



Tumor markers CA15-3, CA125, CEA and breast cancer survival by molecular subtype: a cohort study

Junxian Li¹ · Luyang Liu¹ · Ziwei Feng¹ · Xin Wang¹ · Yubei Huang¹ · Hongji Dai¹ · Liwen Zhang¹ · Fangfang Song¹ · Dezheng Wang² · Pengyu Zhang³ · Baoshan Ma⁴ · Haixin Li^{1,5} · Hong Zheng¹ · Fengju Song¹  · Kexin Chen¹

Received: 9 October 2019 / Accepted: 25 January 2020 / Published online: 10 February 2020
© The Japanese Breast Cancer Society 2020

Abstract

Background The burden of breast cancer has grown rapidly in China during recent decades. However, the association between tumor markers (CA15-3, CA125, and CEA) and breast cancer survival among certain molecular subtypes is unclear; we described this association in a large, population-based study.

Methods We conducted a cohort study including 10,836 women according to the Tianjin Breast Cancer Cases Cohort. Demographic and epidemiologic data were collected by a structured face-to-face questionnaire. Clinico-pathological parameters were abstracted from medical records, and follow-up information was obtained once a year by telephone. The primary endpoints were breast cancer-specific survival (BCSS) and disease-free survival (DFS). We utilized the Cox proportional hazard model to calculate hazard ratios (HRs) and 95% confidence intervals (CI).

Results Among all patients, elevated CA15-3 and CEA exhibited consistently and statistically significant reduced BCSS compared with normal ones (CA15-3: HR 1.54, 95% CI 1.01–2.34; CEA: HR 2.45, 95% CI 1.40–4.30). Similar patterns of association were observed for DFS (CA15-3: HR 2.09, 95% CI 1.44–3.02; CEA: HR 2.71, 95% CI 1.71–4.27). Moreover, in luminal A subtype, high CA15-3 and CEA levels were associated with decreased BCSS (CA15-3: HR 4.47, 95% CI 2.04–9.81; CEA: HR 3.79, 95% CI 1.68–8.55) and DFS (CA15-3: HR 4.06, 95% CI 2.29–7.18, CEA: HR 3.41, 95% CI 1.75–6.64). In basal-like subtype, elevated CEA conferred reduction for BCSS (HR 5.13, 95% CI 1.65–15.9). However, no association was observed between CA125 and breast cancer outcome.

Conclusions Preoperative CA15-3 and CEA levels differ in breast cancer molecular subtypes and yield strong prognostic information in Chinese women with breast cancer. Measuring CA15-3 and CEA levels before surgery may have the potential in predicting breast cancer survival and offering patients' personalized treatment strategy among luminal A and basal-like subtypes.

Keywords Breast cancer · Survival · Tumor marker · Molecular subtype · TBCCC

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12282-020-01058-3>) contains supplementary material, which is available to authorized users.

✉ Fengju Song
songfengju@163.com

✉ Kexin Chen
chenkexin@tjmuch.com

¹ Department of Epidemiology and Biostatistics, National Clinical Research Center for Cancer, Key Laboratory of Molecular Cancer Epidemiology, Tianjin Medical University Cancer Institute and Hospital, Huanhu Xi Road, Tiyuan Bei, Hexi District, Tianjin 300060, People's Republic of China

² Tianjin Centers for Disease Control and Prevention, Tianjin 300011, People's Republic of China

³ Department of Clinical Laboratory, National Clinical Research Center for Cancer, Key Laboratory of Molecular Cancer Epidemiology, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, People's Republic of China

⁴ College of Information Science and Technology, Liaoning Province, Dalian Maritime University, Dalian 116026, People's Republic of China

⁵ Department of Cancer Biobank, National Clinical Research Center for Cancer, Key Laboratory of Molecular Cancer Epidemiology, Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, People's Republic of China

Introduction

In the global area, breast cancer is cancer death's main reason in females, and approximately 2.1 million females will be suspected to be new breast cancer cases in 2018, accounting for 25% of female cancers [1]. Breast cancer ranks the first for incidence and the fifth for mortality of female tumors in China, and the burden of breast cancer has grown rapidly in the past several decades [2]. Breast cancer is a heterogeneous tumor with four major molecular subtypes and distinct in prognosis. Nowadays, the selection of treatment and evaluation of prognosis in clinic are mainly based on distinct molecular subtypes.

The prognoses of breast cancer patients are mainly determined by pathological characteristics of tumors, such as stage, grade, human epidermal growth factor receptor 2 (HER2), and hormone-receptor status [3, 4]. The searching for new prognostic markers to improve the performance of survival prediction has never ended, and recently, the influence of tumor markers on prognosis has attracted increasing attention. The cancer antigen 15-3 (CA15-3), a member of the mucin-1 (MUC-1) family of glycoproteins, is over-expressed in cancers and is identified as a useful tumor marker driven by altered glycosylation of itself [5]. In the blood, the elevation of carcinoembryonic antigen (CEA) levels as a kind of cell adhesion molecule is associated with tumor metastasis [6, 7]. The cancer antigen 125 (CA125), the product of the MUC16 gene, is a key regulator of multiple cell survival pathways in ovarian cancer and breast cancer cells [8]. Utilizing CA15-3, CA125, and CEA for predicting breast cancer risk has been widely incorporated in clinical routine, whereas the association of these markers with breast cancer prognosis has not been well demonstrated.

Considerable effort has gone into the correlation between tumor markers and the prognosis of breast cancer. The European Group on Tumor Markers has recommended utilizing CA15-3 and CEA to evaluate a patient's event in the future course of the disease, early detect the progression of disease, and select therapeutic methods in breast cancer [9]. In contrast, the American Society of Clinical Oncology (ASCO) has not recommended the utilization of CEA and CA15-3 in screening, observing the presence of disease, staging, and choosing treatment regimen of breast cancer [10]. Elevated CA125 levels have been accounting for 84% in metastatic breast patients [11]. CA125 was reported to be proved effective in patients' management with breast cancer and could seem to provide predictive information in the course of disease [12]. Maric et al. [13] recently reviewed serum tumor markers' role in breast cancer, and they indicated that more wide investigations concerning prognostic value were supposed to be needed due to inconsistent studies. Additionally, most studies have been based on breast cancer overall, and

the association between these markers and breast cancer outcome by molecular subtype remains to be elucidated clearly.

Therefore, we established a breast cancer patients' specific cohort in China and assessed the role of tumor markers (CA15-3, CA125, and CEA) in breast cancer outcome overall, and specifically for different molecular subtypes.

Materials and methods

Study population and data sources

Tianjin Breast Cancer Cases Cohort (TBCCC) began in 2007; all the participants were pathologically confirmed breast cancer patients diagnosed and treated in the Tianjin Medical University Cancer Institute and Hospital. Demographic and epidemiologic data of participants were collected by one professional full-time personnel using a structured questionnaire shortly after the confirmation of breast cancer. Clinico-pathological and treatment data were abstracted from medical records. Follow-up information and vital status were attained once a year by telephone. Lost to follow-up rate was 6.04% in the study population. Moreover, the database was annually updated by checking medical records and by linking to the Tianjin Cancer Death Registry System to further ascertain the accuracy and completeness. The last obtainable update of vital status was completed on December 31, 2017. The written informed consent was acquired from each breast cancer patient or patient's guardian in TBCCC, and the current study was approved by the research ethics board of the Tianjin Medical University

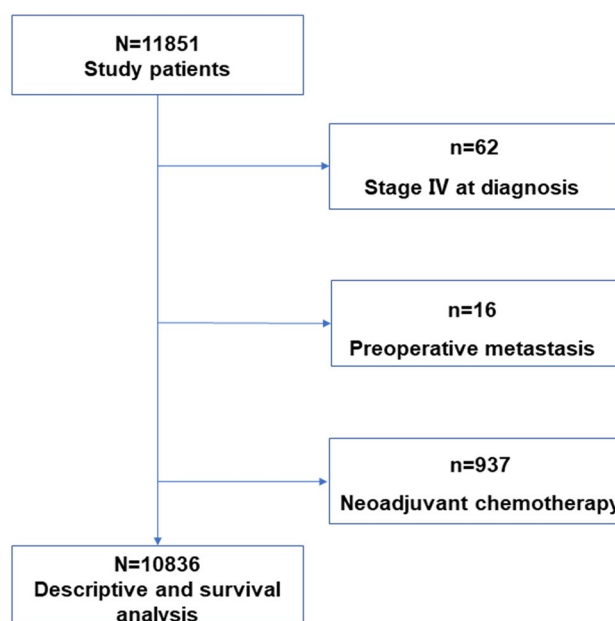


Fig. 1 Diagram of the study population

Cancer Institute and Hospital. The study included 11,851 women identified from TBCCC and diagnosed with invasive breast cancer between 2005 and 2014. We excluded patients with stage IV at diagnosis ($n = 62$), preoperative metastasis ($n = 16$), and receiving neoadjuvant chemotherapy ($n = 937$). A total of 10,836 female cases were included in descriptive and survival analysis (Fig. 1).

Tumor markers and covariates

Peripheral blood samples (5 mL) were collected from all patients before surgery. Then, serum was separated by centrifugation (3000 rpm for 5 min) and kept at $-80\text{ }^{\circ}\text{C}$ for later analysis. Serum CA15-3, CA125, and CEA levels were determined using an automatic electrochemistry luminescence immunoassay system (ROCHE E170; Roche, Germany). The cut-off points for normal and elevated tumor markers were 25 U/mL (CA15-3), 35 U/mL (CA125), and 5 $\mu\text{g/L}$ (CEA). CA15-3 with $> 25\text{ U/mL}$, CA125 with $> 35\text{ U/mL}$, and CEA with $> 5\text{ }\mu\text{g/L}$ were considered elevated levels, respectively. We included demographic, epidemiologic, and clinic-pathological variables in the analyses as adjusting variables. Demographic and epidemiologic factors included age at diagnosis, body mass index (BMI), marital status, education, average monthly income, occupation, age at menarche, parity, menopausal status, duration of breastfeeding, oral contraceptive use, smoking status, physical activity, and family history of breast cancer. Clinico-pathological factors included pT, pN, histological grade, stage, ER status, PR status, molecular subtype, chemotherapy, endocrine treatment, and radiation treatment. BMI was calculated as weight in kg that was divided by height's square in meters. Patients were defined as underweight if $\text{BMI} \leq 18.4\text{ kg/m}^2$, overweight if $\text{BMI} \geq 24\text{ kg/m}^2$ and obese if $\text{BMI} \geq 28\text{ kg/m}^2$. Monthly income of patients was categorized as low, middle, and high. Family history of breast cancer referred to at least one of the relatives (father/mother, brother/sister, son/daughter, and grandparents) with breast cancer. Patients who smoked regularly (at least one cigarette per month, more than three months) were defined as smokers.

Breast cancer molecular subtypes

In our study, part of the patients' molecular subtypes was from the pathology reports, and the rest were evaluated by imputation. The expression of ER, PR, HER2, and Ki-67 status were detected using immunohistochemistry (IHC) method. ER positivity and PR positivity were considered as the presence of $\geq 1\%$ nuclear-stained malignant cells were. The status of HER2 was based on semi-quantitative method by calculating the stained tumor cell nuclear and nuclear staining intensity, and was graded between 0 and 3+. Results of "0" or "1+" or "2+" and HER2-FISH negative

were indicated as HER2 negative and "2+" and HER2-FISH positive or "3+" as positive. Thus, we classified breast cancer intrinsic subtypes as follows: luminal A (ER+ and PR+, Her2-, Ki-67 $< 14\%$), luminal B (ER+ and PR+, Her2-, Ki-67 $\geq 14\%$ or ER+ and PR+, Her2+ [luminal Her2]), HER2-enriched (ER-, PR-, Her2+), and basal-like (ER-, PR-, Her2-, CK5/6+ and EGFR+) [14]. It is worth noting that basal-like subtypes are not completely identical with triple-negative subtypes, with an overlap of approximately 70–80%.

Previously, in the TBCCC, we reported that a subtype classifier was performed by random forest algorithm and the caret R package [15] to predict unknown molecular subtypes [16]. The data set with known molecular subtypes ($n = 1046$) was randomly divided to 80%:20%, with the former used as the primary cohort ($n = 837$) and the latter as validation cohort ($n = 209$). The model was constructed by binary ER, PR, HER2, continuous Ki67, and age at diagnosis and run with five times repeated tenfold cross validation to avoid overfitting. The performance of the imputation was assessed, with an accuracy of more than 99%. At last, we employed the established model to impute patients' subtypes among research population with unknown molecular subtypes.

Statistical analysis

Descriptive analysis was utilized for features of tumor and patients, and results were showed as proportions and frequencies. The Chi-square test was used for presenting the difference of proportions. The end points which we observed were disease progression (defined as the local recurrence, regional, or distant metastatic) and death attributed to breast cancer. Disease-free survival (DFS) was defined as the time from the date of surgery to the date of first recorded progression, the date of last follow-up, or the date of death, whichever came first. Similarly, breast cancer-specific survival (BCSS) was defined as the time from the date of surgery to the date of death attributed to breast cancer or the date of last follow-up, whichever came first. Median follow-up time for BCSS and DFS was calculated by Reverse Kaplan–Meier method. The assumption of proportional hazards was met on the basis of analysis of Schoenfeld residuals. No multicollinearity between the independent variables (tested by variance inflation factor analysis) was found. DFS and BCSS were calculated by the Kaplan–Meier plots, and the log-rank test was used for presenting the differences of survival time among groups. Analyses were performed for breast cancer overall, as well as stratified by molecular subtypes. Prognostic factors that were identified through univariate Cox analysis were subjected to backward stepwise Cox proportional hazard regression analysis to identify statistically significant variables ($P < 0.05$, Wald test) to be incorporated in the multivariable analysis. Risk estimates were shown as

hazard ratios (HRs) with 95% confidence intervals (CIs) for each factor. Interaction between tumor markers levels and molecular subtypes was calculated by including interaction variable to the Cox model. Data analysis was fulfilled by SAS (version 9.4; SAS Institute, Cary, NC, USA) and R software packages. We calculated two-sided *P* values and used a *P* value < 0.05 to express statistical significance.

Results

Patient characteristics

In total, 10,836 female breast cancer cases were included in this study, and the median age at diagnosis for participants was 51.0 (range 13.0–89.0) years. The median follow-up time was 55 (range 0–156) months. The median CA15-3, CA125, and CEA levels before surgery were 9.6 U/mL (range 1.0–395.3), 10.7 U/mL (range 0.5–2988.0), and 1.7 µg/L (range 0.2–1000.0), respectively. Elevated CA15-3, CA125, and CEA levels were identified in 367 (8.3%), 151 (7.6%), and 268 (9.5%) of the participants, respectively. By imputing for unknown molecular subtypes, there were 4513 of luminal A (61.3%), 1174 of luminal B (16.0%), 1172 of HER2-enriched (15.9%), and 498 of basal-like (6.8%) breast cancer cases in this analysis. Adjuvant chemotherapy

was the most common modality (63.5%), in comparison with endocrine treatment (21.4%), and radiation treatment (15.1%). In total, 709 patients (6.5%) displayed disease progression and 571 (5.3%) died of breast cancer (Table 1).

Association between levels of preoperative serum markers and demographic and epidemiologic factors

The distributions of demographic and epidemiologic factors among the participants are shown by tumor markers in Table 2. CA15-3 was significantly related to education ($P < 0.01$), average monthly income ($P < 0.01$), occupation ($P = 0.02$), and parity ($P = 0.02$). CA125 was associated with an average monthly income ($P < 0.01$). Moreover, statistical analysis indicated that CEA was closely correlated with patients with age at diagnosis ($P < 0.01$), education ($P = 0.01$), average monthly income ($P < 0.01$), occupation ($P = 0.03$), menopausal status ($P = 0.02$), and smoking status ($P < 0.01$).

Association between levels of preoperative serum markers and clinico-pathological factors

The distributions of clinico-pathological factors among the participants are shown by tumor markers in Table 3. The elevations of three markers were more tend to be found in larger tumor size group (pT2 or pT3-4) compared with pT0-1 ($P < 0.01$). The elevation of CA125 was more tend to be found in greater lymph-node metastases (pN1 or pN2-3, 17.6%) relative to pN0 (16.9%). Women with greater lymph-node metastases (pN1 or pN2-3, 25.0%) and more advanced stage (AJCC stage II or III, 26.8%) were more likely to be identified as elevated CA15-3 compared with pN0 (9.5%) and stage I (6.9%). The basal-like subtype (17.9%) was most likely to have elevated CA125 levels, followed by luminal B (10.1%), HER2-enriched (9.1%), and luminal A (9.1%) subtypes. The proportion of elevated CEA was similar across luminal B (8.9%), HER2-enriched (7.0%), and basal-like subtype (7.0%), but much higher among luminal A subtype (17.7%).

Association between levels of preoperative serum markers and breast cancer prognosis

For breast cancer overall, univariate analysis showed that female patients with elevated CA15-3 and CEA levels had significantly more adverse BCSS and DFS (Supplementary Tables 1, 2; Supplementary Figs. 1–5). As shown in Figs. 2 and 3, the multivariable analysis further revealed that elevated CA15-3 and CEA values conferred increased hazards of dying as a result of breast cancer (CA15-3: HR 1.54, 95%

Table 1 General characteristics of the study population

Factors	Group	<i>N</i>	%
Age of diagnosis (years)	≤ 50	5211	48.1
	> 50	5625	51.9
	Median	51.0 (13.0–89.0)	
CA15-3 (U/mL)	Normal	4047	91.7
	Elevated	367	8.3
	Median	9.6 (1.0–395.3)	
CA125 (U/mL)	Normal	1849	92.5
	Elevated	151	7.6
	Median	10.7 (0.5–2988.0)	
CEA (µg/L)	Normal	2567	90.5
	Elevated	268	9.5
	Median	1.7 (0.2–1000.0)	
Molecular subtype	Luminal A	4513	61.3
	Luminal B	1174	16.0
	HER2-enriched	1172	15.9
	Basal-like	498	6.8
Treatment	Chemotherapy	8672	63.5
	Endocrine treatment	2922	21.4
	Radiation treatment	2061	15.1
Event	Progression	709	6.5
	Dead	571	5.3

Table 2 Association between serum markers levels and epidemiological factors

Factors	CA15-3			CA125			CEA		
	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>
	<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)	
Age at diagnosis (years)									
≤ 50	1882 (91.8)	169 (8.2)	0.87	834 (91.4)	79 (8.7)	0.09	1196 (92.5)	97 (7.5)	< 0.01
> 50	2165 (91.6)	198 (8.4)		1015 (93.4)	72 (6.6)		1371 (88.9)	171 (11.1)	
Body mass index (kg/m ²)									
≤ 18.4	88 (96.7)	3 (3.3)	0.38	33 (94.3)	2 (5.7)	0.27	53 (89.8)	6 (10.2)	0.52
18.5–23.9	1711 (91.5)	159 (8.5)		782 (91.7)	71 (8.3)		1069 (89.9)	120 (10.1)	
24.0–27.9	1562 (91.7)	141 (8.3)		726 (93.9)	47 (6.1)		1007 (91.6)	92 (8.4)	
≥ 28.0	655 (91.6)	60 (8.4)		295 (91.3)	28 (8.7)		414 (90.0)	46 (10.0)	
Marital status									
Unmarried	52 (89.7)	6 (10.3)	0.82	30 (93.8)	2 (6.3)	0.87	35 (92.1)	3 (7.9)	0.16
Married	3758 (91.8)	337 (8.2)		1693 (92.4)	140 (7.6)		2372 (90.8)	241 (9.2)	
Divorced/widowed	212 (92.2)	18 (7.8)		114 (93.4)	8 (6.6)		138 (86.3)	22 (13.8)	
Education									
Without education	228 (92.3)	19 (7.7)	< 0.01	86 (89.6)	10 (10.4)	0.24	129 (84.3)	24 (15.7)	0.01
Primary school	486 (88.4)	64 (11.6)		217 (90.8)	22 (9.2)		306 (89.0)	38 (11.0)	
Junior high school	1122 (91.7)	102 (8.3)		487 (93.7)	33 (6.4)		669 (89.2)	81 (10.8)	
High school	1644 (93.4)	116 (6.6)		769 (91.7)	70 (8.3)		1053 (92.1)	90 (7.9)	
College and advanced	417 (90.1)	46 (9.9)		229 (94.6)	13 (5.4)		292 (91.5)	27 (8.5)	
Average monthly income									
Low	749 (87.5)	107 (12.5)	< 0.01	234 (83.6)	46 (16.4)	< 0.01	400 (82.0)	88 (18.1)	< 0.01
Middle	1499 (91.0)	148 (9.0)		674 (90.8)	68 (9.2)		963 (88.9)	120 (11.1)	
High	1488 (94.7)	84 (5.3)		802 (96.3)	31 (3.7)		993 (95.6)	46 (4.4)	
Occupation									
No	2064 (90.9)	207 (9.1)	0.02	931 (92.1)	80 (7.9)	0.56	1287 (89.3)	155 (10.8)	0.03
Yes	1863 (92.9)	143 (7.1)		875 (92.8)	68 (7.2)		1182 (91.7)	107 (8.3)	
Age at menarche (years)									
≤ 14	2001 (92.1)	172 (7.9)	0.30	952 (92.3)	80 (7.8)	0.74	1289 (90.9)	129 (9.1)	0.43
> 14	2016 (91.2)	194 (8.8)		881 (92.6)	70 (7.4)		1254 (90.0)	139 (10.0)	
Parity									
≤ 1	2465 (92.5)	200 (7.5)	0.02	1148 (92.2)	97 (7.8)	0.74	1582 (91.2)	152 (8.8)	0.05
> 1	1434 (90.5)	150 (9.5)		628 (92.6)	50 (7.4)		880 (89.0)	109 (11.0)	
Menopausal status									
Premenopausal	1831 (91.8)	163 (8.2)	0.76	788 (91.1)	77 (8.9)	0.05	1152 (91.9)	101 (8.1)	0.02
Postmenopausal	2216 (91.6)	204 (8.4)		1061 (93.5)	74 (6.5)		1415 (89.4)	167 (10.6)	
Breastfeeding duration (months)									
≤ 12	337 (92.6)	27 (7.4)	0.56	137 (90.7)	14 (9.3)	0.45	204 (87.9)	28 (12.1)	0.19
> 12	3552 (91.7)	321 (8.3)		1627 (92.4)	133 (7.6)		2249 (90.6)	234 (9.4)	
Oral contraceptive use									
No	3307 (92.0)	289 (8.0)	0.22	1521 (92.5)	124 (7.5)	0.52	2063 (90.4)	218 (9.6)	0.91
Yes	541 (90.5)	57 (4.5)		242 (91.3)	23 (8.7)		352 (90.3)	38 (9.7)	
Smoking status									
No	3557 (92.0)	308 (8.0)	0.16	1647 (92.3)	137 (7.7)	0.37	2263 (91.6)	209 (8.5)	< 0.01
Yes	401 (90.1)	44 (9.9)		177 (94.2)	11 (5.9)		227 (81.4)	52 (18.6)	
Vigorous physical activity									
No	2555 (91.5)	238 (8.5)	0.29	1141 (64.3)	98 (67.6)	0.43	1581 (90.3)	170 (9.7)	0.71
Yes	1316 (92.4)	108 (7.6)		633 (93.1)	47 (7.0)		852 (90.7)	87 (9.3)	

Table 2 (continued)

Factors	CA15-3			CA125			CEA		
	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>
	<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)	
Family history of breast cancer									
No	3771 (91.5)	351 (8.5)	0.07	1715 (92.8)	146 (96.7)	0.07	2404 (92.2)	250 (7.9)	0.82
Yes	276 (94.5)	16 (5.5)		134 (7.3)	5 (3.3)		163 (96.4)	18 (3.6)	

Significant associations ($P < 0.05$) are in bold

CI 1.01–2.34; CEA: HR 2.45, 95% CI 1.40–4.30). Similar associations were observed for DFS (CA15-3: HR 2.09, 95% CI 1.44–3.02; CEA: HR 2.71, 95% CI 1.71–4.27). However, no association between CA125 and BCSS or DFS of breast cancer was observed (BCSS: $P = 0.89$ and DFS: $P = 0.08$).

When stratified by molecular subtype, univariate analysis showed associations for patients with elevated CA15-3, CA125, and CEA with certain subtypes (Supplementary Tables 1, 2; Supplementary Figs. 1–5). As shown in Figs. 2 and 3, the multivariable analysis further revealed that among those with the luminal A subtype, elevated CA15-3 and CEA exhibited consistently and statistically significant reduced BCSS (CA15-3: HR 4.47, 95% CI 2.04–9.81; CEA: HR 3.79, 95% CI 1.68–8.55) and DFS (CA15-3: HR 4.06, 95% CI 2.29–7.18; CEA: HR 3.41, 95% CI 1.75–6.64). Elevated CEA conferred significant reduction for BCSS in the basal-like group (HR 5.13, 95% CI 1.65–15.9). Nevertheless, elevated CA125 was not found to be a significant indicator of either BCSS or DFS in any molecular subtypes ($P > 0.05$). The P values for the interaction of tumor markers with subtypes were not statistically significant ($P > 0.05$).

Discussion

In line with previous study findings [17–20], the association between high CA15-3 and CEA and the adverse outcome was found, because elevated markers levels are directly tied to tumor burden and the presence of antigens about tumor at the time of diagnosis of breast cancer reveals vascularization in the tumor cell, with the possibility of micrometastases [21]. However, Ebeling et al. conducted a study including 1046 breast cancer cases and did not find CA15-3 an independent prognostic marker [22] and Clinton's study supported this view [23]. A prior study reported no association between breast cancer outcome and CA15-3 among patients younger than 40 years [24], and another research by Maric et al. [13] did not offer prognostic information for tumor markers. Possible reasons for the inconsistent results may include differences in treatment modalities at the time of

measurement, the time to measure tumor markers (preoperative, intraoperative, and postoperative), the prevalence of elevated serum tumors, sample size, and follow-up time.

To our knowledge, little is known about the relationship between tumor markers and the prognosis of breast cancer among specific molecular subtypes in Chinese. The previous study just reported the association of tumor markers and molecular subtypes [17, 25–29], and studies with regard to the association between markers and breast cancer outcome in each subtype were rare. Clarke et al. [30] reviewed that MUC1 is a luminal cell-specific antigen, and can be involved in cell proliferation and differentiation. Our study further showed that patients with elevated CA15-3 and CEA had unfavorable BCSS and DFS among the luminal A subtype. This can be manifested that the elevated levels of CA15-3 and CEA are more tend to be in hormone-receptor positive breast cancer compared to negative ones. Furthermore, CEA seemed to be an independent prognostic marker for BCSS in the basal-like subtype. This may be explained partly by the distinct biological behaviors of distinct molecular subtypes [31].

The preoperative tumor markers are not only easy to obtain, but also the measurement is easily completed. Moreover, clinicians identify credible and comprehensive tumor markers is essential in an attempt to assist in decision-making for female patients, the improvement of therapeutic regimen, as well as predicting patients' outcomes by subtypes. Compared with others' studies, the role of the tumor markers was further validated with a longer follow-up of patients in our study. Although the breast cancer patients' selection criteria were rigorous, our study still guaranteed a large sample size. Furthermore, our study evaluated the association between tumor markers and the prognosis not only in breast cancer overall but also in specific molecular subtypes. The utility of CA15-3 and CEA may be served as effective prognostic indicators for Luminal A breast cancer patients. Basal-like is an important molecular subtype of breast cancer characterized by aggressive behavior and limited therapeutic response [32]. Our study further demonstrated that CEA offer added prognostic value to basal-like subtype patients, whereas it was

Table 3 Association between serum markers levels and clinico-pathological factors

Factors	CA15-3			CA125			CEA		
	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>	Normal	Elevated	<i>P</i>
	<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)		<i>n</i> (%)	<i>n</i> (%)	
pT									
pT0-1	1554 (93.1)	115 (6.9)	< 0.01	740 (94.3)	45 (5.7)	< 0.01	1072 (91.8)	96 (8.2)	< 0.01
pT2	1543 (90.2)	167 (9.8)		605 (89.6)	70 (10.4)		908 (87.8)	126 (12.2)	
pT3-4	111 (76.6)	34 (23.5)		44 (81.5)	10 (18.5)		65 (76.5)	20 (23.5)	
pN									
pN0	1512 (90.5)	158 (9.5)	< 0.01	448 (83.1)	91 (16.9)	< 0.01	937 (85.5)	159 (14.5)	0.12
pN1	773 (90.2)	84 (9.8)		345 (92.0)	30 (8.0)		470 (89.2)	57 (10.8)	
pN2-3	474 (84.8)	85 (15.2)		227 (90.4)	24 (9.6)		305 (86.9)	46 (13.1)	
Histological grade									
I	151 (90.4)	16 (9.6)	0.94	68 (89.5)	8 (10.5)	0.39	101 (89.4)	12 (10.6)	0.69
II	2210 (90.8)	223 (9.2)		1002 (92.7)	79 (7.3)		1397 (89.8)	158 (10.2)	
III	394 (90.4)	42 (9.6)		151 (90.4)	16 (9.6)		266 (91.4)	25 (8.6)	
Stage									
I	1054 (93.1)	78 (6.9)	< 0.01	482 (90.3)	52 (9.7)	0.61	686 (91.3)	65 (8.7)	0.06
II	1940 (91.2)	187 (8.8)		707 (90.1)	78 (9.9)		1119 (88.0)	152 (12.0)	
III	288 (82.1)	63 (18.0)		171 (92.4)	14 (7.6)		204 (87.9)	28 (12.1)	
ER									
Negative	1578 (91.7)	142 (8.3)	0.28	622 (89.2)	75 (10.8)	< 0.01	997 (93.6)	68 (6.4)	< 0.01
Positive	2124 (90.8)	216 (9.2)		1077 (93.7)	72 (6.3)		1331 (87.3)	193 (12.7)	
PR									
Negative	1942 (92.3)	162 (7.7)	< 0.01	844 (91.9)	74 (8.1)	0.87	1290 (92.4)	106 (7.6)	< 0.01
Positive	1761 (90.0)	196 (10.0)		856 (92.1)	73 (7.9)		1037 (86.9)	156 (13.1)	
Molecular subtype									
Luminal A	1201 (88.8)	152 (11.2)	0.12	583 (91.0)	58 (9.1)	< 0.01	749 (82.3)	161 (17.7)	< 0.01
Luminal B	487 (89.4)	58 (10.6)		248 (89.9)	28 (10.1)		286 (91.1)	28 (8.9)	
HER2-enriched	617 (91.3)	59 (8.7)		270 (90.9)	27 (9.1)		399 (93.0)	30 (7.0)	
Basal-like	391 (86.9)	59 (13.1)		138 (82.1)	30 (17.9)		254 (93.0)	19 (7.0)	
Chemotherapy									
No	828 (92.4)	68 (7.6)	0.38	435 (95.8)	19 (4.2)	< 0.01	569 (91.9)	50 (8.1)	0.19
Done	3219 (91.5)	299 (8.5)		1414 (91.5)	132 (8.5)		1998 (90.2)	218 (9.8)	
Endocrine treatment									
No	3045 (92.3)	255 (7.7)	0.02	1394 (93.9)	90 (6.1)	< 0.01	2030 (92.9)	155 (7.1)	< 0.01
Done	1002 (90.0)	112 (10.1)		455 (88.2)	61 (11.8)		537 (82.6)	113 (17.4)	
Radiation treatment									
No	3309 (92.6)	266 (7.4)	< 0.01	1478 (92.8)	114 (7.2)	0.19	2105 (90.8)	213 (9.2)	0.31
Done	738 (88.0)	101 (12.0)		371 (90.9)	37 (9.1)		462 (89.4)	55 (10.6)	
Progression									
No	3784 (93.6)	257 (6.4)	< 0.01	1739 (93.0)	131 (7.0)	< 0.01	2392 (92.3)	199 (7.7)	< 0.01
Yes	263 (70.5)	110 (29.5)		110 (84.6)	20 (15.4)		175 (71.7)	69 (28.3)	
Dead									
No	3885 (92.7)	308 (7.4)	< 0.01	1807 (93.2)	132 (6.8)	< 0.01	2462 (91.9)	217 (8.1)	< 0.01
Yes	162 (73.3)	59 (26.7)		42 (68.9)	19 (31.2)		105 (67.3)	51 (32.7)	

Significant associations ($P < 0.05$) are in bold

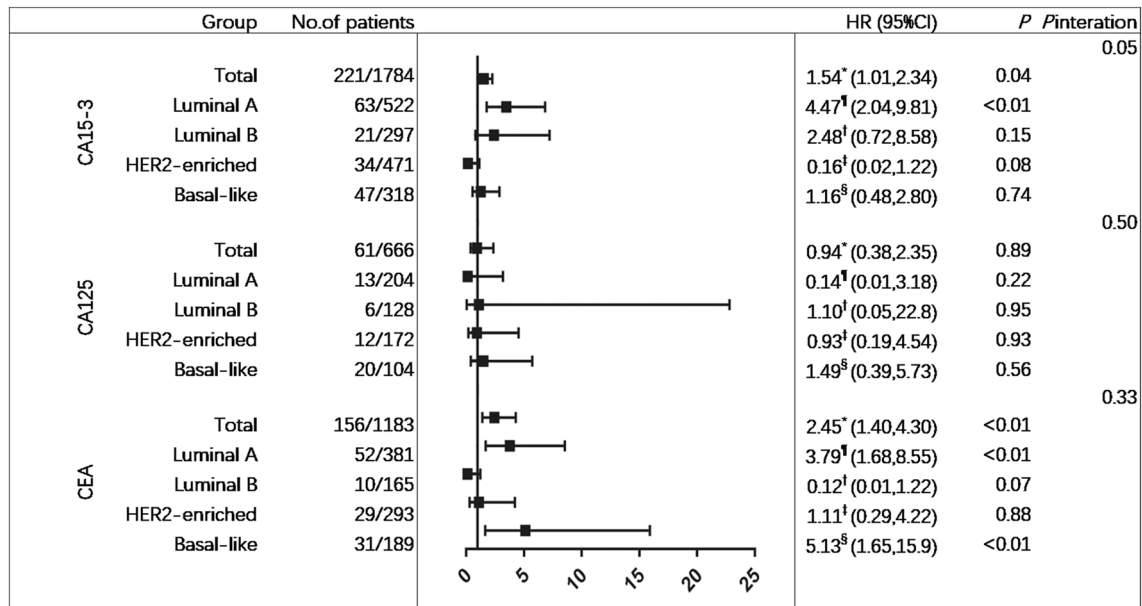


Fig. 2 Multivariate Cox regression analysis for tumor markers and breast cancer-specific survival. *Adjusted for age at diagnosis, BMI, education, average monthly income, occupation, age at menarche, parity, menopausal status, smoking status, pT, pN, histological grade, stage, ER, PR, molecular subtype, chemotherapy, endocrine, and radiation treatment. †Adjusted for age at diagnosis, education, occupation, parity, menopausal status, smoking status, pT, pN, his-

tological grade, stage, PR chemotherapy, endocrine, and radiation treatment. ‡Adjusted for age at diagnosis, marital status, education, occupation, age at menarche, parity, smoking status, pT, pN, stage, and PR. §Adjusted for BMI, Marital status, pT, pN, stage, and radiation treatment. ¶Adjusted for BMI, education, parity, pT, pN, stage, and radiation treatment

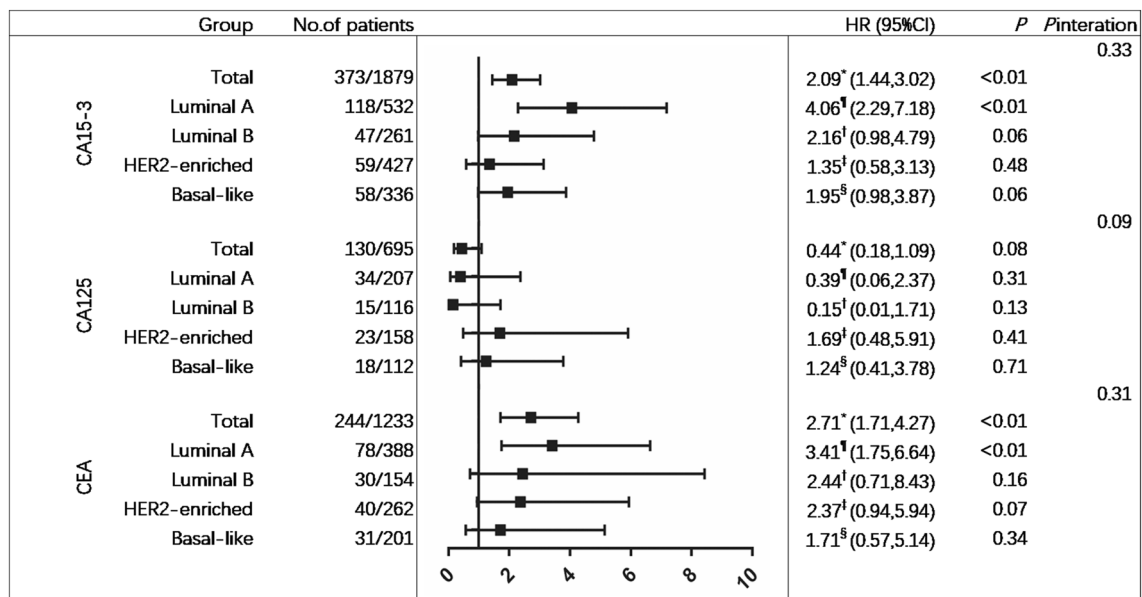


Fig. 3 Multivariate Cox regression analysis for tumor markers and disease-free survival. *Adjusted for BMI, education, parity, breast-feeding duration, menopausal status, pT, pN, histological grade, stage, ER, PR, molecular subtype, chemotherapy, endocrine, and radiation treatment. †Adjusted for education, parity, menopausal sta-

tus, pT, pN, histological grade, stage, PR, chemotherapy, endocrine, and radiation treatment. ‡Adjusted for education, parity, pT, pN, histological grade, stage, PR, and radiation treatment. §Adjusted for pT, pN, stage, and radiation treatment. ¶Adjusted for BMI, smoking status, pN, stage, and radiation treatment

not an independent prognostic factor in HER2-enriched and luminal B subtypes.

Nevertheless, due to the generally good prognosis of breast cancer, the number of people with progression or death was small even after a long follow-up. For newly diagnosed patients, shorter follow-up can lead to bias. Given this limitation, longer term follow-up will be needed in this cohort to update the results, and multi-center prospective studies should be used in an effort to confirm the validity of this study. In addition, experimental studies to investigate the mechanism of the role of tumor markers in future studies are needed.

In conclusion, CA15-3, CA125, and CEA are directly associated with aggressive clinico-pathological characteristics and elevated CA15-3 and CEA before surgery negatively affect breast cancer survival and progression among luminal A and basal-like subtypes. Therefore, CA15-3 and CEA may be combined with known prognostic variables for clinical practice in assessing patients' outcomes, and directing treatment modalities in pursuit of better prognoses, as well as determining personalized treatments for patients with different molecular subtypes.

Funding This work was funded by National Natural Science Foundation of China [Grants 81473039, 81502476, and 81974488], The National Key Research and Development program of China: The Net construction of human genetic resource Bio-bank in North China (2016YFC1201703), the National Key Research and Development Program of China (2016YFC1302703), National Human Genetic Resources Sharing Service Platform (2005DKA21300), and in part by the Program for Changjiang Scholars and Innovative Research Team in University in China (IRT_14R40) and Tianjin Municipal Key Health Research Program Grant 15KG143 and Tianjin Science and Technology Committee Foundation (17JCYBJC25300).

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. 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.
- Li H, Zheng RS, Zhang SW, Zeng HM, Sun KX, Xia CF, et al. Incidence and mortality of female breast cancer in China, 2014. *Zhonghua Zhong Liu Za Zhi*. 2018;40(3):166–71.
- Campbell HE, Gray AM, Harris AL, Briggs AH, Taylor MA. Estimation and external validation of a new prognostic model for predicting recurrence-free survival for early breast cancer patients in the UK. *Br J Cancer*. 2010;103(6):776–86.
- Soerjomataram I, Louwman MW, Ribot JG, Roukema JA, Coebergh JW. An overview of prognostic factors for long-term survivors of breast cancer. *Breast Cancer Res Treat*. 2008;107(3):309–30.
- Perey L, Hayes DF, Maimonis P, Abe M, O'Hara C, Kufe DW. Tumor selective reactivity of a monoclonal antibody prepared against a recombinant peptide derived from the DF3 human breast carcinoma-associated antigen. *Cancer Res*. 1992;52(9):2563–8.
- Falkson HC, Watt JJ, Der V, Portugal MA, Pitout MJ, Falkson G. Carcinoembryonic antigen in patients with breast cancer: an adjunctive tool to monitor response and therapy. *Cancer*. 1978;42:1308–13.
- Mughal AW, Hortobagyi GN, Fritsche HA, Buzdar AU, Yap HY, Blumenschein GR, Serial plasma carcinoembryonic antigen measurements during treatment of metastatic breast cancer. *JAMA*. 1983;249(14):1881.
- Reinartz S, Failer S, Schuell T, Wagner U. CA125 (MUC16) gene silencing suppresses growth properties of ovarian and breast cancer cells. *Eur J Cancer*. 2012;48(10):1558–699.
- Molina R, Barak V, Dalen AV, Duffy MJ, Einarsson R, Gion M, et al. Tumor markers in breast cancer—European group on tumor markers recommendations. *Tumour Biol*. 2005;26(2):281–93.
- Harris L, Fritsche H, Mennel R, Norton L, Ravdin P, Taube S, American Society of Clinical Oncology, et al. 2007 update of recommendations for the use of tumor markers in breast cancer. *J Clin Oncol*. 2007;25(33):5287–312.
- Baskiä D, Ristiä P, Matiä S, Bankoviä D, Popoviä S, Arsenijeviä N. Clinical evaluation of the simultaneous determination of CA 15–3, CA 125 and sHER2 in breast cancer. *Biomarkers*. 2007;12(6):657–67.
- Zhao S, Mei Y, Wang J, Zhang K, Ma R. Different levels of CEA, CA153 and CA125 in milk and benign and malignant nipple discharge. *PLoS ONE*. 2016;11(6):e0157639.
- Marić P, Ozretić P, Levant S, Oresković S, Antunac K, Beketić-Oresković L. Tumor markers in breast cancer—evaluation of their clinical usefulness. *Coll Antropol*. 2011;35(1):241–7.
- Goldhirsch A, Wood WC, Coates AS, Gelber RD, Thürlimann B, Senn HJ, et al. Strategies for subtypes—dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*. 2011;22(8):1736–47.
- Holm J, Eriksson L, Ploner A, Eriksson M, Rantalainen M, Li J, et al. Assessment of breast cancer risk factors reveals subtype heterogeneity. *Cancer Res*. 2017;77(13):3708–17.
- Zhang L, Huang Y, Feng Z, Wang X, Li H, Song F, et al. Comparison of breast cancer risk factors among molecular subtypes: a case-only study. *Cancer Med*. 2019;20:20.
- Lee JS, Park S, Park JM, Cho JH, Kim SI, Park BW. Elevated levels of preoperative CA 15–3 and CEA serum levels have independently poor prognostic significance in breast cancer. *Ann Oncol*. 2013;24(5):1225–311.
- Gion M, Boracchi P, Dittadi R, Biganzoli E, Peloso L, Mione R, et al. Prognostic role of serum CA15.3 in 362 node-negative breast cancers. An old player for a new game. *Eur J Cancer*. 2002;38(9):1181–8.
- Rafael M, Auge JM, Blanca F, Gabriel Z, Jaume P, Montserrat MO, et al. Prospective evaluation of carcinoembryonic antigen (CEA) and carbohydrate antigen 15.3 (CA 15.3) in patients with primary locoregional breast cancer. *Clin Chem*. 2010;56(7):1148.
- Berruti A, Tampellini M, Torta M, Buniva T, Gorzegno G, Dogliotti L. Prognostic value in predicting overall survival of two mucinous markers: CA 15–3 and CA 125 in breast cancer patients at first relapse of disease. *Eur J Cancer*. 1994;30(14):2082–4.
- Gasparini G, Toi M, Gion M, Verderio P, Dittadi R, Hanatani M, et al. Prognostic significance of vascular endothelial growth factor protein in node-negative breast carcinoma. *J Natl Cancer Inst*. 1997;89:139–47.

22. Ebeling FG, Stieber P, Untch M, Nagel D, Konecny GE, Schmitt UM, et al. Serum CEA and CA 15–3 as prognostic factors in primary breast cancer. *Br J Cancer*. 2002;86(8):1217.
23. Clinton SR, Beason KL, Bryant S, Johnson JT, Jackson M, Wilson C, et al. A comparative study of four serological tumor markers for the detection of breast cancer. *Biomed Sci Instrum*. 2003;39:408–14.
24. Li X, Dai D, Chen B, Tang H, Xie X, Wei W. Determination of the prognostic value of preoperative CA15–3 and CEA in predicting the prognosis of young patients with breast cancer. *Oncol Lett*. 2018;16(4):4679–88.
25. Fang C, Cao Y, Liu X, Zeng XT, Li Y. Serum CA125 is a predictive marker for breast cancer outcomes and correlates with molecular subtypes. *Oncotarget*. 2017;8(38):63963.
26. San-Gang W, Zhen-Yu H, Juan Z, Jia-Yuan S, Feng-Yan L, Qin L, et al. Serum levels of CEA and CA15-3 in different molecular subtypes and prognostic value in Chinese breast cancer. *Breast*. 2014;23(1):88–93.
27. Shao Y, Sun X, He Y, Liu C, Liu H. Elevated levels of serum tumor markers CEA and CA15-3 are prognostic parameters for different molecular subtypes of breast cancer. *PLoS ONE*. 2015;10(7):e0133830.
28. Riedinger JM, Goussot V, Desmoulins I, Lorgis V, Coutant C, Beltjens F, et al. CEA and early detection of relapse in breast cancer subtypes: comparison with CA 15–3. *Bull Cancer*. 2016;103(5):434–43.
29. Mudduwa LK, Wijyaratne GB, Peiris HH, Gunasekera SN, Abeywardhana D, Liyanage N. Elevated pre-surgical CA15-3: does it predict the short-term disease-free survival of breast cancer patients without distant metastasis? *Int J Womens Health*. 2018;10:329–35.
30. Clarke RB. Isolation and characterization of human mammary stem cells. *Cell Prolif*. 2005;38(6):375–86.
31. Moldovan L, Mitroi A, Petrescu CM, Aschie M. Classification of breast carcinomas according to gene expression profiles. *J Med Life*. 2013;6(1):14–7.
32. Milioli HH, Tishchenko I, Riveros C, Berretta R, Moscato P. Basal-like breast cancer: molecular profiles, clinical features and survival outcomes. *BMC Med Genom*. 2017;10(1):19.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.