Available online at www.pelagiaresearchlibrary.com

Pelagia Research Library

Advances in Applied Science Research, 2018, 9(2):40-52

ISSN : 0976-8610
CODEN (USA): AASRFC

Diagnosis of Breast Cancer Using Candidate Metabolomic Biomarkers

Faten Z. Mohamed[1], Mohamed M. Omran[2], Hossam Darwish[3], Arafa A. Belal[4], Ramzy E. Rashed[5*]

[1]Faculty of Science, Department of Chemistry, Zagazig University, Zagazig, Egypt
[2]Faculty of Science, Department of Chemistry, Helwan University, Cairo, Egypt
[3]Damietta Cancer Institute, Damietta/Ismailia Teaching Oncology Hospital, Ismailia, Egypt
[4]Faculty of Science, Department of Chemistry, Port Said University, Port Sandaid, Egypt
[5]Laboratory of Toxicology, Central Laboratories, Ministry of Health, Damietta, Egypt

ABSTRACT

Background: *A metabolomic study for determination of some candidate metabolomes derived from nucleosides; 1-methyladenosine (1-MA), 1-methylguanosine (1-MG) and 8-hydroxy-2'-deoxyguanosine (8-OHdG) that suggested as being valuable biomarkers for diagnosis of breast cancer.*

Method: *In this study, the serum samples were collected from 107 female as malignant group, 46 as benign group and 52 as healthy subjects. The serum metabolome were detected and quantified by gas chromatography-mass spectroscopy (GC-MS). The efficiency of candidate metabolomes in diagnosis breast cancer and their combinations with CA15-3 were analyzed using multivariate statistical analysis.*

Result: *The serum concentrations of CA15-3, 1-MA, 1-MG and 8-OHdG were determined; the cutoff values were 31.5 U/l, 23.5 µg/l, 18.5 µg/l and 18.0 µg/l respectively. There are significant differences in mean concentrations of CA 15-3, 1-MA, 1-MG and 8-OHdG in malignant breast diseases patients compared to benign and control group. There are significant correlations between the candidate metabolomes and CA15-3 through different groups of patients with breast malignant diseases. The diagnostic performance analyses of candidate metabolomes and their combinations with CA15-3 gives significant values of area under the curve (AUC) ranged from 0.74 to 0.93 with good efficiency ranged from 60.5% to 88.8%.*

Conclusion: *In this study, the candidate metabolomes had been suggested to be a valuable biomarkers panel for diagnosis of malignant breast diseases.*

Keywords: Breast cancer, Biomarkers, Metabolomics, Mass Spectroscopy.

Abbreviations: 1-MA: 1-Methyladenosine, 1-MG: 1- Methylguanosine, 8-OHdG: 8-Hydroxy-2'-Deoxyguanosine, AUC: Area Under Curve, BC: Breast Cancer, BSTFA: N-Methyl-N- Trimethylsilyltrifluoroacetamide, CA15-3: Cancer Antigen 15-3, CEA: Carcino-Embryonic Antigen, DNA: Deoxyribonucleic Acid, EI: Electron Impact, GC-MS Gas Chromatography- Mass Spectroscopy, LOD: Limit Of Detection, LOQ: Limit Of Quantification, MSD: Mass Selective Detector,, QC: Quality Control, RNA Ribonucleic Acid, ROC: Receiver Operating Characteristic, Rsds, Relative Standard Deviations, SIM: Selected Ion Monitoring, TIC: Total Ion Chromatogram, TNM: T, Tumor-Nodes-Metastasis

INTRODUCTION

In Egypt, breast cancer is the second most prevalent cancer overall, the most recurrent type of cancer in women and the major cause of female death as in the all of the world [1,2]. Mammography and ultrasonography are generally

used for routine surveillance and for staging. However, these tools have limitations such as radiation exposure, and a breast tumor is required to be at least a few millimeters in size for detection [3,4]. Moreover, the common biomarkers of breast cancer for example, estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER-2) have been allied with a poor prognosis [5]. On the other hand, carcinoembryonic antigen (CEA), cancer antigen 15-3 (CA 15-3) and cancer antigen 27-29 (CA 27-29), tissue polypeptide antigen (TPA) and tissue polypeptide specific antigen (TPS) have been associated with low sensitivity and/or low specificity and often respond late to recurrence [6].

The metabolomics studies and its applications becoming more and more an increasingly popular tool in life science and in clinical applications, it is a relatively fast, accurate, cost-efficient and fully automated technique [7]. The metabolomics approach is to use metabolomics profiling through which it can detect diseases based on a band of small molecules in different specimens [8]. Nuclear magnetic resonance (NMR) and mass spectroscopy (MS) are two predominant techniques used in metabolomes detection and quantification in metabolomics studies. NMR may be less sensitive and less readily automated with higher cost. In comparison, MS is more sensitive and highly specific for low molecular weight molecules and it is readily automated with lower costs [9]. MS may often be coupled with gas or liquid chromatography (GC-MS & LC-MS). GC-MS has been regarded as the standard technique for analyzing many small compounds (lipids, drug metabolites, and environmental contaminants), specifically volatile, thermostable and low molecular weight molecules. One of many advantages of GC-MS is that the detected species identification is based on both a retention time and the mass spectra; compounds specific fragmentation pattern (compounds fingerprint), so the GC-MS is a golden tool for metabolomics analysis [10,11].

In malignant diseases, the abnormal levels of metabolomes, derived from nucleosides in body fluid are related oxidative DNA damage or RNA's turnover and higher methyl- transferase activity [12-14]. Therefore, according to several studies, the metabolomes of body fluids have been proposed as possible biomarkers for different cancer diseases such as in colorectal [15], breast [16], bladder [17], Lymphoma [18], thyroid [19] and cervical cancer [20]. In this study, we focused on three candidate metabolomes derived from nucleosides; 1-MA, 1-MG and 8-OHdG, which suggested as valuable biomarkers for the early diagnosis of breast cancer [10,21].

MATERIALS AND METHODS

Patients and samples

The present study was conducted on 107 female patients with malignant breast diseases and 46 female patients with benign breast diseases, in addition 52 healthy female volunteers as a control group who did not undergo antecedent breast biopsy. All patients were referred from the Damietta cancer institute, Damietta, and Ismailia teaching oncology hospital, Ismailia, Egypt. The clinicopathological data for patient population were collected. The patients with malignant breast diseases were classified according to their clinicopathological characteristics and their TNM staging based on the American Joint Committee on Cancer (AJCC). None of the patients with benign breast disease or the healthy controls had a past-history of any malignancy. The levels of CA15-3 were measured by enzyme-linked immunosorbent assay (ELISA) kit, for all patients and healthy individuals. All subjects had normal laboratory data with normal estimated glomerular filtration rate (eGFR). For metabolomic compounds analysis, 3 ml of venous blood sample was withdrawn in a blank tube from patients and healthy control in the morning after overnight fasting. After obtaining the samples, the sera obtained directly by centrifuging the sample at 4000 rpm for 5 min, and then were kept in –20^{o} C until analysis. Blood samples for all patients withdrawn up to 2 days after diagnosis and before starting any cancer-specific treatment. The Ethics and Scientific Damietta cancer institute, Damietta, and Ismailia teaching oncology hospital, Ismailia, Egypt, have approved the study. All patients gave their informed consent prior to enrolment.

Reagents

Certified reference compounds of 1-MA, 1-MG and 8-OHdG, internal standards, and human serum sources purchased from Sigma-Aldrich (St. Louis, MO, USA). HPLC grade methanol and n-hexane purchased from Merck (Darmstadt, Germany). N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA), the derivatizing agent, was purchased from Pierce Chemical Co. The compounds and reagents were all of analytical grade except where stated otherwise.

Metabolomics Analysis

Sample preparation

The serum samples were thawed on the ice at 4°C for 30–60 min. To 200 µ/l serum of each sample, 0.5 ml of methanol, including 1 mg/l of internal standard was added then vortexed for 15 s and centrifuged for 15 min at 5000 rpm and 4°C. Then 150 µ/l of the supernatant was transferred to a separate new tube and dried by nitrogen gas followed

by adding 100 µl MSTFA and mixing for 15 s, then incubated for 1 h (65°C) after that, let it for 15 min at room temperature followed by adding of 1 ml n-hexane in clear analytical vial [22].

GC-MS conditions

The analysis was performed using an Agilent 7890B gas chromatography coupled to an Agilent 5977A mass selective detector (MSD) with inert electron impact ionization source (Agilent Technologies, Santa Clara, CA). Analytes compounds were separated using an HP-5MS (30 m × 250 µm × 0.25 µm) capillary column (Agilent J & W Scientific, Folsom, CA), using helium 99.999% as a carrier gas at a constant flow rate mode (1 ml min^{-1}). The GC parameters were set as the following: 1µ/l injection volume, the front inlet was set in spitless mode at 250°C, the front-line auxiliary temperature was 300°C. For every analysis, the purge time was set to 120 s at a purge flow rate of 20 ml min^{-1}. The time of solvent delay was set to 7.5 min. The initial oven temperature was held at 70°C for 2 min, ramped to 110°C at a rate of 10°C min^{-1}, then to 230°C at a rate of 5°C min^{-1}, and finally to 280°C at a rate of 25°C min^{-1}. The MSD tune parameters were set as the following: Ion source electron impact (EI) set in positive mode at 300°C, transfer line was set at 280°C, the quad temperature was set at 150°C.

Quantitative analysis of candidate metabolomes

A stock solution of the internal standard and mixture of standards was prepared at 1 mg/l and stored at 4°C and used within one month as a working standard solution [23]. Then 1-µ/l of 1 mg/l of mixture standard, including 1 mg/l of internal standard was injected splitless SCAN mode by Autosampler through 36-minute runtime. The obtained retention times (RT) were 13.966, 14.802 and 16.355 for 1-MA, 1-MG and 8-OHdG respectively as shown in the Figure 1(a). The target masses of metabolomic compounds were selected as a highest abundance of three masses that were in agreement NIST library data. According to the retention time and the target masses, the SIM mode method was constructed to determine the trace concentrations of analytes. Then 1-µ/l of different concentrations of mixture standard solutions (5, 10, 25, 50, 100 and 200 µg/l) was injected splitless SIM mode, the calibration curves of 1-MA, 1-MG and 8-OHdG were constructed by plotting the total ion chromatogram (TIC) response area with standard concentrations with correlation coefficients 0.9921, 09946, and 0.9948 respectivily. After the calibration build-up, a 1-µ/l of each blank, reagent blank, sample blank, QC materials and derivatives samples were injected splitless SIM mode and quantified after running time. Instrument control, data acquisition and data analysis were performed under MSD chemstation (E.02.02.1431), Agilent technology software. NIST 08 library was used for peak identification. The the analytes recoveries and limit of quantitation (LOQ) calculated with excellent relative standard deviations (RSDs). The RSDs for 1-MA, 1-MG and 8-OHdG were 8.7%, 8.1%, and 8.2%, and the recoveries were 108%, 96% and 95% respectively. After this, 1-µ/l of each prepared sample of healthy controls benign and malignant breast diseases patients was injected and quantified. Figure 1(b) shows that the TIC of a random sample of a healthy individual and BC patient with a blank.

Statistical Analysis

All the statistical analyses were performed with SPSS 22.0 (SPSS, Chicago, IL, USA). The data are expressed as median values (interquartile range). Nonparametric analyses were performed to assessment the differences of the candidate biomarkers measurements across the different categories. The correlations between studying parameters were evaluated by Spearman correlation coefficient. Receiver operating characteristic (ROC) curves were drawn to discriminate between healthy individual, benign and malignant breast diseases patients. The area under curve (AUC) and cutoff values were derived. Cut-off values were obtained from running the ROC analysis on SPSS software, which normally uses maximum sum of sensitivity and specificity as a cut-off point. To obtain the optimal diagnosis of breast cancer, a multiple linear regression model was run to fit the combinations between each candidate metabolomes and CA15-3 as well as between all candidate metabolomes with CA15-3 (named SCORE). The combined scores were obtained from linear combination of several markers (variables) by running the linear regression. The performance model; sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and the efficiency were derived for individual and combined variables from a 2×2 contingency table. For all tests, the p-values were significant at <0.05 and the confidence level was 95%.

RESULTS

Clinicopathological characteristics of subjects

In this study, 107 patients of malignant breast diseases were enrolled during the period from January 2015 to July 2017 with median age 46 years (ranged from 25 to 64 years). The malignant patients were treated for stages I, II, III and IV accounted 7.5%, 57.9%, 25.2% and 9.3% respectively, and 46 patients of benign breast diseases with median age 44 years (ranged from 25 to 61 years), among of them 34 (73.9%) were fibroadenomas and 12 (26.1%) were papilloma. In addition, 52 healthy controls with median age 40 years (ranged from 26 to 60 years). The clinicopathological features of the patients and healthy control are shown in Table 1.

Table 1: Characteristics of patients and controls, the serum concentrations of CA15-3 and candidate metabolomics depicted as median (IQR) and the significance of mean differences of variables concentrations across different groups.

Variables		n	%	CA15-3(U/l)		1-MA (µg/l)		1-MG (µg/l)		8-OHdG (µg/l)	
				Median (IQR)	*p*	Median (IQR)	*p*	Median (IQR)	*p*	Median (IQR)	*p*
Malignant patients		**107**		**49.0 (28 – 73)**	**<0.001$**	**39.0 (20 – 65)**	**<0.001$**	**27.0 (15 – 39)**	**<0.001$**	**37.0 (17 – 49)**	**<0.001$**
Age, median (range), 46 (25–64) years	≤ 35 years	15	14.0	57.0 (29 – 73)	NS	21.0 (16 – 58)	NS	21.0 (14 - 38)	NS	29.0 (19 – 39)	NS
	≥ 35 years	92	86.0	48.5 (26 – 78)		41.5 (21 – 67)		27.0 (15 – 39)		38.0 (17 – 51)	
Tumor size	≤2 cm	44	41.1	49.0 (27 – 77)	NS	32.5 (19 – 60)	NS	26.5 (13 – 39)	NS	24.5 (14 – 48)	NS
	≥ 2 cm	63	58.9	48.0 (28 – 73)		42.0 (20 – 67)		27.0 (17 – 39)		41.0 (19 - 50)	
T- status	T1	20	18.7	41.5 (20 – 68)	NS	31.5 (14 – 72)	NS	23.5 (13 – 42)	NS	22.5 (14 – 39)	NS
	T2	58	54.2	47.5 (28 – 68)		28.5 (19 – 59)		31.0 (19 – 39)		38.5 (18 – 49)	
	≥T3	29	27.1	70.0 (27 – 118)		57.0 (27 – 76)		21.0 (13 – 32)		42.0 (19 – 61)	
N-status	N0	62	57.9	45.5 (25 – 71)	NS	40.5 (19 – 66)	NS	24.0 (14 – 38)	NS	41.0 (20 – 51)	NS
	N1	26	24.3	49.0 (36 – 73)		32.0 (21 – 72)		32.0 (20 – 41)		30.5 (15 – 47)	
	≥N2	19	17.8	73.0 (39 – 118)		41.0 (21 – 57)		28.0 (15 – 37)		21.0 (12 – 62)	
Metastasis	M0	97	90.7	49.0 (28 – 76)	NS	33.0 (20 – 61)	**0.038**	26.0 (15 – 38)	NS	36.0 (16 – 49)	**0.046**
	M1	10	9.3	37.5 (21 – 73)		61.5 (37 – 85)		46.0 (14 – 52)		45.0 (38 – 65)	
Tumor stages	I	8	7.5	20.5 (17 – 53)	<0.001	14.0 (12 – 28)	**0.002**	13.0 (9 – 20)	**0.033**	18.0 (9 – 31)	**0.024**
	II	62	57.9	46.5 (28 – 67)		37.5 (20 – 66)		27.0 (18 – 38)		37.5 (18 – 49)	
	III	27	25.2	73.0 (49 – 125)		41.0 (26 – 77)		25.0 (13 – 37)		37.0 (14 – 62)	
	IV	10	9.3	37.5 (21 – 73)		61.5 (37 – 85)		46.0 (14 – 52)		45.0 (38 – 65)	
Benign patients		46		31.0 (25 – 36)	**<0.001#**	18.5 (14 – 24)	**<0.001#**	16.0 (11- 19)	**<0.001#**	16.5 (11 – 20)	**<0.001#**
Age, median (range), 44 (25–61) years	≤35 years	9	19.6	33.0 23 – 35)	NS	16.0 (14 – 21)	NS	12.0 (8 – 15)	NS	12.0 (8 – 20)	NS
	≥ 35 years	37	80.4	30.0 (12 – 36)		19.0 (13 – 25)		17.0 (12 -20)		17.0 (12 – 21)	
Benign stages	Fibroadenoma	34	73.9	31.5 (26 – 36)	NS	19.0 (14 – 25)	NS	15.5 (11 – 19)	NS	18.0 (12 – 21)	NS
	Papilloma	12	26.1	28.5 (23 – 36)		18.0 (13 – 23)		16.0 (11 – 19)		14.0 (10 – 17)	
Healthy control		52		31.0 (24 – 34)	NS * **<0.001¥**	18.0 (11 –25)	0.478* **<0.001¥**	12.0 (8 – 18)	NS * **<0.001¥**	13.0 (9 – 17)	**0.014*** **<0.001¥**
Age, median (range), 40 (26–60) years	≤35 years	12	23.1	31.0 (21– 34)	NS	15.0 (12 – 29)		9.0 (7 – 16)	NS	12.0 (7 - 16)	NS
	≥ 35 years	40	76.9	31.0 (24 – 34)		18.0 (11 – 25)	NS	14.0 (8 – 19)		13.0 (9 - 19)	

Note: IQR; interquartile range, T-status; tumour size status, N-status; nodal status, *; Control vs. Benign patients, #; Benign vs. Malignant patients, ¥; Control vs. Malignant patients, $; Difference between three group, NS, Not Significant (p>0.05)

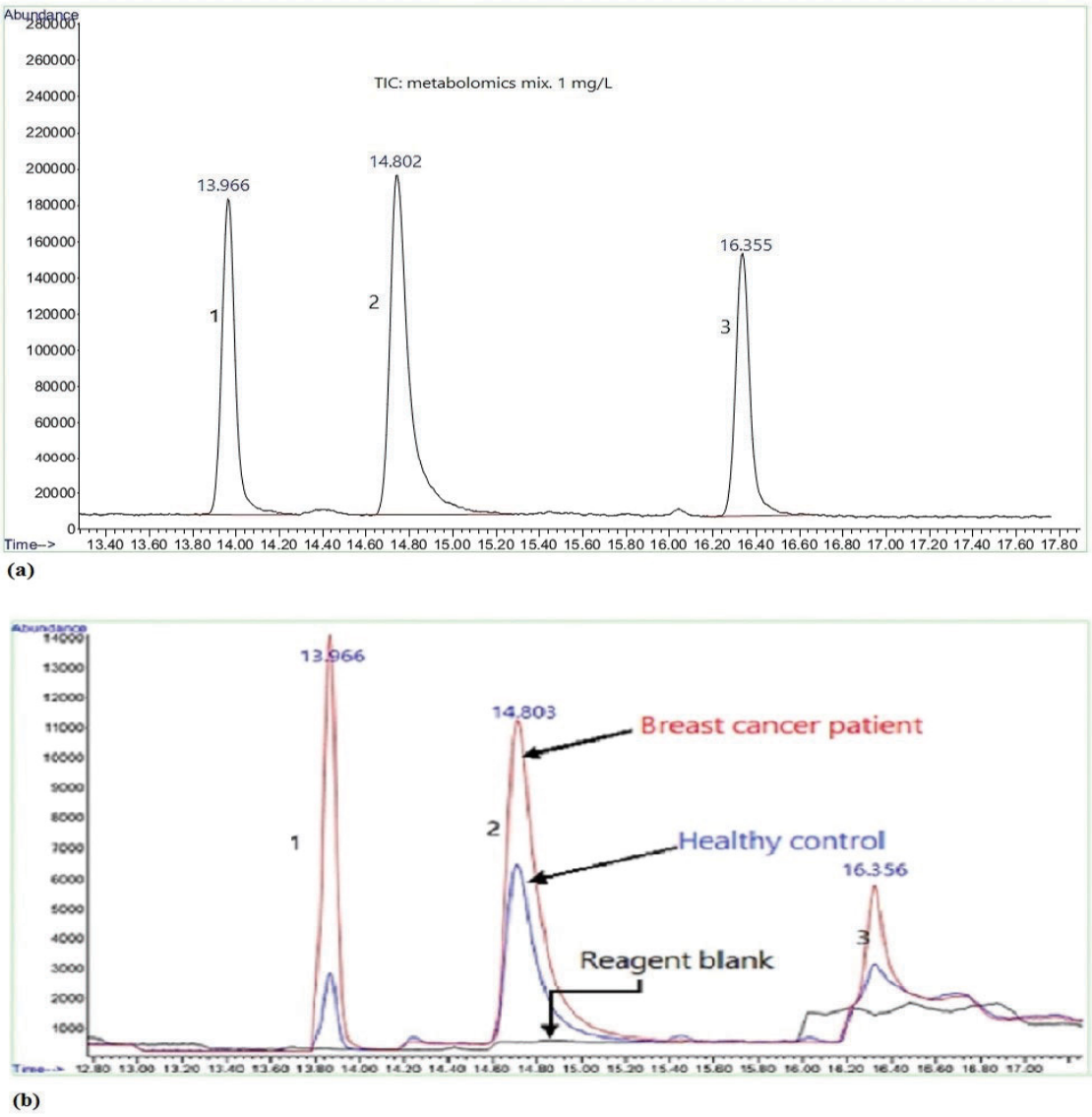


Figure 1: Total ion chromatogram **(a)** SCAN mode analysis for 1 mg/l of mixture standard of metabolomics. 1) 1-MA 2) 1-MG 3) 8-OHdG at retention times 13.966, 14.802, and 16.355 min, respectively. **(b)** SIM mode analysis of blank, random healthy control sample, for which the concentration of the 1-MA, 1-MG, and 8-OHdG was 12, 35 and 7 µg/l, respectively, and abnormal random BC sample, for which the concentration of the 1-MA, 1-MG, and 8-OHdG was 68, 53 and 25 µg/l, respectively.

Levels of candidate markers in serum

As illustrated in Table 1, there are significant differences in concentrations of candidate biomarkers between healthy and malignant patients. There is a significant difference in concentration of the 8-OHdG between healthy and benign group (p=0.014), and between heathy and malignant group ($p \leq 0.001$). Through malignant patients, there are significant differences in concentrations of the CA15-3, 1-MA, 1-MG and 8-OHdG in the different tumor stages ($p \leq 0.001$, 0.002, 0.033 & 0.024 respectively). In addition, there is significant difference in concentrations of 1-MA and 8-OHdG between non- and metastasis stages (p=0.038 & 0.046 respectively).

Diagnostic performance analysis of candidate biomarkers

Optimal metabolomic markers and CA15-3 cutoff levels between patients and healthy controls were determined using the Receiver Operating Characteristic (ROC)., as shown in Table 2. The multivariate analysis of candidate biomarker using linear regression for combinations of three candidate biomarkers with CA15-3 (was always included as the most significant breast cancer biomarkers used in the clinic) was performed to obtain the optimum diagnostic

performance analysis. As illustrated in Table 2, the sensitivity and specificity of combined biomarkers increased with highly efficiency comparing to single biomarkers. For combined biomarkers, the sensitivity ranged from 81.3% to 88.8%, the specificity ranged from 80.6% to 88.8% and the efficiency ranged from 81.0% to 88.8%. The diagnostic performance analysis illustrated in Figure 2. ROC analysis of CA15-3, 1-MA, 1-MG and 8-OHdG in healthy control and malignant breast diseases patients was shown in Figure 2 (a). ROC analysis of CA15-3, 1-MA, 1-MG and 8-OHdG in benign and malignant breast diseases patients was shown in Figure 2 (b). ROC analysis of the combination between the candidate biomarkers in healthy control and malignant breast diseases patients was shown in Figure 2 (c). ROC analysis of the combination between the candidate biomarkers in benign and malignant breast diseases patients was shown in Figure 2 (d).

Table 3 summarizes the diagnostic performance analysis of candidate biomarker to discriminate between healthy and malignant stages (I, II, III, & IV), on one hand and benign patients with malignant stages (I, II, III, & IV) from other hand. The AUC-values of combined biomarker were found to be significant, especially when focusing on advanced disease compared to normal healthy controls and non-malignant breast diseases patient. The diagnostic performance analysis of combined candidate metabolomes with CA15-3 gives significant AUC values ranged from 0.90 to 0.94 in comparing the healthy control with malignant breast diseases and from 0.86 to 0.92 in comparing benign breast diseases to malignant breast diseases patients. The most significant AUC-values (0.99) was for the combination of CA15-3 with three candidate metabolomes biomarkers (SCORE) in comparing the control to stages III and IV and in comparing the benign breast diseases to IV of malignant patients. As shown in Figure 3, the area under carve for combined biomarkers with CA15-3 is higher than that of CA15-3 alone. Figure 3 (a) represents the comparison of control to benign breast diseases and the malignant breast diseases stages I, II, III and IV. Figure 3 (b) represents the comparison of benign breast diseases to malignant breast diseases stages; stage I, II, III and IV.

Correlation between clinicopathological characteristics of patients with candidate biomarkers

As illustrated in Table 4, the elevated serum levels of CA15-3 1-MA, 1-MG and 8-OHdG in malignant patients were identified in 70 (65.4%), 85 (79.4%), 84 (78.5%) and 96 (89.7%) respectively. There are significant correlations

Table 2: The diagnostic performance analysis for each candidate metabolomics and combined candidate metabolomics together with CA15-3 to discriminate between healthy individual and patients with breast malignant diseases.

Test Result Variable(s)	Cutoff	AUC	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Efficiency (%)
CA15-3	31.5	0.74	65.4	59.3	61.4	55.1	60.5
1-MA	23.5	0.80	67.3	67.4	69.2	65.4	65.4
1-MG	18.5	0.77	66.4	75.5	74.4	67.3	70.7
8-OHdG	18.0	0.82	74.7	68.3	72.1	71.3	71.1
CA15-3 + 1-MA	2.08	0.89	81.3	80.6	82.1	79.8	81.0
CA15-3 + 1-MG	2.13	0.87	80.4	80.6	81.9	79.0	80.5
CA15-3 + 8-OHdG	2.05	0.91	84.1	83.7	84.9	82.8	83.9
Score	2.06	0.93	88.8	88.8	89.6	87.8	88.8
Note: SCORE; linear combination between different variables (CA15-3, 1- MA, 1-MG, and 8-OHdG).							

Table 3: The diagnostic performance analysis for each candidate biomarker and multivariate analysis for combined biomarker using logistic regression A.

Test result variables	Healthy vs. Benign breast diseases		Healthy vs. Malignant breast diseases		Healthy vs. Stage I		Healthy vs. Stage II		Healthy vs. Stage III		Healthy vs. Stage IV	
	AUC	*p*	AUC	*p*	AUC	*p*	AUC	*p*	AUC	*p*	AUC	*p*
CA15-3	0.58	0.254	0.75	<0.001	0.48	0.851	0.74	<0.001	0.88	<0.001	0.63	0.485
1-MA	0.54	0.479	0.81	<0.001	0.50	0.723	0.80	<0.001	0.88	<0.001	0.95	<0.001
1-MG	0.59	0.145	0.79	<0.001	0.54	0.215	0.83	<0.001	0.75	0.001	0.85	0.001
8-OHdG	0.64	0.015	0.85	<0.001	0.73	0.041	0.87	<0.001	0.81	<0.001	0.93	<0.001
CA15-3 + 1-MA	0.66	0.006	0.90	<0.001	0.61	0.330	0.89	<0.001	0.97	<0.001	0.99	<0.001
CA15-3 + 1-MG	0.61	0.043	0.88	<0.001	0.57	0.543	0.90	<0.001	0.94	<0.001	0.83	0.001
CA15-3 + 8-OHdG	0.66	0.005	0.93	<0.001	0.65	0.163	0.93	<0.001	0.98	<0.001	0.98	<0.001
Score	0.68	0.002	0.94	<0.001	0.64	0.144	0.95	<0.001	0.99	<0.001	0.99	<0.001

Test result variables	Benign vs. Malignant breast diseases		Benign vs. Stage I		Benign vs. Stage II		Benign vs. Stage III		Benign vs. Stage IV	
	AUC	*p*	AUC	*p*	AUC	*p*	AUC	*p*	AUC	*p*
CA15-3	0.73	<0.001	0.43	0.535	0.72	<0.001	0.87	<0.001	0.61	0.359
1-MA	0.80	<0.001	0.42	0.467	0.78	<0.001	0.84	<0.001	0.94	<0.001
1-MG	0.75	<0.001	0.43	0.551	0.80	<0.001	0.71	0.003	0.81	0.003
8-OHdG	0.78	<0.001	0.57	0.377	0.80	<0.001	0.74	0.001	0.91	<0.001
CA15-3 + 1-MA	0.86	<0.001	0.60	0.377	0.87	<0.001	0.96	<0.001	0.99	<0.001
CA15-3 + 1-MG	0.89	<0.001	0.53	0.793	0.88	<0.001	0.93	<0.001	0.79	0.005
CA15-3 + 8-OHdG	0.95	<0.001	0.52	0.821	0.98	<0.001	0.96	<0.001	0.97	<0.001
Score	0.92	<0.001	0.65	0.321	0.92	<0.001	0.97	<0.001	0.99	<0.001
Score	2.06	0.93	88.8	88.8	89.6	87.8	88.8	--	--	--

Note: SCORE; linear combination between different variables (CA15-3, 1- MA, 1-MG, and 8-OHdG).

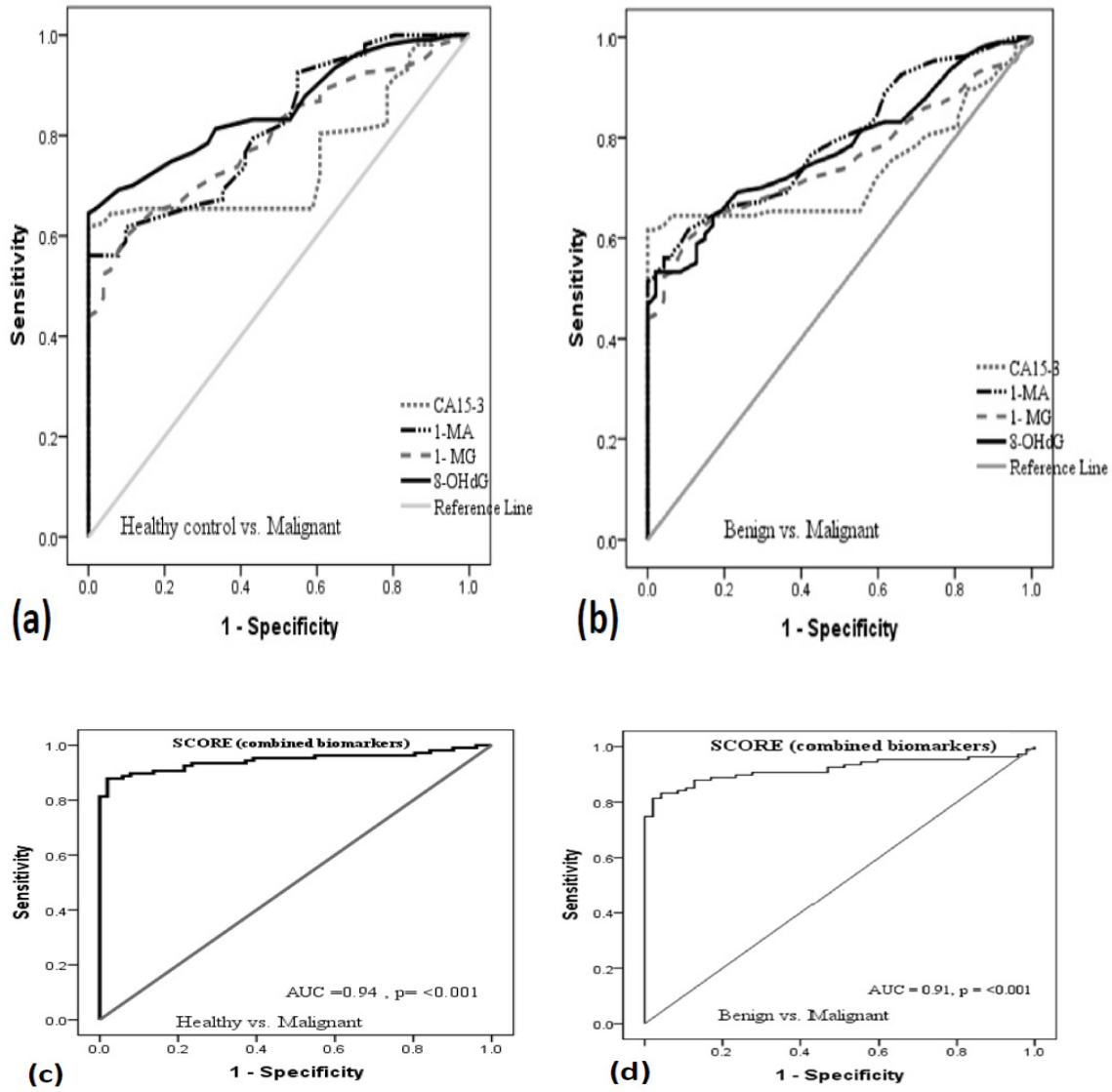
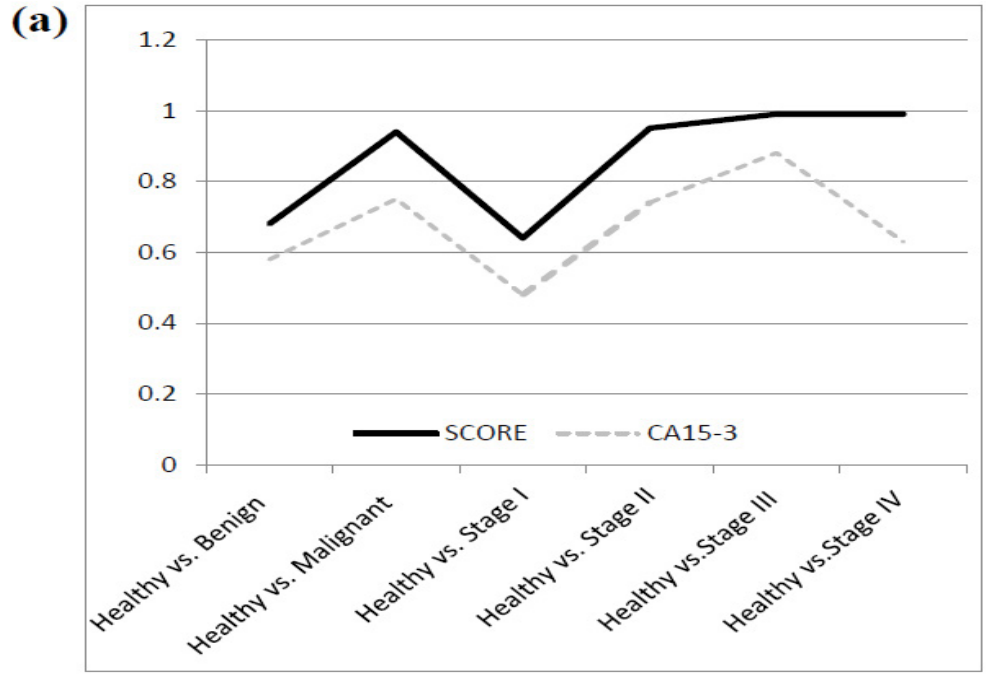


Figure 2: The diagnostic performance analysis. **(a)** ROC analysis of CA15-3, 1-MA, 1-MG and 8-OHdG in healthy control and malignant breast diseases patients. **(b)** ROC analysis of CA15-3, 1-MA, 1-MG and 8-OHdG in benign and malignant breast diseases patients. **(c)** ROC analysis of the combination between the candidate biomarkers in healthy control and malignant breast diseases patients. **(d)** ROC analysis of the combination between the candidate biomarkers in benign and malignant breast diseases patients.

Table 4: Elevated levels of CA15-3 and candidate metabolomics and the correlations between serum CA 15–3 and candidate metabolomics with clinic-pathological data.

Variables		n	CA15-3 (U/l)			1-MA (µg/l)			1-MG (µg/l)			8-OHdG (µg/l)		
			Normal n (%)	Elevated n (%)	*p*	Normal n (%)	Elevated n (%)	p	Normal n (%)	Elevated n (%)	p	Normal n (%)	Elevated n (%)	*p*
Malignant patients		107	37 (34.6)	70 (65.4)		35 (32.7)	72 (67.3)		36 (33.6)	71 (66.4)		27 (25.2)	80 (74.8)	
Age	≤35 years	15	4 (26.7)	11 (73.3)	0.784	8 (53.3)	7 (46.7)	0.178	6 (40.0)	9 (60.0)	**0.001**	3 (20.0)	12 (80.0)	**0.001**
	≥ 35 years	92	33 (35.9)	59 (64.1)		27 (29.3)	65 (70.7)		30 (32.6)	62 (67.4)		24 (26.1)	68 (73.9)	
Tumor size	≤2 cm	44	17 (38.6)	27 (61.4)	**0.005**	14 (31.8)	30 (68.2)	**0.006**	16 (36.4)	28 (63.6)	0.503	16(36.4)	28 (63.6)	**0.005**
	≥ 2cm	63	20 (31.7)	43 (68.3)		21 (33.3)	42 (66.7)		20 (31.7)	43 (68.3)		11 (17.5)	52 (92.5)	
T- status	T1	20	8 (40.0)	12 (60.0)	**0.017**	8 (40.0)	12 (60.0)	0.057	8 (40.0)	12 (60.0)	0.224	7 (35.0)	13 (65.0)	**0.030**
	T2	58	21 (36.2)	37 (63.8)		21 (36.2)	37 (63.8)		14 (24.1)	44 (75.9)		13 (22.4)	45 (77.6)	
	≥T3	29	8 (27.6)	21 (72.4)		6 (20.6)	23 (79.4)		14 (48.3)	15 (51.7)		7 (24.1)	22 (75.9)	
N-status	N0	62	27 (43.5)	35 (56.5)		22 (35.5)	40 (64.5)	0.742	23 (37.1)	39 (62.9)	0.654	11 (17.7)	51 (82.3)	0.176
	N1	26	6 (23.1)	20 (76.9)	**0.023**	8 (30.7)	18 (69.3)		6 (23.1)	20 (76.9)		7 (26.9)	19 (73.1)	
	≥N2	19	4 (21.1)	15 (78.9)		5 (26.3)	14 (73.7)		7 (36.8)	12 (63.2)		9 (47.4)	10 (52.6)	
Metastasis	M0	97	32 (33.0)	65 (67.0)	0.289	34 (35.0)	63 (65.0)	**0.037**	33 (34.0)	64 (66.0)	0.106	26 (26.8)	71 (73.2)	**0.021**
	M1	10	5 (50.0)	5 (50.0)		1 (10.0)	9 (90.0)		3 (30.0)	7 (70.0)		1 (10.0)	9 (90.0)	
Tumor stages	I	8	5 (62.5)	3 (37.5)		6 (75.0)	2 (25.0)		6 (75.0)	2 (25.0)		4 (50.0)	4 (50.0)	
	II	62	23 (37.1)	39 (62.9)	**0.003**	22 (65.5)	40 (64.5)	**0.001**	17 (24.4)	45 (75.6)	0.076	12 (19.3)	50 (80.7)	**0.046**
	III	27	4 (14.8)	23 (85.2)		6 (22.2)	21 (78.8)		10 (37.0)	17 (63.0)		10 (37.0)	17 (63.0)	
	IV	10	5 (50.0)	5 (50.0)		2 (20.0)	8 (80.0)		3 (30.0)	7 (70.0)		2 (20.0)	8 (80.0)	
Benign patients		46	25 (54.3)	21 (45.7)		39 (84.8)	7 (15.2)		37 (80.4)	9 (19.6)		38 (82.6)	8 (17.4)	
Age	≤35 years	9	4 (44.4)	5 (55.6)	0.839	8 (88.9)	1 (11.1)	0.203	1 (11.1)	8 (88.9)	0.108	7 (77.8)	2 (22.2)	0.620
	≥ 35 years	37	8 (21.6)	29 (78.4)		30 (81.1)	7 (18.9)		28 (75.7)	9 (24.3)		30 (81.1)	7 (18.9)	
Benign stages	Fibroadenoma	34	9 (26.5)	25 (72.5)	0.562	26 (76.5)	8 (23.5)	0.570	26 (76.5)	8 (23.5)	0.243	27 (79.4)	7 (20.6)	0.562
	Papilloma	12	10 (83.3)	2 (16.7)		11 (91.7)	1 (8.3)		10 (83.3)	2 (16.7)		10 (83.3)	2 (16.7)	
Healthy control		52	29 (55.8)	23 (44.2)		40 (76.9)	12 (23.1)		41 (78.8)	11 (21.2)		50 (96.2)	2 (3.8)	
Ag	≤35 years	12	7 (58.3)	5 (41.7)	0.808	9 (75.0)	3 (25.0)	0.610	10 (83.3)	2 (16.7)	0.299	11 (91.7)	1(8.3)	0.258
	≥ 35 years	40	22 (55.0)	18 (45.0)		32 (80.0)	8 (20.0)		32 (80.0)	8 (20.0)		39 (97.5)	1 (2.5)	



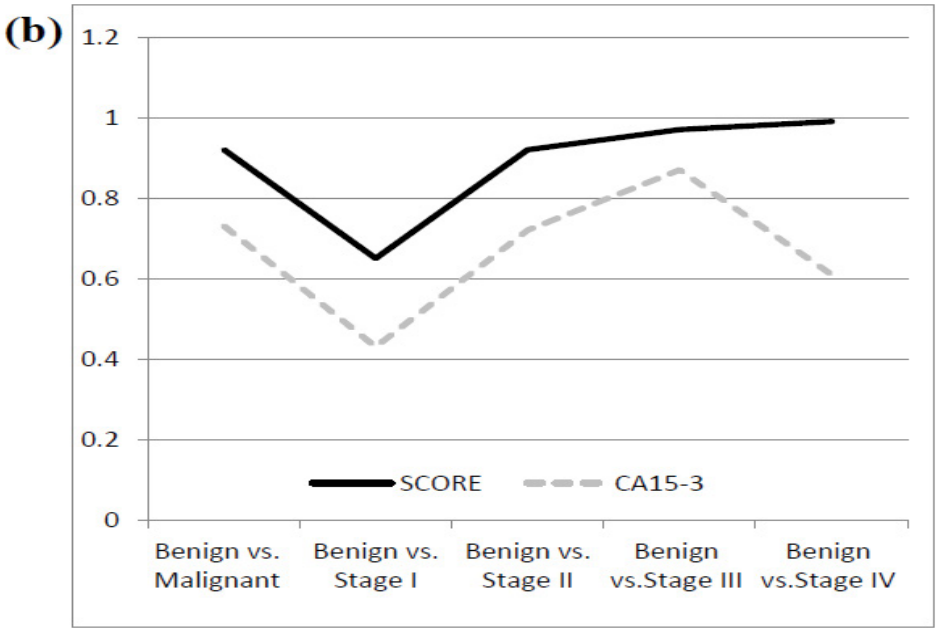


Figure 3: Linear regression AUC-values. **(a)** Comparison of AUC-values for the combination of CA15-3, 1-MA, 1-MG, and 8-OHdG with CA15-3 alone for healthy control compared to benign breast disease, malignant breast diseases patients, stage I, stage II, stage III and stage IV breast cancer patient. **(b)** Comparison of AUC-values for the combination of CA15-3, 1-MA, 1-MG, and 8-OHdG with CA15-3 alone for benign breast diseases patients compared, malignant breast diseases patients, stage I, stage II, stage III and stage IV breast cancer patient.

between CA15-3 and tumor size (p=0.005), T-status (p=0.017), N-status (p=0.023) and tumor stages of malignant breast diseases (p=0.003). In addition, there are significant correlations between serum 1-MA and tumor size (p=0.006), metastasis stages (p=0.036) and tumor stages (p=0.001). Whereas, the 1-MG is correlated only with patients age (p=0.001) and 8-OHdG is correlated with age (p=0.001), tumor size (p=0.005), T-status (p=0.030), metastasis stages (p=0.021) and tumor stage (p=0.046). In benign patients' diseases, the elevated serum levels of the elevated serum levels of CA15-3 1-MA, 1-MG and 8-OHdG were 21 (45.7%), 7 (15.2%), 9 (19.6%) and 8 (17.4%) respectively. On the other hand, there are no significant correlations between CA15-3 and candidate metabolomes with cases of benign

diseases in this study. There are significant positive correlations between CA15-3 in all studied subjects with 1-MA ($p \leq 0.001$), 1-MG ($p \leq 0.001$) and 8-OHdG ($p=0.007$).

DISCUSSION

As in the most malignant diseases, early detection of BC is useful for the easily treatment, enhance patient survival and reducing the rate of death. specificity and often respond late to recurrence [6]. Many recent studies focused on the determination and identification of malignant breast diseases metabolomics in serum and body fluid to emerge its medical significance [16,24-29]. Detailed assignments of the biochemical metabolite compounds in breast tumors have been made using a combination of the different platform of analytical techniques based on mass spectroscopy and nuclear magnetic resonance [9,30-32].

This essay suggests the metabolomics biomarkers for diagnosis of BC by determination of some candidate metabolomes in serum derived from nucleosides by applying the GC-MS analysis technique. In this study, we focused on three candidate metabolomes; 1-MA, 1-MG and 8-OHdG. Due to the low molecular weight and they have no interaction with serum proteins, this metabolomics cannot be retained in the circulation [33] and therefore the serum content of this metabolomics is much less than that in the urine [34]. Nevertheless, the serum specimens are preferred to determinate the metabolomics, this is due to the metabolomics in serum are less affected by exogenous factors than those in the urine, furthermore, the creatinine level measurement is necessary for urinary metabolomics estimation [35].

In this study, the p-value for determination of serum 1-MA, 1-MG and 8-OHdG in patients with breast benign and malignant diseases had demonstrated. We successfully quantified the concentration of serum 1-MA, 1-MG and 8-OHdG in healthy individuals benign and malignant breast diseases patients using GC-MS technique, and then we studied the associations between serum levels of CA15-3 and quantified metabolomics with the different groups. The method of quantification of candidate metabolomes was validated (fit for purpose) by precision studies using internal and external quality control materials. From the fact that the nucleoside metabolomics occurred in the circulation caused by oxidative DNA damage and RNA's metabolites turnover, which increases in malignant diseases [12], we found increased levels of serum 1-MA, 1-MG and 8-OHdG compared to benign and normal healthy control groups. The most noticeable increase in concentrations of 1-MA was detected in stage IV of malignant patients (80.0%). Whereas the most noticeable increase in concentrations of 1-MG and 8-OHdG were detected stage II malignant patients (75.0% and 80.7 respectively). A significantly elevated serum levels of 1-MA, 1-MG and 8-OHdG in patients with malignant diseases in comparison with benign and healthy individuals. In this study, we found 65.4% of malignant patients had elevated levels of CA15-3. The elevated levels of CA15-3 (the cutoff=31.5 U/l) in healthy control and the benign breast diseases were 44.2% with median (interquartile) range 31.0 (24 -34 U/l) and 45.7% with a range 31.0(25 -36 U/l) respectively. The cutoff of the 1-MA, 1-MG and 8-OHdG was 23.5 µg/l, 18.5 µg/l and 18.0 µg/l respectively. The elevated levels of 1-MA, 1-MG and 8-OHdG in malignant breast diseases patients were 67.3%, 66.4% and 74.8% respectively. On the other hand, the mean concentrations of 1-MA and 8-OHdG are approximately two times above the cutoff values. There are significant differences in the mean of concentrations of CA15-3 ($p \leq 0.001$), 1-MA ($p=0.002$), 1-MG ($p=0.033$) and 8-OHdG ($p=0.024$) in stages of patients with malignant breast diseases.

Several biomarker combinations were found to be significant, especially when comparing the advanced stages to healthy controls and benign breast disease patient samples. For example, the most significant AUC values for CA15-3 were found in comparing the healthy control and benign with stage III (AUC=0.88) of malignant diseases. The most significant AUC values for 1-MA, 1-MG and 8-OHdG were found in comparing the healthy control and benign to stage IV of malignant breast diseases patients. In this study, the AUC of CA15-3 ranged from 0.59 to 0.79, the sensitivity ranged from 61.3% to 65.4% and the specificity ranged from 56.4% to 68.0%, whereas in several studies, the sensitivities of CA15-3 ranged from 32% and 90% and specificities between 71% and 100% [36-39]. In the combination between candidate metabolomes and the CA15-3 using the linear regression (CA15-3 is the most significant breast cancer biomarkers currently used in the clinical diagnosis), the sensitivity increases ranged from 80.3% to 88.8% and specificity ranged from 80.6% to 88.8% with good efficiency (ranged from 81.0% to 88.8%) and highly AUC values (ranged from 0.89 to 0.93).

Our study confirms a clear significant correlation in concentrations of 1-MA and 8-OHdG with tumor size, T-status and tumor stages. There are significant positive correlations between CA15-3 and 1-MA, 1-MG and8-OHdG in agreement with studies [40,41].

CONCLUSION

The candidate biomarkers discussed in this study are accurate correlated with tumor burden, especially in advanced stages. Thus, the serum of 1-MA, 1-MG and 8-OHdG metabolomics are golden biomarkers for staging, diagnosis and

treatment monitoring of breast malignant diseases patients attributed to higher sensitivity and specificity especially in combination with CA15-3.

ACKNOWLEDGMENT

We are grateful to the staff of the Damietta Cancer Institute, Damietta, and Ismailia Oncology Teaching Hospital, Ismailia, Egypt. This work has been completely supported financially and carried out at Central Laboratories of Ministry of Health, Damietta, Egypt**.**

CONFLICTS OF INTEREST

None.

REFERENCES

[1] Darwish A, Hilal A, Aly-Eldin N, Soliman L. 1042 Breast cancer in women aging 35 years old and younger: The Egyptian National Cancer institute experience. *Eur J Cancer*, **2015,** 51(3) (2S159.

[2] DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, et al. Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. *CA Cancer J Clin,* **2016,** 66: 31-42.

[3] Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, et al. Swedish two-county trial: impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology*, **2011,** 260: 58-663.

[4] Attallah AM, El-Far M, Omran MM, Abdallah SO, El-desouky MA, et al. Circulating levels and clinical implications of epithelial membrane antigen and cytokeratin-1 in women with breast cancer: can their ratio improve the results? *Tumor Biol,* **2014,** 35: 10737-10745.

[5] Asiago VM, Alvarado LZ, Shanaiah N, Gowda GN, Owusu-Sarfo K, et al. Early detection of recurrent breast cancer using metabolite profiling. *Cancer Res*. 70(21) (2010) 8309-8318.

[6] Nagrath D, Caneba C, Karedath T, Bellance N, Metabolomics for mitochondrial and cancer studies. Biochimica et Biophysica Acta (BBA)-*Bioenerg,* **2011,** 1807: 650-663.

[7] Armitage EG, Barbas C. Metabolomics in cancer biomarker discovery: current trends and future perspectives. *J. Pharm. Biomed Anal,* **2014,** 87: 1-11.

[8] Bujak R, Struck-Lewicka W, Markuszewski MJ, Kaliszan R. Metabolomics for laboratory diagnostics. *J Pharm Biomed Anal,* **2015,** 113: 108-120.

[9] Zhang A, Sun H, Yan G, Wang P, Wang X, et al. Mass spectrometry□based metabolomics: applications to biomarker and metabolic pathway research. *Biomed Chromatogr,* **2016,** 30: 7-12.

[10] Kouremenos KA, Johansson M, Marriott PJ. Advances in gas chromatographic methods for the identification of biomarkers in cancer. *J Cancer,* **2012,** 3: 404-420.

[11] Zhang A, Sun H, Wang P, Han Y, Wang X. Modern analytical techniques in metabolomics analysis. *Analyst,* **2012,** 137: 293-300.

[12] Dutertre M, Lambert S, Carreira A, Amor-Guéret M, Vagner S. DNA damage: RNA-binding proteins protect from near and far. *Trends Biochem Sci*, **2014,** 39: 141-149.

[13] Struck-Lewicka W, Kordalewska M, Bujak R, Mpanga AY, Markuszewski M. Metabolic fingerprinting using LC–MS and GC–MS reveals metabolite changes in prostate cancer: A pilot study. *J Pharm Biomed Anal,* **2015,** 111: 351-361.

[14] Yomtoubian S, Ryu S, Lee S, Markowitz G, Gao D, et al. EZH2 methyltransferase regulates disseminating tumor cells in breast cancer metastasis. *AACR,* **2017.**

[15] Zhang A, Sun H, Yan G, Wang P, Han Y, et al. Metabolomics in diagnosis and biomarker discovery of colorectal cancer. *Cancer Lett,* **2014,** 345: 17-20.

[16] Günther UL, Metabolomics biomarkers for breast cancer. *Pathobiology,* **2015,** 82: 153-165.

[17] Jobu K, Sun C, Yoshioka S, Yokota j, Onogawa M, et al. Metabolomics study on the biochemical profiles of odor elements in urine of human with bladder cancer. *Biol Pharm Bull,* **2012,** 35: 639-642.

[18] Rasmuson T, Bjork GR, Urinary excretion of pseudouridine and prognosis of patients with malignant lymphoma. *Acta Oncol,* **1995,** 34: 61-67.

[19] La S, Cho J, Kim JH, Kim KM, Capillary electrophoretic profiling and pattern recognition analysis of urinary nucleosides from thyroid cancer patients. *Anal Chim Acta,* **2003,** 486: 171-182.

[20] Ye N, Liu C, Shi P. Metabolomics analysis of cervical cancer, cervical intraepithelial neoplasia and chronic cervicitis by 1H NMR spectroscopy. *Eur J Gynaecol Oncol,* **2015,** 36: 174-180.

[21] Berstein L, Poroshina T, Kovalenko I, Vasilyev D. Serum Levels of 8-Hydroxy-2'-Deoxyguanosine DNA in Patients with Breast Cancer and Endometrial Cancer with and without Diabetes Mellitus. *Bull Exp Biol Med,* **2016,** 161: 547-549.

[22] Trygg J, Gullberg J, Johansson AI, Jonsson P, Antti H, et al. Extraction and GC/MS analysis of the human blood plasma metabolome. *Anal Chem,* **2005,** 77: 8086-8094.

[23] Dunn WB, Broadhurst D, Begley P, Zelena E, Francis-McIntyre S, et al. Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nat Protoc,* **2011,** 6: 1060-1083.

[24] Chen Y, Xu J, Zhang R, Abliz Z. Methods used to increase the comprehensive coverage of urinary and plasma metabolomes by MS. *Bioanalysis*, **2016,** 8: 981-997.

[25] Raftery DM, Pan Z, Gu H. Breast cancer biomarkers and identification methods using NMR and gas chromatography-mass spectrometry. Google Patents, 2015.

[26] Oakman C, Tenori L, iganzoli L, Santarpia L, Cappadona S, et al. Uncovering the metabolomic fingerprint of breast cancer. *The International Journal of Biochemistry & Cell Biology*. **2011,** 43: 1010-1020.

[27] Mishra P, Ambs S. Metabolic signatures of human breast cancer. *Molecular & Cellular Oncology.* **2015,** 2: e992217.

[28] Huang JH, Fu L, Li B, Xie HL, Zhang X, et al. Distinguishing the serum metabolite profiles differences in breast cancer by gas chromatography mass spectrometry and random forest method. *RSC Advances,* **2015,** 5: 58952-58958.

[29] Beretov J, Wasinger V, Graham P, Millar E, Kearsley J, et al. Proteomics for breast cancer urine biomarkers. *Adv Clin Chem.* **2014,** 63: 123-167.

[30] Zhang A, Sun H, Wu X, Wang X. Urine metabolomics. *Clin Chim Acta.* **2012,** 414: 65-69.

[31] Zhang AH, Sun H, Qiu S, Wang XJ, Metabolomics in non-invasive breast cancer. *Clin Chim Acta*. **2013,** 424: 3-7.

[32] Zhang S, Raftery D. Headspace SPME-GC-MS metabolomics analysis of urinary volatile organic compounds (VOCs), Mass Spectrometry in Metabolomics: *Methods and Protocols*. **2014,** 265-272.

[33] Chen J, Hou W, Han B, Liu G, Gong J, et al. Target-based metabolomics for the quantitative measurement of 37 pathway metabolites in rat brain and serum using hydrophilic interaction ultra-high-performance liquid chromatography–tandem mass spectrometry. *Anal Bioanal Chem.* **2016,** 408: 2527-2542.

[34] Falegan OS, Ball MW, Shaykhutdinov RA, Pieroraio PM, Farshidfar F, et al. Urine and serum metabolomics analyses may distinguish between stages of renal cell carcinoma. *Metabolites,* **2016,** 7: 6.

[35] Bouatra S, Aziat F, Mandal R, Guo AC, Knox MC, et al. The human urine metabolome. *PLoS One,* **2013,** 8: e73076.

[36] Stieber P, Molina R, Chan DW, Fritsche HA, Beyrau R, et al. Clinical evaluation of the Elecsys CA 15-3 test in breast cancer patients. *Clin Lab,* 2002, 49: 15-24.

[37] Park BJ, Cha MK, Kim IH. Thioredoxin 1 as a serum marker for breast cancer and its use in combination with CEA or CA15-3 for improving the sensitivity of breast cancer diagnoses. *BMC Res,* **2014,** 7: 7.

[38] Nicolini A, Tartarelli G, Carpi A Metelli MR, Ferrari P, et al. Intensive post-operative follow-up of breast cancer patients with tumour markers: CEA, TPA or CA15. 3 vs MCA and MCA-CA15. 3 vs CEA-TPA-CA15. 3 panel in the early detection of distant metastases. *BMC Cancer,* **2006,** 6: 269.

[39] Di Gioia D, Blankenburg I, Nagel D, Heinemann V, Stieber P, et al. Tumor markers in the early detection of tumor

recurrence in breast cancer patients: CA 125, CYFRA 21-1, HER2 shed antigen, LDH and CRP in combination with CEA and CA 15-3. *Clin Chim Acta,* 461 (2016) 1-7.

[40] Jové M, Collado R, Quiles JL, Ramírez-Tortosa MC, Sol J, et al. Granados-Principal, A plasma metabolomic signature discloses human breast cancer. *Oncotarget,* **2017,** 8: 19522.

[41] Hart CD, Vignoli A, Tenori L, Uy GL, Van To T, et al. Love, Serum metabolomic profiles identify ER-positive early breast cancer patients at increased risk of disease recurrence in a multicenter population. *Clin Cancer Res,* **2017,** 23:1422-1431.