

## A Novel Serum Metabolomics-Based Diagnostic Approach to Pancreatic Cancer

Takashi Kobayashi<sup>1</sup>, Shin Nishiumi<sup>1</sup>, Atsuki Ikeda<sup>1</sup>, Tomoo Yoshie<sup>1</sup>, Aya Sakai<sup>1</sup>, Atsuki Matsubara<sup>1</sup>, Yoshihiro Izumi<sup>1</sup>, Hidetaka Tsumura<sup>6</sup>, Masahiro Tsuda<sup>6</sup>, Hogara Nishisaki<sup>6</sup>, Nobuhide Hayashi<sup>5</sup>, Seiji Kawano<sup>5</sup>, Yutaka Fujiwara<sup>2</sup>, Hironobu Minami<sup>2</sup>, Tadaomi Takenawa<sup>4</sup>, Takeshi Azuma<sup>1</sup>, and Masaru Yoshida<sup>1,4,3</sup>

### Abstract

**Background:** To improve the prognosis of patients with pancreatic cancer, more accurate serum diagnostic methods are required. We used serum metabolomics as a diagnostic method for pancreatic cancer.

**Methods:** Sera from patients with pancreatic cancer, healthy volunteers, and chronic pancreatitis were collected at multiple institutions. The pancreatic cancer and healthy volunteers were randomly allocated to the training or the validation set. All of the chronic pancreatitis cases were included in the validation set. In each study, the subjects' serum metabolites were analyzed by gas chromatography mass spectrometry (GC/MS) and a data processing system using an in-house library. The diagnostic model constructed via multiple logistic regression analysis in the training set study was evaluated on the basis of its sensitivity and specificity, and the results were confirmed by the validation set study.

**Results:** In the training set study, which included 43 patients with pancreatic cancer and 42 healthy volunteers, the model possessed high sensitivity (86.0%) and specificity (88.1%) for pancreatic cancer. The use of the model was confirmed in the validation set study, which included 42 pancreatic cancer, 41 healthy volunteers, and 23 chronic pancreatitis; that is, it displayed high sensitivity (71.4%) and specificity (78.1%); and furthermore, it displayed higher sensitivity (77.8%) in resectable pancreatic cancer and lower false-positive rate (17.4%) in chronic pancreatitis than conventional markers.

**Conclusions:** Our model possessed higher accuracy than conventional tumor markers at detecting the resectable patients with pancreatic cancer in cohort including patients with chronic pancreatitis.

**Impact:** It is a promising method for improving the prognosis of pancreatic cancer via its early detection and accurate discrimination from chronic pancreatitis. *Cancer Epidemiol Biomarkers Prev*; 22(4); 571–9. ©2013 AACR.

### Introduction

Pancreatic cancer is characterized by rapid tumor progression and early metastasis, and is one of the leading causes of cancer-related death. Pancreatic cancer is considered to be a lethal solid tumor, and its 5-year survival rate is less than 5%. Although the only curative treatment for pancreatic cancer is surgical resection, more than 80% of patients with pancreatic cancer have a locally advanced or metastatic tumor that is unresectable at the time of

diagnosis (1–3). The clinical symptoms of pancreatic cancer are usually unremarkable until the cancer has progressed to an advanced stage. Furthermore, there is no effective method for its early detection. CA19-9, which is usually used as a tumor marker, is unsuitable for the early detection of pancreatic cancer due to the low sensitivity for resectable stages of disease. In addition, its levels are also increased in other gastrointestinal malignancies and benign pancreaticobiliary diseases such as pancreatitis and cholangitis (4). Imaging examinations are not cost-effective and cannot discriminate pancreatic cancer from benign pancreatic diseases such as tumor-forming pancreatitis. Moreover, exposure to radiation during unnecessary repeated computed tomography (CT) scans might also increase the risk of malignancy (5–7). Endoscopic examinations are not appropriate for screening because of their low throughput and the high risk of complications. Therefore, a novel screening and diagnostic method for pancreatic cancer is required. Screening tests for the general population have to be accurate, safe, and convenient, as well as capable of high throughput. Many researchers have devoted considerable effort to discovering biomarkers of pancreatic cancer (8–10), and CA242 (11), M2-pyruvate kinase (12), PBF-4 (13), MIC-1 (14),

**Authors' Affiliations:** <sup>1</sup>Divisions of Gastroenterology and <sup>2</sup>Medical Oncology/Hematology, and <sup>3</sup>Metabolomics Research, Department of Internal Medicine and <sup>4</sup>The Integrated Center for Mass Spectrometry, Kobe University Graduate School of Medicine; <sup>5</sup>Department of Clinical Laboratory, Kobe University Hospital; and <sup>6</sup>Division of Gastroenterology, Hyogo Cancer Center, Akashi, Hyogo, Japan

**Note:** Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (<http://cebp.aacrjournals.org/>).

**Corresponding Author:** Masaru Yoshida, Division of Gastroenterology, Division of Metabolomics Research, The Integrated Center for Mass Spectrometry, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chu-o-ku, Kobe, Hyogo 6500017, Japan. Phone: 78-382-6305; Fax: 78-382-6309; E-mail: myoshida@med.kobe-u.ac.jp

doi: 10.1158/1055-9965.EPI-12-1033

©2013 American Association for Cancer Research.

PNA-binding glycoprotein (15), hTert (16), MMP-2 (17), and synuclein- $\gamma$  (18) have been reported as candidate biomarkers, but they were not clinically superior to CA19-9 or other clinical examinations.

We have showed the use of metabolomics for diagnosing and evaluating the pathologic conditions of various diseases (19–24). Metabolomics is the comprehensive study of low molecular weight metabolites, which are the endpoint of the omics cascade and hence are the closest point in the cascade to the phenotype. The metabolome represents the metabolic profile of a cell, tissue, organ, or organism. Changes in metabolism result in alterations in the abundance of metabolites, and elucidating the metabolomic changes that occur in a particular disease will increase our understanding of it. Therefore, metabolomics has recently developed rapidly in the medical field (25). Capillary electrophoresis mass spectrometry (CE/MS)-based metabolomic analysis of gastrointestinal cancer tissue revealed that the nutritional conditions in the cancer microenvironment are quite different from those in normal tissue (26). Regarding the pancreas, we previously found that human serum metabolomics using gas chromatography/mass spectrometry (GC/MS) is able to discriminate patients with pancreatic cancer from healthy volunteers (20). Similar results were obtained from the nuclear magnetic resonance (NMR)-based metabolic profiling of rat tissues (27) and NMR-based serum metabolomics in humans (28). In this study, we constructed a diagnostic model for pancreatic cancer using GC/MS-based human serum metabolomics and then confirmed its diagnostic performance via validation analysis and comparisons with conventional tumor markers.

## Materials and Methods

### Subjects

This study was approved by the ethics committee at Kobe University Graduate School of Medicine (Hyogo, Japan) and conducted between Feb 2009 and Feb 2012. The human samples were used in accordance with the guidelines of Kobe University Hospital, and written informed consent was obtained from all subjects. The serum samples from the patients with pancreatic cancer and the chronic pancreatitis were collected at Kobe University Hospital and another institution. The patients with pancreatic cancer and chronic pancreatitis were clinically or pathologically diagnosed by physicians, radiologists, and/or pathologists. The 6th edition of the Union for International Cancer Control (UICC) tumor–node–metastasis (TNM) classification was used to diagnose pancreatic cancer. The serum samples from the healthy volunteers were obtained from Kobe University Hospital and other institutions. No clinical abnormalities were detected in the healthy volunteers during medical check-ups involving physical, blood, urine, imaging, and/or endoscopic examinations. Medical information, that is, height, weight, history of diseases, history of smoking and drinking, histologic findings, and blood biochemical findings, that is, albumin, total cholesterol, triglyceride, aspartate ami-

notransferase, alanine aminotransferase, bilirubin, and amylase, of the subjects were obtained from the clinical records and the reports of medical check-ups on the time of registration (Supplementary Tables S1 and S2).

### Serum samples and preparation

All serum samples were prepared from blood samples collected in the morning using the standard venous blood sampling protocol. The collected blood was centrifuged at  $3,000 \times g$  for 10 minutes at  $4^{\circ}\text{C}$ , and the serum was transferred to a clean tube and stored at  $-80^{\circ}\text{C}$  until use. The extraction of low molecular weight metabolites was done according to the method described in our previous report (24). Oximation and the subsequent derivatization for the GC/MS analysis were also carried out according to the methods described in our previous report (24). The resultant solution was subjected to GC/MS measurement, as described later.

### GC/MS procedure

According to the method described in a previous report (29), the GC/MS analysis was done using a GCMS-QP2010 Ultra (Shimadzu Co.) with a fused silica capillary column (CP-SIL 8 CB low bleed/MS; inner diameter:  $30 \text{ m} \times 0.25 \text{ mm}$ , film thickness:  $0.25 \mu\text{m}$ ; Agilent Co.). The front inlet temperature was  $230^{\circ}\text{C}$ . The flow rate of helium gas through the column was  $39.0 \text{ cm/sec}$ . The column temperature was held at  $80^{\circ}\text{C}$  for 2 minutes and then raised by  $15^{\circ}\text{C}/\text{min}$  to  $330^{\circ}\text{C}$  and held there for 6 minutes. The transfer line and ion-source temperatures were 250 and  $200^{\circ}\text{C}$ , respectively. Twenty scans per second were recorded over the mass range 85 to 500  $m/z$  using the Advanced Scanning Speed Protocol (ASSP; Shimadzu Co.).

Data processing was done according to the methods described in previous reports (29, 30). Briefly, the MS data were exported in netCDF format. The peak detection and alignment were done using the MetAlign software (Wageningen UR). The resultant data were exported in CSV format and then analyzed with in-house analytical software (AIoutput) and an in-house metabolite library for peak identification and semiquantitative analysis. For the semiquantitative analysis, the peak height of each ion was calculated and normalized to the peak height of 2-isopropylmalic acid as an internal standard. Names were assigned to each metabolite peak based on the method of a previous report (30). All data obtained from the serum samples were simultaneously analyzed using the MetAlign software because it was necessary to ensure that the alignment conditions were the same throughout the data analysis. In the GC/MS analysis, multiple peaks were sometimes detected for a particular metabolite because of TMS derivatization, the presence of isomeric forms, etc. In such cases, the peak that most reflected the level of the metabolite was adopted for the semiquantitative evaluation. Furthermore, the metabolites with relative standard deviation (RSD)% values of more than 20% and those that were mainly derived from the experimental background

were not subjected to the semiquantitative evaluation (Supplementary Table S3).

### Statistical analysis

The pancreatic cancer were randomly allocated to the training set or the validation set study of pancreatic cancer. The healthy volunteers were recruited as age- and sex-matched controls for the pancreatic cancer and were randomly allocated to the training set or the validation set. All of the chronic pancreatitis were included in the validation set study to allow us to evaluate whether our model can be used to discriminate cancer from nonmalignant chronic inflammation. In the training set study, the levels of serum metabolites were compared between the pancreatic cancer and healthy volunteers using the Wilcoxon's rank sum test. The metabolites selected via the stepwise method were subjected to multiple logistic regression analysis to construct a diagnostic model. The multicollinearity of the selected variables was assessed by calculating their variance inflation factors (VIF). Receiver operating characteristic (ROC) analysis was used to calculate area under the ROC curve (AUC), sensitivity, and specificity values for the model to evaluate its diagnostic performance. The optimal cut-off value of the model was determined from its ROC curve. In the validation set study, the diagnostic model was evaluated using the AUC, sensitivity, specificity, and accuracy values

observed at the cut-off value obtained in the training set study. In addition, to evaluate the serum metabolic changes in chronic pancreatitis, another set of healthy volunteers was prepared as age- and sex-matched controls for the chronic pancreatitis. The levels of serum metabolites were compared between the chronic pancreatitis and healthy volunteers using the Wilcoxon's rank sum test. In all cases, *P* values of less than 0.05 were considered to indicate a significant difference. These analyses were done using the default conditions of JMP9 (SAS Institute Inc.).

### Results

The clinical characteristics of the patients and healthy volunteers are shown in Tables 1 and 2, and the additional information of the study subjects were exhibited in Supplementary Tables S1 and S2. Using our GC/MS-based metabolomics approach, which mainly targeted water-soluble metabolites, 159 compounds were detected in the subjects' sera (Supplementary Table S3). Among them, 2-isopropylmalic acid was used as an internal standard, and 113 metabolites were excluded because of their instability, contaminants, and so on as described earlier. Therefore, the remaining 45 metabolites were subjected to semiquantitative evaluation. In the training set study, 18 of these metabolites displayed significantly altered levels in the pancreatic cancer compared with the healthy volunteers

**Table 1.** Characteristics of the study subjects

		Training set			Validation set		
		PC	HV	<i>P</i> value	PC	HV+CP	<i>P</i>
<i>N</i>	Total	43	42		42	64 (41+23)	
	Male	23	25		25	41	
	Female	20	17		17	23	
Age, y	Mean	66.2	63.2	0.0862	67.7	62.5	0.0354
	<i>SEM</i>	1.56	1.33		1.55	1.42	
	Range	36–84	39–79		49–93	33–88	
Stage	0	1	—		1	—	
	IA	1	—		2	—	
	IB	2	—		1	—	
	IIA	3	—		2	—	
	IIB	2	—		3	—	
	III	13	—		13	—	
Location	IV	21	—		20	—	
	Head	27	—		27	—	
	Body	9	—		11	—	
	Tail	7	—		4	—	
Histology/cytology	Ductal adenocarcinoma	41	—		40	—	
	Mucinous carcinoma	1	—		1	—	
	Acinar cell carcinoma	0	—		1	—	
	Adenosquamous carcinoma	1	—		0	—	

NOTE: All PC and HV were randomly allocated to the training or validation set. The pancreatic cancer staging was based on the UICC TNM classification. *SEM*, standard error of the mean. *P* values were calculated using the Wilcoxon rank-sum test.

Abbreviations: CP, chronic pancreatitis; HV, healthy volunteers; PC, pancreatic cancer.

**Table 2.** Characteristics of the subjects used for the comparisons between the health volunteers and patients with chronic pancreatitis

		CP	HV	<i>P</i> <sup>a</sup>
<i>N</i>	Total	23	25	
	Male	21	23	
	Female	2	2	
Age, y	Mean	55.7	58.1	0.5218
	SEM	2.44	1.77	
	Range	33–77	39–78	

<sup>a</sup>*P* values were calculated using the Wilcoxon rank-sum test.

(Table 3). The AUC, sensitivity, and specificity of these 18 metabolites were calculated to evaluate their diagnostic performance as individual biomarkers of pancreatic cancer (Supplementary Table S4). 1,5-Anhydro-D-glucitol displayed the highest AUC value as an individual biomarker (0.83499), and its sensitivity and specificity values were 86.0% and 71.4%, respectively. There were no metabolites whose sensitivity and specificity values were both over 80%.

To construct a more effective diagnostic model for pancreatic cancer, we conducted multivariate analysis using the results from the training set study. First, 4 metabolites were selected from the 45 metabolites listed in Table 3 via the stepwise method; that is, xylitol, 1,5-anhydro-D-glucitol, histidine, and inositol. Using these metabolites, which did not display multicollinearity (Supplementary Table S5), a diagnostic model for pancreatic cancer was established using multiple logistic regression analysis (Supplementary Table S6). The diagnostic model produced using these 4 independent variables was as follows:

$$p = 1/[1 + e^{-\{5.48 - 167.57(\text{xylitol}) - 15.21(1,5\text{-anhydro-D-glucitol}) - 282.34(\text{histidine}) + 60.99(\text{inositol})\}}]$$

In the training set study, this model displayed an AUC value of 0.92857 (Fig. 1), and the optimal cut-off value was 0.5137. At this cut-off value, the model's sensitivity and specificity values were 86.0% and 88.1%, respectively. In contrast, the sensitivity and specificity of CA19-9 were 62.8% and 100%, respectively, and those of CEA were 44.2% and 97.6%, respectively (Table 4). The healthy volunteers had their serum CA19-9 and CEA levels measured during health checks, and healthy volunteers whose serum CA19-9 and CEA levels were higher than the relevant clinical cut-off value were excluded from the study. Therefore, the specificity of these conventional tumor markers was markedly high.

The ROC curves for these diagnostic methods obtained in the validation set study are shown in Fig. 2. The AUC of the model was 0.76004. At this cut-off value, its sensitivity and specificity values were 71.4% and 78.1%, respectively.

In the validation set, the AUC, sensitivity, and specificity of CA19-9 were 0.79762, 69.0%, and 85.9%, respectively, whereas those of CEA were 0.66488, 35.7%, and 79.7%, respectively (Fig. 2 and Table 4). Furthermore, in the validation set study, the sensitivities of our model, CA19-9, and CEA in resectable pancreatic cancer (stages 0 to IIB) were 77.8%, 55.6%, and 44.4%, respectively. In the case of chronic pancreatitis, their false-positive rates were 17.4%, 30.4%, and 43.5%, respectively (Table 4). The scatter plots of the predictive values obtained from these diagnostic methods in all 191 cases are shown in Supplementary Fig. S1. The accuracy values of the conventional tumor markers increased with cancer progression, especially in stages III and IV disease. On the contrary, our model even displayed a high accuracy level in the early stage pancreatic cancer cases.

We also conducted comparisons between the chronic pancreatitis and healthy volunteers to assess the influence of nonmalignant chronic inflammation on the serum metabolome. The subject information for this evaluation is shown in Table 2 and Supplementary Table S2. As a result, 16 metabolites were found to display significantly altered levels in chronic pancreatitis, as shown in Supplementary Table S7.

## Discussion

Recently, the possibility of using metabolomics as a diagnostic tool for pancreatic cancer has been investigated in a number of studies (20, 27, 28, 31–34). However, in these reports, principle component analysis (PCA) was used to test the use of metabolomics. PCA cannot be used as a diagnostic method for evaluating individual clinical samples because it requires all samples in a data set to be analyzed simultaneously. Thus, a practical clinical diagnostic tool for pancreatic cancer needs to be developed and evaluated by calculating its sensitivity and specificity using an independent sample set. Therefore, we first constructed and then validated a diagnostic model via the stepwise variable selection method and subsequent multiple logistic regression analysis. All of the metabolites measured in our system are present in a variety of the body's parts including blood in living organisms and are not specific to cancer; hence, they are not suitable as individual biomarkers of specific diseases. Therefore, to develop effective biomarkers several metabolites were combined to produce a multi-biomarker model using multiple logistic regression analysis. Furthermore, to discover reliable biomarkers, well-verified and controlled serum samples were collected from multiple institutions, and then we conducted measurements and data analyses using stringent bioinformatics methods.

Regarding pancreatic disease, there are 3 major problems in the clinical field. The first is the difficulty of detecting pancreatic cancer early in resectable stages. In this study, we confirmed the use of our serum metabolomics-based diagnostic model for detecting resectable pancreatic cancer; that is, the sensitivity of this model for

**Table 3.** Serum metabolites examined in the training set study

		PC		HV		<sup>a</sup> Fold induction	<sup>b</sup> P
		Mean	SEM	Mean	SEM		
1	Pyruvate + Oxalacetic acid	0.041423	0.00307	0.036724	0.00269	1.13	0.1507
2	Lactic acid	0.83882	0.07223	0.73404	0.07704	1.14	0.1830
3	Glycolic acid	0.012996	0.00061	0.013392	0.00057	0.97	0.5297
4	Glycine (2TMS)	0.002317	0.00015	0.002372	0.00008	0.98	0.3227
5	Hydroxybutyrate	0.103307	0.01161	0.106709	0.02022	0.97	0.9474
6	3-Hydroxybutyrate	0.362171	0.06601	0.212465	0.03955	1.70	0.3142
7	Valine (2TMS)	0.699245	0.04142	0.86764	0.03425	0.81	0.0009
8	2-Aminoethanol	0.031801	0.0017	0.035065	0.00126	0.91	0.0283
9	<i>n</i> -Caprylic acid	0.002291	0.00013	0.002987	0.00017	0.77	0.0019
10	Glycerol	0.141685	0.00947	0.164332	0.01202	0.86	0.1918
11	Phosphate	0.22882	0.01806	0.245522	0.01246	0.93	0.0630
12	Threonine (2TMS)	0.032322	0.00145	0.037378	0.00187	0.86	0.0385
13	Proline	0.510797	0.02939	0.578964	0.04047	0.88	0.3447
14	Glycine (3TMS)	0.170047	0.00832	0.168072	0.00517	1.01	0.7284
15	Nonanoic acid (C9)	0.004685	0.00026	0.008784	0.00085	0.53	0.0022
16	Alanine	0.005516	0.00061	0.004501	0.00026	1.23	0.3538
17	Meso-erythritol	0.03754	0.01706	0.015139	0.00205	2.48	0.1049
18	Aspartic acid	0.056614	0.00449	0.05784	0.00265	0.98	0.4847
19	Methionine	0.04934	0.00327	0.057111	0.00276	0.86	0.0056
20	Trans-4-hydroxy-L-proline	0.032874	0.00343	0.028149	0.00257	1.17	0.3963
21	Pyrogallol	0.001075	0.00025	0.001504	0.00034	0.71	0.4112
22	Creatinine	0.008512	0.00088	0.010386	0.00047	0.82	0.0002
23	Glutamic acid	0.20968	0.03276	0.158133	0.01883	1.33	0.3818
24	Phenylalanine	0.194145	0.0106	0.210924	0.01452	0.92	0.4062
25	Arabinose	0.003906	0.00039	0.002238	0.00017	1.75	<.0001
26	Lauric acid	0.010797	0.00079	0.013677	0.00113	0.79	0.1011
27	Ribulose	0.004907	0.00193	0.003541	0.00021	1.39	0.0147
28	Asparagine	0.033292	0.00168	0.03832	0.00145	0.87	0.0162
29	Xylitol	0.003582	0.00056	0.004766	0.0008	0.75	0.8986
30	Arabitol	0.067193	0.0183	0.070207	0.01395	0.96	0.2895
31	Glutamine	0.687485	0.04021	0.795786	0.02593	0.86	0.0072
32	O-phosphoethanolamine	0.000549	0.00002	0.000715	0.00008	0.77	0.0009
33	Glycyl-Glycine_1	0.00081	0.00009	0.00107	0.00006	0.76	0.0063
34	Citric acid + isocitric acid	0.188029	0.01197	0.170173	0.00793	1.10	0.4366
35	Ornithine	0.03155	0.00205	0.028359	0.00121	1.11	0.5828
36	1,5-Anhydro-D-glucitol	0.149795	0.01349	0.274426	0.01443	0.55	<.0001
37	Glucose_1	1.331726	0.09164	1.266738	0.10305	1.05	0.3676
38	Mannose_2	1.463641	0.07925	1.501308	0.06722	0.97	0.7684
39	Lysine (4TMS)	0.124283	0.00587	0.160292	0.00943	0.78	0.0005
40	Histidine	0.019962	0.00095	0.028603	0.00191	0.70	<.0001
41	Glucuronate_1	0.016676	0.00816	0.003648	0.00025	4.57	0.1011
42	Tyrosine	0.361596	0.02023	0.431645	0.02783	0.84	0.0067
43	Inositol	0.090173	0.00683	0.079817	0.00334	1.13	0.5413
44	Uric acid	0.235601	0.01239	0.278596	0.01197	0.85	0.0170
45	Cysteine + cystine	0.044856	0.00286	0.043795	0.0032	1.02	0.6508

NOTE: The serum levels of each metabolite were normalized to the peak intensity of the internal standard; that is, 2-isopropylmalic acid.

<sup>a</sup>Fold induction values were calculated from the PC to HV ratio.

<sup>b</sup>P values were calculated using the Wilcoxon rank-sum test.

resectable stage disease was about 80% (Table 4), which should be acceptable for clinical use because most patients with resectable cancer have no symptoms, and blood

examinations are useful tools for initial screening examinations. The second is the difficulty of discriminating cancer from chronic inflammation. Chronic pancreatitis is

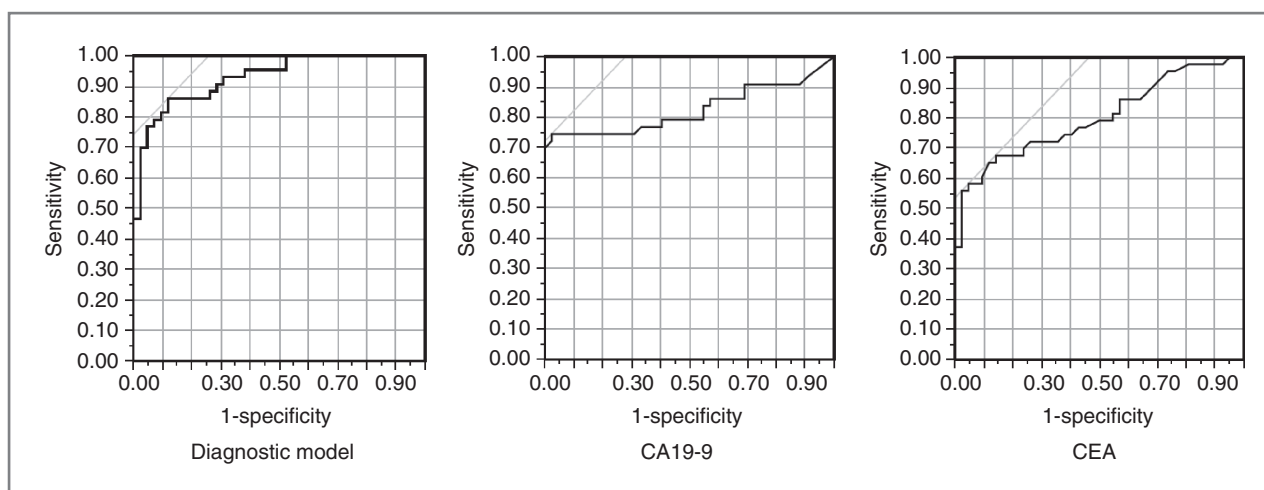


Figure 1. The ROC curve of the diagnostic model and the tumor markers in the training set study. The AUC of the ROC curve, cut-off value, sensitivity, and specificity of each diagnostic method are summarized in Table 4.

one of the major risk factors for pancreatic cancer. Many gastroenterologists follow-up chronic pancreatitis with scheduled CT, magnetic resonance imaging (MRI), and endoscopic ultrasound scans (EUS); tumor marker tests; and endoscopic retrograde cholangiopancreatography (ERCP), but the initial malignant changes are frequently overlooked, and pancreatic tumors can rapidly become unresectable. In addition, unnecessary resections for benign inflammatory lesions are sometimes done because of false-positive results from CA19-9 and/or imaging examinations. In this study, our metabolomics-based diagnostic model even showed the lower false-positive rate than CA19-9 in chronic pancreatitis (Table 4), indicating that using it could reduce the incidences of missed malignant changes in chronic pancreatitis and unnecessary major surgery for benign inflammatory lesions. The third clinical problem is the risk of complications during

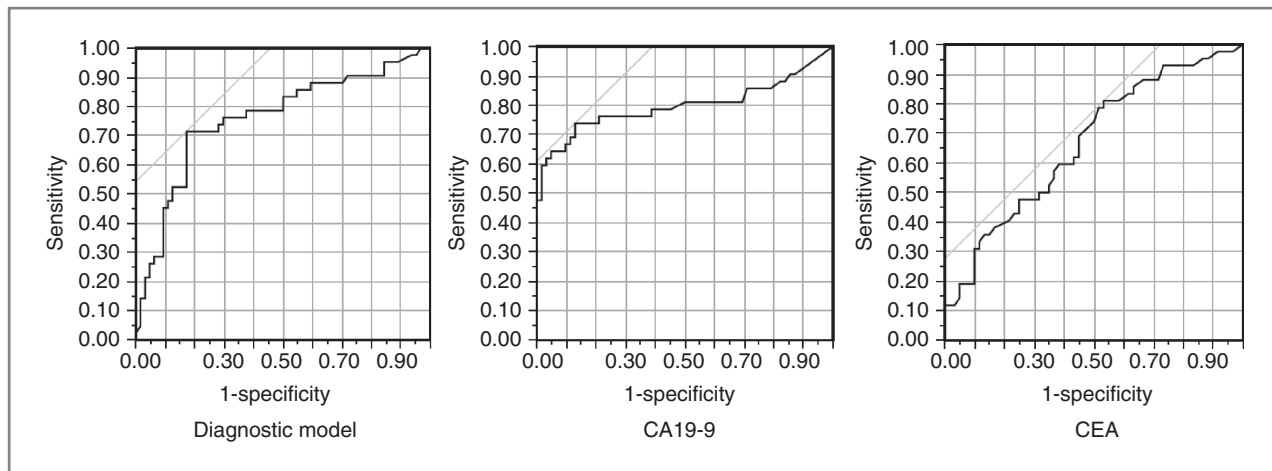
pancreatic examinations. In our metabolomic analysis, serum samples were used as the experimental materials. The samples were taken according to a standard venous blood sampling method. The risk of complications in this procedure was quite low, and its throughput was relatively high. Therefore, our metabolomic approach is acceptable as a screening method for large populations.

Jiang and colleagues suggested tumor-specific growth factor (TSGF) as a candidate serum biomarker for pancreatic cancer and found that it displayed 91.6% sensitivity and 83% specificity (35). However, its sensitivity for stages I and II pancreatic cancer was decreased to 60.0% and 75.0%, respectively. CA242 was showed to be a biomarker candidate for pancreatic cancer by the same group and Kawa and colleagues, but it suffered from the same problem; that is, it displayed low sensitivity for resectable disease (11). Koopmann and colleagues

Table 4. Diagnostic performance of the constructed model and tumor markers

		Diagnostic model	CA19-9	CEA
Training set study	AUC	0.92857	0.82420	0.79956
	(95% confidence interval)	(0.85623–0.96596)	(0.70299–0.90278)	(0.68711–0.87872)
	Cut-off value	0.5137	37 U/mL	5 ng/mL
	Sensitivity (%)	86.0	62.8	44.2
Validation set study	Specificity (%)	88.1	100	97.6
	AUC	0.76004	0.79762	0.66488
	(95% confidence interval)	(0.65673–0.85564)	(0.67501–0.88282)	(0.55120–0.76219)
	Sensitivity (%)	71.4	69.0	35.7
	Specificity (%)	78.1	85.9	79.7
	Sensitivity in PC of stage 0–IIB (%)	77.8	55.6	44.4
	False-positive rate in CP (%)	17.4	30.4	43.5

NOTE: The ROC curves for these evaluations are shown in Figs. 1 and 2. Regarding CA19-9 and CEA, the clinical cut-off values used at our institutions were used.



**Figure 2.** The ROC curves of the diagnostic model and tumor markers in the validation set study. The AUC of the ROC curve, cut-off value, sensitivity, specificity, and false-positive rate of each diagnostic method are summarized in Table 4.

showed that serum MIC-1 possessed 90% sensitivity and 94% specificity in patient cohorts involving patients with pancreatic cancer and healthy controls (14). However, its specificity decreased markedly to 44% in cohorts containing patients with chronic pancreatitis. Carcinoembryonic antigen-related cell adhesion molecule 1 (CEACAM1; ref. 36), REG4 (37), and M2-pyruvate kinase (M2-PK; ref. 12) also suffer from low specificity in cohorts including patients with pancreatitis. Although there are other candidate biomarkers such as platelet factor 4 (PF4; ref. 13), PNA binding glycoprotein (15), cell adhesion molecule 17.1 (CAM 17.1; ref. 38), and serum immune signatures identified by affinity proteomics (39), their accuracy for detecting resectable stage cancer and/or chronic pancreatitis have not been evaluated. Our metabolomics-based diagnostic model was more sensitive at detecting resectable stage cancer (77.8%) with the lower false-positive rate in chronic pancreatitis (17.4%) than conventional markers (Table 4). In addition, the serum metabolites selected as variables in our diagnostic model did not display the significant inter-day and intra-day variances in our previous report (20). This superiority will provide a substantial benefit in clinical use.

The pathogenesis of pancreatic disease causes significant decreases in the serum levels of amino acids and fatty acids in pancreatic cancer (valine, *n*-caprylic acid, threonine, nonanoic acid, methionine, asparagine, glutamine, lysine, histidine, and tyrosine) and chronic pancreatitis (valine, proline, methionine, lauric acid, asparagine, glutamine, and tyrosine). Other research groups have showed similar decreases in plasma samples from pancreatic cancer and chronic pancreatitis using high-performance liquid chromatography (HPLC) analysis (40). It is well known that the uptake and catabolism of amino acids and fatty acids are enhanced to support rapid cell proliferation in cancer tissues (41), and these decreases may be explained as a result of the enhanced usage in tumors. Patients with pancreatic disease are also troubled by

malnutrition because of pancreatic endocrine and exocrine insufficiency. Therefore, there is a possibility that the decreases in their serum metabolite levels also reflect malnutrition. In this study, a reduction in the serum level of 1,5-anhydro-D-glucitol was observed in the pancreatic cancer (Table 3). 1,5-Anhydro-D-glucitol is a serum biomarker of short-term glycemic control (41), and the decreased serum level of 1,5-anhydro-D-glucitol means the presence of hyperglycemia and glycosuria of the past few days. These results suggest that glucose tolerance was impaired in these patients because of pancreatic insufficiency. On the basis of our findings, the serum metabolome seems to reflect not only metabolic changes in focal lesions but also the systemic responses to these pancreatic diseases. However, it is necessary to elucidate the reasons for these alterations in the serum metabolome.

Regarding our study design, there are some limitations. First, other cancers and chronic diseases as autoimmune pancreatitis or tumor forming pancreatitis should be included into the study groups to evaluate the specificity more strictly. The sample size, especially in resectable pancreatic cancer, was not also enough to conduct stringent subgroup analyses. Second, there is a possibility that potential biases and insufficient study size of this study might raise the false significant results. To overcome these problems, development of rigorous standard method and quality control procedures for the biomarker study are needed strongly, and for example Pepe and colleagues proposed that the prospective-specimen-collection and retrospective-blinded-evaluation (PROBE) design is considered as the standard method to discover and evaluate the biomarkers for screening or diagnosis (42).

In conclusion, we developed a serum metabolomics-based diagnostic model for pancreatic cancer using multiple logistic regression analysis. Our model possessed higher accuracy than conventional tumor markers especially at detecting the patients with resectable pancreatic

cancer in cohort including the patients with chronic pancreatitis. Although the increased number of well-defined samples including various diseases and the large-scale multicenter study are still necessary for strict evaluation of specificity and for construction of the robust reliable model, this novel diagnostic approach is expected to improve the prognosis of patients with pancreatic cancer by detecting their cancer early, when it is still in a resectable and curable state.

#### Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

#### Authors' Contributions

**Conception and design:** T. Kobayashi, S. Nishiumi, Y. Izumi, H. Minami, T. Takenawa, T. Azuma, M. Yoshida

**Development of methodology:** T. Kobayashi, S. Nishiumi, T. Yoshie  
**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** T. Kobayashi, S. Nishiumi, A. Ikeda, A. Sakai, A. Matsubara, H. Tsumura, M. Tsuda, H. Nishisaki, N. Hayashi, S. Kawano, H. Minami

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** T. Kobayashi, S. Nishiumi, H. Minami  
**Writing, review, and/or revision of the manuscript:** T. Kobayashi, S. Nishiumi, Y. Fujiwara, H. Minami, M. Yoshida

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** A. Sakai, T. Takenawa  
**Study supervision:** T. Azuma, M. Yoshida

#### References

- Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin* 2012;62:10–29.
- Pannala R, Basu A, Petersen GM, Chari ST. New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer. *Lancet Oncol* 2009;10:88–95.
- Wray CJ, Ahmad SA, Matthews JB, Lowy AM. Surgery for pancreatic cancer: recent controversies and current practice. *Gastroenterology* 2005;128:1626–41.
- Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: an evidence based appraisal. *J Gastrointest Oncol* 2012;3:105–19.
- Rehani MM, Berry M. Radiation doses in computed tomography. The increasing doses of radiation need to be controlled. *BMJ (Clinical research ed)* 2000;320:593–4.
- Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *N Engl J Med* 2007;357:2277–84.
- Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 2012;380:499–505.
- Gonda TA, Saif MW. Early detection and screening of pancreatic cancer. Highlights from the "2011 ASCO Gastrointestinal Cancers Symposium". San Francisco, CA, USA. January 20–22, 2011. *JOP* 2011;12:83–5.
- Eguia V, Gonda TA, Saif MW. Early detection of pancreatic cancer. *JOP* 2012;13:131–4.
- Fry LC, Monkemuller K, Malfertheiner P. Molecular markers of pancreatic cancer: development and clinical relevance. *Langenbecks Arch Surg* 2008;393:883–90.
- Kawa S, Tokoo M, Hasebe O, Hayashi K, Imai H, Oguchi H, et al. Comparative study of CA242 and CA19-9 for the diagnosis of pancreatic cancer. *Br J Cancer* 1994;70:481–6.
- Ventrucci M, Cipolla A, Racchini C, Casadei R, Simoni P, Gullo L. Tumor M2-pyruvate kinase, a new metabolic marker for pancreatic cancer. *Dig Dis Sci* 2004;49:1149–55.
- Fiedler GM, Leichter AB, Kase J, Baumann S, Ceglarek U, Felix K, et al. Serum peptidome profiling revealed platelet factor 4 as a potential discriminating Peptide associated with pancreatic cancer. *Clin Cancer Res* 2009;15:3812–9.
- Koopmann J, Rosenzweig CN, Zhang Z, Canto MI, Brown DA, Hunter M, et al. Serum markers in patients with resectable pancreatic adenocarcinoma: macrophage inhibitory cytokine 1 versus CA19-9. *Clin Cancer Res* 2006;12:442–6.
- Ching CK, Rhodes JM. Enzyme-linked PNA lectin binding assay compared with CA19-9 and CEA radioimmunoassay as a diagnostic blood test for pancreatic cancer. *Br J Cancer* 1989;59:949–53.
- Uehara H, Nakaizumi A, Tatsuta M, Baba M, Takenaka A, Uedo N, et al. Diagnosis of pancreatic cancer by detecting telomerase activity in pancreatic juice: comparison with K-ras mutations. *Am J Gastroenterol* 1999;94:2513–8.
- Yokoyama M, Ochi K, Ichimura M, Mizushima T, Shinji T, Koide N, et al. Matrix metalloproteinase-2 in pancreatic juice for diagnosis of pancreatic cancer. *Pancreas* 2002;24:344–7.
- Hibi T, Mori T, Fukuma M, Yamazaki K, Hashiguchi A, Yamada T, et al. Synuclein-gamma is closely involved in perineural invasion and distant metastasis in mouse models and is a novel prognostic factor in pancreatic cancer. *Clin Cancer Res* 2009;15:2864–71.
- Yoshida M, Hatano N, Nishiumi S, Irino Y, Izumi Y, Takenawa T, et al. Diagnosis of gastroenterological diseases by metabolome analysis using gas chromatography-mass spectrometry. *J Gastroenterol* 2012;47:9–20.
- Nishiumi S, Shinohara M, Ikeda A, Yoshie T, Hatano N, Kakuyama S, et al. Serum metabolomics as a novel diagnostic approach for pancreatic cancer. *Metabolomics* 2010;6:518–28.
- Ooi M, Nishiumi S, Yoshie T, Shiomi Y, Kohashi M, Fukunaga K, et al. GC/MS-based profiling of amino acids and TCA cycle-related molecules in ulcerative colitis. *Inflamm Res* 2011;60:831–40.
- Ikeda A, Nishiumi S, Shinohara M, Yoshie T, Hatano N, Okuno T, et al. Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer. *Biomed Chromatogr* 2012;26:548–58.
- Shiomi Y, Nishiumi S, Ooi M, Hatano N, Shinohara M, Yoshie T, et al. GCMS-based metabolomic study in mice with colitis induced by dextran sulfate sodium. *Inflamm Bowel Dis* 2011;17:2261–74.

#### Acknowledgments

We greatly appreciate the assistance of the staff of the Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine (Hyogo, Japan); the staff of The Integrated Center for Mass Spectrometry, Kobe University Graduate School of Medicine (Hyogo, Japan); T. Bamba, PhD, and E. Fukusaki, PhD, of the Department of Biotechnology, Graduate School of Engineering, Osaka University (Osaka, Japan); and Shimadzu Co. (Kyoto, Japan). We are also grateful to H. Inokuchi, MD, and colleagues of Hyogo Cancer Center (Hyogo, Japan), Y. Tamori, MD, and colleagues of Aijinkai Total Health-care Center (Osaka, Japan), and R. Yamada, MD, and colleagues of Shinkokai Shinko Hospital Health Examination Center (Hyogo, Japan) for collecting sera from the patients and healthy volunteers.

#### Grant Support

This study was supported by a grant for the Global COE Program, Global Center of Excellence for Education and Research on Signal Transduction Medicine in the Coming Generation from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan (T. Kobayashi, T. Yoshie, T. Azuma, and M. Yoshida); a Grant-in-Aid from the Japan Society for the Promotion of Science (JSPS) Fellows (T. Kobayashi); a grant for Research on Applying Health Technology from the Ministry of Health, Labour, and Welfare of Japan (M. Yoshida).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received September 7, 2012; revised December 6, 2012; accepted December 31, 2012; published OnlineFirst March 29, 2013.



24. Nishiumi S, Kobayashi T, Ikeda A, Yoshie T, Kibi M, Izumi Y, et al. A novel serum metabolomics-based diagnostic approach for colorectal cancer. *PLoS One* 2012;7:e40459.
25. Spratlin JL, Serkova NJ, Eckhardt SG. Clinical applications of metabolomics in oncology: a review. *Clin Cancer Res* 2009;15:431–40.
26. Hirayama A, Kami K, Sugimoto M, Sugawara M, Toki N, Onozuka H, et al. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res* 2009;69:4918–25.
27. Fang F, He X, Deng H, Chen Q, Lu J, Spraul M, et al. Discrimination of metabolic profiles of pancreatic cancer from chronic pancreatitis by high-resolution magic angle spinning <sup>1</sup>H nuclear magnetic resonance and principal components analysis. *Cancer Sci* 2007;98:1678–82.
28. Bathe OF, Shaykhtudinov R, Kopciuk K, Weljie AM, McKay A, Sutherland FR, et al. Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiol Biomarkers Prev* 2011;20:140–7.
29. Tsugawa H, Bamba T, Shinohara M, Nishiumi S, Yoshida M, Fukusaki E. Practical non-targeted gas chromatography/mass spectrometry-based metabolomics platform for metabolic phenotype analysis. *J Biosci Bioeng* 2011;112:292–8.
30. Tsugawa H, Tsujimoto Y, Arita M, Bamba T, Fukusaki E. GC/MS based metabolomics: development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC Bioinformatics* 2011;12:131.
31. Urayama S, Zou W, Brooks K, Tolstikov V. Comprehensive mass spectrometry based metabolic profiling of blood plasma reveals potent discriminatory classifiers of pancreatic cancer. *Rapid Commun Mass Spectrom* 2010;24:613–20.
32. Tesiram YA, Lerner M, Stewart C, Njoku C, Brackett DJ. Utility of nuclear magnetic resonance spectroscopy for pancreatic cancer studies. *Pancreas* 2012;41:474–80.
33. OuYang D, Xu J, Huang H, Chen Z. Metabolomic profiling of serum from human pancreatic cancer patients using <sup>1</sup>H NMR spectroscopy and principal component analysis. *Appl Biochem Biotechnol* 2011;165:148–54.
34. Zhang H, Wang Y, Gu X, Zhou J, Yan C. Metabolomic profiling of human plasma in pancreatic cancer using pressurized capillary electrophoresis. *Electrophoresis* 2011;32:340–7.
35. Jiang JT, Wu CP, Deng HF, Lu MY, Wu J, Zhang HY, et al. Serum level of TSGF, CA242 and CA19-9 in pancreatic cancer. *World J Gastroenterol* 2004;10:1675–7.
36. Simeone DM, Ji B, Banerjee M, Arumugam T, Li D, Anderson MA, et al. CEACAM1, a novel serum biomarker for pancreatic cancer. *Pancreas* 2007;34:436–43.
37. Takayama R, Nakagawa H, Sawaki A, Mizuno N, Kawai H, Tajika M, et al. Serum tumor antigen REG4 as a diagnostic biomarker in pancreatic ductal adenocarcinoma. *J Gastroenterol* 2010;45:52–9.
38. Yiannakou JY, Newland P, Calder F, Kingsnorth AN, Rhodes JM. Prospective study of CAM 17.1/WGA mucin assay for serological diagnosis of pancreatic cancer. *Lancet* 1997;349:389–92.
39. Wingren C, Sandstrom A, Segersvard R, Carlsson A, Andersson R, Lohr M, et al. Identification of serum biomarker signatures associated with pancreatic cancer. *Cancer Res* 2012;72:2481–90.
40. Schrader H, Menge BA, Belyaev O, Uhl W, Schmidt WE, Meier JJ. Amino acid malnutrition in patients with chronic pancreatitis and pancreatic carcinoma. *Pancreas* 2009;38:416–21.
41. Yamanouchi T, Akanuma Y. Serum 1,5-anhydroglucitol (1,5 AG): new clinical marker for glycemic control. *Diabetes Res Clin Pract* 1994;24 Suppl:S261–8.
42. Pepe MS, Feng Z, Janes H, Bossuyt PM, Potter JD. Pivotal evaluation of the accuracy of a biomarker used for classification or prediction: standards for study design. *J Natl Cancer Inst* 2008;100:1432–8.