A Validation Study of a Multiple Reaction Monitoring-Based Proteomic Assay to Diagnose Breast Cancer

Yumi Kim 1, Un-Beom Kang 2, Sungsoo Kim 3,4, Han-Byoel Lee 5, Hyeong-Gon Moon 1, Wonshik Han 6, Dong-Young Noh 1

1Department of Surgery, Seoul National University College of Medicine, Seoul, Korea
2Daegu Gyeongbuk Institute of Science & Technology, Daegu, Korea
3Interdisciplinary Graduate Program in Genetic Engineering, Seoul National University College of Natural Science, Seoul, Korea
4Bertis Inc. Korea, Seongnam, Korea

ABSTRACT

Purpose: Currently, the standard screening tool for breast cancer is screening mammography. There have been many efforts to develop a blood-based diagnostic assay for breast cancer diagnosis; however, none have been approved for clinical use at this time. The purpose of this study was to determine the accuracy of a novel blood-based proteomic test for aiding breast cancer diagnosis in a relatively large cohort of cancer patients.

Methods: A blood-based test using multiple reaction monitoring (MRM) measured by mass spectrometry to quantify 3 peptides (apolipoprotein C-1, carbonic anhydrase 1, and neural cell adhesion molecule L1-like protein) present in human plasma was investigated. A total of 1,129 blood samples from 575 breast cancer patients, 454 healthy controls, and 100 patients with other malignancies were used to verify and optimize the assay.

Results: The diagnostic sensitivity, specificity, and accuracy of the MRM-based proteomic assay were 71.6%, 85.3%, and 77%, respectively; the area under the receiver operating characteristic curve was 0.8323. The proteomic assay did not demonstrate diagnostic accuracy in patients with other types of malignancies including thyroid, pancreatic, lung, and colon cancers. The diagnostic performance of the proteomic assay was not associated with the timing of blood sampling before or after anesthesia.

Conclusion: The data demonstrated that an MRM-based proteomic assay that measures plasma levels of three specific peptides can be a useful tool for breast cancer screening and its accuracy is cancer-type specific.

Keywords: Biomarkers; Breast neoplasms; Diagnosis; Blood proteins; Proteomics

INTRODUCTION

Breast cancer is a rapidly growing disease worldwide, with a 20% increase in 2012 compared with 2008 [1]. It is the second most common cancer among women in Korea after thyroid cancer. According to the Ministry of Health and Welfare’s National Cancer Registration Project, breast cancer accounted for 18.9% of all women’s cancers in 2015 [2]. Trends in female age-standardized cancer incidence from 1999 to 2014 revealed that the annual percentage change in most carcinomas tended to decrease except for breast cancer [3].
Breast cancer mortality is strongly influenced by increased rates of early detection and, to achieve this goal, most countries provide screening programs for breast cancer. Breast cancer screening is usually performed using mammography with or without clinical examination. In the West, breast cancer screening was implemented in the 1960s in response to the mortality reduction observed in randomized trials. A previous study by Humphrey et al reported that screening could reduce breast cancer mortality by 15%–30%\(^4\).

However, screening mammography also has many drawbacks. First, it is associated with hazards associated with radiation exposure, especially in younger populations\(^5,6\). Additionally, in women with dense breasts, due to the mammary gland, lesions such as breast cancer may be masked\(^7-10\). Furthermore, breast density among Korean women is different from their counterparts in the West. Up to 70% of Koreans and 40% of Westerners have dense breasts, and younger women tend to exhibit higher density\(^8\). Most studies addressing breast cancer screening have been conducted in the West. Therefore, it has been stressed that sufficient information about the benefits and risks of testing should be provided to make informed decisions for screening mammography\(^11,12\).

To summarize previous studies, early screening is important to improve the survival rate of individuals with breast cancer; however, current standard mammography has several limitations. Thus, we developed a blood test that can detect breast cancer using mass spectrometry to determine quantitative values of three specific proteins in the blood and substitute them into a proprietary algorithm\(^13\). The purpose of this study was to determine whether the blood-based diagnostic techniques we developed in previous studies are suitable for breast cancer screening and useful in clinical practice.

### METHODS

The 3 proteins selected in the previous study (carbonic anhydrase 1 [CAH 1], neural cell adhesion molecule LI-like protein [NCHL1], and apolipoprotein C-I [APOCI]) exhibited differences in concentration between breast cancer patients and healthy individuals\(^13\). Changes in blood concentration of the three proteins in breast and other cancers are presented in Supplemental Table 1\(^14-21\) and the characteristics of each protein are mentioned in Supplemental Table 2. Algorithm optimization and validation using new samples were performed using these three markers. Analysis was performed to confirm the stability of these proteins, their association with other cancers, and anesthesia.

### Patients and study design

A total of 1,129 blood samples (575 breast cancer patients, 454 healthy controls, 100 other cancers) were examined, including an assessment of their association with other types of cancer, and anesthesia. The number of samples in the validation set was calculated by the Medical Research Collaborating Center (MRCC) of Seoul National University Hospital (Seoul, Korea).

All previous experiments were performed using plasma samples, and additional tests were performed with 460 new serum samples (228 breast cancer, 232 healthy controls) not used in the experiments to assess the validity of using serum. At the same time, a total of 64 samples from patients who were prospectively recruited before and after anesthesia were used to confirm associations with anesthesia. The samples were obtained from individuals who visited the Seoul National University Hospital for treatment. Among these patients, those
who agreed to have tissues sampled before treatment were approved by the Institutional Review Board (IRB) of Seoul National University Hospital (IRB No. 1603-053-748); the study protocol adhered to the principles of the Declaration of Helsinki. Blood samples from healthy individuals who visited the Seoul National University Hospital Healthcare System Gangnam Center for routine health check-ups were also obtained. The three proteins selected as early breast cancer biomarkers at this institution are presented in Supplementary Table 1 and further described in Supplementary Table 2. These three biomarkers exhibited changes in concentration in those with breast cancer; however, no specific changes were observed in those with other cancers. Previous studies have identified changes in the blood levels of these three markers in other cancers.

Quantitative protein analysis

Quantitative analysis of the three proteins was performed using commercially available software (Analyst version 1.6, AB SCIEX, Framingham, USA) and reagent solutions (dithioerythritol [DTT], iodoacetamide [IAA], urea, and trypsin). A mass spectrometer (API 5000, AB SCIEX [medical device license No. Seoul, Korea: 10-1245]) was used to perform liquid chromatography with tandem mass spectrometry (LC-MS/MS) in multiple reaction monitoring (MRM) mode.

Blood sample pretreatment preparation

In a tube with 2 mg of synthesized standard peptide, 1 mL of 20% acetonitrile and 0.1% formic acid solution were added, the sample was vortexed for 5 minutes, and centrifuged for 30 seconds. The stock solution (2 mg/mL) was prepared by sonication in an ultrasonic grinder for 5 minutes and centrifuged at 12,000 rpm for 5 minutes. The working solution was diluted 1:50 from the stock solution and centrifuged at 12,000 rpm for 5 minutes. The working solution was diluted 1:50 from the stock solution and stored at −80°C or lower before use. The sample was lysed in lysis buffer (8M urea, 75 mM NaCl, 50 mM alkylation buffer [ABC]), and protein was quantified using the bicinchoninic acid assay to calculate the volume needed for 200 μg of protein (2.5–4 μL). To reduce cysteine disulfide bonds, 100 mM DTT was added. After the reaction, the sample was alkylated with 200 mM IAA to prevent disulfide bonding. The test sample was diluted 1:10 with 50 mM ABC. Trypsin (20 μg/400 μL) was added at a concentration of 1:50, and a digestion reaction was performed in a 37°C incubator for 16 hours. After trypsination, the sample was centrifuged at 4°C for 5 minutes, and 10% TFA was added to the sample to prevent further trypsination. After the sample was cleaned using a spin column (C18), the sample was lyophilized and stored at -80°C or lower before use. To generate a standard calibration curve, 10 pooled samples were mixed with 10 synthetic peptides (0.05–500 ng/mL), 5% acetonitrile (ACN), and 0.1% formic acid (FA). The mixture was pulverized in an ultrasonic mill for 5 min and centrifuged at 12,000 rpm at 4°C for 5 minutes. After transferring 80 μL to a vial, 2 μL of the solution was injected into the LC-MS/MS instrument. Data were analyzed using peak integration in the Analyst program.

Analytical performance tests

Analytical performance tests were performed using pooled plasma samples before sample analysis. The calibration curve was generated using the pooled plasma samples and linearity was evaluated. Quality control of the sample (low, mid, or high) and the lower limit of quantification (LLOQ) concentrations were then determined. From this concentration value, the specificity, precision, and accuracy were determined. Data were analyzed using peak integration and analyst software (AB SCIEX).
Sample analysis
Plasma samples from each group were subjected to a plasma pretreatment method and analyzed using MRM and analyst software. Subsequently, data were evaluated using the breast cancer biomarker algorithm.

RESULTS

Optimization of the proteomic assay
The diagnostic accuracy of the proteomic assay used to quantify levels of three proteins (i.e., CAH1, APOC1 and NCHL1) was optimized using plasma samples from 100 breast cancer patients and 100 healthy controls. After optimization, the sensitivity, specificity, and accuracy of the proteomic assay were 67.0% and 82.0%, and the area under the receiver operating characteristic curve (AUC) was 0.771, respectively (Supplementary Table 3). Based on this, the number of samples needed for the final confirmatory experiment was determined. This result from the optimization process was used as a basis for sample calculation in further validation studies.

Validation of the proteomic assay in an independent cohort
A total of 305 plasma samples, 183 from breast cancer patients and 122 from healthy controls, was used to validate the optimized algorithm. Using the proteomic assay, 131 of 183 breast cancer patients (71.6%) were classified with cancer, and 104 of 122 healthy controls (85.3%) were classified as normal. The positive predictive value was 82.9% and the negative predictive value was 75.0%. The statistic assumption of sensitivity and specificity for validation was satisfied.

Whether the diagnostic accuracy of the proteomic assay varied according to disease severity was also investigated. Among the 183 breast cancer patients, 45 exhibited axillary lymph node metastasis, all of which were pN1. The sensitivity of the proteomic assay in patients without lymph node metastasis was 71.7% and 73.3% in those with metastasis (Table 1). Regarding the molecular subtypes of breast cancer, the sensitivity for luminal A, luminal B, human epidermal growth factor receptor 2, and triple-negative breast cancers were 72.1%, 57.1%, 60.0%, and 73.9%, respectively (Table 1). In addition, the diagnostic accuracy of the proteomic assay when using serum samples from patients was investigated using an additional cohort of 460 serum samples (cancer 228, normal 232). The diagnostic accuracy of the proteomic assay was maintained, with an AUC of 0.8831.

Performance of the proteomic assay in patients with other types of malignancies and its association with anesthesia
To determine whether the performance of the proteomic assay described above was specific to breast cancer or applicable to other types of cancer, blood samples from patients who were diagnosed with other types of malignancies were collected. A total of 100 plasma samples from patients with thyroid, lung, colon, pancreatic, or ovarian cancers were analyzed using the proteomic assay. The proteomic assay failed to identify patients with other types of malignancies, suggesting that the accuracy of the assay was organ-specific and unique to patients with breast cancer (Table 2).

Because samples from the breast cancer patients were collected during the perioperative period, it is possible that the induction of general anesthesia may have affected the circulating levels of the three target proteins. To address this, the levels of the three proteins
before and after anesthesia were compared. Only CAH1 demonstrated a significant difference according to the use of anesthesia; the levels of the other two proteins (i.e., NCHL1, APOC1) were not significantly altered (Table 3).

| Characteristics | Cancer (n = 183) | Diagnosis of cancer by biomarkers (n = 132) | Diagnosed as normal by biomarkers (n = 51) | p-value |
|----------------|-----------------|---------------------------------|---------------------------------|---------|
| Age (yr)       | 51.92 ± 10.54   | 51.57 ± 10.77                   | 52.82 ± 9.97                    | 0.386   |
| AJCC stage     |                 |                                 |                                 | 0.035   |
| 0              | 17 (9.3)        | 13 (9.8)                        | 4 (7.6)                         |         |
| 1              | 91 (49.7)       | 60 (45.5)                       | 31 (60.8)                       |         |
| 2              | 73 (39.9)       | 57 (43.2)                       | 16 (31.4)                       |         |
| 3              | 1 (0.5)         | 1 (0.7)                         | 0 (0.0)                         |         |
| LN             |                 |                                 |                                 | 0.638   |
| pN0            | 137 (74.9)      | 98 (74.2)                       | 29 (56.9)                       |         |
| pN1            | 45 (24.6)       | 33 (25.0)                       | 12 (23.5)                       |         |
| ER             |                 |                                 |                                 | 0.334   |
| Negative       | 54 (29.5)       | 40 (30.3)                       | 14 (27.5)                       |         |
| Positive       | 126 (68.9)      | 89 (67.4)                       | 37 (72.5)                       |         |
| PR             |                 |                                 |                                 | 0.340   |
| Negative       | 72 (39.3)       | 50 (37.9)                       | 22 (43.1)                       |         |
| Positive       | 108 (59.0)      | 79 (59.8)                       | 29 (56.9)                       |         |
| HER2           |                 |                                 |                                 | 0.045   |
| Negative       | 134 (73.2)      | 97 (73.5)                       | 37 (72.5)                       |         |
| Positive       | 42 (22.7)       | 7 (5.3)                         | 5 (9.8)                         |         |
| Subtypes       |                 |                                 |                                 | 0.975   |
| HR+/HER2−      | 111 (60.7)      | 80 (60.6)                       | 31 (60.8)                       |         |
| HR+/HER2+      | 7 (3.8)         | 4 (3.0)                         | 3 (5.9)                         |         |
| HR−/HER2+      | 5 (2.7)         | 3 (2.3)                         | 2 (3.9)                         |         |
| TNBC           | 23 (12.6)       | 17 (12.9)                       | 6 (11.8)                        |         |
| Ki-67 (%)      |                 |                                 |                                 | 0.980   |
| > 10           | 49 (26.8)       | 36 (27.3)                       | 13 (25.5)                       |         |
| < 10           | 129 (70.7)      | 92 (69.7)                       | 37 (72.5)                       |         |
| CAH1           | 13.05 ± 16.25   | 16.10 ± 18.24                   | 5.21 ± 1.97                     | 0.001   |
| NCHL1          | 1.15 ± 0.74     | 1.19 ± 0.74                     | 1.07 ± 0.74                     | 0.972   |
| APOC1          | 19.38 ± 10.27   | 18.60 ± 10.16                   | 21.38 ± 10.40                   | 0.465   |

Values are presented as mean ± standard deviation or number (%).
AJCC = American Joint Committee on Cancer; LN = lymph node; ER = estrogen receptor; PR = progesterone receptor; HR = hormone receptor; HER2 = human epidermal growth factor receptor 2; TNBC = triple negative breast cancer; CAH1 = carbonic anhydrase 1; NCHL1 = neural cell adhesion molecule L1-like protein; APOC1 = apolipoprotein C-1.

| Biomarker result by cancer type | Thyroid cancer (n = 20) | Pancreatic cancer (n = 20) | Lung cancer (n = 20) | Colon cancer (n = 20) | Ovarian cancer (n = 20) |
|--------------------------------|-------------------------|----------------------------|----------------------|-----------------------|------------------------|
|                                | AJCC stage               |                            |                      |                       |                        |
|                                | 0                        | 1                          | 2                    | 3                     | 4                      |
| Thyroid cancer (n = 20)        |                          |                            |                      |                       |                        |
| Normal                         | 0                        | 11                         | 0                    | 6                     | 1                      |
| Cancer                         | 0                        | 1                          | 0                    | 1                     | 0                      |
| Pancreatic cancer (n = 20)     |                          |                            |                      |                       |                        |
| Normal                         | 0                        | 1                          | 11                   | 3                     | 5                      |
| Cancer                         | 0                        | 0                          | 0                    | 0                     | 0                      |
| Lung cancer (n = 20)           |                          |                            |                      |                       |                        |
| Normal                         | 0                        | 2                          | 2                    | 7                     | 9                      |
| Cancer                         | 0                        | 0                          | 0                    | 0                     | 0                      |
| Colon cancer (n = 20)          |                          |                            |                      |                       |                        |
| Normal                         | 1                        | 4                          | 5                    | 6                     | 2                      |
| Cancer                         | 0                        | 0                          | 0                    | 1                     | 1                      |
| Ovarian cancer (n = 20)        |                          |                            |                      |                       |                        |
| Normal                         | 0                        | 2                          | 2                    | 12                    | 0                      |
| Cancer                         | 0                        | 1                          | 0                    | 3                     | 0                      |

AJCC = American Joint Committee on Cancer.
DISCUSSION

The proteomic assay used in the present study is based on MRM of LC-MS/MS optimized for a multi-marker test. Mass spectrometry (i.e., LC-MS/MS) was used to determine quantitative values of the 3 markers (APOC1, CAH1, and NCHL1) present in human plasma and input them into a unique algorithm to determine the presence of breast cancer. Results revealed that the proteomic assay was an effective method for detecting breast cancer and demonstrated reproducibility, with minimal overlapping sensitivity to other types of malignancies.

It has been widely reported that ethnicity-related factors contribute to very large differences in the diagnosis rates of breast cancer. The diagnostic accuracy of mammography is low in women with dense breasts, which makes it a less effective screening tool for breast cancer [7]. Weigel et al. [22] reported the sensitivity of mammography performed in a screening program to be 79.9%, while the sensitivity for grade 1 was 100%, and 83.2% and 80.7% in grades 2 and 3, respectively. For grade 4 cancers, mammography demonstrated only 50% sensitivity. Carney et al. [23] found that the sensitivity of mammography in, 2223 confirmed cancers in mammograms of 463,672 patients tended to decrease as the breast tissue became more dense. In another large study by Rafferty et al. [24], involving 452,320 cases, approximately one-half of all women were classified with dense breasts. For breasts with low density, the positive predictive value of digital mammography is 5.1% compared with 3.8% for women with dense breasts. Because breast density is influence by ethnicity, the role of mammography is further limited in Korean women [25].

Prompted by the current limitations of screening mammography, there have been extensive efforts to develop novel diagnostic markers for breast cancer that can improve the effectiveness of mammography [26]. To date, however, no blood-based test has been validated in a large independent cohort. In this study, we demonstrated that simple, objective, and reproducible blood tests based on proteomic technology can be applied to Korean women to improve the early detection of breast cancer.

Our study, however, had several limitations. First, blood samples were collected at a single institution. Furthermore, we used samples stored in a biorepository, which effectively makes our study retrospective in nature and, as such, vulnerable to the inherent limitations of such a design. To assess the efficacy of the assay, therefore, a larger-scale prospective study is required.

In conclusion, a proteomic assay using plasma or serum samples from women can be an effective tool in aiding the early diagnosis of breast cancer in Korean women.
SUPPLEMENTARY MATERIALS

Supplementary Table 1
Changes in protein concentration in different types of cancer

Click here to view

Supplementary Table 2
Known functions of three proteins

Click here to view

Supplementary Table 3
Comparing parameters from each generated algorithm

Click here to view

REFERENCES

1. GLOBOCAN. 2012. Estimated cancer incidence, mortality and prevalence worldwide in 2012. Lyon: International Agency for Research on Cancer. http://globocan.iarc.fr/Pages/fact_sheets_cancer. Accessed September 1st, 2016.
2. Jung KW, Won YJ, Kong HJ, Lee ES; Community of Population-Based Regional Cancer Registries. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2015. Cancer Res Treat 2018;50:303-16.
3. Statistics Korea. 2019. KOREAN Statistical Information Service. Daejeon: Statistics Korea. http://kosis.kr. Accessed August 14th, 2018.
4. Humphrey LL, Helfand M, Chan BK, Woolf SH. Breast cancer screening: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2002;137:347-60.
5. Berrington de González A. Estimates of the potential risk of radiation-related cancer from screening in the UK. J Med Screen 2011;18:163-4.
6. de Gelder R, Draisma G, Heijnsdijk EA, de Koning HJ. Population-based mammography screening below age 50: balancing radiation-induced vs prevented breast cancer deaths. Br J Cancer 2011;104:1214-20.
7. Shin HJ, Ko ES, Yi A. Breast cancer screening in Korean woman with dense breast tissue. J Korean Soc Radiol 2015;73:279-86.
8. Kim SH, Kim MH, Oh KK. Analysis and comparison of breast density according to age on mammogram between Korean and western women. J Korean Radiol Soc 2000;42:1009-14.
9. Nam SJ. Screening and diagnosis for breast cancers. J Korean Med Assoc 2009;52:946-51.
10. Mandelson MT, Oestreicher N, Porter PL, White D, Finder CA, Taplin SH, et al. Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers. J Natl Cancer Inst 2000;92:1081-7.
11. Bleyer A, Welch HG. Effect of three decades of screening mammography on breast-cancer incidence. N Engl J Med 2012;367:1998-2009.
12. Welch HG, Passow HJ. Quantifying the benefits and harms of screening mammography. JAMA Intern Med 2014;174:448-54.
13. Lee HB, Kang UB, Moon HG, Lee J, Lee KM, Yi M, et al. Development and validation of a novel plasma protein signature for breast cancer diagnosis by using multiple reaction monitoring-based mass spectrometry. Anticancer Res 2015;35:6271-9.

14. Li J, Shi K, Sabet ZF, Fu W, Zhou H, Xu S, et al. New power of self-assembling carbonic anhydrase inhibitor: short peptide-constructed nanofibers inspire hypoxic cancer therapy. Sci Adv 2019;5:eaax0937.

15. Supuran CT, Winum JY. Carbonic anhydrase IX inhibitors in cancer therapy: an update. Future Med Chem 2015;7:1407-44.

16. Drummond F, Sowden J, Morrison K, Edwards YH. The caudal-type homeobox protein Cdx-2 binds to the colon promoter of the carbonic anhydrase 1 gene. Eur J Biochem 1996;236:670-81.

17. He LH, Ma Q, Shi YH, Ge J, Zhao HM, Li SF, et al. CHL1 is involved in human breast tumorigenesis and progression. Biochem Biophys Res Commun 2013;438:433-8.

18. Cohen M, Yossef R, Erez T, Kugel A, Welt M, Karpasas MM, et al. Serum apolipoproteins C-I and C-III are reduced in stomach cancer patients: results from MALDI-based peptidome and immuno-based clinical assays. PLoS One 2011;6:e14540.

19. Sun Y, Zhang J, Guo F, Zhao W, Zhan Y, Liu C, et al. Identification of apolipoprotein C-I peptides as a potential biomarker and its biological roles in breast cancer. Med Sci Monit 2016;22:1152-60.

20. Yang Y, Zhao S, Fan Y, Zhao F, Liu Q, Hu W, et al. Detection and identification of potential biomarkers of non-small cell lung cancer. Technol Cancer Res Treat 2009;8:455-66.

21. Takano S, Yoshitomi H, Togawa A, Sogawa K, Shida T, Kimura F, et al. Apolipoprotein C-1 maintains cell survival by preventing from apoptosis in pancreatic cancer cells. Oncogene 2008;27:2810-22.

22. Weigel S, Heindel W, Heidrich J, Hense HW, Heidinger O. Digital mammography screening: sensitivity of the programme dependent on breast density. Eur Radiol 2017;27:2744-51.

23. Carney PA, Miglioretti DL, Yankaskas BC, Kerlikowske K, Rosenberg R, Rutter CM, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. Ann Intern Med 2003;138:168-75.

24. Rafferty EA, Durand MA, Conant EF, Copit DS, Friedewald SM, Plecha DM, et al. Breast cancer screening using tomosynthesis and digital mammography in dense and nondense breasts. JAMA 2016;315:1784-6.

25. Youn I, Choi S, Kook SH, Choi YJ. Mammographic breast density evaluation in Korean Women using fully automated volumetric assessment. J Korean Med Sci 2016;31:457-62.

26. Núñez C. Blood-based protein biomarkers in breast cancer. Clin Chim Acta 2019;490:113-27.