[4.16]**
IV	18 (0.2)	11 (0.2)	[-]	4 (0.4)	[-]	2 (0.2)	[-]	0 (0.0)	[-]	1 (0.2)	[-]
X/NA	29 (0.4)	12 (0.3)	[-]	4 (0.4)	[-]	7 (0.8)	[-]	1 (0.4)	[-]	5 (1.0)	[-]
pT Category (Tumour Size)											
1	4,532 (62.1)	3,143 (68.7)	**[14.77]**	501 (45.5)	**[-12.45]**	502 (58.8)	[-1.94]	152 (57.4)	[-1.66]	234 (47.1)	**[-6.95]**
2	2,313 (31.7)	1,194 (26.1)	**[-13.55]**	512 (46.5)	**[11.42]**	301 (35.2)	**[2.49]**	96 (36.2)	[1.60]	210 (42.3)	**[5.41]**
3/4	413 (5.7)	225 (4.9)	**[-3.63]**	84 (7.6)	**[3.05]**	42 (4.9)	[-0.96]	16 (6.0)	[0.26]	46 (9.3)	**[3.66]**
x	36 (0.5)	14 (0.3)	[-]	5 (0.5)	[-]	9 (1.1)	[-]	1 (0.4)	[-]	7 (1.4)	[-]
pN Category (Nodal Status)											
0	4,833 (66.3)	3,205 (70.0)	**[8.86]**	636 (57.7)	**[-6.51]**	539 (63.1)	**[-2.07]**	150 (56.6)	**[-3.39]**	303 (61.0)	**[-2.59]**
1	1,593 (21.8)	927 (20.3)	**[-4.24]**	287 (26.0)	**[3.67]**	193 (22.6)	[0.57]	75 (28.3)	**[2.59]**	111 (22.3)	[0.28]
2	384 (5.3)	180 (3.9)	**[-6.60]**	95 (8.6)	**[5.41]**	59 (6.9)	**[2.29]**	12 (4.5)	[-0.55]	38 (7.6)	**[2.46]**
3	202 (2.8)	82 (1.8)	**[-6.60]**	49 (4.4)	**[3.68]**	32 (3.7)	[1.85]	20 (7.5)	**[4.83]**	19 (3.8)	[1.48]
x	282 (3.9)	182 (4.0)	[0.64]	35 (3.2)	[-1.29]	31 (3.6)	[-0.38]	8 (3.0)	[-0.73]	26 (5.2)	[1.64]
pM Category											
1	18 (0.2)	11 (0.2)	[-]	4 (0.4)	[-]	2 (0.2)	[-]	0 (0.0)	[-]	1 (0.2)	[-]
x	7,276 (99.8)	4,565 (99.8)	[-]	1,098 (99.6)	[-]	852 (99.8)	[-]	265 (100.0)	[-]	496 (99.8)	[-]

Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative, PR= Pearson residual value (cells with large positive (negative) Pearson residual values have a higher (lower) proportion than expected under independence. An absolute residual value of two or more (reported in bold) is indicative of lack of fit in that cell under the assumption of independence). Chi² test could not be performed for all levels considered in the table. As a consequence, Pearson residual values could not be reported for these levels [-], NAT= Neo-adjuvant Treatment, X= Unknown, NA= Not Applicable (TNM stage is not applicable, e.g. Phyllodes tumours).

Histological subtype

Although for all molecular subtypes the majority of cases are IDC (between 70.4% and 87.0%), there is some histological variation among the molecular subtypes. The luminal A-like population presents significantly more often as ILC (19.1%; N=1,046), whereas the proportion of IDC is smaller (70.4%; N=3,862) compared to the other subtypes. The HER2 positive non-luminal and the triple negative subtypes are more associated with other carcinoma types, notably with 7.3% (N=34) and 9.2% (N=80), respectively.

Distribution of the clinical stage

Luminal A-like breast carcinomas present significantly more often with cStage 0/I (47.3%; N=2,597), contrasting with the other four molecular subtypes. The luminal B-like HER2 negative and triple negative subtypes are more associated with cStage II-III tumours, while the HER2 positive non-luminal subtype is primarily shifted towards cStage III-IV tumours.

Clinical evaluation of the primary tumour

Tumours measuring 2 cm or less in greatest dimension (cT1) are generally the most frequently diagnosed (40.9%; N=3,905). This proportion is essentially based on luminal A-like breast cancers, in which 49.1% (N=2,693) presents with cT1 tumours. In the four other molecular subtypes cT1 tumours are significantly less represented, with the smallest proportion of 21.8% (N=102) in the HER2 positive non-luminal subgroup. Secondly, in 29.7% (N=2,831) of the breast cancer patients a clinical T2 is diagnosed. This proportion is significantly smaller in the luminal A-like subtype (23.3%; N=1,278) and conversely predominates in all four other subtypes with proportions varying between 37.2% and 40.1%. Furthermore, tumours measuring more than 5 cm in greatest dimension at presentation (cT3) account for 4.7% (N=449) and clinical T4 tumours represent 4.9% (N=464). Both cT3 and cT4 are significantly less often seen in luminal A-like breast cancers. Finally, in 18.9% of the patients (N=1,803) no information on clinical T-category is available (cTx).

Clinical evaluation of the nodal status

Regional nodal metastases are clinically detected (cN+) in 17.6% (N=1,677) of the population, as opposed to 58.9% (N=5,614) who are clinically assessed as cN0. Luminal A-like breast cancer presents significantly more often without clinical suspicion of nodal metastases (64.3% cN0; N=3,526) compared to the other molecular subtypes. In 23.6% (N=2,247) of the population the regional lymph nodes could not be assessed clinically (cNx).

Clinical evaluation of distant metastasis

In 5.0% (N=477) of the population distant metastases are clinically detected (cM1). The luminal B-like HER2 positive and HER2 positive non-luminal subtypes present with significantly larger proportions of cM1 tumours, *i.e.* 7.4% (N=92) and 10.0% (N=47) respectively.

Distribution of the pathological stage (in the absence of neo-adjuvant therapy)

Of the patients with a known molecular subtype, 87.7% (N=8,365) received breast surgery as part of the treatment. In 12.8% (N=1,071) of the operated patients, neo-adjuvant therapy preceded surgery. For the operated patients without neo-adjuvant treatment (N=7,294), pStage and pTNM categories are described comprehensively in *Table 3S*, detailed information on ypStage and ypTNM categories for the operated population treated with neo-adjuvant therapy is provided in *Table 4S* (*Supplementary File*).

The results demonstrated in *Table 4* evaluate the association between the molecular subtype and pathological stage (pTNM), tumour size (pT) and nodal status (pN) for the operated patients without neo-adjuvant therapy. Patients with luminal B-like HER2 positive, HER2 positive non-luminal and triple negative breast cancer significantly more often receive neo-adjuvant treatment, as shown in *Table 3*, subsequently these subtypes are proportionally less represented in the analysis presented in *Table 4*.

Tumours with a pStage I are significantly more represented in the luminal A-like subtype (60.1%; N=2,752) compared to all other subtypes. Luminal B-like HER2 negative and triple negative breast cancer is more associated with pStage II and III (48.1%, N=530 resp. 44.1%, N=219 pStage II and 15.7%, N=173 resp. 15.7%, N=78 pStage III).

Pathological evaluation of the primary tumour (in the absence of neo-adjuvant therapy)

With the exception of the luminal B-like HER2 negative subtype, pT1 tumours represent the majority in all molecular subtypes. Additionally, pT1 tumours count for a significantly larger proportion (68.7%; N=3,143) within the luminal A-like breast carcinomas compared to the other subgroups. Luminal B-like HER2 negative breast carcinomas present almost equally frequent with pT1 (45.5%; N=501) and pT2 tumours (46.5%; N=512). Except for the luminal B-like HER2 negative subtype, pT2 tumours are the second most common in all subtypes, with a proportion ranging from 26.1% (N=1,194) in the luminal A-like subgroup to 42.3% (N=210) in the triple negative subgroup. The proportion of pT2 tumours is significantly higher within both the luminal B-like subtypes as well as in the triple negative subtype. pT3-4 tumours manifest

significantly less often in the luminal A-like subtype (4.9%; N=225), in contrast, they occur more frequently in the luminal B-like HER2 negative (7.6%; N=84) and triple negative breast carcinomas (9.3%; N=46).

Pathological evaluation of the nodal status (in the absence of neo-adjuvant therapy)

Variation between the molecular subtypes regarding pathological nodal status is limited, more than 50% of the cases are classified pN0 in all subtypes. Nonetheless, the proportion of pN0 cases is significantly more dominant in the luminal A-like subtype (70.0%; N=3,205) compared to the other subtypes, while the luminal B-like HER2 negative and HER2 positive non-luminal subtypes are more associated with pN+ cases.

Five-year relative survival

The unadjusted five-year relative survival proportions (RS) are demonstrated in *Table 2*. The RS for the Belgian cohort is 91.4% (95%CI=[90.6,92.2]). RS in the luminal A-like subtype is superior compared to the other subtypes, with a five-year relative survival of 96.8% (95%CI = [95.8,97.8]), opposed to the triple negative subtype having the worst RS proportion (77.4%, 95%CI = [73.9,80.5]). RS proportions for the luminal B-like HER2 negative and HER2 positive non-luminal subtypes are similar, with 86.8% (95%CI = [84.4,89.1]) and 86.9% (95%CI = [82.7,90.5]), respectively.

POPULATION WITH UNKNOWN MOLECULAR SUBTYPE

In 5.4% (N=544) of the Belgian breast cancers the surrogate molecular subtype remains undetermined. The age at diagnosis of this particular population is generally higher, with a mean age of 65 years and 23.7% of the patients (N=129) being 80 years or older (*Table 1S of Supplementary File*). Furthermore, these patients present more often with metastatic disease (14.3%; N=78 cStage IV). Besides information on biomarkers, information concerning differentiation grade, tumour focality and stage is frequently lacking in this patient group (49.4%; N=269 grade unknown, 58.8%; N=320 focality unknown, 31.8%; N=173 cStage unknown, 5.1%; N=12 pStage unknown) (*Table 2S of Supplementary File*). As shown in *Table 2*, this population also has an inferior five-year relative survival proportion (73.1%; 95%CI = [68.1,77.7]) in comparison to the known molecular subtypes.

DISCUSSION

The heterogeneity in the breast cancer population, depicting breast cancer as a compilation of several distinct subtypes with unique prognostic and therapeutic implications has been extensively described.[7,8,10-13]

In this study, the Belgian breast cancer population was successfully characterised in conformity with the 2011 St Gallen surrogate classification, identifying five surrogate molecular subtypes for 95% of the study population.[19] In general, the obtained results are consistent with other published data, assuming a certain variation explained by the inherent variability between populations.[12,23,24] Comparison of the distribution of molecular subtypes in different populations is sometimes hampered by differences in the choice of parameters that are used to classify the cancers, *e.g.* tumour grade *versus* proliferation marker Ki67.[19] In addition, the applied cut-offs to determine either hormone receptor status (*e.g.* positivity equalling at least 1% *versus* at least 10% positive nuclei) or Ki67 status (*e.g.* 14% *versus* 20%) affect the final resulting molecular subtype.[13,19] In Belgium, prognostically favourable subtypes luminal A-like and luminal B-like HER2 negative account for 73% of the population with known molecular subtype (*i.e.* 58% luminal A-like and 16% luminal B-like HER2 negative), the remaining less favourable subtypes (luminal B-like HER2 positive, HER2 positive non-luminal and triple negative) representing 27% (*i.e.* 13% luminal B-like HER2 positive, 5% HER2 positive non-luminal and 9% triple negative). These results are very similar to recently published Dutch data that also used tumour grade to classify the cancers.[25] European studies using Ki67 instead of tumour grade tend to distinguish lower proportions of Luminal A breast cancer and higher proportions of Luminal B in comparison with our results.[26,27]

The association between age at diagnosis and the likelihood of presenting with particular subtypes is also confirmed by our results, and are comparable to other European data.[23,25,26,28,29] More than half of the luminal A-like breast cancers in Belgium are diagnosed between 50 and 69 years of age, a third is aged 70 years or older and only a minority presents at a younger age. In contrast, more aggressive subtypes such as the luminal B-like HER2 positive and the triple negative subtypes are more likely to present at a younger age.

In line with the publications from the Netherlands and Italy, our results establish the association of lobular histology with Luminal A-like breast cancer. [25,26]

As for the initial evaluation of the primary tumour, the results demonstrate an association between the molecular subtype and the initial tumour size, in particular for the luminal A-like subtype. c/pT1 tumours are more frequently of the luminal A-like subtype in contrast to c/pT3-4 tumours that are less likely to be luminal A-like. Also, in patients undergoing surgery without neo-adjuvant therapy, a more advanced tumour size (pT3-4) is more often associated with luminal

KEY MESSAGES FOR CLINICAL PRACTICE

1. Histologically, the overall majority of the Belgian breast cancer population are invasive carcinomas of no special type (NST), formerly called invasive ductal carcinomas (75.2%), 14.5% are invasive lobular carcinomas and 5.8% mixed ductal/lobular invasive carcinomas. Less than five percent of the population harbours less frequently occurring histological subtypes.

2. The mean age at diagnosis for the Belgian breast cancer patients is 62 years and a third of the patients is aged 70 years or older. Nevertheless, one out of five patients is younger than 50 years at the time of diagnosis.

3. The Belgian breast cancer population predominantly presents with luminal A-like carcinomas (54.4%), followed by luminal B-like HER2 negative (14.7%) and luminal B-like HER2 positive subtypes (12.2%).

4. Patients younger than 50 years at diagnosis are significantly more likely to present with triple negative or luminal B-like HER2 positive cancers contrasting with luminal A-like breast cancers that occur significantly less often in this age category.

5. Almost half (47.3%) of the luminal A-like breast carcinomas are diagnosed at a very early stage (clinical stage I).

6. Five percent of the Belgian breast cancers with known molecular subtype are clinically metastasised at the time of diagnosis. This subgroup with poor prognosis is significantly more represented in the HER2 positive non-luminal (10.0%) and luminal B-like HER2 positive subtypes (7.4%).

7. The unadjusted five-year relative survival proportion for the Belgian breast cancer population is 91.4%. Luminal A-like breast cancer opposed to triple negative breast cancer have the best and worst relative survival, with respectively 96.8% and 77.4% 5-year relative survival proportions.

B-like HER2 negative and triple negative breast carcinomas. The results show an association between the molecular subtype and the stage at diagnosis.[23] Luminal A-like breast carcinomas present significantly more often in very early stage (cStage 0/I), in line with the clinical tumour size. The less favourable subtypes HER2 positive non-luminal and luminal B-like HER2 positive present significantly more often in cStage IV.

Thirteen percent of the operated patients were treated with neo-adjuvant therapy, and as expected, the results indicate a clear association with less favourable subtypes (luminal B-like HER2 positive, HER2 positive non-luminal and triple negative).[13]

Finally, the five-year unadjusted relative survival for the Belgian breast cancer population differs according to the molecular subtype and the results are in line with previously published Belgian single centre data.[19] As expected, the luminal A-like and triple negative subtype establish as prognostically most *versus* least favourable. The German five-year unadjusted RS data for breast cancer are in line with our results, performing slightly better with 94.7% (95%CI=[93.4,96.0]) for the total cohort compared to 91.4% (95%CI=[90.6,92.2]) in our population.[27] The variation in five-year RS according to the molecular subtype reported in the German study is also very similar to our results.

CONCLUSION

This study offers an in-depth descriptive analysis of the Belgian breast cancer population, including detailed information on patient and tumour characteristics at initial presentation as well as surrogate molecular classification. Incidence data for primary, unilateral invasive breast carcinomas in females diagnosed in 2014 were selected in the BCR database, and underwent individual manual reviewing of the pathology protocols in search for receptor status. In 95% of the study population a surrogate molecular subtype was successfully derived, using tumour differentiation grade as surrogate for the proliferation marker Ki67 in conformity with the 2011 St Gallen surrogate classification. The various patient and tumour characteristics were analysed for the breast cancer population as a whole, as well as for the different molecular subtypes, highlighting variability between the different mo-

lecular subtypes regarding initial presentation, histopathological parameters and relative survival. This study can be used as a reference for the Belgian breast cancer population.

All Supplementary Files can be consulted online at www.BJMO.be

REFERENCES

1. Globocan: Fact sheets. https://gco.iarc.fr/today/data/factsheets/cancers/20-Breast-fact-sheet.pdf, accessed on 9 December 2019.

2. ECIS - European Cancer Information System https://ecis.jrc.ec.europa.eu, accessed on 9 December 2019 © European Union, 2019.

3. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.

4. Belgian Cancer Registry, Brussels, 2019; www.cancerregistry.org

5. Cancer in an Ageing Population, Belgium 2004-2016. Brussels: Belgian Cancer Registry, 2018.

6. Cancer Burden in Belgium, 2004-2017. Brussels: Belgian Cancer Registry, 2020.

7. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. Nature. 2000;406:747-52.

8. Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A. 2001;19:10869-74.

9. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. JAMA. 2006;295(21):2492-502.

10. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc Natl Acad Sci U S A. 2003;18:10393–8.

11. Cheang MC, Chia SK, Voduc D, et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. J Natl Cancer Inst. 2009;101(10):736–50.

12. Harbeck N, Penault-Llorca F, Cortes J, et al. Breast Cancer. Nat Rev Dis Primers. 2019;5(1):66.

13. Goldhirsh A, Winer EP, Coates AS, et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the primary therapy of early breast cancer 2013. Ann Oncol. 2013;24:2206-23.

14. Elston CW, Ellis IO. Pathologic prognostic factors in breast cancer. I. The value of histological grades in breast cancer. Experience from a large study with long-term follow-up. Histopathology. 1991;19:403-10.

15. Sobin LH, Gospodarowicz MK, Wittekind C. TNM classification of malignant tumours. 7th ed. UICC Wiley-Blackwell, 2009.

16. Fritz A, Percy C, Jack A, et al. International classification of diseases for oncology (ICD-O-3). 3rd ed. Geneva: WHO, 2000.

17. Lambein K, Guiot Y, Galant C, et al. Update of the Belgian guidelines for HER2 testing in breast cancer. Belg J Med Oncol. 2014;8:109-15.

18. Wolff AC, Hammond ME, Hicks DG, et al. Recommendations for Human Epidermal Growth Factor Receptor 2 testing in breast cancer. Arch Pathol Lab Med. 2014;138(2):241-56.

19. Brouckaert O, Laenen A, Vanderhaegen J, et al. Applying the 2011 St Gallen panel of prognostic markers on a large single hospital cohort of consecutively treated primary operable breast cancers. Ann Oncol. 2012;23(10):2578-84.

20. Agresti A. Categorical Data Analysis. 2nd ed. Chapter 3.3: Following-up chi-squared tests. John Wiley & Sons, 2002.

21. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Amer Statist Ass. 1958;53(282):457–81.

22. Ederer F, Axtell LM, Cutler SJ. The relative survival rate: A statistical methodology. National Cancer Institute Monograph. 1962;6:101-21.

23. Howlader N, Altekruse SF, Li CI, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst. 2014;106(5).

24. Cortet M, Bertaut A, Molinié F, et al. Trends in molecular subtypes of breast cancer : description of incidence rates between 2007 and 2012 from three French registries. BMC Cancer. 2018;18:161.

25. van Maaren MC, de Munck L, Strobbe LJ, et al. Ten-year recurrence rates for breast cancer subtypes in the Netherlands: A large population-based study. Int J Cancer. 2019;144:263-72.

26. Minicozzi P, Bella F, Toss A, et al. Relative and disease-free survival for breast cancer in relation to subtype: a population-based study. J Cancer Res Clin Oncol. 2013;139(9):1569-77.

27. Hennigs A, Riedel F, Gondos A, et al. Prognosis of breast cancer molecular subtypes in routine clinical care: a large prospective cohort study. BMC Cancer. 2016;16:734.

28. Jenkins EO, Deal AM, Anders CK, et al. Age-specific changes in intrinsic breast cancer subtypes: a focus on older women. Oncologist. 2014;19:1076-83.

29. Anders CK, Hsu DS, Broadwater G, et al. Young age at diagnosis correlates with worse prognosis and defines a subset of breast cancers with shared patterns of gene expression. J Clin Oncol. 2008;26:3324-30.

SUPPLEMENTARY TABLE 1. Patient and tumour characteristics for female breast cancer patients diagnosed in 2014, bilateral breast cancers and sarcoma excluded.

	Total, all patients (N=10,082)		Total with known molecular subtype (N=9,538)		Unknown molecular subtype (N=544)	
Characteristic	**N**	**%**	**N**	**%**	**N**	**%**
Age at Diagnosis						
Mean (SD)	62	14.2	62	14.0	65	16.1
Median (IQR)	62	52-73	62	52-73	65	53-79
Age Categories						
<40 years	490	4.9	463	4.9	27	5.0
40-49 years	1,547	15.3	1,475	15.5	72	13.2
50-59 years	2,307	22.9	2,204	23.1	103	18.9
60-69 years	2,539	25.2	2,424	25.4	115	21.1
70-79 years	1,860	18.4	1,762	18.5	98	18.0
≥80 years	1,339	13.3	1,210	12.7	129	23.7
Tumour Focality						
Unifocal	5,700	56.5	5,535	58.0	165	30.3
Multifocal	1,476	14.6	1,417	14.9	59	10.8
Unknown	2,906	28.8	2,586	27.1	320	58.8
Tumour Grade						
GI, well differentiated	1,893	18.8	1,843	19.3	50	9.2
GII, moderately differentiated	4,698	46.6	4,577	48.0	121	22.2
GIII, poorly differentiated	3,148	31.2	3,044	31.9	104	19.1
Unknown	343	3.4	74	0.8	269	49.4
Histological Subtype						
IDC	7,580	75.2	7,261	76.1	319	58.6
ILC	1,462	14.5	1,356	14.2	106	19.5
Mixed ductal & lobular	347	3.4	336	3.5	11	2.0
Mixed ductal	186	1.8	177	1.9	9	1.7
Mixed lobular	48	0.5	43	0.5	5	0.9
Mucinous	98	1.0	85	0.9	13	2.4
Papillary	58	0.6	54	0.6	4	0.7
Micropapillary	24	0.2	22	0.2	2	0.4
Medullar	32	0.3	30	0.3	2	0.4
Metaplastic	42	0.4	41	0.4	1	0.2
Tubular	19	0.2	15	0.2	4	0.7
Paget disease	39	0.4	36	0.4	3	0.6
Phyllodes	14	0.1	3	0.0	11	2.0
Other carcinoma types	133	1.3	79	0.8	54	9.9

SD= Standard Deviation, IQR= Inter-quartile Range, IDC= Invasive Carcinoma of no special type (NST), ILC= Invasive Lobular Carcinoma.

SUPPLEMENTARY TABLE 2. Detail of clinical stage (cStage) and cTNM categories for female breast cancer patients diagnosed in 2014, per molecular subtype, bilateral breast cancers and sarcoma excluded.

	Total, all patients (N=10,082)		Total, known mol. subtypes (N=9,538)		LumA (N=5,485)		LumB, HER2- (N=1,484)		LumB, HER2+ (N=1,235)		HER2+, non-lum (N=468)		TN (N=866)		Unknown mol. subtype (N=544)	
cTNM	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Clinical Stage (cStage)																
0 (*)	103	1.0	86	0.9	47	0.9	6	0.4	18	1.5	11	2.4	4	0.5	17	3.1
IA	3,728	37.0	3,623	38.0	2,550	46.5	399	26.9	384	31.1	84	17.9	206	23.8	105	19.3
IB	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
IIA	2,309	22.9	2,225	23.3	1,091	19.9	432	29.1	331	26.8	107	22.9	264	30.5	84	15.4
IIB	853	8.5	830	8.7	312	5.7	172	11.6	161	13.0	63	13.5	122	14.1	23	4.2
IIIA	239	2.4	224	2.3	76	1.4	40	2.7	41	3.3	29	6.2	38	4.4	15	2.8
IIIB	308	3.1	276	2.9	118	2.2	63	4.2	30	2.4	26	5.6	39	4.5	32	5.9
IIIC	82	0.8	76	0.8	21	0.4	21	1.4	12	1.0	11	2.4	11	1.3	6	1.1
IV	555	5.5	477	5.0	208	3.8	94	6.3	92	7.4	47	10.0	36	4.2	78	14.3
X	1,891	18.8	1,718	18.0	1,062	19.4	257	17.3	166	13.4	90	19.2	143	16.5	173	31.8
NA	14	0.1	3	0.0	0	0.0	0	0.0	0	0.0	0	0.0	3	0.3	11	2.0
cT Category																
is	103	1.0	86	0.9	47	0.9	6	0.4	18	1.5	11	2.4	4	0.5	17	3.1
1	1,116	11.1	1,085	11.4	741	13.5	121	8.2	111	9.0	34	7.3	78	9.0	31	5.7
1mi	7	0.1	4	0.0	1	0.0	1	0.1	1	0.1	1	0.2	0	0.0	3	0.6
1a	140	1.4	133	1.4	111	2.0	3	0.2	9	0.7	7	1.5	3	0.3	7	1.3
1b	898	8.9	869	9.1	673	12.3	64	4.3	80	6.5	14	3.0	38	4.4	29	5.3
1c	1,867	18.5	1,814	19.0	1,167	21.3	251	16.9	226	18.3	46	9.8	124	14.3	53	9.7
2	2,947	29.2	2,831	29.7	1,278	23.3	567	38.2	465	37.7	174	37.2	347	40.1	116	21.3
3	468	4.6	449	4.7	183	3.3	94	6.3	78	6.3	39	8.3	55	6.4	19	3.5
4	237	2.4	216	2.3	90	1.6	57	3.8	28	2.3	18	3.8	23	2.7	21	3.9
4a	23	0.2	21	0.2	15	0.3	1	0.1	2	0.2	3	0.6	0	0.0	2	0.4
4b	115	1.1	99	1.0	50	0.9	22	1.5	8	0.6	8	1.7	11	1.3	16	2.9
4c	19	0.2	14	0.1	9	0.2	1	0.1	2	0.2	1	0.2	1	0.1	5	0.9
4d	127	1.3	114	1.2	27	0.5	25	1.7	22	1.8	16	3.4	24	2.8	13	2.4
x	2,015	20.0	1,803	18.9	1,093	19.9	271	18.3	185	15.0	96	20.5	158	18.2	212	39.0
cN Category																
0	5,800	57.5	5,614	58.9	3,526	64.3	788	53.1	690	55.9	193	41.2	417	48.2	186	34.2
1	1,505	14.9	1,419	14.9	535	9.8	274	18.5	263	21.3	130	27.8	217	25.1	86	15.8
2	135	1.3	118	1.2	37	0.7	26	1.8	17	1.4	16	3.4	22	2.5	17	3.1
2a	25	0.2	22	0.2	7	0.1	4	0.3	4	0.3	4	0.9	3	0.3	3	0.6
2b	2	0.0	2	0.0	1	0.0	1	0.1	0	0.0	0	0.0	0	0.0	0	0.0
3	78	0.8	72	0.8	25	0.5	16	1.1	10	0.8	10	2.1	11	1.3	6	1.1
3a	17	0.2	16	0.2	5	0.1	4	0.3	4	0.3	2	0.4	1	0.1	1	0.2
3b	14	0.1	10	0.1	2	0.0	5	0.3	1	0.1	2	0.4	0	0.0	4	0.7
3c	22	0.2	18	0.2	5	0.1	5	0.3	5	0.4	2	0.4	1	0.1	4	0.7
x	2,484	24.6	2,247	23.6	1,342	24.5	361	24.3	241	19.5	109	23.3	194	22.4	237	43.6
cM Category																
0	6,959	69.0	6,719	70.4	3,853	70.2	1,060	71.4	876	70.9	310	66.2	620	71.6	240	44.1
1	555	5.5	477	5.0	208	3.8	94	6.3	92	7.4	47	10.0	36	4.2	78	14.3
x	2,568	25.5	2,342	24.6	1,424	26.0	330	22.2	267	21.6	111	23.7	210	24.2	226	41.5

Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative, X= Unknown, NA= Not Applicable (TNM stage is not applicable, e.g. Phyllodes tumours).

(*) In correspondence with TNM 7th edition, cTis cN0 cM0 tumours are categorised as cStage 0, *i.e.* these tumours were clinically assessed as *in situ* but appeared to be invasive after resection.

SUPPLEMENTARY TABLE 3. Detail of pathological stage (pStage) and pTNM categories for operated female breast cancer patients diagnosed in 2014, who did not receive neo-adjuvant therapy, per molecular subtype, bilateral breast cancers and sarcoma excluded.

	Total, operated patients, no NAT (N=7,528)		Total, known mol. subtypes (N=7,294)		LumA (N=4,576)		LumB, HER2- (N=1,102)		LumB, HER2+ (N=854)		HER2+, non-lum (N=265)		TN (N=497)		Unknown mol. subtype (N=234)	
pTNM	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Pathological Stage (pStage)																
IA	3,802	50.5	3,682	50.5	2,621	57.3	368	33.4	394	46.1	113	42.6	186	37.4	120	51.3
IB	194	2.6	187	2.6	131	2.9	23	2.1	18	2.1	7	2.6	8	1.6	7	3.0
IIA	1,863	24.7	1,806	24.8	1,013	22.1	350	31.8	222	26.0	66	24.9	155	31.2	57	24.4
IIB	833	11.1	821	11.3	439	9.6	180	16.3	102	11.9	36	13.6	64	12.9	12	5.1
IIIA	484	6.4	475	6.5	232	5.1	105	9.5	73	8.5	18	6.8	47	9.5	9	3.8
IIIB	78	1.0	76	1.0	36	0.8	20	1.8	4	0.5	4	1.5	12	2.4	2	0.9
IIIC	206	2.7	200	2.7	81	1.8	48	4.4	32	3.7	20	7.5	19	3.8	6	2.6
IV	19	0.3	18	0.2	11	0.2	4	0.4	2	0.2	0	0.0	1	0.2	1	0.4
X	38	0.5	26	0.4	12	0.3	4	0.4	7	0.8	1	0.4	2	0.4	12	5.1
NA	11	0.1	3	0.0	0	0.0	0	0.0	0	0.0	0	0.0	3	0.6	8	3.4
pT Category (Tumour Size)																
1	13	0.2	9	0.1	4	0.1	4	0.4	1	0.1	0	0.0	0	0.0	4	1.7
1mi	64	0.9	43	0.6	10	0.2	5	0.5	8	0.9	18	6.8	2	0.4	21	9.0
1a	313	4.2	293	4.0	202	4.4	19	1.7	31	3.6	26	9.8	15	3.0	20	8.5
1b	1,228	16.3	1,191	16.3	944	20.6	76	6.9	109	12.8	25	9.4	37	7.4	37	15.8
1c	3,066	40.7	2,996	41.1	1,983	43.3	397	36.0	353	41.3	83	31.3	180	36.2	70	29.9
2	2,363	31.4	2,313	31.7	1,194	26.1	512	46.5	301	35.2	96	36.2	210	42.3	50	21.4
3	324	4.3	318	4.4	184	4.0	61	5.5	33	3.9	11	4.2	29	5.8	6	2.6
4	35	0.5	35	0.5	17	0.4	6	0.5	4	0.5	1	0.4	7	1.4	0	0.0
4a	2	0.0	2	0.0	1	0.0	0	0.0	0	0.0	0	0.0	1	0.2	0	0.0
4b	58	0.8	55	0.8	23	0.5	15	1.4	4	0.5	4	1.5	9	1.8	3	1.3
4c	1	0.0	1	0.0	0	0.0	0	0.0	1	0.1	0	0.0	0	0.0	0	0.0
4d	3	0.0	2	0.0	0	0.0	2	0.2	0	0.0	0	0.0	0	0.0	1	0.4
x	58	0.8	36	0.5	14	0.3	5	0.5	9	1.1	1	0.4	7	1.4	22	9.4
pN Category (Nodal Status)																
0	4,961	65.9	4,833	66.3	3,205	70.0	636	57.7	539	63.1	150	56.6	303	61.0	128	54.7
1	277	3.7	264	3.6	158	3.5	56	5.1	23	2.7	9	3.4	18	3.6	13	5.6
1mi	314	4.2	306	4.2	202	4.4	49	4.4	28	3.3	11	4.2	16	3.2	8	3.4
1a	1,017	13.5	996	13.7	550	12.0	177	16.1	138	16.2	55	20.8	76	15.3	21	9.0
1b	6	0.1	6	0.1	3	0.1	0	0.0	2	0.2	0	0.0	1	0.2	0	0.0
1c	21	0.3	21	0.3	14	0.3	5	0.5	2	0.2	0	0.0	0	0.0	0	0.0
2	66	0.9	64	0.9	31	0.7	17	1.5	6	0.7	0	0.0	10	2.0	2	0.9
2a	325	4.3	319	4.4	149	3.3	77	7.0	53	6.2	12	4.5	28	5.6	6	2.6
2b	1	0.0	1	0.0	0	0.0	1	0.1	0	0.0	0	0.0	0	0.0	0	0.0
3	34	0.5	33	0.5	18	0.4	6	0.5	4	0.5	1	0.4	4	0.8	1	0.4
3a	170	2.3	165	2.3	64	1.4	42	3.8	27	3.2	17	6.4	15	3.0	5	2.1
3b	3	0.0	2	0.0	0	0.0	1	0.1	0	0.0	1	0.4	0	0.0	1	0.4
3c	2	0.0	2	0.0	0	0.0	0	0.0	1	0.1	1	0.4	0	0.0	0	0.0
x	331	4.4	282	3.9	182	4.0	35	3.2	31	3.6	8	3.0	26	5.2	49	20.9
pM Category																
1	19	0.3	18	0.2	11	0.2	4	0.4	2	0.2	0	0.0	1	0.2	1	0.4
x	7,509	99.7	7,276	99.8	4,565	99.8	1,098	99.6	852	99.8	265	100.0	496	99.8	233	99.6

NAT= Neo-adjuvant Treatment, Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative, X= Unknown, NA= Not Applicable (TNM stage is not applicable, e.g. Phyllodes tumours).

SUPPLEMENTARY TABLE 4. Detail of pathological stage (ypStage) and ypTNM categories for operated female breast cancer patients diagnosed in 2014, who received neo-adjuvant therapy, per molecular subtype, bilateral breast cancers and sarcoma excluded.

	Total, operated patients, with NAT (N=1,116)		Total, known mol. subtypes (N=1,071)		LumA (N=263)		LumB, HER2- (N=183)		LumB, HER2+ (N=226)		HER2+, non-lum (N=133)		TN (N=266)		Unknown mol. subtype (N=45)	
pTNM	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
Pathological Stage (ypStage)																
0	101	9.1	95	8.9	2	0.8	14	7.7	18	8.0	15	11.3	46	17.3	6	13.3
is	44	3.9	41	3.8	2	0.8	7	3.8	14	6.2	12	9.0	6	2.3	3	6.7
IA	232	20.8	222	20.7	47	17.9	25	13.7	56	24.8	29	21.8	65	24.4	10	22.2
IB	14	1.3	14	1.3	1	0.4	2	1.1	2	0.9	2	1.5	7	2.6	0	0.0
IIA	249	22.3	239	22.3	65	24.7	43	23.5	62	27.4	20	15.0	49	18.4	10	22.2
IIB	120	10.8	114	10.6	46	17.5	23	12.6	20	8.8	7	5.3	18	6.8	6	13.3
IIIA	144	12.9	138	12.9	50	19.0	35	19.1	22	9.7	9	6.8	22	8.3	6	13.3
IIIB	30	2.7	30	2.8	11	4.2	8	4.4	4	1.8	1	0.8	6	2.3	0	0.0
IIIC	46	4.1	46	4.3	23	8.7	9	4.9	6	2.7	2	1.5	6	2.3	0	0.0
IV	12	1.1	12	1.1	7	2.7	1	0.5	1	0.4	1	0.8	2	0.8	0	0.0
X	124	11.1	120	11.2	9	3.4	16	8.7	21	9.3	35	26.3	39	14.7	4	8.9
ypT Category (Tumour Size)																
0	111	9.9	104	9.7	2	0.8	16	8.7	18	8.0	16	12.0	52	19.5	7	15.6
is	57	5.1	54	5.0	4	1.5	10	5.5	16	7.1	16	12.0	8	3.0	3	6.7
1	1	0.1	1	0.1	0	0.0	0	0.0	1	0.4	0	0.0	0	0.0	0	0.0
1mi	17	1.5	16	1.5	2	0.8	0	0.0	3	1.3	5	3.8	6	2.3	1	2.2
1a	68	6.1	66	6.2	11	4.2	14	7.7	14	6.2	13	9.8	14	5.3	2	4.4
1b	100	9.0	96	9.0	20	7.6	8	4.4	28	12.4	10	7.5	30	11.3	4	8.9
1c	199	17.8	189	17.6	70	26.6	27	14.8	43	19.0	9	6.8	40	15.0	10	22.2
2	304	27.2	293	27.4	95	36.1	52	28.4	67	29.6	19	14.3	60	22.6	11	24.4
3	80	7.2	78	7.3	33	12.5	26	14.2	6	2.7	8	6.0	5	1.9	2	4.4
4	8	0.7	8	0.7	4	1.5	2	1.1	1	0.4	1	0.8	0	0.0	0	0.0
4a	3	0.3	3	0.3	1	0.4	1	0.5	0	0.0	1	0.8	0	0.0	0	0.0
4b	23	2.1	23	2.1	9	3.4	6	3.3	3	1.3	0	0.0	5	1.9	0	0.0
4c	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
4d	3	0.3	3	0.3	0	0.0	1	0.5	1	0.4	0	0.0	1	0.4	0	0.0
x	142	12.7	137	12.8	12	4.6	20	10.9	25	11.1	35	26.3	45	16.9	5	11.1
ypN Category (Nodal Status)																
0	443	39.7	425	39.7	76	28.9	58	31.7	104	46.0	61	45.9	126	47.4	18	40.0
1	63	5.6	61	5.7	18	6.8	15	8.2	7	3.1	4	3.0	17	6.4	2	4.4
1mi	23	2.1	22	2.1	3	1.1	3	1.6	5	2.2	4	3.0	7	2.6	1	2.2
1a	174	15.6	164	15.3	63	24.0	34	18.6	41	18.1	9	6.8	17	6.4	10	22.2
1b	1	0.1	1	0.1	1	0.4	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
1c	3	0.3	3	0.3	1	0.4	2	1.1	0	0.0	0	0.0	0	0.0	0	0.0
2	27	2.4	25	2.3	7	2.7	7	3.8	6	2.7	1	0.8	4	1.5	2	4.4
2a	109	9.8	106	9.9	39	14.8	25	13.7	16	7.1	6	4.5	20	7.5	3	6.7
2b	1	0.1	1	0.1	1	0.4	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
3	8	0.7	8	0.7	4	1.5	1	0.5	1	0.4	1	0.8	1	0.4	0	0.0
3a	39	3.5	39	3.6	21	8.0	8	4.4	4	1.8	1	0.8	5	1.9	0	0.0
3c	1	0.1	1	0.1	0	0.0	0	0.0	1	0.4	0	0.0	0	0.0	0	0.0
x	224	20.1	215	20.1	29	11.0	30	16.4	41	18.1	46	34.6	69	25.9	9	20.0
ypM Category																
1	12	1.1	12	1.1	7	2.7	1	0.5	1	0.4	1	0.8	2	0.8	0	0.0
x	1,104	98.9	1,059	98.9	256	97.3	182	99.5	225	99.6	132	99.2	264	99.2	45	100.0

NAT= Neo-adjuvant Treatment, Lum A= luminal A-like, Lum B HER2-= luminal B-like HER2 negative, Lum B HER2+= luminal B-like HER2 positive, HER2+ non-lum= HER2 positive non-luminal, TN= triple negative, X= Unknown. Note that a distinction was made between ypStage 0, *i.e.* complete pathological response after NAT (ypT0 ypN0,x ypM0,x) and ypStage is, *i.e. in situ* component remains after NAT (ypTis, ypN0,x ypM0,x).