Clinical Significance of Liquid Biopsy through 5-miRNA Signature Profiling in Tumor-derived Extracellular Vesicles to Complement Breast Cancer Diagnosis

Young Kim

The Graduate School
Yonsei University
Department of Medicine

Clinical Significance of Liquid Biopsy through 5-miRNA Signature Profiling in Tumor-derived Extracellular Vesicles to Complement Breast Cancer Diagnosis

A Dissertation Submitted
to the Department of Medicine
and the Graduate School of Yonsei University
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Medical Science

Young Kim

December 2024

This certifies that the Dissertation of Young Kim is approved

Thesis Supervisor Seung Il Kim

Thesis Committee Member Min Jung Kim

Thesis Committee Member Kyoung Hoon Suh

Thesis Committee Member Ja Seung Koo

Thesis Committee Member Ho Kyoung Hwang

The Graduate School
Yonsei University
December 2024

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my heartfelt gratitude to my advisor and mentor, Dr. Seung Il Kim, for his unwavering support, insightful guidance, and boundless patience throughout my PhD journey. His dedication to fostering scientific curiosity and his exceptional mentorship have been instrumental in shaping my professional growth. Furthermore, Dr. Kim's enthusiasm for science and his optimistic spirit have inspired me to remain resilient and motivated, even during the most challenging times.

I am also deeply thankful to Dr. Kyoung Hoon Suh, the founder of E&S Healthcare Co., Ltd., for granting me the opportunity to concurrently pursue my degree and professional career, as well as for his steadfast encouragement and support. My special gratitude extends to my master's advisor, Dr. Kwang Soo Shin, for guiding me as I took my first steps in the world of science.

I would also like to express my appreciation to the esteemed members of my advisory board—Dr. Min Jung Kim, Dr. Ja Seung Koo, and Dr. Ho Kyung Hwang—for their valuable insights and constructive feedback throughout my research. My sincere thanks go to all the members of Dr. Seung Il Kim's and Dr. Jee Ye Kim's labs for their invaluable assistance and thoughtful discussions. In particular, I am grateful to Dr. Min Woo Kim for his technical advice and meaningful discussions, which were pivotal to the progress of this work.

Words cannot fully capture my gratitude to my beloved wife, Jung Hye Yang, for her unconditional love, encouragement, and unwavering belief in me. Your support has been my anchor, giving me the strength to navigate through difficult times and pursue my dreams with determination. Balancing the roles of a husband, employee, and student demanded more than twice the effort and time, and I often fell short in many ways. Yet, this accomplishment would not have been possible without your steadfast support.

This degree represents not just a personal milestone but also a source of nourishment for future growth and endeavors. To all those who stood by me, believed in me, and contributed to this achievement—thank you from the bottom of my heart.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Clinical Significance of Liquid Biopsy through 5-miRNA Signature Profiling in Tumor-derived Extracellular Vesicles to Complement Breast Cancer Diagnosis

Breast cancer is still a major worldwide health concern, which emphasizes the necessity that better diagnostic techniques are needed. While radiological diagnosis has increased early detection rates and decreased the risk of mortality, it is not without limitations, including increased false negatives in dense breasts and unnecessary biopsies due to false positives. Additionally, tissue biopsy may not fully capture the spatiotemporal heterogeneity of tumors. Hence, there's a continuous demand for complementary approaches like liquid biopsy methods. The objective of our research was to establish a liquid biopsy technique that effectively captures the specific traits of breast cancer, offering a valuable addition to current diagnostic tools and improving clinical outcomes. We isolated and analyzed breast cancer-derived extracellular vesicles, targeting EpCAM, ITGA2, and ITGAV to cover molecular subtypes, and utilized an immune-affinity method for isolation. Through miRNA analysis, we sought to establish an optimal diagnostic combination. We tested 211 patients, including breast cancer patients, benign breast disease patients, and normal controls, for a 5-miRNA signature (miR-21, miR-106b, miR-181a, miR-484, and miR-1260b). Our findings demonstrated a clinical sensitivity of 85.83%, a specificity of 84.62%, and an AUC of 0.908. Notably, we observed clinical sensitivity ranging from 82.35% to 91.67% across different TNM stages, highlighting effectiveness, particularly in early-stage breast cancer. When combined analysis of mammography and miRNA signatures yielded a clinical sensitivity of 97.39%. Moreover, we observed promising results in complementary analysis, with high true positive rates for dense breasts and reduced unnecessary biopsies for benign breast diseases. In conclusion, our study presents a novel liquid biopsy method that complements current breast cancer diagnosis. Our approach offers superior clinical performance compared to existing methods, especially in early-stage detection. Additionally, the combined analysis of mammography and miRNA signatures shows potential for improving sensitivity and specificity, thus enhancing diagnostic accuracy. With opportunities for more study and practical application, this multi-component diagnostic strategy using extracellular vesicles derived from breast cancer shows an opportunity as a liquid biopsy method for breast cancer diagnosis.

Key words: **Breast cancer, diagnosis, biomarker, extracellular vesicle, microRNA, liquid biopsy.**

1. INTRODUCTION

Breast cancer is the most prevalent form of cancer among women worldwide (1). According to GLOBOCAN 2020, approximately 2.2 million women were newly diagnosed with breast cancer, resulting in over 684,996 reported deaths (2). The incidence and mortality rates of breast cancer continue to rise annually (2). The World Health Organization (WHO) identifies mammograms as the principal screening technique to mitigate the risk of breast carcinoma (3). The goal of breast screening is to detect the disease at its earliest abnormal status in women who do not exhibit any signs or symptoms, enabling early treatment and reducing the risk of mortality and morbidity associated with the disease (3). However, the current radiological technology, particularly mammography, and the subsequent interpretation of imaging tests through pathological analysis are not flawless (4). Although randomized controlled trials have shown that screening for early detection of breast cancer has clinical benefits in reducing mortality rates, there are also potential risks involved (5). These risks include radiation exposure from X-ray-based imaging methods, unnecessary additional testing, over-diagnosis due to false positives, and the occurrence of interval tumors during the additional evaluation period (5–7). Dense breast tissue poses a particular challenge to early diagnosis as it decreases the sensitivity of mammography in detecting small lesions and may obscure the visualization of underlying cancer (8,9). Despite improvements in digital mammography technology, its clinical sensitivity remains modest, especially when imaging dense breast tissue, with a sensitivity rate of approximately 61.5% (10–12). Even when combining mammography with breast ultrasound, the sensitivity only reaches 81.5% (13–16). False results or re-examinations from screening mammography can lead to over-diagnosis, as there may be a discordance between radiological and pathological results (6,17). This discordance complicates subsequent procedures for clinical physicians and can cause delays in diagnosis, potentially increasing the risk of interval cancer occurrence (3). To give an example, for BI-RADS assessment categories above 4 from mammography, a core needle biopsy is recommended (4). However, these recommendations do not always lead to good practice. It is important to obtain concordance between the pathological result and the imaging findings when performing a needle biopsy. BI-RADS 4 is further classified into 4a, 4b, and 4c, indicating different probabilities of malignant tumors (6). It is worth noting that even with a BI-RADS 4a diagnosis through mammography, the majority of cases (90-98%) do not turn out to be breast cancer, resulting in unnecessary tissue biopsies (6). This can lead to psychological distress and impose time and economic burdens on the individuals involved (6). While radiologic diagnosis confirms the physical presence of a tumor, it has limitations in analyzing tumor characteristics. While artificial intelligence (AI) improves precision in radiologic analysis, it falls short in identifying the molecular uniqueness of malignancies, limiting the depth of information radiologic diagnosis can provide (18–21). Moreover, tissue samples acquired via solid or surgical biopsies represent just a fraction of the tumor, inadequately reflecting its spatiotemporal complexity

(20). Solid biopsies often yield unsuitable quality specimens due to procedural complications, contamination, and human error, introducing bias by collecting tumor samples from specific locations and reflecting only a part of the genetic heterogeneity (20,22,23). In contrast, circulating cancer biomarkers in the blood, present as a result of tumor circulation, can integrate the molecular biological information of the tumor and represent spatial heterogeneity (20,21). These biomarkers can be repeatedly sampled longitudinally during the diagnosis and treatment process, enabling the tracking of temporal heterogeneity and tumor evolution (20,21). According to the results of the recent PAN-CANCER study on circulating tumor DNA(ctDNA) methylation or ctDNA plus protein combination, it was confirmed that the diagnostic performance for breast cancer was not satisfactory (24–28). Even for malignant tumors, the attempt to introduce liquid biopsy through an overall target molecule's analysis that might do not reflect the unique characteristics of each malignant tumor showed a low confidence level in diagnostic accuracy. As such, there is a need to develop a liquid biopsy method that takes into account the unique characteristics of each tumor. Particularly, extracellular vesicles (EVs) are attractive candidates for application of liquid biopsy (29–33). They are secreted by living cells and contain various substances, such as DNA, RNA, proteins, and metabolites, reflecting the characteristics of their parental cell of origin (23,34). While circulating molecules from cancer, particularly nucleotides, may be exposed in the blood and susceptible to removal by restriction enzymes, those present in EVs are more valuable as they are protected from external factors (23). Since it is derived from specific cancer and is protected in the cargo, it may be possible to be distinguished with false signals related to them derived from other cells like clonal hematopoiesis (35,36). All living cells continually generate EVs, which provide several chances for clinically valuable diagnostic information given that they include DNA, RNA, and protein (20,21,29). Through necrosis and apoptosis, dying cancer cells release tiny amounts of "cell-free" DNA (cfDNA), with the majority being ctDNA (23). In order to overcome this unsolved issues regarding circulating biomarkers, it is advantageous to analyze the substances that EVs contain after identifying tumor-derived cell membrane proteins as surface indicators (33,37–39). Prior investigations into extracellular vesicle (EV) isolation and diagnostic application highlighted surface proteins related to tumor heterogeneity, like tetraspanins (CD9, CD81, CD63), which tend to be highly expressed in the EV population (40). Based on miRNA expression pattern research, tumor-derived EVs that include microRNAs (miRNAs) are particularly suggested as an outstanding source for cancer detection. (29,41,42). As primary tumors evolve and undergo cell death, the miRNA expression in cancer undergoes dynamic changes, particularly within EVs (41,42). Cancer cells often exhibit variations in immune evasion mechanisms and proliferation, leading to altered miRNA expression levels (33). These modifications encourage angiogenesis, invasion, proliferation of tumors, and immune evasion. These miRNA changes affect both the tumor's local environment and distant areas, actively promoting tumor growth, tissue invasion, angiogenesis, the formation of metastatic niches, and evasion of the immune system (29,30). Research has demonstrated that during distinct phases of growth and demise, primary tumors leak miRNAs into the bloodstream and the transcription trends in miRNAs within primary tumor tissue and corresponding plasma are comparable (43). Notably, miRNA profiles in tumor tissues and their corresponding EV samples often exhibit considerable similarities, underscoring the diagnostic potential of EV analysis (43). A

crucial stage in evaluating EV-derived miRNAs for liquid biopsy-based cancer diagnosis is the isolation and enrichment of EVs linked to malignancy in plasma (44). Ultracentrifugation (UC), the gold standard for EV isolation, lacks selectivity for EVs associated with cancer, is time-consuming, and has low isolation efficiency (45). Using an approach leading to employing immuno-affinity capture for certain cancer cell surface epitopes is essential in order to get beyond these restrictions (40). The current Total EV isolation method may yield a substantial quantity of EVs; however, it is anticipated that this method might not accurately represent tumor characteristics (37,39,40,46). This discrepancy arises from the mixing of EVs originating from various cells, alongside tumor-derived EVs (29,30,40). Therefore, there is an expectation that by implementing surface biomarkers that specifically reflect tumor characteristics and are overexpressed, an immunoaffinity-based isolation method could effectively isolate tumor-derived EVs (30,37,44). This approach holds promise for enabling more precise and tumor-specific analyses.

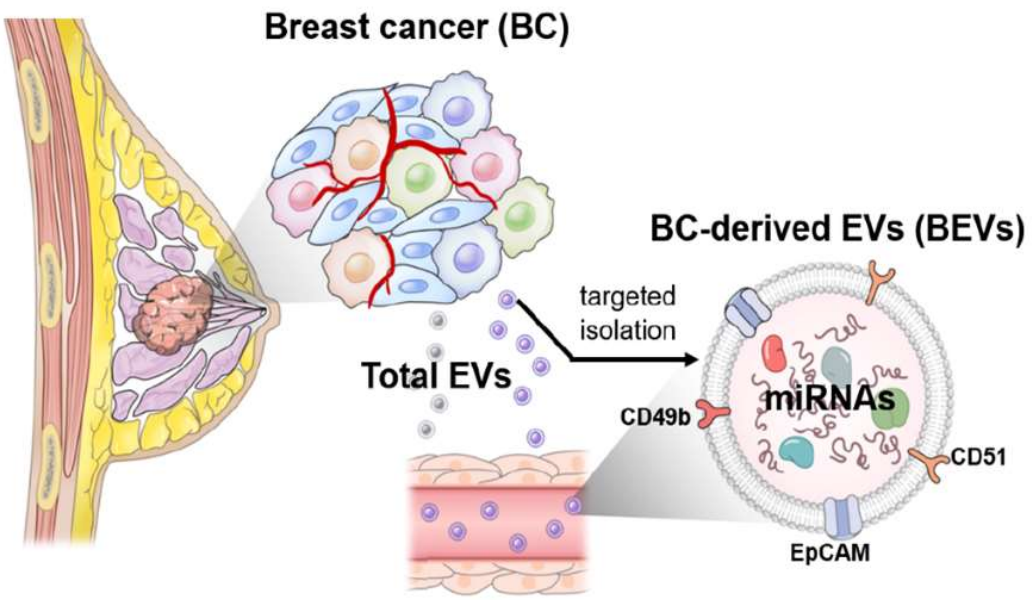

Figure 1. Diagram depicting the workflow for isolating extracellular vesicles from breast cancer (BEVs) and performing miRNA analysis. The schematic depicts the targeted isolation of BEVs from the full pool of extracellular vesicles in the blood, using proteins such as EpCAM (epithelial cell adhesion molecule), ITGAV(integrin αv, CD51), and ITGA2(integrin α2, CD49b), for selective capture.

Recently, the advent of liquid biopsies for precision oncology has marked a significant advancement in cancer detection, attracting considerable interest in the field of medicine. As a result,

we would propose a new complementary diagnostic platform based on liquid biopsy to address the current limitations and challenges of mammography. Our focus is on the fact that surface proteins of EVs share similarities with parental breast cancer cell surface proteins. We anticipate this approach might resolve an inability to fully capture the nuanced and unique characteristics of breast cancer tumors. By successfully isolating breast cancer-derived extracellular vesicles (BEVs) using specific surface proteins and analyzing the miRNA signature enriched in EVs, our goal is to provide clinical evidence to support the intended purpose of utilizing this approach, thus offering valuable information for breast cancer diagnosis (Fig. 1).

2. MATERIALS AND METHODS

2.1 Patients selection and plasma collection

For this study, specimens were collected from subjects in compliance with the guidelines of the Independent Ethics Committee at Yonsei University College of Medicine (IRB approval number 4-2020-1292). All individuals gave informed consent to allow the use of their plasma samples for research purposes. Of the plasma samples collected at Severance Hospital from 2010 to 2021, preoperative samples from 120 breast cancer patients, 45 healthy women, and 46 patients with benign tumors were randomly sampled. These retrospective samples were enrolled in the current study. Subjects qualified for clinical specimen collection if they met the following criteria: 1) Breast cancer or benign disease diagnosis confirmed by radiology and histopathology; 2) No treatments such as adjuvant and/or neo-adjuvant therapy before blood collection; 3) No hemolyzed plasma samples to ensure quality control; 4) No history of other malignant tumors; 5) Control group confirmed as low-risk through mammography. After blood collection in EDTA-treated tubes, the samples were centrifuged at 1500 RCF for 15 minutes. The plasma from the supernatant in the tube was transferred to 1.5ml tubes and stored at -80°C, which was used as a source for EV isolation and purification. To isolate and purify EVs, plasma samples were thawed and centrifuged at 2000 RCF for 10 minutes at 4 ºC, followed by a second centrifugation at 10,000 RCF for 30 minutes. The sample was then filtered through a 0.22 micrometer filter, and 200 microliter of pre-clarified specimen was used in the following EV isolation procedure. The filter used was a Millipore filter (catalog number: SLGPR33RB; Merck Millipore Ltd., Billerica, MA, USA).

2.2 Cell culture and EV enrichment

This study used cell lines representing various subtypes of breast tumor: MCF10a for human epithelial cells from the mammary gland (benign breast tissue), MCF7 for luminal A type, BT-474 for luminal B type, SK-BR-3 for HER2-neu, and, HCC70, HCC1937, HCC1187, MDA-MB-468, MDA-MB-453, MDA-MB-231, and Hs578T for basal-like triple-negative breast cancer (TNBC) subtypes (all obtained from the American Type Culture Collection; Manassas, VA, USA). Each cell line was cultured in Roswell Park Memorial Institute-1640 medium (RPMI-1640; cat. 22400-089), supplemented with 10% fetal bovine serum (FBS; cat. 12483-020) and 1% penicillin-streptomycin (cat. 15140-122), all from Gibco; Thermo Fisher Scientific, Inc., Waltham, MA, USA. Every cell was cultured in a monolayer culture at 37°C with 5% CO_2. Breast cell lines were cultured in RPMI medium containing 10% FBS to 70-80% confluency in order to gather EVs from cell lines. Following medium removal, the cultured cells were washed three times with phosphate-buffered saline (PBS) and re-cultured in serum-free medium at 37°C with 5% CO2 for 72 hours. The EV-rich medium was collected and centrifuged at 600 RCF for 30 minutes to eliminate residual cells and concentrate EVs. A 100K Macrosep centrifugal filter (catalog number: MAP100C37, Pall Life Science, NY, USA) was then used to further focus EVs from the cell-free supernatant.

2.3 Isolation of breast cancer-derived EVs (BEVs)

We optimized the magnetic bead-to-antibody ratio and identified the best incubation time with plasma samples to ensure efficient EV capture before isolating and profiling BEVs. An optimized magnetic bead-to-antibody ratio of 1:80 (weight/weight) was identified, with the ideal reaction conditions being 2 hours at 25˚C (Figure 4). To isolate EVs, 200 micrograms of 3 micrometer streptavidin-coated Mg beads were bound to 2.5 micrograms of biotinylated antibodies that capture essential membrane proteins of breast carcinoma: Epithelial cell adhesion molecule (EpCAM), glypican-1 (GPC1), integrin α2 (ITGA2), integrin αv (ITGAV), and integrin α6 (ITGA6). Through this optimized approach, EV membrane proteins were analyzed to ensure successful separation of BEVs via flow cytometry (FACS LSR Fortessa system, Becton Dickinson, NJ, USA). Immuno-beads were assembled and reacted with EVs obtained from several different cell lines of breast tumor—including MCF10a, MCF7, SK-BR-3, BT-474, MDA-MB-453, MDA-MB-231, MDA-MB-468, HCC70, HCC1187, HCC1937, and Hs578T—for 2 hours at 25˚C. Subsequent rinsing with PBS buffer twice served to prevent excessive binding reactions. The samples were then incubated in darkness for 30 min at 4°C with 5 microliter of anti-CD63-PE-Cy7 antibodies (catalog number: 561982; Becton Dickinson, NJ, USA). The detailed flow cytometry analysis demonstrated that EpCAM, ITGA2, and ITGAV are the most reliable membrane proteins for separating BEVs, which

precisely represent the features of the source breast carcinomas. These surface protein biomarkers were utilized in subsequent analyses in this study.

2.4 Characterization of BEVs

For the morphological analysis of the isolated BEVs, a fixation procedure was applied. BEV-bound immunobeads were incubated for 24 hours in Karnofsky fixative containing 2% paraformaldehyde (catalog number: 818715, Merck KgaA, Darmstadt, Germany) and 2% glutaraldehyde (catalog number: 354400, Merck KGaA, Darmstadt, Germany) in a 0.1 M phosphate buffer (pH 7.4). After the reaction, the solution was washed twice for 30 minutes each in 0.1 M PBS buffer to remove residual fixative. The sample was subsequently incubated with 1% osmium tetroxide (OsO4) for 2 hours to stabilize lipid membranes and improve preservation. Afterward, the samples were dehydrated using a series of ethanol concentrations ranging from 50% to 100% in a Critical Point Dryer (Cat. CPD300; Leica Microsystems, Wetzlar, Germany) and then platinum-coated using an ion sputter coater (Cat. ACE600; Leica Microsystems). The samples were then analyzed with a scanning electron microscope (SEM; Carl Zeiss AG, Oberkochen, Germany, model MERLIN) at x10,000 magnification, allowing for detailed examination of BEV formation. A Nanoparticle Tracking Analyzer (NTA; NanoSight NS300 system, Malvern Panalytical Ltd., Malvern, U.K.) was used to quantitatively evaluate the concentrations and size distributions of the tumor-derived extracellular vesicles after re-suspension in PBS. NTA 3.1 software (Malvern Panalytical Ltd.) was used for this analysis, following the manufacturer's instructions. The NTA system’s camera was precisely calibrated to identify particles with a particular signal corresponding to BEVs, ensuring accurate size distribution and quantification of the vesicles..

2.5 Analysis of gene expression omnibus (GEO) databases

To investigate the microRNAs (miRNAs) implicated in breast cancer (BC) development and progression, we conducted a comprehensive differential expression analysis utilizing ten datasets from the Gene Expression Omnibus (GEO) database in the National Center for Biotechnology Information (NCBI) (as shown in Table 1). Our analysis sought to identify miRNA expression configurations that distinguish BC patients from non-cancer controls, including benign breast tumors patients and healthy individuals. Clinical information from both tissue samples (GSE154255, GSE97811, GSE26659, GSE45666, and GSE44124) and blood samples (GSE98181, GSE118782, GSE110317, GSE73002, and GSE42128) were evaluated in this comparison. The difference in the levels of miRNAs between BC patients and the control group was illustrated using volcano plots,

thereby rendering simpler to figure out miRNAs that were highly elevated or downregulated with regard to breast cancer. Once the differentially expressed miRNAs (DEMs) were identified, the analysis was further enhanced through comparisons made with Venn diagrams. We utilized an online tool available (http://bioinformatics.psb.ugent.be/webtools/Venn/) to perform this step, enabling us to intersect the DEM lists from the previously referenced datasets.

2.6 Tumor-derived exosomal miRNA analysis

To confirm miRNA profiles in BEVs, we followed the manufacturer's protocol to extract miRNAs from cell media or plasma samples using the Total Exosome RNA and Protein Kit (cat. 4478545; Thermo Fisher Scientific Inc., Waltham, MA, USA). RNA level was determined using the Qubit™ microRNA Assay Kit (cat. no. Q32880; Thermo Fisher Scientific Inc.) in combination with a Qubit® 2.0 Fluorometer (cat. no. Q32866; Thermo Fisher Scientific Inc.). The TaqMan microRNA Reverse Transcription Kit (cat. no. 4366597; Thermo Fisher Scientific Inc.) was used to reverse transcribe the extracted RNA. From the public datasets, we selected microRNA candidates, including hsa-miR-181a, hsa-miR-106b-5p, hsa-miR-155, hsa-miR-21-5p, hsa-miR-1290, and hsa-miR-484. Using the TaqMan Universal PCR Master Mix, No AmpErase UNG (catalog no.: 4324018; Thermo Fisher Scientific Inc.) and the TaqMan microRNA Assay Kit (catalog no.: 4440887; Thermo Fisher Scientific Inc.) on a CFX96 Real-Time PCR System (catalog no.: 3600037; Bio-Rad Laboratories Inc., Hercules, CA, USA), The miRNA candidates' differential expression patterns were measured using cDNA amplification. The following parameters were followed for the reverse transcription of individual miRNAs: 30 minutes at 16°C for primer annealing, 30 minutes at 42°C for extension, and 5 minutes at 85°C to stop the reaction. The cDNA was subsequently employed for real-time PCR, which was run in 40 cycles: 10 minutes of enzyme activation at 95°C, 15 seconds of denaturation at 95°C, and 10 minutes of annealing and elongation at 60°C. To normalize miRNA expression levels, hsa-miR-16-5p was used as the reference control for exosomal miRNAs. Every experiment was conducted in duplicate and adhered to the manufacturer's instructions. The relative expression of miRNAs in BEVs was determined using the 2-ΔΔCT method.

2.7 Statistics analysis

In this study, we employed MedCalc software (version: 20.014, MedCalc Software Ltd.) to analyze each miRNA and microRNA signature using Receiver Operating Characteristic (ROC) and Precision-Recall Curves (PRC). These analyses were conducted using data from both breast cancer patients and a non-breast cancer control group. Statistical analyses, including ANOVA, t-tests, and

correlation analysis as well as data visualizations, were carried out using GraphPad Prism 10 (version: 10.2.0 (392), GraphPad Software, LLC.). Clinical information for each patient was collected retrospectively from the Electronic Medical Records (EMR). A bivariate, binormal model was used to compare each miRNA's AUC values in order to determine the most effective miRNA combination for diagnostic performance. To pinpoint the most effective miRNA signature, logistic regression analysis was performed. Selection criteria for the miRNA signature in breast cancer derived extraceulluar vesicles (BEV) included selecting the variable with the highest AUC value and the largest difference from total extracellular vesicles (TEV).

To find out the optimal combination among miRNAs, logistic regression analysis was carried out, with the miRNAs treated as independent factors. Coefficients of each biomarker from logistic regression, accompanied by significance levels and standard errors, were used to derive a prediction equation for the logit transformation of the probability (Logit(p)) of the function of interest (47–49). The predicted probability index (PI) was used to express the dichotomous dependent factor in this study. MicroRNAs were incorporated as independent factors in the logistic regression model. The significance level (α) was established at 0.05, with any p-values exceeding 0.1 excluded from the analysis. Subsequently, in order to establish a binary dependent variable indicating the impact of each microRNA on the outcome, logistic regression coefficients were calculated. Overall model fit statistics were used to evaluate the validity of the model of logistic regression. An independent variable was deemed to significantly predict the outcome if its *p*-value was less than 0.05. The MedCalc software was used to perform the statistical analyse. To determine the best miRNA combination, we compared the AUC of all predicted PI calculated through logistic regression.

2.8 MicroRNA combination and analysis with radiological data

In evaluating the clinical applicability of the microRNA combination with screening mammograms, we first categorized all mammography and ultrasonography data following international guidelines for breast radiological reporting (50). The BI-RADS system uses a seven-scoring system (0 to 6) for radiographic breast examination results. A score of 1 indicates negative findings, 2 reflects benign breast disease, and 3 suggests a probable benign condition. Category 4 suggests suspicion of malignancy, with 4a indicating low suspicion, 4b moderate suspicion, and 4c high suspicion. A score of 5 reflects a high likelihood of malignancy, while 6 is reserved for patients diagnosed with breast cancer who have undergone a biopsy. To avoid bias, a score of 6, which pertains exclusively to patients with biopsy-confirmed breast cancer, was excluded from our analysis. In addition, BI-RADS 0 was defined as incomplete assessments that are re-evaluated and recoded as BI-RADS 1–5 after further radiological examinations. Therefore, BI-RADS 0 was categorized as 'negative' since it represents an incomplete result that does not conclusively diagnose breast cancer and requires additional investigation.

In summary, our analysis focused on participants as the primary unit of analysis, with the endpoint being the breast imaging assessment obtained from mammography. For comparison with the diagnostic accuracy of the miRNA signature in clinical applications, we categorized mammography results as "positive" if they received a BI-RADS assessment of 4a, 4b, 4c, or 5, indicating findings that are "suspicious." Conversely, mammography results were considered "negative" if they received a BI-RADS assessment of 0, 1, 2, or 3, indicating no significant findings or findings not suspicious for malignancy(51). In addition, we recorded overall breast density from mammograms based on ACR BI-RADS categories: A (almost entirely fatty, < 25%), B (scattered fibroglandular density, 25–50%), C (heterogeneously dense, 51–75%), and D (extremely dense, >75%). This information allowed us to consider the impact of breast density on mammographic performance and the potential applicability of microRNA combination.

3. RESULTS

3.1 Optimization of immune-affinity methods to isolate BEVs

3.1.1 Selection of surface protein biomarkers using cell lines for each breast cancer molecular subtypes

The development of liquid biopsy assays for isolating and analyzing BEVs requires focusing on membrane proteins that are continuously elevated in multiple molecular subtypes of breast carcinoma. To identify suitable candidates, a comprehensive literature review was conducted, focusing on surface proteins with documented overexpression in breast cancer. Based on this review, EpCAM, CD49f, CD51, CD49b, and GPC1 were selected as potential targets (52–65). Subsequently, protein expression profiles of these candidates in breast cancer cell lines were assessed across different molecular subtypes using fluorescence-activated cell sorting (FACS) analysis.

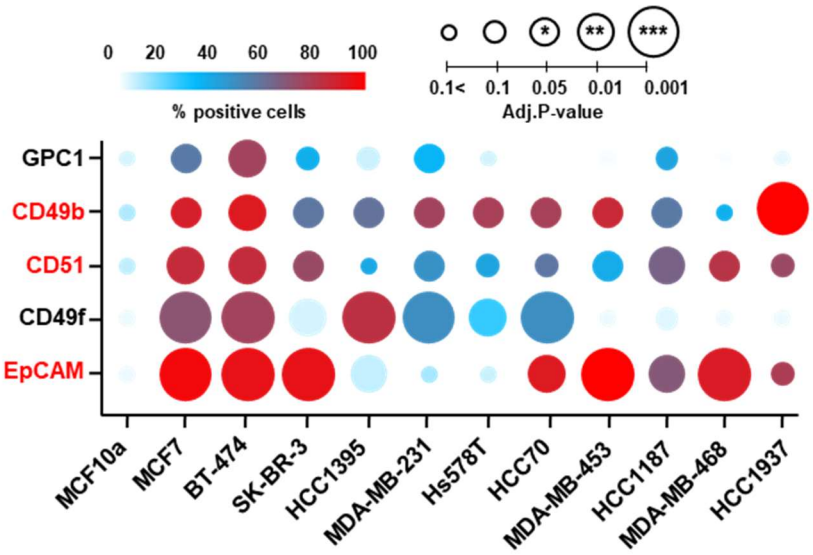

Figure 2. Profiling of surface protein biomarker candidates in cell lines by breast cancer molecular type. This figure illustrates the comparison and analysis of the expression of breast cancer surface protein biomarker candidates across 11 different breast cancer cell lines based on molecular subtypes. The x-axis represents the cell lines categorized by molecular subtype, while the y-axis indicates the breast cancer surface biomarker candidates. The size of each circle represents the associated p-value, where larger circles denote higher statistical significance. The color of each circle indicates the proportion of positive cells, with red indicating a higher proportion and blue indicating a lower proportion of positive cells.

The study utilized a panel of breast cell lines representing various molecular subtypes, including MCF10a (normal breast cell line), MCF7, SK-BR-3, and BT-474 (luminal type breast cancer cell lines), MDA-MB-453, HCC1187, MDA-MB-468, HCC70, and HCC1937 (basal-like triple-negative breast cancer cell lines), and MDA-MB-231, HCC1395, and Hs578T(mesenchymal-like TNBC cell lines). Surface protein expression in each cell line was evaluated and selected based on two criteria: 1) strong or moderate expression rates, and 2) statistical significance. GPC1 was excluded due to its relatively low expression in most breast cancer cell lines, while CD49f was omitted because of its low expression rate in TNBC cell lines and lack of statistical significance(Fig. 2, 3). Figure 3 depicted the distribution of surface protein levels across different breast cancer cell lines.

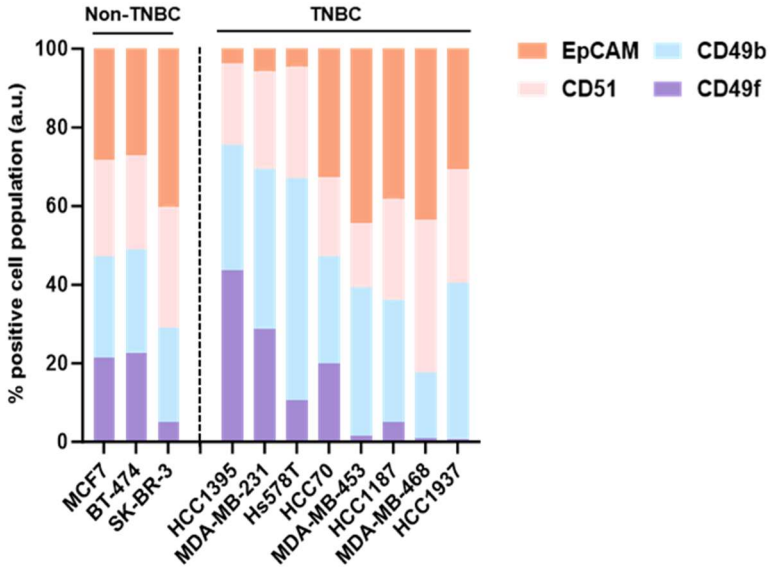

Figure 3. Distribution of overexpressed surface protein biomarkers candidates in cell lines by breast cancer molecular type. This graph illustrates the distribution of expression ratios of each surface protein in breast cancer cell lines by molecular subtype. The y-axis represents the population of positive cells, while the x-axis indicates the cell lines categorized by molecular subtype. The vertical dashed line distinguishes between TNBC (Triple-Negative Breast Cancer) and non-TNBC cell lines.

EpCAM exhibited low expression levels in mesenchymal-like TNBC cell lines, with percentages of 3.79% in HCC1395, 5.64% in MDA-MB-231, and 4.51% in Hs578T cell lines (Fig. 3). While EpCAM serves as a well-established surface marker for isolating EVs, its expression was notably lower in mesenchymal-type TNBC compared to luminal and basal-type TNBC (Fig 3). Thus, targeting EpCAM alone proved inefficient for isolating BEVs from mesenchymal-type TNBC cell lines. Conversely, both CD49b and CD51 exhibited high expression rates across luminal breast cancer and basal and mesenchymal forms of TNBC (Fig. 3). Therefore, targeting cells expressing a combination of EpCAM, CD49b, and CD51 was recommended for efficient isolation of BEVs. This multi-marker approach was confirmed to yield successful separation, particularly in mesenchymal-type TNBC (Fig. 3). These findings suggest that the overexpression of surface proteins in breast cancer may have a significant impact on its development and could potentially be valuable biomarkers. Further exploration into the functional significance of selectively isolating EVs with these overexpressed surface proteins in breast cancer could offer valuable insights into its biology and contribute to the development of novel diagnostic and therapeutic approaches.

3.1.2 Optimization of immune-affinity method to isolate BEVs

To establish optimal conditions for isolating BEVs using immuno-affinity methods, we optimized test specifications to determine the ideal concentration ratio, reaction time, and reaction temperature of antibodies and magnetic beads.

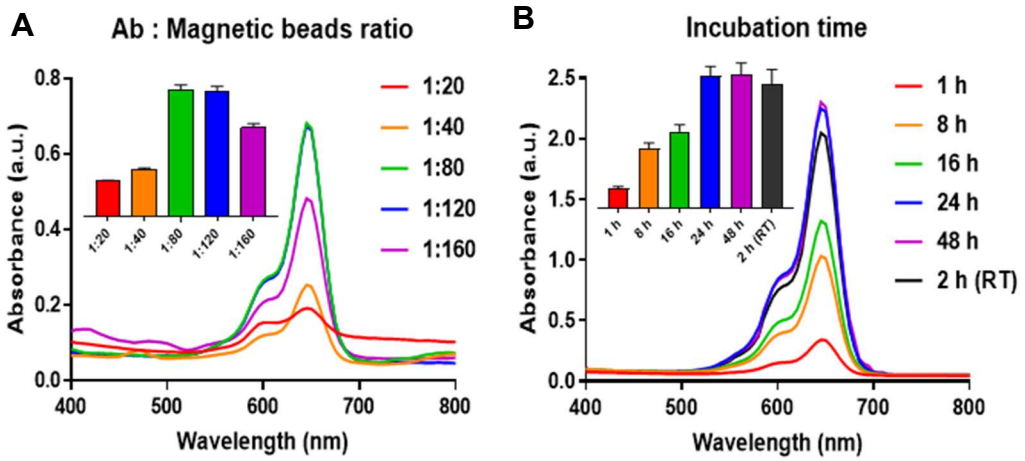

Figure 4. Optimization of immunoaffinity-based BEVs isolation technology. The x-axis of both graphs represents wavelength (nm), while the y-axis represents absorbance. (A) Graph depicting the analysis conducted to determine the optimal ratio of antibodies to magnetic beads. (B) Graph illustrating the analysis performed to establish the optimal reaction time.

The results of the specifications indicated that the optimal ratio between each antibody and magnetic beads was 1:80 (Fig. 4A). Additionally, the most favorable reactivity was achieved with a reaction time of 2 hours at room temperature (Fig. 4B). Consequently, we refined the immuno-affinity method to isolate BEVs using these validated conditions.

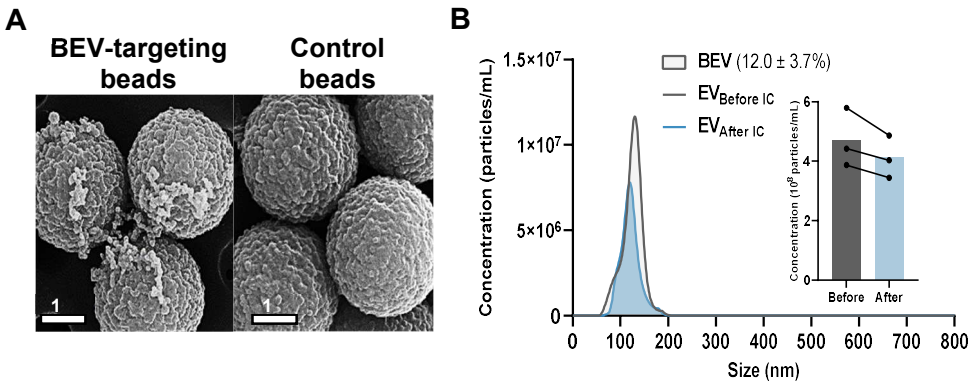

Figure 5. Verification of immunoaffinity-based BEVs isolation technology. (A) Image obtained through scanning electron microscopy after binding BEVs to magnetic beads. (B) Graph showing the size and concentration of EVs analyzed using a nanoparticle tracking analyzer. The x-axis represents the size of analyzed EVs (nm), while the y-axis represents the concentration of EVs (particles/ml). "Before" denotes the total number of EVs in plasma before BEV isolation, while "After" represents the remaining total EVs in plasma after BEV isolation. The number of BEVs was determined by subtracting "After" from "Before".

The figure above depicts the results of scanning electron microscope (SEM) observation following the optimized isolation method for BEVs (Fig. 5A). The images showed the presence of antibodies bound to magnetic beads, validating the effectiveness of the immuno-affinity isolation process. Additionally, the isolated BEVs were observed to fall within the size range of 100 to 200 nm by using nanoparticle tracking analyzer(NTA), consistent with established characteristics of such EVs (Fig. 5B). These findings underscore the successful implementation of the optimized isolation method, ensuring reliable and efficient extraction of BEVs for further analysis. Furthermore, the effectiveness of the BEV isolation in plasma was confirmed by NTA, with an estimated yield of 12.0%±3.7% based on EV concentration measurements before and after the immuno-capture process (Fig. 5B).

3.2 MicroRNA candidate selection

3.2.1 MicroRNA candidate screening from public dataset

To discover potential candidates of microRNAs for diagnosing breast carcinoma, the database of Gene Expression Omnibus (GEO) was examined. Our approach involved a comprehensive examination of datasets from publicly available repositories to identify candidate miRNAs present in BEVs. Specifically, we compared datasets from repositories analyzing tissue-derived and circulating miRNAs. By cross-referencing these databases, we aimed to identify miRNAs consistently implicated in breast cancer pathogenesis and those particularly enriched in circulating tumor-derived EVs. This comparative analysis allowed us to select strong candidate miRNAs for further investigation.

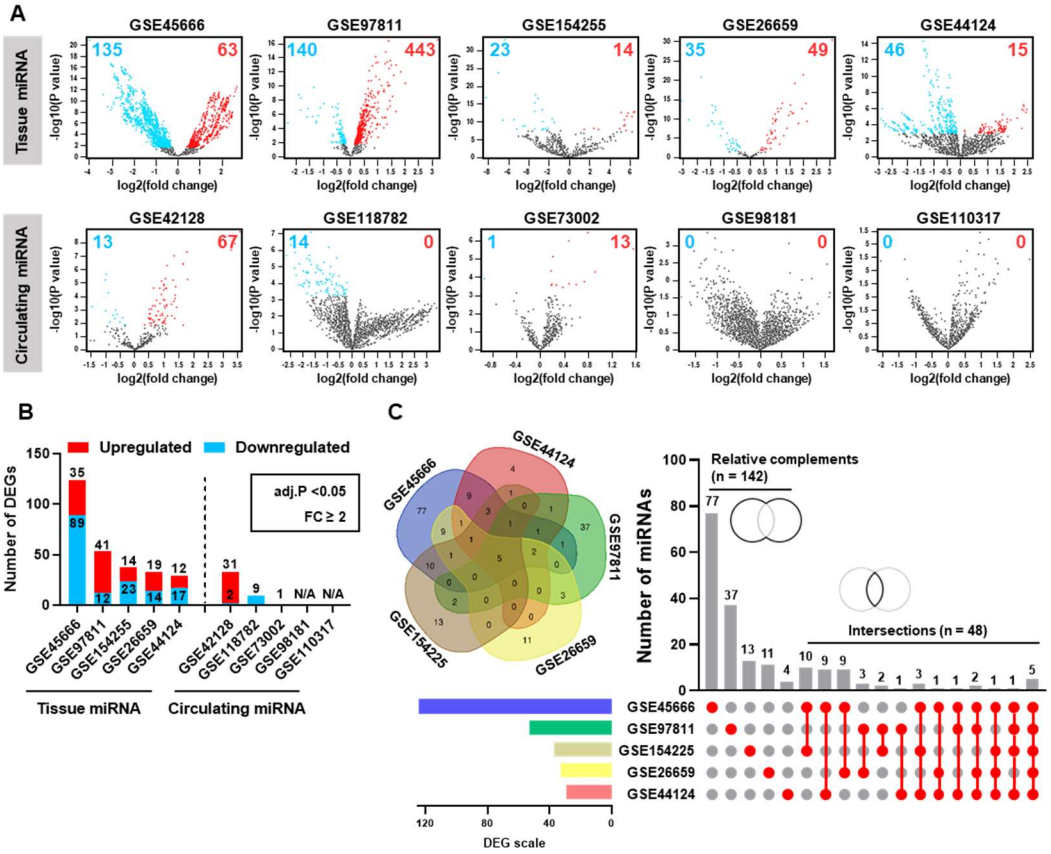

Figure 6. Comparison of tumor-derived and circulating miRNA studies based on public database. (A) Differentially expressed miRNAs (DEMs) in the tissue miRNA dataset displayed as volcano plots (GSE26659, GSE97811, GSE44124, GSE154255, and GSE45666) and circulating miRNA dataset(GSE42128, GSE118782, GSE73002, GSE98181, and GSE110317). The volcano plots showed the DEMs in tumor tissues of breast cancer patients compared with adjacent tissues; red dots represent highly expressed miRNAs in breast tumors, and blue dots represent low expressed miRNAs. (B) Graph showing up-regulated DEGs and down-regulated DEGs for each dataset when strictly adjusting the criteria of DEMs in the database (fold change ≥2, P<0.05) (C) Graph showing DEG Scale analysis of 142 independently overexpressed and 48 commonly overexpressed DEMs in tissue-derived miRNA datasets (GSE44124, GSE45666, GSE154255, GSE26659, and GSE97811)

The public datasets analyzed are summarized in Table 1. In five separate GEO datasets (GSE26659, GSE97811, GSE44124, GSE154255, and GSE45666) containing breast cancer tissue samples, our study identified a large number of differentially expressed microRNAs (DEMs). Across the five GEO datasets, we found 584 upregulated and 379 downregulated DEMs in the tissue

samples (Fig. 6A). In contrast, circulating miRNA databases (GSE42128, GSE118782, GSE73002, GSE98181, GSE110317) showed a relatively smaller number of miRNA candidates with statistically significant differences compared to tissue-derived studies (66–75). Only 28 down-regulated and 80 up-regulated DEMs were found in blood samples across five additional GEO datasets. Applying more stringent filtering criteria (adjusted fold change ≥ 2 and p-value < 0.05) further narrowed the range, leading to sharper contrasts (Fig. 6B). Our analysis revealed 13 candidate miRNAs— miR-16, miR-9, miR-21, miR-10b, miR-106b, miR-128, miR-96, miR-181a, miR-429, miR-484, miR-1260b, miR-155, and miR-1290—among the relative complements (142 DEMs) and intersections (48 DEMs) shared across the five GEO datasets on breast cancer, with significant likelihood of expression in BEVs (Fig. 6C). Considering that EVs originate from parental cells, share a membrane, and contain real-time intracellular information of molecules, this study prioritized candidate miRNAs identified through a tissue-derived miRNA database. This decision was based on the premise that tissue-derived miRNAs would more accurately reflect the miRNA composition of BEVs.

Table 1. Summary of clinical studies on tissue-derived and circulating miRNA for selection of miRNA candidates in public databases.

GSE ID	Cohorts (BC *vs* N)	Total	Samples	Platform	Year
Tissue miRNA analysis					
GSE26659	77 *vs* 17	94	Ductal breast carcinoma biopsies & normal tissues from mammoplastic reductions	Agilent human miRNA microarray (V2)	2012
GSE45666	101 *vs* 15	116	Breast tumor tissues & adjacent normal tissues	Agilent human miRNA microarray (V3)	2013
GSE44124	50 *vs* 30	80	Breast tumor tissues & adjacent normal tissues	Agilent human miRNA microarray (V3)	2014
GSE97811	45 *vs* 16	61	Breast tumor tissues & adjacent normal tissues	3D-Gene® Human miRNA (V21)	2017
GSE154255	10 *vs* 10	20	Breast tumor tissues & adjacent normal tissues	Agilent human miRNA microarray (V19)	2023
Circulating miRNA analysis					
GSE42128	32 *vs* 22	54	Serum samples from BC patients & healthy individuals	Exiqon miRCURY LNA microRNA array	2013
GSE73002	1,290 *vs* 54	1,344	Serum samples from BC patients & patients with benign breast disease	3D-Gene® Human miRNA (V20)	2016
GSE98181	24 *vs* 24	48	Serum samples from BC patients & cancer-free women	Affymetrix Multispecies miRNA	2018
GSE118782	30 *vs* 10	40	Plasma samples from BC patients & healthy women	Affymetrix Multispecies miRNA	2019
GSE110317	921 *vs* 37	958	Serum samples from BC patients & patients with benign breast disease	3D-Gene® Human miRNA (V21)	2023

3.2.2 MiRNA candidate screening through cell lines covering molecular subtypes of breast cancer

Figure 7 illustrates the relative quantities of thirteen miRNA candidates in BEV from cell lines, categorized by molecular subtype. To further analyze the expression patterns of these miRNA candidates in BEVs, qRT-PCR was employed. This assay compared each miRNA's CT values in 11 types of breast cancer cell lines to those in the healthy controls, MCF-10A. The values presented in the figure represent the ratio of CT values in each cell line to the expression of the respective miRNA in the normal control group.

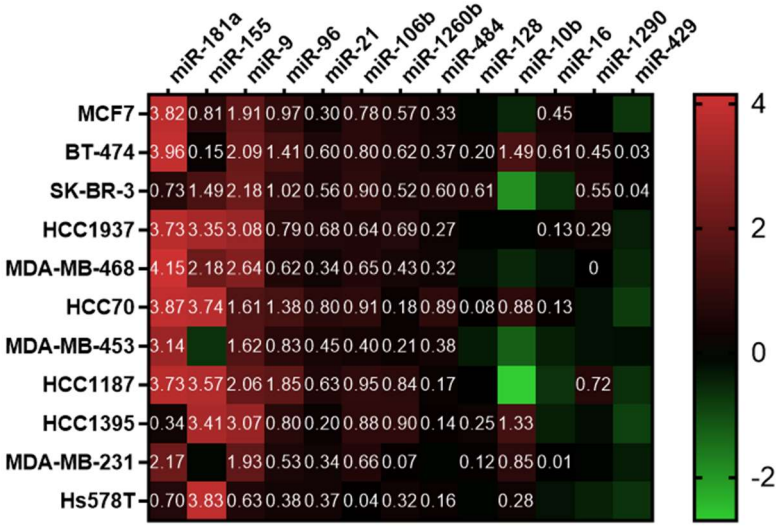

Figure 7. Comparison of miRNA candidates in cell lines by breast cancer molecular type. This heatmap illustrates the expression levels of 13 miRNA candidates across 11 different breast cancer cell lines categorized by molecular subtype. Black indicates a fold change of 0, while red represents higher fold changes and green indicates lower fold changes. Fold change values were calculated relative to the CT values of the MCF-10A cell line used as the control group.

Among the 11 breast cancer cell lines classified by molecular subtype, miRNAs with a fold change exceeding 0 in at least 8 cell lines were prioritized as candidate miRNAs for further investigation. As a result, miR-181a, miR-155, miR-9, miR-96, miR-21, miR-106b, miR-1260b, and miR-484 were found to be up-regulated in the majority of breast cancer cell lines. Therefore, based on these findings, we conducted an integrated analysis with the results of tissue-derived miRNA assays.

3.2.3 MicroRNA candidate screening through breast cancer tissue

We hypothesized that miRNA expression profiles in BEVs isolated using breast cancer-specific membrane proteins may differ from those in TEVs and could potentially offer greater specificity for breast cancer. To explore this possibility, we undertook an integrated analysis of miRNA expression patterns using both cell line test results and tissue analysis results. The figure below presents the results of analyzing 13 types of miRNA across 20 breast cancer tissues and 20 matched normal tissues.

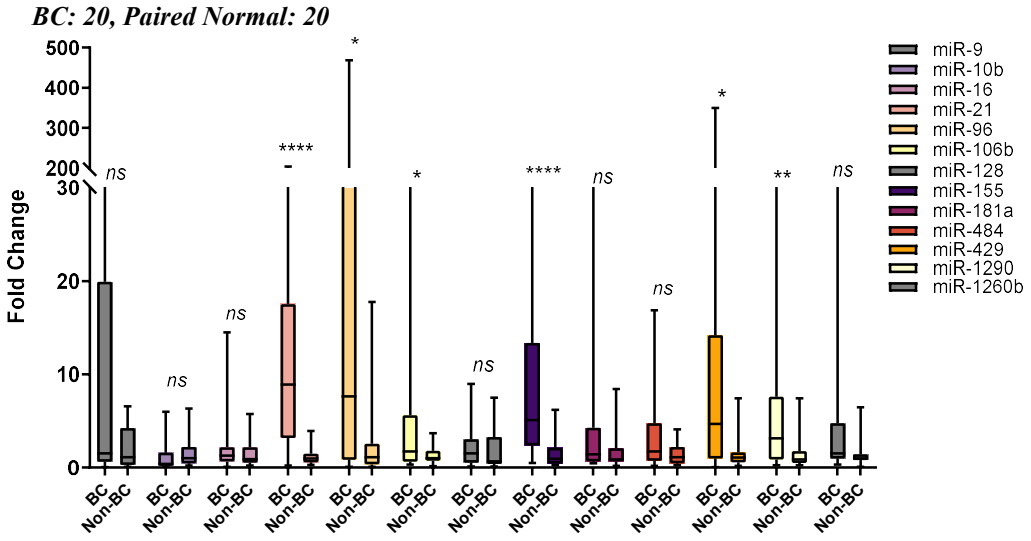

Figure 8. Comparison of miRNA candidates in BEV from breast cancer tissues. This graph illustrates the expression of 13 miRNA candidates using breast cancer tissues and paired normal tissues. The x-axis labels "BC" represent breast cancer tissues, while "Non-BC" represents paired normal tissues. The y-axis represents each miRNA candidate's fold change, calculated relative to the MCF10A reference. "ns" indicates statistically non-significant results, "*" denotes $p<0.05$, "**" denotes $p<0.01$, and "***" denotes $p<0.0001$.

The analysis revealed significant differences in the expression levels of certain miRNAs between breast cancer patients and matched normal tissues (Fig. 8). Specifically, microRNA-21, -106b, -96, -429, -155, and -1290 exhibited significantly higher expression levels in breast cancer

patients compared to normal tissues (with p-values ranging from < 0.0001 to < 0.05). On the other hand, while microRNA-9, -181a, -484, and -1260b were also highly expressed in breast cancer patients, the differences in expression levels compared to normal tissues were not statistically significant (Fig. 8). These findings suggest that microRNA-21, -106b, -96, -429, -155, and -1290 may serve as potential biomarkers for breast cancer diagnosis or prognosis due to their differential expression in breast cancer patients. However, because the number of patient tissues participating in the test is very small (40 cases in total), a very careful approach is required in interpreting the results. Therefore, we attempted to select the final candidates by integrating the cell line test and tissue test results. Seven miRNAs' expression profiles were determined considering integrating the outcomes of cell line testing and tissue sample analysis, indicating that these miRNAs may serve as surrogate markers for the diagnosis of breast cancer.

3.3 Discovery set of miRNAs candidate

The selection process for candidate miRNAs involved a comprehensive approach combining results from cell line and tissue analyses. From this integrated analysis, microRNA-155, -21, and -106b merged as commonly overexpressed microRNAs in both breast cancer cell lines and tissues. Additionally, microRNA-181a, -484, -1290, and -1260b, which showed high expression levels in cell line tests and/or tissue analyses, were included for further validation. Consequently, the final set of selected miRNA candidates for validation in plasma comprised microRNA-21, -155, -181a, -106b, -484, -1290, and -1260b. To analyze the expression patterns of these seven microRNAs in BEVs, qRT-PCR was performed using plasma samples from a total of 40 individuals (20 breast cancer patients and 20 non-breast cancer women). The study aimed to compare and analyze the expression of these miRNAs in tissues, TEVs obtained by PEG precipitation, and BEVs obtained using an immuno-affinity method which we developed (Fig. 9).

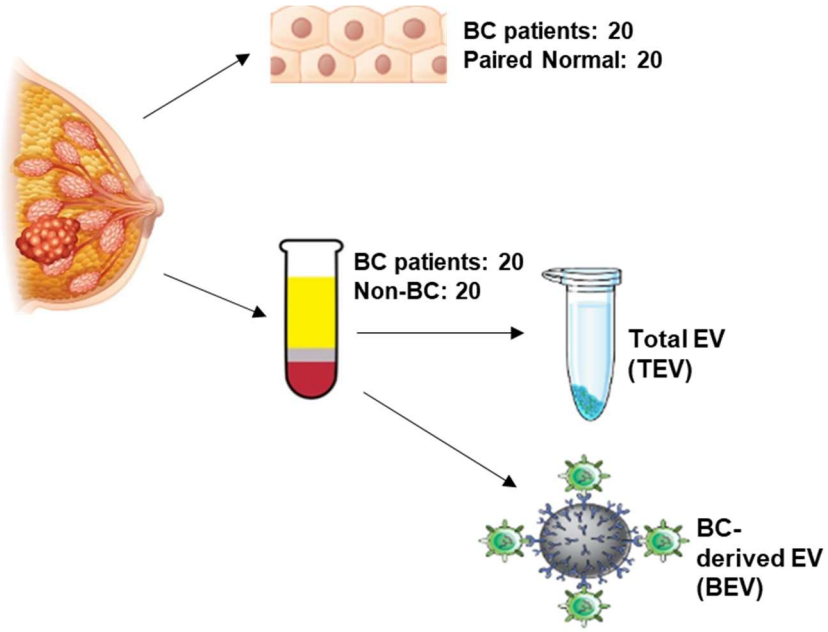

Figure 9. Scheme of discovery set to compare BEV with TEV between tissue and plasma. This schematic outlines the research overview to compare tissue-derived miRNA with TEV (Tissue Extracellular Vesicle)-derived miRNA and BEV (Blood Extracellular Vesicle)-derived miRNA to assess BEV performance. The experimental groups for tissue-derived miRNA analysis included breast cancer tissues and paired normal tissues, while plasma samples from breast cancer patients and patients with benign breast conditions were used to analyze TEV and BEV-derived miRNA.

This approach allows for a comprehensive assessment of miRNA expression patterns across different sample types, providing insights into the potential utility of these miRNAs as biomarkers for breast cancer diagnosis.

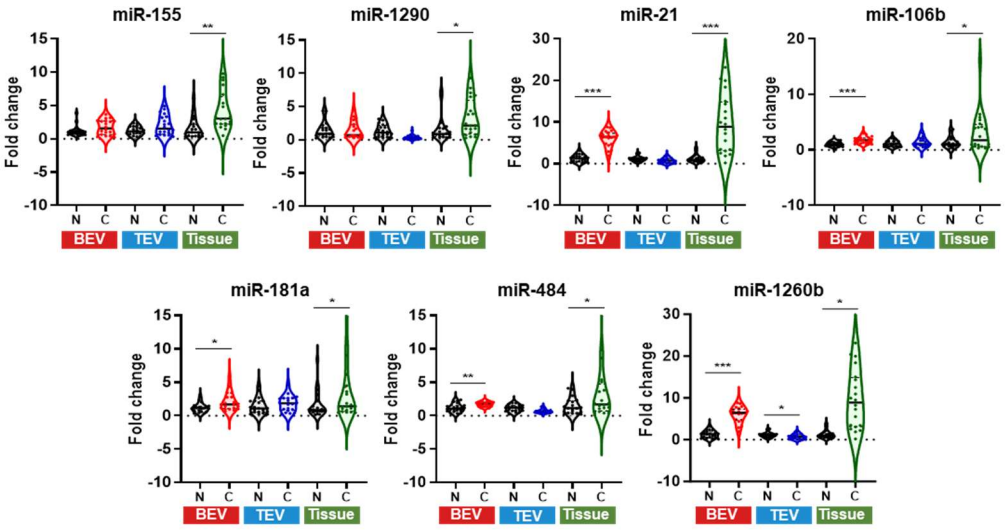

Figure 10. Selection of miRNA candidates among tissue, BEVs, and TEVs. This graph compares the expression of miRNA candidates derived from tissue, BEVs, and TEVs between breast cancer patients and non-breast cancer females. The x-axis labels "N" represent healthy controls or paired normal tissues, while "C" represents breast cancer tissues. The y-axis represents each miRNA's fold change. "*" denotes $p<0.05$, "**" denotes $p<0.01$, and "***" denotes $p<0.0001$.

To evaluate the clinical value of candidate miRNAs in BEVs obtained through an immuno-affinity method using surface proteins that can cover molecular subtypes of breast cancer cell lines, it was compared with TEVs and tissue-derived miRNAs. Among the 7 BEV-derived miRNAs, 5 miRNAs (miR-21, -106b, -484, and -1260b) excluding microRNA-1290 and -155 were upregulated and showed statistically significant differences (Fig. 10). MicroRNA-155 and -1290 were upregulated in BEV, but the difference was not statistically significant (Table 2).

Table 2. Comparison of mean fold change values and *P*-values for miRNA candidates in BEVs.

microRNA	Mean fold change in BEV	SEM	P-value
miR-21	6.08	0.42	< 0.001
miR-106b	1.73	0.13	< 0.001
miR-181a	2.05	0.30	< 0.05
miR-484	1.72	0.08	< 0.01
miR-1260b	2.31	0.25	< 0.001
miR-155	1.77	0.26	0.1482
miR-1290	0.30	0.17	0.7936

Given the heightened expression of these miRNAs in BC patients, we evaluated ROC curve analysis to compare the diagnostic ability of five miRNA candidates to detect breast cancer. According to the findings of this study, the candidate miRNAs' AUC values in tumor tissues ranged from 0.640 to 0.920, in TEVs from 0.577 to 0.785, and in BEVs from 0.790 to 0.987 (Fig. 11). It is noteworthy that candidate miRNAs performed better in BEVs for diagnostic purposes than in TEVs or tumor tissues overall, with all AUC values above 0.7. MiR-21 and miR-106b, in particular, demonstrated considerable potential as diagnostic markers, with significant AUC value differences between the TEV and BEV sets (p-value < 0.01). As a result, microRNA -21, -181a, -106b, -484, and -1260b in BEVs are recognized as trustworthy microRNA candidates for improving diagnostic methods for breast cancer. These outcomes imply that effective EV isolation from tumors enhances miRNA analysis, providing a more precise representation of the tumor's molecular profile. The results of ROC analysis for the expression of five candidate-derived and BEV-derived miRNAs are shown below.

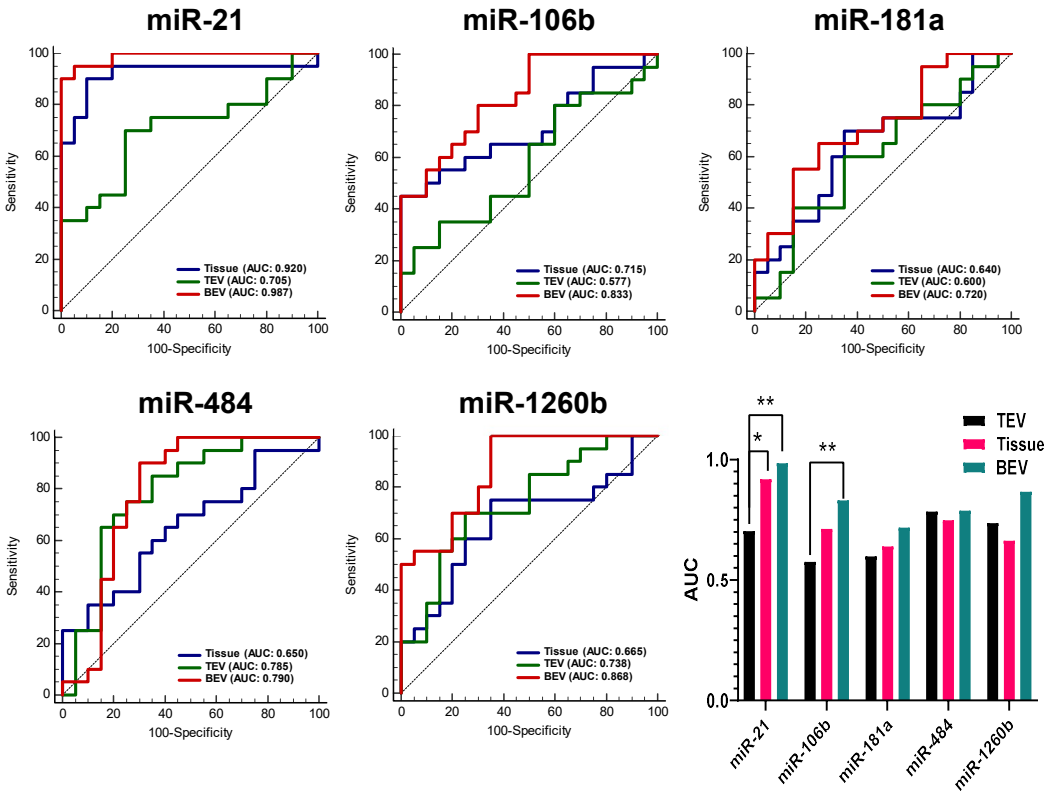

Figure 11. Comparison of ROC analysis of miRNA candidates between tissue-derived, BEV, and TEV. This graph presents the ROC (Receiver Operating Characteristic) analysis comparing miRNA candidates derived from tissue, BEVs, and TEVs. "**" denotes $p<0.01$, and "*" denotes $p<0.05$.

3.4 Validation set of BEV compared to TEV

3.4.1 Comparative analysis of miRNA candidate in TEV and BEV

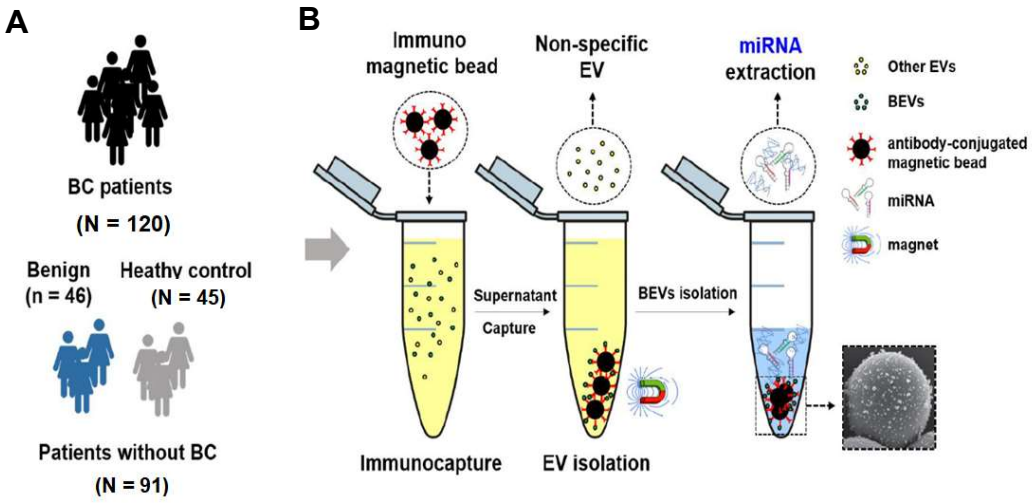

Figure 12. Study design of clinical performance test to validate miRNA analysis in BEV. (A) This section represents the subjects recruited for the analysis of miRNA in BEVs. Between May 2010 and August 2021, a total of 120 patients with breast cancer, 45 women in healthy condition, and 46 patients with benign tumors had their pre-operative plasma samples collected at Severance Hospital. The study was registered retrospectively (approved date: January 4, 2021; Seoul, South Korea; IRB approval no. 4-2020-1292). (B) Overview of isolating BEVs and analyzing miRNA from plasma samples of the subjects.

The study included a total of 211 subjects: 120 BC patients, 46 patients with benign breast disease, and 45 women with no breast cancer-related conditions (Fig. 12A). The process of isolating BEV and extracting target miRNA is shown in Figure 12B. The study aimed to confirm the clinical effectiveness of BEV-derived miRNAs for breast cancer diagnosis. Comparative analysis between BEV-derived and TEV-derived miRNAs was carried out. Due to limited plasma volume in 6 breast cancer patients, the analysis used data from the remaining 114 patients (Fig. 13).

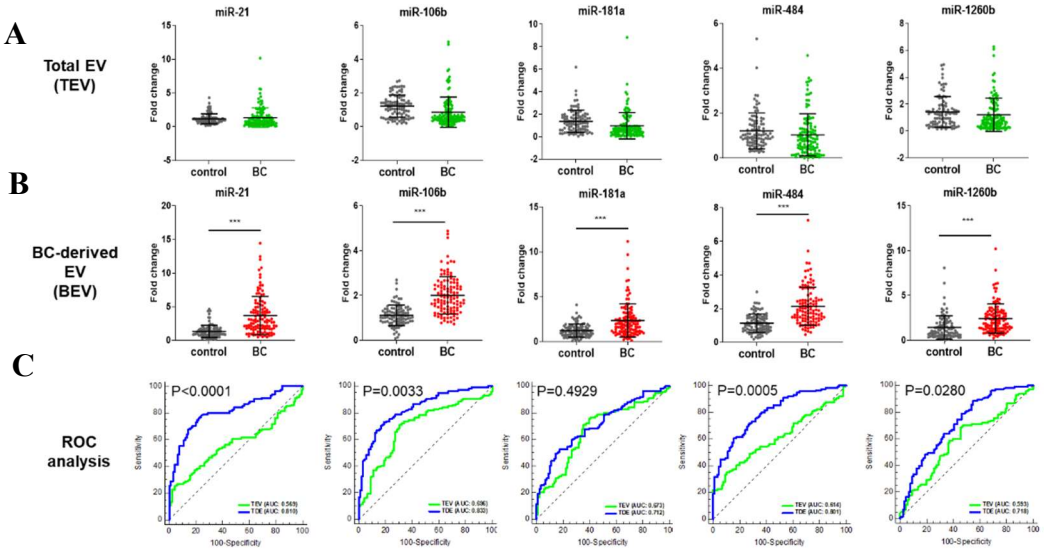

Figure 13. Comparison of miRNA candidate in clinical specimens in validation set. (A) Graph comparing the expression of TEV-derived miRNA candidates between the control group and breast cancer patients' plasma. (B) Graph comparing the expression of BEV-derived miRNA candidates between the control group and breast cancer patients' plasma. (C) ROC curve comparing the AUC difference between BEVs and TEVs.

As shown in Figure 13, each miRNA in BEVs was found to be relatively overexpressed compared to TEVs. Notably, all miRNA candidates did not show statistically significant differences between breast cancer patients and controls in TEVs (Fig. 13A), whereas all BEV-derived miRNA candidates showed statistically significant differences ($P < 0.001$, Fig. 13B). ROC analysis was performed to compare the differences in clinical performance between BEV and TEV. The analysis showed that the AUC of each miRNA candidate in TEV ranged from 0.569 to 696, while in BEV it ranged from 0.718 to 0.833 (Fig. 13C). The difference in AUC of all miRNAs except miR-181a ($p = 0.4929$) was statistically significant. Even though the difference in AUC between BEV and TEV did not exhibit statistical significance, miR-181a was nevertheless included in the miRNA candidates due to its statistically significant difference in the plasma analysis comparing breast cancer and control groups.

3.5 Validation set of 5-miRNA signature

3.5.1 miRNA signature selection

The evaluation of each microRNA's clinical performance was supported by a thorough comparison between TEVs and BEVs. Subsequently, an extensive examination involving 26 different miRNA combinations was conducted through logistic regression analysis. Table 3 outlines the logistic regression equation and clinical performance of combinations of five types of miRNA candidates.

Table 3. Logistic regression comparison for discovering miRNA combinations.

	Combination		Logistic regression equation	Sensi.[a]	Speci.[b]	AUC	SE	95% CI
1	**miR-21**	**miR-106b**	**-4.181+0.748*LN(miR_21)+1.998*LN(miR_106b)**	**81.58%**	**78.02%**	**0.881**	**0.0235**	**0.828 to 0.922**
2	miR-21	miR-181a	-1.987+0.813*LN(miR_21)+0.351*LN(miR_181a)	70.18%	78.02%	0.806	0.0302	0.745 to 0.857
3	miR-21	miR-484	-3.003+0.677*LN(miR_21)+1.230*LN(miR_484)	78.07%	75.82%	0.846	0.0265	0.789 to 0.893
4	miR-21	miR-1260b	-1.650+0.904*LN(miR_21)0.013*LN(miR_1260b)	71.93%	80.22%	0.810	0.03	0.749 to 0.861
5	miR-106b	miR-181a	-3.390+2.158*LN(miR_106b)+0.298*LN(miR_181a)	74.56%	78.02%	0.837	0.0275	0.779 to 0.884
6	miR-	miR-	-4.25+1.847*LN(miR_106b)+1.	79.82%	78.02%	0.865	0.0253	0.810 to 0.909

	106b	484		180*LN(miR_484)					
7	miR-106b	miR-1260b		-3.340+2.240*LN(miR_106b)+0.191*LN(miR_1260b)	77.19%	74.73%	0.838	0.0275	0.781 to 0.886
8	miR-181a	miR-484		-2.636+0.379*LN(miR_181a)+1.489*LN(miR_484)	72.81%	68.13%	0.808	0.0295	0.747 to 0.859
9	miR-181a	miR-1260b		-1.533+0.704*LN(miR_181a)+0.335*LN(miR_1260b)	68.42%	65.93%	0.732	0.0347	0.666 to 0.792
10	miR-484	miR-1260b		-2.413+2.035*LN(miR_484)-0.225*LN(miR_1260b)	73.68%	72.53%	0.809	0.0297	0.749 to 0.861
11	miR-21	miR-106b	miR-181a	-4.199+0.744*LN(miR_21)+1.973*LN(miR_106b)+0.032*LN(miR_181a)	79.82%	78.02%	0.880	0.0236	0.828 to 0.921
12	**miR-21**	**miR-106b**	**miR-484**	**-4.780+0.630*LN(miR_21)+1.716*LN(miR_106b)+0.822*LN(miR_484)**	**82.46%**	**80.22%**	**0.893**	**0.0224**	**0.842 to 0.931**
13	**miR-21**	**miR-106b**	**miR-1260b**	**-4.14+0.809*LN(miR_21)+2.031*LN(miR_106b)-0.117*LN(miR_1260b)**	**81.58%**	**80.22%**	**0.881**	**0.0235**	**0.829 to 0.922**
14	miR-21	miR-181a	miR-484	-3.087+0.647*LN(miR_21)+0.140*LN(miR_181a)+1.180*LN(miR_484)	74.56%	73.63%	0.844	0.0267	0.787 to 0.891

15	miR-21	miR-181a	miR-1260b		-1.983+0.817*LN(miR_21)+0.352*LN(miR_181a)-0.007*LN(miR_1260b)	70.18%	78.02%	0.806	0.0301	0.745 to 0.858
16	miR-21	miR-484	miR-1260b		-3.296+0.886*LN(miR_21)+1.860*LN(miR_484)-0.567*LN(miR_1260b)	78.95%	79.12%	0.865	0.025	0.810 to 0.908
17	miR-106b	miR-181a	miR-484		-4.262+1.795*LN(miR_106b)+0.084*LN(miR_181a)+1.149*LN(miR_484)	79.82%	78.02%	0.864	0.0254	0.810 to 0.908
18	miR-106b	miR-181a	miR-1260b		-3.516+2.057*LN(miR_106b)+0.267*LN(miR_181a)+0.171*LN(miR_1260b)	75.44%	75.82%	0.842	0.0273	0.784 to 0.889
19	miR-106b	miR-484	miR-1260b		-4.286+1.855*LN(miR_106b)+1.450*LN(miR_484)-0.206*LN(miR_1260b)	80.70%	76.92%	0.870	0.025	0.816 to 0.913
20	miR-181a	miR-484	miR-1260b		-2.659+0.372*LN(miR_181a)+1.771*LN(miR_484)-0.211*LN(miR_1260b)	72.81%	71.43%	0.815	0.029	0.755 to 0.866
21	**miR-21**	**miR-106b**	**miR-181a**	**miR-484**	**-4.763+0.640*LN(miR_21)+1.751*LN(miR_106b)-0.083*LN(miR**	**82.46%**	**81.32%**	**0.892**	**0.0224**	**0.841 to 0.931**

22	**miR-21**	**miR-106b**	**miR-181a**	**miR-1260b**		**_181a)+0.848*LN(miR_484)-4.166+0.805*LN(miR_21)+2.014*LN(miR_106b)+0.040*LN(miR_181a)-0.118*LN(miR_1260b)**	**80.70%**	**80.22%**	**0.881**	**0.0236**	**0.829 to 0.922**
23	**miR-21**	**miR-106b**	**miR-484**	**miR-1260b**		**-5.039+0.822*LN(miR_21)+1.677*LN(miR_106b)+1.407*LN(miR_484)-0.494*LN(miR_1260b)**	**84.21%**	**82.42%**	**0.906**	**0.0212**	**0.857 to 0.942**
24	miR-21	miR-181a	miR-484	miR-1260b		-3.336+0.867*LN(miR_21)+0.080*LN(miR_181a)+1.818*LN(miR_484)-0.560*LN(miR_1260b)	78.07%	79.12%	0.864	0.0251	0.809 to 0.908
25	miR-106b	miR-181a	miR-484	miR-1260b		-4.291+1.810*LN(miR_106b)+0.070*LN(miR_181a)+1.418*LN(miR_484)-+0.202*LN(miR_1260b)	80.70%	76.92%	0.870	0.0251	0.816 to 0.913
26	**miR-21**	**miR-106b**	**miR-181a**	**miR-484**	**miR-1260b**	**-5.027+0.841*LN(miR_21)1.737*LN(miR_106b)-0.133*LN(miR_181a)+1.466*LN(miR_484)-0.506*LN(miR_1260b)**	**85.09%**	**84.62%**	**0.905**	**0.0213**	**0.856 to 0.941**

[a] Sensi.: Sensitivity, [b]Speci.: Specificity

Figure 14 presents a comparative analysis of the top 5 ranking miRNA combinations, identified for their excellent clinical performance, among the combinations of miRNAs identified through logistic regression.

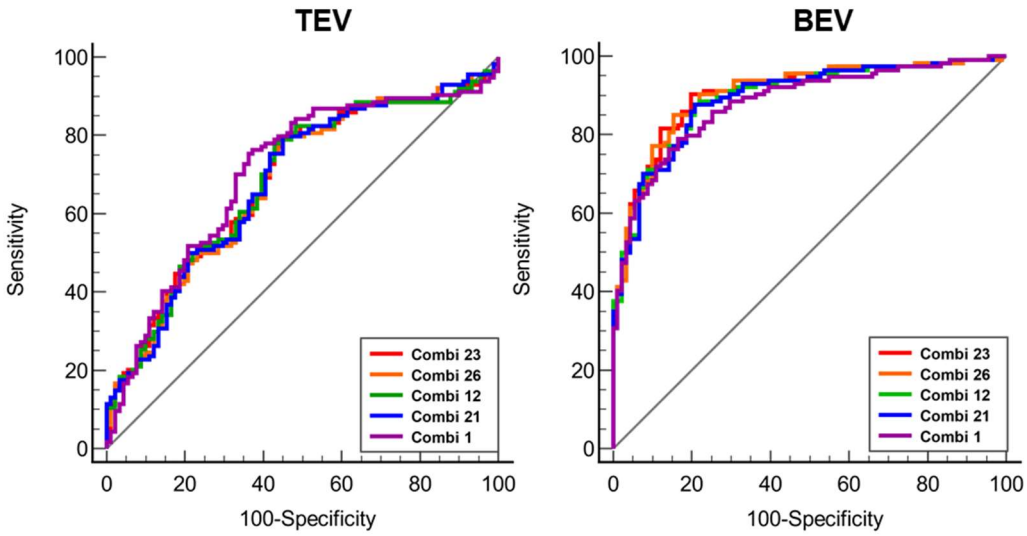

Figure 14. Comparison of clinical performance of TEV and BEV for selection of miRNA signature. This graph compares the ROC curves of the top 5 miRNA combinations within BEV-derived miRNA signatures with those of TEV-derived miRNA signatures for breast cancer diagnosis.

When selecting BEV-based miRNA signatures, the highest area under the curve(AUC) value and the difference from TEV are the most important factors. Among the top five BEV combinations, combination 23 (microRNA 21, -106b, -484, and -1260b) exhibited the highest AUC value (0.906), while combination 26 (microRNA 21, -106b, -181a, -484, and -1260b) demonstrated the most significant difference between BEV and TEV (Table 4). Therefore, combination 26 was selected as the final microRNA signature.

Table 4. Comparison of top 5 BEV-derived miRNA combinations with TEV-derived miRNA Combinations.

Rank	Variable	AUC (95% Confidential interval)		Difference	Remark
		TEVs	BEVs		
1	Combination 23	0.682 (0.614 - 0.745)	0.906 (0.857 - 0.942)	32.84%	
2	**Combination 26**	**0.680 (0.612 - 0.744)**	**0.905 (0.856 - 0.941)**	**33.09%**	**miRNA signature**
3	Combination 21	0.682 (0.613 - 0.745)	0.892 (0.841 - 0.931)	30.79%	
4	Combination 12	0.684 (0.615 - 0.747)	0.893 (0.842 - 0.931)	30.56%	
5	Combination 1	0.703 (0.635 - 0.764)	0.881 (0.828 - 0.922)	25.32%	

3.5.2 Clinical performance of miRNA signature

As the miRNA signature, selected miRNA combination 26th was defined, and ROC analysis was conducted to validate its final clinical performance using data from a total of 211 subjects, including 120 breast cancer patients, 46 patients with benign breast disease, and 45 healthy donors. The control group was further divided into subgroups, including the patients with benign breast disease, and healthy donors, and ROC curve analysis was performed specifically for breast cancer diagnosis (Fig. 15). It is important to acknowledge that the reliability of clinical outcome measures can be influenced by the number of participants recruited and potential selection bias. This means that there could be imbalances resulting from the selection of subjects expected to yield the best results. To address this concern and enhance the validity of the breast cancer diagnostic accuracy of the miRNA signature, we also calculated the Area Under the Precision-Recall Curve (AUPRC) and the F1max value through precision-recall curve analysis.

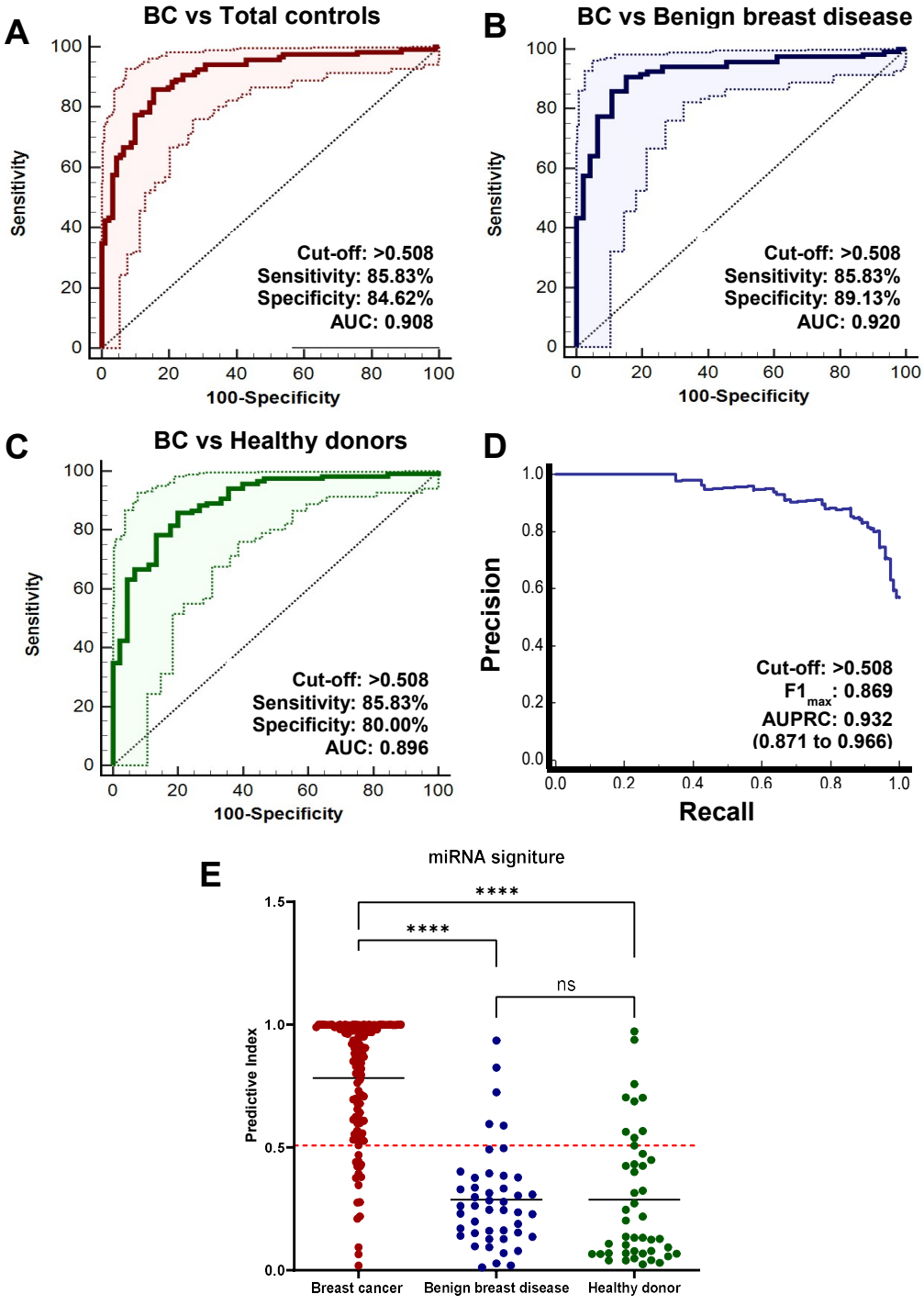

Figure 15. Diagnostic performance of microRNA signature derived from BEV. (A) ROC curve analysis comparing breast cancer patients with the overall control group (patients with benign breast conditions and healthy donors). (B) ROC curve analysis comparing breast cancer patients with patients with benign breast conditions. (C) ROC curve analysis comparing breast cancer patients with healthy donors. (D) Precision-recall curve analysis comparing breast cancer patients with the overall control group (patients with benign breast conditions and healthy donors). (E) Box and whisker plot showing the miRNA signature results of breast cancer patients compared to the overall control group (patients with benign breast conditions and healthy donors). The subjects are shown on the x-axis, while the miRNA signature values are shown on the y-axis. $P<0.0001$ is indicated by ****.

Remarkably, this miRNA signature exhibited a clinical sensitivity of 85.83% (95% CI: 78.3% to 91.5%) and a clinical specificity of 84.62% (95% CI: 75.5% to 91.5%) at a cut-off threshold of 0.508 from the overall analysis (Table 5). Moreover, the AUC analysis revealed a robust value of 0.908 (95% CI: 0.861 to 0.943), underscoring the diagnostic potential of this miRNA signature (Fig. 15A). As shown in Table 5, the positive likelihood ratio was calculated to be 5.58 (95% CI: 3.43 to 9.09), meaning that individuals with a 'positive' result are about 5.58 times more likely to have the condition compared to those with a negative result. Additionally, the negative likelihood ratio was calculated as 0.17 (95% CI: 0.11 to 0.27), suggesting that individuals with a negative test result are approximately 0.17 times as likely to have the condition compared to those with a positive test result. These likelihood ratios are crucial for understanding the diagnostic precision and usefulness of the test. In the analysis of the subgroup within the control group, each clinical specificity for patients with benign breast disease and healthy donors was calculated at the same cut-off threshold of 0.508. The clinical specificity for patients with benign breast disease was 89.13% with a 95% CI of 76.4% to 96.4% (Fig 15B), while for healthy donors, it was 80.00% with a 95% CI of 65.4% to 90.4% (Fig. 15C). Given the potential bias in interpreting clinical performance indicators evaluated in the controlled population, the AUPRC value was calculated using positive predictive value and clinical sensitivity. As a result, the AUPRC was determined as 0.932 (95% CI: 0.871 to 0.966), and the $F1_{max}$ value was determined as 0.869 (Fig. 15D). The miRNA signature detected in BEVs from BC patients was significantly upregulated compared to those from non-BC controls (Fig. 15E). Crucially, no significant difference was found between patients with benign breast disease and normal controls (Fig. 15E). This suggests that the miRNA signature in BEVs could serve as a potential biomarker for distinguishing breast cancer patients from individuals with benign breast disease or normal breast tissue.

Table 5. Results of miRNA signature clinical performance parameter.

		Breast cancer		Total
		Positive	Negative	
miRNA signature	Elevated	103	14	117
	Not elevated	17	77	94
	Total	120	91	211
Clinical Performance Parameters		**N=211**	**95% CI**	
Clinical Sensitivity		85.83%	78.3 - 91.5	
Clinical Specificity		84.62%	75.5 - 91.3	
Likelihood Ratio Negative (LRN)		0.17	0.11 -0.27	
Likelihood Ratio Positive (LRP)		5.58	3.43 - 9.09	
Area Under Curve (AUC)		0.908	0.861 - 0.943	

Considering the intended use of breast screening, it is essential to ensure high clinical specificity to control false positives. Conversely, for aiding in breast diagnosis purposes, high clinical sensitivity is crucial to control false negatives and ensure reliability for true positives. The threshold of this miRNA signature was adjusted accordingly and compared (Table 6). Analysis revealed that when clinical specificity was fixed at 99%, the clinical sensitivity was 35.00% (95% CI: 26.5% to 44.2%), and when clinical sensitivity was fixed at 99%, the clinical specificity was 10.99% (95% CI: 1.10% to 47.25%). Considering that the recall rate of mammography is 5-12% (76–78), when clinical specificity was fixed at 95%, clinical sensitivity was found to be 63.33% (95% CI: 40.00 to 81.67%).

Table 6. Clinical performance when sensitivity and specificity are fixed at 99%.

Estimated specificity at fixed sensitivity			
Sensitivity	Specificity	95% CI	Criterion
80	85.71	71.43 - 92.31	>0.549
90	75.82	56.04 - 85.71	>0.422
95	57.14	23.08 - 80.22	>0.275
97.5	46.15	7.69 - 75.82	>0.202
99	10.99	1.10 - 47.25	>0.058
Estimated sensitivity at fixed specificity			
Specificity	Sensitivity	95% CI	Criterion
80	86.67	73.32 - 93.33	>0.448
90	77.50	63.33 - 90.00	>0.595

95	63.33	40.00 - 81.67	>0.743
97.5	43.33	34.3 - 52.7	>0.935
99	**35.00**	**26.5 - 44.2**	**>0.971**

3.6 Clinical feasibility analysis

The clinical information of the subjects participating in this study was retrospectively collected through Electronic Medical Records (EMR), and correlation analysis with the miRNA signature was conducted. This analysis sought to assess the stability of the miRNA signature as a biomarker by investigating its correlation with different clinical parameters. If the observed increase in the level of the miRNA signature in breast cancer patients could be attributed to other clinical factors, it would not be appropriate to consider it as a specific biomarker for breast cancer. Therefore, we systematically compared and analyzed the expression levels of the miRNA signature in relation to the collected clinical information. The data collected from breast cancer patients included results from mammography and breast ultrasound reporting results (BI-RADS), breast density, pathological results, TNM stage classification, molecular subtyping results, Ki-67 expression levels, information on recurrence, survival status, CA15-3 and CEA biomarker results that were measured before surgery (Table 7). Similarly, data from patients with benign breast diseases encompassed mammography and breast ultrasound reporting results, breast density, and pathology. Analysis was conducted using these collected datasets, and subjects with incomplete or unverifiable data were excluded from the analysis to ensure the reliability of the results (Table 8).

Table 7. Patient clinical details concerning breast cancer

Breast cancer patients

Variables	N (%)	miR signature	95% CI	*P* value
BI-RADS , Mammography				
0	27 (22.50)	0.833	0.730 - 0.935	
1	4 (3.33)	0.926	0.818 - 1.033	
2	6 (5.00)	0.694	0.353 - 1.034	
3	18 (15.00)	0.798	0.688 - 0.908	0.200
4	32 (26.67)	0.695	0.601 - 0.789	
5	28 (23.33)	0.850	0.784 - 0.917	
N/A	5 (4.17)	0.623	0.108 - 1.138	

BI-RADS , Ultrasonography				
0				
1				
2				
3	1 (0.83)	0.443	-	
4	77 (64.17)	0.732	0.671 - 0.793	0.003
5	42 (35.00)	0.883	0.830 - 0.936	
N/A	-	-	-	
Breast Density Grade				
A	4 (3.33)	0.65	0.246 - 1.053	
B	14 (11.67)	0.864	0.742 - 0.986	
C	84 (70.00)	0.790	0.738 - 0.843	0.241
D	14 (11.67)	0.762	0.622 - 0.903	
N/A	4 (3.33)	0.596	0.546 - 0.646	
Pathology				
DCIS	5 (4.17)	0.597	0.087 - 1.110	
IDC	102 (85.00)	0.792	0.744 - 0.840	
ILC	5 (4.17)	0.741	0.539 - 0.942	0.451
Mucinous	3 (2.50)	0.899	0.469 - 1.330	
Metaplastic	4 (3.33)	0.743	0.211 - 1.280	
Phyllodes tumor, malignant	1 (0.83)	0.715	-	
TNM Stage				
0	5 (4.17)	0.597	0.087 - 1.110	
I	46 (38.33)	0.774	0.699 - 0.849	
II	57 (47.50)	0.811	0.749 - 0.873	0.463
III	10 (8.33)	0.76	0.584 - 0.937	
IV	2 (1.67)	0.727	-1.530 - 2.980	
T Stage				
Tis	5 (4.17)	0.597	0.087 - 1.11	
T1	64 (53.33)	0.793	0.730 - 0.856	0.297
T2	44 (36.67)	0.788	0.717 - 0.858	
T3	5 (4.17)	0.696	0.470 - 0.921	

T4	2 (1.67)	0.993	0.955 - 1.03	
N Stage				
N0	83 (69.17)	0.769	0.714 - 0.825	
N1	26 (21.67)	0.834	0.738 - 0.931	0.215
N2	3 (2.50)	0.971	0.857 - 1.086	
N3	8 (6.67)	0.678	0.473 -0.883	
M Stage				
M0	118 (98.33)	0.783	0.738 - 0.829	0.751
M1	2 (1.67)	0.727	-1.529 - 2.982	
Molecular subtype				
Luminal A	21 (17.50)	0.87	0.770 - 0.970	
Luminal B	21 (17.50)	0.793	0.700 - 0.886	
HER-2/neu	20 (16.67)	0.838	0.740 - 0.936	0.037
TNBC	53 (44.17)	0.739	0.667 - 0.811	
N/A	5 (4.17)	0.607	0.083 - 1.132	
Ki67				
<15%	33 (27.50)	0.806	0.724 - 0.888	
≥ 15%	77 (64.17)	0.792	0.738 - 0.846	0.385
N/A	10 (8.33)	0.632	0.375 - 0.889	
Recurrence				
Yes	44 (36.67)	0.841	0.779 - 0.903	0.234
No	76 (63.33)	0.748	0.687 - 0.809	
Survival				
Death	19 (15.83)	0.842	0.743 - 0.941	0.356
Survive	101 (84.17)	0.771	0.721 - 0.822	

Table 8. Clinical information of patients with benign breast disease.

Benign breast disease patients

Variables	N (%)	miR signature	95% CI	*P* value
Mammography (BI-RADS)				
0	14 (30.43)	0.232	0.097 - 0.366	0.351

1	3 (6.52)	0.240	-0.003 - 0.484	
2	5 (10.87)	0.288	0.126 - 0.450	
3	4 (8.70)	0.427	0.191 - 0.664	
4	6 (13.04)	0.272	0.084 - 0.461	
5	-	-	-	
N/A	14 (30.43)	0.311	0.178 - 0.444	
Ultrasonography (BI-RADS)				
0				
1				
2				
3	12 (26.09)	0.296	0.150 - 0.443	
4	29 (63.04)	0.289	0.215 - 0.363	0.560
5	1 (2.17)	0.079	-	
N/A	4 (8.70)	0.249	0.085 - 0.413	
Breast Density Grade				
A				
B				
C	31 (67.39)	0.275	0.200 - 0.350	0.822
D	8 (17.39)	0.270	0.096 - 0.445	
N/A	7 (15.22)	0.344	0.147 - 0.542	
Pathology				
Atypical hyperplasia	4 (8.70)	0.466	0.305 - 0.626	
Fibroadenoma	10 (21.74)	0.29	0.135 - 0.445	
Intraductal papilloma	15 (32.61)	0.237	0.126 - 0.348	
Phyllodes tumor	9 (19.57)	0.331	0.141 - 0.521	
Sclerosing adenosis	3 (4.35)	0.162	-0.011 - 0.335	0.161
Usual ductal hyperplasia	2 (4.35)	0.273	-0.255 - 0.800	
Cystic and papillary apocrine metaplasia	1 (2.17)	0.492	-	
Lymphoid follicular hyperplasia	1 (2.17)	0.136	-	
Ulcer	1 (2.17)	0.246	-	

3.6.1 Correlation with demographic information (Age)

According to demographic information analysis, the average age of breast cancer patients was determined to be 59.6 ± 11.4 years, while the average age of control group (patients with benign breast diseases, and health donors) was 42.22 ± 10.15 and 48.8 ± 9.08 years (Table 9). Remarkably, there was a statistically significant age difference between the breast cancer patients and the control group (P value < 0.0001). Given the statistically significant results observed between the ages of the test and control groups, it was imperative to ascertain the stability of miRNA signature expression with age.

Table 9. Demographic information of participants.

	BC (N=120)	**Benign (N=46)**	**Healthy donor (N=45)**
Age			
Mean±SD	59.6±11.4	42.22±10.15	48.8±9.08
95% CI	57.54 to 61.66	39.20 to 45.23	46.07 to 51.53
Median	58	44.5	46
Min/Max	26 to 85	23 to 65	33 to 69
25th-75th perc.	51 to 68	34.75 to 48.00	42 to 56.5
P-value	<0.0001	<0.05	<0.0001

Therefore, linear regression analysis was conducted on the miRNA signatures of each group (Fig. 16). Furthermore, the correlation between clinical information and miRNA signature expression was examined to elucidate the relationship between disease and biomarkers. Suppose the miRNA signature demonstrates significant changes due to specific clinical information. In that case, it may not conclusively indicate that the miRNA signature has increased solely due to the occurrence of the disease. Hence, the stability of the miRNA signature as a diagnostic algorithm for breast cancer was evaluated by comparing it with clinical information. The clinical information subject to analysis encompassed the comparison of miRNA signature expression based on age, pathological results, and molecular biological subtype.

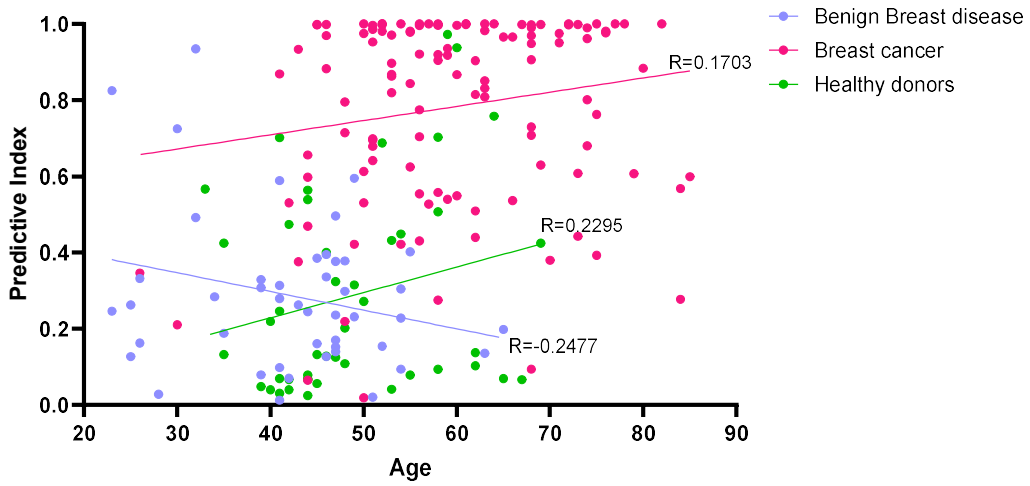

Figure 16. Correlation of age with miRNA signature. The Pearson correlation coefficient (R) for the miRNA signature in relation to age was indicated.

To assess the relationship between age and miRNA signature expression, correlation analysis was performed on a cohort consisting of 120 breast cancer patients, 46 patients with benign breast diseases, and 45 healthy donors. The Pearson correlation coefficient (R) for the miRNA signature in relation to age was calculated as 0.1703 in breast cancer patients, 0.2295 in patients with benign breast diseases, and -0.2477 in healthy donors (Fig. 16). Furthermore, to explore potential age-related trends in miRNA signature expression, the subjects were subdivided into age groups: those in their 40s or younger, 50s, 60s, 70s, and 80s (Table 10). However, the analysis did not reveal statistically significant differences in miRNA signature values across these age groups (each P-value: breast cancer patients, 0.082; patients with benign breast diseases, 0.3; healthy donors, 0.142). These findings suggest that the expression of the miRNA signature remained consistent across different age groups within each subject group.

Table 10. Comparison of miRNA expression in subjects by age group.

Age

Breast cancer patients	**N (%)**	**miR signature**	**95% CI**	***P* value**
≤40's	19 (15.83)	0.634	0.488 - 0.781	
50's	46 (38.33)	0.804	0.735 - 0.873	
60's	30 (25.00)	0.833	0.751 - 0.916	0.082
70's	20 (16.67)	0.825	0.720 - 0.931	
80's	5 (4.17)	0.666	0.312 - 1.019	
Patients with benign breast disease	**N (%)**	**miR signature**	**95% CI**	***P* value**
≤40's	38 (82.61)	0.307	0.238 - 0.376	
50's	6 (13.04)	0.200	0.053 - 0.347	0.300
60's	2 (4.35)	0.167	-0.227 - 0.561	
70's	-	-	-	
80's	-	-	-	
Healthy donors	**N (%)**	**miR signature**	**95% CI**	***P* value**
≤40's	28 (62.22)	0.221	0.143 - 0.299	
50's	10 (22.22)	0.424	0.203 - 0.644	0.142
60's	7 (15.56)	0.357	0.022 - 0.691	
70's	-	-	-	
80's	-	-	-	

3.6.2 Correlation with recurrence and survival

To determine the miRNA signature's potential as a prognostic factor, we conducted a comparative analysis of recurrence and survival outcomes in accordance with the expression of the miRNA signature.

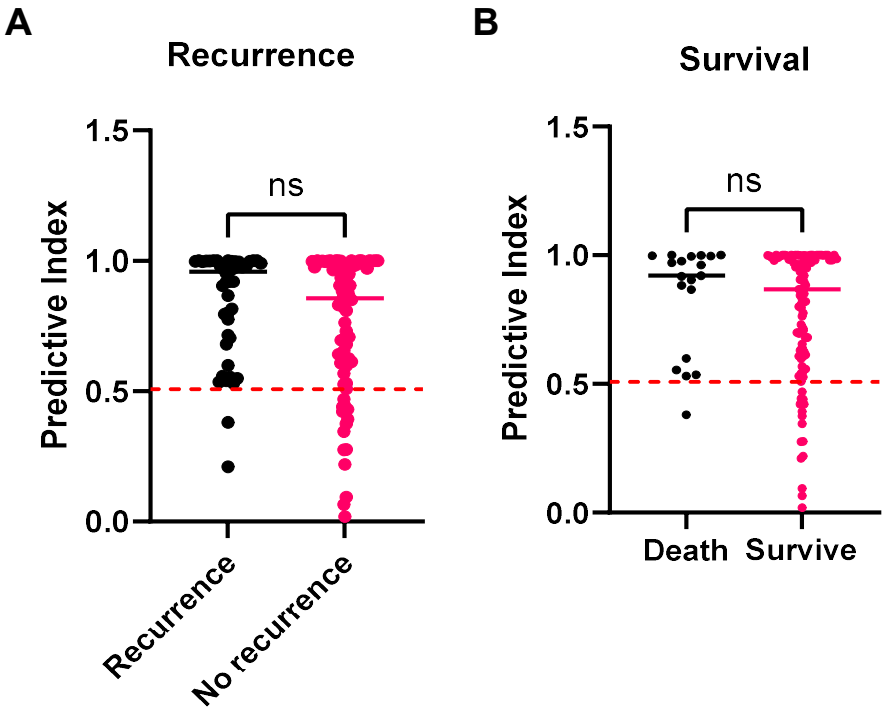

Figure 17. Correlation of recurrence and survival with miRNA signature. (A) Graph comparing the expression of miRNA signature based on recurrence status. (B) Graph comparing the expression of miRNA signature based on survival status. "ns" indicates statistically non-significant results.

Among 120 breast cancer patients, 44 (36.67%) experienced recurrence, while 76 (63.33%) did not (Table 7). Upon analysis, the average expression of the miRNA signature was found to be 0.748 (95% CI: 0.687 to 0.809) in patients without recurrence and 0.841 (95% CI: 0.779 to 0.903)

in patients with recurrence (Table 7). However, this difference was not statistically significant (P = 0.234, Fig. 17A). Similarly, at the time of data collection, 19 (15.83%) of the 120 breast cancer patients had deceased, while 101 (84.17%) were alive (Table 7). Analysis revealed that the average expression of the miRNA signature in surviving patients was 0.771 (95% CI: 0.721 to 0.822), whereas in deceased patients, it was 0.842 (95% CI: 0.743 to 0.941) (Table 7). Again, this difference was not statistically significant (P = 0.356, Fig. 17B). These findings clearly show that there was no statistically significant correlation between survival or recurrence and the miRNA signature. Therefore, its performance as a prognostic factor was not conclusively confirmed in this analysis. Further studies with larger sample sizes and longitudinal follow-up are necessary to elucidate the prognostic potential of the miRNA signature in breast cancer patients.

3.6.3 *Clinical sensitivity of miRNA signature for early breast cancer*

A comparison of clinical sensitivity was conducted to explore the potential correlation between the expression of different groups of miRNA candidates in BEVs and the miRNA signature of breast cancer TNM stages, aiming to assess their role in aiding early diagnosis. Given that the control group was repeatedly applied to the analysis for each stage, only clinical sensitivity was analyzed to evaluate the clinical effectiveness for early breast cancer patients.

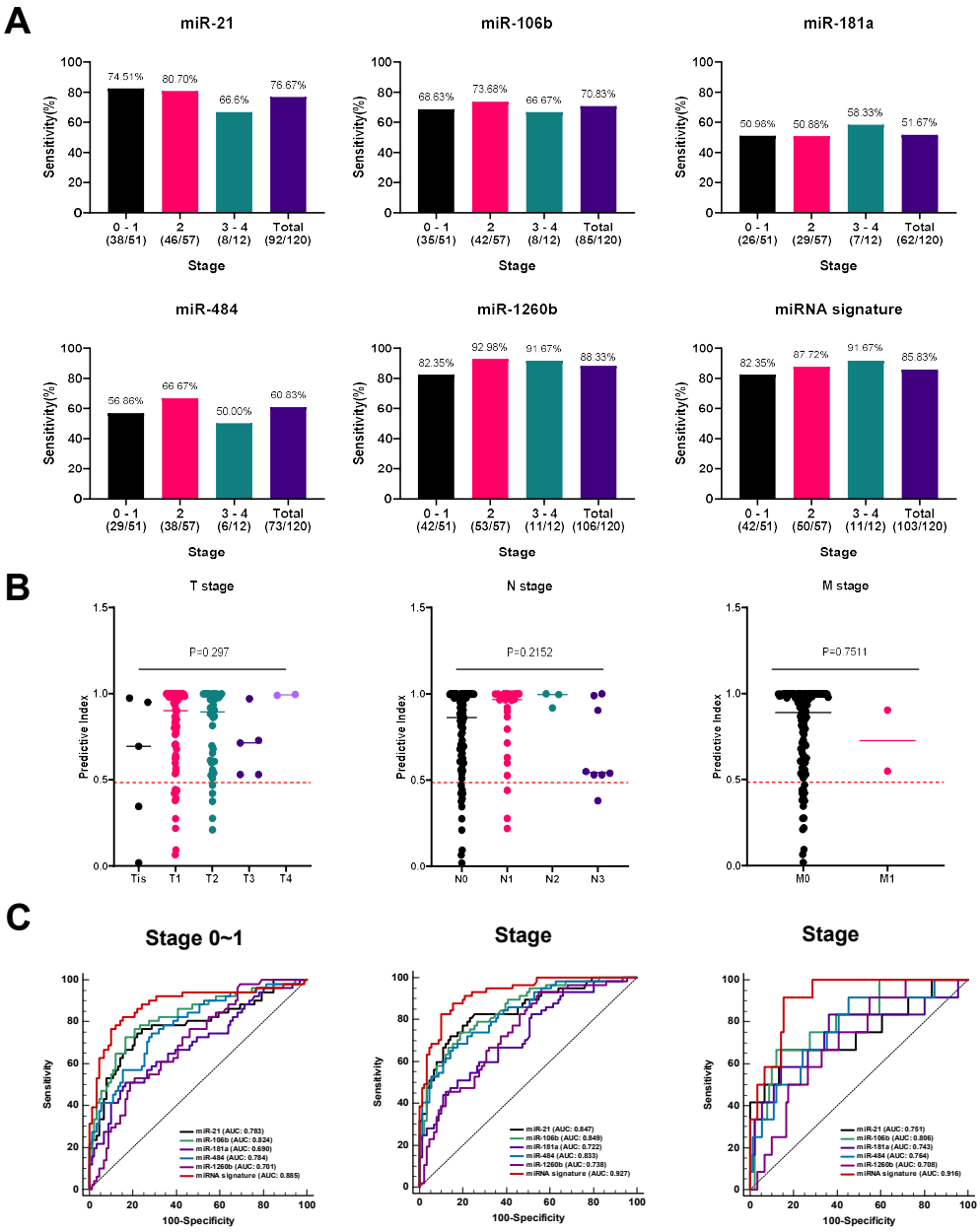

Figure 18. Correlation between miRNA signature and pathological diagnosis results. (A) Comparison of clincial sensitivity of each miRNA candidate group and miRNA signature from TNM stage (B) Correlation of among T stage, N stage, M stage with miRNA signature (c) Comparison of ROC curves for each miRNA candidate and miRNA signature according to breast cancer TNM stage

Among the analyzed cases, 51 were diagnosed with early breast cancer, with 5 cases (4.17%) classified as stage 0 and 46 cases (38.33%) as stage 1 (Table 7). Additionally, 57 cases (47.50%) were classified as stage 2 breast cancer, and 12 cases (10%) were considered stage 3 or 4 (8.33% in stage 3, 1.67% in stage 4, Table 7). Evaluation of the miRNA signature's sensitivity according to each breast cancer stage revealed 82.35% sensitivity in early breast cancer, 87.72% in stage 2, and 91.67% in stage 3 or 4, indicating high clinical sensitivity for breast cancer (Fig. 18A). When considering breast cancer stage 1 or lower, the clinical sensitivities of microRNA-21, -106b, -181a, -484, and -1260b were 74.51%, 68.63%, 50.98%, 56.86%, and 82.35%, respectively (Fig. 18A). The area under the curve (AUC) for the miRNA signature at each breast cancer stage was evaluated, resulting in 0.885 for stages 0 to 1, 0.927 for stage 2, and 0.916 for stages 3 to 4 (Fig. 18C). In early breast cancer stages 0 to 1, the AUC values of miRNAs were 0.783, 0.824, 0.690, 0.784, and 0.701 for miR-21, miR-106b, miR-181a, miR-484, and miR-1260b, respectively (Fig. 18C). The AUC values of miRNA in stage 2 breast cancer were 0.847, 0.849, 0.722, 0.833, and 0.738, and in stage 3-4 breast cancer, they were 0.751, 0.806, 0.743, 0.764, and 0.708 (Fig. 18C). The AUC of miR-1260b for stage 1 or lower breast cancer was 0.701, while the AUC of the miRNA signature was 0.885, showing a statistically significant difference ($P<0.0001$). Although miR-1260b demonstrated similar clinical sensitivity as the miRNA signature, the evaluation of AUC values combined with specificity demonstrated the diagnostic effectiveness of the miRNA combination for early-stage breast cancer. However, it's essential to acknowledge that the number of patients with stages 0 and 3 or higher is relatively small, which may introduce bias. Therefore, it's prudent to exercise caution when claiming the validity of breast cancer early detection performance based solely on the analyzed clinical sensitivity. Additionally, the lack of statistically significant differences in the miRNA signature across TNM stages 0 to 4 suggests consistent expression regardless of tumor size, lymph node metastasis, or distant metastasis ($P=0.297$, 0.2152, 0.7511, Fig. 18B). Large-scale cohort studies are needed to further verify clinical effectiveness.

3.6.4 Correlation with histopathological results of breast cancer and benign breast disease

We conducted a comparative analysis of the expression patterns of miRNA signatures according to histopathological findings in breast cancer patients and patients with benign breast diseases. Among the subjects of this clinical performance test, the pathological distribution of the 120 breast cancer patients revealed that 107 individuals (89.17%) were diagnosed with invasive carcinoma, while 5 individuals (4.17%) had carcinoma in situ (Table 7). Thus, the distribution among breast cancer patients recruited for this trial mirrored the crude incidence rate of patients with invasive cancer. Notably, regardless of the pathological results among the 120 breast cancer patients, the miRNA signature demonstrated consistent expression ($P = 0.451$, Fig. 19A). According to the

breast cancer statistics in South Korea(79) , breast cancer is predominantly classified into invasive cancer and carcinoma in situ. The crude incidence rate of all breast cancer patients in 2019 was analyzed as 115.6, with the crude incidence rate of invasive carcinoma reported as 96.5, and the crude incidence rate of carcinoma in situ as 19.1. Consequently, among all breast cancer patients, approximately 83.5% were diagnosed with invasive cancer, while 16.5% were diagnosed with carcinoma in situ. It was confirmed that a somewhat smaller number of patients with carcinoma in situ were recruited than the breast cancer statistics, but the number of invasive carcinoma patients was similar to the breast cancer statistics.

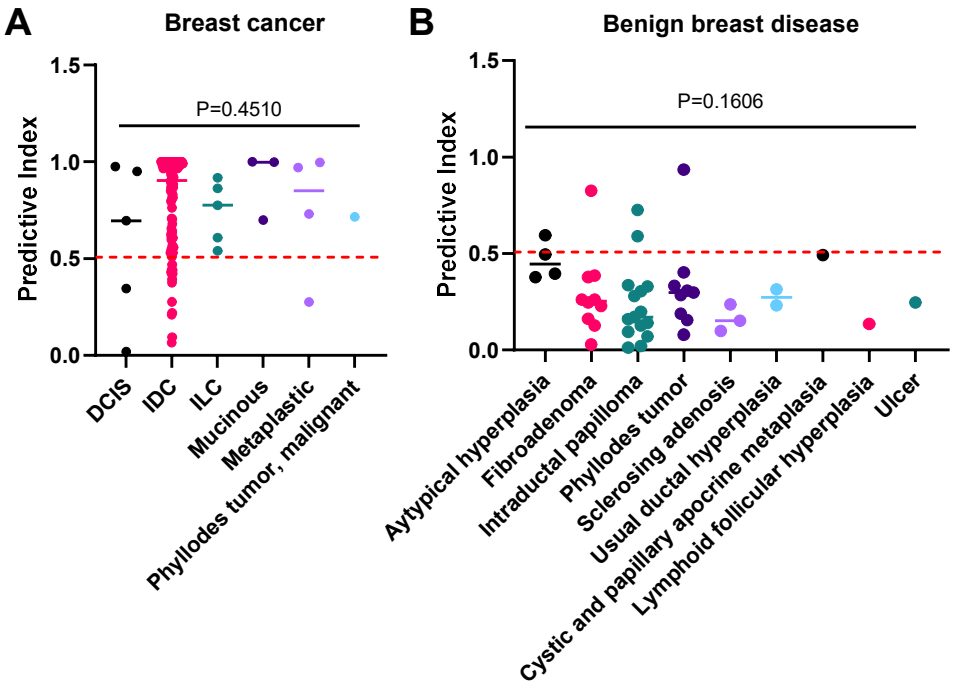

Figure 19. Correlation of histopathological classification with miRNA. (A) Distribution of miRNA signatures in plasma of breast cancer patients DCIS, Ductal carcinoma in situ; IDC, Invasive ductal carcinoma; ILC, Invasive lobular carcinoma; Mucinous, Mucinous carcinoma; Metaplastic, Metaplastic carcinoma. (B) Distribution of miRNA signatures in plasma of patients with benign breast diseases by histopathological classification

In Table 7, the histopathological classification of breast cancer patients revealed 5 cases of DCIS (4.17%), 102 cases of IDC (85.00%), 5 cases of ILC (4.17%), 3 cases of Mucinous (2.50%), 4 cases of Metaplastic (3.33%), and 1 case of Phyllodes tumor (0.83%) The average miRNA

signature for each histopathological subtype ranged from 0.597 to 0.899, with no statistically significant difference observed (P = 0.451). Patients with benign breast diseases included various pathologies such as atypical hyperplasia, fibroadenoma, intraductal papilloma, phyllodes tumor, sclerosing adenosis, ductal hyperplasia, cystic and papillary apocrine metaplasia, lymphoid follicular hyperplasia, and ulcer (Table 8). The miRNA signature for each pathology result did not exhibit statistical significance (P= 0.161, Fig. 19B).

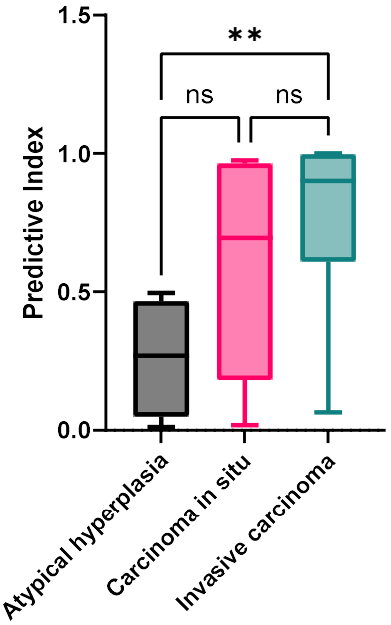

Figure 20. Comparison of miRNA signatures at different stages of breast cancer development. The number of cases at each stage of breast cancer development were as follows: Atypical hyperplasia (n=4), Carcinoma in situ (n=5), and Invasive carcinoma (n=107). "ns" indicates statistically non-significant results, "**" denotes p<0.01.

We further analyzed the miRNA signatures according to the pathological stage, focusing on atypical hyperplasia, a precursor to breast cancer. The miRNA signature for atypical hyperplasia was 0.466 (95% CI: 0.305 to 0.626), compared to 0.597 (95% CI: 0.087 to 1.110) for carcinoma in situ and 0.792 (95% CI: 0.744 to 0.840) for invasive carcinoma (Table 7, 8). A statistically significant difference was observed between atypical hyperplasia and invasive carcinoma (P < 0.01, Fig. 20). However, the imbalance in the number of subjects between the different pathological stages

necessitates a cautious interpretation of the results. Future studies will focus on securing a statistically significant number of samples evenly distributed across pathological stages to validate the association between miRNA signature expression and breast cancer development stages.

3.6.5 Correlation with molecular subtype and Ki-67

In analyzing the correlation between molecular subtype and miRNA signature, statistically significant differences were observed only in luminal A and triple-negative breast cancer (TNBC) types ($P < 0.05$, Fig. 21A). However, among the 120 breast cancer patients, 53 (44.17%) were diagnosed with TNBC, 21 (17.50%) with Luminal A, 21 (17.50%) with Luminal B, and 20 (16.67%) with HER2-enriched breast cancer (Table 7). Given the potential for over-bias due to differences in recruitment rates between groups, it is inconclusive to assert that the miRNA signature is statistically significantly under-expressed in TNBC based solely on these results (Fig. 21A). In addition, as a result of correlation analysis between the proliferation factor ki-67 and the miRNA signature, there was no statistically significant difference ($P= 0.385$, Fig. 21B).

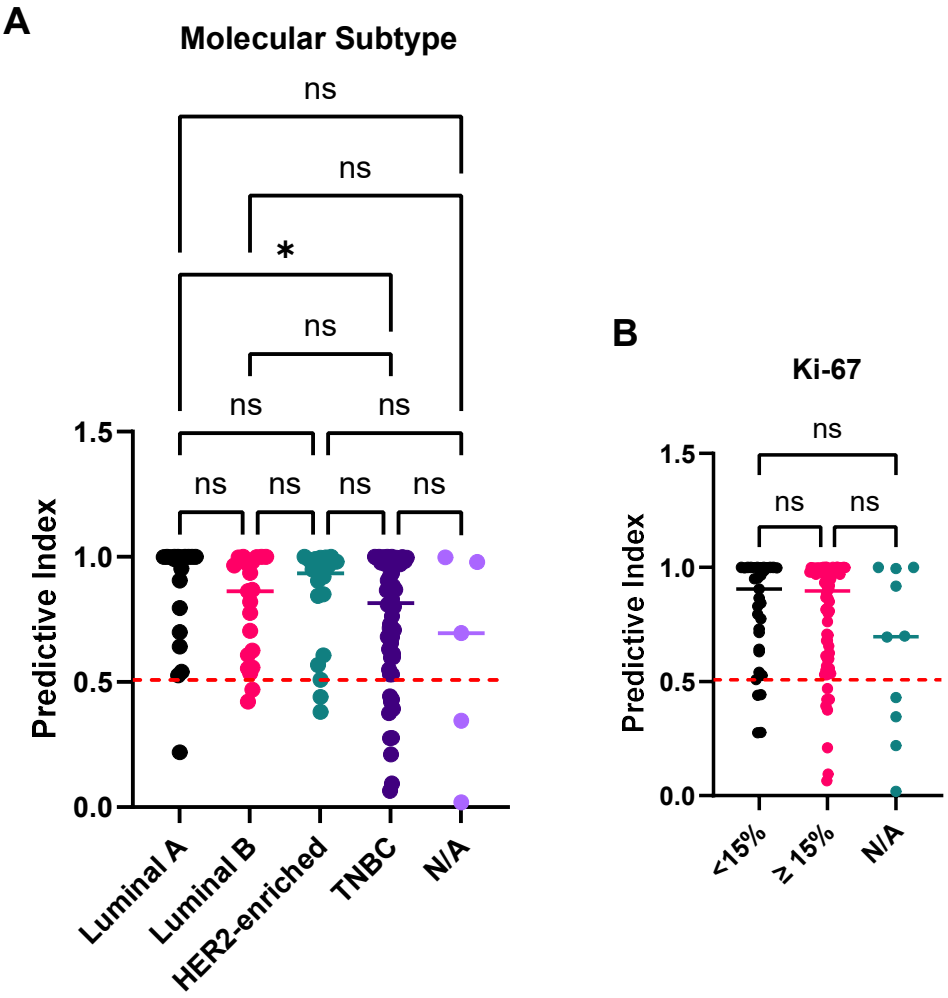

Figure 21. Correlation of molecular subtype with miRNA signature. (A) Molecular subtyping result with miRNA Signature. (B) Ki-67 expression with miRNA Signature. <15% is defined as a high proliferation rate, and≥15 % is defined as vice versa. "ns" denotes statistically non-significant results, "*" denotes $p<0.05$.

3.6.6 *Clinical performance of miRNA signature with mammography.*

This study examined the diagnostic utility of mammography alongside miRNA combination analysis for the diagnosis of breast cancer. Mammography, the primary method for breast cancer screening, often faces limitations in sensitivity, especially in dense breast tissue. This necessitates further tests or risks missing additional diagnostics due to false negatives. To address these challenges, we sought to determine if miRNA signature analysis could complement mammography findings. We categorized mammography results using the BI-RADS classification system, focusing on positive screening outcomes. BI-RADS category 0, indicating inconclusive findings requiring additional testing, was deemed 'Negative' for our study. To summarize, in our study, we defined 'positive screening' as classifying only BI-RADS categories that recommend tissue biopsy. If a BI-RADS category is determined as 0, indicating an 'Incomplete'. Since additional imaging rather than biopsy is recommended as an additional test in this case, it was not determined to be suspected of breast cancer at that time. Therefore, BI-RADS 0 was included in the 'negative' category for our analysis. Mammography results receiving BI-RADS assessments of 4a, 4b, 4c, or 5 were considered "positive," indicating suspicious or seen abnormalities. Conversely, assessments of 0, 1, 2, or 3 were deemed "negative," suggesting incomplete or findings not suspicious of malignancy. Our analysis centered on individual participants, with mammography results serving as the primary endpoint for comparison with miRNA signature clinical performance. This approach enabled us to evaluate the potential of miRNA signature analysis in augmenting mammography's diagnostic accuracy for breast cancer.

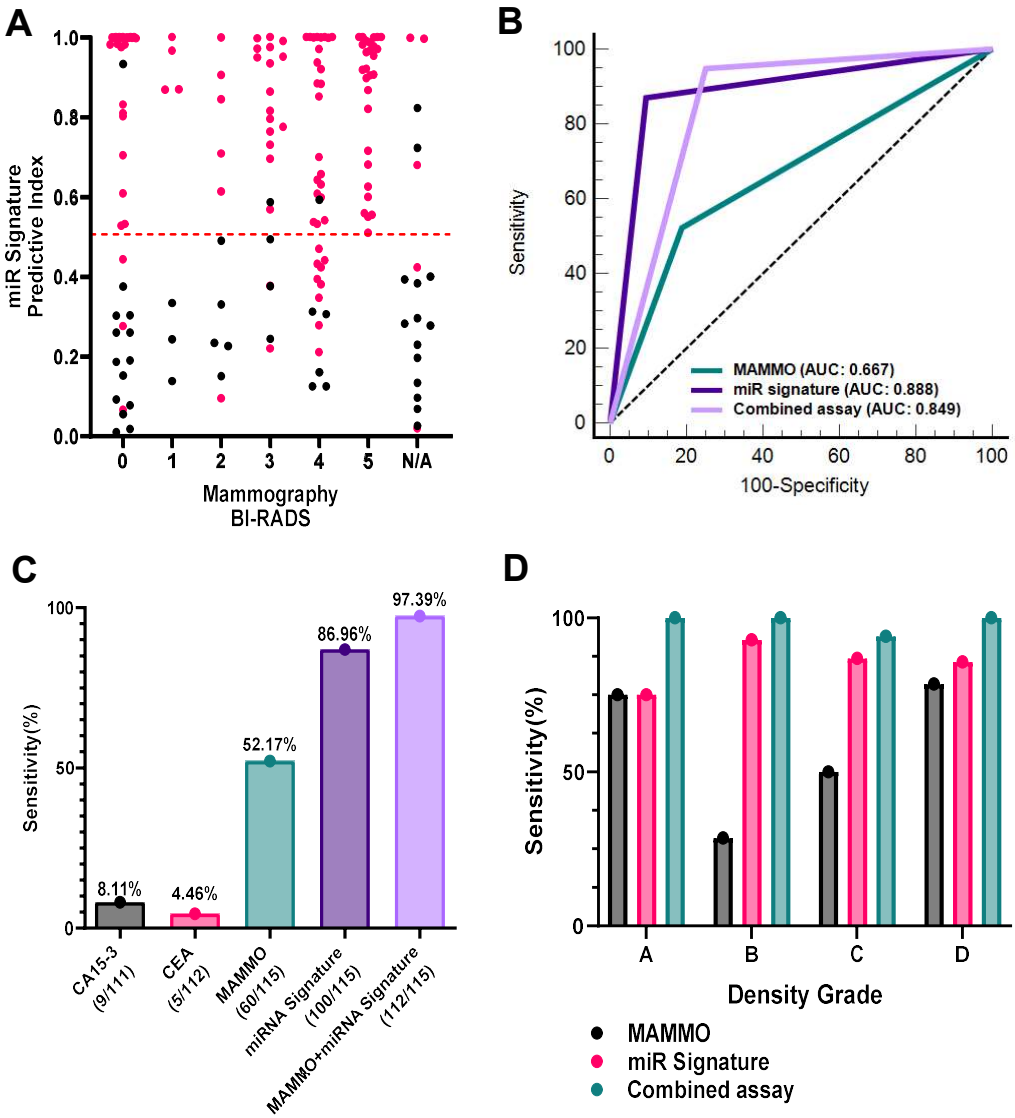

Figure 22. Comparison of clinical performance between miRNA signature and mammography results. (A) The dot plot illustrates the analysis of mammography results alongside the criteria for determining the miRNA Signature. The horizontal dotted line represents the threshold for determining the miRNA Signature. Mammography results were categorized as positive when classified as BI-RADS 4 or 5. (B) ROC curve comparing the diagnostic performance of mammogram, miRNA combnation, and miRNA combination plus mammography. For determining the performance of miRNA + mammography, a positive result was considered if either result was positive. (C) Box and whisker plot comparing the clinical sensitivity of CA15-3, CEA,

mammography, miRNA signature, and mammography + miRNA signature in 82 breast cancer patients with confirmed mammography results. Four patients with unconfirmed CA15-3 results and three patients with unconfirmed CEA results were excluded from the analysis. (D) Graph comparing the clinical sensitivity based on breast density in 116 breast cancer patients. The x-axis represents breast density, with "Fatty" indicating Grade A, "Scattered" indicating Grade B, "Heterogeneous dense" indicating Grade C, and "Extremely dense" indicating Grade D. The x-axis represents clinical sensitivity.

To ascertain the potential mutual benefit of miRNA combination analysis for the cases of BI-RADS 0, we assessed the recovery rates. A total of firty participants were assigned a BI-RADS 0 classification, consisting of 27 BC patients and 14 benign patients (Fig. 22A). By analyzing the microRNA combination, 24 breast cancer patients (88.89%) were correctly identified as true positives, and 13 cases with benign tumors (92.86%) were classified as true negatives (Table 11). Following these results, even among confirmed breast cancer patients, the miRNA signature reassessed these subjects based on mammography results, indicating mammography's potential utility for subjects diagnosed with BI-RADS 0.

Table 11. Determination results of miRNA signatures for each subject according to mammography results (BI-RADS).

	Breast cancer patients			**Patients with benign breast disease**		
		miRNA Signature expression			**miRNA Signature expression**	
Mammography (BI-RADS)	**N (%)**	**Low, N (%)**	**High, N (%)**	**N (%)**	**Low, N (%)**	**High, N (%)**
0	27 (22.50)	**3 (11.11)**	**24 (88.89)**	14 (30.43)	**13 (92.86)**	**1 (7.14)**
1	4 (3.33)	0 (0.00)	4 (100.00)	3 (6.52)	3 (100.00)	0 (0.00)
2	6 (5.00)	1 (16.67)	5 (83.33)	5 (10.87)	5 (100.00)	0 (0.00)
3	18 (15.00)	2 (11.11)	16 (88.89)	4 (8.70)	3 (75.00)	1 (25.00)
4	32 (26.67)	9 (28.13)	23 (71.88)	6 (13.04)	5 (83.33)	1 (16.67)
5	28 (23.33)	0 (0.00)	28 (100.00)	-		
NA	5 (4.17)	2 (40.00)	3 (60.00)	14 (30.43)	12 (85.71)	2 (14.29)

Furthermore, the diagnostic performance of mammogram, miRNA combination, and combined assay was evaluated in 115 BC patients and 32 benign patients. ‘BI-RADS 4 and 5’ were classified

as 'positive' for mammography, and in the integrated assay, a 'positive' result from either test was interpreted as an 'overall positive'. Results revealed an AUC value of 0.667 (95% confidence interval: 0.585 to 0.743) for mammography, 0.888 (95% confidence interval: 0.825 to 0.934) for the miRNA signature, and 0.849 (95% confidence interval: 0.781 to 0.903) for the combined assay (Fig. 22B). Figure 22C presents a comprehensive examination of diagnostic sensitivity, including CEA and CA15-3. CEA and CA15-3 demonstrated very low sensitivities, both under 10%, for breast cancer patients, while mammography showed a sensitivity of 52.17% (Fig. 22C). However, the miRNA combination showed a significantly higher sensitivity of 86.96%, and the integrated analysis reached 97.39%, affirming its diagnostic accuracy for breast cancer (Fig. 22C). Additionally, our study investigated the relationship between the miRNA signature and dense breast tissue. A retrospective analysis of breast density data from 120 breast cancer patients was conducted to evaluate the diagnostic sensitivity of mammography, miRNA combinations, and integrated assays in dense breast tissue. Breast density was confirmed in 116 out of 120 breast cancer patients. Subjects with BI-RADS 6 or no information based on mammography findings were excluded from the analysis. Accordingly, 116 breast cancer patients were compared (Fig. 22D). Our results underscore significant differences in the sensitivity of mammography, miRNA combinations, and integrated analyses across varying breast density grades. In dense breast grades C and D, the sensitivities were analyzed as 54.08%, 86.73%, and 94.90%, respectively (Table 12). These findings suggest the potential for mitigating the low sensitivity of mammography attributed to dense breast tissue through the complementary use of miRNA signature analysis and combined assay.

Table 12. Clinical sensitivity of mammography, miRNA signature, and combined assay according to breast density.

Breast Density	N	miR Signature	*P* value	Sensitivity (%)		
				MAMMO	miR Signature	Combined assay
Each grade						
A (Fatty)	4	0.650±0.254	0.382	75.00 (3/4)	75.00 (3/4)	100.00 (4/4)
B (Scattered)	14	0.864±0.211		28.57 (4/14)	92.86 (13/14)	100.00 (14/14)
C (Heterogeneously dense)	84	0.790±0.244		50.00 (42/84)	86.90 (73/84)	94.05 (79/84)
D (Extremely dense)	14	0.762±0.243		78.57 (11/14)	85.71 (12/14)	100.00(14/14)
Non-dense breast vs Dense breast						
A, B	18	0.816±0.232	0.580	38.89 (7/18)	88.89 (16/18)	100.00 (18/18)
C, D	98	0.786±0.242		54.08 (53/98)	86.73 (85/98)	94.90 (93/98)
Total	116	0.791±0.240		51.72 (60/116)	87.07 (101/116)	95.69111/116)

3.6.7 miRNA signature with ultrasound

The retrospective analysis of breast ultrasound data from 42 patients with confirmed benign breast disease aimed to assess the potential of miRNA signature analysis in reducing unnecessary tissue biopsies. The total number of participants with benign breast disease was 46, but information on breast ultrasound was missing for 4 participants.

Table 13. Breast ultrasound results for patients with benign breast tumors who underwent tissue biopsy.

Ultrasound (BI-RADS)	N (%)	miRNA signature expression	
		Low, N (%)	High, N (%)
0			
1			
2			
3	12 (26.09)	11 (91.67)	1 (8.33)
4 or 5	30 (65.21)	27 (90.00)	3 (10.00)
NA	4 (8.70)	3 (75.00)	1 (25.00)
Total	46 (100.00)	41 (89.13)	5 (10.87)

Among the 42 patients with benign breast disease, the results of breast ultrasound were retrospectively found that there were 30 cases (65.21%) with BI-RADS 4 or 5 readings, while 12 cases (26.09%) with BI-RADS 3 readings (Table 13). When breast ultrasound results were set as the endpoint and evaluated from the perspective of positive diagnostic evaluation, the cutoff of breast ultrasound was set to BI-RADS 4. Comparing the true negative and false positive rates of breast ultrasound and miRNA signature analysis among the 42 patients with benign breast disease, it was observed that 11 out of 12 benign patients with BI-RADS 3 readings were classified as true negatives (91.67%) based on miRNA signature analysis. For the 30 subjects who underwent unnecessary biopsies due to BI-RADS readings of 4 or higher, 27 cases (90.00%) were identified as true negatives through miRNA signature analysis.

Table 14. Comparison of clinical specificity of breast ultrasound and miRNA signature in subjects confirmed to have benign breast disease through biopsy.

		Patients with benign breast tumor
BI-RADS from Ultrasound	**≥ Category 4**	30
	< Category 4	12
Total		42

		Patients with benign breast tumor
miRNA Signature	**Elevated**	4
	Not elevated	38
Total		42

Parameters	**Ultrasound**	**miRNA signature**
True negative rate	28.57%	90.48%
False positive rate	71.43%	9.52%

Considering the clinical utility, the introduction of miRNA signature analysis in the diagnostic process for breast cancer could potentially classify 90.48% of patients diagnosed with benign breast disease, who underwent unnecessary biopsies based on breast ultrasound, as true negatives (Table. 14). This suggests that miRNA signature analysis could significantly aid in the diagnostic process, reducing unnecessary invasive procedures and improving patient care.

4. DISCUSSION

Recent years have witnessed numerous attempts at liquid biopsies utilizing circulating biomarkers like ctDNA and CTCs, often through whole-genome sequencing or DNA methylation patterns (20,21,23,34,80). Despite advancements in technologies with improving detection limits, developing a liquid biopsy that accurately reflects specific cancer characteristics, especially in early diagnosis, remains an unresolved challenge (20,21,34). EVs, actively secreted and concentrated in living tumor cells, offer a promising avenue (45). Their cargo, enriched and protected by a lipid double-layer structure, presents them as potential circulating tumor biomarkers (29,30). Unlike other liquid biopsy biomarkers, they are relatively immune to stability limitations and noise, making them attractive for diagnostic purposes (40). Previous research has primarily concentrated on membrane proteins, particularly tetraspanins (CD9, CD63, CD81, and so forth), which frequently become overexpressed in EVs (40), our study stands out for discovering and analyzing surface biomarkers that were not previously explored. Given the overexpression of specific membrane proteins in diverse molecular subtypes of breast carcinoma, BEVs were successfully isolated using immune-affinity technology and demonstrated significant potential. Their diagnostic performance in BEV microRNA analysis surpassed that of miRNAs derived from tissue and TEV. The relatively high sensitivity for early-stage breast cancer, particularly contrasting with previous studies, is highly encouraging. Furthermore, our research went beyond traditional diagnostic analysis by confirming the potential clinical utility of breast cancer diagnosis through complementary analysis of breast imaging and miRNA signature. This not only supplements the sensitivity of breast imaging but also enhances the specificity for benign breast conditions. Thus, our study underscores the potential clinical utility of breast cancer diagnosis through these combined approaches. In this study, a miRNA combination comprised of miR-21, miR-106b, miR-181a, miR-484, and miR-1260b in BEVs exhibited upregulation in breast cancer patients compared to those with benign breast disease and healthy individuals. The validity of the BEV-derived miRNA signature is supported by prior studies on the distinct functions of these miRNAs in breast cancer, even if separation and analysis of EVs derived from tumors are not yet ideally established. Breast cancer progression and metastasis are influenced by miR-21, which suppresses tumor-suppressor genes like PTEN and PDCD4, thereby triggering the MEK/ERK and PI3K/AKT signaling pathways (81–84). Like miR-21, miR-106b facilitates breast carcinoma growth by inhibiting PTEN, which enhances tumor invasion, cell proliferation, and migration (85–89). Additionally, it activates the Rho/ROCK1 pathway, accelerating tumor progression (88). Breast cancer progression is associated with miR-181a-5p, which targets PTEN, promotes myeloid-derived suppressor cell expansion, and drives tumor growth (90,91). It also upregulates p-AKT levels, promoting cell proliferation and S phase entry, while dysregulated TGF-β signaling increases its expression, leading to reduced apoptosis, increased cell migration, and invasion (91). *Yang et al.* (92) conducted a review highlighting the conflicting findings regarding the role of miR-181a-5p as a dual regulator. Specifically, the expression levels of miR-181a in breast cancer patients relative to healthy individuals have been the subject of

conflicting findings in several studies investigating miRNAs in blood. While some researchers reported lower miR-181a level in breast carcinomas (93–96), others reported elevated levels (97). Although there are reports indicating sufficient presence of miR-181a in EVs (92), research on its function and overexpression in both blood and EVs is still in its nascent stages, and definitive conclusions remain elusive. Thus, the expression of miR-181a-5p in the serum of breast cancer patients remains controversial. However, it is well-established that miR-181a-5p is overexpressed in breast cancer tissues, where it serves various functions, and is abundant in EVs (92). This suggests that BEVs may provide a more accurate reflection of the microenvironment of breast cancer tissue compared to TEVs. In our investigation, patients with benign breast diseases had considerably less levels of miR-181a in their BEVs than did patients with breast cancer. Therefore, it is plausible that miR-181a-5p is enriched to a greater extent in BEVs. By targeting KLF4 expression, miR-484 re-sensitizes breast cancer cells that are resistant to tamoxifen (98). It is enriched in key pathways implicated in breast carcinoma growth and is significantly upregulated in the plasma and serum of breast cancer patients, suggesting its promise as a biomarker (98–100). High levels of miR-1260b expression are associated with reduced overall survival, greater invasiveness, higher tumor burden, accelerated proliferation, and distant metastasis (101,102). MiR-2160b activates the MAPK signaling pathway by targeting CCDC134, thereby promoting cancer cell migration, invasion, and immune system evasion (101,103). In summary, these miRNAs are key factors targeting essential genes and pathways involved in breast cancer development and invasiveness. The robust diagnostic performance of miRNA candidates derived from BEV validates the BEV analysis method, reinforcing its accuracy in detecting and characterizing onco-miRNAs associated with breast cancer

Beyond evaluating the diagnostic ability of the microRNA combination via the isolation of BEV, we explored its potential clinical feasibility to enhance current breast cancer diagnostic methods. Mammography is a vital tool for breast cancer screening because of its excellent clinical specificity. (4,6,14,17). Mammography frequently produces false-negative outcomes in women with dense breast tissue, considerably obstructing breast cancer detection and assessment and increasing the likelihood of interval cancer (3,4). Therefore, our findings suggest that combining miRNA analysis with mammography could address the issue of low clinical sensitivity in individuals with dense breast tissue. Additionally, the clinical specificity of the microRNA combination in patients with benign breast tumors referred for unnecessary biopsy confirmed its potential to compensate for false positives due to radiological diagnosis. This disparity between radiological evaluation and miRNA analysis proved problematic, given the lack of precedent for addressing such situations in this area. In particular, comparing and analyzing clinical performance by focusing on the BI-RADS 0 category required a highly cautious approach. In our approach, we prioritized classifying imaging findings as 'positive' if they were clearly malignant during the initial screening. This decision was driven by the clinical context, where any positive result would prompt additional diagnostic evaluations. As a result, the combined analysis of mammograms and microRNA was considered 'Positive' if either test yielded a positive results. Anticipating its potential application as a liquid biopsy for screening purposes, we conducted a sensitivity-specificity analysis with a focus on maintaining 99% clinical specificity. Under these stringent conditions, the clinical sensitivity was

slightly reduced, a strategic measure aimed at minimizing false positives in real-world scenarios. However, considering the reported mammography recall rate of approximately 5-12%(76–78), adjusting the clinical specificity to 90% would result in a confirmed clinical sensitivity of 77.50% (95% CI: 63.33 to 90.00). This suggests a promising potential for its use as a screening test. It is worth noting that significant scope exists for improvement by incorporating other analytes such as DNA, proteins, or metabolites of tumor-derived extracellular vesicle cargo along with this miRNA signature. However, these findings have limitations that must be addressed through large-scale clinical studies, so reaching a conclusion based solely on the study results is insufficient. Consequently, we evaluated the diagnostic performance of the miRNA combination from BEV isolates, reviewed its clinical feasibility, and evaluated its applicability in clinical situation. To our knowledge, the combination of a single type of analyte has resulted in the best clinical diagnostic performance reported. There is room for improvement if EVs are analyzed with various analytes, and there is also the possibility of expansion to other tumors. We aimed to evaluate the clinical feasibility of microRNA combinations, with confirmed performance through BEV, to determine their potential in enhancing existing diagnostic methods. Our study's primary intent was not to promote the direct implementation of this integrated assay in breast screening or to propose further diagnostic interventions. Instead, our study validates the diagnostic value of miRNA combination derived from BEV, demonstrates their possibility for diagnosis of breast cancer, and outlines future study and clinical application possibilities

To solidify the 5-microRNA combination in BEV as a biomarker of breast carcinoma, certain key discussions have remained. To begin, it is crucial to conduct an in-depth examination of miRNA expression patterns in breast cancer. This investigation should clarify whether elevated microRNA combinations are uniquely associated with breast cancer, or if additional contributing factors are also involved. While this study limited demographic analysis to age, other factors—such as BMI, smoking, alcohol use, menopausal status, and obstetric history—should also be included to explore their possible contribution to the elevated microRNA signature in breast cancer patients. Second, there is a bias in the patient population distribution. The breast cancer patients in this study were spread across different TNM stages, histological types, and molecular subtypes, which could introduce biases in the interpretation of results, warranting careful analysis. Third, the study's approach to molecular subtypes presents limitations. The goal was to reflect tumor heterogeneity, which is often overlooked in studies with lower sensitivity for breast cancer detection. By incorporating EpCAM, CD49b, and CD51 in BEV separation through immunoaffinity, the study aimed to capture this heterogeneity. While this approach showed higher clinical performance than TEV-based diagnostic methods, it does not guarantee the capture of all breast cancer-specific EVs, leaving room for missed EVs. In conclusion, the study suggests that selectively isolating EVs that capture breast cancer heterogeneity for miRNA analysis offers a valuable method to enrich tumor-derived EVs within the broader EV population in blood. This enables a multifaceted diagnostic window, allowing for the analysis of nucleic acids, lipids, metabolites, and proteins within BEV cargo.

5. CONCLUSION

Our study focused on isolating EVs that precisely mirror the unique features of breast cancer, aiming to complement current diagnostic strategies and enhance clinical performance over TEVs and tissue-based methods. The miRNA signature from BEVs, consisting of oncogenic miRNAs, outperformed those from TEVs and tissue in clinical performance, suggesting the strong potential of our BEV-derived analysis. In contrast to ctDNA analysis, which faces certain limitations, our approach offers several improvements. Firstly, our method enhances the detection limit, particularly in early-stage tumors where ctDNA expression is minimal. EV-based testing capitalizes on the enrichment of tumor-related molecules within EVs, overcoming the sensitivity limitations of ctDNA assays in early-stage cancer detection. Secondly, our approach offers greater stability. Unlike ctDNA, which is prone to degradation in the bloodstream due to clearance mechanisms, EVs are protected by a double-layer phospholipid membrane and exhibit prolonged stability during sampling and storage. This enhanced stability facilitates practical clinical applications, positioning EV-based testing as a more reliable option compared to ctDNA-based methods. Thirdly, our method enhances specificity. ctDNA assays often suffer from a high false positive rate due to the presence of mutations originating from clonal hematopoiesis of indeterminate potential (CHIP) in the blood, particularly in aging individuals. However, EVs encapsulate intracellular molecules and are shielded by endosomes during release, minimizing the presence of CHIP-related mutations in circulating samples and thereby reducing the risk of false positives.

Indeed, leveraging the advantages offered by EVs, we plan to broaden our approach to encompass multi-omics analysis, aiming to identify additional biomarkers that are enriched and protected within BEVs. By expanding our research in this direction, we will aim to enhance sensitivity while maintaining a specificity of 99%. To ensure the dependability of our data, we will implement a research strategy that minimizes sampling bias and validates our results over numerous patient cohorts. Through these efforts, we seek to strengthen the robustness of our data and establish the diagnostic value of BEV-derived biomarkers with confidence. This study highlights the clinical significance of tumor-specific liquid biopsy strategies, demonstrating improved results via the collection and evaluation of tumor-derived extracellular vesicles (EVs).

REFERENCES

1. Lei S, Zheng R, Zhang S, Wang S, Chen R, Sun K, et al. Global patterns of breast cancer incidence and mortality: A population-based cancer registry data analysis from 2000 to 2020. Cancer Commun. 2021 Nov;41(11):1183–94.

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.

3. Ginsburg O, Yip CH, Brooks A, Cabanes A, Caleffi M, Dunstan Yataco JA, et al. Breast cancer early detection: A phased approach to implementation. Cancer. 2020 May 15;126 Suppl 10(Suppl 10):2379–93.

4. Bevers TB, Niell BL, Baker JL, Bennett DL, Bonaccio E, Camp MS, et al. NCCN Guidelines® Insights: Breast Cancer Screening and Diagnosis, Version 1.2023: Featured Updates to the NCCN Guidelines. J Natl Compr Canc Netw. 2023 Sep 1;21(9):900–9.

5. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013 Jun 11;108(11):2205–40.

6. Bevers TB, Helvie M, Bonaccio E, Calhoun KE, Daly MB, Farrar WB, et al. Breast Cancer Screening and Diagnosis, Version 3.2018, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2018 Nov;16(11):1362–89.

7. Hendrick RE. Obligate Overdiagnosis Due to Mammographic Screening: A Direct Estimate for U.S. Women. Radiology. 2018 May;287(2):391–7.

8. Kerlikowske K. Comparative Effectiveness of Digital Versus Film-Screen Mammography in Community Practice in the United States: A Cohort Study. Ann Intern Med. 2011 Oct 18;155(8):493.

9. Berg WA, Blume JD, Cormack JB, Mendelson EB, Lehrer D, Böhm-Vélez M, et al. Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer. JAMA. 2008 May 14;299(18):2151–63.

10. Shin HJ, Ko ES, Yi A. Breast Cancer Screening in Korean Woman with Dense Breast Tissue. J Korean Soc Radiol. 2015;73(5):279.

11. Freer PE. Mammographic Breast Density: Impact on Breast Cancer Risk and Implications for Screening. RadioGraphics. 2015 Mar;35(2):302–15.

12. Lee CI, Chen LE, Elmore JG. Risk-based Breast Cancer Screening. Med Clin North Am. 2017 Jul;101(4):725–41.

13. Buchberger W, Geiger-Gritsch S, Knapp R, Gautsch K, Oberaigner W. Combined screening with mammography and ultrasound in a population-based screening program. Eur J Radiol. 2018 Apr;101:24–9.

14. Berg WA, Zhang Z, Lehrer D, Jong RA, Pisano ED, Barr RG, et al. Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk. JAMA. 2012 Apr 4;307(13):1394–404.

15. Kuhl CK, Schrading S, Leutner CC, Morakkabati-Spitz N, Wardelmann E, Fimmers R, et al. Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer. J Clin Oncol. 2005 Nov 20;23(33):8469–76.

16. Lee JM, Arao RF, Sprague BL, Kerlikowske K, Lehman CD, Smith RA, et al. Performance of Screening Ultrasonography as an Adjunct to Screening Mammography in Women Across the Spectrum of Breast Cancer Risk. JAMA Intern Med. 2019 May 1;179(5):658–67.

17. Heywang-Köbrunner SH, Hacker A, Sedlacek S. Advantages and Disadvantages of Mammography Screening. Breast Care. 2011 Jun;6(3):199–207.

18. Yoon JH, Kim EK. Deep Learning-Based Artificial Intelligence for Mammography. Korean J Radiol. 2021 Aug 1;22(8):1225–39.

19. Lång K, Josefsson V, Larsson AM, Larsson S, Högberg C, Sartor H, et al. Artificial intelligence-supported screen reading versus standard double reading in the Mammography Screening with Artificial Intelligence trial (MASAI): a clinical safety analysis of a randomised, controlled, non-inferiority, single-blinded, screening accuracy study. Lancet Oncol. 2023 Aug 1;24(8):936–44.

20. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic — implementation issues and future challenges. Nat Rev Clin Oncol. 2021 May;18(5):297–312.

21. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet. 2019 Feb;20(2):71–88.

22. Cescon DW, Bratman SV, Chan SM, Siu LL. Circulating tumor DNA and liquid biopsy in oncology. Nat Cancer. 2020 Mar 20;1(3):276–90.

23. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nat Rev Cancer. 2017 Apr;17(4):223–38.

24. Klein EA, Richards D, Cohn A, Tummala M, Lapham R, Cosgrove D, et al. Clinical validation of a targeted methylation-based multi-cancer early detection test using an independent validation set. Ann Oncol. 2021 Sep;32(9):1167–77.

25. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV, Liu MC, et al. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020 Jun;31(6):745–59.

26. Jamshidi A, Liu MC, Klein EA, Venn O, Hubbell E, Beausang JF, et al. Evaluation of cell-free DNA approaches for multi-cancer early detection. Cancer Cell. 2022 Dec;40(12):1537-1549.e12.

27. Cohen JD, Li L, Wang Y, Thoburn C, Afsari B, Danilova L, et al. Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science. 2018 Feb 23;359(6378):926–30.

28. Lennon AM, Buchanan AH, Kinde I, Warren A, Honushefsky A, Cohain AT, et al. Feasibility of blood testing combined with PET-CT to screen for cancer and guide intervention. Science. 2020 Jul 3;369(6499):eabb9601.

29. Yu W, Hurley J, Roberts D, Chakrabortty SK, Enderle D, Noerholm M, et al. Exosome-based liquid biopsies in cancer: opportunities and challenges. Ann Oncol Off J Eur Soc Med Oncol. 2021 Apr;32(4):466–77.

30. Yu D, Li Y, Wang M, Gu J, Xu W, Cai H, et al. Exosomes as a new frontier of cancer liquid biopsy. Mol Cancer. 2022 Feb 18;21(1):56.

31. Meng Y, Sun J, Wang X, Hu T, Ma Y, Kong C, et al. Exosomes: A Promising Avenue for the Diagnosis of Breast Cancer. Technol Cancer Res Treat. 2019 Jan 1;18:153303381882142.

32. Wang X, Sun C, Huang X, Li J, Fu Z, Li W, et al. The Advancing Roles of Exosomes in Breast Cancer. Front Cell Dev Biol. 2021 Nov 1;9:731062.

33. Zhou B, Xu K, Zheng X, Chen T, Wang J, Song Y, et al. Application of exosomes as liquid biopsy in clinical diagnosis. Signal Transduct Target Ther. 2020 Aug 3;5(1):1–14.

34. Siravegna G. How liquid biopsies can change clinical practice in oncology. Ann Oncol. 2019;30(10).

35. Chan HT, Chin YM, Nakamura Y, Low SK. Clonal Hematopoiesis in Liquid Biopsy: From Biological Noise to Valuable Clinical Implications. Cancers. 2020 Aug 14;12(8):2277.

36. Grenier-Pleau I, Tyryshkin K, Le TD, Rudan J, Bonneil E, Thibault P, et al. Blood extracellular vesicles from healthy individuals regulate hematopoietic stem cells as humans age. Aging Cell. 2020 Nov;19(11):e13245.

37. Rahbarghazi R, Jabbari N, Sani NA, Asghari R, Salimi L, Kalashani SA, et al. Tumor-derived extracellular vesicles: reliable tools for Cancer diagnosis and clinical applications. Cell Commun Signal. 2019 Dec;17(1):73.

38. Piombino C, Mastrolia I, Omarini C, Candini O, Dominici M, Piacentini F, et al. The Role of Exosomes in Breast Cancer Diagnosis. Biomedicines. 2021 Mar 18;9(3):312.

39. Ekström K, Crescitelli R, Pétursson HI, Johansson J, Lässer C, Olofsson Bagge R. Characterization of surface markers on extracellular vesicles isolated from lymphatic exudate from patients with breast cancer. BMC Cancer. 2022 Dec;22(1):50.

40. Soung YH, Ford S, Zhang V, Chung J. Exosomes in Cancer Diagnostics. Cancers. 2017 Jan 12;9(1):8.

41. Tan S, Xia L, Yi P, Han Y, Tang L, Pan Q, et al. Exosomal miRNAs in tumor microenvironment. J Exp Clin Cancer Res. 2020 Apr 16;39(1):67.

42. Li C, Zhou T, Chen J, Li R, Chen H, Luo S, et al. The role of Exosomal miRNAs in cancer. J Transl Med. 2022 Jan 3;20(1):6.

43. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A. 2008 Jul 29;105(30):10513–8.

44. Kim MW, Park S, Lee H, Gwak H, Hyun KA, Kim JY, et al. Multi-miRNA panel of tumor-derived extracellular vesicles as promising diagnostic biomarkers of early-stage breast cancer. Cancer Sci. 2021 Dec;112(12):5078–87.

45. Yvonne C, Edit B, Dolores DV, Yong Song G, Paul H, Andrew H, et al. A brief history of nearly EV-erything - The rise and rise of extracellular vesicles. J Extracell Vesicles [Internet]. 2021 Dec [cited 2024 Jan 18];10(14). Available from: https://pubmed.ncbi.nlm.nih.gov/34919343/

46. Lee Y, Ni J, Beretov J, Wasinger VC, Graham P, Li Y. Recent advances of small extracellular vesicle biomarkers in breast cancer diagnosis and prognosis. Mol Cancer. 2023 Feb 16;22(1):33.

47. Applied Logistic Regression | Wiley Series in Probability and Statistics [Internet]. [cited 2024 Mar 5]. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118548387

48. SAGE Publications Inc [Internet]. 2024 [cited 2024 Mar 5]. Regression Models for Categorical and Limited Dependent Variables. Available from: https://us.sagepub.com/en-us/nam/regression-models-for-categorical-and-limited-dependent-variables/book6071

49. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996 Dec 1;49(12):1373–9.

50. D'Orsi C, Sickles E, Mendelson E, Morris E. ACR BI-RADS® Atlas, Breast Imaging

Reporting and Data System. American College of Radiology. 2013.

51. Rosenberg RD, Yankaskas BC, Hunt WC, Ballard-Barbash R, Urban N, Ernster VL, et al. Effect of variations in operational definitions on performance estimates for screening mammography. Acad Radiol. 2000 Dec;7(12):1058–68.

52. Gostner JM, Fong D, Wrulich OA, Lehne F, Zitt M, Hermann M, et al. Effects of EpCAM overexpression on human breast cancer cell lines. BMC Cancer. 2011 Jan 31;11(1):45.

53. Soysal SD, Muenst S, Barbie T, Fleming T, Gao F, Spizzo G, et al. EpCAM expression varies significantly and is differentially associated with prognosis in the luminal B HER2+, basal-like, and HER2 intrinsic subtypes of breast cancer. Br J Cancer. 2013 Apr 16;108(7):1480–7.

54. Osta WA, Chen Y, Mikhitarian K, Mitas M, Salem M, Hannun YA, et al. EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer Res. 2004 Aug 15;64(16):5818–24.

55. Bemmerlein L, Deniz IA, Karbanová J, Jacobi A, Drukewitz S, Link T, et al. Decoding Single Cell Morphology in Osteotropic Breast Cancer Cells for Dissecting Their Migratory, Molecular and Biophysical Heterogeneity. Cancers. 2022 Jan 25;14(3):603.

56. Taherian A, Li X, Liu Y, Haas TA. Differences in integrin expression and signaling within human breast cancer cells. BMC Cancer. 2011 Jul 13;11(1):293.

57. Hellinger JW, Schömel F, Buse JV, Lenz C, Bauerschmitz G, Emons G, et al. Identification of drivers of breast cancer invasion by secretome analysis: insight into CTGF signaling. Sci Rep. 2020 Oct 21;10(1):17889.

58. Cooper J, Giancotti FG. Integrin Signaling in Cancer: Mechanotransduction, Stemness, Epithelial Plasticity, and Therapeutic Resistance. Cancer Cell. 2019 Mar;35(3):347–67.

59. Desgrosellier JS, Cheresh DA. Integrins in cancer: biological implications and therapeutic opportunities. Nat Rev Cancer. 2010 Jan;10(1):9–22.

60. Yousefi H, Vatanmakanian M, Mahdiannasser M, Mashouri L, Alahari NV, Monjezi MR, et al. Understanding the role of integrins in breast cancer invasion, metastasis, angiogenesis, and drug resistance. Oncogene. 2021 Feb;40(6):1043–63.

61. Wong NC, Mueller BM, Barbas CF, Ruminski P, Quaranta V, Lin ECK, et al. αv Integrins mediate adhesion and migration of breast carcinoma cell lines. Clin Exp Metastasis. 1998 Jan 1;16(1):50–61.

62. Zhang DX, Dang XTT, Vu LT, Lim CMH, Yeo EYM, Lam BWS, et al. αvβ1 integrin is enriched in extracellular vesicles of metastatic breast cancer cells: A mechanism mediated by galectin-3. J Extracell Vesicles. 2022 Aug;11(8):e12234.

63. Alshammari FOFO, Al-Saraireh YM, Youssef AMM, AL-sarayra YM, Alrawashdeh HM. Glypican-1 Overexpression in Different Types of Breast Cancers. OncoTargets Ther. 2021 Jul 30;14:4309–18.

64. Matsuda K, Maruyama H, Guo F, Kleeff J, Itakura J, Matsumoto Y, et al. Glypican-1 is overexpressed in human breast cancer and modulates the mitogenic effects of multiple heparin-binding growth factors in breast cancer cells. Cancer Res. 2001 Jul 15;61(14):5562–9.

65. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. Nature. 2015 Jul 9;523(7559):177–82.

66. Shiino S, Matsuzaki J, Shimomura A, Kawauchi J, Takizawa S, Sakamoto H, et al. Serum miRNA-based Prediction of Axillary Lymph Node Metastasis in Breast Cancer. Clin Cancer Res Off J Am Assoc Cancer Res. 2019 Mar 15;25(6):1817–27.

67. Shimomura A, Shiino S, Kawauchi J, Takizawa S, Sakamoto H, Matsuzaki J, et al. Novel combination of serum microRNA for detecting breast cancer in the early stage. Cancer Sci. 2016 Mar;107(3):326–34.

68. Farina NH, Ramsey JE, Cuke ME, Ahern TP, Shirley DJ, Stein JL, et al. Development of a predictive miRNA signature for breast cancer risk among high-risk women. Oncotarget. 2017 Dec 22;8(68):112170–83.

69. Cimino D, De Pittà C, Orso F, Zampini M, Casara S, Penna E, et al. miR148b is a major coordinator of breast cancer progression in a relapse-associated microRNA signature by targeting ITGA5, ROCK1, PIK3CA, NRAS, and CSF1. FASEB J Off Publ Fed Am Soc Exp Biol. 2013 Mar;27(3):1223–35.

70. Chan M, Liaw CS, Ji SM, Tan HH, Wong CY, Thike AA, et al. Identification of circulating microRNA signatures for breast cancer detection. Clin Cancer Res Off J Am Assoc Cancer Res. 2013 Aug 15;19(16):4477–87.

71. Feliciano A, Castellvi J, Artero-Castro A, Leal JA, Romagosa C, Hernández-Losa J, et al. miR-125b Acts as a Tumor Suppressor in Breast Tumorigenesis via Its Novel Direct Targets ENPEP, CK2-α, CCNJ, and MEGF9. PLoS ONE. 2013 Oct 3;8(10):e76247.

72. Lee CH, Kuo WH, Lin CC, Oyang YJ, Huang HC, Juan HF. MicroRNA-Regulated Protein-Protein Interaction Networks and Their Functions in Breast Cancer. Int J Mol Sci. 2013 May 30;14(6):11560–606.

73. Zhao XG, Hu JY, Tang J, Yi W, Zhang MY, Deng R, et al. miR-665 expression predicts poor survival and promotes tumor metastasis by targeting NR4A3 in breast cancer. Cell Death Dis. 2019 Jun 17;10(7):1–21.

74. Matamala N, Vargas MT, González-Cámpora R, Miñambres R, Arias JI, Menéndez P, et al. Tumor microRNA expression profiling identifies circulating microRNAs for early breast cancer detection. Clin Chem. 2015 Aug;61(8):1098–106.

75. Hironaka-Mitsuhashi A, Matsuzaki J, Takahashi RU, Yoshida M, Nezu Y, Yamamoto Y, et al. A tissue microRNA signature that predicts the prognosis of breast cancer in young women. PloS One. 2017;12(11):e0187638.

76. Gur D, Sumkin JH, Rockette HE, Ganott M, Hakim C, Hardesty L, et al. Changes in Breast Cancer Detection and Mammography Recall Rates After the Introduction of a Computer-Aided Detection System. JNCI J Natl Cancer Inst. 2004 Feb 4;96(3):185–90.

77. Rauscher GH, Murphy AM, Qiu Q, Dolecek TA, Tossas K, Liu Y, et al. The “Sweet Spot” Revisited: Optimal Recall Rates for Cancer Detection With 2D and 3D Digital Screening Mammography in the Metro Chicago Breast Cancer Registry. Am J Roentgenol. 2021 Apr;216(4):894–902.

78. Yankaskas BC, Cleveland RJ, Schell MJ, Kozar R. Association of Recall Rates with Sensitivity and Positive Predictive Values of Screening Mammography. Am J Roentgenol. 2001 Sep;177(3):543–9.

79. Choi JE, Kim Z, Park CS, Park EH, Lee SB, Lee SK, et al. Breast Cancer Statistics in Korea, 2019. J Breast Cancer. 2023;26(3):207.

80. Tellez-Gabriel M, Knutsen E, Perander M. Current Status of Circulating Tumor Cells, Circulating Tumor DNA, and Exosomes in Breast Cancer Liquid Biopsies. Int J Mol Sci. 2020 Dec 11;21(24):9457.

81. Bautista-Sánchez D, Arriaga-Canon C, Pedroza-Torres A, De La Rosa-Velázquez IA, González-Barrios R, Contreras-Espinosa L, et al. The Promising Role of miR-21 as a Cancer Biomarker and Its Importance in RNA-Based Therapeutics. Mol Ther Nucleic Acids. 2020 Mar 13;20:409–20.

82. Feng YH, Tsao CJ. Emerging role of microRNA-21 in cancer (Review). Biomed Rep. 2016 Oct 1;5(4):395–402.

83. Hamam R, Hamam D, Alsaleh KA, Kassem M, Zaher W, Alfayez M, et al. Circulating microRNAs in breast cancer: novel diagnostic and prognostic biomarkers. Cell Death Dis. 2017 Sep 7;8(9):e3045–e3045.

84. Najjary S, Mohammadzadeh R, Mokhtarzadeh A, Mohammadi A, Kojabad AB, Baradaran B. Role of miR-21 as an authentic oncogene in mediating drug resistance in breast cancer. Gene. 2020 May;738:144453.

85. Lee J, Kim HE, Song YS, Cho EY, Lee A. miR-106b-5p and miR-17-5p could predict

recurrence and progression in breast ductal carcinoma in situ based on the transforming growth factor-beta pathway. Breast Cancer Res Treat. 2019 Jul;176(1):119–30.

86. Li N, Miao Y, Shan Y, Liu B, Li Y, Zhao L, et al. MiR-106b and miR-93 regulate cell progression by suppression of PTEN via PI3K/Akt pathway in breast cancer. Cell Death Dis. 2017 May 18;8(5):e2796–e2796.

87. Li N, Liu Y, Miao Y, Zhao L, Zhou H, Jia L. MicroRNA-106b targets FUT6 to promote cell migration, invasion, and proliferation in human breast cancer. IUBMB Life. 2016 Sep;68(9):764–75.

88. Sagar SK. miR-106b as an emerging therapeutic target in cancer. Genes Dis. 2022 Jul;9(4):889–99.

89. Yang C, Dou R, Yin T, Ding J. MiRNA-106b-5p in human cancers: diverse functions and promising biomarker. Biomed Pharmacother. 2020 Jul;127:110211.

90. Zhai Z, Mu T, Zhao L, Li Y, Zhu D, Pan Y. MiR-181a-5p facilitates proliferation, invasion, and glycolysis of breast cancer through NDRG2-mediated activation of PTEN/AKT pathway. Bioengineered. 13(1):83–95.

91. Li J, Shen J, Zhao Y, Du F, Li M, Xu X, et al. Role of miR-181a-5p in cancer (Review). Int J Oncol. 2023 Aug 3;63(4):108.

92. Yang C, Tabatabaei SN, Ruan X, Hardy P. The Dual Regulatory Role of MiR-181a in Breast Cancer. Cell Physiol Biochem. 2017;44(3):843–56.

93. Ferracin M, Lupini L, Salamon I, Saccenti E, Zanzi MV, Rocchi A, et al. Absolute quantification of cell-free microRNAs in cancer patients. Oncotarget. 2015 Jun 10;6(16):14545–55.

94. Guo LJ, Zhang QY. Decreased serum miR-181a is a potential new tool for breast cancer screening. Int J Mol Med. 2012 Sep;30(3):680–6.

95. McDermott AM, Miller N, Wall D, Martyn LM, Ball G, Sweeney KJ, et al. Identification and Validation of Oncologic miRNA Biomarkers for Luminal A-like Breast Cancer. Weisz A, editor. PLoS ONE. 2014 Jan 31;9(1):e87032.

96. Zhu J, Yao K, Wang Q, Guo J, Shi H, Ma L, et al. Circulating miR-181a as a Potential Novel Biomarker for Diagnosis of Acute Myocardial Infarction. Cell Physiol Biochem. 2016;40(6):1591–602.

97. Godfrey AC, Xu Z, Weinberg CR, Getts RC, Wade PA, DeRoo LA, et al. Serum microRNA expression as an early marker for breast cancer risk in prospectively collected samples from the Sister Study cohort. Breast Cancer Res. 2013 Jun;15(3):R42.

98. Jia Y zhao, Liu J, Wang G qiao, Song Z fang. miR-484: A Potential Biomarker in Health and Disease. Front Oncol. 2022 Mar 9;12:830420.

99. Holubekova V, Kolkova Z, Grendar M, Brany D, Dvorska D, Stastny I, et al. Pathway Analysis of Selected Circulating miRNAs in Plasma of Breast Cancer Patients: A Preliminary Study. Int J Mol Sci. 2020 Oct 2;21(19):7288.

100. Zearo S, Kim E, Zhu Y, Zhao JT, Sidhu SB, Robinson BG, et al. MicroRNA-484 is more highly expressed in serum of early breast cancer patients compared to healthy volunteers. BMC Cancer. 2014 Mar 18;14:200.

101. Park S, Kim J, Cho Y, Ahn S, Kim G, Hwang D, et al. Promotion of tumorigenesis by miR-1260b-targeting CASP8: Potential diagnostic and prognostic marker for breast cancer. Cancer Sci. 2022 Jun;113(6):2097–108.

102. Kim M, Moon S, Lee S, Lee H, Kim Y, Kim J, et al. Exploring miRNA-target gene profiles associated with drug resistance in patients with breast cancer receiving neoadjuvant chemotherapy. Oncol Lett. 2024 Feb 16;27(4):158.

103. Huang Z, Zhen S, Jin L, Chen J, Han Y, Lei W, et al. miRNA-1260b Promotes Breast Cancer Cell Migration and Invasion by Downregulating CCDC134. Curr Gene Ther. 23(1):60–71.

APPENDIX

Kim, Y.; Kim, J.Y.; Moon, S.; Lee, H.; Lee, S.; Kim, J.Y.; Kim, M.W.; Kim, S.I. Tumor-derived EV miRNA signatures surpass total EV miRNA in supplementing mammography for precision breast cancer diagnosis. Theranostics 2024, 14 (17), 6587-6604. DOI: 10.7150/thno.99245.

Abstract in Korean

유방암 진단을 보완하기 위한 종양 유래 세포외 소포체 내 5-miRNA 시그니처 프로파일링 액체 생검의 임상적 의의

전세계적으로 유방암은 영상의학적 진단 도입을 통해 암의 조기 진단율이 높아졌고 사망의 위험이 명백히 감소하였다. 그러나, 영상의학적 진단은 치밀유방으로 인한 위음성 증가, 위양성 결과로 인한 불필요한 생검 증가 등의 한계가 존재하며 조직 생검은 시공간적 종양 이질성을 반영하는데 한계가 있다. 이를 보완하기 위한 액체생검법의 개발은 지속적으로 요구되고 있다. 극소량의 분자를 검출할 수 있는 기술의 발전에도 불구하고, 특히 조기 진단에서 특정 암의 특성을 정확히 반영하는 액체 생검법의 개발은 아직 해결되지 않은 과제이다. 그러므로 우리는 유방암의 고유한 특성을 보다 정확하게 반영하는 유방암 유래 세포외 소포체 분리 및 분석을 목표로 현재 진단 방법을 보완하고 기존 방법 대비 임상적 성능을 향상시키는 것을 목표로 하였다. 유방암 분자아형을 포괄할 수 있는 EpCAM, ITGA2, ITGAV 를 도입하여 면역친화법을 통해 유방암 유래 세포외 소포체를 분리하고 소포체 내 miRNA 를 분석하여 최적의 진단 조합을 구축함으로써, 상술한 현 유방암 진단 및 기존 액체생검 연구 접근법의 한계를 보완하고자 하였다.

총 211 명(유방암 환자 120 례, 유방 양성질환 환자 46 례, 정상대조군 45 례)을 대상으로 5-miRNA 시그니처(miR-21, miR-106b, miR-181a, miR-484, miR-1260b)의 임상적 성능을 검증한 결과, 임상적 민감도는 85.83%, 임상적 특이도는 84.62%였으며, AUC 는 0.908 로 분석되었다. 그리고 TNM 병기에 따라 임상적 민감도는 82.35%, 87.72%, 91.67%로 확인됨에 따라, 초기 유방암 병기에 대한 임상적 유효성이 확인되었다. 유방촬영술과 miRNA 시그니처를 병행 분석하였을 때, 임상적 민감도는 97.39%로 분석되었다. 또한 유방촬영술 결과, BI-RADS 카테고리 0 로 판정되었던 각 대상자에 대한 miRNA 시그니처의 진양성률은 88.89%이었고, 진음성률은 92.86%로 분석되었다. 이와 더불어 치밀 유방에 대한 miRNA 시그니처와 유방촬영술 병행 분석의 임상적 민감도는 94.90%로 분석되었으며, 불필요한 조직검사를 받은 유방양성질환 환자에 대해 miRNA 시그니처의 진음성률은 90.48%로 분석되었다. 결론적으로 우리 연구는 현재의 유방암 진단을 보완하기 위해 종양의 고유 특성을 반영한 액체생검법을 개발하여 임상적 근거를 제시하였다. 유방암 유래 세포외 소포체 분리 분석방법을 통해 기존 방법 대비 높은 임상적

성능을 달성했고, 특히 기존 연구와 대비되는 초기 유방암에 대한 상대적으로 높은 민감도를 달성했다. 또한 유방촬영술과 miRNA 시그니처의 상호보완적 분석을 통해 유방촬영술의 민감도 보완과 유방 양성 질환에 대한 특이도 보완의 가능성을 확인함으로써, 유방암 진단에 대한 잠재적 유용성을 제시하였다. 이를 통해 유방암 유래 세포외 소포체 내에서 핵산, 지질, 대사산물, 단백질 등을 활용하는 다중체학적 진단 기회를 제공하며 유방암 진단 액체 생검법으로서의 잠재력을 제공하였다.

핵심되는 말 : 유방암, 액체 생검, 세포외 소포체, 엑소좀, 면역친화방법, 마이크로 RNA, 진단