

# BMJ Open Characteristics and prognostic values of traditional pathological parameters and advanced molecular subtypes in women in Beijing with operable breast cancer: a retrospective analysis

Qin Li,<sup>1</sup> Li Li,<sup>1</sup> Xiaoyue Jiang,<sup>1</sup> Qi Du,<sup>1</sup> Yingrui Li,<sup>2</sup> Teng Li,<sup>1</sup> Hong Gong,<sup>3</sup> Bangwei Cao<sup>1</sup>

**To cite:** Li Q, Li L, Jiang X, *et al.* Characteristics and prognostic values of traditional pathological parameters and advanced molecular subtypes in women in Beijing with operable breast cancer: a retrospective analysis. *BMJ Open* 2018;**8**:e021819. doi:10.1136/bmjopen-2018-021819

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2018-021819>).

Received 21 January 2018  
Revised 3 September 2018  
Accepted 4 October 2018



© Author(s) (or their employer(s)) 2018. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

<sup>1</sup>Department of Cancer center, Beijing Friendship Hospital, Capital Medical University, Beijing, China

<sup>2</sup>Department of Biochemistry and Molecular biology, Basic Medical College, Shanxi Medical University, Taiyuan, Shanxi, China

<sup>3</sup>Department of Surgery, Beijing Changping Hospital of Traditional Chinese Medicine, Beijing, China

## Correspondence to

Dr Bangwei Cao;  
[oncology@ccmu.edu.cn](mailto:oncology@ccmu.edu.cn)

## ABSTRACT

**Objectives** This study investigated the characteristics and prognostic values of traditional pathological parameters and advanced molecular subtypes in women with operable breast cancer in Beijing.

**Design** A retrospective study through case information enquiry or telephonic follow-up.

**Setting** Beijing Friendship Hospital.

**Participants** 1042 patients with primary operable breast cancer between 2008 and 2012 were enrolled in the study.

**Measures** The characteristics and 5-year relapse rates according to the Nottingham Prognosis Index (NPI) and molecular subtypes were analysed.

**Results** In 1042 patients, the percentages of high histological grade, N1+N2, T2+T4 were 7.3%, 24.2%, 46.9%, respectively. In patients with invasive breast cancer, the percentages of auxiliary staging, positive margins, vascular invasion and nerve infiltration were 65.0%, 2.8%, 10.5% and 1.1%, respectively. The missing percentages of auxiliary staging, margins, vascular tumour invasion and nerve infiltration were 14.2%, 31.4%, 46.5% and 97.4%, respectively. The percentages of ER-positive, PR-positive, HER2-positive and Ki-67 high expression were 64.3%, 43.8%, 18.8% and 62.7%, respectively. The percentages of luminal A, luminal B, HER2-overexpression and basal-like breast cancers were 10.5%, 54.2%, 8.2% and 11.2%, respectively. Luminal A, luminal B and basal-like breast cancer subtypes were more common in the >60 years group, the 41–60 years group and the 20–40 years group, respectively. The 5-year relapse rates according to NPI were as follows: 6.2% in the low recurrence risk group, 10.4% in the moderate recurrence risk group and 12.9% in the high recurrence risk group. The 5-year relapse rates according to molecular subtypes were as follows: luminal A 4.0%, luminal B 7.0%, HER2-overexpression 14.2%, basal-like 15.6%.

**Conclusions** Reasonable analysis of traditional pathological parameters and advanced molecular subtypes in women with operable breast cancer in Beijing may be useful to guide precise treatment and predict prognosis.

## Strengths and limitations of this study

- The characteristics of traditional pathological parameters and advanced molecular subtypes of operable breast cancer in women in Beijing were compared.
- The 5-year relapse rates according to the Nottingham Prognosis Index were reported.
- The 5-year relapse rates according to molecular subtypes were reported.
- The study was retrospective, and a perspective study is expected.
- It was conducted in a single institution, and multi-centre studies are ongoing.

## INTRODUCTION

Breast cancer is the most common cause of cancer death in women, with approximately 1.67 million cases diagnosed worldwide in 2012.<sup>1</sup> Breast cancer is a highly heterogeneous disease. Rational analysis of pathological characteristics is useful for judging the prognosis of patients with breast cancer. Traditional pathological markers including node staging,<sup>2–3</sup> positive margin,<sup>4,5</sup> vascular tumour invasion,<sup>6</sup> differentiation grade<sup>3,7</sup> and lymph vessel tumour embolus grade<sup>3,8</sup> have been verified as independent risk factors for the recurrence and prognosis. Oestrogen receptor (ER) and progesterone receptor (PR) have been included in routine pathological practice, and are used to predict the patients' course of disease and response to adjuvant hormonal therapy.<sup>9–11</sup> The Nottingham Prognosis Index (NPI) integrates the size of the lesion, the number of involved lymph nodes and the grade of the tumour, which is often used to determine the prognosis of postoperative patients with breast cancer,<sup>12–14</sup> although it is sometimes controversial.

In recent years, more and more research supports the detection of multiple genes (21-gene signature, 70-gene signature, TP53 mutation-correlated genes) in patients with breast cancer.<sup>5–18</sup> Multigene assays could subdivide patients into high-risk and low-risk cohorts thereby providing prognostic and predictive decisions. However, the cost of these multigene assays remains prohibitive for many societies, and it can't be carried out on a large scale.<sup>19</sup> So experts propose that molecular subtypes can be replaced by pathology parameters. In 2013, the St Gallen Consensus Conference and European Society for Medical Oncology (ESMO) Clinical Practice Guidelines recommended surrogate definitions of intrinsic subtypes of breast cancer.<sup>20</sup> According to the ER, PR, HER2 and ki67 status, breast cancer is divided into four subtypes: luminal A, luminal B, HER2-overexpression and basal-like. Understanding these molecular subtypes means a big step forward for the individual precise treatment and prediction of recurrence risk.<sup>21–23</sup> Although the immunohistochemical parameters are not as accurate as multigene assays, the simpler detection method and lower cost are easily accepted by most patients.

Although these molecular subtypes have been theoretically accepted, large-scale data on molecular subtype classification and pathological characteristics associated with different age groups in the population of Beijing have not been systematically studied. Therefore, the present study was carried out to investigate traditional pathological markers and advanced molecular subtypes in women in Beijing with operable breast cancer.

## MATERIALS AND METHODS

### Study design

We retrospectively collected data on all patients (n=1042) with primary operable breast cancer between January 2008 and December 2012 at the Beijing Friendship Hospital. Patients with benign diseases of the breast or metastatic breast cancer were excluded. Biopsies or surgical resection specimens were pathologically examined and histologically confirmed, and complete clinical and pathological records were available. Pathological parameters included tumour location, operation type, distance from the cutting edge, positive margins, vascular tumour invasion, nerve infiltration, histological grade (G), primary tumour (T), lymph nodes (N), histopathological type, and ER, PR, HER2 and Ki67 status. Written informed consent was obtained from all participants.

### The observation end points

All the patients were followed up and 5-year relapse rates were calculated in some of the patients. The follow-up approach involved checking of hospital medical records and outpatient medical records, and contacting the patients/family members for recurrence information. All the patients with primary operable breast cancer were retrospectively collected at Beijing Friendship Hospital. Informed consent was signed by all patients.

### Patient and public involvement

The patients and/or public were not involved in the study design or conduct of the study. Free clinical medical support was provided to all the patients during the follow-up process, for example, the related medical questions were answered, the patients were given guidance on the follow-up plan, and were advised on the next-step therapeutic regimen if recurrence occurred.

### Diagnosis criterion of traditional pathological markers

T, N, G and histopathological type were collected and classified according to the American Joint Committee on Cancer TNM Staging System for Breast Cancer (National Comprehensive Cancer Network Guidelines V.2.2015 for Breast Cancer). G was centrally performed on whole sections according to the recommendations of Nottingham combined with histological grade (Elston-Ellis modification of Scarff-Bloom-Richardson grading system).<sup>24 25</sup>

Vascular tumour invasion was assessed on H&E-stained whole sections of primary tumours. Blood/lymph vessels were identified morphologically and were carefully differentiated from breast ducts/retraction tissue. Tumour cells within vessels mostly formed clusters of various sizes. However, at least one single tumour within a vessel was scored as vascular tumour cell infiltration, if conclusive tumour cell morphology was present.

ER, PR and Ki67 statuses were determined by immunohistochemical staining. Tumours were considered HER2-positive if they were scored 3+ by immunohistochemical staining or if they were 2+ by immunohistochemical staining and also HER2-amplified (ratio >2.0) on the basis of fluorescence in situ hybridisation.

### Surrogate definitions for molecular subtypes of breast cancer

Four molecular subtypes (luminal A, luminal B, HER2-overexpression and basal-like) were classified. Table 1 shows the surrogate definitions of molecular subtypes of breast cancer according to the 2013 St Gallen Consensus Conference and ESMO Clinical Practice Guidelines.<sup>20</sup>

### The judgement criterion for the recurrence risk

For each eligible patient, NPI was calculated using the formula  $NPI = (0.2 \times S) + N + G$ . In this formula, S is the tumour size in cm, N is the number of lymphatic nodes involved (>4 = 3, 4–1=2, 0=1) and G is the degree of malignancy of the tumour (degree 3=3, degree 2=2, degree 1=1). Based on the numerical score obtained from the formula, patients are placed in one of the prognosis groups, good prognostic/low recurrence risk: 2.00–3.40, moderate prognostic/moderate recurrence risk: 3.41–5.40, poor prognostic/high recurrence risk: >5.41.<sup>12–14</sup>

### Follow-up and statistical analysis

The actual 5-year relapse rates have been recorded for 203 patients. The deadline of follow-up was 31 December 2016. Disease free survival (DFS) was defined as the period from the date of diagnosis to occurrence of

**Table 1** Surrogate definitions of molecular subtypes of breast cancer

Molecular subtypes	Luminal A	Luminal B	HER2-overexpression	Basal-like
Histopathological surrogate definition	<ul style="list-style-type: none"> <li>▶ ER-positive</li> <li>▶ HER2-negative</li> <li>▶ Ki67 low</li> <li>▶ PR high*</li> </ul>	<ul style="list-style-type: none"> <li>HER2-negative</li> <li>▶ ER-positive</li> <li>▶ HER2-negative</li> <li>▶ and either               <ul style="list-style-type: none"> <li>– Ki67 high† or</li> <li>– PR low</li> </ul> </li> <li>▶ HER2-positive</li> <li>▶ ER-positive</li> <li>▶ HER2-positive</li> <li>▶ any Ki67</li> <li>▶ any PR</li> </ul>	<ul style="list-style-type: none"> <li>HER2-positive (non-luminal)</li> <li>▶ HER2-positive</li> <li>▶ ER and PR absent</li> </ul>	<ul style="list-style-type: none"> <li>Triple-negative (ductal)</li> <li>▶ ER and PR absent</li> <li>▶ HER2-negative</li> </ul>

\*The cut-off value is 20% for PR high expression.

†The cut-off value is 14% for Ki67 high expression.

ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

any event such as progression, recurrence, metastasis or death. Only patients with invasive breast cancer were included in the prognostic analysis. All data were analysed using the SPSS Statistics software (V.13.0; Chicago, Illinois, USA). Comparisons were determined using the  $\chi^2$  test, Fisher's exact test or independent *t*-test. A *p* value <0.05 was considered statistically significant.

## RESULTS

### Distribution features of age

In the study, the average age of the patients was 55.56±12.37 years (range 22 to 92 years). Of these patients 115 (11.0%) were 20–40 years, 599 (57.5%) patients were 41–60 years and 328 (31.5%) patients were older than 61 years.

### Distribution features of pathological parameters

In 1042 patients the percentages of high histological grade, N1+N2, T2+T4 were 7.3%, 24.2%, 46.9%, respectively. In patients with invasive breast cancer, the percentages of auxiliary staging, positive margins, vascular invasion and nerve infiltration were 20.8%, 2.8%, 10.5% and 1.1%, respectively. The missing percentages of auxiliary staging, margins, vascular tumour invasion and nerve infiltration were 14.2%, 31.4%, 46.5% and 97.4%, respectively. There were significant differences in neoadjuvant chemotherapy, auxiliary staging, tumour size and lymph nodes in patients among the three age groups (20–40 years, 41–60 years and ≥61 years, [table 2](#)). Neoadjuvant chemotherapy was much less in the 41–60 years group. Auxiliary staging, T2+T4 and N1+N2 was much less in the 20–40 years group. There were no statistically significant differences in tumour location, margins, vascular tumour invasion, nerve infiltration, grade (all *p*>0.05). Features of traditional pathological parameters in patients with operable breast cancer are shown in [table 2](#). With regard to histopathological types, 104 (10.0%) patients had ductal carcinoma in situ (DCIS) and 938 (90.0%) patients had invasive carcinoma. There were no significant differences

in histopathological types in patients among the three age groups (20–40 years, 41–60 years and ≥61 years).

### Distribution features of ER/PR/HER2/Ki67 and molecular subtypes

Of 1042 patients, 670 (64.3%) patients were ER-positive and 196 (18.8%) patients were HER2-positive ([figure 1](#)). With a cut-off value of 20%, high expression and low expression of PR were detected in 456 (43.8%) and 105 (10.1%) patients, respectively. With a cut-off value of 14%, high expression and low expression of Ki-67 were detected in 653 (62.7%) and 170 (16.3%) patients, respectively. There was significant difference of Ki67 status among the three age groups (20–0 years, 41–60 years and ≥61 years, *p*=0.025). In HER2-positive tumours, 15.2% of patients were ER-positive and 24% of patients highly expressed Ki-67.

In the population with complete data, 109 (10.5%) patients had luminal A, 565 (54.2%) patients had luminal B, 85 (8.2%) patients had HER2-overexpression and 117 (11.2%) patients has basal-like ([table 3](#), [figure 1](#)) molecular subtype. There was a statistically significant difference in the molecular subtypes among the three age groups (20–40 years, 41–60 years and ≥61 years; *p*=0.038). Luminal A was more common in the >60 years age group, luminal B was more common in the 41–60 years age group and basal-like was more common in the 20–40 years age group ([figure 2](#)).

### Distribution of recurrence risk

Recurrence risk was evaluated based on NPI. Among the 623 patients evaluated, 263 (42.2%) should have good prognostic/low recurrence risk, 312 (50.1%) should have moderate prognostic/moderate recurrence risk and 48 (7.7%) should have poor prognostic/high recurrence risk. However, there was no significant difference in recurrence risk among the three age groups.

The actual 5-year relapse rates of the patients with invasive breast cancers have been recorded in 193 patients. The 5-year relapse rates according to NPI were as

**Table 2** The characteristics of traditional pathological parameters in the different age groups

Pathological parameters	No. of patients (%)				X <sup>2</sup>	P values
	All patients (n=1042)	20–40 years (n=115)	41–60 years (n=599)	≥61 years (n=328)		
<b>In situ and invasive breast cancer</b>						
Tumour size					23.32	0.010
TX	202 (19.4)	39 (33.9)	115 (19.2)	48 (14.6)		
T1	352 (33.7)	34 (29.6)	203 (33.9)	114 (34.8)		
T2	420 (40.3)	38 (33.0)	241 (40.2)	141 (43.0)		
T3	32 (3.1)	2 (1.7)	19 (3.2)	11 (3.3)		
T4	36 (3.5)	2 (1.7)	21 (3.5)	13 (4.0)		
Lymph nodes					22.27	0.001
NX	382 (36.7)	62 (53.9)	206 (34.4)	114 (34.8)		
N0	408 (39.2)	33 (28.7)	230 (38.4)	145 (44.2)		
N1	250 (24.0)	20 (17.4)	162 (27.0)	68 (20.7)		
N2	2 (0.2)	0 (0.0)	1 (0.2)	1 (0.3)		
Grade					8.37	0.212
Not detected	399 (38.3)	53 (46.1)	217 (36.2)	129 (39.3)		
High histological grade	76 (7.3)	9 (7.8)	49 (8.2)	18 (5.5)		
Intermediate histological grade	478 (45.9)	45 (39.1)	286 (47.8)	147 (44.8)		
Low histological grade	89 (8.5)	8 (7.0)	47 (7.8)	34 (10.4)		
<b>Invasive breast cancer</b>						
Auxiliary staging					15.12	0.004
No description	133 (14.2)	20 (19.8)	78 (14.4)	35 (11.9)		
With auxiliary staging	610 (65.0)	49 (48.5)	363 (67.0)	198 (67.1)		
Without auxiliary staging	195 (20.8)	32 (31.7)	101 (18.6)	62 (21.0)		
Margins					9.63	0.055
Not detected	294 (31.4)	44 (43.6)	168 (31.1)	82 (27.8)		
No residual cancer	617 (65.8)	56 (55.4)	358 (66.2)	203 (68.8)		
With residual cancer	26 (2.8)	1 (1.0)	15 (2.8)	10 (3.4)		
Vascular tumour invasion					7.47	0.102
Not detected	436 (46.5)	54 (53.5)	239 (44.2)	143 (48.5)		
No	403 (43.0)	35 (34.7)	238 (44.0)	130 (44.1)		
Yes	98 (10.5)	12 (11.9)	64 (11.8)	22 (7.5)		
Nerve infiltration					4.19	0.380
Not detected	913 (97.4)	98 (97.0)	528 (97.6)	287 (97.3)		
No	14 (1.5)	3 (3.0)	8 (1.5)	1 (1.0)		
Yes	10 (1.1)	0 (0.0)	5 (0.9)	5 (1.7)		

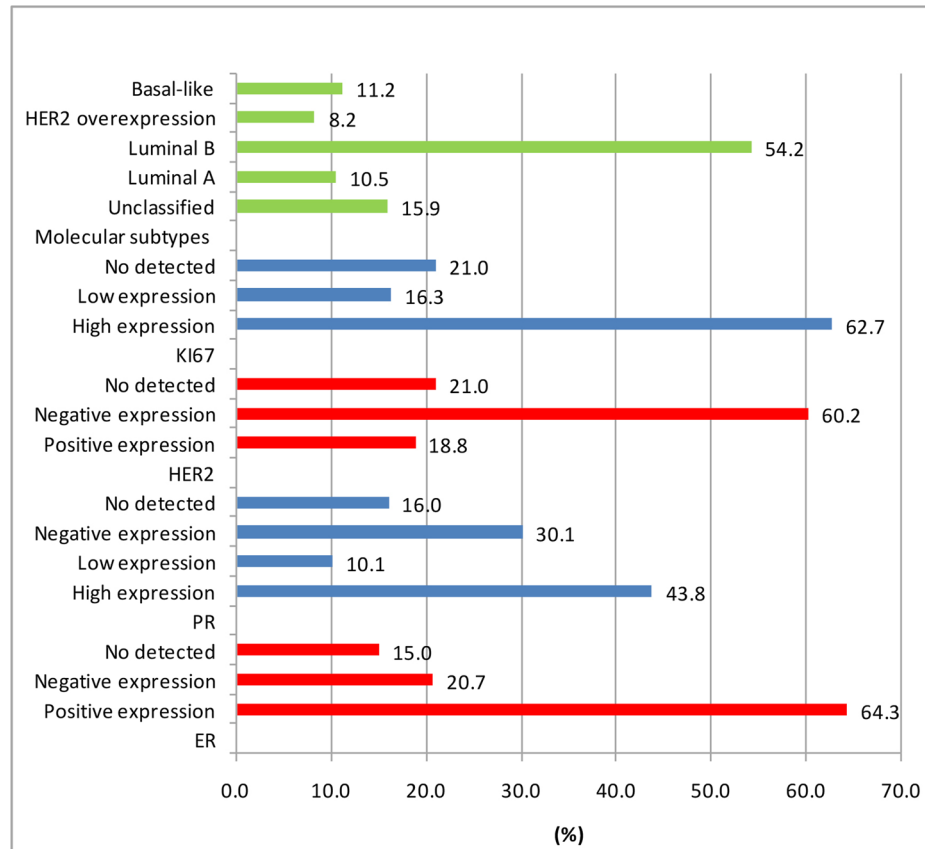
follows: 6.2% in the low recurrence risk group, 10.4% in the moderate recurrence risk group and 12.9% in high recurrence risk group. The 5-year relapse rates according to molecular subtypes were as follows: luminal A 4.0%, luminal B 7.0%, HER2 overexpression 14.2%, basal-like 15.6%.

## DISCUSSION

Traditional pathological parameters including positive margin, vascular tumour invasion, high histological grade

and lymph node staging have been verified as independent risk factors for recurrence and as markers of prognosis.<sup>2–7</sup> Tumour size has been demonstrated to be closely related to relapse-free survivals.<sup>26</sup> Sarsenov *et al* reported that younger age (<40 years), large tumour size (>2 cm), high grade and triple-negative phenotype were identified as independent prognostic factors with a negative impact on overall survival of patients with recurrent breast cancer.<sup>27</sup> In our analysis, the percentages of positive margins, vascular tumour invasion, high histological





**Figure 1** The overall distribution features of ER/PR/HER2/Ki67 and molecular subtypes in all patients. ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

grade, N1+N2 and T2+T4 were 2.8%, 10.5%, 7.3%, 24.2% and 46.9%, respectively. These indicators reflect the percentages of patients with poor prognosis from different perspectives. In our study, the missing percentages of positive margins, vascular tumour invasion, nerve infiltration and grade were up to 31.4%, 46.51%, 97.4% and 38.3%, respectively. The missing data are at random. Accurate analysis and diagnosis of preoperative staging, standardised surgical operation, standardised pathological slice making and handling, comprehensive and accurate interpretation of pathological findings, and comprehensive detection of prerequisite markers will greatly reduce the missing data. These startling missing data raise the strict demands to the surgeons, physicians and pathologists.

DCIS and invasive ductal cancer were the two main histopathological types of breast cancer in patients in Beijing. Julian *et al's* study showed that auxiliary nodal dissection in DCIS is not recommended.<sup>28</sup> In our study, 50% of patients with DCIS received auxiliary staging. Whether patients with DCIS should receive auxiliary staging is a question worthy of discussion. Although patients with DCIS have a favourable prognosis, recurrence risk was increased in high-grade DCIS (OR, 4.39).<sup>29</sup> The DCIS score (12-gene) assay provides clinically relevant information on recurrence risk and may facilitate decision making by clinicians.<sup>30</sup> The percentage of invasive ductal cancers was 90.0% in the entire patient sample, and Hasebe *et al's*

study showed that type 2 invasive ductal cancer was one of the best factors for accurately predicting locoregional recurrence.<sup>8</sup>

ER, PR, Ki67 and HER2 have been routinely applied in clinical practice. ER and PR are associated with good response to hormonal therapy and better clinical outcomes. In our study, the ER-positive rate was 75.6%, which coincided with the results reported by other studies.<sup>31–33</sup> PR-positive rates were 53.9% in all cases and 81.0% in ER-positive patients, which is in agreement with the results reported by Liu *et al.*<sup>34</sup> It has been shown that 5-year adjuvant tamoxifen reduces annual breast cancer death rate by 31% for ER-positive patients.<sup>28</sup> In our study, the high and low expressions of Ki67 were 62.7% and 16.3%, respectively. Ki67 is closely related to cellular proliferation,<sup>35</sup> and a larger decrease in Ki67 indicates better responsiveness to chemotherapy.<sup>36–37</sup> Borderline distribution of Ki67 indicated significantly more distant bone and liver metastases and worse disease-specific survival.<sup>38</sup> In patients with complete data, 23.8% were HER2-positive, which is similar to 25.5% reported by Zhu *et al.*<sup>33</sup> HER2 overexpression is directly proportional to relapse.<sup>39–40</sup> Trastuzumab, a powerful HER2-targeted agent, has dramatically improved the outcomes of patients with HER2-overexpression breast cancer.<sup>41–42</sup>

The distribution features of molecular subtypes were luminal B>basal like>luminal A>HER2 overexpression. Luminal B, HER2-overexpression and basal-like

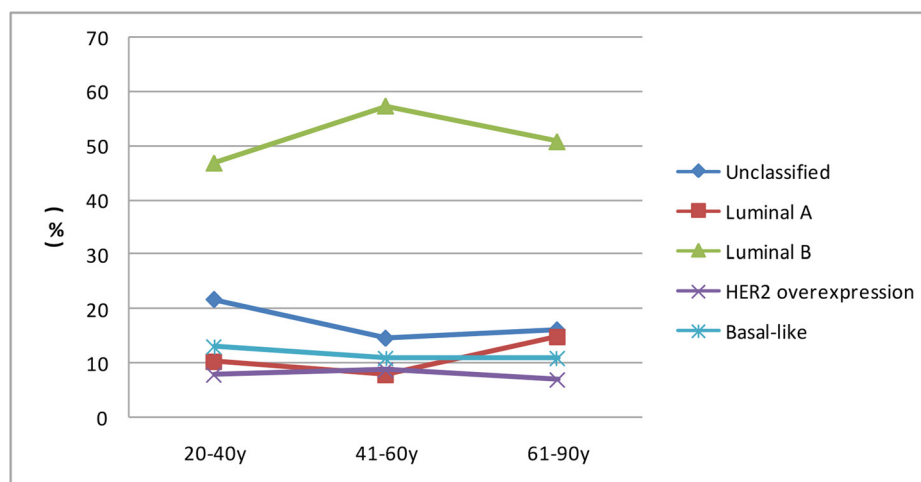
**Table 3** The distribution features of ER/PR/HER2/Ki67 and molecular subtypes in the different age groups

Parameters	No. of patients (%)				X <sup>2</sup>	P values
	All patients (n=1042)	20–40 years (n=115)	41–60 years (n=599)	≥61 years (n=328)		
ER status					<b>3.293</b>	<b>0.510</b>
Positive expression	670 (64.3)	68 (59.1)	384 (64.1)	218 (66.5)		
Negative expression	216 (20.7)	24 (20.9)	128 (21.4)	64 (19.5)		
Not detected	156 (15.0)	23 (20.0)	87 (14.5)	46 (14.0)		
PR status					<b>9.411</b>	<b>0.152</b>
High expression	456 (43.8)	54 (47.0)	257 (42.9)	145 (44.2)		
Low expression	105 (10.1)	5 (4.3)	64 (10.7)	36 (10.8)		
Negative expression	314 (30.1)	30 (26.1)	189 (31.6)	95 (29.0)		
Not detected	167 (16.0)	26 (22.6)	89 (14.9)	52 (15.9)		
HER2 status					<b>10.380</b>	<b>0.110</b>
Positive expression	196 (18.8)	17 (14.8)	128 (21.4)	51 (15.5)		
Negative expression	627 (60.2)	65 (56.5)	351 (58.6)	211 (64.3)		
Not detected	219 (21.0)	33 (28.7)	120 (20.0)	66 (20.1)		
Ki67 status					<b>11.302</b>	<b>0.023</b>
High expression	653 (62.7)	67 (58.3)	398 (66.4)	188 (57.3)		
Low expression	170 (16.3)	17 (14.8)	86 (14.4)	67 (20.4)		
Not detected	219 (21.0)	31 (27.0)	115 (19.2)	73 (22.3)		
Molecular subtype					<b>16.93</b>	<b>0.031</b>
Unclassified	166 (15.9)	25 (21.7)	88 (14.7)	53 (16.2)		
Luminal A	109 (10.5)	12 (10.4)	48 (8.0)	49 (14.9)		
Luminal B	565 (54.2)	54 (47.0)	344 (57.4)	167 (50.9)		
HER2-overexpression	85 (8.2)	9 (7.8)	53 (8.8)	23 (7.0)		
Basal-like	117 (11.2)	15 (13.0)	66 (11.0)	36 (11.0)		

ER, oestrogen receptor; HER2, human epidermal growth factor receptor 2; PR, progesterone receptor.

subtypes were more common in the 41–60 years age group. The distribution of molecular subtypes in our study is consistent with that reported by Si *et al.*<sup>43</sup> Molecular subtypes, as advanced pathological indications, are critical

for predicting prognosis and guiding treatment.<sup>21,22</sup> Voduc *et al* reported that patients with the luminal A subtype have better prognosis than those with HER2-overexpression and basal-like subtypes, as indicated by the relatively



**Figure 2** The distribution features of molecular subtypes in the different age groups. HER2, human epidermal growth factor receptor 2.

low rates of local and regional relapse.<sup>39</sup> The luminal A subtype is very sensitive to endocrine therapy, the luminal B (HER2-) subtype benefits from endocrine therapy or chemotherapy, the luminal B (HER2+) subtype benefits from endocrine therapy or chemotherapy combined with anti-HER2-targeted therapy<sup>43 44</sup> and the HER2-overexpression subtype benefits from chemotherapy combined with anti-HER2-targeted therapy.<sup>40 42 45</sup> The target is lacking in basal-like breast cancer, and combined chemotherapy is the standard treatment option.

NPI is usually used to determine the prognosis of post-operative patients with breast cancer. NPI was calculated using tumour size, positive lymphatic nodes and grade. In our study, the 5-year relapse rates increased with the rise of NPI, suggesting significance of traditional pathological parameters in prognosis. The 5-year relapse rates according to molecular subtypes were as follows: basal-like >HER2 overexpression >luminal B >luminal A, and this is consistent with the results reported by Shim *et al.*<sup>46</sup> However, Arvold *et al* showed that the 5-year cumulative incidence of local relapse was 0.8% in patients with luminal A, 4.4% in those with luminal B, 10.8% in those with HER2-overexpression and 6.7% in those with basal-like subtypes,<sup>47</sup> and the patients with the HER2-overexpression subtype had the worst prognosis. Both the evaluated methods are able to predict recurrence risk and prognosis, however, the latter shows its unique advantages in guiding specific treatment schemes.

In conclusion, our study has shown the features of traditional pathological parameters and advanced molecular subtypes in women in Beijing with operable breast cancer. In-depth understanding of the biological behaviour of breast cancer would be beneficial for oncologists to guide treatment, identify recurrence risk and make reasonable follow-ups. However, our study has several limitations. It was a retrospective study conducted in a single institution with a relatively small sample. At present, we are carrying out a study on molecular subtypes and recurrence risk in a larger population in China, and the results are awaited.

**Acknowledgements** The authors thank the family members of patients for providing the related information on follow-up.

**Contributors** QL, BC designed the study; QD and YL developed the methodology and performed the analyses; XJ, LL, TL, QD collected the data; QL and HG analysed the data; and QL wrote the first draft. All the authors contributed to the review and revision of the manuscript, and all authors read and approved the final manuscript.

**Funding** This work was supported by National Natural Science Foundation of China (Grant NO. 81301912 and 81272615), Beijing Municipal Health System High-level Health Person Foundation Project (Grant NO. 2014-3-005), Beijing Municipal Science and Technology Commission (Capital Features, Z161100000516083, to QL).

**Competing interests** None declared.

**Patient consent** Not required.

**Ethics approval** The study was approved by the ethics committee of Beijing Friendship Hospital.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** All data are available from author QL.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which

permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

1. Ferlay J, Soerjomataram I, Dikshit R, *et al.* Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 2015;136:E359–E386.
2. Solak M, Turkoz FP, Keskin O, *et al.* The lymph node ratio as an independent prognostic factor for non-metastatic node-positive breast cancer recurrence and mortality. *J Buon* 2015;20:737–45.
3. Joshi S, Watkins J, Gazinska P, *et al.* Digital imaging in the immunohistochemical evaluation of the proliferation markers Ki67, MCM2 and Geminin, in early breast cancer, and their putative prognostic value. *BMC Cancer* 2015;15:546.
4. Voguet L, Hébert T, Levêque J, *et al.* Patient age and positive margins are predictive factors of residual tumor on mastectomy specimen after conservative treatment for breast cancer. *Breast* 2009;18:233–7.
5. Houssami N, Macaskill P, Marinovich ML, *et al.* The association of surgical margins and local recurrence in women with early-stage invasive breast cancer treated with breast-conserving therapy: a meta-analysis. *Ann Surg Oncol* 2014;21:717–30.
6. Kurz C, Obermair A, Nagele F, *et al.* [The value of the "vascular invasion" factor for prognosis of breast cancer]. *Geburtshilfe Frauenheilkd* 1994;54:295–9.
7. Contesso G, Mouriessse H, Friedman S, *et al.* The importance of histologic grade in long-term prognosis of breast cancer: a study of 1,010 patients, uniformly treated at the Institut Gustave-Roussy. *J Clin Oncol* 1987;5:1378–86.
8. Hasebe T, Iwasaki M, Hojo T, *et al.* Histological factors for accurately predicting first locoregional recurrence of invasive ductal carcinoma of the breast. *Cancer Sci* 2013;104:1252–61.
9. Rakha EA, Reis-Filho JS, Ellis IO. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res Treat* 2010;120:293–308.
10. Althuis MD, Fergenbaum JH, Garcia-Closas M, *et al.* Etiology of hormone receptor-defined breast cancer: a systematic review of the literature. *Cancer Epidemiol Biomarkers Prev* 2004;13:1558–68.
11. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet* 2005;365:1687–717.
12. Todd JH, Dowle C, Williams MR, *et al.* Confirmation of a prognostic index in primary breast cancer. *Br J Cancer* 1987;56:489–92.
13. Haybittle JL, Blamey RW, Elston CW, *et al.* A prognostic index in primary breast cancer. *Br J Cancer* 1982;45:361–6.
14. Blamey RW, Pinder SE, Ball GR, *et al.* Reading the prognosis of the individual with breast cancer. *Eur J Cancer* 2007;43:1545–7.
15. King TA, Lyman JP, Gonen M, *et al.* Prognostic Impact of 21-Gene Recurrence Score in Patients With Stage IV Breast Cancer: TBCRC 013. *J Clin Oncol* 2016;34:2359–65.
16. Epstein AJ, Wong YN, Mitra N, *et al.* Adjuvant Chemotherapy Use and Health Care Costs After Introduction of Genomic Testing in Breast Cancer. *J Clin Oncol* 2015;33:4259–67.
17. Kuijter A, Bommel A, Drukker CA, *et al.* Using a gene expression signature when controversy exists regarding the indication for adjuvant systemic treatment reduces the proportion of patients receiving adjuvant chemotherapy: a nationwide study[J]. *Genet Med* 2015;60:187–200.
18. Györfy B, Bottai G, Lehmann-Che J, *et al.* TP53 mutation-correlated genes predict the risk of tumor relapse and identify MPS1 as a potential therapeutic kinase in TP53-mutated breast cancers. *Mol Oncol* 2014;8:508–19.
19. Ward S, Scope A, Rafia R, *et al.* Gene expression profiling and expanded immunohistochemistry tests to guide the use of adjuvant chemotherapy in breast cancer management: a systematic review and cost-effectiveness analysis. *Health Technol Assess* 2013;17:1–302.
20. Senkus E, Kyriakides S, Penault-Llorca F, *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24 Suppl 6(Suppl 6):vi7–vi23.
21. Anderson WF, Luo S, Chatterjee N, *et al.* Human epidermal growth factor receptor-2 and estrogen receptor expression, a demonstration project using the residual tissue repository of the Surveillance, Epidemiology, and End Results (SEER) program. *Breast Cancer Res Treat* 2009;113:189–96.

22. Schwartz AM, Henson DE, Patel A. Re: Age-specific incidence of breast cancer subtypes: understanding the Black-White crossover. *J Natl Cancer Inst* 2013;105:368–70.
23. Prat A, Ellis MJ, Perou CM. Practical implications of gene-expression-based assays for breast oncologists. *Nat Rev Clin Oncol* 2011;9:48–57.
24. Harris L, Fritsche H, Mennel R, *et al*. American Society of Clinical Oncology 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol* 2007;25:5287–312.
25. Singletary SE, Allred C, Ashley P, *et al*. Revision of the American Joint Committee on Cancer staging system for breast cancer. *J Clin Oncol* 2002;20:3628–36.
26. Huang S, Yee C, Ching T, *et al*. A novel model to combine clinical and pathway-based transcriptomic information for the prognosis prediction of breast cancer. 2014;10:e1003851.
27. Sarsenov D, Ilgun S, Ordu C, *et al*. True Local Recurrences after Breast Conserving Surgery have Poor Prognosis in Patients with Early Breast Cancer. *Cureus* 2016;8:e541.
28. Julian TB, Land SR, Fourchotte V, *et al*. Is sentinel node biopsy necessary in conservatively treated DCIS? *Ann Surg Oncol* 2007;14:2202–8.
29. Sue GR, Chagpar AB. Predictors of recurrence in patients diagnosed with ductal carcinoma in situ. *Am Surg* 2015;81:48–51.
30. Wood WC, Alvarado M, Buchholz DJ, *et al*. The current clinical value of the DCIS Score. *Oncology* 2014;28 Suppl 2:1C3–18.
31. Dowsett M, Houghton J, Iden C, *et al*. Benefit from adjuvant tamoxifen therapy in primary breast cancer patients according oestrogen receptor, progesterone receptor, EGF receptor and HER2 status. *Ann Oncol* 2006;17:818–26.
32. Caldarella A, Buzzoni C, Crocetti E, *et al*. Invasive breast cancer: a significant correlation between histological types and molecular subgroups. *J Cancer Res Clin Oncol* 2013;139:617–23.
33. Zhu X, Ying J, Wang F, *et al*. Estrogen receptor, progesterone receptor, and human epidermal growth factor receptor 2 status in invasive breast cancer: a 3,198 cases study at National Cancer Center, China. *Breast Cancer Res Treat* 2014;147:551–5.
34. Liu S, Chia SK, Mehl E, *et al*. Progesterone receptor is a significant factor associated with clinical outcomes and effect of adjuvant tamoxifen therapy in breast cancer patients. *Breast Cancer Res Treat* 2010;119:53–61.
35. Denkert C, Budczies J, von Minckwitz G, *et al*. Strategies for developing Ki67 as a useful biomarker in breast cancer. *Breast* 2015;24 Suppl 2(Suppl 2):S67–S72.
36. Burcombe RJ, Makris A, Richman PI, *et al*. Evaluation of ER, PgR, HER-2 and Ki-67 as predictors of response to neoadjuvant anthracycline chemotherapy for operable breast cancer. *Br J Cancer* 2005;92:147–55.
37. Burcombe R, Wilson GD, Dowsett M, *et al*. Evaluation of Ki-67 proliferation and apoptotic index before, during and after neoadjuvant chemotherapy for primary breast cancer. *Breast Cancer Res* 2006;8:R31.
38. Gong P, Wang Y, Liu G, *et al*. New insight into Ki67 expression at the invasive front in breast cancer. *PLoS One* 2013;8:e54912.
39. Voduc KD, Cheang MC, Tyldesley S, *et al*. Breast cancer subtypes and the risk of local and regional relapse. *J Clin Oncol* 2010;28:1684–91.
40. Perez EA, Romond EH, Suman VJ, *et al*. Four-year follow-up of trastuzumab plus adjuvant chemotherapy for operable human epidermal growth factor receptor 2-positive breast cancer: joint analysis of data from NCCTG N9831 and NSABP B-31. *J Clin Oncol* 2011;29:3366–73.
41. Valachis A, Mauri D, Polyzos NP, *et al*. Trastuzumab combined to neoadjuvant chemotherapy in patients with HER2-positive breast cancer: a systematic review and meta-analysis. *Breast* 2011;20:485–90.
42. Si C, Jin Y, Wang H, *et al*. Association between molecular subtypes and lymph node status in invasive breast cancer. *Int J Clin Exp Pathol* 2014;7:6800–6.
43. Kaufman B, Mackey JR, Clemens MR, *et al*. Trastuzumab plus anastrozole versus anastrozole alone for the treatment of postmenopausal women with human epidermal growth factor receptor 2-positive, hormone receptor-positive metastatic breast cancer: results from the randomized phase III TAnDEM study. *J Clin Oncol* 2009;27:5529–37.
44. Giordano SH, Temin S, Kirshner JJ, *et al*. Systemic therapy for patients with advanced human epidermal growth factor receptor 2-positive breast cancer: American Society of Clinical Oncology clinical practice guideline. *J Clin Oncol* 2014;32:2078–99.
45. Perez EA, Romond EH, Suman VJ, *et al*. Trastuzumab plus adjuvant chemotherapy for human epidermal growth factor receptor 2-positive breast cancer: planned joint analysis of overall survival from NSABP B-31 and NCCTG N9831. *J Clin Oncol* 2014;32:3744–52.
46. Shim HJ, Kim SH, Kang BJ, *et al*. Breast cancer recurrence according to molecular subtype. *Asian Pac J Cancer Prev* 2014;15:5539–44.
47. Arvold ND, Taghian AG, Niemierko A, *et al*. Age, breast cancer subtype approximation, and local recurrence after breast-conserving therapy. *J Clin Oncol* 2011;29:3885–91.