

Serum biomarker panels for diagnosis of gastric cancer

Weihua Tong¹
 Fei Ye²
 Liang He¹
 Lifeng Cui¹
 Miao Cui²
 Yuan Hu²
 Wei Li¹
 Jing Jiang³
 David Y Zhang²
 Jian Suo¹

¹Department of Gastrointestinal Surgery, The First Hospital, Jilin University, Changchun, Jilin, People's Republic of China; ²Department of Pathology, Icahn School of Medicine at Mount Sinai, New York, NY, USA; ³Division of Clinical Epidemiology, The First Hospital, Jilin University, Changchun, Jilin, People's Republic of China

Correspondence: Jian Suo
 Department of Gastrointestinal Surgery,
 The First Hospital, Jilin University, 71
 Xinmin Street, Changchun, Jilin 130021,
 People's Republic of China
 Email suojian0066@126.com

David Y Zhang
 Department of Pathology, Icahn School
 of Medicine at Mount Sinai, Icahn
 (East) Building Floor 9th Room 9-52,
 1425 Madison Avenue, New York,
 NY 10029, USA
 Email david.zhang@mssm.edu

Purpose: Currently, serum biomarkers that are sufficiently sensitive and specific for early detection and risk classification of gastric adenocarcinomas are not known. In this study, ten serum markers were assessed using the Luminex system and enzyme-linked immunosorbent assay for the diagnosis of gastric cancer and analysis of the relation between prognosis and metastases.

Patients and methods: A training set consisting of 228 gastric adenocarcinoma and 190 control samples was examined. A Luminex multiplex panel with nine biomarkers, consisting of three proteins discovered through our previous studies and six proteins previously reported to be cancer-associated, was constructed. One additional biomarker was detected using a commercial kit containing EDTA. Logistic regression, random forest (RF), and support vector machine (SVM) were used to identify the panel of discriminatory biomarkers in the training set. After selecting five proteins as candidate biomarkers, multivariate classification analyses were used to identify algorithms for diagnostic biomarker combinations. These algorithms were independently validated using a set of 57 gastric adenocarcinoma and 48 control samples.

Results: Serum pepsinogen I, serum pepsinogen II, A Disintegrin And Metalloproteinase domain-containing protein 8 (ADAM8), vascular endothelial growth factor (VEGF), and serum IgG to *Helicobacter pylori* were selected as classifiers in the three algorithms. These algorithms differentiated between the majority of gastric adenocarcinoma and control serum samples in the training/test set with high accuracy (RF 79.0%, SVM 83.8%, logistic regression 76.2%). These algorithms also differentiated the samples in the validation set (accuracy: RF 82.5%, SVM 86.1%, logistic regression 78.7%).

Conclusion: A panel of combinatorial biomarkers comprising VEGF, ADAM8, IgG to *H. pylori*, serum pepsinogen I, and pepsinogen II were developed. The use of biomarkers is a less invasive method for the diagnosis of gastric adenocarcinoma. They may supplement clinical gastroscopic evaluation of symptomatic gastric cancer patients and enhance the diagnostic accuracy.

Keywords: gastric adenocarcinoma, cancer diagnosis, cancer screening, Luminex

Introduction

Gastric adenocarcinoma accounts for 40% of all new gastric cancer cases worldwide. It is the fifth most common cancer globally and the third most common cause of cancer-related deaths (estimated 723,000 deaths, 8.8% of the total in 2012),¹ mainly because of a late diagnosis. The disease is particularly prevalent in Far Eastern countries. Gastric adenocarcinoma is the most common malignancy in the People's Republic of China, and was the second leading cause of death by malignancy (14.33%) in 2012.² Gastric cancer patients have considerably lower survival rates than those with other gastroenterological tract cancers, except for cancer of the esophagus. Five-year survival rates of patients with localized disease (stages I, II, N0) (63.2%) decreased

to 28.4% after the cancer spread to regional lymph nodes (stages II, III, N1–N3), and to 3.9% following the detection of distant metastases (stage IV).³

An early diagnosis coupled with good treatment strategy for gastric cancer can significantly improve the survival rates. Esophagogastroduodenoscopy is the most reliable method for diagnosis of gastric adenocarcinoma. It is used broadly in gastroscopy in the People's Republic of China, South Korea, and Japan because of the high incidence of gastric adenocarcinoma in these countries. South Korea and Japan are the only two countries that provide a government-sponsored screening program for gastric cancer. These resulted in low mortality-to-incidence ratio (0.43 in Japan, 0.35 in South Korea), indicating the benefits of population-based screening in high-prevalence regions.¹ The screening programs run in these countries include gastroscopy and serology tests. However, the feasibility and effectiveness of gastroscopy is questionable for the countries with low incidence rates for gastric cancers.

Hence, serum biomarkers that identify patients with high risk for gastric adenocarcinoma would increase the effectiveness of endoscopy and improve the early diagnosis rate. In past decades, studies have revealed several serum biomarkers for gastric cancer which comprise carcinoembryonic antigen, cancer antigen (CA) 19-9, and CA 72-4, among others. However, the sensitivity of these serum biomarkers used during diagnosis of gastric adenocarcinoma (20%–30%) is lower as compared to other cancers.^{4–6} Resent proteomic-based research identified several significantly more selective and specific biomarkers.⁷ Also, miRNAs are emerging as a new potentially important tool in diagnostics of gastric cancer.⁸ These new approaches, however, often require specific technique and expensive instrumentation and cannot be used for routine screening tests.

Traditionally, the detection of serum markers was done using the enzyme-linked immunosorbent assay (ELISA). However, the disadvantage of this method is that it requires larger volumes of sample and incurs high costs compared to the recent method of multiplex detection. The Luminex® xMAP® technology enables the large numbers of biological tests to be conducted and analyzed quickly, cost-effectively, and accurately as compared to the ELISA tests.

In this study, we developed a diagnostic serum biomarker panel using Luminex multiplex technology and validated its performance by differentiating the patients with gastric adenocarcinoma from the healthy controls.

Materials and methods

Patient samples

During 2008–2010, serum samples were collected from patients (n=285) with newly diagnosed primary gastric

adenocarcinoma without the presence of other cancers at The First Hospital of Jilin University, Changchun, People's Republic of China. All gastric cancer patients underwent radical tumor resection with histologically confirmed diagnosis of gastric adenocarcinoma. Control samples (n=238) were collected from attendees of the health check-up program from the same hospital between 2008 and 2010. Patients from the control group had no history of cancer. All participants underwent: 1) routine medical history and physical examination; 2) routine blood and *Helicobacter pylori* IgG tests; 3) chest X-ray; 4) abdominal ultrasound investigation or computed tomography scan; 5) gastroduodenoscopy; and 6) mammography in women. Two hundred and thirty-eight control samples without inflammatory conditions, or confirmed or suspected cancer that needed medical treatment, were included. All blood samples were collected before any medical treatment in the early morning. Peripheral blood was collected and stored in SST II tubes (BD, Franklin Lakes, NJ, USA) at room temperature for 1 hour, then centrifuged at 3,000×g for 5 minutes. Supernatants were collected, separated in aliquots, and stored at –80°C. All participants signed an informed consent form. The study was approved by the Institutional Review Board at The First Hospital of Jilin University.

Clinical data on demographics and pathological characteristics were available for all patients whose serum samples were used for this study (Table 1). The T status, N status, and TNM stage of each tumor were classified according to the seventh edition of the AJCC classification.⁹

Samples were randomly divided into two datasets. Serum samples from 228 patients with gastric adenocarcinoma and 190 control serum samples were grouped as a training set. Serum samples from 48 patients with gastric adenocarcinoma and 47 control serum samples were grouped as an independent validation set. Patient demographics and clinical profiles are presented in Table 2.

Construction of 9-plex bead array panel

The xMAP bead-based technology (Luminex Corp., Austin, TX, USA) permits simultaneous analysis of multiple analytes in a single run. It is an antibody-based microarray, with high sensitivity and high throughput advantage, which has been used extensively in clinical diagnosis and research work. Among the nine target proteins (Table 3) chosen for this project, COX2, ADAM8, and beta-catenin were identified in our previous studies.^{10,11} There was difference in the expression of beta-catenin between the gastric cancer tissue and normal gastric mucosa tissue. COX2 and ADAM8 were higher in the gastric cancer tissues with lymph node metastases than in those without metastasis. Serum pepsinogen I (PGI) and

Table 1 Demographic data and pathological characteristics of patients with gastric cancer and healthy controls

	Gastric adenocarcinoma, n=285	Control, n=238
Age, years	59.89±11.499	59.14±9.33
Sex		
Males	204	157
Females	81	81
Stage		
I	66 (23.1%)	
II	93 (32.6%)	
III	126 (44.2%)	
Alcohol and smoking history		
Alcohol	86 –199	
Smoking	130 –155	
Early stage	48	
Advanced stage	237	
Differentiation type		
High	7	
Medium	120	
Low	158	
Tumor		
T1	48	
T2	55	
T3	3	
T4	179	
Lymph node		
N0	120	
N1	40	
N2	47	
N3	78	
Tumor ≤2 cm	52	
Tumor >2 cm	233	
Tumor position		
Antrum	188	
Body	65	
Cardia	32	

serum pepsinogen II (PGII), as well as PGI/PGII ratio were used in the screening of gastric cancer in the countries with high and moderate incidence rates, such as the People's Republic of China, Japan, and South Korea. Most studies have demonstrated that low concentrations of PGI and low

PGI/PGII ratios in the serum or plasma are indicators of atrophic gastritis, which is linked to elevated risk of gastric cancer.^{12–15} *H. pylori* and gastric atrophy have both been identified as etiological factors for gastric cancer. *H. pylori* was declared as a class I carcinogen by the International Agency for Research on Cancer in 1994.¹⁶ Epidemiological surveys support a positive association between *H. pylori* infection and the development of gastric cancer.^{17–19} In this study, all patients and health controls had been tested for *H. pylori* IgG with ELISA kit (Biohit ELISA kit; Biohit, Helsinki, Finland). Intercellular cell adhesion molecule (ICAM), vascular endothelial growth factor (VEGF), epidermal growth factor receptor (EGFR), and hemoglobin were also associated with gastric cancer development and metastasis in previous studies.^{20–26}

All antibodies have been validated in ELISA and were used according to the manufacturer's instructions. The antigen protein has been validated for ELISA too. We developed all kits for Luminex according to The Luminex Cookbook.²⁷ All nine protein kits were divided into a 5-plex panel and two 2-plex panels with different dilution rates for each patients' samples. All samples were randomly divided into different plates to perform the Luminex test. The random number allocation and grouping was done using STATA 12 software.

Multiplex assay procedure

Multiplex assay was performed on a COSTAR™ 96-well round bottom plate (Corning Incorporated, Corning, NY, USA). Standard curve samples and prediluted patient samples (50 µL) were dispensed into the wells. Capture antibody-beads mixtures (50 µL) were added into each well and incubated at room temperature for 2 hours. Next, magnetic beads were suspended in 50 µL of 1% bovine serum albumin (BSA) in phosphate-buffered saline (PBS), then 50 µL of detection antibodies with biotin were added to each well and the mixtures incubated at room temperature for 2 hours. Finally, 100 µL of streptavidin-phycoerythrin

Table 2 Patients' demographic data and clinical profiles in training and validation groups

	Training group		Validation group	
	Gastric carcinoma (228)	Control (190)	Gastric carcinoma (57)	Control (48)
Age, years	59.82±11.32	59.15±9.27	60.14±10.73	59.11±7.93
Sex				
Male	164	120	40	36
Female	64	70	17	12
Stage				
I	47		19	
II	78		15	
III	103		23	

Table 3 List of markers

Biomarker	Antigen	Capture antibody	Detection antibody
Growth/angiogenic factor			
VEGF	R&D System	R&D System	R&D System
EGFR	R&D System	R&D System	R&D System
Coagulation/thrombosis			
Hemoglobin	Sigma-Aldrich	Pierce	Pierce
Adhesion			
ADAM8	R&D System	R&D System	R&D System
Beta-catenin	R&D System	R&D System	R&D System
sICAM-I	R&D System	R&D System	R&D System
Inflammation			
COX2	R&D System	R&D System	R&D System
Pepsinogen			
Pepsinogen-I	R&D System	Abcam	Abcam
Pepsinogen-II	Biohit	NOVUS	NOVUS

Notes: R&D System: R&D System Inc., Minneapolis, MN, USA; Sigma-Aldrich: Sigma-Aldrich Co., St Louis, MO, USA; Pierce: ; Biohit: Biohit, Helsinki, Finland; Abcam: Abcam, MA, USA; NOVUS: Novus Biologicals, Littleton, CO, USA.

Abbreviations: ICAM, intercellular cell adhesion molecule; VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor; ADAM8, A Disintegrin And Metalloproteinase domain-containing protein 8.

(4 µg/mL) were added and the mixtures incubated at room temperature for 30 minutes. Each step was followed by a double washing step using 1% BSA in PBS (Sigma-Aldrich Co., St Louis, MO, USA) on the magnetic plate (Alpaqua Engineering, LLC, Beverly, MA, USA). All incubations were done using the Digene Hybrid Capture System at 800 rpm.

Subsequent to the final washing step, samples were resuspended in 100 µL of 1% BSA in PBS and read using Luminex-200™ (Luminex Corp.). The standard curve was constructed using a seven-parametric-curve fitting, and the results were calculated using Xponent 3.1 (Luminex Inc.). Markers were grouped together according to the dilution factor after the cross-reactivity was checked across all analytes. During the experiment, intra-assay precision ranged from 2% to 16%, and interassay precision ranged from 6% to 19%. The acceptance criteria for each individual run followed the Westgard rule.²⁸

Statistical analysis

The raw data of biomarkers from Luminex combined with the *H. pylori* IgG concentrations were analyzed by comparison between the gastric cancer patients group and the control group using χ^2 , the z-statistic, and the Student's *t*-test. Further, the univariate and multivariate logistic regression analysis and correlations between biomarkers were analyzed. The analysis of correlations between biomarkers was performed as well. For the identification of algorithms distinguishing controls from patients with gastric carcinoma, three classification analysis methods – random forest (RF), support vector

machine (SVM), and logistic regression – were performed with the R program package. The R project of statistical computing (Wirtschaftsuniversität, Wien, Austria) was used for classification analysis to develop the algorithms.

The RF ensemble learning method, proposed by Breiman, is a combination of tree predictors where each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest.²⁹ The idea behind SVMs is to construct a separating hyperplane or set of hyperplanes in a high or infinite dimensional space for classification.³⁰ After obtaining results with the training set, each classification algorithm with different groups of classifiers was cross-validated with the validation set. The prediction performances with accuracy (number of patients or controls identical to the result of the classification/examined total number), sensitivity, and specificity were analyzed. The receiver operating characteristic curves were constructed and the areas under the curves were calculated. A *P*-value of 0.05 was considered statistically significant.

Results

The results of nine Luminex test biomarkers and the ELISA result of IgG to *H. pylori* are listed in Table 4 for the gastric cancer group and the control group. With the exception of EGFR (*P*=0.1951) and ICAM (*P*=0.2627), there were significant statistical differences in the biomarkers' levels (*P*<0.05). The levels of eight markers were higher in the gastric cancer group compared to the control group.

The correlation analysis for ten biomarkers showed that there were significant correlations between beta-catenin, COX2, and VEGF (*r*>0.7 for all three pairs, Table 5).

Since ten proteins are too many to run a single algorithm, the feature selection process was performed using RF. This function shows the cross-validated prediction performance of models with sequentially reduced numbers of predictors (ranked by variable importance) via a nested cross-validation procedure (Figure 1). The figure shows that five proteins (IgG to *H. pylori*, ADAM8, PGI, PGII, VEGF) are the best for the prediction. Biomarker selection is shown in Table 6, and the important scores are plotted in Figure 2. The results for three classification analysis methods – RF, SVM, and logistic regression – are shown in Table 7. The receiver operating characteristics of RF, SVM, and logistic regression algorithm of the five biomarkers are shown in Figure 3. In the test group, the accuracy of the RF and SVM algorithms for this five-biomarker panel was 82.5% and 86.1%, respectively, which was higher than the accuracy of logistic regression (78.7%).

Table 4 The description of variables in gastric cancer patients and controls

Variable	–	Control (N=238)	Gastric cancer (N=285)	Statistic	P-value
IgG to <i>Helicobacter pylori</i>	Mean ± SD	35.14±37.98	66.96±42.19	z=8.90	<0.0001
	M (Q1–Q3)	16.67 (6.31–60.64)	66.44 (27.94–107.41)		
	Min–Max	0.46–142.71	0.60–155.67		
PGI	Mean ± SD	101.15±36.49	4,740.89±22,768.80	z=4.59	<0.0001
	M (Q1–Q3)	94.05 (77.37–120.19)	116.66 (74.28–211.51)		
	Min–Max	13.54–286.47	0.24–197,160.00		
PGII	Mean ± SD	13.96±9.98	3,422.14±10,146.89	z=9.12	<0.0001
	M (Q1–Q3)	10.91 (7.38–17.38)	21.66 (12.13–38.19)		
	Min–Max	1.60–66.94	0.14–90,759.61		
PGI/PGII	Mean ± SD	9.13±4.07	5.71±4.07	z=10.26	<0.0001
	M (Q1–Q3)	8.56 (6.26–10.93)	5.23 (3.03–7.61)		
	Min–Max	2.22–36.76	0.04–24.49		
Beta-catenin	Mean ± SD	409.20±473.89	767.73±854.38	z=6.88	<0.0001
	M (Q1–Q3)	269.98 (135.91–460.89)	483.61 (263.08–883.18)		
	Min–Max	1.00–3,210.81	1.00–5,456.21		
COX2	Mean ± SD	190.31±210.51	308.00±320.50	z=4.95	<0.0001
	M (Q1–Q3)	146.60 (38.79–260.89)	221.51 (97.94–422.12)		
	Min–Max	0.34–1,325.78	1.00–2,481.92		
ADAM8	Mean ± SD	500.78±492.02	1,146.98±1,161.13	z=10.11	<0.0001
	M (Q1–Q3)	337.12 (199.81–627.00)	803.82 (466.89–1,467.16)		
	Min–Max	2.65–3,168.90	1.00–9,645.67		
VEGF	Mean ± SD	242.79±277.96	548.92±490.40	z=10.23	<0.0001
	M (Q1–Q3)	164.49 (77.28–290.32)	416.47 (236.06–707.36)		
	Min–Max	1.00–1,962.88	1.00–3,653.43		
EGFR	Mean ± SD	9,308.89±4,302.37	9,837.98±5,024.02	t=1.30	0.1951
	M (Q1–Q3)	8,211.83 (6,565.46–10,832.58)	9,029.69 (6,786.17–11,484.39)		
	Min–Max	3,897.56–41,378.60	1.00–47,739.98		
ICAM	Mean ± SD	174,870.10±86,585.45	165,658.82±95,597.97	t=1.15	0.2527
	M (Q1–Q3)	151,346.57 (119,159.57–213,194.11)	144,412.52 (107,495.51–200,256.92)		
	Min–Max	56,918.07–627,365.84	1.00–1,101,730.81		
HB	Mean ± SD	369,038.51±470,630.63	939,135.01±3,663,960.29	z=2.42	0.0157
	M (Q1–Q3)	235,096.52 (135,491.09–417,276.58)	268,581.00 (144,175.20–724,694.68)		
	Min–Max	31,780.94–5,466,994.12	1.00–50,057,376.68		

Abbreviations: HB, hemoglobin; M (Q1–Q3), median and interquartile range; Min–Max, minimum to maximum; PGI, serum pepsinogen I; PGII, serum pepsinogen II; SD, standard deviation; ICAM, intercellular cell adhesion molecule; VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor; ADAM8, A Disintegrin And Metalloproteinase domain-containing protein 8.

The Cox model was used to model the relationship between the survival time, disease stage, and covariates, including the five proteins identified earlier. Conditional censoring distribution bootstrapping was used to construct the 95% confidence interval. The data show that proteins have no effect on survival time, while pathology N stage pN and pathology T stage pT had significant effect on survival (Table 8).

Discussion

This study used Luminex and ELISA methods to identify ten serum biomarkers to distinguish between gastric cancer

patients and healthy controls. Among the biomarkers included in the panel, eight biomarkers demonstrated statistically significant differences between the gastric cancer group and the healthy control group. Additionally, in this study we tested three proteins that we have earlier reported to have higher expression in the gastric cancer tissue compared to the normal gastric mucosa (beta-catenin), or higher expression in the gastric cancer tissue with metastases compared to the gastric cancer tissue without metastases (COX2, ADAM8).^{10,11} The levels of these three proteins were higher in the gastric cancer patients' serum than in the healthy controls. The proteomics studies provide the basics for high-sensitivity tests like the Luminex. In this study, using the computing algorithm combined with high-throughput method, we identified five biomarkers for gastric cancer among the ten selected proteins.

A number of research studies have been aimed toward identifying the early diagnostic biomarkers for gastric cancers

Table 5 The results of correlation analysis for three variables

	Beta-catenin	COX2	VEGF
Beta-catenin	1.0000		
COX2	0.8370*	1.0000	
VEGF	0.7893*	0.7291*	1.0000

Note: *The correlation is statistically significant.

Abbreviation: VEGF, vascular endothelial growth factor.

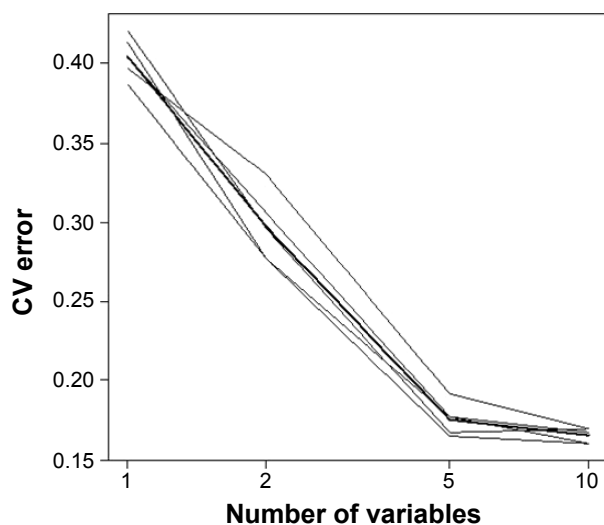


Figure 1 Random forest cross-validation for feature selection.
Abbreviation: CV, cross-validation.

in blood serum, gastric juice, and even in the exhalants from the lungs.^{31–33} However, to date, there is no single known biomarker unique for diagnosis of gastric cancer. The traditional biomarkers such as carcinoembryonic antigen, CA 19-9, CA 12-5, and CA 7-24 are more sensitive in the diagnostics of other gastroenterological cancers rather than for gastric cancer.^{4–6} For this reason, they were excluded from this study. Instead, we tried to test other proteins that are known to be associated with tumorigenesis and development of gastric cancer. We used three algorithms to combine five biomarkers (IgG to *H. pylori*, ADAM8, PGI, PGII, and VEGF) in a panel cross-validated in a training group of 418 samples and a test group of 105 samples. SVM and RF resulted in higher accuracy and sensitivity compared to the logistic regression method. For SVM, the accuracy, sensitivity, and specificity in the test group were 86.1%, 88.6%, and 83.2%, respectively. Both PGI and PGII are being used in the gastric

Table 6 Biomarker selection in the training group

Marker	Avg.imp
ADAM8	42.6065
PGI	38.7232
PGII	37.7563
VEGF	35.1600
IgG to <i>Helicobacter pylori</i>	34.7993
EGFR	10.8508
Beta-catenin	10.3391
HB	8.4268
ICAM	6.8035
COX2	2.1472

Abbreviations: HB, hemoglobin; ICAM, intercellular cell adhesion molecule; VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor; ADAM8, A Disintegrin And Metalloproteinase domain-containing protein 8; PGI, serum pepsinogen I; PGII, serum pepsinogen II.

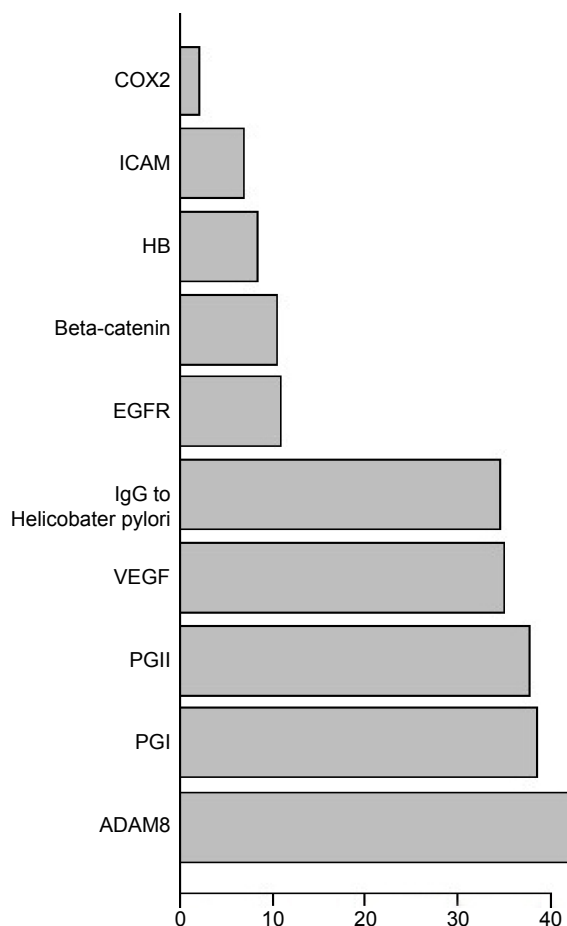


Figure 2 Importance scores of biomarkers in random forest.

Abbreviations: ICAM, intercellular cell adhesion molecule; VEGF, vascular endothelial growth factor; EGFR, epidermal growth factor receptor; ADAM8, A Disintegrin And Metalloproteinase domain-containing protein 8; HB, hemoglobin; IgG, immunoglobulin G; PGII, serum pepsinogen II; PGI, serum pepsinogen I.

cancer screening in the countries with high incidence of this disease. *H. pylori* infection is known to be a major cause of peptic ulcer development and, eventually, gastric cancer development in Asian countries. The infection rate and the carrier rate are higher in Asian populations than elsewhere in the world. With the *H. pylori* infection, gastric mucosa cells can produce more COX2, and beta-catenin is known to be involved in this process. In turn, COX2 and beta-catenin direct the production of VEGF. We detected significant correlation between protein levels of beta-catenin, COX2, and VEGF in our analysis. Moreover, VEGF and ADAM8 were associated with tumor metastases. Thus, the five selected biomarkers have a significant role to play during gastric carcinogenesis.

In this study, SVM algorithm demonstrated the best performance. It had the highest accuracy, sensitivity, and specificity as compared to RF and logistic regression analysis. The five selected biomarkers were not associated with the prognosis of gastric cancer patients. In addition, the overall

Table 7 Diagnostic performance of classification algorithms with five-biomarker panel

	Training group			Test group		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
RF	0.790	0.842	0.729	0.825	0.860	0.784
SVM	0.838	0.877	0.792	0.861	0.886	0.832
Logistic	0.762	0.807	0.708	0.787	0.829	0.737

Abbreviations: Logistic, logistic regression; RF, random forest; SVM, support vector machine.

sensitivity was sustained regardless of TNM stage and tumor size. Although none of the individual markers showed sufficient diagnostic power independently, the biomarker panel identified in this study performed well. SVM algorithm demonstrated the best performance in this study. The

concentrations of proteins that we identified in this study are much higher than the concentrations of other cytokines and low level biomarkers that are usually detected with mass spectroscopy in serum samples. Therefore, the panel of biomarkers may have clinical applications and can be

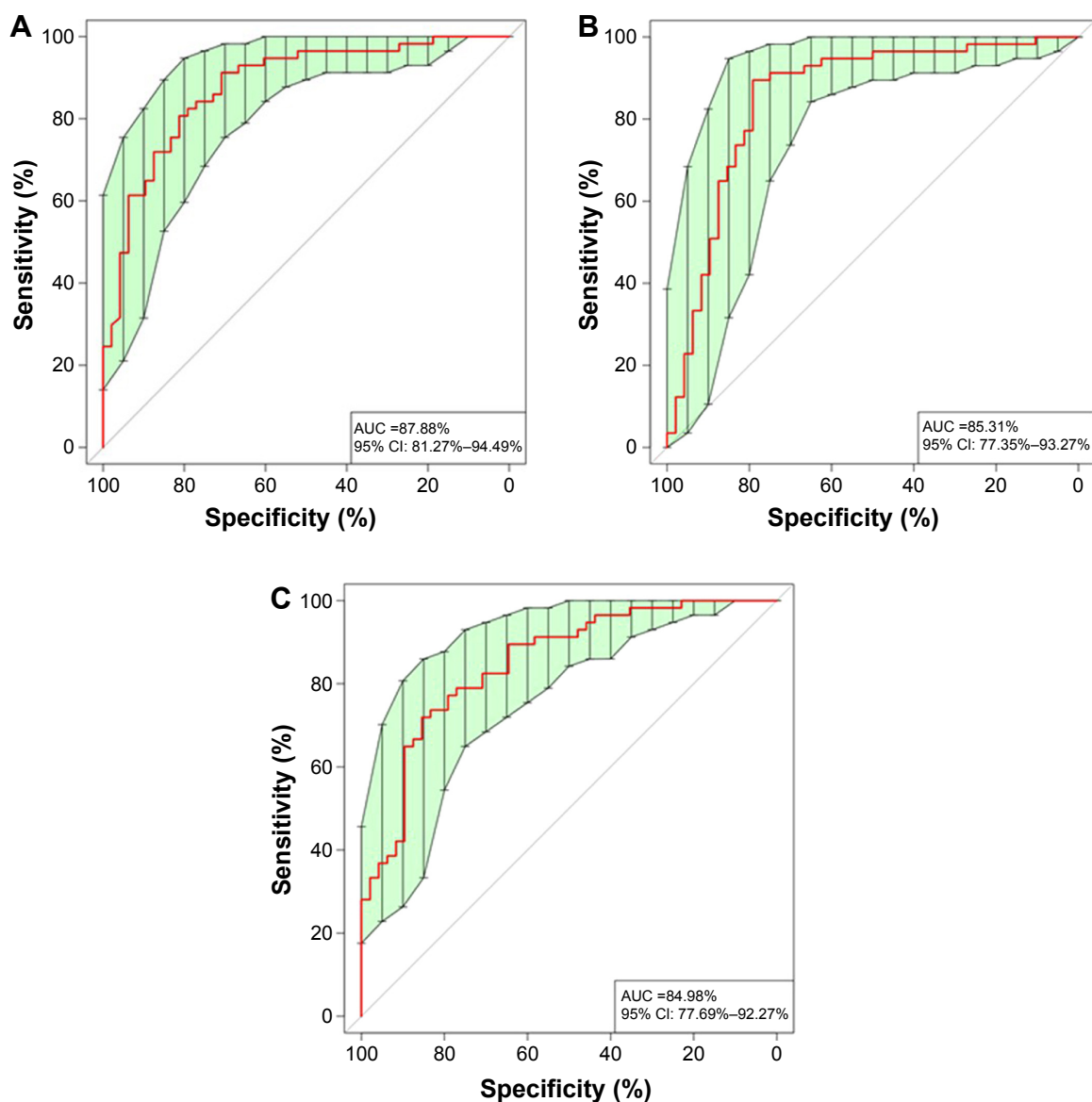


Figure 3 The ROC of five biomarkers.

Notes: (A) RF. (B) SVM. (C) Logistic regression.

Abbreviations: AUC, area under the curve; CI, confidence interval; ROC, receiver operating characteristic; RF, random forest; SVM, support vector machine.

Table 8 Cox model coefficients and 95% confidence interval

	OR	2.5% OR	97.5% OR
ADAM8	1.118	0.754	1.591
PGI	0.935	0.640	1.444
PGII	1.035	0.733	1.392
VEGF	1.077	0.795	1.541
IgG to <i>Helicobacter pylori</i>	1.170	0.937	1.544
N1	0.726	0.000	2.086
N2	3.817	1.841	10.506
N3	5.998	3.110	15.813
T1	1.365	0.180	38,189,896.926
T2	15.098	0.000	789,835,105.698
T3	5.680	1.800	141,247,321.593
T4	11.395	2.777	391,254,179.091

Abbreviations: OR, odds ratio; PGI, serum pepsinogen I; PGII, serum pepsinogen II; VEGF, vascular endothelial growth factor; ADAM8, A Disintegrin And Metalloproteinase domain-containing protein 8.

particularly useful for the preliminary diagnostics in patients with reported symptoms.

Conclusion

Using three diagnostic algorithms – RF, SVM, and logistic regression – we created and validated a panel of biomarkers that might prove useful in the clinical diagnostics and/or screening of gastric cancer. The panel includes five biomarkers: IgG to *H. pylori*, ADAM8, PGI, PGII, and VEGF. Among the diagnostic algorithms, SVM out-performed the others regardless of TNM stage or tumor size. In future, the developed panel of biomarkers may supplement clinical gastroscopic evaluation of symptomatic patients and enhance diagnostic accuracy.

Disclosure

The authors report no conflicts of interest in this work.

References

- GLOBOCAN 2012: Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 [webpage on the Internet]. Lyon: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr>. Accessed June 14, 2015.
- 2012 Chinese Cancer Registry Report. The Military Publishing House, China; 2013.
- Cancer Stat Fact Sheets [webpage on the Internet]. Bethesda: National Cancer Institute. Available from: <http://seer.cancer.gov/statfacts/>. Accessed April 4, 2015.
- Shimada H, Noie T, Ohashi M, Oba K, Takahashi Y. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association. *Gastric Cancer*. 2014;17:26–33.
- Pectasides D, Mylonakis A, Kostopoulou M, et al. CEA, CA 19-9, and CA-50 in monitoring gastric carcinoma. *Am J Clin Oncol*. 1997;20:348–353.
- Fan B, Xiong B. Investigation of serum tumor markers in the diagnosis of gastric cancer. *Hepatogastroenterology*. 2011;58:239–245.
- Lin LL, Huang HC, Juan HF. Discovery of biomarkers for gastric cancer: a proteomics approach. *J Proteomics*. 2012;75:3081–3097.

- Wu HH, Lin WC, Tsai KW. Advances in molecular biomarkers for gastric cancer: miRNAs as emerging novel cancer markers. *Expert Rev Mol Med*. 2014;16:e1.
- Washington K. 7th edition of the AJCC cancer staging manual: stomach. *Ann Surg Oncol*. 2010;17:3077–3079.
- Li W, Ye F, Wang D, et al. Protein predictive signatures for lymph node metastasis of gastric cancer. *Int J Cancer*. 2013;132:1851–1859.
- Wang D, Ye F, Sun Y, et al. Protein signatures for classification and prognosis of gastric cancer a signaling pathway-based approach. *Am J Pathol*. 2011;179:1657–1666.
- Kudo T, Kakizaki S, Sohara N, et al. Analysis of ABC (D) stratification for screening patients with gastric cancer. *World J Gastroenterol*. 2011;17:4793–4798.
- Leung WK, Wu MS, Kakugawa Y, et al. Screening for gastric cancer in Asia: current evidence and practice. *Lancet Oncol*. 2008;9:279–287.
- Cao XY, Jia ZF, Jin MS, et al. Serum pepsinogen II is a better diagnostic marker in gastric cancer. *World J Gastroenterol*. 2012;18:7357–7361.
- Barh D, Carpi A, Verma M, Gunduz M, editors. *Cancer Biomarkers: Minimal and Noninvasive Early Diagnosis and Prognosis*. CRC Press, Taylor & Francis Group; 2014.
- IARC. Schistosomes, liver flukes and *Helicobacter pylori*. IARC Working Group on the evaluation of carcinogenic risks to humans, Lyon, 7–14. *IARC Monogr Eval Carcinog Risks Hum*. 1994;61:1–241.
- Malfertheiner P, Sipponen P, Naumann M, et al. *Helicobacter pylori* eradication has the potential to prevent gastric cancer: a state-of-the-art critique. *Am J Gastroenterol*. 2005;100:2100–2115.
- Kim SS, Ruiz VE, Carroll JD, Moss SF. *Helicobacter pylori* in the pathogenesis of gastric cancer and gastric lymphoma. *Cancer Lett*. 2011;305:228–238.
- Lamb A, Chen LF. Role of the *Helicobacter pylori*-induced inflammatory response in the development of gastric cancer. *J Cell Biochem*. 2013;114:491–497.
- Jung WC, Jang YJ, Kim JH, et al. Expression of intercellular adhesion molecule-1 and e-selectin in gastric cancer and their clinical significance. *J Gastric Cancer*. 2012;12:140–148.
- Nakata B, Hori T, Sunami T, et al. Clinical significance of serum soluble intercellular adhesion molecule 1 in gastric cancer. *Clin Cancer Res*. 2000;6:1175–1179.
- Alexiou D, Karayiannakis AJ, Syrigos KN, et al. Clinical significance of serum levels of E-selectin, intercellular adhesion molecule-1, and vascular cell adhesion molecule-1 in gastric cancer patients. *Am J Gastroenterol*. 2003;98:478–485.
- Liu L, Ma XL, Xiao ZL, Li M, Cheng SH, Wei YQ. Prognostic value of vascular endothelial growth factor expression in resected gastric cancer. *Asian Pac J Cancer Prev*. 2012;13:3089–3097.
- Zhao R, Liu XG, Wu XP, et al. Vascular endothelial growth factor (VEGF) enhances gastric carcinoma invasiveness via integrin alpha(v) beta6. *Cancer Lett*. 2010;287:150–156.
- Jung DH, Lee HJ, Han DS, et al. Impact of perioperative hemoglobin levels on postoperative outcomes in gastric cancer surgery. *Gastric Cancer*. 2012;16:377–382.
- Terashima M, Kitada K, Ochiai A, et al. Impact of expression of human epidermal growth factor receptors EGFR and ERBB2 on survival in stage II/III gastric cancer. *Clin Cancer Res*. 2012;18:5992–6000.
- Luminex. The xMAP® Cookbook, 2nd Edition. Austin: Luminex. Available from: <http://info2.luminexcorp.com/download-the-xmap-cookbook>. Accessed April 4, 2015.
- Westgard JO, Barry PL, Hunt MR, Groth T. A multi-rule Shewhart chart for quality control in clinical chemistry. *Clin Chem*. 1981;27:493–501.
- Breiman L. Random Forests. *Mach Learn*. 2001;45:5–32.
- Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M, Haussler D. Support vector machine classification and validation of cancer tissue samples using microarray expression data. *Bioinformatics*. 2000;16:906–914.

31. Cui L, Zhang X, Ye G, et al. Gastric juice MicroRNAs as potential biomarkers for the screening of gastric cancer. *Cancer*. 2013;119:1618–1626.
32. Liu HS, Xiao HS. MicroRNAs as potential biomarkers for gastric cancer. *World J Gastroenterol*. 2014;20:12007–12017.
33. Xu ZQ, Broza YY, Ionsecu R, et al. A nanomaterial-based breath test for distinguishing gastric cancer from benign gastric conditions. *Br J Cancer*. 2013;108:941–950.

OncoTargets and Therapy

Publish your work in this journal

OncoTargets and Therapy is an international, peer-reviewed, open access journal focusing on the pathological basis of all cancers, potential targets for therapy and treatment protocols employed to improve the management of cancer patients. The journal also focuses on the impact of management programs and new therapeutic agents and protocols on

Submit your manuscript here: <http://www.dovepress.com/oncotargets-and-therapy-journal>

patient perspectives such as quality of life, adherence and satisfaction. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Dovepress