

# Study on the small breast epithelial mucin tumor tissue expression difference and its relationship with survival in triple-negative breast cancer patients in Han, Miao and Buyi ethnic

Liu Liang (✉ [jinshatanwo@163.com](mailto:jinshatanwo@163.com))

Third Affiliated Hospital of Guizhou Medical University

**Ye Weilin**

Third Affiliated Hospital of Guizhou Medical University

**Liu Hui**

Third Affiliated Hospital of Guizhou Medical University

**Lu Qili**

Third Affiliated Hospital of Guizhou Medical University

**Wang Jun**

Third Affiliated Hospital of Guizhou Medical University

**Yang Yan**

Third Affiliated Hospital of Guizhou Medical University

**He Yang**

Third Affiliated Hospital of Guizhou Medical University

---

## Research

**Keywords:** small breast epithelial mucin; breast cancer; DFS; OS; ethnic difference

**Posted Date:** August 11th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-56422/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Objective:** To investigate Small breast epithelial mucin(SBEM) expression difference in mammary tissues and relationship with survival in the triple-negative breast cancer(TNBC) patients in Han, Buyi and Miao ethnic.

**Methods:** In our study, SBEM protein expressions were detected by means of immunohistochemistry in 297 patients diagnosed from 2014 to 2015, including 99 normal breast tissue specimens, 99 cases of breast benign tumor tissue specimens and 99 tissues specimens from TNBC patients. Each set of tissue specimens contains 33 samples from Han, Miao and Buyi ethnics, respectively. We analyze the expression of SBEM in different mammary tissues in different ethnics and the association of different SBEM expression levels in tissue of TNBC patients with clinical- pathological features, DFS and OS in Han, Miao and Buyi ethnics, respectively.

**Results:** SBEM expression in breast cancer tumor cells were related to the ki67 in the Han, Miao and Buyi ethnic ( $P=0.034$ – $0.027$ – $0.047$ ), respectively. There was a marked associations between the SBEM expression level and lymphatic metastasis ( $P= 0.042$ – $0.039$ ) in the Han and Miao ethnic, while the same results were not found in the Buyi people ( $P=0.072$ ). There was significant difference in DFS ( $P =0.028$ – $0.013$ ) and OS ( $P=0.09$ – $0.037$ ) between the high expression group and low expression group in the Han and Miao ethnic. But there was no significant difference in DFS ( $P= 0.053$ ) and OS ( $P= 0.088$ ) in the Buyi people.

**Conclusion:** The SBEM positive detection rate was no significant difference in the Han, Miao and Buyi ethnic groups. SBEM was associated with DFS and OS in the Han and Miao ethnic, while no correlation in the Buyi ethnic.

## Introduction

Breast cancer is the most common malignancy and is second only to lung cancer in mortality among women worldwide<sup>[1, 2]</sup>. While the incidence rate of breast cancer continues to rise over the last few decades, its mortality rate, on the other hand, declines every year thanks for the development of diagnosis and targeting medication<sup>[3, 4]</sup>. Triple-negative breast cancer (TNBC), one of five molecular subtypes recognized in 2000<sup>[5, 6]</sup>, lacks estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) expression. TNBC is characterized by high incidence in young women, early recurrence and shows a relative sensitivity of chemotherapy. Most studies<sup>[7–9]</sup> showed that the prognosis of TNBC was less favorable than that of non-TNBC.

Micro metastasis is considered to be a key factor for poor prognosis of breast cancer patients<sup>[10]</sup>. Despite the improvement in detection of breast cancer, about 30% of patients are still detected with micro metastasis at their first visit. The micro metastasis of breast cancer before primary treatment is considered to be one of the recurrence reasons that will directly impact the survival of patients. The

development of distant metastasis is the major cause of their death<sup>[11]</sup>. So, it is crucial to find specific markers to detect micro metastasis and provide useful information to guide early therapeutic methods of breast cancer patients, especially TNBC.

Small breast epithelial mucin (SBEM) was a tissue specific protein and only expresses in mammary and salivary glands<sup>[12]</sup>. High SBEM expression was found to be strongly associated with the histopathological detection of lymph node metastasis<sup>[12]</sup>. SBEM could serve as a useful marker for breast micro metastasis, also as a specific target for the treatment of breast cancer<sup>[12]</sup>. O'Brien<sup>[13]</sup> confirmed that rates of detection of SBEM in protein levels were 51% (52/103) and 4% (1/26) in breast cancer and non-breast cancer tissue, respectively, which further confirmed the SBEM expression was significantly higher in breast cancer tissue than non-breast tissue. Skliris GP<sup>[14]</sup> study showed that SBEM expression rate was 18% in 300 patients with breast cancer using the method of immunohistochemical analysis and the SBEM expression rate was significantly higher in ER negative (22%) than in ER positive (13%) breast cancer tissue. Several laboratories showed that SBEM expression correlated with higher tumor grade, TNM staging and lymph node metastasis at both mRNA and protein levels.

SBEM has significant guidance for clinical activities. Liu ZZ<sup>[15]</sup> reported that SBEM has the potential to be a specific marker for predicting hematogenous micro metastasis and response to neo-adjuvant chemotherapy in breast cancer. Valladares-Ayerbes M et al. <sup>[16]</sup> studied the expression profiles of SBEM gene in silico and in vitro, and demonstrated that SBEM-mRNA could serve as a marker for bone marrow micro metastasis in breast cancer patients and SBEM-mRNA could serve as a marker for bone marrow micro metastasis in breast cancer patients <sup>[15]</sup>. Research of Liang Liu et al <sup>[17]</sup> suggested SBEM 3+ score was cut-off value of prognosis and significantly correlated with decreased DFS and OS in TNBC patients. SBEM is an independent risk predictor and may offer utility as a prognostic marker in TNBC patients <sup>[17]</sup>.

Based on the results above, SBEM might play an important role in progression and metastasis of breast cancer in different races, especially in TNBC. But humans have genetic diversity and every race and nationality's genome has its own characteristics. Only to figure out what each nation's genetic structure characteristics and the change rule, we can analyze their origins and the genetic relationship with other ethnic groups. This kind of population genetics research is a difficult project. The genetics investigation need to rely on all sorts of genetic markers, especially those with racial specificity and individual specific genetic markers, which can provide enough valuable genetic information.

The Southern GuiZhou is a multi-ethnic area of China, of which Han, Miao and Buyi ethnics are its main components. The breast cancer incidence here is increasing year by year as well as other regions in China. Because the different ethnic groups have different genetic background and ways of life, we are not sure whether there is difference between different ethnic groups, including the SBEM expression level and the prognostic significance in TNBC. And to the best of our knowledge, this is the first study to investigate the differences of the expression of SBEM in mammary tissue of Han, Miao and Buyi ethnic

and prognosis in TNBC patients in different ethnics. The aim of this study was to analyze whether there is the same expression, prediction results or the clinical meaning for SBEM in different ethnics.

## Materials And Methods

The study examined cases from 297 patients diagnosed from 2014 to 2015 in Third Affiliated Hospital of Guizhou Medical University, including 99 normal breast tissue specimens, 99 cases of breast benign tumor tissue specimens and 99 available formalin-fixed paraffin-embedded tissues specimens from TNBC patients by means of immunohistochemistry (IHC). Each set of tissue specimens contains 33 samples from Han, Miao and Buyi ethnics, respectively. Clinical data of all the cases were reviewed retrospectively from medical records in our hospital. All patients were females and had a minimum 4 years' follow-up records. The follow-up period end at 2020.06.30. All the TNBC patients underwent operational treatment according to clinical practice guidelines of National Comprehensive Cancer Network (NCCN) of the United States. None of the patients received neoadjuvant therapy. Statistic and analysis of clinicopathological parameters, including age at diagnosis, disease stage, tumor size, tumor grade, lymph node status, P53 and Ki67, were listed in Table 2–4. All TNBC cases examined were ER and PR negative by IHC. HER2 status was considered negative if the immunohistochemical score was 0 or 1+, or if the score was 2+ but non-amplification by fluorescence in situ hybridization (FISH), and positive if the score was 3+. Based on the past research results, SBEM expression with SBEM 3+ score is the SBEM cut-off value of prognosis. We divided the cases into two groups, one is the SBEM < 3+ group defined as a high expression group, the other is SBEM = 3+ group defined as a high expression group.

The study examined cases from 297 patients diagnosed from 2014 to 2015 in Third Affiliated Hospital of Guizhou Medical University, including 99 normal breast tissue specimens, 99 cases of breast benign tumor tissue specimens and 99 available formalin-fixed paraffin-embedded tissues specimens from TNBC patients by means of immunohistochemistry (IHC). Each set of tissue specimens contains 33 samples from Han, Miao and Buyi ethnics, respectively. Clinical data of all the cases were reviewed retrospectively from medical records in our hospital. All patients were females and had a minimum 4 years' follow-up records. The follow-up period end at 2020.06.30. All the TNBC patients underwent operational treatment according to clinical practice guidelines of National Comprehensive Cancer Network (NCCN) of the United States. None of the patients received neoadjuvant therapy. Statistic and analysis of clinicopathological parameters, including age at diagnosis, disease stage, tumor size, tumor grade, lymph node status, P53 and Ki67, were listed in Table 2–4. All TNBC cases examined were ER and PR negative by IHC. HER2 status was considered negative if the immunohistochemical score was 0 or 1+, or if the score was 2+ but non-amplification by fluorescence in situ hybridization (FISH), and positive if the score was 3+. Based on the past research results, SBEM expression with SBEM 3+ score is the SBEM cut-off value of prognosis. We divided the cases into two groups, one is the SBEM < 3+ group defined as a high expression group, the other is SBEM = 3+ group defined as a high expression group.

### Sbem Expression And Evaluation Of Ihc

All tissues were collected surgically under the supervision of an experienced pathologist. SBEM expression was measured by IHC. Streptavidinperoxidase (S-P) IHC staining was performed using SBEM antibody of mouse monoclonal (diluted 1/800, Abcam plc. Cambridge, UK). The detailed procedures were done as described by Jennbacken<sup>[18]</sup>. PBS was used to replace the primary antibody in negative controls. SBEM was a secreted protein and it was mainly located in cell membrane, secondly in cytoplasm. Normal breast tissues were in general weakly or negative. So it was evaluated and scored if cell membrane and/or cytoplasm reactivity were observed<sup>[8]</sup>. According to our data and TMA IHC grading method by Serrero G<sup>[19]</sup> Pan<sup>[20]</sup> and Liu Liang<sup>[17]</sup>, our scoring was semiquantitatively categorized as:  $\leq 5\%$  of tumor cells staining with/without weakly stained was negative (0), followed by a score of 1 ( $> 5\%$  of tumor cells and with weak/focal positive staining or  $\leq 5\%$  of tumor cells with strongly stained), 2 ( $> 5\%$  of tumor cells and with moderate/focal positive staining), 3 ( $> 5\%$  of tumor cells and with strong/diffuse positive staining).

### Statistical analysis

The correlation between SBEM, Clinicopathological characteristics and survival outcomes was compared by Pearson's  $X^2$  test. Survival analyses, including DFS and OS, were performed with the log rank test and all results were displayed in Kaplan–Meier. DFS was defined as the time interval from date of diagnosis to the time of last disease free follow-up or at death for those patients who died without a previous recurrence<sup>[21]</sup>. OS was defined as the time interval from date of diagnosis to time of last follow-up or death<sup>[21]</sup>. Statistical significance was defined as P value  $< 0.05$ . SPSS17.0 software package was used for all statistical analyses.

## Results

Figure 1 Immunohistochemical staining for SBEM. A: 0+:  $\leq 5\%$  of tumor cells staining with/without weakly stained was negative (original magnification  $\times 200$ ); B: 1+:  $> 5\%$  of tumor cells and with weak/focal positive staining or  $\leq 5\%$  of tumor cells with strongly stained (original magnification  $\times 100$ ); C: 2+:  $> 5\%$  of tumor cells and with moderate/focal positive staining (original magnification  $\times 200$ ); D: 3+:  $> 5\%$  of tumor cells and with strong/diffuse positive staining (original magnification  $\times 200$ ).

Table 1 showed the SBEM positive detection rate in the breast cancer tissue, breast benign tumor tissue and normal breast tissue in the han, miao and buyi ethnic (Table 1). SBEM positive detection rate of breast cancer tissue is significantly higher than the breast benign tumor tissue and normal breast tissue ( $P = 0.005$ ). But there was no significant difference in the han, miao and buyi ethnic in breast cancer tissue, breast benign tumor tissue and normal breast tissue ( $P > 0.05$ ) (Table 1).

Table 1  
SBEM positive detection rate in the breast cancer tissue, breast benign tumor tissue and normal breast tissue in the han, miao and buyi ethnic

Tissue type	SBEM positive detection rate (%)			
	Han ethnic	Miao ethnic	Buyi ethnic	P $\square$
Breast cancer tissue	49	47	45	0.982
Breast benign tumor tissue	4	3	3	0.923
Normal breast tissue	3	2	2	0.889
P $\square$	0.005	0.005	0.005	

The clinicopathological characteristics of patients were described in Table 2–4. SBEM expression in TNBC tumor cells were related to ki67 in the Han, Miao and Buyi ethnic ( $P = 0.034 \square 0.027 \square 0.047$ ), respectively. There was a marked association between the low expression of SBEM group and high expression group in lymphatic metastasis ( $P = 0.042 \square 0.039$ ) in the Han and Miao ethnic, while the same results were not found in the Buyi people ( $P = 0.072$ )(Table 2,3,4).

Table 2  
Clinicopathological characteristics of patients in Han ethnic

Parameter	Number (n)	Subgroup cut-offs	SBEM < 3		SBEM = 3		P <sup>†</sup>
			Number (n)	%	Number (n)	%	
Age	33	X > 35	26	73.1	7	57.1	0.614
		X ≤ 35	19	26.9	4	42.8	
			7		3		
TNM staging	33	Ⅰ(1)	26	23.1	7	14.3	0.075
		Ⅰ(2)	6	53.8	1	42.8	
		Ⅰ(3)	14	23.1	3	42.8	
			6		3		
P53	31	Mutated	24	45.8	7	57.1	0.801
		No-mutated	11	44.2	4	42.8	
			13		3		
Lymph Node	33	+	26	38.5	7	71.4	0.042
			10	61.5	5	28.6	
			16		2		
Grade	30	Low (1)	23	39.1	7	28.6	0.212
		Mod(2)	9	34.8	2	42.8	
		High(3)	8	26.1	3	28.6	
			6		2		
Size	33	X > 20 mm	26	53.8	7	57.1	0.872
		X ≤ 20 mm	14	46.2	4	42.8	
			12		3		
Ki67	32	X > 35	26	42.3	6	83.3	0.034
		X ≤ 35	11	57.7	5	16.7	
			15		1		

Table 3  
Clinicopathological characteristics of patients in Miao ethnic

Parameter	Number (n)	Subgroup cut-offs	SBEM < 3		SBEM = 3		P <sup>1</sup>
			Number (n)	%	Number (n)	%	
Age	33	X > 35	26	84.6	7	71.4	0.523
		X ≤ 35	22	15.4	5	28.6	
			4		2		
TNM staging	33	Ⅰ(1)	26	26.9	7	14.3	0.092
		Ⅰ(2)	7	53.8	1	63.6	
		Ⅰ(3)	13	19.3	4	28.6	
			6		2		
P53	27	Mutated	21	47.6	6	50.0	0.901
		No-mutated	10	52.4	3	50.0	
			11		3		
Lymph Node	33	+	26	42.3	7	85.7	0.039
			11	57.7	6	14.3	
			15		1		
Grade	30	Low (1)	24	37.5	6	33.3	0.361
		Mod(2)	9	33.3	2	33.3	
		High(3)	8	29.2	2	33.3	
			7		2		
Size	33	X > 20 mm	26	65.4	7	63.6	0.849
		X ≤ 20 mm	17	34.6	4	36.4	
			9		3		
Ki67	31	X > 35	24	19.2	7	71.4	0.027
		X ≤ 35	7	70.8	5	28.6	
			17		2		

Table 4  
Clinicopathological characteristics of patients in Buyi ethnic

Parameter	Number (n)	Subgroup cut-offs	SBEM < 3		SBEM = 3		P <sup>†</sup>
			Number (n)	%	Number (n)	%	
Age	33	X > 35	25	72.4	8	63.6	0.806
		X ≤ 35	21	27.6	6	36.4	
			4		2		
TNM staging	33	Ⅰ(1)	25	24.0	8	12.5	0.101
		Ⅰ(2)	6	48.0	1	37.5	
		Ⅰ(3)	12	32.0	3	50.0	
		8		4			
P53	29	Mutated	22	59.1	7	63.6	0.678
		No-mutated	13	40.9	4	36.4	
			9		3		
Lymph Node	33	+	25	40.0	8	62.5	0.072
			10	60.0	5	37.5	
			15		3		
Grade	26	Low (1)	20	40.0	6	50.0	0.201
		Mod(2)	8	35.0	3	33.3	
		High(3)	7	25.0	2	16.7	
			5		1		
Size	33	X > 20 mm	25	56.0	8	62.5	0.615
		X ≤ 20 mm	14	44.0	5	37.5	
			12		3		
Ki67	30	X > 35	23	39.1	7	63.6	0.047
		X ≤ 35	9	60.9	4	36.4	
			14		3		

Figure 2 Kaplan-Meier estimates for DFS by SBEM scores in Han, Miao and Buyi ethnic

## Discussion

Breast cancer is pushed into first place in the United States and many other parts of world. Breast cancer alone is expected to account for 29% (226,870) of all new cancer cases among women [22]. Although incidence rate of breast cancer remains relatively stable in recent 5 years, its death rate declines by 34% because of the development of diagnosis and targeting medication [4, 10]. The micro metastasis of breast cancer before primary treatment is considered to be one of the recurrence reasons that will directly impact the survival of patients. So, it is crucial to find specific markers to detected micro metastasis and provide useful information to guide early therapeutic methods of TNBC patients.

The definition of tumor markers is broad. Tumor cells may express certain molecules at a different rate from that of normal cells, and these substances are released in the bloodstream or in other biologic fluids. Although various biological markers had been proposed for the detection of breast cancer cells, they were often affected by tumor differentiation, lower specificity and detection rate. But almost all the current clinical application of tumor markers can not reach ideal level in the identification of the tumor specificity and sensitivity, especially in TNBC at present.

SBEM was a tissue specific protein, a member of the MUC family, and its expression is highly specific to mammary gland tissue. High SBEM expression was found to be strongly associated with the histopathological detection of lymph node metastasis [12]. SBEM was a secreted protein and it was mainly located in cell membrane. Normal breast tissues were in general weakly or negative. SBEM could serve as a useful marker for breast micro metastasis, also as a specific targets for the treatment of breast cancer [12].

However, as the genomes of different races and nationalities have their own characteristics, it is of guiding significance to explore the genetic characteristics of different ethnics in the diagnosis and treatment of TNBC. The Buyi and Miao ethnic groups are the main components in Guizhou province in China. With different genetic backgrounds and lifestyles from those of the Han ethnic. As far as we knew the difference in the expression level of SBEM among different ethnic groups and its prognosis significance for survival have not been reported. The aim of this research was to analyze whether there is the same expression in different mammary tissues in different ethnics and the association of SBEM expression in tissue of TNBC patients with clinical- pathological features, DFS and OS, and to identify whether there is the same prediction results or the clinical meaning for SBEM in Han, Miao and Buyi ethnics.

Researchers have conducted many studies on SBEM as a tumor marker of breast cancer micro metastasis worldwide. RT-PCR and immunohistochemistry are used to detect the expression of gene and protein nowadays [15, 23]. The research of O'Brien et al [13] has shown that SBEM are expressed relatively exclusively in breast tissue than non-breast tissue and are potential new markers for breast cancer [14]. Skliris GP [14] study showed that the SBEM expression rate was significantly higher in ER

negative(22%) than in ER positive (13%) breast cancer tissue.Zhong Lei et al<sup>[24]</sup>study showed that SBEM expression rate was 53.3% in 60 patients with breast cancer the negative SBEM- mRNA results were found in peripheral blood in 20 patients with mammary gland fibroma and 10 cases of healthy people. SBEM expression in TNBC tumor cells were related to TNM staging and axillary lymph node metastasis. The similar results were found by Yang Hua-wei et al<sup>[25]</sup>.This research shows that SBEM is mainly expressed in the cell membrane, the second expression in the cytoplasm, few expression in the nucleus. SBEM positive detection rate of breast cancer tissue is significantly higher than the breast benign tumor tissue and normal breast tissue. But there was no significant difference in the han, miao, buyi ethnic in breast cancer tissue, breast benign tumor tissue and normal breast tissue.

Many factors can affect the prognosis of breast cancer recurrence and death, including tumor histologic features, clinical and pathological features of primary tumor, lymph node status, tumor hormone levels,Her-2 state, complications, age, menopausal status, tumor size, tumor grade, TNM staging, Ki67 and P53, which can evaluate DFS and OS<sup>[26-27]</sup>. Ki67 is also a important marker of cell proliferation, and it has important significance for prognosis judgment in breast cancer<sup>[28]</sup>. In our study, SBEM expression in TNBC tumor cells were related to the ki67 in the Han, Miao and Buyi ethnic (P = 0.034,0.027,0.047), respectively. There was a marked associations between the low expression of SBEM group and high expression group in lymphatic metastasis (P = 0.042,0.039) in the Han and Miao ethnic, while the same results were not found in the Buyi people (P = 0.072)).Thus it can be seen that there is difference in different ethnic groups between SBEM different expression levels and metastatic tumor recurrence and the prognosis.

The past studies have shown that SBEM correlated with the prognosis of patients.

Breast cancer stem cell can obtain higher resistance and high invasive due to their high plasticity<sup>[29]</sup>. The previous studies have not been reported whether there is the same prediction results or the clinical meaning for the same tumor marker in different ethnics.The experiment has carried on the preliminary exploration to this. In our study, we found that DFS and OS function curves showed the large separation between the high expression group and low expression group. There was significant difference in DFS ( P = 0.028,0.013)and OS ( P = 0.09,0.037)between the high expression group and low expression group in the Han and Miao ethnic.But there was no significant difference in DFS (P = 0.053 ) and OS (P = 0.088 )in the Buyi people.Thus it can be seen there is different prognostic significance in different ethnic groups for SBEM different expression levels.

In conclusion, this research shows that SBEM positive detection rate was no significant difference in the han, miao and buyi ethnic in breast cancer tissue, breast benign tumor tissue and normal breast tissue.The SBEM expression level is related to the prognosis of TNBC patients and its clinical significance is not exactly the same in different ethnics ,which can provide a more reasonable choice for clinical individualized treatment and prognostic judgement and reduce a certain amount of medical costs. However, it needs more centers to participate in further study because this is just a retrospective single-

center study and the small sample size limits to some extent the generalization of the findings made in the study.

## Conclusion

We have demonstrated that the SBEM positive detection rate was no significant difference in the han, miao and buyi ethnic groups in breast cancer tissue, breast benign tumor tissue and normal breast tissue. The expression of SBEM does significantly correlate with a DFS and an OS of TNBC patients in the han, miao and buyi ethnic. But it is not obvious for prognostic significance in Buyi ethnic. The distinction in different ethnic groups shall certainly provide a more reasonable choice for clinical individualized treatment and prognostic judgement in TNBC patients.

## Abbreviations

SBEM: Small breast epithelial mucin; TNBC: Triple-negative breast cancer; DFS: Disease-free survival; OS: Overall survival; ER: Estrogen receptor; PR: Progesterone receptor; HER2: Human epidermal growth factor receptor-2; TNM: Tumor node metastasis; IHC: immunohistochemical; PB: Peripheral blood; RT-PCR: Reverse transcription polymerase chain reaction; FISH: Fluorescence in situ hybridization; NCCN: National comprehensive cancer network; CEA: Carcinoembryonic antigen.

## Declarations

### Consent

Written informed consent was obtained from the patient for publication of this report and any accompanying images.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

LL contributed to the study design, data acquisition and analysis and drafted the manuscript; WL was involved in data acquisition and revision of the manuscript; QL and YY worked on aspects of data acquisition and analysis; HY contributed to the study design and data-analysis, developed the algorithm and coordinated the study; LH conceived and coordinated the study. All authors read and approved the final manuscript.

### Acknowledgments

We thank Dr. Xie Xiaodogn and Dr. Jing Peng for expert technical assistance with IHC. We thank Dr. Yang Jie, Dr. Liu Hui and Dr. You Hui for secretarial and organizational support in our experiments. We also thank Dr. Liu Hongshuo and Dr. Wang Xu for critical revision of the manuscript.

## Author details

Department of Oncology, Third Affiliated Hospital of Guizhou Medical University

## References

- [1] Siegel R L, Miller K D, Jemal A. Cancer statistics, 2020. *CA Cancer J Clin*, 2020, 70(1):7-30.
- [2] Ahmedin J, Rebecca S, Elizabeth W, *et al*. Cancer statistics 2007. *CA Cancer J Clin*, 2007, 57:43-66.
- [3] Jemal A, Siegel R, Xu J, *et al.*. Cancer statistics, 2010. *CA Cancer J Clin* 2010; 60: 277-300.
- [4] Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomized trials. *Lancet* 2005; 365: 1687-1717.
- [5] Perou CM, Sørlie T, Eisen MB, *et al.* Molecular portraits of human breast tumors. *Nature* 2000, 406:747–752.
- [6] Sørlie T, Perou CM, Tibshirani R, *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 2001, 98:10869–10874.
- [7] Carey LA, Perou CM, Livasy CA, *et al.* Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006, 295:2492–2502.
- [8] Liedtke C, Mazouni C, Hess KR, *et al.* Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. *J Clin Oncol* 2008, 26:1275–1281.
- [9] Millar EK, Graham PH, O'Toole SA, *et al.* Prediction of local recurrence, distant metastases, and death after breast conserving therapy in early-stage invasive breast cancer using a five-biomarker panel. *J Clin Oncol* 2009, 27:4701–4708.
- [10] Gao J D, Wang J, Zhang B N, *et al*. Correlation of peripheral blood micrometastasis to distant metastasis of breast cancer. *Chinese Journal of Cancer*, 2007, 26(12):1385-1387.
- [11] Bundred NJ. Prognostic and predictive factors in breast cancer [J]. *Cancer Treat Rev*, 2001, 27(3): 137-142.
- [12] Miksicek R J, Myal Y, Watson P H, *et al*. Identification of a novel breast-and salivary gland specific mucin-like gene strongly expressed in normal and tumor human mammary epithelium. *Cancer Res*, 2002, 62(10):2736-2740.

- [13] O'Brien N, O'donovan N, Foley D, *et al*. Use of a panel of novel genes for differentiating breast cancer from nonbreast tissues [J]. *Tumor Biol*. 2007; 28:312-317.
- [14] Skliris GP, Hube F, Gheorghiu I, *et al*. Expression of small breast epithelial mucin (SBEM) protein in tissue microarrays (TMAs) of primary invasive breast cancers. *Histopathology*. 2008; 52(3):355-369.
- [15] Liu ZZ, Xie XD, Qu SX, *et al*. Small breast epithelial mucin (SBEM) has the potential to be a marker for predicting hematogenous micrometastasis and response to neoadjuvant chemotherapy in breast cancer. *Clinical & Experimental Metastasis*, 2010, 27:251-259.
- [16] Valladares-ayerbes M, Iglesias-Diaz P, Diaz-Prado S, *et al*. Diagnostic accuracy of small breast epithelial mucin mRNA as a marker for bone marrow micrometastasis in breast cancer: a pilot study. *J Cancer Res Clin Oncol*. 2009 Sep; 135(9):1185-95.
- [17] Liang Liu, Zhaozhe Liu, Shuxian Qu, *et al*. SBEM tumor tissue expression is associated with increased risk of recurrence in Triple-negative breast cancer patients. *Diagnostic Pathology*. 2013,5(1).
- [18] Jennbacken K, Vallbo C, Wang W, *et al*. Expression of Vascular Endothelial Growth Factor C (VEGF-C) and VEGF Receptor-3 in human prostate cancer are associated with regional lymph node metastasis. *Prostate* 2005, 65(2):110–116.
- [19] Serrero G, Ioffe O. Expression of the novel autocrine growth factor PC-Cell Derived Growth Factor in human breast cancer tissue. *Hum Pathol* 2003, 34:1148–1154.
- [20] Pan AP, Huang GY, Chen J. Relationship between hepatitis B virus covalently closed circular DNA and HBx protein expression in hepatocellular carcinoma and its significance. *World Chin J Digestol* 2009, 17:712–715.
- [21] Serrero G, Hawkins DM, Yue Bet *et al*. Progranulin (GP88) tumor tissue expression is associated with increased risk of recurrence in breast cancer patients diagnosed with estrogen receptor positive invasive ductal carcinoma. *Breast Cancer Res* 2012, 14:R26
- [22] Rebecca Siegel, Deepa Naishadham, Ahmedin Jemal. Cancer statistics, 2012. *CA: A Cancer Journal for Clinicians* 2012, 62(1):10-29.
- [23] Wu Fanglan, Fan Chengxian. SBEM mRNA detection and its clinical meaning in breast cancer patients with bone marrow. *Modern Chinese medicine*. 2009, 22(19): 3473-3476.
- [24] Zhong Lei, Zhang Jianguo, Guo Baoliang, *et al*. The clinical significance of the peripheral blood expression of SBEM - mRNA and CEA in breast cancer patients. *Chinese journal of clinical oncology*, 2008,35(23): 1344-1347.
- [25] Yang HW, Cao J, Yang NW, *et al*. Expression of small breast epithelial mucin mRNA in peripheral blood of breast cancer patients and its clinical significance. *Ai Zheng*. 2005, 24(7), 842-845.

- [26] Ravdin PM, Siminoff LA, *et al.* Computer program to assist in making decision about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001, 19: 980-991.
- [27] Olivotto IA, Bajdik CD, Ravdin PM, *et al.* Population-Based Validation of the Prognostic Model Adjuvant for early breast cancer. *J Clin Oncol* 2005, 23: 2716-2725.
- [28] Weigel MT, Dowsett M. Current and emerging biomarkers in breast cancer: prognosis and prediction. *Endocrine-Related Cancer* 2010, 17: R245-262.
- [29] Lindstrom L S, Yau, C, Czene K, *et al.* Intratumor heterogeneity of the estrogen receptor and the long-term risk of fatal breast cancer. *J Natl Cancer Inst*, 2018, 110(7):726-733.

## Figures

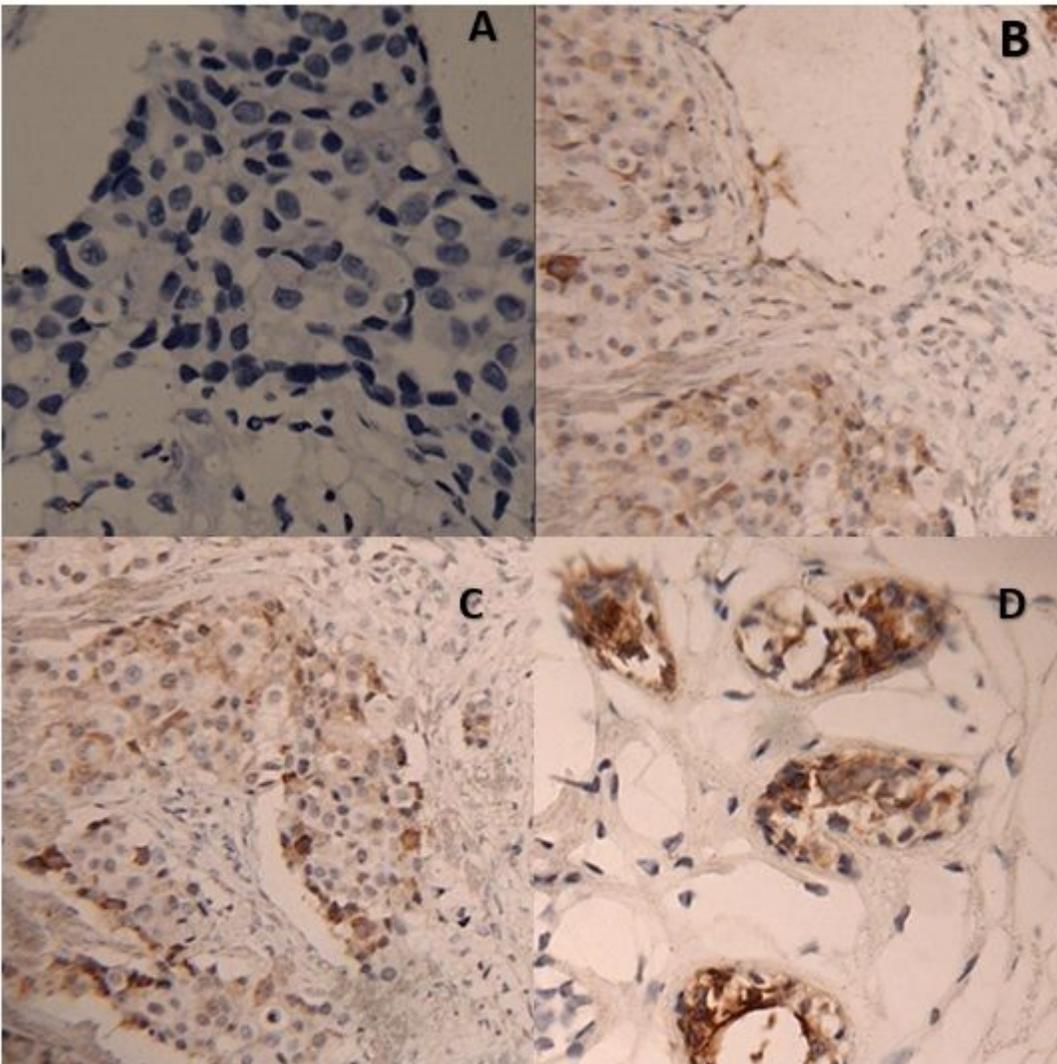


Figure 1

Immunohistochemical staining for SBEM. A: 0+:  $\leq 5\%$  of tumor cells staining with/without weakly stained was negative (original magnification  $\times 200$ ); B: 1+:  $>5\%$  of tumor cells and with weak/focal positive staining or  $\leq 5\%$  of tumor cells with strongly stained (original magnification  $\times 100$ ); C: 2+:  $>5\%$  of tumor cells and with moderate/focal positive staining (original magnification  $\times 200$ ); D: 3+:  $>5\%$  of tumor cells and with strong/diffuse positive staining (original magnification  $\times 200$ ).

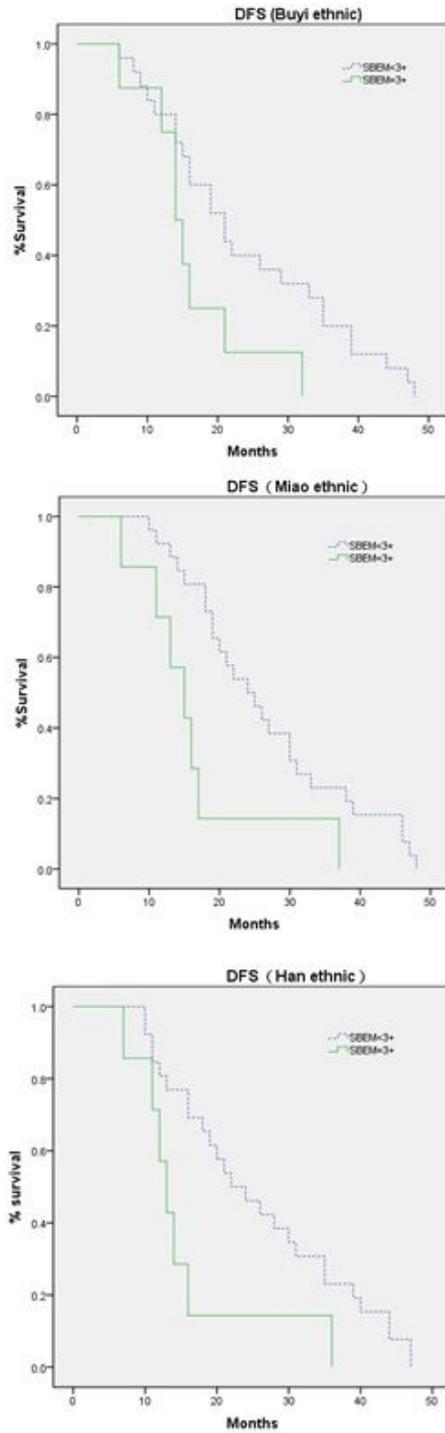


Figure 2

Kaplan-Meier estimates for DFS by SBEM scores in Han, Miao and Buyi ethnic

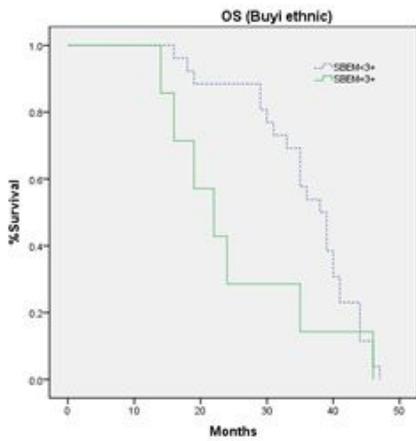
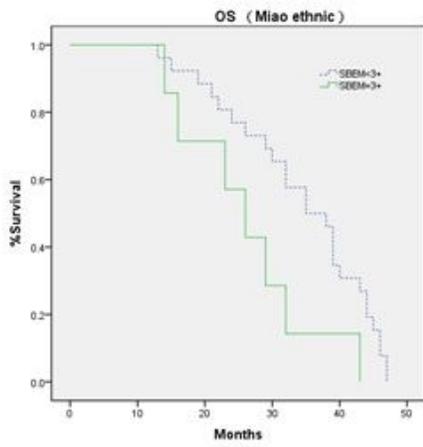
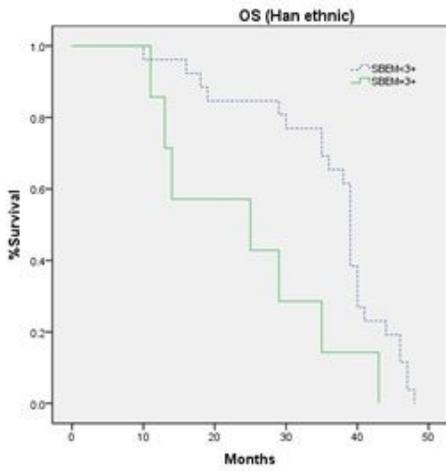


Figure 3

Kaplan-Meier estimates for OS by SBEM scores in Han, Miao and Buyi ethnic