Breast Cancer Research and Treatment (2019) 174:375–385
https://doi.org/10.1007/s10549-018-05069-9

PRECLINICAL STUDY

KRT19 and *CEACAM5* mRNA-marked circulated tumor cells indicate unfavorable prognosis of breast cancer patients

Xi-Mei Wang[1,2,3] · Zhen Zhang[1,2,3] · Li-Hui Pan[1,2,3] · Xu-Chen Cao[1,2,3] · Chunhua Xiao[1,2,3]

Received: 14 September 2018 / Accepted: 23 November 2018 / Published online: 8 December 2018

Abstract

Aim To investigate the clinical and prognostic significance of circulated tumor cells (CTC) marked by cytokeratin 19 coding gene *KRT19* mRNA and carcinoembryonic antigen coding gene *CEACAM5* mRNA in preoperative peripheral blood of breast cancer patients and provide molecular markers for breast cancer metastasis risk.

Methods The mRNA levels of *KRT19* and *CEACAM5* in preoperative peripheral blood of breast cancer patients without ($n=603$) and with ($n=76$) distant metastases at the time of initial diagnosis were detected by reverse transcription-quantitative polymerase chain reaction (RT-qPCR). The relationship between CTC_{KRT19}, $CTC_{CEACAM5}$ and clinicopathological features, local recurrence-free survival (LRFS), distant metastasis-free survival (DMFS), or overall survival (OS) was statistically analyzed.

Results In different pathological stages of breast cancer, the rates of $CTC_{KRT19-pos}$ and $CTC_{CEACAM5-pos}$ increased with the increase of the stages ($P=0.077$ and $P=0.004$). Preoperative $CTC_{KRT19-pos}$ in breast cancer patients was closely related to the lymph node metastasis statues ($P<0.0001$), and had no significant correlation with other clinicopathological features. There was no significant correlation between $CTC_{CEACAM5}$ and the clinicopathological features. Patients with high levels of CTC double-marked by *KRT19* and *CEACAM5* mRNA had shorter DMFS ($P<0.0001$) and OS ($P=0.016$) for patients with breast cancer. The 7-year DMFS rates for the low-, intermediate-, and high-risk groups were 90.7%, 67.5%, and 59.1%, respectively ($P<0.0001$). The prognosis of patients with decreased *KRT19* and *CEACAM5* mRNA after treatment is better than that of patients who have not decreased, and the combination of the two indicators is better than the single one for predicting PFS ($P=0.002$ compare with $P=0.036$ or $P=0.047$).

Conclusion Double-marked CTC by *KRT19* and *CEACAM5* mRNA is a prognostic index of breast cancer patients before surgery and after chemotherapy. Single-marked CTC by *KRT19* mRNA indicates lymph node statues of preoperative patients. Therefore, the RT-qPCR-based molecular diagnosis of CTC could be used for prognostic prediction of breast cancer patients and guiding clinical treatment.

Keywords *KRT19* · *CEACAM5* · RT-qPCR · Circulated tumor cells · Lymph node metastasis · Prognosis · Breast cancer

Introduction

As one of the most common malignant tumors, breast cancer has the highest incidence and is the second leading cause of mortality among females [1]. Although important advances in surgical techniques and adjuvant chemotherapy have improved patient outcomes, recurrence and metastasis after surgery always remain the main causes of treatment failure and poor prognosis [2]. To establish molecular biomarkers that can detect the early metastasis and improve survival is necessary for breast cancer patients.

Circulating tumor cells (CTCs) are derived from tumors as an early step in blood-borne metastasis, which is transient

✉ Chunhua Xiao
XCHwxm2019@163.com

1 First Department of Breast Cancer, National Clinical Research Center of Cancer, Tianjin Medical University Cancer Institute and Hospital, 1 Huan-Hu Xi Road, Ti-Yuan Bei, He Xi, Tianjin 300060, People's Republic of China

2 Key Laboratory of Cancer Prevention and Therapy, Tianjin, People's Republic of China

3 Key Laboratory of Breast Cancer Prevention and Therapy, Ministry of Education, Tianjin Medical University, Tianjin 300060, People's Republic of China

in blood with a half-life of 1–2.4 h [3–6]. The presence of CTCs at levels ≥ 5/7.5 ml peripheral blood is associated with poor prognosis of breast cancer patients, and it might be another breast cancer prognosis assessment tool following the four biological indicators of ER/PR, HER2, Ki67, and tumor histological grade [7, 8]. Cytokeratin 19 coding gene *KRT19* is derived from epithelial cells and is expressed in normal epithelial and epithelial primary tumors and metastatic tumor cells. *KRT19* is absent in normal peripheral blood and lymphoid tissues [9]. Therefore, the detection of *KRT19* transcripts in the peripheral blood of breast cancer patients indicates the presence of breast cancer cells [10]. Carcinoembryonic antigen coding gene *CEACAM5* is overexpressed in gastrointestinal cancer, breast cancer, etc. It plays an important role in tumor metastasis and may be related to the prognosis of breast cancer [11].

Recently, reverse transcription-quantitative polymerase chain reaction (RT-qPCR) technology has been shown to have a sensitivity 10- to 100-fold higher than the routine immunological methods and to be a reliable method for detecting circulating tumor cells and measuring the mRNA expression of tumor markers [12, 13]. However, there have been few reports on the combined detection and clinical significance of *KRT19* and *CEACAM5* mRNA expression in peripheral blood of patients with breast cancer.

In our study, we detected the *KRT19* and *CEACAM5* mRNA levels before surgery and after chemotherapy in peripheral blood of patients with breast cancer using RT-qPCR and estimated the clinical and prognostic value of these biomarkers in breast cancer patients.

Table 1 Clinicopathological characteristics of breast cancer patients

Characteristics	Cases (%)
Age (years)	
≤ 45	179 (29.8)
45–55	228 (37.9)
> 55	196 (32.3)
Tumor size (cm)	
≦ 2	274 (45.4)
2–5	306 (50.8)
> 5	23 (3.8)
Lymph node status	
0	352 (58.4)
1–3	131 (21.7)
4–9	68 (11.3)
≥ 10	52 (8.6)
Histological grade	
I	76 (12.6)
II	420 (69.7)
III	107 (17.7)
pTNM	
0	55 (9.1)
I	148 (24.5)
II	297 (49.3)
III	103 (17.1)
ER status	
Negative	160 (26.5)
Positive	443 (73.5)
PR status	
Negative	211 (34.9)
Positive	392 (65.1)
HER2 status	
Negative	476 (78.9)
Positive	127 (21.1)

Materials and methods

Patients and follow-up

A total of 603 breast cancer patients who underwent radical mastectomy between May 2010 and April 2011 at Tianjin Medical University Cancer Institute and Hospital (TMUCIH) were recruited for the study. We also recruited 76 patients with advanced breast cancer who had distant metastases at the same time as control. The inclusion criteria for the study were as follows: (1) complete clinical pathology data and follow-up data; (2) no neoadjuvant chemotherapy before surgery; (3) no distant metastasis before surgery; (4) all patients underwent radical mastectomy; (5) standardized treatment after surgery; (6) no other synchronous malignancies. The histological diagnosis and tumor-node-metastasis (TNM) staging were based on 7th edition of the American Joint Committee on Cancer (AJCC)/Union for International Cancer Control (UICC) TNM staging system [14]. The clinicopathological characteristics of breast cancer patients are shown in Table 1. The study and acquisition of blood specimens were approved by the Research Ethics Committee of Tianjin Medical University Cancer Institute and Hospital. In addition, all of the patients provided written informed consent. After radical mastectomy, all patients were followed regularly every 3 month during the first year, every 6 month during the second to third years, and every 1 year thereafter until death or the last follow-up. The final follow-up date was December 2017. Local recurrence-free survival (LRFS) was calculated from the date of surgery to the date of local chest wall recurrence or final follow-up date; distant metastatic free survival (DMFS) was calculated from the date of surgery to the date of distant metastasis or the date of final follow-up; overall survival (OS) is calculated from the date of surgery to the date of death or the date of final follow-up; progression-free survival (PFS) is calculated from the date of surgery to the disease progression or death or the date of final follow-up.

Samples collection

Five millilambert peripheral blood samples were obtained through a catheter inserted into a peripheral vessel and collected into EDTA tubes from each breast cancer patient. Nucleated cells were then separated by hypotonic lysis of red blood cells. Briefly, three volumes of hypotonic lysis buffer (10 mM Tris-HCl, 10 mM NaCl, 5 mM MgCl2, pH 7.6) was mixed with one volume of blood sample and centrifuged at 2000 rpm for 15 min a 4 °C with the supernatant discarded. The hypotonic lysis was repeated one more time and the resulting cell pellet was stored at − 80 °C.

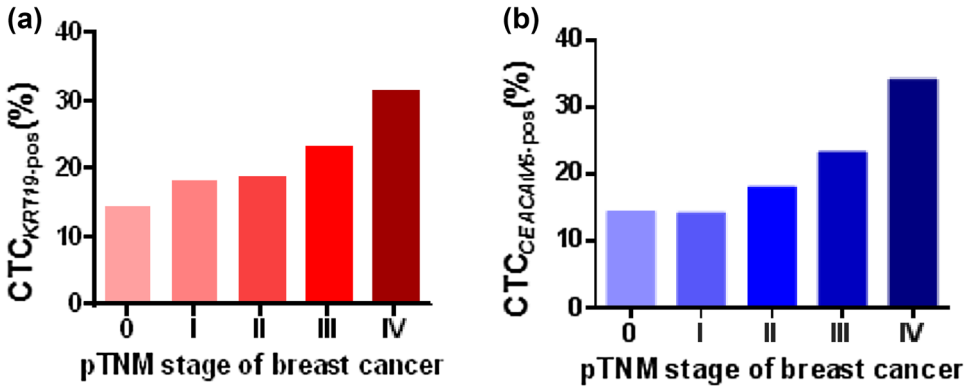

Fig. 1 The relationship between the CTC_{KRT19}, $CTC_{CEACAM5}$ and the pathological stages was calculated using the χ^2 test. **a** The rates of $CTC_{KRT19-pos}$ in different pathological stages of breast cancer, $P = 0.077$. **b** The rates of $CTC_{CEACAM5-pos}$ in different pathological stages of breast cancer, $P = 0.004$

Table 2 Correlation between the preoperative *KRT19* and *CEACAM5* mRNA-marked CTC levels and the clinicopathological features of 603 patients with breast cancer

Variables	Cases	$CTC_{KRT19-pos}$ (%)	$CTC_{KRT19-neg}$ (%)	χ^2	*P*	$CTC_{CEACAM5-pos}$ (%)	$CTC_{CEACAM5-neg}$ (%)	χ^2	*P*
Age (years)				5.342	0.069			0.799	0.671
≤ 45	179	40 (22.3)	139 (77.7)			35 (19.6)	144 (80.4)		
45–55	228	47 (20.6)	181 (79.4)			37 (16.2)	191 (83.8)		
> 55	196	27 (13.8)	169 (86.2)			35 (17.9)	161 (82.1)		
Tumor size (cm)				1.793	0.408			1.775	0.412
≦ 2	274	58 (21.2)	216 (78.8)			44 (16.1)	230 (83.9)		
2–5	306	51 (16.7)	255 (83.3)			57 (18.6)	249 (81.4)		
>5	23	5 (21.7)	18 (78.3)			6 (26.1)	17 (73.9)		
Lymph node status				85.749	< 0.0001			1.888	0.596
0	352	60 (17.0)	292 (83.0)			58 (16.5)	294 (83.5)		
1–3	131	26 (19.8)	105 (80.2)			28 (21.4)	103 (78.6)		
4–9	68	15 (22.0)	53 (78.0)			13 (19.1)	55 (80.9)		
≥ 10	52	13 (25.0)	39 (75.0)			8 (15.4)	44 (84.6)		
Histological grade				0.942	0.624			1.248	0.536
I	76	16 (21.1)	60 (78.9)			10 (13.2)	66 (86.8)		
II	420	81 (19.3)	339 (80.7)			77 (18.3)	343 (81.7)		
III	107	17 (15.9)	90 (84.1)			20 (18.7)	87 (81.3)		
ER status				0.355	0.551			0.411	0.521
Negative	160	33 (20.6)	127 (79.4)			31 (19.4)	129 (80.6)		
Positive	443	81 (18.3)	362 (81.7)			76 (17.2)	367 (82.8)		
PR status				1.605	0.205			0.655	0.418
Negative	211	45 (21.3)	166 (78.7)			41 (19.4)	170 (80.6)		
Positive	392	69 (17.6)	323 (82.4)			66 (16.8)	326 (83.2)		
HER2 status				2.191	0.139			0.190	0.663
Negative	476	84 (17.6)	392 (82.4)			75 (15.8)	401 (84.2)		
Positive	127	30 (23.6)	97 (76.4)			32 (25.2)	95 (74.8)		

Total RNA isolation and cDNA synthesis

Total RNA was extracted from the cell lysate using TRIzol reagent (Invitrogen, Gaithersburg, MD, United States) according to the manufacturer's instructions. The concentration and purity of RNA were determined using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE, United States). A260/A280 ratios in the range of 1.8–2.0 were considered satisfactory for purity standards in this study. First-strand cDNA was synthesized using the SuperScript First-Strand cDNA Synthesis kit (Invitrogen) according to the manufacturer's instructions and then stored at − 20 °C for subsequent quantitative polymerase chain reaction experiments.

Real-time qPCR

The mRNA expression levels of *KRT19* and *CEACAM5* were detected by real-time PCR using the ABI 7500 Real-Time PCR system (Applied Biosystems, Foster City, CA, United States). Quantification of mRNA levels of target genes was accomplished by measuring the fractional cycle number at which the amount of expression reached a fixed threshold (CT). All of the reactions were run in triplicate. The primers targeting *KRT19* (NM_002276.4) and *CEACAM5* (NM_001291484.2) were optimized using Oligo 6.0 software and synthesized by Shanghai Shenggong Bioengineering Co., Ltd. The following thermocycling conditions were used under the standard mode according to the manufacturer's recommendations: 50 °C 2 min, 95 °C pre-denaturation 3 min; 95 °C 30 s, 62 °C 1 min, 40 cycles.

CTC-positive criteria

The receiver operating characteristic (ROC) curves were made based on *KRT19* mRNA levels of samples and the corresponding overall survival (OS) status of patients. The optimal cut-off value of 32.5 of *KRT19* mRNA CT level was selected according to ROC curve analyses. It is with higher sensitivity and specificity to separate all patients into high *KRT19* mRNA level group and low *KRT19* mRNA level group with distinguished OS status. Patients with a CT value of *KRT19* mRNA less than or equal to 32.5 were defined as CTC-positive patients, and with a CT value higher than 32.5 were defined as CTC-negative patients ($CTC_{KRT19-pos}$, *KRT19* mRNA CT ≤ 32.5, $n = 114$; $CTC_{KRT19-neg}$, *KRT19* mRNA CT > 32.5, $n = 489$).

According to the ROC curve, the optimal cut-off value of the *CEACAM5* mRNA CT level was set at 35.0. Based on this cut-off value in predicting 5-year survival, it is with higher

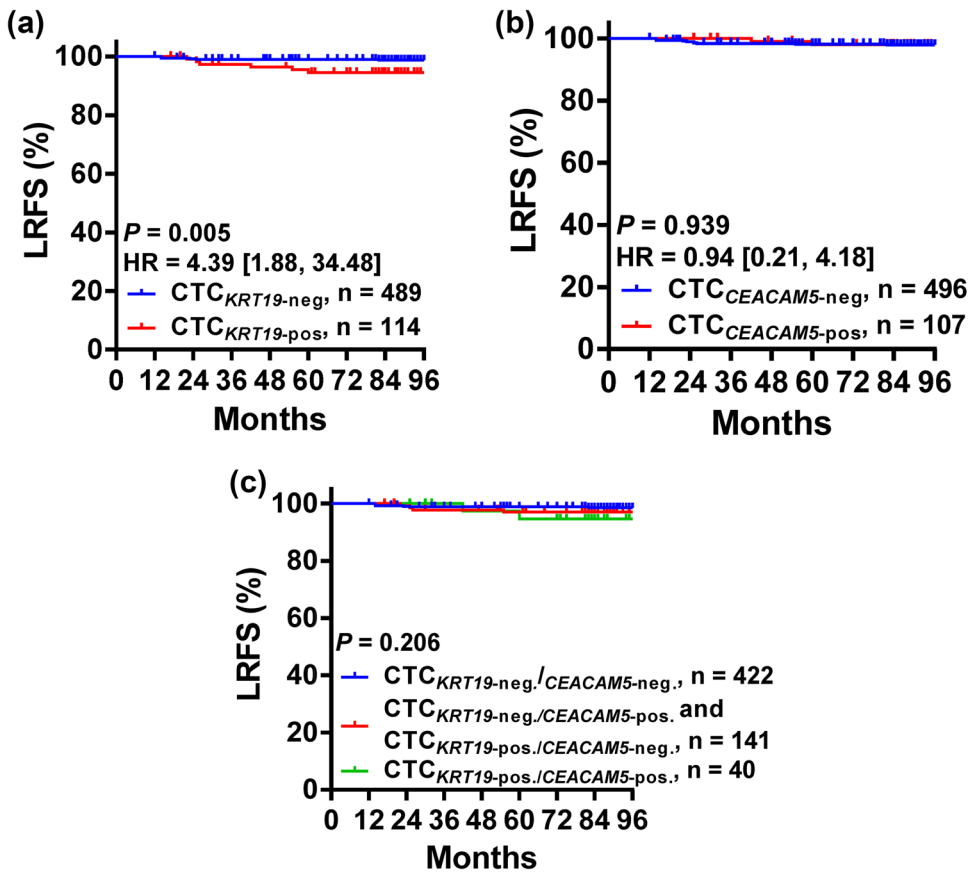

Fig. 2 Kaplan–Meier survival curves according to preoperative CTC_{KRT19} and $CTC_{CEACAM5}$ levels in patients with breast cancer. **a** CTC_{KRT19} level for LRFS; **b** $CTC_{CEACAM5}$ level for LRFS; **c** combined CTC_{KRT19} and $CTC_{CEACAM5}$ levels for LRFS; *P* values were calculated by the log-rank test and $P < 0.05$ denoted significance

sensitivity and specificity to separate all patients into high *CEACAM5* mRNA level group and low *CEACAM5* mRNA level group with distinguished OS status. CTC-positive patients were defined as those whose CT value of *CEACAM5* mRNA was less than or equal to 35.0, while CTC-negative patients were defined as those whose CT value was higher than 35.0 ($CTC_{CEACAM5-pos}$, *CEACAM5* mRNA CT ≤ 35.0, $n = 107$; $CTC_{CEACAM5-neg}$, *CEACAM5* mRNA CT > 35.0, $n = 496$).

We also defined CTC status by combining *KRT19* mRNA and *CEACM5* mRNA levels. We defined *KRT19* mRNA CT value higher than 32.5 and *CEACAM5* mRNA CT value higher than 35.0 as $CTC_{negtive}$; *KRT19* mRNA CT value less than or equal to 32.5 or *CEACAM5* mRNA CT value less than or equal to 35.0 was defined as $CTC_{weak-positive}$; *KRT19* mRNA CT value was less than or equal to 32.5 and *CEACAM5* mRNA CT value was less than or equal to 35.0 defined as $CTC_{positive}$.

Statistical analysis

All statistical analyses were performed using SPSS 24.0 statistical software (SPSS Inc, Chicago, IL, United States). Based on *KRT19, CEACAM5* mRNA CT levels in breast cancer tissues, and corresponding OS status of patients, the receiver operating characteristic (ROC) curve was made to identify the optimized cut-off value of the two marker's mRNA levels. The relationship between the CTC_{KRT19}, $CTC_{CEACAM5}$ and the clinical pathological features was calculated using the χ^2 test and are displayed in cross-tables. Survival analysis was performed using the Kaplan–Meier method and compared by the log-rank test. The variables significantly affecting LRFS, DMFS, OS, and PFS were investigated by multivariate analysis according to the Cox proportional hazard model. *P* values < 0.05 were considered statistically significant.

Results

Difference of CTC-positive rate in preoperative peripheral blood of breast cancer patients with different pTNM

The rates of $CTC_{KRT19-pos}$ in preoperative peripheral blood of stage 0–IV breast cancer were 14.5% (8/55), 18.2% (27/148), 18.8% (56/298), 23.3% (24/103), and 31.6% (24/76), respectively, and the difference was statistically significant ($\chi^2 = 8.429$, $P = 0.077$, Fig. 1a). The rates of $CTC_{CEACAM5-pos}$ in preoperative peripheral blood of stage 0–IV breast cancer were 14.5% (8/55), 14.2% (21/148), 18.1% (54/298), 23.3% (24/103), and 34.2% (26/76), respectively, and the difference was statistically significant ($\chi^2 = 15.269$, $P = 0.004$, Fig. 1b).

Correlation between the preoperative CTC_{KRT19}, $CTC_{CEACAM5}$ levels and clinicopathological factors

The correlation between the preoperative CTC_{KRT19} levels and the clinicopathological characteristics is shown in Table 2. Preoperative $CTC_{KRT19-pos}$ in peripheral blood of patients with breast cancer was closely related to the metastasis of lymph nodes ($P < 0.0001$), and had no significant correlation with other clinicopathological features ($P > 0.05$). The correlation between the preoperative $CTC_{CEACAM5}$ levels and the clinicopathological characteristics is shown in Table 2. There was no significant correlation between $CTC_{CEACAM5}$ and the clinicopathological features ($P > 0.05$).

Univariate and multivariate survival analyses of clinicopathological characteristics for LRFS, DMFS, and OS of patients with Breast Cancer

The median follow-up period for the entire cohort was 84 months (range 12–107 months).

For local recurrence-free survival (LRFS), the 7-year LRFS rates in the preoperative $CTC_{KRT19-neg}$ and $CTC_{KRT19-pos}$ groups were 98.8% and 94.8%, respectively

Table 3 Univariate and multivariate analyses of the clinicopathological variables for LRFS in breast cancer patients

Variables	HR (95% CI)	*P*
Univariate		
CTC_{KRT19}		
Positive versus negative	4.39 (1.88–34.48)	0.005
$CTC_{CEACAM5}$		
Positive versus negative	0.94 (0.21–4.30)	0.939
Tumor size (cm)		
> 2 versus ≤ 2	1.01 (0.37–2.76)	0.984
Pathological type		
DCIS versus invasive carcinoma	1.13 (0.15–8.73)	0.909
Histological grade		
III versus I–II	1.95 (0.70–5.44)	0.200
Lymph node status		
Positive versus negative	1.72 (1.07–2.76)	0.026
ER status		
Positive versus negative	0.71 (0.21–2.35)	0.574
PR status		
Positive versus negative	1.07 (0.32–3.54)	0.918
HER2 status		
Positive versus negative	0.76 (0.17–3.46)	0.720
Multivariate		
CTC_{KRT19}		
Positive versus negative	4.05 (1.30–12.59)	0.016
Lymph node status		
Positive versus negative	1.65 (1.02–2.67)	0.041

($P = 0.005$, Fig. 2a). The 7-year LRFS rates in the CTC $_{CEACAM5-neg}$ and CTC$_{CEACAM5-pos}$ groups were 98.0% and 98.1%, respectively ($P = 0.939$, Fig. 2b). By combining *KRT19* with *CEACAM5* mRNA, we divided patients into CTC$_{positive}$, CTC$_{weak-positive}$, and CTC$_{negtive}$ groups. The 7-year LRFS rates of the three groups are 95%, 97.2%, and 98.6%, respectively ($P = 0.206$, Fig. 2c).

Univariate analysis showed that preoperative CTC $_{KRT19-pos}$ ($P = 0.005$) and lymph node metastasis ($P = 0.026$) were associated with local recurrence-free survival (LRFS) (Table 3).

The multivariate survival analysis indicated that CTC $_{KRT19-pos}$ in peripheral blood before surgery ($P = 0.016$) and lymph node metastasis ($P = 0.041$) were independent risk factors for local recurrence-free survival (LRFS) in breast cancer patients (Table 3).

For distant metastatic free survival (DMFS), the 7-year DMFS rates in the preoperative CTC$_{KRT19-neg}$ and CTC $_{KRT19-pos}$ groups were 88.5% and 78.3%, respectively ($P = 0.002$, Fig. 3a). The 7-year DMFS rates in the preoperative CTC$_{CEACAM5-neg}$ and CTC$_{CEACAM5-pos}$ groups were 89.5% and 72.9%, respectively ($P < 0.0001$, Fig. 3b). The 7-year DMFS rates of the CTC$_{positive}$, CTC$_{weak-positive}$, and CTC$_{negtive}$ groups defined by combining *KRT19* with *CEACAM5* mRNA levels are 67.5%, 80.1%, and 90.5%, respectively ($P < 0.0001$, Fig. 3c).

Univariate analysis showed that preoperative CTC $_{KRT19-pos}$ ($P = 0.002$), CTC$_{CEACAM5-pos}$ ($P < 0.0001$), tumor size ($P < 0.0001$), and lymph node metastasis ($P < 0.0001$) were related to distant metastatic free survival (DMFS) (Table 4).

The four factors with prognostic potential for DMFS were subsequently subjected to multivariate analysis using the Cox proportional hazards model. As shown in Table 5, preoperative CTC$_{KRT19-pos}$ ($P = 0.010$), CTC$_{CEACAM5-pos}$ ($P = 0.000$), and lymph node metastasis ($P = 0.040$) were the independent prognostic factors for distant metastasis-free survival (DMFS) in breast cancer patients (Table 4).

For overall survival (OS), the 7-year OS rate of preoperative CTC$_{KRT19-neg}$ and CTC$_{KRT19-pos}$ groups were 94.1% and 87.8%, respectively ($P = 0.019$, Fig. 4a). The 7-year OS rate of the preoperative CTC$_{CEACAM5-neg}$ and CTC $_{CEACAM5-pos}$ groups were 93.8% and 88.8%, respectively ($P = 0.066$, Fig. 4b). The 7-year OS rates of the CTC$_{positive}$, CTC$_{weak-positive}$, and CTC$_{negtive}$ groups defined by combining *KRT19* with *CEACAM5* mRNA levels are 82.5%, 91.5%, and 94.3%, respectively ($P = 0.016$, Fig. 4c).

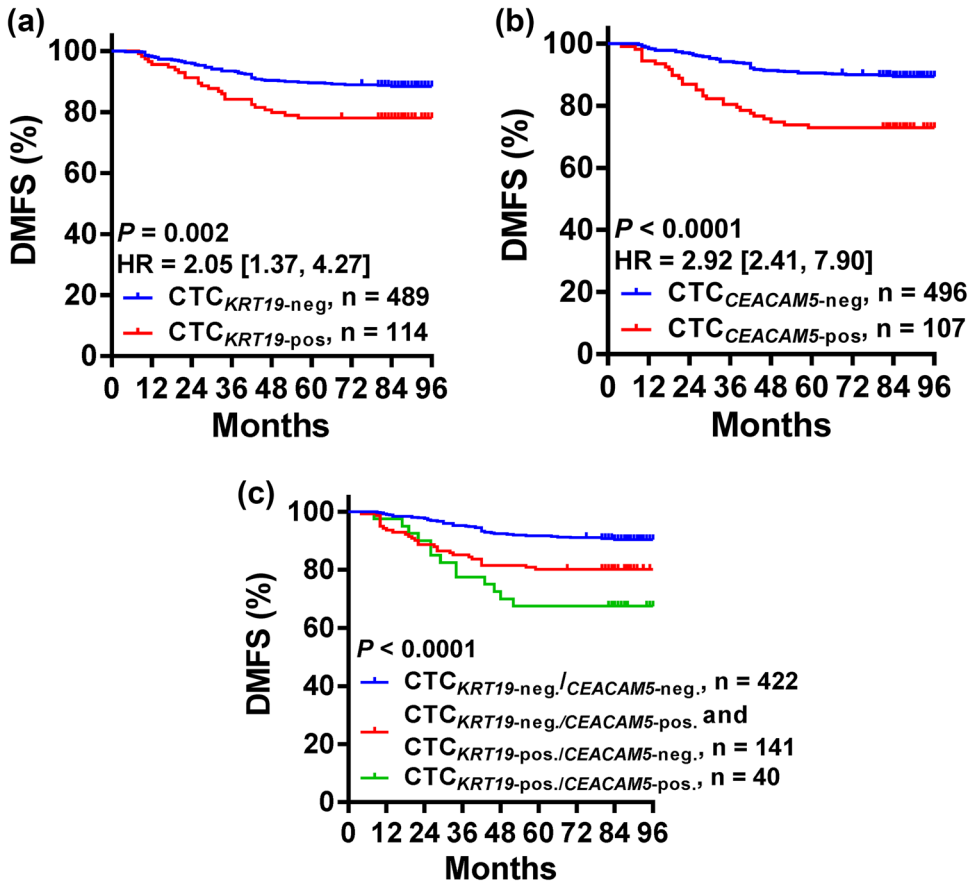

Fig. 3 Kaplan–Meier survival curves according to preoperative CTC$_{KRT19}$ and CTC$_{CEACAM5}$ levels in patients with breast cancer. **a** CTC$_{KRT19}$ level for DMFS; **b** CTC$_{CEACAM5}$ level for DMFS; **c** combined CTC $_{KRT19}$ and CTC$_{CEACAM5}$ levels for DMFS; *P* values were calculated by the log-rank test and $P < 0.05$ denoted significance

Table 4 Univariate and multivariate analyses of the clinicopathological variables for DMFS in breast cancer patients

Variables	HR (95% CI)	*P*
Univariate		
CTC_{KRT19}		
Positive versus negative	2.05 (1.38–4.28)	0.002
$CTC_{CEACAM5}$		
Positive versus negative	2.92 (2.41–7.90)	< 0.0001
Tumor size (cm)		
> 2 versus ≤ 2	2.03 (1.39–2.95)	< 0.0001
Pathological type		
DCIS versus Invasive carcinoma	23.51 (0.94–588.87)	0.055
Histological grade		
III versus I–II	1.29 (0.87–1.92)	0.212
Lymph node status		
Positive versus negative	1.96 (1.63–2.36)	< 0.0001
ER status		
Positive versus negative	0.70 (0.44–1.12)	0.134
PR status		
Positive versus negative	0.73 (0.47–1.13)	0.161
HER2 status		
Positive versus negative	1.55 (0.96–2.51)	0.075
Multivariate		
CTC_{KRT19}		
Positive versus negative	4.24 (1.50–6.45)	0.01
$CTC_{CEACAM5}$		
Positive versus negative	3.03 (1.92–4.80)	< 0.0001
Lymph node status		
Positive versus negative	1.44 (1.02–2.04)	0.040

Univariate analysis showed that preoperative $CTC_{KRT19-pos}$ ($P = 0.019$), tumor size ($P = 0.003$), and lymph node metastasis ($P < 0.0001$) were related to overall survival (OS) (Table 5).

The three factors with prognostic potential for OS were subsequently subjected to multivariate analysis using the Cox proportional hazards model. As shown in Table 3, preoperative $CTC_{KRT19-pos}$ ($P = 0.046$) was the independent prognostic factor for overall survival in breast cancer patients (Table 5).

Prognostic model and risk groups

We proposed a new prognostic model to better identify breast cancer patients at high risk for metastasis and poor prognosis. This new prognostic model was proposed by risk factors for distant metastasis-free survival (DMFS) ($CTC_{KRT19-pos}$, $CTC_{CEACAM5-pos}$, and lymph node metastasis) and stratified patients into three groups as follows: the low-risk group, comprising patients with 0 or 1 risk factor; the intermediate-risk group, comprising patients with 2 factors; and the high-risk group, comprising patients with all 3 factors. Finally, there were 504, 77, and 22 patients in the low-, intermediate-, and high-risk groups, respectively. The 7-year DMFS rates for the low-, intermediate-, and high-risk groups were 90.7%, 67.5%, and 59.1%, respectively, and a statistically significant difference was observed ($P < 0.0001$; Fig. 5). This new prognostic model may potentially help clinicians make a more accurate judgment for prognosis and individual therapeutic treatment to improve survival and quality of life.

Relationship between changes of *KRT19*, *CEACAM5* mRNA levels and prognosis before and after treatment

We randomly selected 100 patients from 604 patients to receive *KRT19* and *CEACAM5* mRNA expression levels after treatment, and explored the relationship between *KRT19*、*CEACAM5* mRNA levels and prognosis before and after treatment. All patients underwent a chemotherapy regimen containing anthracyclines and/or paclitaxel. HER2-positive patients were treated with trastuzumab and the number of median chemotherapy was 5 times. The expression levels of *KRT19* and *CEACAM5* mRNA after treatment were obtained by detecting peripheral blood on the second day after the end of the last chemotherapy. We defined post-treatment reduction in *KRT19* and/or *CEACAM5* mRNA as therapeutically effective. Among them, 65 patients had a decrease in *KRT19* mRNA after treatment, and 2 patients developed disease progression (local recurrence, distant metastasis, or death); the 7-year PFS rate was 96.9%; There were 35 patients with unchanged or elevated *KRT19* mRNA after treatment, and 5 patients developed disease progression; the 7-year PFS rate was 85.7% ($P = 0.036$; Fig. 6a).

There were 63 patients with decreased *CEACAM5* mRNA after treatment, and 2 patients with disease progression (local recurrence, distant metastasis, or death); the 7-year PFS rate was 96.8%. There were 37 patients with unchanged or elevated *CEACAM5* mRNA after treatment and 5 cases of disease progression; the 7-year PFS rate of 86.5% ($P = 0.047$; Fig. 6b).

There were 83 patients with decreased *KRT19* and/or *CEACAM5* mRNA after treatment, and 3 patients developed disease progression (local recurrence, distant metastasis, or death), and the 7-year PFS rate was 96.4%. There were 17 patients with no reduction in *KRT19* and *CEACAM5* mRNA after treatment, and 4 patients developed disease progression. The 7-year PFS rate was 76.5% ($P = 0.002$; Fig. 6c).

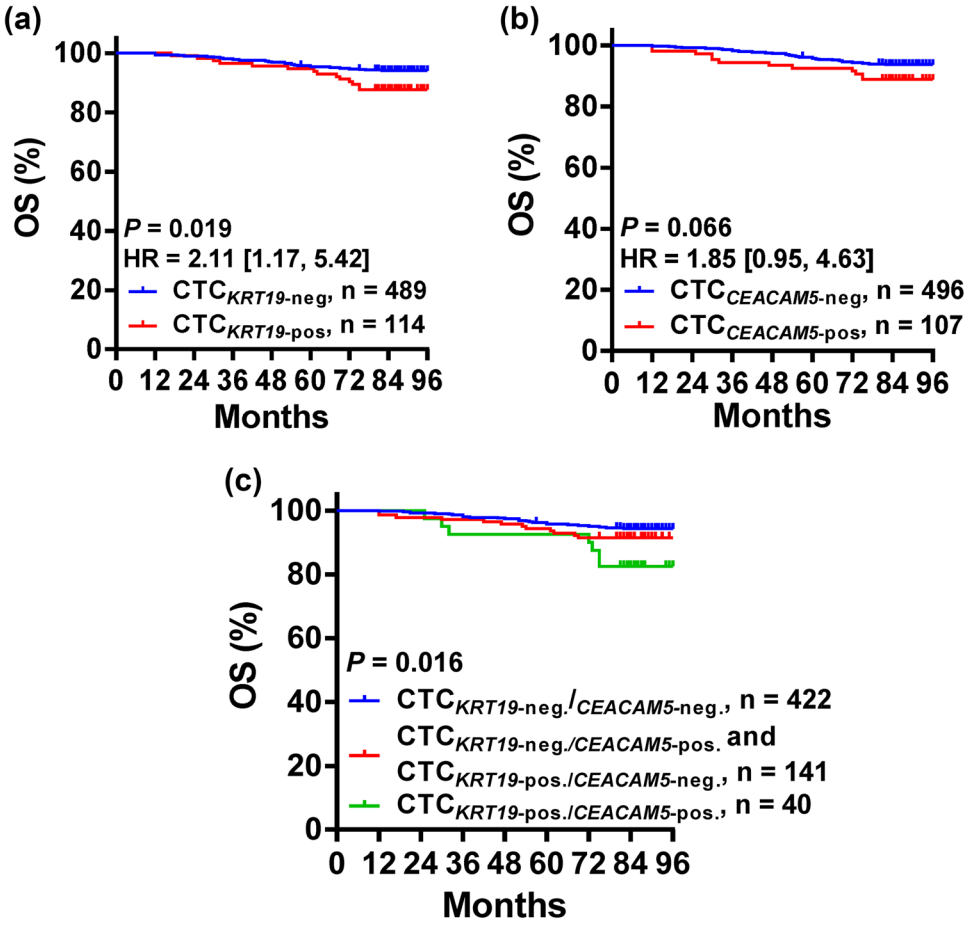

Fig. 4 Kaplan–Meier survival curves according to preoperative CTC_{KRT19} and $CTC_{CEACAM5}$ levels in patients with breast cancer. **a** CTC_{KRT19} level for OS; **b** $CTC_{CEACAM5}$ level for OS; **c** combined CTC_{KRT19} and $CTC_{CEACAM5}$ levels for OS; *P* values were calculated by the log-rank test and $P < 0.05$ denoted significance

Discussion

Tumor recurrence and distant metastasis are the leading factors influencing the clinical outcome of patients with breast cancer. Many breast cancer patients with resectable tumors died of postoperative distant metastasis and had a poor prognosis [2]. Therefore, identifying one method or marker to determine the potential of cancer cell spreading and disease progression in patients with breast cancer patients will certainly help to tailor postoperative adjuvant therapies and improve clinical outcomes.

Recently, gene-based biomarkers have demonstrated clinical value for the diagnosis, prognosis, and prediction of drug responses in a variety of cancers [15]. For breast cancer, several tumor-specific markers, including *KRT19* and *CEACAM5* in either serum or tumor tissue, have been used to detect metastasis and predict prognosis [9, 11, 16, 17]. RT-qPCR, a sensitive, specific, and rapid method, has been widely used to detect the presence of circulating cancer cells in peripheral blood, the expression of tumor markers, and micrometastases, as well as predict prognoses [8].

In the present study, we detected the *KRT19* mRNA-marked CTC and *CEACAM5* mRNA-marked CTC levels in preoperative peripheral blood of breast cancer patients using RT-qPCR and analyzed the differences of these biomarkers expression in different pathological stages of breast cancer using the χ^2 test. In metastatic breast cancer, the rate of $CTC_{KRT19-pos}$ is significantly higher than that of intraductal carcinoma of the breast and invasive breast cancer (stage I–III), although the *P* value ($P = 0.077$) is not statistically significant. And this result was similar to those of a previous study [18, 19]. The same is true for the rate of $CTC_{CEACAM5-pos}$ in these five groups of breast cancer patients and the *P* value ($P = 0.004$) obtained is statistically significant. Our results were similar to the reports by Michiko Imamura et al. [20] and Shao et al. [16]. The result suggested that quantitative monitoring of peripheral blood *KRT19* and *CEACAM5* mRNA-marked CTC levels may suggest metastasis of breast cancer, and *KRT19* and *CEACAM5* may be potential biomarkers for distant metastasis in breast cancer patients.

We also analyzed the correlation between CTC_{KRT19}, $CTC_{CEACAM5}$ and clinicopathological variables. Our results showed significant correlation between the preoperative $CTC_{KRT19-pos}$ levels and the metastasis of lymph nodes. Lymph node metastasis is one of the most important prognostic factors for breast cancer patients. However, it is difficult to

Table 5 Univariate and multivariate analyses of the clinicopathological variables for OS in breast cancer patients

Variables	HR (95% CI)	*P*
Univariate		
CTC_{KRT19}		
Positive versus negative	2.11 (1.17–5.42)	0.019
$CTC_{CEACAM5}$		
Positive versus negative	1.85 (0.95–4.60)	0.066
Tumor size (cm)		
> 2 versus ≤ 2	2.19 (1.31–3.66)	0.003
Pathological type		
DCIS versus Invasive carcinoma	23.37 (0.26–2084.69)	0.169
Histological grade		
III versus I–II	1.25 (0.73–2.16)	0.417
Lymph node status		
Positive versus negative	1.73 (1.35–2.23)	< 0.0001
ER status		
Positive versus negative	0.82 (0.43–1.58)	0.560
PR status		
Positive versus negative	0.75 (0.41–1.37)	0.341
HER2 status		
Positive versus negative	1.32 (0.66–2.61)	0.431
Multivariate		
CTC_{KRT19}		
Positive versus negative	1.92 (1.01–3.63)	0.046

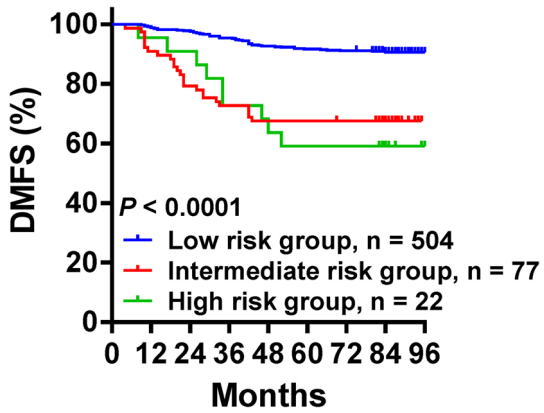

Fig. 5 Cumulative survival curves for 603 patients with breast cancer according to risk groups by Kaplan–Meier survival analysis. The 7-year DMFS rates for the low-, intermediate-, and high-risk groups were 90.7%, 67.5%, and 59.1%, respectively ($P = 0.000$)

detect lymph node micrometastasis by routine pathological examination. In our study, the result showed that $CTC_{KRT19-pos}$ strongly correlated with lymph node metastasis. Our results were similar to the reports by Yu et al. [21]. *KRT19*, a cytoskeletal protein, is derived from epithelial cells and is expressed in normal epithelial and epithelial primary tumors and metastatic tumor cells. *KRT19* expression is absent in normal peripheral blood and lymphoid tissues [9]. If the *KRT19* transcript is detected in lymph nodes of patients with epithelial tumors, the presence of disseminated cancer cells could be considered [22]. Therefore, our result indicated that preoperative *KRT19* mRNA levels may serve as potential marker to detect lymph node metastasis in breast cancer patients.

KRT19, a tissue-specific marker for epithelial tumor micrometastasis, has been reported to be a prognostic marker in esophageal cancer [23], hepatocellular carcinoma [24], gastric cardia cancer [22], and lung cancer [25]. Several studies have suggested the potential role of *CEACAM5* in monitoring disease recurrence and treatment response as well as in predicting the prognosis of many malignancies [17, 26]. Based on the previously published data of *KRT19* and *CEACAM5* in other malignancies, we further explored the relationships between preoperative *KRT19*/ *CEACAM5* mRNA-marked CTC levels with LRFS, DMFS, and OS in patients with breast cancer. To analyze the prognostic value of preoperative *KRT19* and *CEACAM5* mRNA-marked CTC levels in breast cancer patients, Kaplan–Meier survival analysis and multivariate analysis were performed in the following analysis. The Kaplan–Meier survival analysis showed that high preoperative $CTC_{KRT19-pos}$ level was correlated with a poor LRFS and OS in breast cancer patients. Our results were similar to the reports by Konstantinos Tryfonidis et al. [25], who demonstrated that preoperative *KRT19* mRNA level in peripheral blood was correlated with poor prognosis. The Univariate analysis showed that preoperative $CTC_{KRT19-pos}$ and $CTC_{CEACAM5-pos}$ levels were independent risk factors for DMFS in breast cancer patients. In a word, high level of double-marked CTC by *KRT19* and *CEACAM5* mRNA indicated unfavorable survival of breast cancer patients.

In our study, we have a conclusion that preoperative $CTC_{KRT19-pos}$, $CTC_{CEACAM5-pos}$, and lymph node metastasis were the independent prognostic factors for distant metastasis-free survival (DMFS). And in clinical practice, we noticed that breast cancer patients with these risk factors tended to have a poorer prognosis. Therefore, we proposed a prognostic model based on these risk factors and classified breast cancer patients into low-, intermediate- and high-risk groups. We further compared the survival curves of the three groups and found that there was a significant difference in DMFS among the three different risk groups. This prognostic model can be easily constructed, and it may potentially help clinicians make a more accurate judgment for prognosis and individual therapeutic treatment to improve survival and quality of life based on the risk stratification. Breast cancer patients with a high-risk score may benefit from closer monitoring or more aggressive postoperative adjuvant therapy.

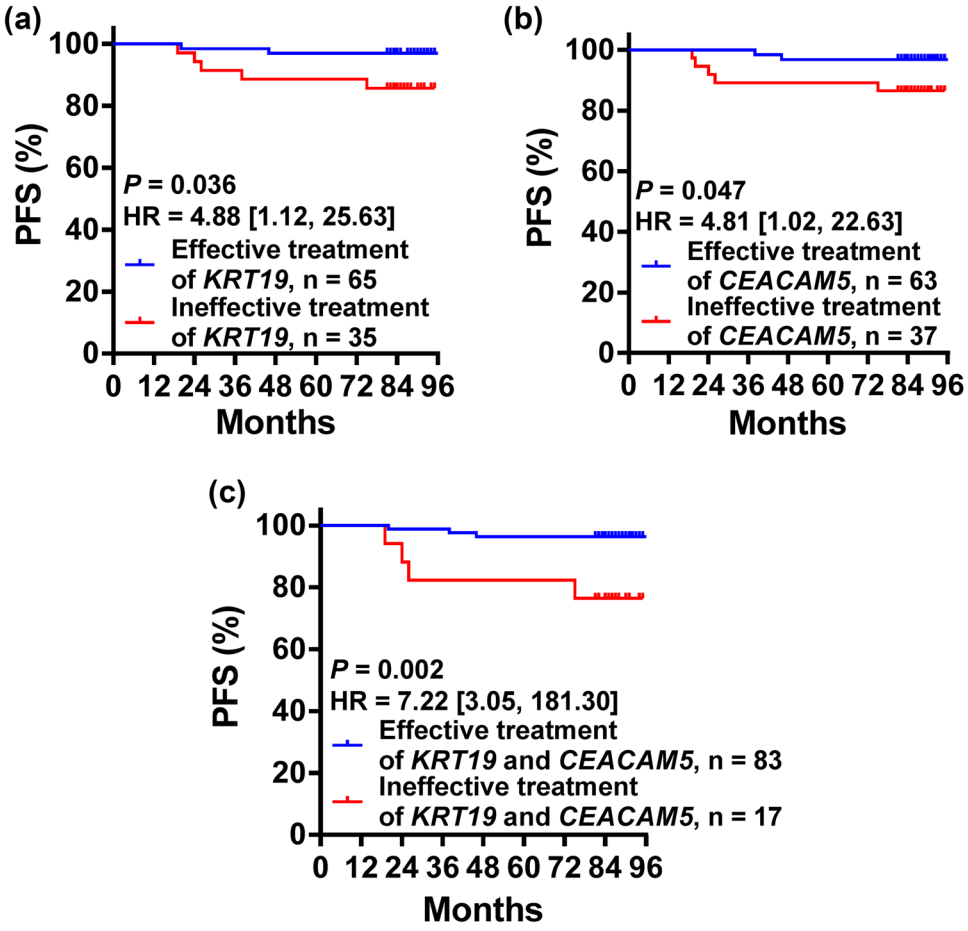

Fig. 6 Kaplan–Meier survival curves according to changes of *KRT19* and *CEACAM5* mRNA levels before and after treatment. **a** *KRT*19 mRNA level for PFS; **b** *CEACAM5* mRNA level for PFS; **c** combined *KRT*19 and *CEACAM5* mRNA levels for PFS; *P* values were calculated by the log-rank test and $P < 0.05$ denoted significance

In our study, we found that patients with a significant reduction in *KRT19* and *CEACAM5* mRNA in peripheral blood after treatment had a better prognosis than patients who did not change or increased. And combining these two indicators, the prediction ability of PFS is better than single *KRT19* or *CEACAM5* ($P = 0.002$ compare with $P = 0.036$ or $P = 0.047$). Therefore, measuring the value of *KRT19* and *CEACAM5* mRNA in peripheral blood after the end of the last chemotherapy and comparing these two indicators with the preoperative peripheral blood, we can measure the therapeutic effect and predict the prognosis, thus guiding the clinical jobs.

In summary, double-marked CTC by *KRT19* and *CEACAM5* mRNA is a prognostic index of breast cancer patients before surgery and after chemotherapy. CTC $_{KRT19-pos}$ indicates lymph node statues of preoperative patients. Therefore, the RT-qPCR-based molecular diagnosis of CTC could be used for prognostic prediction of breast cancer patients and guiding clinical treatment. More importantly, the new proposed prognostic model may help clinicians provide better individual therapeutic approaches and improve the outcome of patients with breast cancer.

Acknowledgements The authors would like to thank Dr. Yumei Feng (Tianjin Medical University Cancer Research Institute and Hospital Department of Biochemistry and Molecular Biology, Tianjin, China) for the useful discussion and insightful observation during the research process. Thanks to Tianjin Medical University Cancer Research Institute and Hospital laboratory of Biochemistry and Molecular Biology for technical assistance in RT-qPCR.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the respective institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Howell A, Anderson AS, Clarke RB, Duffy SW, Evans DG, Garcia-Closas M, Gescher AJ, Key TJ, Saxton JM, Harvie MN (2014) Risk determination and prevention of breast cancer. Breast Cancer Res 16:446. https://doi.org/10.1186/s13058-014-0446-2
2. Yates LR, Knappskog S, Wedge D, JHR F, Gonzalez S, Martincorena I, Alexandrov LB, Van Loo P, Haugland HK, Lilleng PK, Gundem G, Gerstung M, Pappaemmanuil E, Gazinska P, Bhosle SG, Jones D, Raine K, Mudie L, Latimer C, Sawyer E, Desmedt C, Sotiriou C, Stratton MR, Sieuwerts AM, Lynch AG, Martens JW, Richardson AL, Tutt A, Lønning PE, Campbell PJ (2017) Genomic evolution of breast cancer metastasis and relapse. Cancer Cell 32:169–184.e7. https://doi.org/10.1016/j.ccell.2017.07.005
3. Bednarz-Knoll N, Alix-Panabières C, Pantel K (2011) Clinical relevance and biology of circulating tumor cells. Breast Cancer Res 13:228. https://doi.org/10.1186/bcr2940
4. Haber DA, Velculescu VE (2014) Blood-based analyses of cancer: circulating tumor cells and circulating tumor DNA. Cancer Discov 4:650–661. https://doi.org/10.1158/2159-8290.CD-13-1014
5. Lowes LE, Bratman SV, Dittamore R, Done S, Kelley SO, Mai S, Morin RD, Wyatt AW, Allan AL (2016) Circulating tumor cells (CTC) and cell-free DNA (cfDNA) Workshop 2016: scientific opportunities and logistics for cancer clinical trial incorporation. Int J Mol Sci. https://doi.org/10.3390/ijms17091505
6. Alix-Panabières C, Pantel K (2014) Challenges in circulating tumour cell research. Nat Rev Cancer 14:623–631. https://doi.org/10.1038/nrc3820
7. Peeters DJ, van Dam PJ, Van den Eynden GG, Rutten A, Wuyts H, Pouillon L, Peeters M, Pauwels P, Van Laere SJ, van Dam PA, Vermeulen PB, Dirix LY (2014) Detection and prognostic significance of circulating tumour cells in patients with metastatic breast cancer according to immunohistochemical subtypes. Br J Cancer 110:375–383. https://doi.org/10.1038/bjc.2013.743
8. Wallwiener M, Riethdorf S, Hartkopf AD, Modugno C, Nees J, Madhavan D, Sprick MR, Schott S, Domschke C, Baccelli I, Schönfisch B, Burwinkel B, Marmé F, Heil J, Sohn C, Pantel K, Trumpp A, Schneeweiss A (2014) Serial enumeration of circulating tumor cells predicts treatment response and prognosis in metastatic breast cancer: a prospective study in 393 patients. BMC Cancer 14:512. https://doi.org/10.1186/1471-2407-14-512
9. Saloustros E, Perraki M, Apostolaki S, Kallergi G, Xyrafas A, Kalbakis K, Agelaki S, Kalykaki A, Georgoulias V, Mavroudis D (2011) Cytokeratin-19 mRNA-positive circulating tumor cells during follow-up of patients with operable breast cancer: prognostic relevance for late relapse. Breast Cancer Res 13:R60. https://doi.org/10.1186/bcr2897
10. Terrenato I, D'Alicandro V, Casini B, Perracchio L, Rollo F, De Salvo L, Di FS, Di FF, Pescarmona E, Maugeri-Saccà M, Mottolese M, Buglioni S (2017) A cut-off of 2150 cytokeratin 19 mRNA copy number in sentinel lymph node may be a powerful predictor of non-sentinel lymph node status in breast cancer patients. PLoS ONE 12:e0171517. https://doi.org/10.1371/journal.pone.0171517
11. Zhao S, Mei Y, Wang Y, Zhu J, Zheng G, Ma R (2015) Levels of CEA, CA153, CA199, CA724 and AFP in nipple discharge of breast cancer patients. Int J Clin Exp Med 8:20837–20844
12. Hioki T, Sugimura Y (1999) Detection of circulating cancer cells by nested reverse transcription-polymerase chain reaction of cytokeratin-19 in patients with renal cell carcinoma. Hinyokika Kiyo 45:577–581
13. Andergassen U, Hofmann S, Kölbl AC, Schindlbeck C, Neugebauer J, Hutter S, Engelstädter V, Ilmer M, Friese K, Jeschke U (2013) Detection of tumor cell-specific mRNA in the peripheral blood of patients with breast cancer—evaluation of several markers with real-time reverse transcription-PCR. Int J Mol Sci 14:1093–1104
14. Edge SB, Compton CC (2010) The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 17:1471–1474. https://doi.org/10.1245/s10434-010-0985-4
15. Scott A, Ambannavar R, Jeong J, Liu ML, Cronin MT (2011) RT-PCR-based gene expression profiling for cancer biomarker discovery from fixed, paraffin-embedded tissues. Methods Mol Biol 724:239–257. https://doi.org/10.1007/978-1-61779-055-3_15
16. Shao Y, Sun X, He Y, Liu C, Liu H (2015) Elevated levels of serum tumor markers CEA and CA15-3 are prognostic parameters for different molecular subtypes of breast cancer. PLoS ONE 10:e0133830. https://doi.org/10.1371/journal.pone.0133830
17. Yang Y, Zhang H, Zhang M, Meng Q, Cai L, Zhang Q (2017) Elevation of serum CEA and CA15-3 levels during antitumor therapy predicts poor therapeutic response in advanced breast cancer patients. Oncol Lett 14:7549–7556. https://doi.org/10.3892/ol.2017.7164
18. Stathopoulou A, Vlachonikolis I, Mavroudis D, Perraki M, Ch K, Apostolaki S, Malamos N, Kakolyris S, Kotsakis A, Xenidis N, Reppa D, Georgoulias V (2002) Molecular detection of cytokeratin-19-positive cells in the peripheral blood of patients with operable breast cancer: evaluation of their prognostic significance. J Clin Oncol 20:3404–3412. https://doi.org/10.1200/JCO.2002.08.135
19. Park HS, Han HJ, Lee S, Kim GM, Park S, Choi YA, Lee JD, Kim GM, Sohn J, Kim SI (2017) Detection of circulating tumor cells in breast cancer patients using cytokeratin-19 real-time RT-PCR. Yonsei Med J 58:19–26. https://doi.org/10.3349/ymj.2017.58.1.19
20. Imamura M, Morimoto T, Nomura T, Michishita S, Nishimukai A, Higuchi T, Fujimoto Y, Miyagawa Y, Kira A, Murase K, Araki K, Takatsuka Y, Oh K, Masai Y, Akazawa K, Miyoshi Y (2018) Independent prognostic impact of preoperative serum carcinoembryonic antigen and cancer antigen 15-3 levels for early breast cancer subtypes. World J Surg Oncol 16:26. https://doi.org/10.1186/s12957-018-1325-6
21. Yu XF, Yang HJ, Lei L, Wang C, Huang J (2016) CK19 mRNA in blood can predict non-sentinel lymph node metastasis in breast cancer. Oncotarget 7:30504–30510. https://doi.org/10.18632/oncotarget.8851
22. Qiao YF, Chen CG, Yue J, Ma MQ, Ma Z, Yu ZT (2017) Prognostic significance of preoperative and postoperative CK19 and CEA mRNA levels in peripheral blood of patients with gastric cardia cancer. World J Gastroenterol 23:1424–1433. https://doi.org/10.3748/wjg.v23.i8.1424
23. Qiao YF, Chen CG, Yue J, Ma Z, Yu ZT (2016) Clinical significance of preoperative and postoperative cytokeratin 19 messenger RNA level in peripheral blood of esophageal cancer patients. Dis Esophagus 29:929–936. https://doi.org/10.1111/dote.12403
24. Yang XR, Xu Y, Shi GM, Fan J, Zhou J, Ji Y, Sun HC, Qiu SJ, Yu B, Gao Q, He YZ, Qin WZ, Chen RX, Yang GH, Wu B, Lu Q, Wu ZQ, Tang ZY (2008) Cytokeratin 10 and cytokeratin 19: predictive markers for poor prognosis in hepatocellular carcinoma patients after curative resection. Clin Cancer Res 14:3850–3859. https://doi.org/10.1158/1078-0432.CCR-07-4338
25. Ge MJ, Wu QC, Wang M, Zhang YH, Li LB (2005) Detection of disseminated lung cancer cells in regional lymph nodes by assay of CK19 reverse transcriptase polymerase chain reaction and its clinical significance. J Cancer Res Clin Oncol 131:662–668. https://doi.org/10.1007/s00432-005-0009-0
26. Zhang Q, Qu H, Sun G, Li Z, Ma S, Shi Z, Zhao E, Zhang H, He Q (2017) Early postoperative tumor marker responses provide a robust prognostic indicator for N3 stage gastric cancer. Medicine (Baltimore) 96:e7560. https://doi.org/10.1097/MD.0000000000007560