Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

MicroRNAs detect breast cancer in nipple aspirate fluid: results from a Dutch multicenter cross-sectional study

Laura A. de Rooij

l.a.derooij-4@umcutrecht.nl

University Medical Center Utrecht

Susana I.S. Patuleia

University Medical Center Utrecht

Bas B.L. Penning de Vries

Julius Center for Health Sciences and Primary Care

Sjoerd G. Elias

Julius Center for Health Sciences and Primary Care

Joost M. H. H. van Gorp

St. Antonius Ziekenhuis

Julia E.C. van Steenhoven

Diakonessenhuis hospital

Thijs van Dalen

Erasmus MC

Liesbeth M. Veenendaal

Alexander Monro Ziekenhuis

Carmen C. van der Pol

Alrijne Ziekenhuis

Weronika Szczecińska

Medical University of Gdańsk

Jacek Zieliński

Medical University of Gdańsk

Milena Deptuła

Medical University of Gdańsk

Aneta Skoniecka

Medical University of Gdańsk

Michał Pikuła

Medical University of Gdańsk

Elsken van der Wall

University Medical Center Utrecht
Paul J. van Diest
University Medical Center Utrecht
Cathy B. Moelans
University Medical Center Utrecht

Research Article

Keywords: Breast cancer, microRNAs, nipple aspirate fluid, screening, diagnosis

Posted Date: July 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10201025/v1

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

Additional Declarations: Competing interest reported. P.J. van Diest, E. van der Wall, C.B. Moelans, L.A. de Rooij and S. G. Elias are listed as inventors on a patent application filed by Utrecht Holdings that describes the method of the diagnostic test described in this study. All remaining authors have declared no conflict of interest.

Abstract

Background

Breast cancer is the most common type of cancer and the leading cause of cancer-related deaths in women worldwide. Novel (personalized) treatments and detection as early as possible are key to improving prognosis. We hypothesized that microRNA (miRNA) levels in nipple aspirate fluid (NAF) can be used to detect breast cancer.

Methods

NAF samples were acquired from 135 breast cancer patients and 114 healthy controls using a non-invasive collection method. miRNA levels were analyzed with quantitative polymerase chain reaction using Taqman Advanced miRNA assays. The discriminative value of a 20-miRNA NAF panel for breast cancer detection was evaluated with LASSO regression and cross-validation.

Results

Six individual miRNAs demonstrated significant odds ratio ranging between 0.86–1.38. After LASSO regression with cross-validation, a 20-miRNA panel resulted in an area under the curve (AUC) of 0.70. A second analysis, where within women the affected breast was compared to the contralateral breast, resulted in an AUC of 0.54.

Conclusion

A 20-miRNA panel in NAF is able to discriminate well between breast cancer patients and healthy controls. With a good discriminative value and little discomfort, our method shows great promise for future personalized screening applications. One group that can especially benefit from this personalized strategy is women with dense breasts, where a positive miRNA test could stratify for supplemental MRI. Future studies are still necessary to determine if early stage detection is possible using the miRNA panel.

Trial registration

CCMO NL57343.041.16, Dutch Trial Register NTR6162/NL6031, Registered 17 August 2016, https//onderzoekmetmensen.nl/nl/trial/50271 CCMO NL57343.041.16, Dutch Trial Register NTR6162/NL6031, Registered 20 February 2013, https//onderzoekmetmensen.nl/nl/trial/47370

Introduction

Breast cancer is the most common type of cancer and the leading cause of cancer-related deaths in women worldwide[1]. In the Netherlands, 1 in 8 women develop breast cancer during their lifetime[2]. Fortunately, the five year overall survival rate of breast cancer has increased to 88% in recent decades[3]. This increase is mainly due to nationwide screening programs and the development of better treatment options, where screening has contributed to a 33–58% decrease in mortality rates since 1988[2]. Nonetheless, still nine Dutch breast cancer-related deaths are reported every day, necessitating better early detection and further improvements in therapy[3].

In the Netherlands, women with an average lifetime risk (< 20%) of developing breast cancer are screened through biennial mammograms between the ages of 50–75[4]. Women with an increased lifetime risk (≥ 20%) may be screened with mammography and magnetic resonance imaging (MRI)[5]. Mammography and MRI have contributed to an increase in survival rates, but have limitations[4, 6, 7]. Mammography has high specificity (~ 85%, range of 63–99%), but low sensitivity (~ 55%, range of 27–87%)[8], with an area under the curve (AUC) ranging between 0.76 and 0.91[9–11]. High false negative rates are especially seen in women with dense breasts[12]. In contrast, MRI has lower specificity (~ 74%, range of 25–100%), but high sensitivity (~ 95%, range of 86–100%)[8], with an AUC of 0.92[13]. The lower specificity leads to high false positives rates (52–97/1000 cases[14]). False positive screening results cause anxiety and increase healthcare costs[15]. In addition, discomfort is present for both mammography (breast compression) and MRI (noise, claustrophobia), that may result in women not adhering to screening guidelines[16]. Therefore, new personalized screening techniques are needed.

One such technique could be the detection of biomarkers in biofluids, such as nipple aspirate fluid (NAF). NAF is present in the ductal system of the breast and originates from luminal epithelial cells[17]. Nipple fluid aspiration is a non-invasive collection method, best supported by oxytocin nasal spray [16, 18]. Due to their close proximity to the origin of carcinogenesis, biomarkers in NAF will likely indicate subtle pathological changes occurring in the breast microenvironment with high sensitivity and specificity[17]. MicroRNAs (miRNAs) are a promising class of biomarkers in this respect[19, 20]. miRNAs are small non-coding RNAs that regulate gene expression at the post-translational level through binding of messenger RNAs[21]. Compelling evidence has demonstrated that miRNA expression is dysregulated in cancer[21], and that miRNAs are attractive biomarkers due to their unparalleled stability in liquid biopsies[21].

In this study, we evaluated the biomarker potential of miRNAs in NAF for the detection of breast cancer, as a first step towards the development of a novel personalized screening tool. Moreover, we studied the associations between miRNA expression and participants characteristics to gain more insight into the biological mechanisms underlying breast cancer development. A case-control study was set up to investigate the differences in miRNA expression between breast cancer patients and healthy controls. Additionally, a second analysis was conducted to asses whether the affected breast could be distinguished from the contralateral (non-affected) breast within patients.

Material & Methods

A detailed explanation of the material and methods can be found in supplementary material 1. In short, all invasive breast cancer samples were obtained from the ORNAMENT study: a multicenter, crOss sectional, study to assess microRNA expression in Nipple Aspirate Fluid, blood and tuMor material in women with primary brEast caNcer compared with healthy conTrols (CCMO NL57343.041.16; Dutch Trial Register NTR6162/NL6031)[22]. All healthy control (women > 45 years old) samples were obtained from another study: Breast Cancer Biomarkers in Nipple Aspiration Fluid in Healthy Controls (CCMO NL41845.041.12; Dutch Trial Register NL8987)[23]. Both studies were approved by the Institutional Review Board of the University Medical Center Utrecht and other participating hospitals (ORNAMENT: study number 16–337; healthy controls: study number 12–495). Written informed consent was obtained from all participating women.

NAF samples were acquired as previously described by Patuleia et al.[16]. To extract (mi)RNA, and DNA, the Allprep Universal DNA/RNA/miRNA kit was used. The Taqman Advanced miRNA cDNA Synthesis kit was used for cDNA synthesis (with 19 reaction cycles). Next, a quantitative polymerase chain reaction (qPCR) was performed in custom pre-coated 384 well plates containing 23 TaqMan Advanced miRNA Assays (Tables S1 and S2), with the Taqman Fast Advanced Mastermix. For relative quantification, miR-99b-5p was used as endogenous control.

R version 4.3.2 was used for the statistical analysis. Linear regression was used to determine the difference in individual miRNA expression (fold change) between healthy controls and breast cancer patients. Logistic regression was used to calculate adjusted odds ratios (ORs) for individual miRNAs, with breast cancer status as the outcome variable. Graphpad Prism v10 was used to plot the fold changes and ORs. To assess associations between miRNA expression and participant characteristics, appropriate statistical tests (two-sided t-test, ANOVA with post-hoc Tukey HSD test, Mann-Whitney U, and Kruskal Wallis) and correlation analysis (Pearson or Spearman correlation) were applied. A Bonferroni Holm correction for multiple comparisons was used. To assess the ability of the 20-miRNA panel (and age) to discriminate between patients and healthy controls, a multivariable model was developed using LASSO regression together with an external cross-validation step (10-folds, 20 iterations). A second model was developed using LASSO-penalized conditional logistic regression for the paired comparison between the affected breast and the contralateral breast.

Results

Participant and tumor characteristics

In this case-control study, we compared miRNA expression in NAF between breast cancer patients (n = 154 in total / n = 135 excluding samples without successful nipple fluid aspiration) and healthy controls (n = 116 / n = 114 excluding participants with a breast cancer diagnosis after NAF collection). Specific inclusion and exclusion criteria were set (Supplement 1), enabling unbiased comparison between the

two cohorts. The most relevant participant characteristics for each cohort are presented in Table 1 (overview of all characteristics in tables S3, S4 and S5). The variables age at NAF collection, body mass index (BMI), and parity did not significantly differ between breast cancer patients and healthy controls ($P = 0.059$, $P = 0.084$, and $P = 0.663$, respectively). As expected, having a first-degree relative with breast cancer was significantly more frequent in the breast cancer population ($P < 0.001$).

Table 1

Participant and tumor characteristics of breast cancer and healthy control cohorts. Variables presented are age at NAF collection, body mass index, having a first-degree relative with breast cancer, and parity. Tumor characteristics from the breast cancer cohort include molecular subtype (luminal A-like, luminal B-like, Human Epidermal growth factor Receptor 2 (HER2))-driven, and triple negative), receptor status (Estrogen receptor (ER), Progesterone receptor (PR), and HER2), and the modified Bloom and Richardson grade. This table excludes participants from whom NAF collection was unsuccessful in both breasts. For continuous variables, the median and interquartile range (IQR) interval are given. Significance for continuous variables was calculated using a t-test or Mann-Whitney U test, depending on their distribution. For categorical variables a Chi-square test was used.

	Breast Cancer (N = 135)			**Healthy Controls (N = 114)**			**Comparison**
	Median	**IQR**	**Missing**	**Median**	**Range**	**Missing**	***P* value**
Age at NAF collection	56	49–66	0	53	48–59	0	0.061
Body mass index	26	23–29	12	24	22–28	0	0.085
	Yes	**No**	**Missing**	**Yes**	**No**	**Missing**	**P value**
First-degree relative	41 (43%)	55 (57%)	39	19 (17%)	93 (83%)	2	< 0.001
Parous	104 (84%)	20 (16%)	11	98 (87%)	15 (15%)	1	0.663
	n	**%**	**Missing**	**n**	**%**	**Missing**	**P value**
Luminal A-like	91	75%	11	NA	NA	NA	NA
Luminal B-like	24	18%		NA	NA	NA	NA
HER2-driven	4	3%		NA	NA	NA	NA
Triple negative	5	4%		NA	NA	NA	NA
ER+	121	93%	5	NA	NA	NA	NA
PR+	93	73%	7	NA	NA	NA	NA
HER2+	11	8%	6	NA	NA	NA	NA
Grade 1	39	30%	4	NA	NA	NA	NA
Grade 2	71	54%		NA	NA	NA	NA
Grade 3	21	16%		NA	NA	NA	NA

Nipple fluid characteristics

From the total cohort, 244 NAF samples were cloudy (Breast cancer (BC) affected side = 61; Contralateral (CL) side = 55; Healthy control (HC) = 128, see Table S5) and 218 samples were clear (BC = 68; CL = 68; HC = 82). This variable is associated with miRNA expression levels, as evidenced by separate grouping of cloudy and clear samples (Figure S1). miRNA expression in clear NAF samples was significantly higher compared to the expression in cloudy samples ($P < 0.001$ for all interrogated miRNAs, except for miR-21-5p ($P = 0.077$).

Another known factor that is associated with miRNA expression is the color of the NAF sample[19]. This association was validated in our data, where in general green samples showed a significantly lower miRNA expression, whereas colorless, beige, yellow, and orange/bloody samples tended to have a higher miRNA expression level (Table S6). The distribution of different NAF colors was similar within the breast cancer cohort and the healthy controls (Table S5 & Figure S2). In addition, NAF color is associated with sample turbidity (Table S7).

MicroRNA expression levels in relation to participant and tumor characteristics

Pearson or Spearman correlation analysis was performed to evaluate the relation between multiple continuous variables and NAF miRNA expression. No significant correlations were observed with BMI, tumor mitotic count, and tumor size after Bonferroni Holm correction for multiple comparisons (Table S6). For age at NAF collection, miR-125a-5p ($P < 0.001$, r = -0.175), miR-145-5p ($P < 0.001$, r = -0.310), miR-155-5p ($P < 0.001$, r = 0.202), miR-21-5p ($P < 0.001$, r = 0.316) and miR-222-3p ($P = 0.001$, r = -0.154) showed a significant correlation, although there is a high degree of variability in the data (Figure S3).

The categorical variables with significant differences in miRNA expression levels between categories after Bonferroni Holm correction were: parity combined with breastfeeding, and breastfeeding (Supplement 2 and Table S6). No significant differences in miRNA expression levels between groups were observed after Bonferroni Holm correction for the following variables: breast density, history of breast disease (benign or invasive), previous benign breast lump, previous breast surgery, first-degree relative, parity, age of first birth, smoking, alcohol use, breast cancer laterality, histological subtype, estrogen receptor (ER) status, progesterone receptor (PR) status, human epidermal growth factor receptor 2 (HER2) status, lymph node status, Bloom and Richardson grade, tumor size (diameter), molecular subtype and tumor focality (Table S6, *P* values before correction also shown).

Differential microRNA expression in breast cancer patients versus healthy controls

A 23-miRNA panel including one endogenous control miRNA, was assessed in NAF samples from both cohorts. The data of miR-126-5p, miR-151a-5p, and miR-18a-5p were excluded due to insufficient quality following laboratory processing, leaving a panel of 20 miRNAs for the final data analysis. As shown in Supplementary Table 8, before normalization, all miRNAs, except miR-145-5p and miR-205-5p, showed a slightly higher average raw Ct value (lower expression) in the healthy control group compared to the breast cancer group (both affected and contralateral side). A comparison of ΔCt values between affected, contralateral, and healthy, after correction for RNA input and cloudiness, is shown in Supplementary Table 9. Moreover, a linear regression identified fold changes of the miRNAs in breast cancer patients versus healthy controls ranging from 0.54–1.57 (Table S10 and Figure S4).

Associations between miRNA expression levels and breast cancer status

A logistic regression was performed to estimate adjusted ORs for breast cancer based on individual miRNA expression levels (Table S11). Similar to the observed fold changes, five miRNAs had an increased adjusted OR: miR-153-3p (OR = 1.31, *P* = 0.002, 95% confidence interval (CI) [1.11, 1.57]), miR-155-5p (OR = 1.13, *P* = 0.023, 95% CI [1.02, 1.27]), miR-21-5p (OR = 1.38, *P* = 0.004, 95% CI [1.12, 1.72]), miR-29c-5p (OR = 1.27, *P* = 0.012, 95% CI [1.06, 1.55]), and miR-374b-5p (OR = 1.33, *P* = 0.011, 95% CI [1.07, 1.66]) (Fig. 1). miR-145-5p showed a decreased adjusted OR (OR = 0.86, *P* = 0.010, 95% CI [0.77, 0.96]). After correction for multiple comparison using Bonferroni Holm only miR-153-3p remained significant.

Discriminative value of a 20-miRNA panel for the detection of breast cancer

A LASSO regression analysis was performed to determine whether the 20 interrogated miRNAs were able to differentiate between breast cancer patients (n = 128) and healthy controls (n = 113). Age was also included as a discriminative variable. Model performance was investigated by means of a receiver operating characteristic curve, resulting in an AUC of 0.804 (95% CI [0.749, 0.858]) (Fig. 2). After correction for overfitting through cross-validation, the median AUC was 0.702 (95% of cross-validation estimates fall between 0.485 and 0.896). Because LASSO regression enhances predictive accuracy on unseen data by shrinking the coefficients of weak predictors to zero, we used it to select the miRNAs that contribute most to out-of-sample model performance. The miRNAs that contributed the most and that were taken along in more than 90% of the folds and iterations were: miR-145-5p (100%), miR-153-3p (100%), miR-29c-5p (100%), miR-221-3p (98.5%), miR-222-3p (96.5%), miR-320a-3p (95.5%), miR-374b-5p (95.0%), miR-92a-3p (94.5%), miR-205-5p (92.0%), and miR-19a-3p (90.5%). All other miRNAs were included in at least half of the folds/iterations, with miR-16-5p showing the lowest inclusion percentage at 75.5%. The variable age was also incorporated into the model for every fold (100%).

Another LASSO regression was performed as described above excluding the variable age, to solely look at the discriminative potential of the miRNA panel. This resulted in an apparent AUC of 0.797 (95% CI

[0.741, 0.852]) (Figure S5), and a median AUC of 0.707 after cross-validation (95% of cross-validation estimates fall between 0.464 and 0.893).

A subtype analysis for Luminal A-like and Luminal B-like patients was also performed, where the model based on the 20-miRNA panel and age was validated within these subgroups. In luminal A-like patients (n = 89), an apparent AUC of 0.765 (95% CI [0.698, 0.833]) was obtained (Figure S6), while in Luminal B-like patients (n = 21) an apparent AUC of 0.827 (95% CI [0.735, 0.920]) was obtained (Figure S7). No subtype analysis could be performed for HER2-driven or triple negative breast cancer due to low sample size within these groups.

Due to the high number of variables in our model, we also tried to reduce the number of dimensions by including a principal component analysis (PCA) of our 20-miRNA panel. The principal components (PCs) that explained 80% of the variance were included in the model (PC1-5). All miRNAs contributed within the different PCs, but the degree of contribution differed (Figure S8). PC1, PC2, PC3, and age were included at least 95% times in the regression model under cross-validation. The PCA-based AUC with LASSO regression was 0.689 (95% CI [0.623, 0.755]) and 0.646 (95% of cross validation estimated fall between 0.423 and 0.836) before and after cross-validation, respectively (Figure S9).

Comparative analysis of affected breast versus the contralateral breast

Another LASSO regression was performed to determine whether our 20-miRNA panel was able to differentiate between the affected breast and the contralateral breast (n = 112) within each breast cancer patient. An AUC of 0.591 (95% CI [0.518, 0.664]) was obtained, and after correction for overfitting through cross-validation, the AUC was 0.535 (95% of cross-validation estimates fall between 0.388 and 0.678) (Figure S10). Additionally, a comparison between the contralateral breast versus the healthy controls was performed, where an AUC of 0.689 (95% of cross-validation estimates fall between 0.381 and 0.875) was obtained after cross-validation (Figure S11).

Discussion

There is a great clinical need for new personalized methods that can detect breast cancer at an early stage[24]. One promising avenue is the identification of biomarkers within liquid biopsies. This study identified a 20-miRNA panel within NAF that can distinguish breast cancer patients from healthy individuals with an AUC of 0.70. Moreover, most miRNAs showed an increased expression, which is in line with their oncogenic function within breast cancer[25]. The only downregulated miRNA was miR-145-5p, a well-established tumor suppressor[26]. Despite the modest fold changes of individual miRNAs (ranging from 0.54–1.57), their combined discriminative power was evident. This miRNA panel was, however, not able to discriminate between the affected breast and the contralateral breast within one patient.

This study expands on previous literature on biomarkers in NAF. A comprehensive review of previous research on this topic was published by Patuleia et al.[17]. Most groups within this field have investigated proteins (individual and/or panels) for the detection of breast cancer[27]. Only one previous study examined ten miRNAs in NAF, including miR-145-5p, miR-155-5p, miR-181-5p, miR-16-5p, miR-205-5p, and miR-21-5p, that overlap with our panel[28]. In contrast to our data, they found no significant dysregulation of these miRNAs in breast cancer patients (n = 20; 18 invasive tumors and 2 ductal carcinoma in situ) compared to a non-cancer group (n = 12; benign disease). This is probably due to the low sample size, and/or because control subjects had benign disease rather than an absence of disease. It is of importance that future research will examine the influence of benign disease on miRNA expression, to make sure false positives are avoided when using miRNAs in NAF for breast cancer detection.

A few other studies evaluated miRNA panels in nipple discharge or ductal lavage samples. Zhang et al.[29] examined 3,100 miRNAs in nipple discharge (n = 21), of which six were significantly altered between patients with invasive breast cancer and benign controls. Two of these miRNAs, miR-3646 (strand designation not defined on miRbase) and miR-4732-5p, showed excellent discriminative power with an individual AUC of 1. Moreover, an upregulation of miR-155-5p was seen (cut-off at 1.5 fold), which we also saw in our data (FC = 1.67). Except for miR-155-5p, none of the other dysregulated miRNAs in the study by Zhang et al.[29] overlap with our miRNA panel. This could be due to the low sample size, or because nipple fluid aspiration differs from spontaneous nipple discharge[30], making comparison with our study difficult. In a second study, Do Canto et al.[31] measured 742 miRNAs in ductal lavage samples (n = 22), of which 17 showed differential expression between tumor and paired normal samples (contralateral breast). Of these 17 miRNAs, only one, miR-181a-5p, is also part of our panel. In contrast to Do Canto et al.[31], no significant upregulation of this miRNA was seen in our study between groups (including affected versus healthy, and affected versus contralateral). Again, this could be due to the difference in obtaining NAF and obtaining ductal lavage samples, but probably also because our fold changes were corrected for NAF characteristics[30]. Additionally, many studies have examined miRNA levels in blood, including the prospective Sister Study and mammography cohorts, which significantly contributed to the development of our target miRNA panel[32, 33]. Surprisingly, among the miRNAs included in our study based on these cohorts, only miR-145-5p showed significant downregulation, consistent with findings from the mammography cohort. This difference may be due to the measurement of miRNAs in blood rather than NAF. As part of our research, we have also gathered blood samples and will investigate these miRNA levels in a future study.

Comparison between the affected breast and the contralateral breast in unilateral breast cancer patients resulted in a low AUC of 0.535 (after cross-validation, Figure S10). This indicates that the miRNA expression pattern in the affected breast, containing the tumor, is highly similar to the contralateral (“healthy”) breast. Moreover, when we performed an analysis for the contralateral breast versus the healthy controls, an AUC of 0.689 was obtained after cross-validation (Figure S11). Even though this is a slightly lower discriminative value compared to the affected breast, it does indicate that the contralateral

sample is also able to distinguish healthy from diseased. This would allow tumor detection using either breast's NAF, thereby eliminating the need for successful nipple fluid collection from both breasts.

The similar miRNA patterns in both breasts suggests that we might be measuring a systemic response to the tumor, instead of local miRNA secretion by the tumor cells into the surrounding milk ducts. This hypothesis is supported by the knowledge that NAF can contain exogenous and endogenous components[17, 34, 35]. For example, Petrakis et al.[34] found nicotine in NAF of smokers. Moreover, Chatterton et al.[36] found altered hormone levels in NAF that were associated with oral contraceptives or hormone replacement therapy, indicating that hormones can diffuse into NAF too. Cross-talk is also possible in the other direction, from NAF into other biofluids, such as the lymphatic system. Petrakis et al.[34] identified that after injecting India ink into mammary ducts of rabbits, this ink moved out of the lumen and reached the lymphatics. Moreover, our group identified systemic side effects after cisplatin administration through the nipple in mice, indicating that local injection leads to systemic exposure[37]. Taken together, there is a clear cross-talk between NAF and the lymphatic and/or circulatory system, which could provide an explanation of why the affected and contralateral breast show similar miRNA expression levels.

Based on the hypothesis that a systemic response to breast cancer is being measured, it is possible that the miRNA levels in our data do not originate from the tumor cells. This does not influence the discriminative potential of our miRNA panel, since the miRNAs secreted by other cell types can also have diagnostic potential. One very likely source of the observed changes in miRNA expression are immune cells, known to be present in NAF[17]. Previous research has shown that multiple miRNAs from our panel are expressed and secreted by immune cells[38–42]. The extensively studied miR-155-5p is secreted by dendritic cells[38], macrophages[39] and T cells[40] in different *in vitro* and *in vivo* studies. Moreover, macrophages also release miR-21-5p[39] and miR-221-3p[41]. Similarly, miR-16-5p and miR-92a-3p were shown to be released by natural killer cells in 2D cell culture conditions[42]. All these miRNAs were upregulated in our study, although a significant fold change was only obtained for miR-155-5p, miR-21-5p and miR-29c-5p. This information suggests that multiple of the miRNAs detected in NAF may originate from immune cells rather than breast cancer cells. The positive side to measuring a systemic response, is that the origin of NAF from the specific tumor duct is not a requirement anymore. This point has been used to argue against using NAF as a biofluid, since different ducts can contain distinct components, and atypical cells are probably absent in ducts not affected by breast cancer[43–45].

In the present study, multiple variables showed an association with miRNA expression, thereby potentially affecting the discriminative value of the panel. For instance, NAF color is known to affect miRNA expression[19], but this can be partially addressed by correcting for NAF cloudiness, which is associated with color (Table S7). Although correction was partly possible, not all factors could be taken along in our analysis, mainly due to a high number of missing information and the risk of overfitting[46]. Moreover, many of these variables are known risk factors for breast cancer development, such as having a first-degree relative with breast cancer[47], and could therefore be predictors instead of covariates[48]. A

second limitation within our cohort is the skewed distribution of breast cancer subtypes with an overrepresentation of luminal B-like breast cancer and underrepresentation of triple negative breast cancer. Due to the sample size of the groups, it was not possible to perform subclass analysis using cross-validation.

While some limitations exist, miRNAs remain a promising biomarker for breast cancer detection evidenced by the AUC of 0.70. Additionally, the NAF collection method is significantly less uncomfortable compared to other screening options[16]. We envision multiple ways for implementing our new screening method, which would require integration with imaging techniques to accurately determine tumor location[17] (Figure S12). Firstly, our miRNA panel could be used for stratification, for example in women with dense breast tissue, women with increased risk of developing breast cancer, and/or women in the age group of 45–50 years that fall outside of the nationwide screening population. In women with dense breast tissue, where mammography is less accurate[12, 49], a positive liquid biopsy result with a high positive predictive value could stratify for additional MRI. Moreover, if our test has a high negative predictive value, it could reduce unnecessary follow-up after a false-positive MRI result. Furthermore, for the high-risk population that has to undergo breast cancer screening from a young age[5], a more accurate and less painful screening method would be especially beneficial. Secondly, our miRNA panel could be used for triage, where imaging techniques would only be indicated when our NAF miRNA panel is positive, reducing the burden and costs of current imaging methods[17]. Self-sampling NAF could be of great value here, and seems to be a realistic option in our pilot studies. Furthermore, this approach could possibly increase participation in breast cancer screening, particularly among women who avoid mammography due to physical discomfort[50]. Offering a less painful screening alternative may increase attendance and even shorten screening intervals, ultimately leading to better overall survival[7]. Future validation is required for clinical implementation, but we consider these first results very promising. Our group is currently analyzing miRNA levels in NAF samples of a longitudinal cohort of women with increased risk of breast cancer. This could provide insight into the timing of breast cancer detection in relation to current clinical and imaging based diagnosis, and potentially contribute to the early and less discomforting personalized detection of early breast cancer.

Conclusion

A 20-miRNA panel in NAF is able to discriminate between breast cancer patients and healthy controls with a good discriminative value. The similar miRNA patterns in both breasts suggest that we might be measuring a systemic response to the tumor, thereby broadening the diagnostic applicability of our approach. Taken together, with an AUC of 0.70 and minimal discomfort, our method shows great promise for future personalized screening applications. This strategy may be particularly beneficial for women with dense breasts, where a positive miRNA test could stratify for supplemental MRI. Nonetheless, further validation in larger cohorts is necessary.

Abbreviations

AUC Area under the curve

BMI Body mass index

ER Estrogen receptor

FC Fold change

HER2 Human epidermal growth factor receptor 2

IQR Interquartile range

miRNA microRNA

MRI Magnetic resonance imaging

NAF Nipple aspirate fluid

OR Odds ratio

PCA Principal component analysis

PC Principal component

PR Progesterone receptor

qPCR Quantitative polymerase chain reaction

ROC Receiver operating characteristic curve

Declarations

Ethical approval and consent to participate

This study was conducted in accordance with the Declaration of Helsinki and was approved by the Institutional Review Board of the UMCU and other participating hospitals (ORNAMENT: 16-337; Healthy controls: 12-495). Written informed consent was obtained from all participating women.

Consent for publication

Not applicable.

Availability of data and materials

All original code is available in this paper's supplemental information.

Declaration of competing interests

P.J. van Diest, E. van der Wall, C.B. Moelans, L.A. de Rooij and S. G. Elias are listed as inventors on a patent application filed by Utrecht Holdings that describes the method of the diagnostic test described in this study. All remaining authors have declared no conflict of interest.

Funding

This work was supported by A Sister's Hope; the Schumacher-Kramer Foundation and one private foundation via UMC Utrecht & Wilhelmina Kinderziekenhuis Foundation; and one private funder. The funders played no role in study design, data collection, analysis and interpretation of data, or the writing of this manuscript.

Author contribution

Conceptualization: C.B.M., P.J.v.D., E.v.d.W., L.A.d.R., S.I.S.P.; Formal analysis: B.B.L.P.d.V., S.G.E., L.A.d.R., C.B.M.; Funding acquisition: C.B.M., P.J.v.D., E.v.d.W.; Investigation: C.B.M., L.A.d.R.; Methodology: C.B.M., P.J.v.D., E.v.d.W, L.A.d.R., S.I.S.P.; Resources: S.I.S.P.; J.v.G., J.E.C.v.S., T.v.D., L.M.V., C.C.v.d.P., W.S., J.Z., M.D., A.S., M.P.; Supervision: C.B.M., P.J.v.D., E.v.d.W.; Visualization: L.A.d.R.; Writing – original draft: L.A.d.R.; Writing – review & editing: C.B.M., P.J.v.D., E.v.d.W, S.I.S.P., B.B.L.P.d.V., S.G.E., T.v.D., M.P.

Acknowledgements

The authors are very thankful to all the women who participated in our study. We would also like to thank our funders. The UMCU samples were collected by M. Voorhorst, A. Balfoort, S. Samuels, M. van Opdorp, M. van Dijk, and E. van Staveren, whom we would like to thank for their hard work. We would also like to acknowledge Ms. Izabela Drążkowska from University Clinical Centre, Gdansk, the nurse who provided care and assistance to the patients.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used copilot in order to receive suggestions aimed to improve readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229–63.
2. van Ravesteyn N, Kregting L, Heijnsdijk E, de Bruijn AE, de Koning H, Broeders M, Otten J, Verbeek A, Pijnappel R. National evaluation of breast cancer screening in the Netherlands: 2023 (XV) - The fifteenth evaluation report. 2023. Accessed 17 August, 2023. https://www.rivm.nl/documenten/National-evaluation-of-breast-cancer-screening-in-the-Netherlands-2023

3. Dutch Cancer Registry (NCR). IKNL. https://nkr-cijfers.iknl.nl/, Accessed on August 11 2023.
4. National Institute for Public Health and the Environment (RIVM.). https://www.rivm.nl/en/breast-cancer-screening-factsheet, Accessed on August 11 2023.
5. van Winkel SL, Samperna R, Loehrer EA, Drukker CA, Sessink K, Karssemeijer N et al. Breast cancer in women after withdrawing from an increased-risk MRI screening program. Eur Radiol. 2025.
6. Lubinski J, Kotsopoulos J, Moller P, Pal T, Eisen A, Peck L, et al. MRI Surveillance and Breast Cancer Mortality in Women With BRCA1 and BRCA2 Sequence Variations. JAMA Oncol. 2024;10(4):493–9.
7. Zuley ML, Bandos AI, Duffy SW, Logue D, Bhargava R, McAuliffe PF et al. Breast Cancer Screening Interval: Effect on Rate of Late-Stage Disease at Diagnosis and Overall Survival. J Clin Oncol. 2024:JCO2400285.
8. Aristokli N, Polycarpou I, Themistocleous SC, Sophocleous D, Mamais I. Comparison of the diagnostic performance of Magnetic Resonance Imaging (MRI), ultrasound and mammography for detection of breast cancer based on tumor type, breast density and patient's history: A review. Radiography. 2022;28(3):848–56.
9. Fitzjohn J, Zhou C, Chase JG. Critical Assessment of Mammography Accuracy. IFAC-PapersOnLine. 2023;56(2):5620–5.
10. Boita J, van Engen RE, Mackenzie A, Tingberg A, Bosmans H, Bolejko A et al. Validation of a candidate instrument to assess image quality in digital mammography using ROC analysis. Eur J Radiol. 2021;139.
11. van der Waal D, Abbey CK, Tetteroo E, Geertse TD, Smid-Geirnaerdt MJA, Sechopoulos I et al. Finding the optimal recall rate in breast cancer screening: results from the ROCS study. Eur Radiol. 2026.
12. Veenhuizen SGA, de Lange SV, Bakker MF, Pijnappel RM, Mann RM, Monninkhof EM, et al. Supplemental Breast MRI for Women with Extremely Dense Breasts: Results of the Second Screening Round of the DENSE Trial. Radiology. 2021;299(2):278–86.
13. Xiang W, Rao H, Zhou L. A meta-analysis of contrast-enhanced spectral mammography versus MRI in the diagnosis of breast cancer. Thorac Cancer. 2020;11(6):1423–32.
14. Orlando AAM, Clauser P, Zarcaro C, Ferraro F, Curatolo C, Marino MA et al. MRI Insights in Breast Imaging. Curr Med Imaging Former Curr Med Imaging Reviews. 2024;20.
15. Lee JM, Lowry KP, Cott Chubiz JE, Swan JS, Motazedi T, Halpern EF, et al. Breast cancer risk, worry, and anxiety: Effect on patient perceptions of false-positive screening results. Breast. 2020;50:104–12.
16. Patuleia SIS, Moelans CB, Koopman J, van Steenhoven JEC, van Dalen T, van der Pol CC, et al. Patient-centered research: how do women tolerate nipple fluid aspiration as a potential screening tool for breast cancer? BMC Cancer. 2022;22(1):705.
17. Patuleia SIS, Suijkerbuijk KPM, van der Wall E, van Diest PJ, Moelans CB. Nipple Aspirate Fluid at a Glance. Cancers. 2021;14(1).

18. Suijkerbuijk KPM, van der Wall E, Meijrink H, Pan X, Borel Rinkes IHM, Ausems MGEM, et al. Successful oxytocin-assisted nipple aspiration in women at increased risk for breast cancer. Fam Cancer. 2010;9(3):321–5.
19. Patuleia SIS, van der Wall E, van Gils CH, Bakker MF, Jager A, Voorhorst-Ogink MM, et al. The changing microRNA landscape by color and cloudiness: a cautionary tale for nipple aspirate fluid biomarker analysis. Cell Oncol. 2021;44(6):1339–49.
20. Patuleia SIS, van Gils CH, Oneto Cao AM, Bakker MF, van Diest PJ, van der Wall E et al. The Physiological MicroRNA Landscape in Nipple Aspirate Fluid: Differences and Similarities with Breast Tissue, Breast Milk, Plasma and Serum. International Journal of Molecular Sciences. 2020;21(22).
21. de Rooij LA, Mastebroek DJ, ten Voorde N, van der Wall E, van Diest PJ, Moelans CB. The microRNA Lifecycle in Health and Cancer. Cancers. 2022;14(23).
22. Overview of Medical-Scientific Research in the Netherlands (OMON), NL6031. The ORNAMENT study: A multicenter, crOss sectional, study to assess microRNA expression in Nipple Aspirated Fluid, blood and tuMor material in women with primary brEast caNcer compared with healthy conTrols, https://onderzoekmetmensen.nl/nl/trial/28422, Accessed on March 17 2025.
23. Overview of Medical-Scientific Research in the Netherlands (OMON), NL8987. Breast Cancer Biomarkers in Nipple Aspirate Fluid and Blood in Healthy Women, https://www.onderzoekmetmensen.nl/nl/trial/24678, Accessed on March 17 2025.
24. Li J, Guan X, Fan Z, Ching L-M, Li Y, Wang X et al. Non-Invasive Biomarkers for Early Detection of Breast Cancer. Cancers. 2020;12(10).
25. Loh H-Y, Norman BP, Lai K-S, Rahman NMANA, Alitheen NBM, Osman MA. The Regulatory Role of MicroRNAs in Breast Cancer. Int J Mol Sci. 2019;20(19).
26. Lv P, Zhang Z, Hou L, Zhang Y, Lu L, Wang C et al. Meta-analysis of the clinicopathological significance of miRNA-145 in breast cancer. Biosci Rep. 2020;40(9).
27. Pant A, Anjankar AP, Shende S, Dhok A, Jha RK, Manglaram AV. Early detection of breast cancer through the diagnosis of Nipple Aspirate Fluid (NAF). Clin Proteomics. 2024;21(1).
28. Qin W, Zhang K, Sauter ER. Exosomal miRNAs in nipple aspirate fluid and breast cancer. Translational Cancer Res. 2017;6(S8):S1304–10.
29. Zhang K, Zhao S, Wang Q, Yang H-S, Zhu J, Ma R. Identification of microRNAs in Nipple Discharge as Potential Diagnostic Biomarkers for Breast Cancer. Ann Surg Oncol. 2015;22(S3):536–44.
30. de Groot JS, Moelans CB, Elias SG, Hennink A, Verolme B, Suijkerbuijk KPM et al. Repeated Nipple Fluid Aspiration: Compliance and Feasibility Results from a Prospective Multicenter Study. PLoS ONE. 2015;10(5).
31. Do Canto LM, Marian C, Willey S, Sidawy M, Da Cunha PA, Rone JD, et al. MicroRNA analysis of breast ductal fluid in breast cancer patients. Int J Oncol. 2016;48(5):2071–8.
32. 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;15(3).

33. Kodahl AR, Lyng MB, Binder H, Cold S, Gravgaard K, Knoop AS, et al. Novel circulating microRNA signature as a potential non-invasive multi-marker test in ER-positive early-stage breast cancer: A case control study. Mol Oncol. 2014;8(5):874–83.
34. Petrakis NL. Physiologic, biochemical, and cytologic aspects of nipple aspirate fluid. Breast Cancer Res Treat. 1986;8(1):7–19.
35. Sartorius OW. The biochemistry of breast cyst fluids and duct secretions. Breast Cancer Res Treat. 1995;35(3):255–66.
36. Chatterton RT, Geiger AS, Mateo ET, Helenowski IB, Gann PH. Comparison of Hormone Levels in Nipple Aspirate Fluid of Pre- and Postmenopausal Women: Effect of Oral Contraceptives and Hormone Replacement. J Clin Endocrinol Metabolism. 2005;90(3):1686–91.
37. de Groot JS, van Diest PJ, van Amersfoort M, Vlug EJ, Pan X, ter Hoeve ND, et al. Intraductal cisplatin treatment in a BRCA-associated breast cancer mouse model attenuates tumor development but leads to systemic tumors in aged female mice. Oncotarget. 2017;8(37):60750–63.
38. Lindenbergh MFS, Koerhuis DGJ, Borg EGF, van 't Veld EM, Driedonks TAP, Wubbolts R et al. Bystander T-Cells Support Clonal T-Cell Activation by Controlling the Release of Dendritic Cell-Derived Immune-Stimulatory Extracellular Vesicles. Front Immunol. 2019;10.
39. Lan J, Sun L, Xu F, Liu L, Hu F, Song D, et al. M2 Macrophage-Derived Exosomes Promote Cell Migration and Invasion in Colon Cancer. Cancer Res. 2019;79(1):146–58.
40. Guay C, Kruit JK, Rome S, Menoud V, Mulder NL, Jurdzinski A, et al. Lymphocyte-Derived Exosomal MicroRNAs Promote Pancreatic β Cell Death and May Contribute to Type 1 Diabetes Development. Cell Metabol. 2019;29(2):348–e616.
41. Li X, Tang M. Exosomes released from M2 macrophages transfer miR-221-3p contributed to EOC progression through targeting CDKN1B. Cancer Med. 2020;9(16):5976–88.
42. Di Pace AL, Pelosi A, Fiore PF, Tumino N, Besi F, Quatrini L et al. MicroRNA analysis of Natural Killer cell-derived exosomes: the microRNA let-7b-5p is enriched in exosomes and participates in their anti-tumor effects against pancreatic cancer cells. OncoImmunology. 2023;12(1).
43. Bhandare D, Nayar R, Bryk M, Hou N, Cohn R, Golewale N, et al. Endocrine Biomarkers in Ductal Lavage Samples from Women at High Risk for Breast Cancer. Cancer Epidemiol Biomarkers Prev. 2005;14(11):2620–7.
44. Khan SA, Wiley EL, Rodriguez N, Baird C, Ramakrishnan R, Nayar R, et al. Ductal Lavage Findings in Women With Known Breast Cancer Undergoing Mastectomy. JNCI J Natl Cancer Inst. 2004;96(20):1510–7.
45. King BL, Love SM. The intraductal approach to the breast: raison d'être. Breast Cancer Res. 2006;8(2).
46. Harrell FE, Lee KL, Mark DB. Multivariable Prognostic Models: Issues in Developing Models, Evaluating Assumptions and Adequacy, and Measuring and Reducing Errors. Stat Med. 1996;15(4):361–87.

47. Ripping TM, Hubbard RA, Otten JDM, den Heeten GJ, Verbeek ALM, Broeders MJM. Towards personalized screening: Cumulative risk of breast cancer screening outcomes in women with and without a first-degree relative with a history of breast cancer. Int J Cancer. 2015;138(7):1619–25.
48. Obeagu EI, Obeagu GU. Breast cancer: A review of risk factors and diagnosis. Medicine. 2024;103(3).
49. Bakker MF, de Lange SV, Pijnappel RM, Mann RM, Peeters PHM, Monninkhof EM, et al. Supplemental MRI Screening for Women with Extremely Dense Breast Tissue. N Engl J Med. 2019;381(22):2091–102.
50. National Institute for Public Health and the Environment (RIVM). https://www.rivm.nl/en/documenten/monitoring-breast-cancer-screening-2023, Accessed on March 17 2025.

Figures

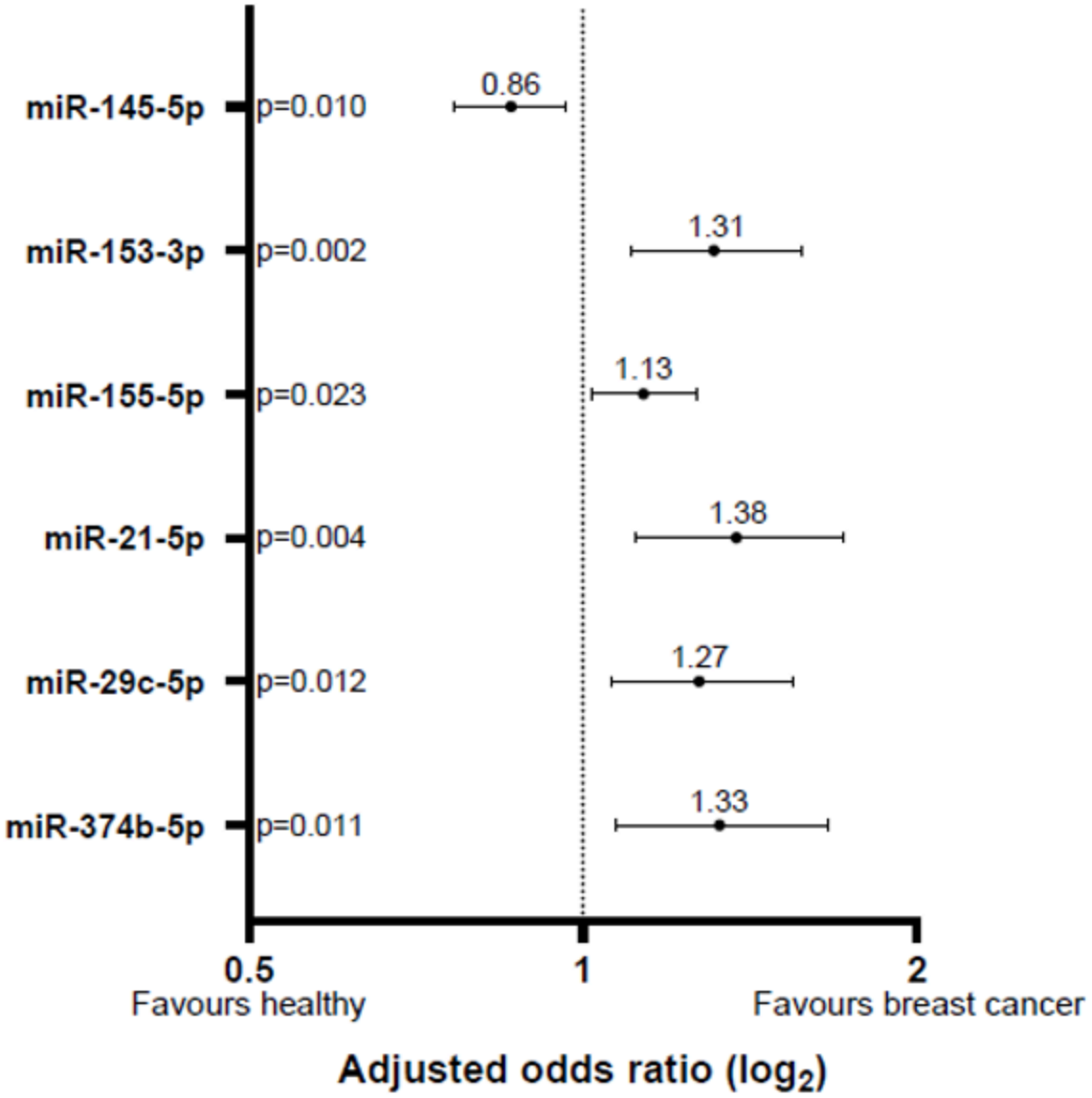


Figure 1

Breast cancer adjusted odds ratios for individual miRNA expression levels. *The y-axis indicate the significant miRNAs (P value given). The x-axis represents the adjusted odds ratio (logarithmic scale), with 95% confidence interval. This analysis was performed with 129 breast cancer samples (except for miR-145-5p that contained 1 missing) and 113 healthy control samples.*

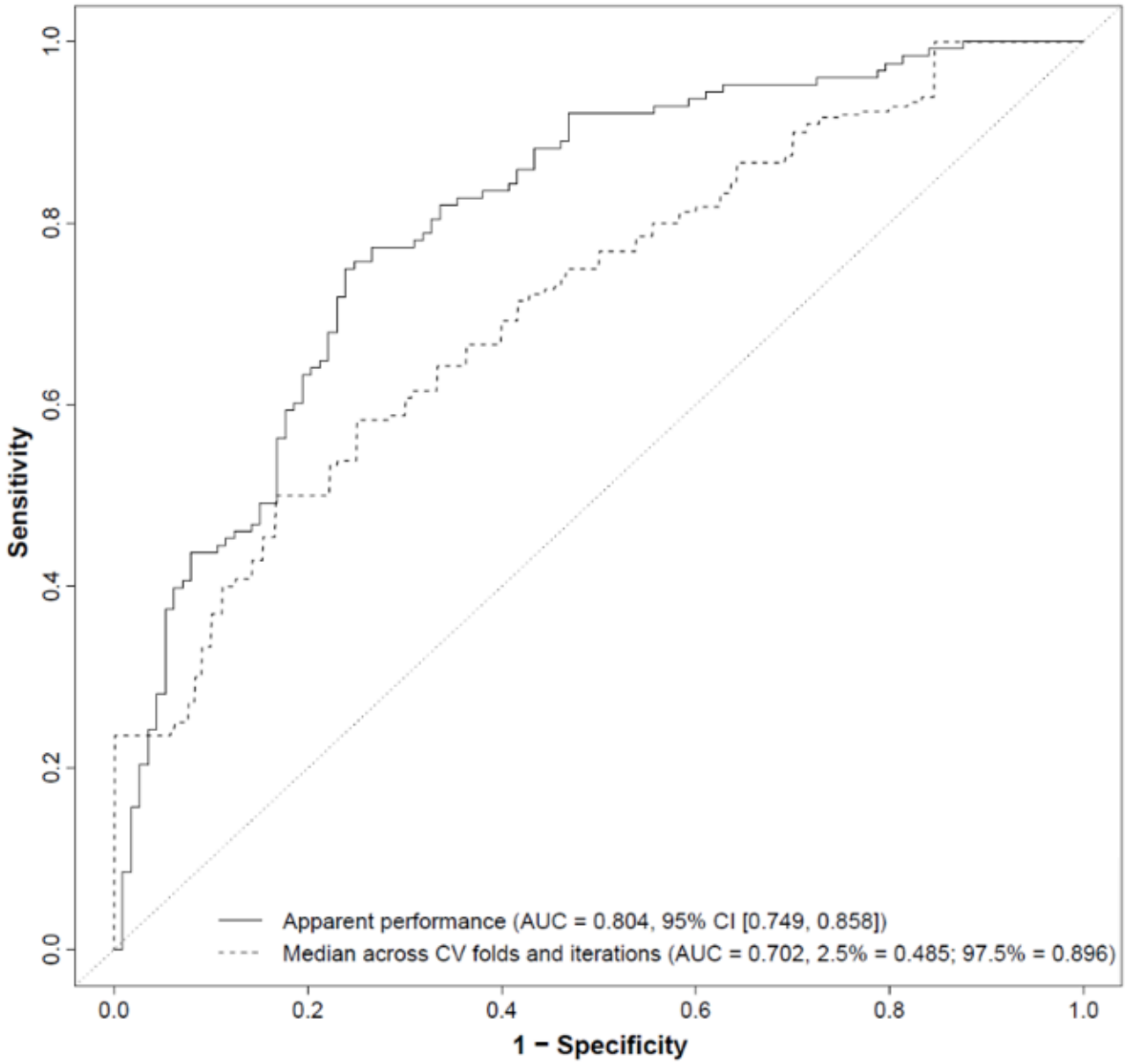


Figure 2

Receiver operating characteristics curve illustrating the performance of a 20-miRNA model to differentiate breast cancer patients and healthy controls. *This model also includes the variable age. The x-axis indicates 1 minus specificity, and the y-axis indicates sensitivity. The solid line indicates the receiver operating curve (ROC) curve of the LASSO regression model without cross-validation, while the dotted line shows cross-validation results. Area under the curve (AUC) values together with the 95% confidence interval of the apparent AUC and the 2.5th and 97.5th percentile of the cross-validated AUC are given in the right corner. This analysis was based on 128 breast cancer samples and 113 healthy control samples.*

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigure11.pdf
- Supplementaryfigure2.pdf
- Supplementaryfigure5.pdf
- Supplementaryfigure4.pdf
- Supplementaryfigure12.pdf
- Supplementaryfigure1.pdf
- Supplementaryfigure6.pdf
- Supplementaryfigure7.pdf
- Supplementaryfigure9.pdf
- Supplementaryfigure3.pdf
- Supplementaryfigure8.pdf
- Supplementaryfigure10.pdf
- Supplement1MaterialandMethods.docx
- Supplement3Questionnaire.docx
- SupplementaryMaterial.docx
- Supplement2AssociationmiRNAsandparitybreastfeeding.docx
- SupplementaryTable6.xlsx
- FinalscriptORNNAF2026.r