Kobe University Repository : Kernel

PDF issue: 2025-12-04

Identification of highly sensitive biomarkers that can aid the early detection of pancreatic cancer using GC/MS/MS-based targeted metabolomics

Hirata, Yuichi ; Kobayashi, Takashi ; Nishiumi, Shin ; Yamanaka, Kodai ; Nakagawa, Takashi ; Fujigaki, Seiji ; Iemoto, Takao ; Kobayashi,…

(Citation)
Clinica Chimica Acta,468:98-104

(Issue Date)
2017-05

(Resource Type)
journal article

(Version)
Accepted Manuscript

(Rights)
©2017 Elsevier B.V.
This manuscript version is made available under the CC-BY-NC-ND 4.0 license
http://creativecommons.org/licenses/by-nc-nd/4.0/

(URL)
https://hdl.handle.net/20.500.14094/90004501

*Abstract

Abstract

Background: To improve prognosis of pancreatic cancer (PC) patients, the discovery of more reliable biomarkers for the early detection is desired.

Methods: Blood samples were collected by two independent groups. The 1st set was included 55 early PC and 58 healthy volunteers (HV), and the 2nd set was included 16 PC and 16 HV. The 16 targeted metabolites were quantitatively analyzed by gas chromatography/tandem mass spectrometry together with their corresponding stable isotopes. In the 1st set, the levels of these metabolites were evaluated, and diagnostic models were constructed via multivariate logistic regression analysis, leading to validation using the 2nd set.

Results: In the 1st set, model X consisting of 4 candidates based on our previous report possessed higher sensitivity (74.1%) than carbohydrate antigen 19-9 (CA19-9). Model Y, consisting of 2 metabolites newly selected from 16 metabolites via stepwise method possessed higher sensitivity (70.4%) than CA19-9. Furthermore, combining model Y with CA19-9 increased its sensitivity (90.7%) and specificity (89.5%). In the 2nd set, combining model Y with CA19-9 displayed high sensitivity (81.3%) and specificity (93.8%). In particular, it displayed very high sensitivity (100%) for resectable PC.

Conclusions: Quantitative analysis confirmed that metabolomics-based diagnostic

methods are useful for detecting PC early.

Keywords: biomarker; GC/MS/MS; pancreatic cancer; early detection; metabolomics.

***Highlights**

・GC/MS/MS revealed significant change of serum metabolites in early pancreatic cancer.

・Histidine and xylitol were selected for diagnostic model via the stepwise method.

・Diagnostic models by metabolomics possessed higher sensitivity than CA19-9.

・Combination model of metabolites and CA19-9 increased its diagnostic accuracy.

*Manuscript

Identification of highly sensitive biomarkers that can aid the early detection of pancreatic cancer using GC/MS/MS-based targeted metabolomics

Yuichi Hirata[a], Takashi Kobayashi[a, *], Shin Nishiumi[a, *], Kodai Yamanaka[a], Takashi Nakagawa[a], Seiji Fujigaki[a], Takao Iemoto[a], Makoto Kobayashi[b], Takuji Okusaka[c], Shoji Nakamori[d], Masashi Shimahara[e], Takaaki Ueno[e], Akihiko Tsuchida[f], Naohiro Sata[g], Tatsuya Ioka[h], Yohichi Yasunami[i], Tomoo Kosuge[j], Takashi Kaneda[k], Takao Kato[l], Kazuhiro Yagihara[m], Shigeyuki Fujita[n], Tesshi Yamada[b], Kazufumi Honda[b, o], Takeshi Azuma[a], and Masaru Yoshida[a, p, q, **]

Authors' Affiliations: [a]Division of Gastroenterology, Department of Internal Medicine, Kobe University Graduate School of Medicine, Hyogo, Japan; [b]Division of Chemotherapy and Clinical Research, National Cancer Center Research Institute, Tokyo, Japan; [c]Department of Hepatobiliary and Pancreatic Oncology, National Cancer Center Hospital, Tokyo, Japan; [d]Departments of Hepato-Biliary-Pancreatic Surgery, Osaka National Hospital, National Hospital Organization, Osaka, Japan; [e]Department of Oral Surgery, Osaka Medical College, Osaka, Japan; [f]Department of Gastrointestinal and Pediatric Surgery, Tokyo Medical University, Tokyo, Japan; [g]Department of Surgery,

19 Jichi Medical University, Tochigi, Japan; [h]Department of GI Cancer Screening and
20 Surveillance, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka,
21 Japan; [i]Islet Institute, Fukuoka University, Fukuoka, Japan; [j]Department of
22 Gastrointestinal Surgery, JR Tokyo General Hospital, Tokyo, Japan; [k]Department of
23 Radiology, Nihon University School of Dentistry at Matsudo, Chiba, Japan;
24 [l]Department of Oral Implant, Nihon University School of Dentistry at Matsudo, Chiba,
25 Japan; [m]Department of Oral Surgery, Saitama Cancer Center, Saitama, Japan;
26 [n]Department of Oral and Maxillofacial Surgery, Wakayama Medical University,
27 Wakayama, Japan; [o]Japan Agency for Medical Research and Development (AMED)
28 CREST, Tokyo, Japan; [p]Metabolomics Research, Department of Internal Related, Kobe
29 University Graduate School of Medicine, Hyogo, Japan; [q]AMED-CREST, AMED, Hyogo,
30 Japan
31 * Authors contributed equally
32
33 **Corresponding Author: Masaru Yoshida
34 Division of Gastroenterology, Department of Internal Medicine, Kobe University
35 Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chu-o-ku, Kobe, Hyogo 650-0017,
36 Japan; Division of Metabolomics Research, Department of Internal Related, Kobe

37 University Graduate School of Medicine, 7-5-1 Kusunoki-cho, Chu-o-ku, Kobe, Hyogo
38 650-0017, Japan; AMED-CREST, AMED, 7-5-1 Kusunoki-cho, Chu-o-ku, Kobe, Hyogo
39 650-0017, Japan; Phone: +81-78-382-6305; Fax: +81-78-382-6309; E-mail:
40 myoshida@med.kobe-u.ac.jp

1. Introduction

Pancreatic cancer (PC) is one of the leading causes of cancer-related death. PC is considered to be a lethal solid tumor, and its 5-year survival rate is <8% [1] because of the difficulty of its early detection [2]. The only curative treatment for PC is surgical resection, but >80% of PC patients are diagnosed with unresectable advanced disease [3-5]. The clinical symptoms of PC patients are usually unremarkable until the cancer has progressed to an advanced stage. Furthermore, there are no effective screening methods for the early detection of PC. Carbohydrate antigen 19-9 (CA19-9), which is usually used as a tumor marker of PC, is unsuitable for aiding the early detection of the disease because of its low sensitivity for the resectable early stages. In addition, the levels of CA19-9 are also increased in other gastrointestinal malignancies and benign pancreaticobiliary diseases, such as pancreatitis and cholangitis [6]. Imaging examinations, such as computed tomography, magnetic resonance imaging, and endoscopic ultrasonography, are not suitable screening methods in terms of cost-effectiveness. Therefore, novel screening methods for the early detection of PC are required and many researchers have studied for the development of candidate biomarkers of PC [7-10].

59 Our study group has shown the usefulness of metabolomics for diagnosing various
60 diseases and gaining an enhanced understanding of their underlying mechanisms
61 [11-19]. Regarding PC, using gas chromatography/mass spectrometry (GC/MS) we have
62 identified candidate metabolites that have potential as biomarkers that could aid the
63 early detection of PC [18]. We reported that there are some serum metabolites whose
64 levels differ significantly between PC patients and healthy volunteers (HV), and a
65 diagnostic model established from 4 selected metabolites was shown to be useful for
66 distinguishing PC patients from HV. These results indicate that the metabolome is
67 closer to the phenotype than other stages in the omics cascade and reflects the
68 conditions found in particular diseases. Metabolomics involves the comprehensive
69 analysis of low molecular weight metabolites. Metabolic changes result in alterations in
70 various metabolites, and elucidating the metabolic changes that occur in a particular
71 disease will increase our understanding of its underlying mechanisms. Therefore, the
72 use of metabolomics in the medical field has recently developed rapidly [20].
73
74 However, our previous study had several limitations [18]. First, our previous model
75 targeted all stages of PC, and the number of patients with early stage PC (stage 0-IIB)
76 was small. Second, the blood samples in the validation set were collected at the same

77 hospitals as those in the training set and were not completely independent. Third, the
78 qualitative evaluation was insufficient because the metabolites were analyzed
79 comprehensively. There are two main methodological approaches that can be used for
80 metabolomics, non-targeted metabolomics and targeted metabolomics. Non-targeted
81 metabolomics aims to analyze the global metabolite profile of each sample. In contrast,
82 targeted metabolomics focuses on analyzing a predefined set of molecules and usually
83 provides quantitative information. During the analysis of biospecimens, such as blood
84 samples, which contain various types of metabolites that exhibit a wide range of
85 concentrations, the analytical reliability of GC/MS performed with a single quadrupole
86 mass spectrometer is often adversely affected by interference from contaminating
87 substances in the sample. On the other hand, performing GC/MS/MS with a triple
88 quadrupole mass spectrometer can eliminate interference and enables the selection of
89 more specific ions. Therefore, GC/MS/MS could result in improvements in both
90 qualitative and quantitative performance and higher sensitivity than GC/MS [21, 22].
91 Fourth, the quantitative evaluation was also insufficient because we conducted a
92 semi-quantitative analysis based on relative concentrations compared with a single
93 internal standard. Analyses using stable isotopes that correspond to the targeted
94 metabolites are able to produce more accurate quantitative data than the use of a single

internal standard. Due to these limitations, the reproducibility of our previous results has not been fully confirmed.

This study is a follow-up study of our previous study [18]. In this study, in order to aid the early detection of PC, we focused on early stage PC. We strictly validated our findings by preparing two independent sets from several institutions, focusing on 16 candidate metabolites whose levels differed significantly between PC patients and HV in our previous study, and by performing quantitative analyses using GC/MS/MS and corresponding stable isotopes.

2. Materials and Methods

2.1. Chemicals and isotopes

We focused on 16 metabolites that have previously been reported to be biomarker candidates; i.e., valine, 2-aminoethanol, nonanoic acid, threonine, methionine, creatinine, arabinose, asparagine, xylitol, glutamine, 1,5-anhydro-D-glucitol (1,5-AG), lysine, histidine, tyrosine, inositol, and uric acid. Stable isotope products for each metabolite; i.e., L-valine [D8, 98%], 2-aminoethanol [D4, 98%], nonanoic acid [U-13C9, 98%], L-threonine [13C4, 97-99%], L-methionine [13C5, 99%], creatinine [N-methyl-D3,

113 98%], D-arabinose [U-13C5, 99%], L-asparagine:H2O [13C4, 99%], D-xylitol [U-13C5,

114 99%], L-glutamine [13C5, 99%], 1,5-AG [U-13C6, 98%+], L-lysine:2HCl [13C6, 99%],

115 L-histidine:HCl:H2O [13C6, 97-99%], L-tyrosine [ring-13C6, 99%], myo-inositol

116 [1,2,3,4,5,6-D6, 98%], and uric acid [1,3-15N2, 98%], were obtained from Cambridge

117 Isotope Laboratories, Inc. (Tewksbury, MA). Methanol was acquired from Kanto

118 Chemical Co., Inc. (Tokyo, Japan). 2-isopropylmalic acid and methoxyamine

119 hydrochloride were purchased from Sigma-Aldrich (St. Louis, MO). Pyridine was

120 obtained from Wako Pure Chemical Industries, Ltd (Osaka, Japan).

121 *N*-methyl-*N*-(trimethylsilyl)trifluoroacetamide (MSTFA) was acquired from GL Sciences

122 Inc. (Tokyo, Japan). Human standard plasma (as a quality control) was obtained from

123 Kohjin Bio Co., Ltd. (Osaka, Japan).

124

125 2.2. Subjects

126 2.2.1. The 1st set

127 Serum samples were obtained from PC patients and age- and sex-matched HV at the

128 National Cancer Center Hospital (Tokyo, Japan) and its related hospitals between

129 October 2006 and January 2015. All of the PC patients had stage I or II disease. The

130 selection of the matched controls was carried out by a third party that was not directly

131 involved in this study. The serum samples were separated from blood using the
132 standard clinical method in both groups. After the sampling procedure, the blood was
133 left for 30 minutes at room temperature and then centrifuged at 3,000 x g for 15
134 minutes at 4°C within 8 hours of the sampling procedure. The separated serum was
135 transferred to a clean Eppendorf tube and stored at -80°C until the analysis.
136
137 2.2.2. The 2nd set
138 Blood samples were prospectively collected in the morning after >8 hours' fasting at
139 Kobe University Hospital and related hospitals between January 2015 and January
140 2016. All plasma samples were collected using the standard venous blood sampling
141 protocol and EDTA-2Na tubes. After the sampling procedure, the blood was rapidly
142 cooled to a temperature of 0°C to 5°C and then centrifuged at 3,000 x g for 15 minutes at
143 4°C within 8 hours of the sampling procedure. The separated plasma was transferred to
144 a clean Eppendorf tube and stored at -80°C until the analysis.
145
146 This study was approved by the ethics committees of the National Cancer Center (Tokyo,
147 Japan) and Kobe University Graduate School of Medicine (Hyogo, Japan). Written
148 informed consent was obtained from all subjects of both groups. The 7th edition of the

149 Union for International Cancer Control (UICC) tumor-node-metastasis (TNM)
150 classification was used to diagnose PC. Examinees with no history of cancer and no
151 symptoms were adopted as HV in the 1st set study. In the 2nd set study, HV were defined
152 as asymptomatic individuals who had been confirmed to be free from malignancies of
153 the pancreas or other organs during medical check-ups that included physical
154 examinations, blood tests, chest X-rays, and abdominal ultrasonography. We also
155 confirmed that the HV in this study did not exhibit any marked pancreatic clinical
156 symptoms or findings at 6 months after the sampling procedure.
157
158 2.3. Preparation
159 The extraction and derivatization of low molecular weight metabolites were performed
160 as follows: Fifty μL of each sample were dispensed into a 1.5-mL Eppendorf tube. The
161 samples were then extracted with 250 μL methanol containing the stable isotopes and
162 shaken in a vortex. Next, 10 μL of 2-isopropylmalic acid (0.5 mg/mL) were added as an
163 internal standard. The mixture was incubated at 1,200 rpm for 30 minutes at 37°C.
164 After the mixture had been centrifuged at 19,300×g for 5 minutes at 4°C, 225 μL of the
165 supernatant were transferred to a new Eppendorf tube capped with a pierced cap. After
166 being centrifuged for 40 minutes in a vacuum concentrator (Thermo SpeedVac), the

167 mixture was freeze-dried overnight. For the derivatization, 80 μL of methoxyamine
168 hydrochloride in pyridine (20 mg/mL) were added as the first derivatizing agent. The
169 mixture was then incubated at 1,200 rpm for 90 minutes at 30°C. The second
170 derivatizing agent, 40 μL of MSTFA, was added, and then the mixture was incubated at
171 1,200 rpm for 30 minutes at 37°C. After the mixture had been centrifuged at 19,300 x g
172 for 5 minutes at room temperature, the supernatant was transferred to a vial for
173 analysis by GC/MS/MS.
174
175 2.4. GC/MS/MS procedure
176 The GC/MS/MS analysis was carried out on a GCMS-TQ8040 GC/MS/MS system
177 (Shimadzu Co.). Each sample was injected with a split ratio of 1:10, and then the
178 separation was performed on a fused silica capillary column (BPX5; inner diameter: 30
179 m x 0.25 mm, film thickness: 0.25 μm; Agilent Co.). The front inlet temperature was
180 250°C. Helium gas was used as the GC carrier gas, and argon gas was used as the
181 MS/MS collision gas. The flow rate of helium gas through the column was 39.0
182 cm/second. The column temperature was held at 80°C for 2 minutes and then raised by
183 15°C/min to 330°C, before being held there for 3 minutes. The transfer line and
184 ice-source temperatures were 280°C and 200°C, respectively.

185

186 Multiple reaction monitoring optimization was performed as follows. Each stable
187 isotope was analyzed by GC/MS/MS and the optimal transition, which consisted of the
188 precursor ion, collision energy, and product ion, was selected. The transitions for all
189 stable isotopes were added to the software, and target and reference ions were selected.
190 The method file created by the software was used to analyze the samples.

191

192 2.5. Data processing

193 The MS data were exported to a personal computer with the GCMSsolution software
194 (Shimadzu Co.), and the peaks for the targeted metabolites and stable isotopes were
195 detected by the software and then checked manually. The relative standard deviation
196 (RSD)% values for the peak area values of each stable isotope and the internal standard
197 were calculated. The metabolites with RSD% values of >20% were excluded. The
198 samples that exhibited outlying or saturated peak area values were excluded. The
199 concentrations of the targeted metabolites in each sample were calculated based on the
200 peak area ratios of the targeted metabolites to the corresponding stable isotopes. All
201 metabolite concentrations are shown in μM.

202

203 We prepared a large stock of plasma mixture as a quality control (QC) standard, and
204 two aliquots of the stock were processed simultaneously with every batch. The
205 concentrations of the targeted metabolites in the QC samples and the associated RSD%
206 values were also calculated. If the RSD% was >20%, the batch was excluded as an
207 outlier.
208
209 2.6. Statistical analyses
210 Univariate analyses were performed by conducting comparisons between the PC
211 patients and HV using Mann-Whitney's U test. P-values of <0.003125 were considered
212 to indicate a significant difference, and Bonferroni's method was used to adjust for
213 multiple comparisons. All 16 metabolites were subjected to univariate regression
214 analysis to confirm their diagnostic performance as single diagnostic markers. The
215 stepwise method was used to select variables from among the 16 metabolites. The
216 multicollinearity of the selected metabolites was assessed by calculating their variance
217 inflation factors. The multivariate analysis was performed via multivariate logistic
218 regression analysis. P-values of <0.05 were considered to indicate significant differences
219 in the multivariate analysis. Receiver operating characteristic (ROC) curves were
220 generated to calculate area under the curve (AUC) values, and sensitivity and

specificity values for the models to evaluate their diagnostic performance. The optimal cut-off values of the models were determined from their ROC curves. All analyses were performed using the default conditions of JMP11 (SAS Institute, Inc.).

3. Results

The clinical characteristics of the PC patients and HV are shown in Table 1. One sample from the PC patients was excluded because of target metabolite peak area saturation. The RSD% values of all stable isotopes, the internal standard, and all QC samples were <20%. In the univariate analyses of the 1st set study, the levels of 11 of the 16 targeted metabolites differed significantly between the PC patients and HV ($p<0.003125$). In the univariate regression analysis (Supplementary Table S1), histidine displayed the highest ROC-AUC value (0.80702) for discriminating between the PC patients and HV. Its sensitivity and specificity were 72.2% and 86.0%, respectively.

To validate the diagnostic performance of our previous model, we constructed a diagnostic model for PC via multivariate logistic regression analysis based on the results of the 4 metabolites identified in our previous study; i.e., 1,5-AG, histidine, inositol, and xylitol, which did not display multicollinearity (Supplementary Table S2).

239 The formula for this diagnostic model was as follows:

240

241 Model X

242 $p=1/[1+\exp^{-\{9.28-0.06(histidine)-0.05(inositol)-0.002(1,5\text{-}AG)-0.07(xylitol)\}}]$

243

244 In the 1st set, the AUC value of model X was 0.86062, and its optimal cut-off value was

245 0.5677. At this cut-off value, model X exhibited sensitivity and specificity values of

246 74.1% and 86.0%, respectively (Figure 1). These results were almost the same as those

247 reported previously. However, as shown in Supplementary Table S2, inositol, 1,5-AG,

248 and xylitol did not display significance in the multivariate analysis. To construct a more

249 effective diagnostic model for PC, we selected variables from among the 16 metabolites

250 in the 1st set via the stepwise method. As a result, 2 metabolites; i.e., histidine and

251 xylitol, were selected, and the formula for this multivariate logistic regression model

252 was as follows (Supplementary Table S3):

253

254 Model Y

255 $p=1/[1+\exp^{-\{8.15-0.08(histidine)-0.10(xylitol)\}}]$

256

257 Model Y displayed an AUC of 0.83008, and its optimal cut-off value was 0.6048. At this
258 cut-off value, the sensitivity and specificity of model Y were 70.4% and 89.5%,
259 respectively. We defined the cut-off value for CA19-9 as 37 U/ml, which is typically used
260 in the general clinical field. In the 1st set, CA19-9 demonstrated sensitivity and
261 specificity values of 68.5% and 93.0%, respectively (Figure 1). The GC/MS/MS-based
262 diagnostic models for PC exhibited higher sensitivity values than CA19-9 (Table 2).
263
264 Furthermore, it was revealed that the combination of model Y and CA19-9
265 demonstrated increased sensitivity and specificity for PC. The combination of model Y
266 and CA19-9 was constructed via multivariate logistic regression analysis
267 (Supplementary Table S4). This model displayed an AUC value of 0.93112 and an
268 optimal cut-off value of 0.4172. At this cut-off value, the sensitivity and specificity of
269 this model were 90.7% and 89.5%, respectively (Table 2). Similar results were obtained
270 for the combination of model X and CA19-9 (Supplementary Fig. S1).
271
272 Next, we further validated the diagnostic performance of these models in the 2nd set
273 study (Table 3). To eliminate the influence of slight differences between preparation
274 procedures at some institutions, the 2nd set was performed. In the 2nd set, sample

275 preparation methods including the blood sampling, plasma preparation and stored
276 methods were completely unified. The results of the ROC curve analysis of the 2nd set
277 are shown in Figure 2. Model X displayed an AUC value of 0.84375, and its optimal
278 cut-off value was 0.9173. At this cut-off value, model X displayed sensitivity and
279 specificity values of 81.3% and 87.5%, respectively. Model Y displayed an AUC of
280 0.89453, and its optimal cut-off value was 0.9227. At this cut-off value, model Y
281 demonstrated sensitivity and specificity values of 75.0% and 100.0%, respectively. The
282 sensitivity and specificity of CA19-9 were 75.0% and 81.3%, respectively. We established
283 combined models involving CA19-9 via multivariate logistic regression analysis. The
284 combination of model X with CA19-9 displayed an AUC value of 0.89063, and its
285 optimal cut-off value was 0.8860. At this cut-off value, its sensitivity and specificity
286 were 81.3% and 93.8%, respectively. The combination of model Y with CA19-9 displayed
287 an AUC value of 0.91797, and its optimal cut-off value was 0.8369. At this cut-off value,
288 its sensitivity and specificity were 81.3% and 93.8%, respectively (Supplementary Table
289 S5). Although there were no significant differences between the performance of the
290 models with and without CA19-9 in the 2nd set, the combined models tended to exhibit
291 increased sensitivity and specificity. Furthermore, in the 2nd set model X, model Y,
292 CA19-9, model X with CA19-9, and model Y with CA19-9 exhibited sensitivity values for

resectable PC (stages 0 to IIB) of 100%, 85.7%, 85.7%, 100%, and 100%, respectively.

4. Discussion

In the 1st set study, model X exhibited lower sensitivity for PC than it did in our previous study. There are a number of possible reasons for this. First, the distribution of the clinical stage of PC differed between the two populations. As advanced stage PC predominated in our previous study population, model X was greatly affected by advanced PC. Second, differences in the measurement methods, such as the use of GC/MS/MS and the internal standards employed, might also have contributed to these discrepancies. Regarding the internal standard, in the semi-quantitative analysis performed in our previous study the peak height of each ion was measured and normalized to the peak height of 2-isopropylmalic acid alone as an internal standard. In contrast, in this study the peak area values of the targeted metabolites were normalized to those of the corresponding stable isotopes, and then the blood concentration of each metabolite was calculated. Normalization using corresponding stable isotopes produces more accurate quantitative data than normalization using a single representative internal standard. As some of the results obtained in our previous study might not have been accurate, our previous model might not have been appropriate. Therefore, we

311 constructed a new model (model Y) based on the results of this study.

312

313 In our previous study, 1,5-AG was the substance whose levels differed most significantly

314 between the PC patients and HV. However, in this study 1,5-AG was not selected as a

315 variable in model Y. 1,5-AG is a metabolically stable metabolite that is mainly derived

316 from the diet and absorbed in the intestinal tract [23]. It is commonly used as a marker

317 of short-term glycemic control. In healthy individuals, the blood level of 1,5-AG is

318 relatively constant because of the balance between the absorption of 1,5-AG, the

319 reabsorption of urinary 1,5-AG, and the secretion of 1,5-AG in the intestinal tract, and

320 little or no biochemical transformation of the molecule occurs in the body. In

321 hyperglycemia, the secretion of 1,5-AG is increased by competitive inhibition of urinary

322 1,5-AG reabsorption by glucose, so the blood 1,5-AG level is inversely correlated with

323 the blood glucose concentration [24]. It was also reported that insulin resistance is

324 associated with the aggressiveness of PC [25]. In summary, the blood level of 1,5-AG

325 decreases as PC progresses. In addition, decreased blood 1,5-AG levels might be caused

326 by a reduction in food intake in patients with advanced PC. These reasons might

327 explain the non-selection of 1,5-AG as a variable for model Y.

328

329 In this study, it was suggested that histidine might play an important role in the early
330 detection of PC. Histidine is an essential amino acid in humans and other mammals,
331 and has anti-oxidant, anti-inflammatory, and anti-secretory properties in mice [26]. Low
332 serum histidine levels are observed in arthritis, chronic renal failure, and obesity in
333 humans [27, 28]. Histidine is a precursor of histamine. Histamine is known to have the
334 ability to accelerate cancerous cells into cell cycle arrest and inhibit the proliferation of
335 PC cells [29, 30]. Therefore, histidine decarboxylase (HDC), which converts histidine to
336 histamine, has recently become an important focus of study. In normal pancreatic islet
337 cells, HDC is predominantly found in glucagon cells, but in pancreatic neuroendocrine
338 tumors, HDC is found in all types of islet cells; i.e., glucagon-, insulin-, somatostatin-,
339 pancreatic polypeptide-, and serotonin-producing enterochromaffin cells. On the other
340 hand, HDC is not expressed in PC cells derived from the pancreatic duct epithelium [31].
341 Therefore, in PC cells it is unlikely that the histidine concentration is decreased by its
342 conversion to histamine. Recently, it was also reported that a multivariate index
343 composed of plasma free amino acids is useful for the early detection of PC. In the latter
344 report, decreased plasma levels of tryptophan and histidine were observed [32]. It was
345 suggested that there are several possible mechanisms that can influence the plasma
346 free amino acid levels of cancer patients, including marked metabolic changes in local

347 cancerous lesions, the induction of remote organ metabolic changes caused by factors
348 emitted from cancer cells, and the involvement of the immune system. However, several
349 points regarding the mechanisms responsible for changes in plasma free amino acid
350 levels in PC remain unclear. In addition, PC patients can also be affected by
351 malnutrition because of insufficient pancreatic endocrine and exocrine function.
352 Regarding exocrine function, the pancreas secretes lipase, amylase, and trypsinogen,
353 which are the enzymes that degrade lipids, carbohydrates, and proteins. The failure of
354 the production or secretion of trypsinogen by PC can cause impaired absorption of
355 histidine from the intestinal tract. It is considered that the decreased levels of histidine
356 seen in PC might be the result of malnutrition. Further research is necessary to
357 examine these issues.
358
359 Xylitol is not only found in the human body from exogenous sources (such as the diet)
360 but it is also an endogenous metabolite that is produced and metabolized in the liver
361 [33]. Some possible endogenous mechanisms may be through reduced food intake and
362 liver disorders caused by obstructive jaundice, but its pathway remains unclear. In the
363 future, we will perform a larger scale study using a larger number of samples, to try to
364 elucidate the possible mechanisms.

365

366 CA19-9 is used as a tumor marker of PC. In the 1st set, the sensitivity of CA19-9 was

367 high in spite of the fact that this study focused on early stage PC. However, the level of

368 CA19-9 is also often increased in some benign diseases, especially obstructive jaundice

369 [34]. In the 1st set, 15 of the 54 PC patients had obstructive jaundice that required

370 biliary drainage. In 13 of these 15 cases, elevated CA19-9 levels were observed. Thus, it

371 was considered that the high sensitivity of CA19-9 is affected not only by PC, but also by

372 obstructive jaundice. In the patients with obstructive jaundice, the sensitivity of model

373 X, model Y and CA 19-9 were 85.6%, 86.7% and 86.7%, respectively. On the other hand,

374 in the patients without obstructive jaundice, the sensitivity of model X, model Y, and CA

375 19-9 were 70.0%, 64.1%, and 61.5%, respectively. However, the number of samples is

376 small to evaluate the influence of obstructive jaundice, so it is needed to perform the

377 larger scale study using a larger number of samples in the future.

378

379 In this study, the optimal cut-off values obtained in the 2nd set were not the same as

380 those calculated using the 1st set. The 1st set was composed of serum samples, whereas

381 the 2nd set was composed of plasma samples. Due to the differences in the preparation of

382 these two blood components, metabolite levels are expected to differ between serum and

383 plasma. Serum is obtained from whole blood after the clotting process, whereas plasma
384 is obtained from whole blood in the presence of anticoagulant so coagulation factors are
385 not activated, and thus, no blood clots form. It has previously been reported that serum
386 exhibits significantly higher metabolite concentrations than plasma [35]. Since the
387 same result was confirmed in our preliminary study (data not shown), it can be said
388 that the procedure used for the optimization of the cut-off value was appropriate.
389
390 This study had several limitations. First, the sample type should be unified to evaluate
391 the reproducibility of the GC/MS/MS-based diagnostic models. Second, although the HV
392 were chosen based on the criteria described above, it is almost impossible to prove that
393 these patients did not have tiny cancerous lesions, especially in the pancreas. Third, the
394 number of samples included in the 2nd set was relatively small. A further large and
395 global study is needed to confirm these promising results using the layering analysis,
396 cross validation, bootstrapping and so on. In addition, it is also needed to take the
397 quality of human samples into account in the large and global study.
398
399 In conclusion, we established a sophisticated diagnostic model for PC via multivariate
400 logistic regression analysis and GC/MS/MS-based targeted metabolomics. Our model

possessed higher sensitivity than CA19-9. Furthermore, it was revealed that the combination of the selected metabolites and CA19-9 exhibited increased sensitivity and specificity for PC, and the reproducibility of these results was confirmed in another independent set. It was suggested that combining information about the selected metabolites with data regarding CA19-9, which is a conventional tumor marker of PC, might contribute to the early detection of PC. In this study, the combination of model X; i.e., histidine, inositol, 1,5-AG and xylitol, and CA19-9 demonstrated higher sensitivity and specificity than model X alone. The combination of model Y; i.e., histidine and xylitol, and CA19-9 demonstrated higher sensitivity and specificity than model Y alone. This finding is expected to improve the prognosis of patients with PC by aiding the early detection of the disease.

Acknowledgements, Financial support

This study was supported in part by a Grant-in-Aid for Scientific Research (B) from the Japan Society for the Promotion of Science [M.Y.]; a Grant-in-Aid for Young Scientists (B) from the Japan Society for the Promotion of Science [T.K.]; Practical Research for Innovative Cancer Control from the Japan Agency for Medical Research and Development [16ck0106101h0103 to T.K., S.N., T.A., M.Y.; 15ck0106101h0002 to KH];

the Project for the Development of Innovative Research on Cancer Therapeutics (P-DIRECT) from the Japan Agency for Medical Research and Development [M.Y.]; , the Project for Cancer Research And Therapeutic Evolution (P-CREATE) from the Japan Agency for Medical Research and Development (16cm0106403h0001 to KH); and AMED-CREST from the Japan Agency for Medical Research and Development [T.K., S.N., K.H., T.A., M.Y.].

5. References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2016, CA: a cancer journal for clinicians 66(1) (2016) 7-30.

[2] K. Hanada, A. Okazaki, N. Hirano, Y. Izumi, Y. Teraoka, J. Ikemoto, K. Kanemitsu, F. Hino, T. Fukuda, S. Yonehara, Diagnostic strategies for early pancreatic cancer, Journal of gastroenterology 50(2) (2015) 147-54.

[3] R. Pannala, A. Basu, G.M. Petersen, S.T. Chari, New-onset diabetes: a potential clue to the early diagnosis of pancreatic cancer, The Lancet. Oncology 10(1) (2009) 88-95.

[4] C.J. Wray, S.A. Ahmad, J.B. Matthews, A.M. Lowy, Surgery for pancreatic cancer: recent controversies and current practice, Gastroenterology 128(6) (2005) 1626-41.

[5] K.Y. Bilimoria, D.J. Bentrem, C.Y. Ko, J. Ritchey, A.K. Stewart, D.P. Winchester, M.S. Talamonti, Validation of the 6th edition AJCC Pancreatic Cancer Staging System: report from the National Cancer Database, Cancer 110(4) (2007) 738-44.

[6] U.K. Ballehaninna, R.S. Chamberlain, The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal, Journal of gastrointestinal oncology 3(2) (2012) 105-19.

[7] K. Honda, S. Srivastava, Potential usefulness of apolipoprotein A2 isoforms for screening and risk stratification of pancreatic cancer, Biomarkers in medicine (2016).

[8] T. Yoneyama, S. Ohtsuki, K. Honda, M. Kobayashi, M. Iwasaki, Y. Uchida, T. Okusaka, S. Nakamori, M. Shimahara, T. Ueno, A. Tsuchida, N. Sata, T. Ioka, Y. Yasunami, T. Kosuge, T. Kaneda, T. Kato, K. Yagihara, S. Fujita, W. Huang, T. Yamada, M. Tachikawa, T. Terasaki,

Identification of IGFBP2 and IGFBP3 As Compensatory Biomarkers for CA19-9 in Early-Stage Pancreatic Cancer Using a Combination of Antibody-Based and LC-MS/MS-Based Proteomics, PloS one 11(8) (2016) e0161009.

[9] K. Honda, M. Ono, M. Shitashige, M. Masuda, M. Kamita, N. Miura, T. Yamada, Proteomic approaches to the discovery of cancer biomarkers for early detection and personalized medicine, Japanese journal of clinical oncology 43(2) (2013) 103-9.

[10] K. Honda, M. Kobayashi, T. Okusaka, J.A. Rinaudo, Y. Huang, T. Marsh, M. Sanada, Y. Sasajima, S. Nakamori, M. Shimahara, T. Ueno, A. Tsuchida, N. Sata, T. Ioka, Y. Yasunami, T. Kosuge, N. Miura, M. Kamita, T. Sakamoto, H. Shoji, G. Jung, S. Srivastava, T. Yamada, Plasma biomarker for detection of early stage pancreatic cancer and risk factors for pancreatic malignancy using antibodies for apolipoprotein-AII isoforms, Scientific reports 5 (2015) 15921.

[11] S. Nishiumi, M. Shinohara, A. Ikeda, T. Yoshie, N. Hatano, S. Kakuyama, S. Mizuno, T. Sanuki, H. Kutsumi, E. Fukusaki, T. Azuma, T. Takenawa, M. Yoshida, Serum metabolomics as a novel diagnostic approach for pancreatic cancer, Metabolomics 6(4) (2010) 518-528.

[12] M. Ooi, S. Nishiumi, T. Yoshie, Y. Shiomi, M. Kohashi, K. Fukunaga, S. Nakamura, T. Matsumoto, N. Hatano, M. Shinohara, Y. Irino, T. Takenawa, T. Azuma, M. Yoshida, GC/MS-based profiling of amino acids and TCA cycle-related molecules in ulcerative colitis, Inflammation research : official journal of the European Histamine Research Society ... [et al.] 60(9) (2011) 831-40.

[13] A. Ikeda, S. Nishiumi, M. Shinohara, T. Yoshie, N. Hatano, T. Okuno, T. Bamba, E. Fukusaki, T. Takenawa, T. Azuma, M. Yoshida, Serum metabolomics as a novel diagnostic approach for gastrointestinal cancer, Biomedical chromatography : BMC 26(5) (2012) 548-58.

[14] Y. Shiomi, S. Nishiumi, M. Ooi, N. Hatano, M. Shinohara, T. Yoshie, Y. Kondo, K. Furumatsu, H. Shiomi, H. Kutsumi, T. Azuma, M. Yoshida, GCMS-based metabolomic study in mice with colitis induced by dextran sulfate sodium, Inflammatory bowel diseases 17(11) (2011) 2261-74.

[15] S. Nishiumi, T. Kobayashi, A. Ikeda, T. Yoshie, M. Kibi, Y. Izumi, T. Okuno, N. Hayashi, S. Kawano, T. Takenawa, T. Azuma, M. Yoshida, A novel serum metabolomics-based diagnostic approach for colorectal cancer, PloS one 7(7) (2012) e40459.

[16] M. Kohashi, S. Nishiumi, M. Ooi, T. Yoshie, A. Matsubara, M. Suzuki, N. Hoshi, K. Kamikozuru, Y. Yokoyama, K. Fukunaga, S. Nakamura, T. Azuma, M. Yoshida, A novel gas chromatography mass spectrometry-based serum diagnostic and assessment approach to ulcerative colitis, Journal of Crohn's & colitis 8(9) (2014) 1010-21.

[17] M. Yoshida, N. Hatano, S. Nishiumi, Y. Irino, Y. Izumi, T. Takenawa, T. Azuma,

483 Diagnosis of gastroenterological diseases by metabolome analysis using gas
484 chromatography-mass spectrometry, Journal of gastroenterology 47(1) (2012) 9-20.
485 [18] T. Kobayashi, S. Nishiumi, A. Ikeda, T. Yoshie, A. Sakai, A. Matsubara, Y. Izumi, H.
486 Tsumura, M. Tsuda, H. Nishisaki, N. Hayashi, S. Kawano, Y. Fujiwara, H. Minami, T.
487 Takenawa, T. Azuma, M. Yoshida, A novel serum metabolomics-based diagnostic approach to
488 pancreatic cancer, Cancer epidemiology, biomarkers & prevention : a publication of the
489 American Association for Cancer Research, cosponsored by the American Society of
490 Preventive Oncology 22(4) (2013) 571-9.
491 [19] A. Sakai, M. Suzuki, T. Kobayashi, S. Nishiumi, K. Yamanaka, Y. Hirata, T. Nakagawa,
492 T. Azuma, M. Yoshida, Pancreatic cancer screening using a multiplatform human serum
493 metabolomics system, Biomarkers in medicine 10(6) (2016) 577-86.
494 [20] J.L. Spratlin, N.J. Serkova, S.G. Eckhardt, Clinical applications of metabolomics in
495 oncology: a review, Clinical cancer research : an official journal of the American Association
496 for Cancer Research 15(2) (2009) 431-40.
497 [21] R. Wei, G. Li, A.B. Seymour, High-throughput and multiplexed LC/MS/MRM method for
498 targeted metabolomics, Analytical chemistry 82(13) (2010) 5527-33.
499 [22] H. Tsugawa, Y. Tsujimoto, K. Sugitate, N. Sakui, S. Nishiumi, T. Bamba, E. Fukusaki,
500 Highly sensitive and selective analysis of widely targeted metabolomics using gas
501 chromatography/triple-quadrupole mass spectrometry, Journal of bioscience and
502 bioengineering 117(1) (2014) 122-8.
503 [23] T. Yamanouchi, Y. Tachibana, H. Akanuma, S. Minoda, T. Shinohara, H. Moromizato, H.
504 Miyashita, I. Akaoka, Origin and disposal of 1,5-anhydroglucitol, a major polyol in the
505 human body, The American journal of physiology 263(2 Pt 1) (1992) E268-73.
506 [24] T. Yamanouchi, N. Ogata, T. Tagaya, T. Kawasaki, N. Sekino, H. Funato, L. Akaoka, H.
507 Miyashita, Clinical usefulness of serum 1,5-anhydroglucitol in monitoring glycaemic control,
508 Lancet (London, England) 347(9014) (1996) 1514-8.
509 [25] E. Dugnani, G. Balzano, V. Pasquale, M. Scavini, F. Aleotti, D. Liberati, G. Di Terlizzi, A.
510 Gandolfi, G. Petrella, M. Reni, C. Doglioni, E. Bosi, M. Falconi, L. Piemonti, Insulin
511 resistance is associated with the aggressiveness of pancreatic ductal carcinoma, Acta
512 diabetologica (2016).
513 [26] J.W. Peterson, I. Boldogh, V.L. Popov, S.S. Saini, A.K. Chopra, Anti-inflammatory and
514 antisecretory potential of histidine in Salmonella-challenged mouse small intestine,
515 Laboratory investigation; a journal of technical methods and pathology 78(5) (1998) 523-34.
516 [27] R.N. Feng, Y.C. Niu, X.W. Sun, Q. Li, C. Zhao, C. Wang, F.C. Guo, C.H. Sun, Y. Li,
517 Histidine supplementation improves insulin resistance through suppressed inflammation in
518 obese women with the metabolic syndrome: a randomised controlled trial, Diabetologia 56(5)

(2013) 985-94.
[28] D.A. Gerber, Low free serum histidine concentration in rheumatoid arthritis. A measure of disease activity, The Journal of clinical investigation 55(6) (1975) 1164-73.
[29] G. Cricco, G. Martin, V. Medina, M. Nunez, N. Mohamad, M. Croci, E. Crescenti, R. Bergoc, E. Rivera, Histamine inhibits cell proliferation and modulates the expression of Bcl-2 family proteins via the H2 receptor in human pancreatic cancer cells, Anticancer research 26(6B) (2006) 4443-50.
[30] G. Martin, G. Cricco, Z. Darvas, M. Croci, M. Nunez, R. Bergoc, A. Falus, E. Rivera, Histamine inhibits proliferation of a pancreatic carcinoma cell line without inducing apoptosis significantly, Inflammation research : official journal of the European Histamine Research Society ... [et al.] 51 Suppl 1 (2002) S67-8.
[31] A. Tanimoto, Y. Matsuki, T. Tomita, T. Sasaguri, S. Shimajiri, Y. Sasaguri, Histidine decarboxylase expression in pancreatic endocrine cells and related tumors, Pathology international 54(6) (2004) 408-12.
[32] N. Fukutake, M. Ueno, N. Hiraoka, K. Shimada, K. Shiraishi, N. Saruki, T. Ito, M. Yamakado, N. Ono, A. Imaizumi, S. Kikuchi, H. Yamamoto, K. Katayama, A Novel Multivariate Index for Pancreatic Cancer Detection Based On the Plasma Free Amino Acid Profile, PloS one 10(7) (2015) e0132223.
[33] O. Hamberg, T.P. Almdal, Effects of xylitol on urea synthesis in normal humans: relation to glucagon, JPEN. Journal of parenteral and enteral nutrition 20(2) (1996) 139-44.
[34] M.S. Lin, J.X. Huang, H. Yu, Elevated serum level of carbohydrate antigen 19-9 in benign biliary stricture diseases can reduce its value as a tumor marker, International journal of clinical and experimental medicine 7(3) (2014) 744-50.
[35] Z. Yu, G. Kastenmuller, Y. He, P. Belcredi, G. Moller, C. Prehn, J. Mendes, S. Wahl, W. Roemisch-Margl, U. Ceglarek, A. Polonikov, N. Dahmen, H. Prokisch, L. Xie, Y. Li, H.E. Wichmann, A. Peters, F. Kronenberg, K. Suhre, J. Adamski, T. Illig, R. Wang-Sattler, Differences between human plasma and serum metabolite profiles, PloS one 6(7) (2011) e21230.

550 Tables

551

552 Table 1. Characteristics of the subjects in the 1st set

		PC	HV	*P* value
N	Total	54	58	
	Male	34	27	
	Female	20	31	
Age, y	Median	69	65	0.054
	SEM	2.73	1.26	
	Range	63-85	45-86	
Stage	IA	4		
	IB	3		
	IIA	12		
	IIB	35		
	III	0		
	IV	0		
Histology/cytology	Ductal adenocarcinoma	53		
	Adenosquamous carcinoma	1		

553 NOTE: The pancreatic cancer staging was based on the UICC TNM classification.

554 SEM, standard error of the mean; P-values were calculated using Mann-Whitney's U

555 test.

556 Abbreviations: PC, pancreatic cancer; HV, healthy volunteers

557

558

559

560

561 Table 2. Diagnostic performance of the constructed models and CA19-9 in the 1st set

		Model X	Model Y	CA19-9	Model Y with CA19-9
1st set	AUC	0.86062	0.83008	0.87508	0.93112
	(95% confidence interval)	(0.7766-0.9164)	(0.7376-0.8946)		(0.8561-0.9685)
	Cut-off value	0.5677	0.6048	37 U/ml	0.4172
	Sensitivity (%)	74.1	70.4	68.5	90.7
	Specificity (%)	86	89.5	93	89.5

562 NOTE: The ROC curves for these models are shown in Figure 1. Regarding CA19-9, the

563 clinical cut-off value used at our institutions was employed.

564

565

566

567

568

569

570

571

572

573

574

575

576 Table 3. Characteristics of the subjects in the 2nd set

		PC	HV	*P*-value
N	Total	16	16	
	Male	10	10	
	Female	6	6	
Age, y	Median	66	66	0.478
	SEM	2.74	2	
	Range	46-85	48-79	
Stage	IA	0		
	IB	0		
	IIA	4		
	IIB	3		
	III	6		
	IV	3		
Histology/cytology	Ductal adenocarcinoma	16		

577 NOTE: The pancreatic cancer staging was based on the UICC TNM classification.

578 SEM, standard error of the mean; P-values were calculated using Mann-Whitney's U

579 test.

580 Abbreviations: PC, pancreatic cancer; HV, healthy volunteers

581

Figure(s)

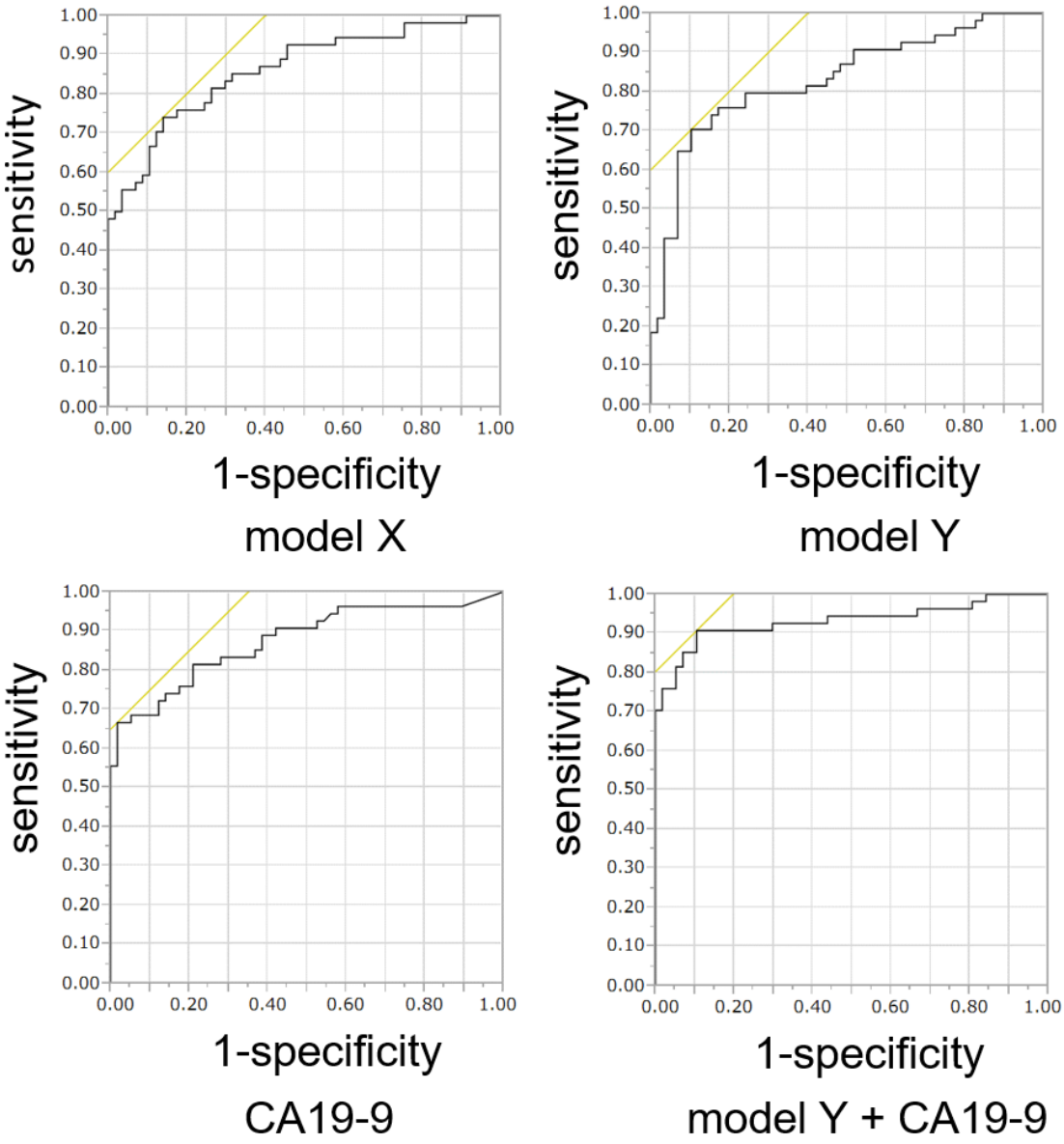


Figure 1: The ROC curves for model X, model Y, CA19-9 and the combination of model Y and CA19-9 in the 1st set

The AUC of the ROC curve, cut-off value, sensitivity, and specificity of each diagnostic model are summarized in Table 2.

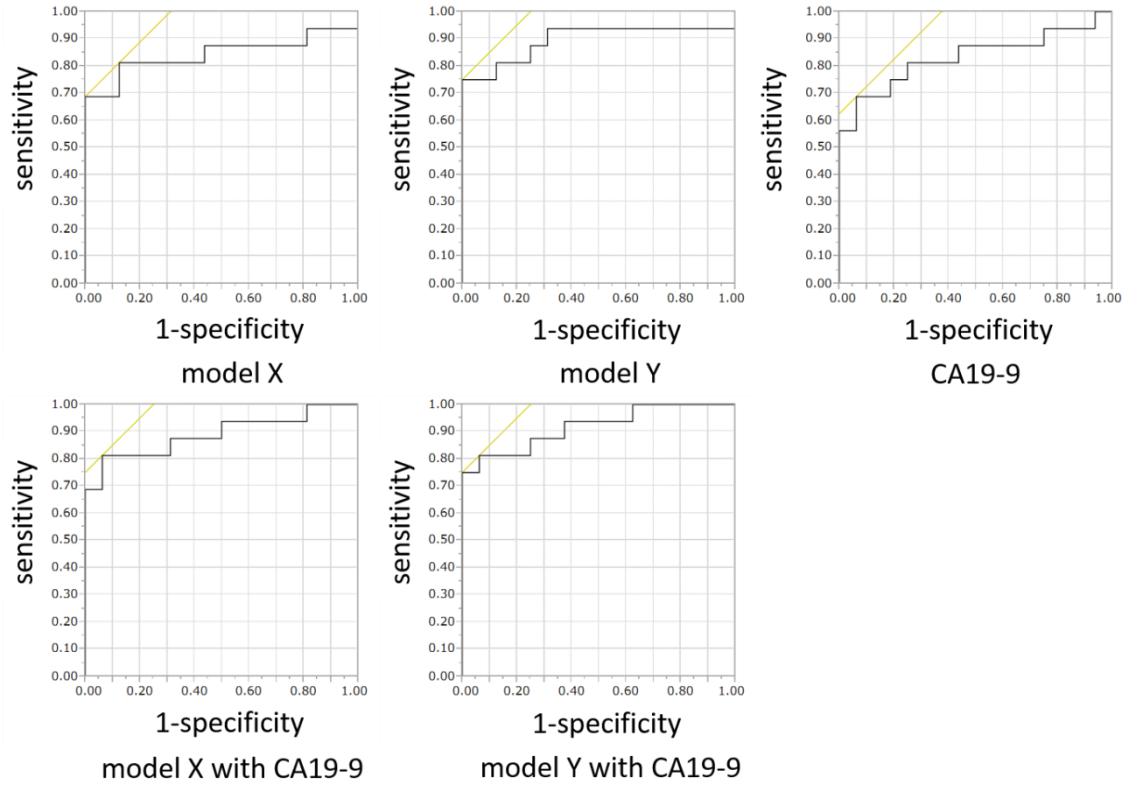


Figure 2: The ROC curves for the diagnostic models, CA19-9, and the combined models in the 2nd set

The AUC of the ROC curve, and the cut-off, sensitivity, and specificity values of each diagnostic method are summarized in Supplementary Table 5.

Figure legends

Figure 1: The ROC curves for model X, model Y, CA19-9 and the combination of model Y and CA19-9 in the 1st set

The AUC of the ROC curve, cut-off value, sensitivity, and specificity of each diagnostic model are summarized in Table 2.

Figure 2: The ROC curves for the diagnostic models, CA19-9, and the combined models in the 2nd set

The AUC of the ROC curve, and the cut-off, sensitivity, and specificity values of each diagnostic method are summarized in Supplementary Table 5.

Supplementary Table 1. The diagnostic performance of the metabolites as individual biomarkers of PC in the 1st set

Metabolite	AUC	Cut-off value (μM)	Sensitivity (%)	Specificity (%)	*P*-value
Valine	0.72612	232.6426	51.85	91.23	<0.0001
2-Aminoethanol	0.70435	3.95206	62.96	73.68	<0.0001
Nonanoic acid	0.58317	2.283348	40.74	80.7	0.1095
Threonine	0.68486	143.2763	59.26	73.68	0.0003
Methionine	0.65302	28.11616	48.15	84.21	0.0051
Creatinine	0.69396	100.0639	57.41	77.19	0.002
Arabinose	0.58934	2.349835	75.93	49.12	0.2807
Asparagine	0.73879	76.0426	70.37	73.68	<0.0001
Xylitol	0.69038	2.7624	59.26	73.68	0.0009
Glutamine	0.5731	548.0356	31.48	91.23	0.1928
1,5-Anhydro-D-glucitol	0.74464	195.53	59.26	92.98	0.0005
Lysine	0.66732	178.9633	48.15	84.21	0.0021
Histidine	0.80702	94.366	72.22	85.96	<0.0001
Tyrosine	0.67316	73.71	70.37	63.16	0.021
Inositol	0.78947	42.69775	75.93	77.19	<0.0001
Uric acid	0.59487	263.6015	44.44	75.44	0.0695

NOTE: The AUC and optimal cut-off values were obtained from the ROC curve for each metabolite (data not shown).

Supplementary Table 2. Data for model X

Metabolite	Coefficient	Standard error	*P*-value	Lower 95% CI	Upper 95% CI
(Intercept)	9.28401438	1.8801164	<0.0001	5.9153265	13.350226
Histidine	-0.064642	0.0174991	0.0002	-0.1012822	-0.032005
Inositol	-0.0461278	0.0249312	0.0643	-0.0974238	0.00359945
1,5-Anhydro-D-glucitol	-0.0023751	0.0017899	0.1845	-0.006229	0.00087884
Xylitol	-0.0715707	0.0664941	0.2818	-0.2754049	-0.0096064

The *p*-values were calculated using the Wald test.

Supplementary Table 3. Data for model Y

Metabolite	Coefficient	Standard error	*P*-value	Lower 95% CI	Upper 95% CI
(Intercept)	8.15358969	1.6585247	<0.0001	5.14553825	11.6974441
Histidine	-0.0790614	0.0166341	<0.0001	-0.1144571	-0.0487372
Xylitol	-0.0956385	0.0849777	0.2604	-0.3212372	-0.0117317

The *p*-values were calculated using the Wald test.

Supplementary Table 4. Data for the combination of model Y and CA19-9

Metabolite	Coefficient	Standard error	*P*-value	Lower 95% CI	Upper 95% CI
(Intercept)	5.96517284	2.0346642	0.0034	2.2603558	10.3427555
Histidine	-0.0718331	0.0213236	0.0008	-0.1182578	-0.0334605
CA19-9	0.04040841	0.0138489	0.0035	0.01862254	0.0736808
Xylitol	-0.1373177	0.1913994	0.4731	-0.6207619	-0.0010367

The *p*-values were calculated using the Wald test.

Supplementary Table 5. Diagnostic performance of the constructed models and CA19-9 in the 2nd set

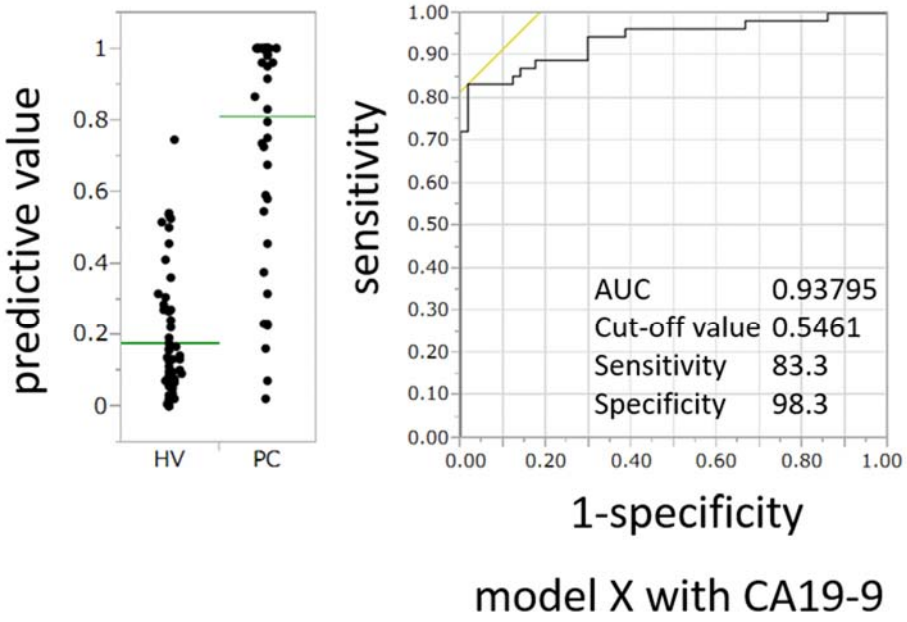
		Model X	Model Y	CA19-9	Model X with CA19-9	Model Y with CA19-9
2nd set	AUC	0.84375	0.89453	0.83203	0.89063	0.91797
	(95% confidence interval)	(0.6180-0.9474)	(0.8788-0.9715)		(0.6984-0.9663)	(0.7534-0.9762)
	Cut-off value	0.9173	0.9227	37 U/ml	0.886	0.8369
	Sensitivity (%)	81.3	75	75	81.3	81.3
	Specificity (%)	87.5	100	81.3	93.8	93.8

NOTE: The ROC curves for these evaluations are shown in Supplementary Figure 2.

Regarding CA19-9, the clinical cut-off value used at our institutions was employed.

Figure legends

Supplementary Figure 1: The ROC curve for the combination of model X and CA19-9 in the 1st set



Supplementary Figure 1: The ROC curve for the combination of model X and CA19-9 in the 1st set