Characterization of triple negative breast cancer gene expression profiles in Mexican patients

ERIC ORTIZ VALDEZ[1], CLAUDIA RANGEL-ESCAREÑO[2], JUAN ANTONIO MATUS SANTOS[1], RAFAEL VÁZQUEZ ROMO[1], ALBERTO GUIJOSA[3], CYNTHIA VILLARREAL-GARZA[4], OSCAR ARRIETA[5], RUBÉN RODRÍGUEZ-BAUTISTA[5], CLAUDIA H. CARO-SÁNCHEZ[6] and ALETTE ORTEGA GÓMEZ[7]

[1]Breast Tumors Department, Mexico's National Institute of Cancer, Sección XVI, Tlalpan, Mexico City 14080; [2]Computational Genomics Department, Instituto Nacional de Medicina Genómica, Arenal Tepepan, Tlalpan, Mexico City 14610; [3]School of Medicine, Universidad Panamericana, Benito Juárez, Mexico City 03920; [4]Breast Cancer Center, Hospital Zambrano Hellion TecSalud, Tecnológico de Monterrey, Real San Agustín, San Pedro Garza García, Nuevo León 66278; [5]Thoracic Oncology Unit, Mexico's National Institute of Cancer; [6]Pathology Department, Mexico's National Institute of Cancer; [7]Laboratory of Translational Medicine, Mexico's National Institute of Cancer, Sección XVI, Tlalpan, Mexico City 14080, Mexico

Received February 3, 2022; Accepted October 5, 2022

DOI: 10.3892/mco.2022.2601

Abstract. Triple negative breast cancer (TNBC) is an aggressive type of cancer that accounts for ~23% of breast tumors in Mexico. In an attempt to understand in an improved way the behavior of TNBC, throughout the years, gene expression in these tumors has been studied. Lehman *et al* identified 6 subtypes of gene expression in TNBC with distinct characteristics. In the present study, it was aimed to assess clinical, pathological and prognostic characteristics of TNBC in a Mexican-based cohort. A total of 55 patients diagnosed with TNBC at Mexico's National Institute of Cancer (INCan) were included. Tumor needle biopsy samples were obtained and subjected to microarray analysis. Patients were thus classified into one of the 6 TNBC molecular subtypes. The prognostic, clinical and pathological information of patients was obtained, and differences across molecular subtypes were sought. Out of the 55 included patients, the following subtypes were identified: 9 basal-like-1, 11 basal-like-2 (BSL2), 16 immunomodulatory (IM), 12 mesenchymal, 6 androgen receptor-like and 1 mesenchymal stem-like. Mean follow-up time was 47.1 months. The IM molecular subtype had the best overall survival (OS) (median OS was not reached). BSL2 had the worst OS (15 months). A complete pathologic response to neoadjuvant chemotherapy was obtained more often in the IM subtype (P=0.032). No significant associations were found between any of the clinical or pathological characteristics and the TNBC molecular subtypes. The results obtained from the present study should be considered when seeking to implement a clinical-molecular model for TNBC patient care, particularly in Hispanic-based populations, as they have been frequently underrepresented in clinical studies assessing TNBC molecular subtypes.

Correspondence to: Dr Alette Ortega Gómez, Laboratory of Translational Medicine, Mexico's National Institute of Cancer, 22 San Fernando Avenue, Sección XVI, Tlalpan, Mexico City 14080, Mexico
E-mail: ortega.alette@gmail.com

Key words: triple negative breast cancer, gene expression profiles, survival, prognosis, clinical, hispanics

Introduction

Breast cancer is the most commonly occurring neoplasm in women. In 2018 alone, 2.1 million new cases were reported worldwide, rendering breast cancer the malignancy with the second highest overall incidence and the first cause of cancer-related death in women. Risk factors for breast cancer range from hormonal and anthropometric to dietary (1).

Triple negative breast cancer (TNBC) does not express hormone receptors and lacks human epidermal growth factor receptor 2 (HER2) overexpression (2). It accounts for 12-17% of all breast tumors (3). In a study of breast cancer patients in Mexico, TNBC was reported to represent between 16 and 23% of breast tumors. The frequency of diagnosis is 72% in stages II and III, and 12.9% in stage IV (4,5).

Clinically, triple negative tumors are more common in young, Hispanic and African-American patients. They frequently present as high-grade tumors with a high risk of recurrence (HR of 4.2 for recurrence compared with non-triple negative tumors), and, similarly, with a high prevalence of BRCA mutations (up to 20% of TNBCs) (2,6-8).

TNBC, along with HER2 positive tumors, present a greater risk of developing distant metastasis at the time of diagnosis (7,9). These metastases preserve the molecular characteristics of the primary tumor that originated them (10).

Main sites for TNBC metastasis are the liver, lung and brain. These are associated with a worse prognosis and have a median survival of 9.0 (3.0-17.0), 11.0 (4.0-20.0) and 6.0 (2.0-13.0) months, respectively. A distinctive feature of TNBC is a significantly lower rate of bone metastasis (10,11).

The risk of brain metastasis in TNBC is up to 3.5 times higher than in luminal tumors. A total of ~3.5 to 4.7% of TNBC patients will develop brain metastases as the first site of recurrence, compared with 1.3% in patients with non-TNBC. Moreover, TNBC has the shortest time interval between diagnosis of early stage disease and brain metastasis (12). Certain clinical characteristics, including young age, lymph node disease, large or high grade tumors and multiple visceral metastases, have been associated with a higher incidence of brain metastasis (13,14). The cerebellum and basal ganglia are the most common brain metastasis locations, representing 33% of cases. This may be due to the high blood flow in these areas (13). This is relevant as patients with symptomatic brain metastases present a worse prognosis, with a median survival of 3 to 9 months after diagnosis (15).

TNBC presents a higher risk of leading to death whenever a recurrence takes place (6). A distinctive pattern of recurrence has been identified among TNBCs: in the first two years after diagnosis, there is a rapid increase in the rate of recurrence with a peak at 3 years, followed by a rapid decrease in the following 5 years, and a very low risk of subsequent recurrences (16).

All the aforementioned factors contribute to the fact that TNBC has a clearly lower overall survival (OS) and cancer-specific survival, as well as a worse prognosis, with an increase in mortality after 2 years of diagnosis, compared with other subtypes of breast cancer (6-9). Furthermore, treatment of TNBC is often complex. The lack of tumor markers to direct treatment, along with an increased resistance to conventional treatments, make TNBC a challenge for clinicians (8). There is consequently a growing need for the identification and development of risk profiles to guide management.

TNBC tumors have distinct patterns of genetic expression. A previous study showed that levels of COX1, COX2, ALOX5 and ALOX5AP expression were high in TNBC, but low in other subtypes. It should be noted that this report also showed that there is overlap in gene expression across breast cancer subtypes. For instance, CYP19A1, which encodes for aromatase, is expressed in all subtypes. However, its expression is correlated to different genes (17).

A study published in 2011 by Lehman *et al* (18) identified 6 subtypes of gene expression in TNBC: basal-like-1 (BSL1), basal-like-2 (BSL2), immunomodulatory (IM), mesenchymal (M), mesenchymal stem-like (MSL) and luminal androgen receptor (LAR).

The BSL subtypes (BSL1 and BSL2) represent 47% of TNBC cases and express high levels of cell cycle genes. In the BSL1 subtype there are high levels of DNA damage response gene expression, while the BSL2 subtype displays unique gene ontologies involving growth factor signaling. The IM subtype is enriched for gene ontologies involved in immune cell processes. M and MSL subtypes both express genes involved in cell motility and cell differentiation. However, the MSL subtype additionally expresses genes linked to growth factor signaling pathways. The LAR subtype expresses genes involved in hormonally regulated pathways, such as those related to the androgen receptor (AR) (18).

Regarding the clinicopathological characteristics, the Lehman *et al* (18) study suggested that tumor size and histological type did not differ significantly between TNBC subtypes, while the age at diagnosis was higher in the LAR subtype. In terms of prognosis, the study showed that recurrence-free survival (RFS) differed significantly between subtypes. RFS was lower in the M subtype compared with BSL1 and IM. In another study by Masuda *et al* (19), BSL1 had the highest pathologic complete response (pCR) rate, while BSL2 and LAR had the lowest.

Hispanics are more likely to develop TNBC compared with non-Hispanic whites (20). However, little is known about the role of TNBC molecular subtypes in this population. Hispanics have often been underrepresented in studies defining TNBC subtypes and their clinical, pathological and prognostic characteristics (21).

In the present study, it was aimed to identify differences in prognosis across TNBC molecular subtypes in a Mexican based cohort. Additionally, it was aimed to describe the clinical and pathological characteristics and identify differences in treatment response and incidence of metastasis per TNBC molecular subtype in this population. Finally, the present study examined the behavior of metastatic [central nervous system (CNS) and visceral non-CNS] and non-metastatic TNBC.

Materials and methods

The present study retrospective cohort study was approved (approval no. CLAVE SALUD-2013-01-201336) by the bioethics and scientific committee of Mexico's National Institute of Cancer (INCan; Mexico City, Mexico) and conducted at the aforementioned institute. Written consent was obtained from all included patients before being included in the study.

Patients. Female patients (n=55) with a histopathological diagnosis of TNBC and a viable tissue sample were included. TNBC diagnosis was defined as having an immunohistochemical report (IHC) indicating estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER-2-negative in the initially performed Tru-Cut biopsy at INCan (Mexico City, Mexico). A viable tissue sample was defined as the availability of a formalin-fixed paraffin-embedded breast tumor specimen with 70% or higher neoplastic cellularity and 200 ng of RNA.

For hematoxylin-eosin staining, Tissue-Tek® Glas™ g2 Glass Coverslipper was used. The wax was first dissolved with xylene, which was later removed by passing the slide through ethanol and thoroughly rinsing with water. The slide was first stained for 3 min at room temperature with Harris hematoxylin and was 'blued' by treatment with a weakly alkaline solution. The section was later stained for 30 sec at room temperature with eosin and was lastly rinsed with alcohol and xylene.

Samples were assessed by IHC according to the 2020 ASCO/CAP guidelines. The antibodies used were ER (clone SP1; cat. no. 760-4324), PR (clone 1E2; cat. no. 790-4296), HER2 (4B5; cat. no. 760-4324; all from Ventana Medical Systems, Inc.) and Ki-67 (clone SP6; cat. no. CRM 325B

Biocare Medical, LLC;). An independent batch of tumor tissues, processed in the same manner as our samples, was used for determining staining specificity. Serial titrations were performed in order to obtain optimal concentrations for every antibody: anti-CK14 (1:500; clone SP53; cat. no. 760-4805), anti-CK17 (1:150; clone EP98; cat. no. 317R-16; both from Cell Marque; MilliporeSigma), anti-AR (1:80; clone 441; cat. no. Mob245; Diagnostic BioSystems, Inc.), anti-p63 (1:100; clone cm163c; cat. no. CM 163C; Biocare Medical, LLC) and anti-CK5/6 (1:500; clone 16B4; cat. no. GA780; DAKO; Agilent Technologies, Inc.). Additionally, antigen retrieval using Tris-EDTA or citrate was performed.

For chromogenic immunodetection, DAKO Envision systems or MACH 1 Universal HRP Polymer and diaminobenzidine were used. Afterwards, samples were counterstained for 3 min at room temperature with hematoxylin. All samples were reviewed with an Olympus BX53 microscope (phase contrast and fluorescence) at low magnification by a breast cancer pathologist who was blinded to patient characteristics and outcomes, and the positivity/negativity was determined following the College of American Pathologists guidelines (22): Any given biomarker was reported as positive when its expression was ≥1% in neoplastic cells even at low intensity in either cytoplasm, nucleus, or cytoplasmic membrane. On the contrary, when the expression was <1%, the biomarker was reported as negative.

All included patients were first treated at Mexico's National Institute of Cancer between 2007 and 2011. Follow-up for each patient began at the date when treatment was first administered, and continued until: i) Loss of follow-up, ii) Death or iii) Last visit before our cut-off date (June 2019). Patient data was obtained from electronic medical records. Collected information included administered drugs, treatment response, presence of metastatic disease and/or recurrence, as well as other clinical and pathological characteristics.

Gene expression profile analysis. Ultrasound-guided Tru-Cut needle biopsy samples were obtained, formalin fixed for 1 h at 56˚C and paraffin embedded. These biopsies were later subjected to microarray analysis using the Human Gene ST 2.0 microarray platform (Affymetrix; Thermo Fisher Scientific, Inc.). It should be noted that the corresponding microarray data was used in a previous study (23) and is publicly available (accession information is available in the 'Availability of data and materials' section). Microarray data was analyzed with the R software tool (24) and Bioconductor libraries (25). Based on the gene expression results, each patient was classified into one of the 6 TNBC molecular subtypes reported by Lehman *et al* in 2011 (18).

For technical quality control, results were processed before performing the differential expression analysis. This so-called low-level processing is performed at the probe level and corresponds to background correction using Robust Multiarray Average. This eliminates nonspecific hybridization, normalizes to remove systematic variations, and allows fluorescence intensity signals to be comparable with each other (Quantile Normalization method) (26,27).

RNA was extracted with TRIzol (Ambion; Thermo Fisher Scientific, Inc.) from paraffin-embedded samples. Later, cDNA was obtained using the AffymetrixTM SensationPlus kit following the manufacturer's protocol. This technique allows for a high and good quality of pure cDNA for microarray analysis.

In order to increase the number of samples used for subgroup identification, previously processed and classified (by PAM50) Mexican women breast cancer samples (106 tumor samples and 35 controls) from the National Institute of Genomic Medicine (INMEGEN) in Mexico City, were utilized in addition to our collected samples (66 samples). These samples were assessed in a previously published work and their data are publicly available (in the dbGaP repository, accession no. phs000369.v1.p1) (28). In order to combine all samples, data was normalized using the batch effect adjustment method with the ComBat algorithm (29) implemented in the Bioconductor library.

Statistical analysis. Chi-squared tests or Fisher's exact tests were used to compare the distribution of categorical variables (patient characteristics and treatment-related characteristics) between groups. For continuous variables, differences were analyzed using unpaired Student's t-test. These tests were performed as two-tailed, and $P<0.05$ was considered to indicate a statistically significant difference. Survival curves were generated using the Kaplan-Meier method, and differences between groups were analyzed with the log-Rank test. All analyses were performed using SPSS version 20.0 (IBM Corp.).

Results

Our cohort initially comprised 80 female patients with a diagnosis of TNBC, of whom, 66 had a viable paraffin-embedded sample (as aformentioned). Representative images of the pathological TNBC specimens used in the present study are shown in Fig. 1.

After molecular testing, 11 samples were reported as 'unspecified' molecular subtype. Therefore, the final cohort was made up of 55 patients. Based on the gene expression results, each patient was classified into one of the TNBC molecular subtypes reported by Lehman *et al* (18) in 2011, obtaining the following results: 16 patients with the IM subtype, 12 of M subtype, 11 of BSL2 subtype, 9 of BSL1 subtype, 6 of LAR subtype and 1 of MSL (Table I). Mean patient follow-up time was 47.1 months (range, 3-137 months).

Gene expression profile analysis. A heat map demonstrating hierarchical unsupervised clustering is observed in Fig. 2. Centroids, which were defined by Parker *et al* (30) using Caucasian women samples, were used. Therefore, slight differences may be encountered compared with Hispanic women. In the dendrogram, it is revealed that samples are clustered by their molecular profile; very few samples are clustered in intermediate positions.

All samples were analyzed using the TNBCtype web-based tool (developed by Vanderbilt University). This algorithm is based on 3,247 gene expression profiles from 21 breast cancer data sets, from which the 6 aforementioned TNBC subtypes were discovered. The 55 samples which were classified into TNBC subtypes (9 BL1, 11 BL2, 16 IM, 6 LAR, 12 M and 1 ML) are revealed in Fig. 3.

Table I. Triple-negative breast cancer molecular subtypes prevalence in a 55-patient cohort at Mexico's National Institute of Cancer between 2007 and 2011.

Molecular subtype	Number of patients (%)
Basal-like 1	9 (16.4)
Basal-like 2	11 (20)
Immunomodulatory	16 (29.1)
Mesenchymal	12 (21.8)
Androgen-like receptor	6 (10.9)
Mesenchymal Stem-Like	1 (1.8)

Table II. Triple-negative breast cancer molecular subtypes median OS.

Molecular subtype	Median OS, months
Immunomodulatory	NA
Mesenchymal Stem-Like	29
Basal-like 1	38
Androgen-like receptor	27
Mesenchymal	18
Basal-like 2	15

OS, overall survival; NA, median overall survival was not reached at the cut-off date.

OS and survival analysis by TNBC molecular subtype. As aforementioned, the cut-off date for follow-up was June 2019. Up to that moment, only 14 of the 55 patients (25%) were reported alive. The entire cohort's median OS was 29 months. OS by molecular subtype is presented in Table II. The best OS was observed in the IM subtype, as median OS was not reached at the cut-off date. The worst median OS was observed in the BSL2 subtype (15 months). Differences in OS between molecular subtypes did not reach statistical significance (P=0.064).

Most patients reported alive at the cut-off date (64.3%) had the IM subtype. In the survival analysis, the difference in survival between the IM subtype and other molecular subtypes met statistical significance (P=0.034; Fig. 4).

OS according to metastatic disease status. At the cutoff date, 14 patients were alive, all without metastatic disease. All patients who had metastasis succumbed. These patients were stratified according to their metastatic disease status (CNS, visceral non-CNS, no metastasis), and a survival analysis was performed (Fig. 5). It was observed that the group of patients with CNS metastasis had a lower OS than patients with visceral non-CNS metastatic activity: 12 months vs. 27 months. Among patients who did not develop metastatic disease, the median OS was not reached. More than half of the patients from the latter group (56.25%) belonged to the IM molecular subtype.

Clinicopathological characteristics of TNBC molecular subtypes. The main analyzed clinicopathological characteristics per molecular subtype are summarized in Table III. It should be noted that several of the patient characteristics initially intended to be analyzed were not recorded in the medical notes and therefore could not be examined. None of the assessed characteristics met statistical significance. As for BMI, most molecular subtypes had a majority of patients that were either obese or overweight, as opposed to the BSL2 molecular subtype, where 63% had a low to normal BMI. However, this difference did not meet statistical significance (P=0.498). Regarding the use of hormones, BSL1 was the molecular subtype where a higher proportion of patients was exposed (P=0.162). In the present study, 66.7% of the patients reported to have used hormone-containing drugs, mainly in the form of oral contraceptives.

More than half of the patients were premenopausal (56.3%), and this proportion was increased in the IM (75%) and BSL2 (72.7%) subtypes. Although a trend was identified, statistical significance was not met (P=0.125). Most patients had a parity of 3 or greater. However, 63.6% of the BSL2 subtype patients presented a parity of 2 or less (P=0.681). Breastfeeding was not assessable, since more than half of patient files did not report it.

In addition to clinical characteristics, pathological variables and their association with TNBC molecular subtypes were also analyzed. Histological type (luminal or ductal), histological grade, lymphovascular infiltration, CK 5/6, CK 17, CK14 and p63 were not associated with any molecular subtype. Ki-67 was not significantly associated with any subtype (P=0.27). However, it was lower in the M subtype; 75% of patients presented Ki-67 <20. Only 5 patients were reported as positive for the androgen receptor, but positivity for this receptor was not assessed for most patients. BRCA status was non-assessable since its status was only reported for 3 patients (one each for the BSL1, BSL2 and IM subtypes).

Response to treatment by TNBC molecular subtype. Out of the 55 patients, only 36 (65.5%) received neoadjuvant chemotherapy (CT), with the platinum-based scheme being the one most frequently administered (58.3% of CTs). This was followed by the anthracycline- and taxane-based scheme (38.2%). A total of 12 patients undergoing neoadjuvant CT (33.3%) had a pCR, of whom, 9 (75%) were identified as IM subtype. This association was statistically significant (P=0.011). A total of 8 of the patients who had a pCR (66.6%) belong to the group of patients who were reported to have no metastatic activity and who are currently alive. This association was also statistically significant (P=0.032). There was no statistically significant association between achieving a pCR and the type of scheme used, nor between prognosis and type of scheme: out of the 14 patients who were reported to still be alive, 8 received a platinum scheme, 2 did not receive neoadjuvant and 4 received a scheme with anthracyclines and taxanes.

Development of metastatic disease by TNBC molecular subtype. Patients who had CNS metastases were mainly of the BSL1 and Ml subtypes, 30.4 and 21.7%, respectively. Patients who had non-CNS visceral metastasis were mainly of the BSL2 and M subtypes, 37.5 and 31.2%, respectively. Finally, patients who did not have metastases belonged mainly to the

Table III. Triple-negative breast cancer molecular subtypes main clinical and pathological characteristics.

	BSL1	BSL2	IM	M	LAR	MSL	P-value
	n (%)						
BMI							0.498
<25	2 (22.2)	7 (63.6)	6 (37.5)	6 (50)	1 (16.7)	0	
25-29.9	4 (44.4)	2 (18.2)	6 (37.5)	2 (16.7)	3 (50)	0	
>30	3 (33.3)	2 (18.2)	4 (25)	4 (33.3)	2 (33.3)	1 (100)	
Hormone use							0.194
Yes	6 (66.7)	1 (9.1)	6 (37.5)	5 (41.7)	2 (33.3)	0	
No	2 (22.2)	9 (81.8)	8 (50)	7 (58.3)	2 (33.3)	1 (100)	
Unknown	1 (11.1)	1 (9.1)	2 (12.5)	0	2 (33.3)	0	
Parity							0.681
<2	2 (22.2)	7 (63.6)	4 (25)	4 (33.3)	3 (50)	0	
>2	7 (77.8)	4 (36.4)	12 (75)	8 (66.6)	3 (50)	1 (100)	
Hormonal status							0.125
Postmenopausal	4 (44.4)	3 (27.3)	4 (25)	8 (66.7)	4 (66.7)	1 (100)	
Premenopausal	5 (55.6)	8 (72.7)	12 (75)	4 (33.3)	2 (33.3)	0	
Ki 67 (%)							0.275
0	4 (44.4)	6 (54.5)	5 (31.2)	7 (58.3)	3 (50)	0	
10-19	0	0	0	2 (16.7)	0	0	
>20	5 (55.6)	5 (45.5)	11 (68.8)	3 (25)	3 (50)	1 (100)	
Androgen Receptor							0.076
Negative	8 (88.9)	6 (54.5)	6 (37.5)	7 (58.3)	3 (50)	0	
Positive	1 (11.1)	1 (9.1)	0	1 (8.3)	2 (33.3)	0	
Unknown	0	4 (36.4)	10 (62.5)	4 (33.3)	1 (16.7)	1 (100)	
CK 17							0.171
Negative	8 (88.9)	6 (54.5)	4 (54.5)	6 (50)	4 (66.7)	0	
Positive	1 (11.1)	1 (9.1)	2 (12.5)	2 (16.7)	1 (16.7)	0	
Unknown	0	4 (36.4)	10 (62.5)	4 (33.3)	1 (16.7)	1 (100)	
CK 14							0.194
Negative	7 (77.8)	5 (45.5)	4 (54.5)	5 (41.7)	5 (41.7)	0	
Positive	2 (22.2)	2 (18.2)	2 (12.5)	3 (25)	1 (16.7)	0	
Unknown	0	4 (36.4)	10 (62.5)	4 (33.3)	1 (16.7)	1 (100)	
p63							0.127
Negative	9 (100)	6 (54.5)	6 (37.5)	5 (41.7)	4 (66.7)	0	
Positive	0	0	0	1 (8.3)	0	0	
Unknown	0	5 (45.5)	10 (62.5)	6 (50)	2 (33.3)	1 (100)	
pCR							0.011
Yes	1 (11.1)	1 (9.1)	9 (56.2)	1 (8.3)	0	0	
Neoadjuvant CT scheme							0.092
Anthracycline and taxane	3 (33.3)	6 (54.5)	3 (18.8)	8 (66.7)	2 (33.3)	0	
Platinum	6 (66.7)	4 (36.4)	9 (56.2)	3 (25)	1 (16.7)	0	
Platinum and taxane	0	0	1 (6.2)	0	0	0	
No neoadjuvant CT	0	1 (9.1)	3 (18.8)	1 (8.3)	3 (50)	1 (100)	
Status							0.034
Alive	2 (22.2)	1 (9.1)	9 (56.2)	1 (8.3)	1 (16.7)	0	
Dead	7 (77.8)	10 (90.9)	7 (43.8)	11 (91.7)	5 (83.3)	1 (100)	

BSL1, Basal-like 1; BSL2, Basal-like 2; IM, immunomodulatory; M, Mesenchymal; LAR, Luminal androgen receptor; MSL, Mesenchymal stem-like; BMI, Body Mass Index; CT, Chemotherapy; CK, cytokeratin; pCR, pathologic complete response.

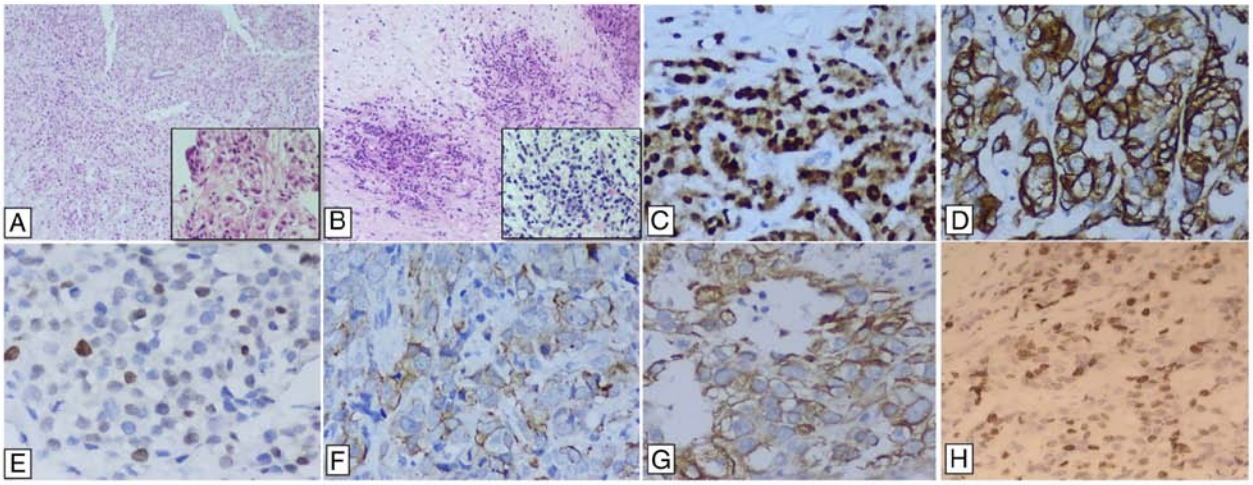

Figure 1. Representative images of immunohistochemical assay by biomarker. (A) H&E staining, which shows a ductal carcinoma (magnification, x10) and tumor detail (square) (magnification, x100). (B) H&E staining of lobular carcinoma (magnification, x100) and tumor detail (square) (magnification, x100). (C) Immunoperoxidase staining of androgen receptor with nuclear expression in 100% of neoplastic cells with high intensity (magnification, x400). (D) Immunoperoxidase staining of CK of low molecular weight (CK14) with cytoplasm membrane staining and high intensity (magnification, x400). (E) p63 nuclear staining (magnification, x400). (F) CK17 staining (magnification, x400). (G) Immunoperoxidase staining of CK 5/6, with cytoplasm membrane staining 60% of neoplastic cells with high intensity (magnification, x400). (H) Immunoperoxidase staining of ki-67, nuclear expression in the 30% of neoplastic cells with high intensity (magnification, x400). CK, cytokeratin.

IM molecular subtype (56%). Although this association was not statistically significant, it showed a trend (P=0.057).

Discussion

Breast cancer is the neoplasm with highest incidence among women in Mexico. Hormone receptors and/or HER2 overexpression in the surface of the tumor's cells allow for breast cancer characterization. In turn, breast cancer classification per receptor expression implies different behaviors and prognoses, which now has been exploited through therapeutic targets.

Globally, between 13 and 20% of all cases of breast cancer correspond to triple negative tumors, that is, tumors without expression of either hormone or HER2 receptors. Through Mexico's National Institute of Cancer reports, it has been identified that the incidence of TNBC in Mexico is ~23% (5).

TNBC is well known for being an aggressive disease with a greater capacity to develop metastasis. This, and particularly CNS metastases, deteriorates prognosis of patients. Unlike hormone-sensitive or HER2 overexpressing breast tumors, in TNBC there is still no biomarker that enables the use of targeted therapy, which would improve prognosis. For this reason, CT continues to be the mainstay of treatment.

In an attempt to gain improved understanding of the behavior of TNBC, throughout the years, gene expression in these tumors has been studied. Based on this, different gene expression profiles have been identified as possible prognosticators for the variability among different TNBC tumors. This way, TNBC could be regarded as a set of diseases, which can have different behaviors, prognoses and treatment sensitivities. One of the best described and most complete models is the one proposed by Lehman *et al* in 2011 (18), in which 6 different molecular subtypes were described, each one with a specific behavior, prognosis and treatment response profile. Since this is a rather complete and practical model, in the present study, our patients were classified according to it.

The present study builds upon a previous assessment of TNBC patients being treated between 2007 and 2011 at Mexico's National Institute of Cancer. That study aimed at identifying the role of gene expression profiles in the CNS metastasis process. The current study, by contrast, has a clinical approach. Herein, the information regarding the gene expression profiles, which was obtained previously, was used to find an association with clinical characteristics. This focus places the present project within the realm of translational medicine.

The intention of the current study was to identify the differences in survival across TNBC molecular subtypes. Additionally, it aimed to define clinical and pathological characteristics to help characterize them. This could eventually lead to an increasing effectiveness of treatments by aiding in the development of personalized treatment guidelines.

In the present study, differences in OS across TNBC molecular subtypes were identified. The best OS was reported in the IM subtype (median OS not reached), followed by BSL1 (38 months). Furthermore, until June 2019, 64.3% of patients reported to be alive belonged to this molecular subtype. By contrast, the worst OS was reported in patients identified with the BSL2 subtype (15 months), followed by the M subtype (18 months) (P=0.64). Up to the cut-off date, none of the patients reported as alive (14 of the 55) had metastatic disease and/or recurrence. One of these patients, however, had a double metachronous primary in the contralateral breast, which was successfully treated.

In total, 36 patients received neoadjuvant CT. A third of them (12 patients) had a pCR. Remarkably, the IM molecular subtype represented 75% of these patients (P=0.032). A total of 8 of these 12 patients belonged to the group of patients who did not develop metastasis.

Our analyses showed a tendency for BSL2 and M molecular subtypes to present with a higher rate of metastatic disease. The BSL2 subtype tended to present with CNS metastasis

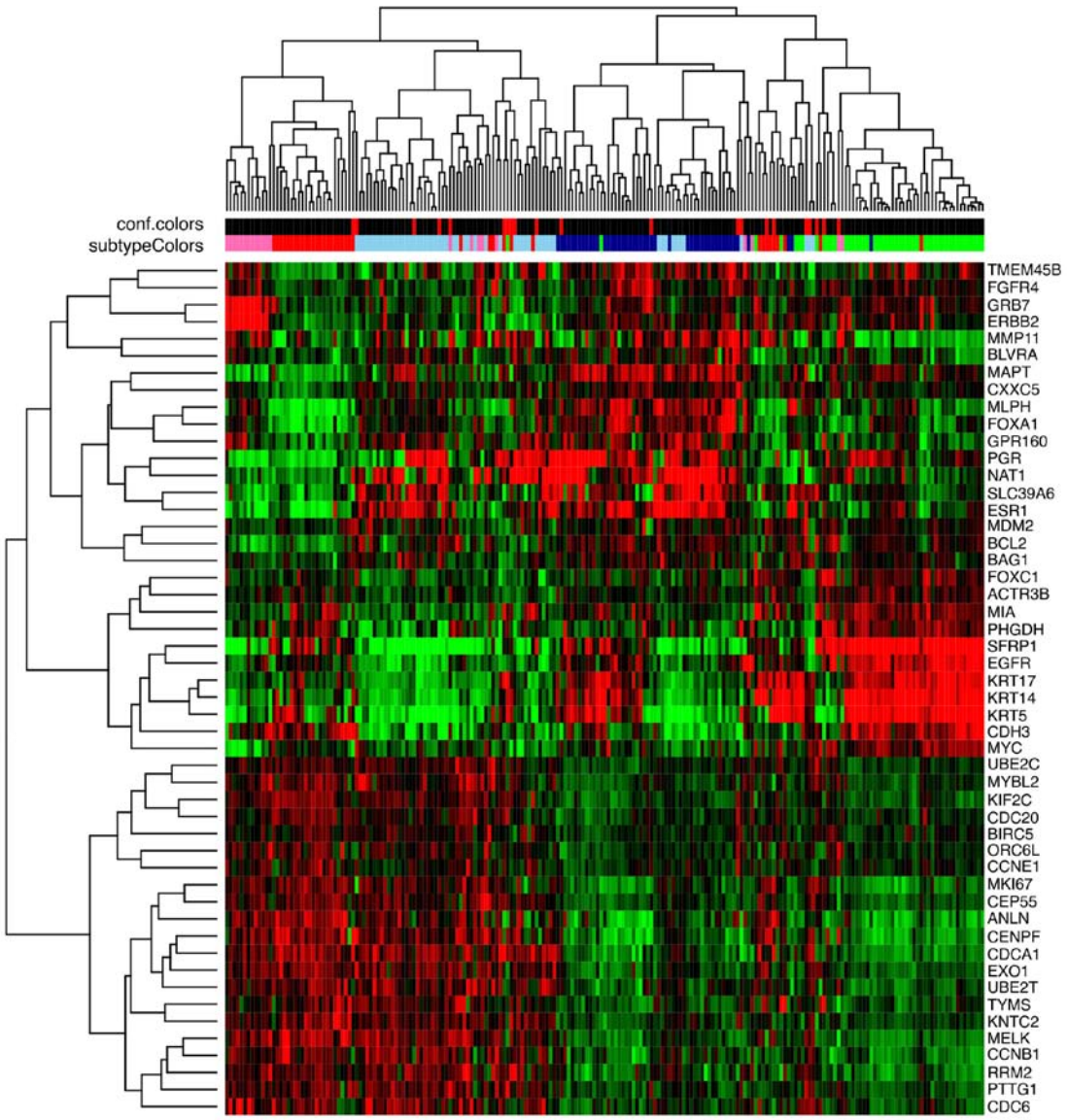

Figure 2. Hierarchical Unsupervised Cluster Heatmap. The 207 samples used are represented in the X axis and the 50 genes used are shown in the Y axis. Samples identified as luminal A are represented in dark blue, luminal B in light blue, HER2 positive in pink, Basal or triple-negative breast cancer in red and normal-like in green. Unspecified molecular subtype samples are represented with the color of the subtype they are more similar to.

more often, and the M subtype with non-CNS visceral metastatic disease. On the contrary, the IM molecular subtype had a lower probability of presenting metastasis.

Previous studies have found BSL2 and M subtypes to have the worst prognosis (18,21). These two subtypes were reported to have the worst OS and distant metastasis-free survival (19). This is in accordance with the present findings. On the other hand, current knowledge suggests that either BSL1 or IM have the best prognosis, with studies favoring one or the other (18,30). In our population, IM consistently showed the best prognosis, followed by BSL1.

The IM subtype is highly enriched for genes involved in the immune cell signaling process (18). Recently, this subtype was found to present with a higher proportion of intratumoral infiltrating lymphocytes (23). This is relevant as different studies have demonstrated that tumors with high lymphocytic infiltrate have an improved prognosis. The presence of tumor-infiltrating lymphocytes is associated with an improved OS and decreased metastasis incidence (32). These phenomena could be a source of rationalization for our findings.

A total of 31 variables (clinical and pathological) were assessed in an attempt to clinically characterize the TNBC molecular subtypes in our population. Unfortunately, electronic medical records were highly heterogeneous and often did not report these variables. Hence, the elucidation of the relationship between certain of these variables and the TNBC molecular subtypes could not be accomplished.

Mean age in our cohort was 49 years, which is in accordance with studies of TNBC in the Hispanic population (5). It should be noted that the age range was wide (30-80 years old). There was no significant relation between age and either development of metastatic disease or molecular subtype.

Having a high body mass index (BMI) has been linked to the development of this disease. In this cohort, 60% of patients were either overweight or obese. However, there was

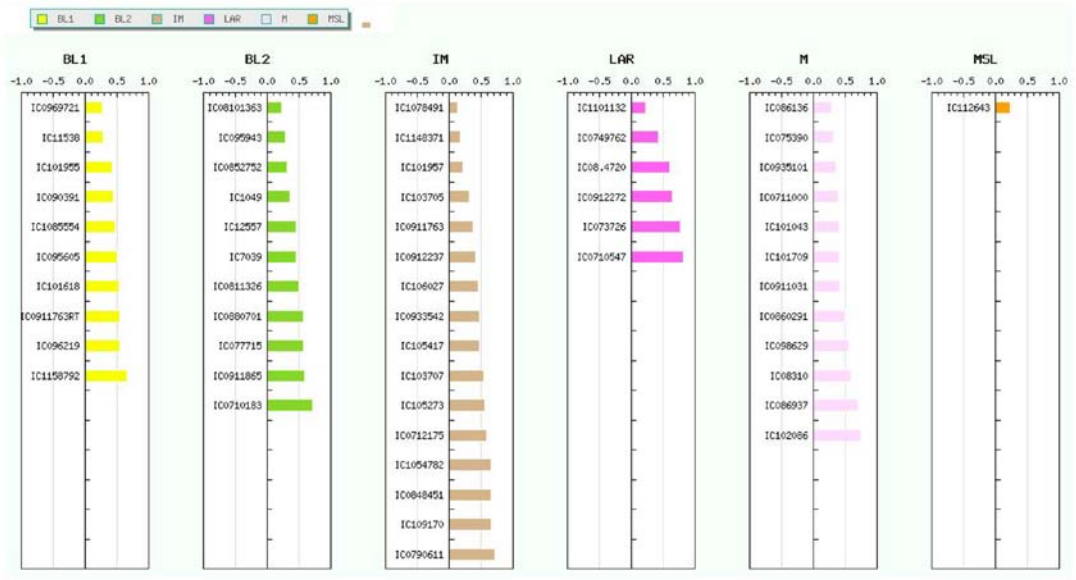

Figure 3. TNBC subtype classification of 55 samples using the TNBCType tool. TNBC, triple-negative breast cancer. BL1, Basal-like 1; BL2, Basal-like 2; IM, Immunomodulatory; LAR, Androgen-like receptor; M, Mesenchymal; MSL, Mesenchymal Stem-Like.

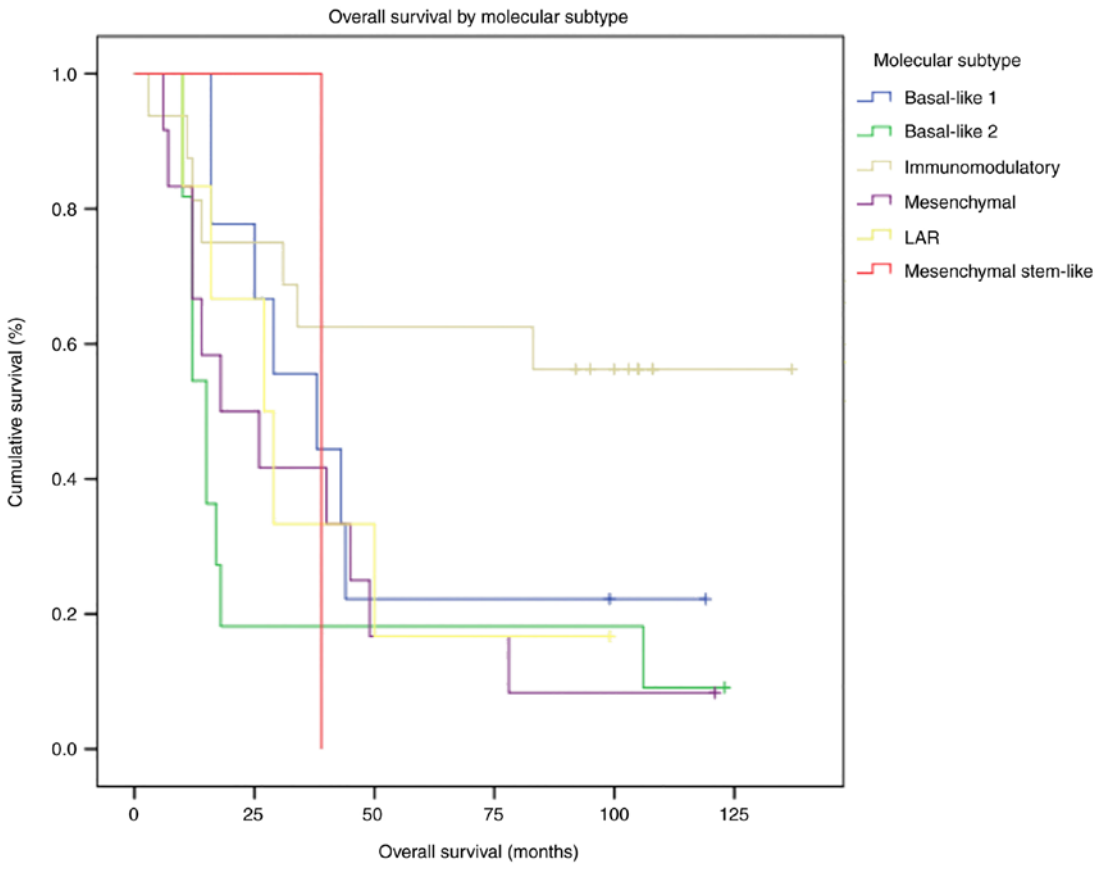

Figure 4. Kaplan-Meier survival curve for overall survival by triple-negative breast cancer molecular subtype. The difference in survival between the IM subtype and other molecular subtypes met statistical significance (P=0.034).

no significant association between BMI and the development of metastasis, or BMI and molecular subtype.

Hormone exposure (either exogenous or endogenous) does not seem to have a relationship with the development of TNBC over other subtypes of breast cancer. In the present study, the age of menarche and the use of hormonal drugs were not associated with metastasis or a molecular subtype. Hormonal status showed a slight tendency towards an association with the development of metastasis (P=0.149). Most patients with CNS metastasis (56.5%) were postmenopausal, while most patients

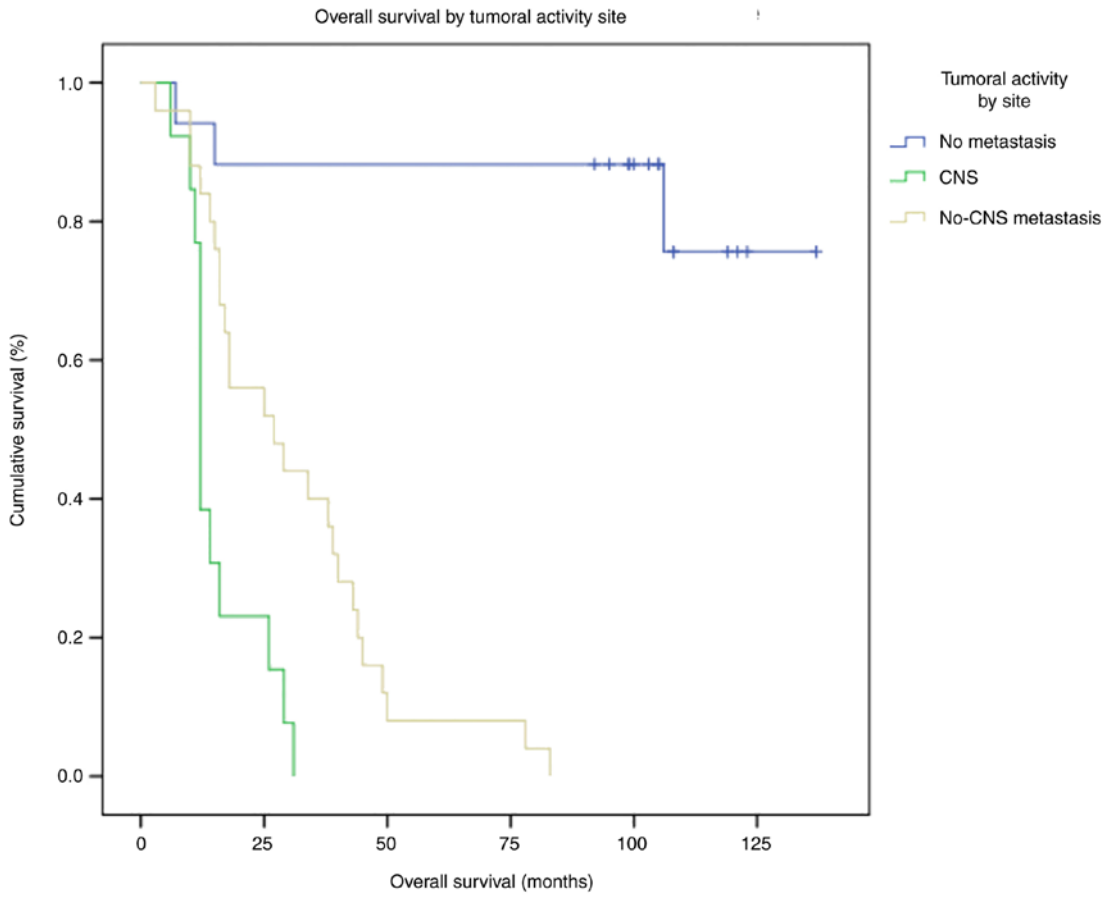

Figure 5. Kaplan-Meier survival curve for overall survival by metastasis status in triple-negative breast cancer patients. CNS, central nervous system.

with non-CNS visceral metastasis (56%) and without metastasis (75%) were premenopausal. Furthermore, premenopausal patients consisted mainly of patients who had the BSL2 or IM molecular subtype (25 and 38% of this group, respectively) (P=0.125).

Parity was not significantly associated with any molecular subtype. Neither was the age of the first and last deliveries, or breastfeeding. The BSL2 molecular subtype presented the lowest parity (63.6% had 2 or less deliveries).

As part of the study, a possible relationship between 10 histopathological characteristics and the development of metastasis and molecular subtype was also investigated. However, this analysis was limited due to lack of reporting of these variables in the medical records.

The development of high grade tumors, with a high completion rate (determined by Ki-67, mostly >20%) are well-known features of TNBC. Accordingly, in the present study these same features were observed, exempting the M molecular subtype, where 75% of the patients had tumors with a Ki 67 <20%.

Metastatic disease resulted in lower survival regardless of the molecular subtype. CNS metastasis in particular represented the worst prognosis. Of the 55 studied patients, 23 had metastatic CNS disease, 16 patients had non-CNS visceral metastatic disease and 16 patients did not develop metastatic disease.

Patients with CNS metastasis belonged mainly to the molecular subtypes BSL1 and M (30.4 and 21.7%, respectively). The non-CNS visceral metastasis group consisted mainly of patients with the BSL2 and M molecular subtypes (37.5 and 31.2%, respectively). Lastly, the group of patients without metastases was made up mainly of patients with the IM molecular subtype (56%). Even though statistical significance was not met, differences between groups showed a trend.

Patients with metastatic CNS disease had the worst median OS (12 months), followed by patients with non-CNS visceral metastatic disease (27 months). Even when only analyzing patients with metastatic disease, differences across TNBC molecular subtypes were observed. The best survival still occurred in patients with the IM subtype. Patients without metastatic disease had the best survival; up to the cut-off date, the survival median had not yet been reached.

Overall, these results clearly revealed that CNS metastasis is associated with the worst prognosis in TNBC patients, which is consistent with previous studies (33). Factors such as brain inflammation and edema have classically been proposed to lead to potentially deadly complications (brain herniation, neurological deficit and seizures) that may account, at least partially, to the worsening of prognosis (34). Certain breast cancer genes have been individually associated with CNS metastasis, and therefore, it is reasonable to consider that gene expression profiles may affect CNS metastasis incidence (35). Our findings suggested that patients with the BSL-1 subtype have a higher likelihood of developing CNS metastasis. Consequently, future research should evaluate BSL-1 genes as mechanistic factors leading to CNS metastasis. Moreover,

efforts to detect brain metastasis opportunely should be made in breast cancer patients with this subtype.

It was not possible to analyze the implications of BRCA mutations in TNBC molecular subtypes in the present study, as these mutations were seldom investigated and reported in medical records. It would be interesting to study elsewhere the implications of these mutations in TNBC subtypes in Hispanics.

One strength of the current study is the fact that a large number of clinical and pathological variables were collected and considered for each patient. No variable met statistical significance for its association with a molecular subtype. Another strength is that included patients have maintained adherence to the surveillance consultations throughout the study period. These patients were closely monitored by a multidisciplinary team in Mexico's National Institute of Cancer.

The main limitation of the present study was the heterogeneity and occasional lack of reporting in the electronic medical records. This limited the power to analyze all the desired variables. This problem was encountered since numerous of the collected variables were previously not considered a standard in patient evaluation at the Institute, hence they were often not collected and reported in medical records. For instance, the acquisition of negative surgical margins or the evidence of lymphatic spread at diagnosis, which could have been useful to compare groups who did and did not develop metastasis, were not evaluated. However, it should be mentioned that patients from stages II to III were included to ensure that the progression of disease was evaluated without being influenced by the time of diagnosis. Additionally, no clinical differences were identified between patients who did and did not develop metastasis, apart from the molecular subtype.

Another limitation to the current study was the diversity in administered treatments between patients. This hindered our ability to analyze results regarding individual patient response to treatment. It should also be mentioned that our microarray data were not validated using RT-qPCR, which would be ideal, due to lack of sampled tissue. Future studies assessing TNBC subtypes should aim to validate their microarray results using methods such as RT-qPCR in order to ensure accuracy. One final limitation is the relatively small gathered sample size in the present study. This may have limited the ability to elucidate associations and to discuss interesting but rare subtypes such as MSL (only 1 patient with MSL was included in our study).

Information obtained in the present study may help to incorporate a clinical-molecular model for TNBC patient care. Particularly, this information is valuable for decision making in Hispanic-based populations, where TNBC represents a larger proportion of cases and information is lacking. This way, treatment and monitoring can be individualized to the patient's risk. For instance, whenever managing a patient with the BSL2 or M molecular subtypes in populations similar to ours, a more rigorous diagnostic approach could be initiated to rule out metastatic disease. Subclinical disease could be diagnosed in an early manner, highly benefiting patients. The implementation of measures such as these, which personalize management, utilizing existing and forthcoming information, may ultimately contribute to a substantial improvement in prognosis and quality of life of patients with TNBC.

In conclusion, patients with the TNBC IM molecular subtype had a longer OS, with a median survival not reached during the span of our study. By contrast, the lowest OS was observed in patients with the BSL2 molecular subtype, followed by the M subtype. No significant associations were found between any of the clinical or pathological characteristics and TNBC molecular subtypes. Patients with the IM molecular subtype presented a pCR to neoadjuvant CT more often (75% of patients who had a pCR belonged to this subtype). Patients with the BSL2 and M molecular subtypes had a higher probability of developing CNS metastatic disease and visceral non-CNS metastatic disease, respectively. The IM molecular subtype had the lowest probability of developing metastatic disease. Patients with CNS metastatic disease had the lowest survival. The results obtained in the present study should be considered to implement a clinical-molecular model for TNBC patient care in Hispanic-based populations.

Acknowledgements

Not applicable.

Funding

The present study was supported by the Fondo SS/IMSS/ISSSTE-CONACYT (grant no. SALUD-2013-1-201336).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request. Microarray data are available at the Gene Expression Omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE176128).

Authors' contributions

AOG, OA, CVG, JAMS and EOV conceptualized and designed the study. EOV, CRE, RRB and CHCS performed data collection. RVR performed the statistical analysis. AOG, EOV, CRE, JAMS, RVR, AG, RRB and CHCS analyzed and interpreted data. EOV, AOG and AG drafted the manuscript. JAMS, AOG, EOV, CVG and OA reviewed and revised the manuscript. AOG, EOV, CHCS and JAMS confirm the authenticity of all the raw data. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

The present study was approved (approval no. CLAVE SALUD-2013-01-201336) by the local bioethics and scientific committee of Mexico's National Institute of Cancer (INCan, Mexico City, Mexico). Written consent was obtained from all included patients before being included in the study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA and Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68: 394-424, 2018.
2. Sharma P: Biology and management of patients with triple-negative breast cancer. Oncologist 21: 1050-1062, 2016.
3. Foulkes WD, Smith IE and Reis-Filho JS: Triple-negative breast cancer. N Engl J Med 363: 1938-1948, 2010.
4. Reynoso-Noverón N, Villarreal-Garza C, Soto-Perez-de-Celis E, Arce-Salinas C, Matus-Santos J, Ramírez-Ugalde MT, Alvarado-Miranda A, Cabrera-Galeana P, Meneses-García A, Lara-Medina F, *et al*: Clinical and epidemiological profile of breast cancer in mexico: Results of the seguro popular. J Glob Oncol 3: 757-764, 2017.
5. Lara-Medina F, Pérez-Sánchez V, Saavedra-Pérez D, Blake-Cerda M, Arce C, Motola-Kuba D, Villarreal-Garza C, González-Angulo AM, Bargalló E, Aguilar JL, *et al*: Triple-negative breast cancer in Hispanic patients. Cancer 117: 3658-3669, 2011.
6. de Ruijter TC, Veeck J, de Hoon JPJ, van Engeland M and Tjan-Heijnen VC: Characteristics of triple-negative breast cancer. J Cancer Res Clin Oncol 137: 183-192, 2011.
7. Kennecke H, Yerushalmi R, Woods R, Cheang MC, Voduc D, Speers CH, Nielsen TO and Gelmon K: Metastatic behavior of breast cancer subtypes. J Clin Oncol 28: 3271-3277, 2010.
8. Li X, Yang J, Peng L, Sahin AA, Huo L, Ward KC, O'Regan R, Torres MA and Meisel JL: Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. Breast Cancer Res Treat 161: 279-287, 2017.
9. Martin AM, Cagney DN, Catalano PJ, Warren LE, Bellon JR, Punglia RS, Claus EB, Lee EQ, Wen PY, Haas-Kogan DA, *et al*: Brain metastases in newly diagnosed breast cancer: A population-based study. JAMA Oncol 3: 1069-1077, 2017.
10. Xiao W, Zheng S, Yang A, Zhang X, Zou Y, Tang H and Xie X: Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: A population-based study. Cancer Manag Res Volume 10: 5329-5338, 2018.
11. Ignatov A, Eggemann H, Burger E and Ignatov T: Patterns of breast cancer relapse in accordance to biological subtype. J Cancer Res Clin Oncol 144: 1347-1355, 2018.
12. Jin J, Gao Y, Zhang J, Wang L, Wang B, Cao J, Shao Z and Wang Z: Incidence, pattern and prognosis of brain metastases in patients with metastatic triple negative breast cancer. BMC Cancer 18: 446, 2018.
13. Rostami R, Mittal S, Rostami P, Tavassoli F and Jabbari B: Brain metastasis in breast cancer: A comprehensive literature review. J Neurooncol 127: 407-414, 2016.
14. Kim YJ, Kim JS and Kim IA: Molecular subtype predicts incidence and prognosis of brain metastasis from breast cancer in SEER database. J Cancer Res Clin Oncol 144: 1803-1816, 2018.
15. Heitz F, Harter P, Lueck HJ, Fissler-Eckhoff A, Lorenz-Salehi F, Scheil-Bertram S, Traut A and du Bois A: Triple-negative and HER2-overexpressing breast cancers exhibit an elevated risk and an earlier occurrence of cerebral metastases. Eur J Cancer 45: 2792-2798, 2009.
16. Minami CA, Chung DU and Chang HR: Management options in triple-negative breast cancer. Breast Cancer (Auckl) 5: 175-199, 2011.
17. Kennedy BM and Harris RE: Cyclooxygenase and lipoxygenase gene expression in the inflammogenesis of breast cancer. Inflammopharmacology 26: 909-923, 2018.
18. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y and Pietenpol JA: Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest 121: 2750-2767, 2011.
19. Masuda H, Baggerly KA, Wang Y, Zhang Y, Gonzalez-Angulo AM, Meric-Bernstam F, Valero V, Lehmann BD, Pietenpol JA, Hortobagyi GN, *et al*: Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. Clin Cancer Res 19: 5533, 2013.
20. Hines LM, Risendal B, Byers T, Mengshol S, Lowery J and Singh M: Ethnic disparities in breast tumor phenotypic subtypes in hispanic and non-hispanic white women. J Womens Health 20: 1543-1550, 2011.
21. Yin L, Duan JJ, Bian XW and Yu S: Triple-negative breast cancer molecular subtyping and treatment progress. Breast Cancer Res 22: 61, 2020.
22. Fitzgibbons PL, Dillon DA, Alsabeh R, Berman MA, Hayes DF, Hicks DG, Hughes KS and Nofech-Mozes S: Template for reporting results of biomarker testing of specimens from patients with carcinoma of the breast. Arch Pathol Lab Med 138: 595-601, 2014.
23. Rodríguez-Bautista R, Caro-Sánchez CH, Cabrera-Galeana P, Alanis-Funes GJ, Gutierrez-Millán E, Ávila-Ríos S, Matías-Florentino M, Reyes-Terán G, Díaz-Chávez J, Villarreal-Garza C, *et al*: Immune Milieu and genomic alterations set the triple-negative breast cancer immunomodulatory subtype tumor behavior. Cancers (Basel) 13: 6256, 2021.
24. RStudio Team: RStudio: Integrated Development of R., 2021.
25. Marini F, Linke J and Binder H: Ideal: An R/Bioconductor package for Interactive differential expression analysis. BMC Bioinformatics 21: 565, 2020.
26. Bolstad BM, Irizarry RA, Åstrand M and Speed TP: A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19: 185-193, 2003.
27. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U and Speed TP: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 4: 249-264, 2003.
28. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, *et al*: Sequence analysis of mutations and translocations across breast cancer subtypes. Nature 486: 405-409, 2012.
29. Johnson WE, Li C and Rabinovic A: Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8: 118-127, 2007.
30. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, *et al*: Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol 27: 1160-1167, 2009.
31. Bareche Y, Venet D, Ignatiadis M, Aftimos P, Piccart M, Rothe F and Sotiriou C: Unravelling triple-negative breast cancer molecular heterogeneity using an integrative multiomic analysis. Ann Oncol 29: 895-902, 2018.
32. Hubalek M, Czech T and Müller H: Biological subtypes of triple-negative breast cancer. Breast Care 12: 8-14, 2017.
33. Brosnan EM and Anders CK: Understanding patterns of brain metastasis in breast cancer and designing rational therapeutic strategies. Ann Transl Med 6: 163-163, 2018.
34. Quattrocchi CC, Errante Y, Gaudino C, Mallio CA, Giona A, Santini D, Tonini G and Zobel BB: Spatial brain distribution of intra-axial metastatic lesions in breast and lung cancer patients. J Neurooncol 110: 79-87, 2012.
35. Lv Y, Ma X, Du Y and Feng J: Understanding patterns of brain metastasis in triple-negative breast cancer and exploring potential therapeutic targets. Onco Targets Ther 14: 589-607, 2021.