Libertas Academica
FREEDOM TO RESEARCH

Open Access: Full open access to this and thousands of other papers at http://www.la-press.com.

Biomarkers in Cancer

An Antibody-based Blood Test Utilizing a Panel of Biomarkers as a New Method for Improved Breast Cancer Diagnosis

Galit Yahalom[1], Daria Weiss[2], Ilya Novikov[1], Therese B. Bevers[2], Laszlo G. Radvanyi[2], Mei Liu[2], Benjamin Piura[1], Stefano Iacobelli[3], Maria T. Sandri[4], Enrico Cassano[4], Tanir M. Allweis[5], Arie Bitterman[6], Pnina Engelman[7], Luis M. Vence[2] and Marvin M. Rosenberg[8]

[1]Eventus Diagnostics (Israel) Ltd., Ora, Israel. [2]University of Texas, MD Anderson Cancer Center, Houston, TX, USA. [3]Media Pharma Srl, Chieti, Italy. [4]Laboratory Medicine Division, Istituto Europeo di Oncologia, Milan, Italy. [5]Kaplan Medical Center, Rehovot, Israel. [6]Carmel Medical Center, Surgery A, Haifa, Israel. [7]Zvulun Breast Center, Clalit Organization, Israel. [8]Eventus Diagnostics Inc. Miami, FL, USA.

ABSTRACT: In order to develop a new tool for diagnosis of breast cancer based on autoantibodies against a panel of biomarkers, a clinical trial including blood samples from 507 subjects was conducted. All subjects showed a breast abnormality on exam or breast imaging and final biopsy pathology of either breast cancer patients or healthy controls. Using an enzyme-linked immunosorbent assay, the samples were tested for autoantibodies against a predetermined number of biomarkers in various models that were used to determine a diagnosis, which was compared to the clinical status. Our new assay achieved a sensitivity of 95.2% [CI = 92.8–96.8%] at a fixed specificity of 49.5%. Receiver-operator characteristic curve analysis showed an area under the curve of 80.1% [CI = 72.6–87.6%]. These results suggest that a blood test which is based on models comprising several autoantibodies to specific biomarkers may be a new and novel tool for improving the diagnostic evaluation of breast cancer.

KEYWORDS: breast cancer, autoantibodies, diagnostic testing, biomarkers

CITATION: Yahalom et al. An Antibody-based Blood Test Utilizing a Panel of Biomarkers as a New Method for Improved Breast Cancer Diagnosis. *Biomarkers in Cancer* 2013:5 71–80 doi:10.4137/BIC.S13236.

TYPE: Original Research

FUNDING: This work was supported in part by a grant from The University of Texas MD Anderson Cancer Center Duncan Family Institute for Cancer Prevention and Risk Assessment. This work was supported by Eventus Diagnostics (Israel) LTD.

COMPETING INTERESTS: Some of the authors of this article are employees of EventusDx, who funded this work.

COPYRIGHT: © the authors, publisher and licensee Libertas Academica Limited. This is an open-access article distributed under the terms of the Creative Commons CC-BY-NC 3.0 License.

CORRESPONDENCE: galit@eventusdx.com

Introduction

Traditional diagnostic tools for breast cancer detection, including clinical breast examination and mammography, are only moderately effective for accurately detecting early stage breast cancer. Mammography has limited sensitivity, a high rate of false-positive results, and cumulative radiation exposure as significant risk factors. The mean sensitivity of mammography has been estimated to be 77% (range: 29–97%),[1,2] with the rate of false-positive mammographic findings as high as 35%.[2,3] Ultrasound (US) and magnetic resonance imaging (MRI) are additional diagnostic tests for questionable lesions of the breast as well as potential screening modalities for high-risk women. There is a general consensus among oncologists and the public that there is an urgent and unmet need to develop more accurate, non-invasive, simple, and low-risk alternative modalities for the screening and diagnosis of breast cancer.[4]

It is generally accepted that there is a humoral immune response to intracellular or cell surface tumor-associated antigens (TAAs) released at the site of tumor genesis. With the development of new technologies, studies have profiled serum from cancer patients for the detection of autoantibodies (AAbs) to TAAs.[5,6] AAbs represent an attractive biomarker for diagnostic assays, principally due to the stability of immunoglobulins in cancer patient serum facilitating measurements with conventional assays. Expression levels of AAbs related to cancer are altered in cancer patients, whereas the disease does not alter other non-cancer related AAbs. Thus, the change in cancer-specific AAbs can indicate the presence or absence of

a specific cancer. This may be detectable well in advance of clinically detected disease using current conventional diagnostic techniques.[7] AAbs to TAAs could represent novel biomarkers for cancer screening, diagnosis, prognosis, monitoring, and prediction of response to chemotherapy. The challenge is how to measure and interpret these changes among cancer specific AAbs and develop an assay and algorithm for an accurate, low-risk tool for the diagnosis of cancer.

In breast carcinoma, as in other malignancies, the use of larger panels of TAAs, rather than individual TAAs, enhances the likelihood of accurately detecting cancer-associated AAbs with more accurate diagnostic value. Thus far, only a small number of circulating AAbs specific to breast carcinoma TAAs have been identified and investigated.[8,9] The most familiar are Her2[10] and Muc1,[11] both of which are known to be over-expressed in breast cancer tissues and involved in the production of specific autoantibodies. Current efforts to predict or diagnose breast cancer based on autoimmunity to either an individual TAA, or groups of TAAs, have so far not resulted in clinically applicable serologic biomarkers with accurate and definitive predictive and diagnostic capabilities. In this study, we tested a new enzyme-linked immunosorbent assay (ELISA)-based method for measuring the ratio of blood-based AAbs against a selected panel of breast TAAs for its diagnostic potential in distinguishing breast cancer patients from a cohort of healthy controls.

Materials and Methods

Study subjects and blood samples. All blood samples were obtained from female subjects over the age of 18 years with a breast abnormality detected by a clinical breast examination, mammogram, ultrasound, or breast MRI. Additionally, each subject must have had a final pathologic diagnosis of the breast abnormality determined by needle or surgical biopsy. Women were not eligible to participate if they had a previous or concurrent malignancy including hematologic malignancies, were receiving active chemotherapy, or had chemotherapy or steroid-based therapy in the past 6 months. Women undergoing immunosuppressive treatments or women with an autoimmune disorder were also excluded.

A positive breast cancer diagnosis was defined as women having invasive ductal or lobular carcinoma and ductal carcinoma in situ as verified by pathologic evaluation. Women diagnosed with either fibroadenoma, fibrocystic changes (including sclerosingadenosis, and benign papilloma), atypical ductal hyperplasia, atypical lobular hyperplasia and lobular carcinoma in situ were included in the control (healthy) group. As stated, to determine the clinical status of the women, a biopsy confirmation was required. The control group had negative biopsy samples and was not dependent upon imaging interpretations. In situations where there was a difference between the results of the needle biopsy and surgery, the pathologic findings at surgery overruled the needle biopsy results. Verification of a definitive breast cancer diagnosis was dependent on both imaging and pathology concordance. Patients with biopsy samples with no pathological report, or with no final diagnosis, were excluded from the study. A total of 546 samples were obtained from five centers worldwide (Israel, Italy, and the USA) (Table 1). All blood samples were collected with local Institutional Review Board (IRB) approval after each participant signed an informed consent. The trial registration ID is NCT00331942.

Plasma was collected from whole blood using heparin tubes (cat. No. 455084, Greiner Bio-One, Frickenhausen, Germany) centrifuged at 3,000×*g* for 10 min at room

Table 1. Study population information—number of samples collected at each site according to final diagnosis as verified by biopsy of the lesion. Patients were considered cases with either invasive cancer or DCIS and ADH, ALH, LCIS, and other lesions were considered to be healthy. Average age (sd) is shown according to diagnosis for all population and according to menopausal status.

DIAGNOSIS	PATIENT		HEALTHY				TOTAL
	INVASIVE BC	DCIS	OTHER LESIONS	ADH	ALH	LCIS	
Israel[1]	54		92				146
Italy[2]	82	12	127	4	4	1	230
USA[3]	46	7	104	7	5	1	170
Total	182	19	323	11	9	2	546
		201				345	
	PATIENT		HEALTHY				TOTAL
Average age (sd)	59.2 (13.4)		46.4 (11.6)				51.1 (13.8)
Average age (sd) pre-menopause	44.9 (5.4)		40.1 (8.8)				41.2 (8.4)
Average age (sd) post-menopause	66.6 (10.1)		58.2 (7.2)				62.8 (9.8)

[1]Carmel Medical Center, Haifa; Kaplan Medical Center, Rehovot.
[2]Mediterranean School of Oncology, Chietti (MSO), European Institute of Oncology, Milan (IEO).
[3]MD Anderson Cancer Center, Houston, Texas (MD).

Table 2. List of the 15 tumor-associated antigens used in the study.

ANTIGEN NUMBER	ANTIGEN CODE[a]	PEPTIDE SEQUENCE OR PROTEIN NAME	REMARKS	REFERENCE
1	1	IISAVVGI	Her2neu aa655-aa661	10
2	8	TAPLQPEQLQVFETLEEI	Her2neu aa389-aa406, ALH epitope	10
3	11	SGSGHGVTSAPDTR	Derived from Muc1 tandem repeat	11
4	12	HGVTSAPDTRPAPGSTAPPA HGVTSAPDTRPAPGSTAPPA HGVTSAPDTRPAPGSTAPPA HGVTSAPDTRPAPGSTAPPA	Derived from Muc1 tandem repeat	11
5	16	KAAELIPLHKLAAK	Derived from Cathepsin D, aa24-aa28 and additional stabilizing aa	12
6	18	NGTSFDIHYGSGSLSGYLS	Derived from Cathepsin D, aa135-aa152	12
7	19	VGFAEAAR	Derived from Cathepsin D, aa494-aa451	12
8	80	Endostatin		13
9	41	RPA32		14
10	76	HOXB7		15
11	78	90K		16
12	114	CEA human		17
13	95	Cathepsin D		12
14	85, 103	Erbb2		10
15	115, 116	P53		18

[a]Coding was used to ensure a blinded analysis at each site.

temperature (RT), and aliquots were stored frozen at −80°C until ELISA analysis. At the MD Anderson Cancer Center only, plasma was centrifuged at 1,300 RPM for 30 min at 4°C and aliquots were stored frozen at −80°C until ELISA analysis. Data forms were completed by each site to obtain clinical information and final pathological diagnosis.

Antigen selection for AAb assay. Antigens were chosen from the current literature according to their known involvement in the humoral response against breast cancer (Table S1 in Supplementary Data on-line—Details of antigens used in this study). An initial set of 15 different antigens, all showing the ability to elicit antibody production in breast cancer patients (and some, to a smaller extent, in healthy populations as well) were chosen for initial testing (Table 2). All proteins and peptides were purchased from different suppliers (Table S1 in Supplementary Data on-line). Each antigen was calibrated with specific antibodies for best-coating concentration.

ELISA methodology. ELISA was used to measure the humoral immune response in the serum or plasma of participating women to the various peptides or whole proteins antigens (Table 2). At each location, a specific standardized ELISA protocol was followed (described below) on local samples to ensure assay consistency across sites. Each sample was given a barcode identifier in the laboratory to ensure a blinded analysis. White "Maxiorp" 96 wells plates (Nunc, Roskilde, Denmark) were coated with commercial antigens at concentrations ranging 2–6 µg/mL for proteins, and 0.25–1 mg/mL for peptides in phosphate-buffered saline (PBS) and blocked with Well Champion reagent (Kem-En-Tec, Taastrup, Denmark) according to the manufacturer's instructions. Serum or plasma samples (100 µL) were loaded in 6 serial dilutions starting at 1:40–1:320 in 1% skim milk in PBS (Fluka, St. Louis, MO, USA) for each of the coated antigens in the plates and incubated at 37°C for 1.5 h with gentle agitation. The plates were washed 8 times with 300 µL of Dulbecco's PBS, 0.05% Tween 20 (PBST), and 1:10,000 horseradish peroxidase conjugated goat anti-human IgG (Chemicon, Temecula, CA, USA) was added for 1 h at 37°C, followed by 4 washes with 0.025% PBST. EZ-ECL (Biological Industries, Beit-Haemek, Israel) was used for luminescent development according to the manufacturer's instructions. Luminescence was measured with Luminoscan Ascent (Thermo Scientific, Waltham, MA, USA) using Ascent software (Thermo Scientific). Results were loaded into an internet database in a secure server according to the barcodes.

Statistical methods. All statistical analyses were performed using STATA 12 SE (StataCorp, College Station, TX, USA). All *P*-values were two-sided. *P*-values below 0.05 were considered significant. No corrections for multiple comparisons were performed. The initial data for each sample consisted of 6 measurements of AAb relative luminescence units (RLU)

for each antigen in 6 consecutive dilutions. In the first step, we fitted the $\log_{10}$[RLU] as a linear function of the $\log_{10}$[dilution]. If the goodness of fit was not satisfactory, we excluded one outlier and refit the data by linear regression with the remaining 5 points. Reference values of the dilution were fixed for each AAb. If the goodness of fit was high, the fitted value at the reference point was calculated for each AAb. Otherwise the value was classified as "missing" for this antigen only, meaning that the data did not pass the quality control. A missing value was only given to a specific antigen and not to a sample.

Thus, each sample was left with a set of maximum 15 values of AAb $\log_{10}$[RLU] for all antigens, each at the reference dilution point chosen for the antigens. We used the data as described to develop a classification scheme for all the samples into "patients" and "controls" as follows. First, we applied a main effect logistic regression for all possible models that included age, and 4 antigens out of the 15. Each model consisted of samples with full data only (all 4 antigens present). The models were sorted according to the sensitivity at 50% specificity, conditioned upon the fact that the model can be applied for a sufficient number of the samples (no less than 80 samples per model). Next, we established a combined decision rule whereby for each sample, the final decision as to "patient" or "control" was accepted according to the highest ranked model that could be used (ie, that all antigens in the highest sorted model were simultaneously "not-missing" for this sample, otherwise, the next highest model, with all "not-missing" values was applied to this subject).

Results

Theoretical considerations of the assay and data analysis approach. Current diagnostic methods generally rely upon observing one TAA against which the amount of AAbs in patients is higher than in controls. Such a method uses a "cut-off" criterion with subjects above the cut-off designated as "patients" and those below the cut-off designated as "healthy". This premise is typically true for external antigens such as bacteria and viruses. When an individual is infected, there is an immune response and a specific antibody response. In such a scenario, using a specific cut-off to score positive or negative or "infected" or "uninfected" is applicable.

However, when examining AAbs, the situation is different because AAbs are found in serum in the absence of overt disease among all populations. The constitutive or "natural" levels of AAbs differ among individuals, which has no correlation to specific diseases. Using a cut-off criterion for AAbs will result in a distortion of the diagnostic results, as many false-positives (those with high amounts of AAbs), and false negatives (those with low amounts of AAbs), will occur. An example is shown in Figure 1A. Alternatively, if absolute values are not considered and if the ratio in the amount of cancer-specific AAbs relative to the presence of non-cancer specific AAbs is calculated, a more accurate distinction can be made between patients and healthy subjects (Fig. 1B). As illustrated in Figure 1B, to determine and analyze the ratio between normal AAbs and cancer-specific AAbs, at least two AAbs should be used. This would include one normal occurring AAb unique to the individual (AAbA), and a second cancer-specific AAb (AAbB). Comparing the amount of the "non-relevant" normal occurring AAbs (AAb A) to the amount of cancer-related AAbs (AAb B), whose amounts are higher than the normal amounts of AAbs (AAb A), produces the following decision rule: a cancer patient is defined when AAb B > AAb A, and a healthy individual is defined when AAb B < AAb A. Using this method prevents both false-positive and false-negative results compared to the cut-off method, as it eliminates the misrepresentation created by absolute values.

Initial sample analysis using logarithmic transformation of ELISA data. The first objective of the study was to obtain reliable measurements of the amounts of each breast cancer specific AAb for its corresponding breast cancer specific

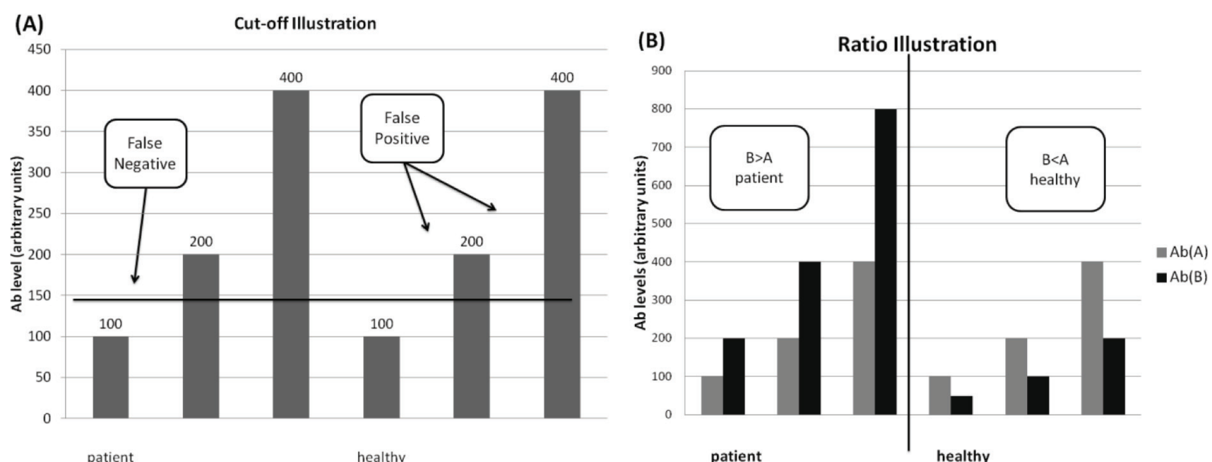

Figure 1. (**A**) Using "cut-off" criteria (AAb = 150) results with false-positive and false-negative results because different immune systems have different AAb levels. (**B**) Measuring relative amounts of antibodies (AAb A relative to AAb B) eliminates both false-positive and false-negative results that emerge when using a "cut-off" criterion. Using a ratio of AAb B > AAb A as a criteria for patients eliminates all false results.

antigen in each blood sample. We did not use a standard approach of averaging several measurements at a fixed dilution because, as described above, each woman has specific and different initial levels of AAbs. Alternatively, using a serial dilution approach with different starting dilutions for each antigen eliminates this problem. With this serial dilution approach, a curve of RLU measurements as a function of the dilution was obtained. The range of dilutions was chosen in a manner that for most women and antigen combinations, at least 5 measurements could be approximated to a linear curve (in logarithmic scale).

Following logarithmic transformations, those antigens that were saturated following a serial dilution were classified as an antigen with a missing value. The resultant curve was approximated to a linear curve using linear regression after excluding potential outliers. Applying this regression resulted with an estimate of log[RLU] at a pre-defined fixed dilution for each antigen. In cases where the quality of the linear fitting was not satisfactory (R^2 of the linear line was lower than 0.95), this predicted value was removed from the analysis and this antigen was assigned with a missing value. However, this resulted with the exclusion of only 7.1% of the samples (39/546 cases). A blood sample was qualified for inclusion in the study only when it showed detectable cancer specific AAb levels against all antigens included in at least one of the different models, which was 92.9% of the samples (507/546 cases).

Figure 2 is an example representing the analysis of the raw data. As seen in Figure 2, each raw data set was transformed into log-log scale, and linear regression was applied. The new line replaced the original, and the middle of the line (corresponding to dilutions between 1:160–1:320) was chosen as the final value for each antigen-women pair of data (final value, shown in Supplementary Data online Table S2—Data after smoothing procedure for all antigens sorted by samples participating in the study). For each linear regression, R^2 was calculated and data generating lines with $R^2 < 0.95$ were omitted from further analysis (for example, antigen 016 [$R^2 = 0.85$]).

This smoothing analysis was performed for all samples. After applying this analysis, significant differences between the two populations (healthy and breast cancer patients) were not observed for any of the antigens, as shown in the box graph in Figure 3. These results indicate that a simple average RLU determination of the data after logarithmic

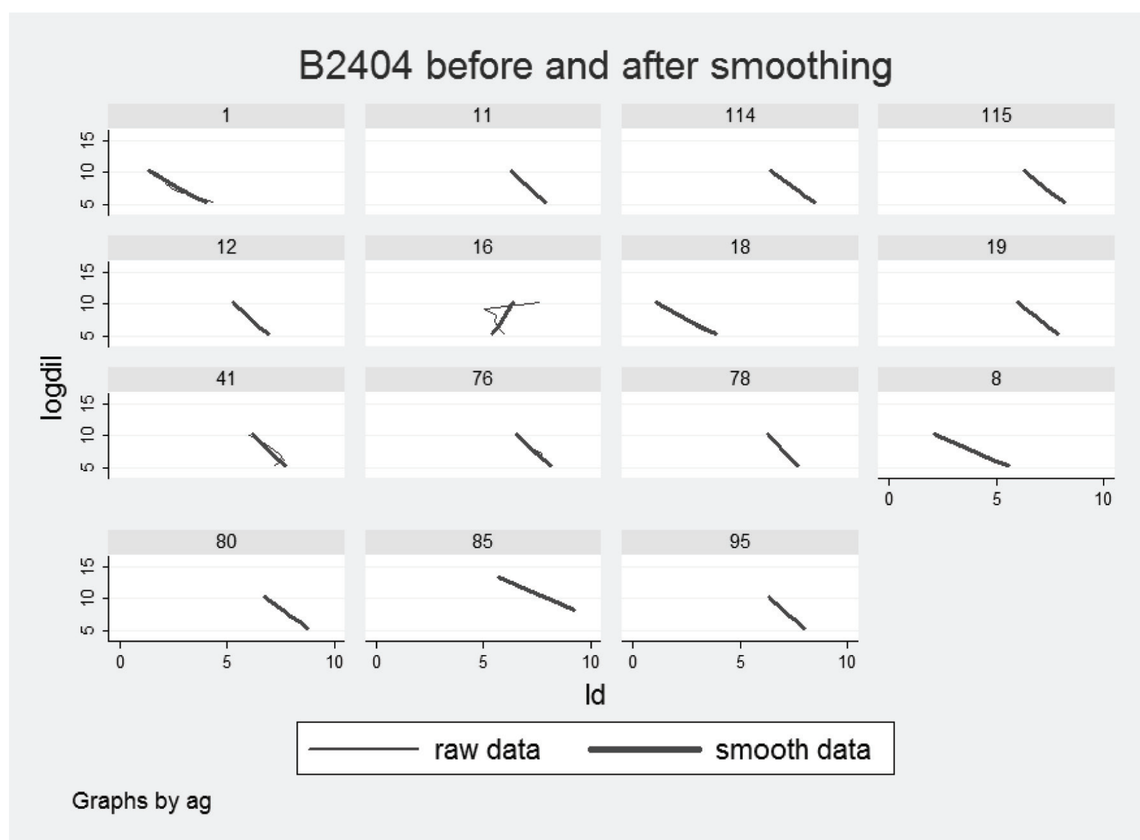

Figure 2. An example of the smoothing procedure. Each graph shows data corresponding to antigens of sample B2404, with the raw data shown as the thin line and data after smoothing as the thick line. In most cases, the raw data dilution curves yielded high linear correlation ($R^2 > 0.95$). When data could not be replaced by a straight line with good fitting (such as antigen #16 and antigen #41), the specific antigen was replaced by a missing value for this specific sample. All other antigens of the sample received a value corresponding to the value at a specific reference point of dilution in the middle of the theoretical line.

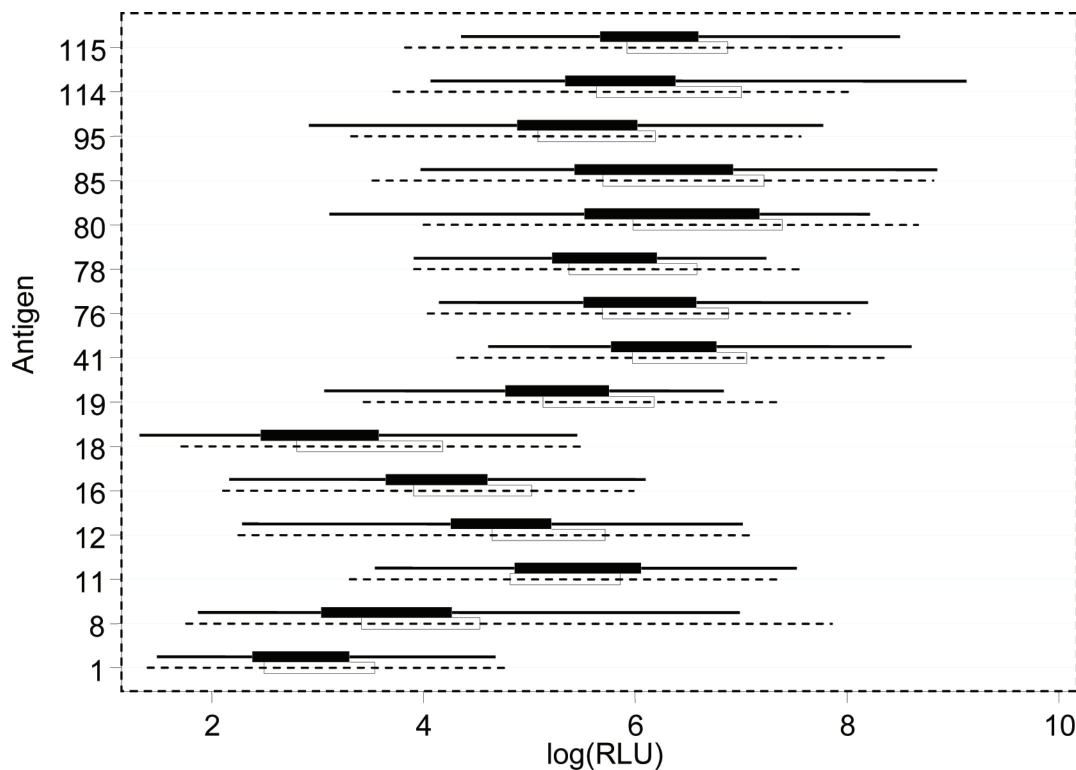

Figure 3. Box plots of average of $\log_{10}$ [RLU] of all antigens after the smoothing procedure.The two clinical groups are represented in the graph are breast cancer (filled bars) and healthy (empty bars). No statistically significant separation could be achieved between the groups using any one of these antigens.

transformation and smoothing did not reveal differences between patients and healthy controls. We therefore applied a more sophisticated method of data analysis using the ratio concept and "separation models" based on relevant clinical, demographic, or epidemiological parameters.

Classification of different samples. Before starting this second-tier analysis, we evaluated which relevant clinical or demographic parameter can be incorporated into the analysis. We checked the following parameters: age, menopausal status, and familial breast cancer history (as given by the patients verbally). We tested the distribution of these parameters between the two groups. Using the Mann-Whitney test for age gave a $P < 0.0001$, and Fisher's exact test for menopause ($P < 0.001$) and for family history ($P = 0.005$) (see Table S7 B-D in Supplementary Data online Analysis of clinical variables as "stand alone" predictors, for detailed analysis). We only used age for the entire population and performed a separate analysis for post-menopausal women. We did not use the family history parameter because this notion could not be rigorously defined, making it less reliable (the information is not always available to the subjects) and less significant. We also performed the logistic regression of the outcome (health status) on age and menopause. In this analysis, only age retained its significance ($P < 0.001$), while menopause became non-significant ($P = 0.076$) after age adjustment (see Table S7-A in Supplementary Data online—Analysis of clinical variables as "stand alone" predictors, for detailed analysis).

To further use the AAbs results to discriminate between patient samples and control samples, we used logistic regression of the disease status ("patient" or "control") on age and 4 antigens testing all possible combinations of 4 antigens out of 15. A classification model is defined as the set of antigens, as well as clinical data (age), and their corresponding coefficients obtained after logistic regression is performed. All sub-sets of theoretical combinations of the antigens (ie, all classification models) were tested for their sensitivities at the level of 50% specificity. Models created with at least 80 samples, resulting in a specificity of at least 50%, were ranked according to their sensitivities. For final classification, only models with sensitivity of at least 95% were used, after sorting according to the sensitivity and the area under the Receiver-operator characteristic (ROC) curve. The set of final models used for further analysisis shown in Table S3 (Supplementary Data online—List of the final models used for separation). After the set of models was established, the overall sensitivity was calculated as follows: Each sample was assigned a final diagnosis according to only one model, which was the highest model that could be assigned to the sample, ie, the highest model with no missing value for either of the antigens in the model for this specific sample, and was given a calculated value of "1" to indicate a positive or a "0" as a negative. After assignment, a comparison between biopsy diagnosis and the calculated result was conducted. A true positive (TP) was a sample with "patient" biopsy diagnosis designated

as a "1" in the test results. True negative (TN) was a sample with "control" biopsy diagnosis designated as a "0" in the test results. A false positive (FP) was defined as a sample with "control" biopsy diagnosis and "1" in the test results, while a false negative (FN) was defined as a sample with "patient" biopsy diagnosis and "0" in the test results. Some of the samples (n = 37) had too many missing values, and could not be applied to any of the models used (ie, each had less than 4 values in any of the models with high sensitivity). Those samples could not be assigned final test results and were not a part of the final analysis. The overall sensitivity of the test was then determined as the highest sensitivity with at least 50% specificity. Results of this analysis are shown in Supplementary Data online (Table S4—Prediction given to each sample after applying the models).

The total number of blood samples used from all sites was 546, which included 201 "patients", according to their final positive breast cancer diagnosis. In total, 345 healthy "controls" were used. The classification models each containing 4 antigens and age, were sorted according to the area under the curve (AUC). The final decision was according to 16 models with sensitivity above 95% at fixed specificity of 50% (models shown in Table S3 in the Supplementary Data online). Of the 546 samples, 507 showed definitive diagnostic results (final classification as well as the model used for each sample is shown in Table S4 in the Supplementary Data online). Of the 507 women with definitive diagnostic results, 339 were classified as positive ("1") and 168 as negative ("0"). When compared to biopsy diagnosis, 177 samples were true-positive, 159 true-negative, 162 false-positive and 9 false-negative. Thus, the sensitivity of this set of 507 samples was 95.2% and the specificity was 49.5% (Table 3), the calculated AUC of the ROC curve was 80.1% (Fig. 4A).

Effect of menopausal status. To explore the effect of menopausal status on the algorithm, we divided the population according to menopausal status that was given by the women when the sample was obtained. Of the 238 women reported to be post-menopausal, 131 were "patients", and 107 were "controls" (total of 238). Using this subpopulation only, new models were created using 4 antigens and age (see Table S5 in the Supplementary Data online—List of the final models used for separation for post-menopausal women). Of the 238 samples in the data set, only 193 samples remained with non-missing values, and resulted in 96.2% sensitivity and 52.8% specificity (Table 3). The total AUC for this sub-population was 84% (Fig. 4B). Final classification as well as the model used for each sample is shown in Table S6 (Supplementary Data online—Prediction given to each sample after applying the models for post-menopausal sub population).

Using the method for clinical status prediction. Our objective was to further validate the method in order to predict the status of blinded samples. To achieve this objective, we utilized the largest subset that contained the same antigens and had no missing values. A total of 252 samples, with 143 patients and 109 healthy controls, all shared the same 4 antigens (Antigens no. 016; 080; 095; 115). We divided the set into two separate groups, a training set containing 94 patients and 110 healthy controls, and a prediction set containing 15 patients and 33 healthy controls. We used only one model to establish a cutoff point for separation between the groups in the training set and applied separating criteria on the prediction subset. The training set, tuned to 94.7% sensitivity and 61.8% specificity, resulted in a cutoff point of 0.4, above which the subject was considered as a patient. This cutoff criteria was applied to the prediction set, giving a sensitivity of 100% and specificity of 45.4%, as shown in Table 4 and in Table S8 (Supplementary Data online—Blinded samples prediction using a single model). The ROC for this subset is shown in Figure 5, for the training set (Fig. 5A) and the predictions set (Fig. 5B).

Discussion

In this study, we tested 546 samples using ratios between different AAbs to a panel of biomarkers for each sample rather than using traditional cut-off thresholds for AAbs. Using this approach we differentiated between healthy controls and individual breast cancer patients with high sensitivity (>95%) and moderate specificity (up to 50%) with a total AUC of 0.8. The results were slightly better for a sub-population of post-menopausal women (AUC = 0.84). The subset analysis of post-menopausal women suggests that this approach could be particularly useful for identifying patients with breast cancer in this population. Increasing the number of age-similar controls, however, is needed to further test this possibility. The results achieved with a smaller and more homogenous subset, used for blinded samples clinical status predictions, is further validation of the proof of concept of this method, which will allow blind predictions using a single logistic regression-based model.

Table 3. Summary of tests results compared to clinical status.

WHOLE POPULATION	TEST NEGATIVE	TEST POSITIVE	
Clinical negative	159	162	Specificity = 49.5%
Clinical positive	9	177	Sensitivity = 95.2% (CI = 92.8–96.8)
Post-menopausal			
Clinical negative	47	42	Specificity = 52.8%
Clinical positive	4	100	Sensitivity = 96.2% (CI = 92.6–98.5)

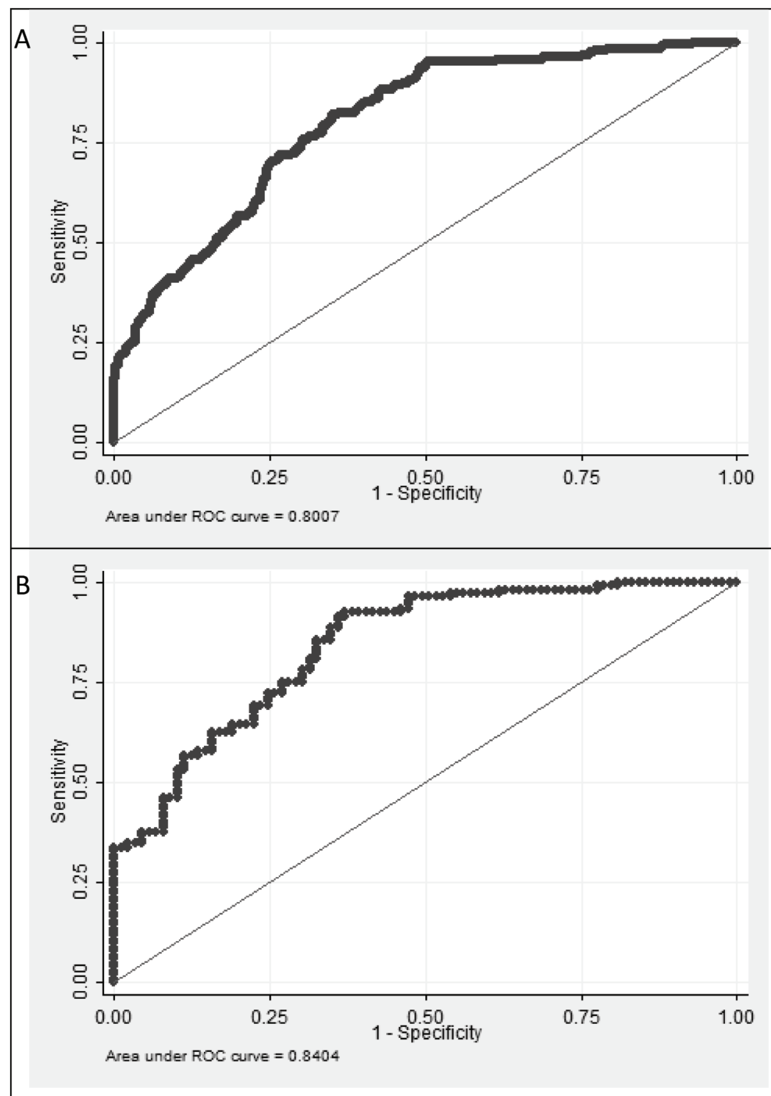

Figure 4. (**A**) ROC curve (sensitivity versus 1—specificity) of the 507 samples in the data set. The AUC is 80.1% (CI = 72.6%–87.6%). At specificity of 49.5%, sensitivity is 95.2% (CI = 92.8–96.8%). (**B**) ROC curve (sensitivity versus 1—specificity) of the 193 samples in the data set of post-menopausal women. The AUC is 84% (CI = 66.1–93.4%). At specificity of 52.8%, sensitivity is 96.2% (CI = 92.6–98.5%).

These results also suggest that better separations can be achieved for pre-defined sub-groups of either the malignant or pre-malignant population (such as atypical ductal or lobular hyperplasia, etc). Other sub-populations, such as high-risk and different ethnic origins, may also be suitable for specially designed diagnostics, with dedicated antigens used in the ratio approach. This hypothesis could not be tested in this study, as a much larger sample size of these sub-populations would be needed in order to achieve statistically significant outcomes. In order to further validate the use of this method for clinical status predictions on larger sample sizes, and eliminate heterogeneity problems of the system, a better and

Table 4. Summary of tests results for blind predictions.

TRAINING SET	TEST NEGATIVE	TEST POSITIVE	
Clinical negative	68	42	Specificity = 61.8%
Clinical positive	5	89	Sensitivity = 94.7% (CI = 88.0–98.3)
Prediction set			
Clinical negative	15	18	Specificity = 45.4%
Clinical positive	0	15	Sensitivity = 100.0% (CI = 78.2–100.0)

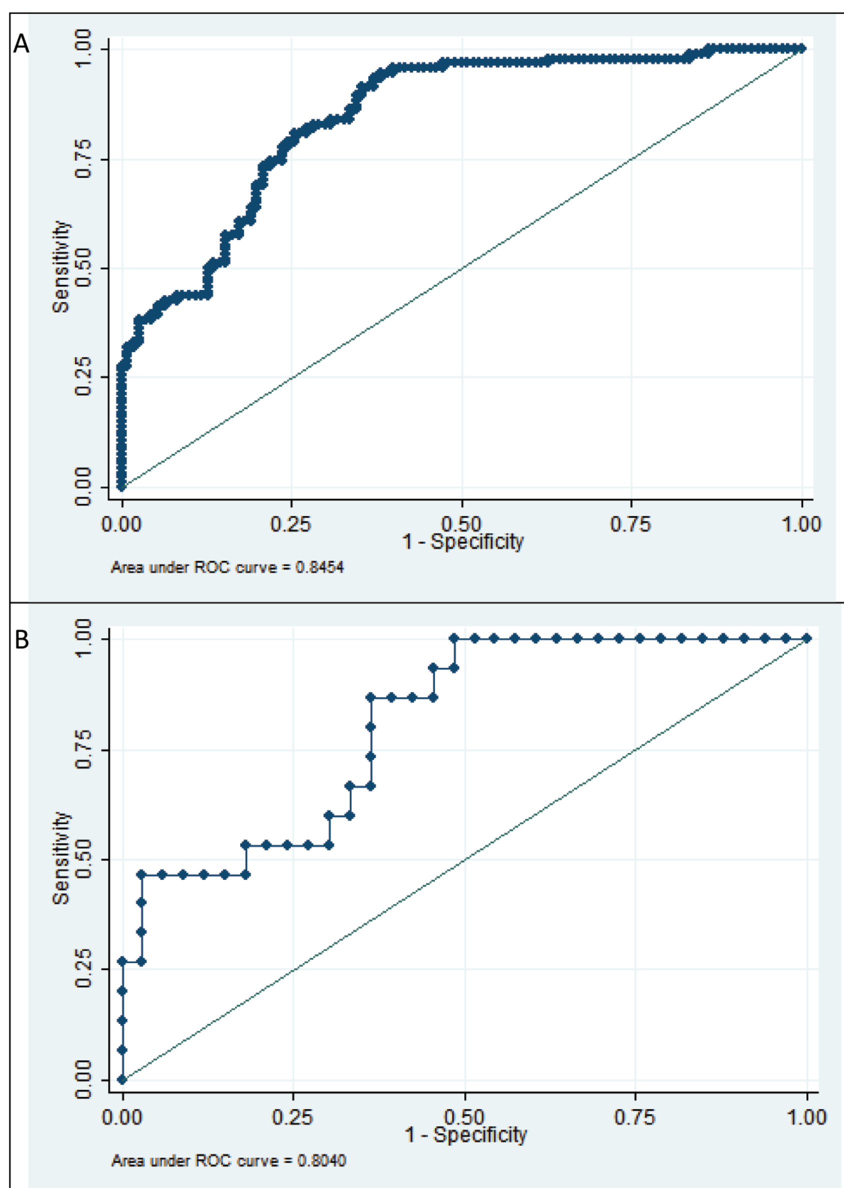

Figure 5. (**A**) ROC curve (sensitivity versus 1—specificity) of the 152 samples in the training set. The AUC is 84.5% (CI = 78.6–89.0%). At specificity of 61.8%, sensitivity is 94.7% (CI = 88.0–98.3%). (**B**) ROC curve (sensitivity versus 1—specificity) of the 48 samples in the data set of post-menopausal women. The AUC is 80.4% (CI = 67.4–91.0%). At specificity of 52.8%, sensitivity is 96.2% (CI = 78.2–100.0%).

more precise and sensitive method, such as protein microarray should be used, which will enable better prediction on statistically significant populations. This is currently being developed.

More significantly, in order to further improve the results in terms of higher specificity, a new category of antigens which are specifically designed for this outcome should be identified and utilized. To perform our study, we relied on a group of antigens relevant in breast cancer chosen from previously published studies that used traditional "cut-off" criteria that were not specifically designed for use in the ratio approach we developed here. It is assumed that identifying special biomarkers, whose amounts of AAbs are increased during cancer, could be used in the ratio approach and will result with better diagnostic capabilities such as higher specificities without compromising the high sensitivity.

In conclusion, we demonstrated the proof-of-concept that measuring the ratio between the levels of AAbs against a panel of previously identified breast cancer TAAs provides an accurate and low-risk confirmatory aid for the diagnosis of breast cancer with high sensitivity and moderate specificity. Our data supports the premise that an assay incorporating calculations of the ratio between the levels of serum AAbs has powerful diagnostic potential. We therefore propose and plan to further improve this concept and outcome by using more accurate and powerful biological techniques such as analyzing dedicated antigens for specific sub-groups and more sensitive diagnostic modalities such as microarray fluorescent

platforms. Further unbiased identification of AAb targets as new TAAs using larger-scale human protein arrays may also be needed using our approach to further enhance the predictive value of this approach.

Acknowledgements

The authors greatly appreciate the cooperation and efforts of M.D. Anderson Cancer Center, University G. D'Annunzio, Division of Senology, European Institute of Oncology, Carmel Medical Center and Kaplan Medical Center, who have generously provided all the specimens for this study. We also thank M.D. Anderson's research personnel Diane M Weber (RN, BSN) and Valerie O Sepeda (RN, BSN), for their contribution to the study, for help in obtaining the breast cancer samples used in this study. The authors would like to acknowledge the huge contribution of the late Prof. Donny Strosberg for his scientific involvement in this project.

Abbreviations

TAA, tumor-associated antigen; AAb, autoantibody; RLU, relative luminescence units; ELISA, enzyme-linked immunosorbent assay; ROC, receiver-operator characteristic; AUC, area under the curve.

Author Contributions

GY and IN contributed equally to the methodology and algorithm development and to the analysis of the data. GY and DW designed the experiments; DW performed all of the laboratory work that established the final assay. TB, LR, LV, ML, MR, BP, and GY contributed to the writing and revisions of this manuscript. TB, SI, MTS, PE, AB and TA provided the plasma and serum specimens used in this study for data acquisition, and helped editing this manuscript. All authors reviewed and approved of the final manuscript.

DISCLOSURES AND ETHICS

As a requirement of publication the authors have provided signed confirmation of their compliance with ethical and legal obligations including but not limited to compliance with ICMJE authorship and competing interests guidelines, that the article is neither under consideration for publication nor published elsewhere, of their compliance with legal and ethical guidelines concerning human and animal research participants (if applicable), and that permission has been obtained for reproduction of any copyrighted material. This article was subject to blind, independent, expert peer review. The reviewers reported no competing interests.

Supplementary Data

Supplementary data table 1. Details of antigens used in this study.

Supplementary data table 2. Results—Data after smoothing procedure for all antigens sorted by samples participating in the study.

Supplementary data table 3. List of the final models used for separation (with at least 80 samples in each model, and sensitivity of at least 95%).

Supplementary data table 4. Prediction given to each sample after applying the models.

Supplementary data table 5. List of the final models used for separation for post menopausal women.

Supplementary data table 6. Prediction given to each sample after applying the models for post menopausal sub population.

Supplementary data table 7. Analysis of clinical variables as "stand alone" predictors.

Supplementary data table 8. Blinded samples prediction using a single model.

REFERENCES

1. Coolen A, Leunen K, Menten J, van Steenbergen W, Neven P. False-negative tests in breast cancer management. *Neth J Med*. 2011;69(7):324–329.
2. Smith-Bindman R, Chu P, Miglioretti DL, et al. Physician predictors of mammographic accuracy. *J Natl Cancer Inst*. 2005;97(5):358–367.
3. Brewer NT, SalzT, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med*. 2007;146(7):502–510.
4. Weigelt B, Geyer FC, Reis-Filho JS. Histological types of breast cancer: how special are they? *Mol Oncol*. 2010;4(3):192–208.
5. Murphy MA, O'Leary JJ, Cahill DJ. Assessment of the humoral immune response to cancer. *J Proteomics*. 2012;75(15):4573–4579.
6. Ran Y, Hu H, Zhou Z, et al. Profiling tumor-associated autoantibodies for the detection of colon cancer. *Clin Cancer Res*. 2008;14(9):2969–2700.
7. Murphy MA, O'Connell DJ, O'Kane SL, et al. Epitope presentation is an important determinant of the utillity of antignes in identified from protein arrays in the developmen of autoantibody diagnostic assays. *J Proteomics*. 2012;75(15):4668–4675.
8. Piura E, Piura B. Autoantibodies to tummor-associated antigens in breast carcinoma. *J Oncol*. 2010:264926.
9. Piura E, Piura B. Autoantibodies to tailor-made panels of tumor associated antigens in breast carcinoma. *J Oncol*. 2011:982425.
10. Baselga J, Seidman AD, Rosen PP, Norton L. HER2 overexpression and paclitaxel sensitivity in breast cancer: therapeutic implications. *Oncology (Williston Park)*. 1997;11(3 Suppl 2):43–48.
11. Yang E, Hu XF, Xing PX. Advances of MUC1 as a target for breast cancer immunotherapy. *Histol Histopathol*. 2007;22(8):905–922.
12. Chinni SR, Falchetto R, Gercel-Taylor C, Shabanowitz J, Hunt DF, Taylor DD. Humoral immune responses to cathepsin D and glucose-regulated protein 78 in ovarian cancer patients. *Clin Cancer Res*. 1997;3(9):1557–1564.
13. Bachelot T, Ratel D, Menetrier-Caux C, Wion D, Blay JY, Berger F. Autoantibodies to endostatin in patients with breast cancer: correlation to endostatin levels and clinical outcome. *Br J Cancer*. 2006;(94):1066–1070.
14. Tomkiel JH, Alansari H, Tang N, et al. Autoimmunity to the Mr 32,000 subunit of replication protein A in breast cancer. *Clin Cancer Res*. 2002;(8):752–758.
15. Erkanli A, Taylor DD, Dean D, et al. Application of Bayesian modeling of autologous antibody responses against ovarian tumor-associated antigens to cancer detection. *Cancer Res*. 2006;(66):1792–1798.
16. Natoli C, Iacobelli S, Kohn L. The immune stimulatory protein 90K increases major histocompatibility complex class I expression in a human breast cancer cell line. *Biochem Biophys Res Commun*. 1996;225(2):617–620.
17. Gold P, Freedman SO. Specific carcinoembryonic antigens of the human digestive system. *J Exp Med*. 1965;(122):467–468.
18. Nicolini A, Carpi A, Tarro G. Biomolecular markers of breast cancer. *Front Biosci*. 2006;(11):1818–1843.