의학박사 학위논문

Development of biomarker complex for early diagnosis of pancreatic cancer

췌장암 조기 진단을 위한 복합 바이오마커 개발

2020 년 12 월

서울대학교 대학원

의학과 외과학 전공

김 홍 범

Development of biomarker complex for early diagnosis of pancreatic cancer
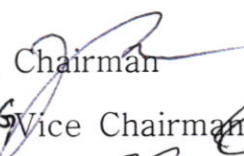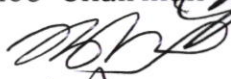
by Hongbeom Kim, M.D.

(Directed by Jin−Young Jang, M.D., PhD.)

A Thesis Submitted to the Department of Surgery in Partial Fulfilment of the Requirements for the Degree of Doctor of Philosophy in Medicine (Surgery) at the Seoul National University College of Medicine

December, 2020

Approved by thesis committee

Professor ___Ryu, Ji Kon___ Chairman

Professor ___JIN−YOVNG JANG___ Vice Chairman

Professor ___Haeryong Khu___

Professor ___NAM−JOON YI___

Professor ___Hyeon Kook Lee___

췌장암 조기 진단을 위한
복합 바이오마커의 개발

지도교수 장 진 영

이 논문을 의학박사 학위논문으로 제출함

2020 년 12 월

서울대학교 대학원

의학과 외과학 전공

김 홍 범

김홍범의 박사 학위논문을 인준함

2020 년 12 월

위 원 장	류지곤	(인)
부위원장	장 진 영	(인)
위 원	김 혜 정	(인)
위 원	이 남 준	(인)
위 원	이 현 국	(인)

Abstract

Development of biomarker complex for early diagnosis of pancreatic cancer

Hongbeom Kim

Medicine

The Graduate School

Seoul National University

Background: The diagnostic value of a single tumor marker for pancreatic cancer is not high. To overcome this limitation, combinations of multiple biomarkers and unique analysis algorithms can be applied. The purpose of this study was to develop multiple biomarker panel and to validate its performance diagnosing pancreatic ductal adenocarcinoma (PDAC).

Methods: From July 2010 to May 2015, 180 PDAC and 573 normal control serum samples were used. As benign pancreatic disease, low grade intradctal papillary mucinous neoplasm (IPMN) samples were used. Candidate markers consisted of 11 markers. Two-thirds of the samples were used as a training set. Validation was performed with the remaining one-third. Five linear or nonlinear classification methods were used to select the optimal model.

Results: There were differences in 10 out of 11 markers (ApoA1, CA125, CA19-9, CRP, CYFRA21.1, LRG1, CEA, ApoA2, TTR and D.Dimer). Using 11 markers, 2047 combinations

were made and all combinations were applied to 5 algorithms. The optimal multiple biomarker combination consisted of 6 marker ApoA1, CA125, CA19-9, CEA, ApoA2, and TTR with the random forest classification algorithms method. The AUC, sensitivity and specificity were 0.992, 95% and 96% and in training set, respectively. In the validation set the AUC, sensitivity, and specificity were 0.993, 96% and 93%, respectively. Diagnostically, CA9-9 alone was 71% accurate in stage 1 and 83% accurate in stage 2. However, the diagnostic accuracy of new panel were 89% and 92% in validation set. When comparing IPMN and PDAC, the AUC, sensitivity and specificity were 0.964, 91.1% and 81.4%

Conclusion: This study demonstrated the usefulness of the combination of multiple biomarkers for early detection of PDAC. Diagnostic biomarker panel algorithms that included 6 biomarkers were developed and validated. These algorithms help to diagnosis pancreatic cancer, especially stage 1, 2 PDAC.

Contents

List of Tables

List of Figures

Introduction

Pancreatic cancer is the third leading cause of cancer-related death. In the future, it is expected to become the second leading cause of cancer-related death, following lung cancer (1). Despite advances in surgical techniques and the introduction of new treatment strategies, the prognosis of pancreatic cancer remains poor (2), mainly because patients frequently present in an inoperable metastatic state or locally advanced state at the time of diagnosis. There are no pancreatic cancer-specific symptoms, therefore early detection is difficult. Only 20% of all patients with pancreatic cancer are eligible for surgery (3). Early diagnosis is required to improve pancreatic cancer survival.

Pancreatic cancer is diagnosed through imaging technologies such as computed tomography (CT) or magnetic resonance imaging (MRI). However, there are many obstacles to adopting these modalities for initial screening. The ideal initial screening test should be efficient, with high sensitivity and specificity, as well as being safe, readily available, convenient, and affordable (4). Carbohydrate antigen 19-9 (CA159-9) is currently the most effective and widely used biomarker for pancreatic cancer (5,6). The median diagnostic sensitivity of CA 19-9 is 79% and the median specificity is approximately 80%, limiting the utility of CA19-9 in the screening of pancreatic cancer [6]. The diagnostic value of a single tumor marker is not high in pancreatic cancer. Therefore CA19-9 is more valuable in the detection of recurrence or the assessment of the response to adjuvant treatment (7,8).

To overcome the limitations of single tumor markers in the screening of pancreatic cancer, diagnostic methods utilizing combinations of multiple biomarkers can be applied. The necessity for multiple biomarkers in the diagnosis of pancreatic cancer is due to tumor

heterogeneity and the cancer microenvironment. Even among single tumors, differences exist in differentiation or evolutionary steps among intra-tumor cells, resulting in intra-tumor heterogeneity within solid cancers (9). In order to comprehensively assess the status of the tumor microenvironment, a number of markers should be analyzed (10). An in vitro diagnostic multivariate index assay (IVDMIA), which combines multiple biomarkers and adds unique analysis algorithms, is useful for the diagnosis of cancer (11). Multiple biomarker panels consisting of 19 serum proteins have previously been constructed via an extensive screening process using serum samples from patients with a variety of cancers as well as healthy controls (12–15).

The purpose of this study was to develop diagnostic algorithms using multiple biomarker panels and to validate their performance in the diagnoses of pancreatic ductal adenocarcinoma (PDAC). To the knowledge of the authors, this article is the first to evaluate PDAC diagnostic ability in a cancer panel that has already been commercialized and used for various cancer diagnosis.

.

Materials and methods

Patient Samples

From July 2010 to May 2015, PDAC samples were collected from patients who underwent surgery at Seoul National University Hospital. Whole blood samples were collected in 10-ml syringes prior to surgery using standard blood collection techniques and were stored in EDTA tubes at room temperature for 1 hour. Samples were centrifuged at 3000 g for 5 min, after which supernatants were collected and stored at -80 oC. Control blood samples were obtained from 573 healthy individuals who visited the hospital for medical check-ups and agreed to participate in the study. Control subjects with confirmed cancer, suspected cancer, or inflammatory conditions that needed medical management were excluded through the following examinations: (1) medical history, (2) physical examination, (3) routine blood analysis, (4) chest X-ray, (5) abdominal sonography or computed tomography, (6) esophagogastroduodenoscopy, (7) colonoscopy, sigmoidoscopy with stool hemoglobin, or computed tomographic colonoscopy, and (8) mammography or breast sonography in women and/or thyroid sonography. As benign pancreatic disease, low grade intraductal papillary mucinous neoplasm (IPMN) samples were used as to determine whether it was possible to differentiate between benign pancreatic disease and PDAC. Clinico-pathologic demographics and tumor characteristics were obtained for each patient included in this study. The T status, N status, and TNM stage of each tumor were classified according to the 8th edition of the American Joint Committee on Cancer (AJCC) classification. PDAC samples and control samples were randomly assigned to either the training set or the validation set. Two-thirds of the samples were used as the training set and validation was performed with the remaining one-third of samples. This study was waived from consent. Including waiving informed consent, this study was approved by the Institutional Review Board at Seoul National University

Hospital (H-1703-005-835).

Selection of Candidate Markers

The commercial Korean pan-cancer panel consisting of 19 biomarkers has been studied in the context of hepatocellular carcinoma, breast cancer, lung cancer, gastric cancer, colon cancer, and prostate cancer [12–15]. Of the 19 biomarkers in the panel, 10 markers (Apolipoprotein A1 [ApoA1], cancer antigen 125 [CA 125], CA19-9, C-reactive protein [CRP], cytokeratin 19 fragment 21.1 [CYFRA21.1], carcinoembryonic antigen [CEA], ApoA2, transthyretin [TTR], beta-2 microglobulin [B2M] and D.Dimer) were selected, for which an automated immunological and clinical chemistry testing platform was completed. Based on findings reported in the Korean pancreatic cancer diagnostic marker study, leucine rich alpha-2-glycoprotein 1 (LRG1) was added for a final panel of 11 candidate markers (16).

ApoA1, ApoA2, B2M, CRP, D-Dimer, and TTR were measured on the Cobas c501 (Hoffmann-La Roche AG., Basel, Switzerland) by the immunoturbidimetric method, and CA 125, CA 19-9, CEA, and CYFRA21.1 were measured on the Cobas e601 (Hoffmann-La Roche AG., Basel, Switzerland) using the electrochemiluminescent detection method, according to the manufacturer's instructions. LRG-1 was measured by enzyme-linked immunosorbent assay (ELISA) using recombinant LRG1 protein and anti-Human LRG1 antibody (R&D Systems, Minneapolis, MN, USA).

Data analysis

The Mann–Whitney U test and Wilcoxon rank-sum test were used for the analysis of the 11 candidate protein biomarkers to detect statistically significant differences in biomarker expression between PDAC samples and control samples. A p-value less than 0.01 was considered statistically significant. The data was then log transformed to minimize the influence of outliers among the biomarker measurements, and data pre-treatment was

performed to convert the age data to categorical data in order to address bias in the distribution of the numerical values for age.

A classification model was generated based on the training data set using linear classification methods (Generalized Linear Model [GLM] algorithm and Ridge regression algorithm), non-linear classification methods (Support Vector Machine [SVM] algorithm and Random Forest [RF] algorithm), and a combination of a linear classification method and a nonlinear classification method (GLM+RF algorithm), which has the advantages of both the linear classification method (i.e., ease of interpretation) and the nonlinear classification method (i.e., robust performance). The model was then verified, and its performance was evaluated using a 10-fold cross-validation technique, to confirm the stability of the model.

The criteria for evaluating the performance of a classification model are as follows: the area under the curve (AUC) of the receiver operating characteristic (ROC) produced at model generation should be large, and the protein marker panel should demonstrate excellent performance with respect to all 5 of the classification algorithms.

All analysis was performed using R statistical package (version 3.5.1) and SPSS version 25.0 (IBM SPSS Statistics, Armonk, NY: IBM Corp).Sixty patients who underwent deceased donor kidney transplantation between March 2015 and October 2017 and whose donor blood sample and/or renal tissues were available were enrolled in this study. This study was approved by the institutional review board of Seoul National University Hospital (IRB No.H-1611-048-807). The study was performed in accordance with the Declaration of Helsinki, and written informed consent was obtained from all participants. Blood samples for this study were provided by the Biobank of Seoul National University Hospital, a member of the Korea Biobank Network.

Results

Clinical Characteristics

A total of 180 PDAC samples and 573 healthy control samples were included in this study. The mean age of the PDAC patients was 64 years (range: 28-87), and 65.0% were male. Pancreaticoduodenectomy, pylorus preserving or not, was performed in 55.5% of the PDAC cases. Classified according to the AJCC 8th edition, 29.4% were stage 1, 41.1% were stage 2, 16.7% were stage 3, and 12.8% were stage 4. The mean age of the healthy control group was 57 years, and 58.3% were male. The samples were divided into a training data set for selecting optimal marker panels (120 pancreatic cancer samples and 382 normal control samples) and a validation data set for verifying the selected marker panels (60 pancreatic cancer samples and 191 normal control samples). Clinico-pathologic data were evenly balanced between the training and validation sets (Table 1).

Biomarker selection and model development

The overall study process is shown in Figure 1. Among the 11 candidate biomarkers, all except B2M showed a statistical difference between PDAC and healthy control samples (Figure 2). The marker panels used in the generation of the model consisted of 2,047 combinations, which is the total number of all possible combinations ($_{11}C_1 + _{11}C_2 + --- + _{11}C_{11}$) of the 11 candidate biomarkers. After adding age and gender variables to each panel, the combination was then applied to the 5 classification algorithms

Out of the top 10% of the initial 2047 sets, we selected 137 sets containing CEA and CA19-9, as these are used as tumor markers in PDAC and digestive system cancer. The validation

data set was then applied to the classification model that had been generated using the selected candidate marker panels, to assess whether the model performed similarly for both the validation and training data sets. We selected 32 sets that demonstrated excellent performance and minimal differences between the training and validation sets. Of these, a marker set with excellent performance independent of the linear and nonlinear methods was selected as the new marker set. The AUC in the validation set was 0.993 for RF, 0.983 for GLM, 0.986 for GLM +RF, 0.985 for RIDGE and 0.991 for SVM. The final marker panel consisted of ApoA1, CA125, CA19-9, CEA, ApoA2, and TTR with the RF classification algorithm method.

Diagnostic performance of new biomarker combination set

The AUC, specificity and sensitivity were 0.992, 95% and 96% in the training set, and 0.993, 96% and 93% in the validation set. Table 2 shows the diagnostic values when applied to the other statistical algorithms. Comparing CA19-9, CEA and the combination of CA19-9 + CEA, the diagnostic performance of the new model was excellent.

The AUC of the new model was 0.993 in the validation set, and that of the CEA + CA19-9 combination was 0.960. CEA alone had the lowest diagnostic ability for PDAC, and even when combined with CA19-9, diagnostic performance did not increase (Figure 3). In the validation set, the diagnostic accuracy (sensitivity) was 89% in stage 1, 92% in stage 2, and 100% in stages 3 and 4.

Particularly in stages 1 and 2, the new model improved diagnostic ability compared to CA19-9 alone. The diagnostic accuracy (sensitivity) of CA19-9 alone in stages 1 and 2 were 72 % and 83%; however, those of the new model were 89% and 92% in the validation set (Figure 4).

Discriminate between PDAC and benign pancreatic disease

Eighty low grade IPMNs were used to determine if it was possible to differentiate between benign pancreatic disease and PDAC. The 80 samples consisted of main duct type (n=1), branch duct type (n=41), and mixed type (n=38). The concentrations of 6 markers measured in the IPMN sample were between the normal control and PDAC. The AUC, specificity and sensitivity were 0.9638, 91.1% and 81.4% in the training set, and 0.9403, 91.7% and 86.7% in the test set. (Figure 5)

.

Discussion

In this study, we identified a combination of 6 biomarkers (ApoA1, CA125, CA19-9, CEA, ApoA2, and TTR) through an RF classification algorithm method that increased the diagnostic accuracy of PDAC to 95%.

In general, a single tumor marker is used to screen for each type of cancer, but the rate of false positives and false negatives is high. Cancer cells do not always secrete tumor markers or do not secret the same tumor marker even within a single tumor. As well, tumor markers may increase in chronic diseases or other cancers (9, 17, 18). To overcome the limitation of single tumor marker screening tests, diagnostic methods using combinations of multiple biomarkers can be used. IVDMIA, which combines multiple biomarkers and adds a unique analysis algorithm, is helpful for the diagnosis of cancer (11). The representative multiple biomarker set currently used as a diagnostic method in the clinical setting is Ova1 in ovarian cancer. In September 2009, the FDA approved a serum-based test called OVA1 (Vermillion, Inc., Austin, TX), combining 5 measured proteins (CA 125-II, TTR, ApoA1,B2M and transferrin) as an adjunct to clinical decision making for women planning surgery for an adnexal mass (19).

There is also a diagnostic antibody microarray platform in pancreatic cancer. This platform, consisting of 29 markers, was able to distinguish patients with stage I and II PDAC from controls with an ROC AUC value of 0.96 (20). However, due to the high cost, it has limited utility as a screening test. In Korea, the multi-marker panel (CA19-9, LRG1, and TTR) that has been developed and validated in large-scale cohorts by multiple reaction monitoring-mass spectrometry (MRM-MS) and immunoassay has clinical applicability in the early detection of PDAC. The triple-marker panel exceeded the diagnostic performance of CA19-9 alone by more than 10% in all PDAC samples and by more than 30% in patients with a normal range of CA19-

9. However, an automated system is still being established and has not yet been used in clinical practice (16). The candidate markers in the present study consisted of 11 markers used in the pan-cancer diagnostic kit which is commercially available in Korea. This cancer panel can be applied in real clinical practice, so commercialization can progress quickly. This cancer panel is already used for screening 7 cancers; hepatocellular carcinoma, breast cancer, lung cancer, gastric cancer, colon cancer, prostate cancer and ovarian cancer. The sensitivity, specificity and AUC of 7 cancers are 85~90%, 90~95% and 0.920~0.992. If pancreatic cancer is included, 8 cancers can be screened for 300 USD.

Serum CA19-9 is one of the most widely used serum tumor biomarkers for the detection of PDAC. Serum CEA and CA125 are 2 other biomarkers that are associated with the tumor burden of PDAC (21,22). These tumor markers could be applied not only to diagnosis but also to predicting prognosis and assessing treatment response. Xu et al. reported that the combination of postoperative serum CA19-9, CA125, and CEA served to determine a subgroup of patients benefiting from adjuvant chemo-radiotherapy (23). There have been several reports on the relationship between PDAC and Apo. Liu et al. discovered 5 biomarker combinations that can diagnose PDAC using the MS-based pipeline method, and 3 out of 5 contained Apo (ApoA1, ApoL1, and ApoE) (24). In the prospective evaluation, when using the combination of ApoA2 isoform (ApoA2-ATQ/AT) and CA19-9, the diagnostic rate was higher than that of CA19-9 alone (25). ApoA2 was reported to have an important role in the metastatic process in a study of serum-derived exosomes using iTRAQ-based proteomic analysis (26).

There are several reports on the diagnosis of cancers other than pancreatic cancer using the multiple biomarker panel in Korea. This panel was developed from a serum bank containing approximately 4500 samples from 5 types of cancer: breast, colon, stomach, liver, and lung. Kim et al. initially demonstrated the utility of the antibody-bead array approach in identifying

x

signatures specific for primary non-metastatic breast cancer with a high accuracy (91.8%) (12).

In non-small cell lung cancer patients, the highest accuracy of multivariate classification algorithms was observed when using the 5 highest-ranked biomarkers (alpha-1 antitrypsin [A1AT], CYFRA 21.1, insulin-like growth factor [IGF-1], regulated upon activation normal T cell expressed and secreted [RANTES], and alpha-fetoprotein [AFP]). In the validation set, the diagnostic accuracy was 88.2%–91.8% according to the analysis algorithm (14). These multiple panels were also validated in gastrointestinal tract cancer. Ahn et al. identified marker combinations of epidermal growth factor receptor (EGFR), TTR, RANTES, and vitronectin (VN) in gastric cancer with a diagnostic accuracy of 85.9%–89.2% (13).

The reason for comparison with single tumor markers is that the tumor markers used in clinical practice or screening of pancreatic cancer are CA19-9 and CEA. Therefore, we tried to show how the diagnostic rate is improved when a new combination of markers is used in actual clinical practice or screening system. Table 3 showed the results of the 4 and 5 marker combination models, the marker combination with the highest discrimination performance seems to outperform the selected panel. However, the reason we didn't select the 5 marker combinations was because they didn't fit our selection criteria. The criteria for selecting the optimal marker panel set by the authors was to select marker panel that showed excellent stability within the combination panel group showing excellent discrimination performance. Stability was evaluated in two ways: 1) The AUC difference between training and validation set should be small, and 2) The selected panel should show similar stability in 5 classification methods. For this reason, among the 4 and 5 combination marker models, the combination with the highest discrimination performance was judged to have lower stability than the panel, and therefore was not selected as the optimal marker panel.

In this study, we identified a new combination of protein markers that distinguish between PDAC samples and control samples. This panel has been shown to include markers that were previously not known to be related to pancreatic cancer and demonstrated improved classification performance compared to conventional cancer-specific markers. In addition to identifying new characteristics of previously unknown markers through statistical analysis, this study can be expanded to develop customized models for various purposes such as early diagnosis of pancreatic cancer or developing prognostic models, or to improve performance by developing complex marker models that combine protein biomarkers with new biomarkers such as DNA and RNA.

This study has limitations. The PDAC patient group contains patients who had surgery for pancreatic cancer. Although 29.5% of all patients were stage 3 or 4, this is a small fraction when considering the staging distribution in the diagnosis of PDAC patients. Additionally, the patients included in this study all received their operations in a single center. Therefore, a large-scale multicenter follow-up study is needed. The next issue concerned the normal control. We excluded patients with inflammatory diseases from the "normal" patient group. However, there were many patients with pancreatic disease, and tumor markers are frequently elevated in these patients. We regret not obtaining samples from patients with pancreatitis. To overcome this, a comparative study was conducted with a low grade IPMN samples. The last issue concerns efficiency in actual clinical application. Although there are already completed test platforms, measurement of the 6 markers is costly. Since the prevalence of PDAC is low in the general population, it is not cost-effective to use it as a routine check, which lessens its clinical utility. However, it can be useful in patients at increased risk of pancreatic cancer, such as those with family histories of pancreatic cancer, chronic pancreatitis, and new-onset diabetes.

This study demonstrated the utility of a combination of multiple biomarkers for the early detection of PDAC. Diagnostic biomarker panel algorithms that included six biomarkers (ApoA1, CA125, CA19-9, CEA, ApoA2, and TTR) were developed and validated. These algorithms will assist in the diagnosis of early pancreatic cancer, particularly stage 1 & 2 PDAC.

.

References

1.   Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. Cancer Res 2014,74(11), 2913-21.

2.   Saad AM, Turk T, Al-Husseini MJ, Abdel-Rahman O. Trends in pancreatic adenocarcinoma incidence and mortality in the United States in the last four decades; a SEER-based study. BMC Cancer 2018,18(1), 688.

3.   Gillen S, Schuster T, Zum Büschenfelde CM, Friess H, Kleeff J. Preoperative/neoadjuvant therapy in pancreatic cancer: a systematic review and meta-analysis of response and resection percentages. PLoS Med 2010, 7(4), e1000267.

4.   Issa IA, Noureddine M. Colorectal cancer screening: An updated review of the available options. World J Gastroenterol 2017, 23(28), 5086-96.

5.   Humphris JL, Chang DK, Johns AL, Scarlett CJ, Pajic M, Jones MD, et al. The prognostic and predictive value of serum CA19. 9 in pancreatic cancer. Ann Oncol 2012, 23(7), 1713-22.

6.   Goonetilleke K, Siriwardena A. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. Eur J Surg Oncol (EJSO) 2007, 33(3), 266-70.

7.   Rieser CJ, Zenati M, Hamad A, Al Abbas AI, Bahary N, Zureikat AH, et al. CA19-9 on postoperative surveillance in pancreatic ductal adenocarcinoma: predicting recurrence and changing prognosis over time. Ann Surg Oncol 2018, 25(12), 3483-91.

8.   Tsai S, George B, Wittmann D, Ritch PS, Krepline AN, Aldakkak M, et al. Importance of Normalization of CA19-9 Levels Following Neoadjuvant Therapy in Patients With Localized Pancreatic Cancer. Ann Surg 2020, 1(4), 740-47.

9.   Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012, 366(10), 883-92.

10.   Marusyk A, Almendro V, Polyak K. Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 2012,12(5), 323-34.

11.   Zhang Z. An in vitro diagnostic multivariate index assay (IVDMIA) for ovarian cancer: harvesting the power of multiple biomarkers. Rev Obstet Gynecol 2012, 5(1), 35-41.

12.   Kim BK, Lee JW, Park PJ, Shin YS, Lee WY, Lee KA, et al. The multiplex bead array approach to identifying serum biomarkers associated with breast cancer. Breast Cancer Res 2009,11(2), R22.

13.   Ahn HS, Shin Y, Park P, Kang K, Kim Y, Lee HJ, et al. Serum biomarker panels for the diagnosis of gastric adenocarcinoma. Br J Cancer 2012,106(4), 733-9.

14.   Lee HJ, Kim YT, Park PJ, Shin YS, Kang KN, Kim Y, et al. A novel detection method of non–small cell lung cancer using multiplexed bead-based serum biomarker profiling. J Thorac Cardiovasc Surg 2012,143(2), 421-7.

15.   Yoon HI, Kwon OR, Kang KN, Shin YS, Shin HS, Yeon EH, et al. Diagnostic value of combining tumor and inflammatory markers in lung cancer. J Cancer Prev 2016, 21(3), 187-193.

16.   Park J, Choi Y, Namkung J, Yi SG, Kim H, Yu J, et al. Diagnostic performance enhancement of pancreatic cancer using proteomic multimarker panel. Oncotarget 2017, 8(54), 93117-30.

17.   Ohtsuka T, Sato S, Kitajima Y, Tanaka M, Nakafusa Y, Miyazaki K. False-positive findings for tumor markers after curative gastrectomy for gastric cancer. Dig Dis Sci 2008, 53(1), 73-9.

18.   Bates SE, Longo DL. Tumor markers: value and limitations in the management of cancer patients. Cancer Treat Rev 1985, 12(3), 163-207.

19.   Fung ET. A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. Clin Chem 2010, 56(2), 327-9.

20.   Mellby LD, Nyberg AP, Johansen JS, Wingren C, Nordestgaard BG, Bojesen SE, et al. Serum biomarker signature-based liquid biopsy for diagnosis of early-stage pancreatic cancer. J Clin Oncol 2018, 36(28), 2887-94.

21.   Coppin L, Benomar K, Corfiotti F, Cattan S, Renaud F, Lapere C, et al. CA-125, but not galectin-3, complements CA 19-9 for discriminating ductal adenocarcinoma versus non-malignant pancreatic diseases. Pancreatology 2016, 16(1), 115-20.

22.   O'Brien DP, Sandanayake NS, Jenkinson C, Gentry-Maharaj A, Apostolidou S, Fourkala EO, et al. Serum CA19-9 is significantly upregulated up to 2 years before diagnosis with pancreatic cancer: implications for early disease detection. Clin Cancer Res 2015, 21(3), 622-31.

23.   Xu H-X, Li S, Wu C-T, Qi Z-H, Wang W-Q, Jin W, et al. Postoperative serum CA19-9, CEA and CA125 predicts the response to adjuvant chemoradiotherapy following radical resection in pancreatic adenocarcinoma. Pancreatology 2018, 18(6), 671-7.

24.   Liu X, Zheng W, Wang W, Shen H, Liu L, Lou W, et al. A new panel of pancreatic cancer biomarkers discovered using a mass spectrometry-based pipeline. Br J Cancer 2018, 118(6), e15.

25.   Honda K, Katzke VA, Hüsing A, Okaya S, Shoji H, Onidani K, et al. CA19-9 and apolipoprotein-A2 isoforms as detection markers for pancreatic cancer: a prospective evaluation. Int J Cancer 2019, 144(8), 1877-87.

26.   Tang P, Tao L, Yuan C, Zhang L, Xiu D. Serum Derived Exosomes From Pancreatic

Cancer Patients Promoted Metastasis: An iTRAQ-Based Proteomic Analysis. Onco Targets Ther 2019, 12, 9329-39.

Tables

Table 1. Demographics and clinicopathological characteristics of pancreatic ductal adenocarcinoma patients and healthy controls used in the training and validation sets

Pancreatic ductal adenocarcinoma

		Total N=180	Training set n=120	Validation set n= 60	*p*-value
Age	mean (range)	64.4 (28-87)	63.6 (28-82)	66.0 (37-87)	0.109
Sex	M	117 (65.0)	75 (62.5)	42 (70)	0.320
	F	63 (35.0)	45 (37.5)	18 (30)	
Operation	PPPD	62 (34.4)	44 (36.7)	18 (30)	0.996
	PD	38 (21.1)	25 (20.8)	13 (21.7)	
	DP	50 (27.8)	32 (26.7)	18 (30)	
	TP	15 (8.3)	11 (9.2)	4 (6.7)	
	Others[*]	15 (8.3)	8 (6.7)	7 (11.7)	
Stage	1	53 (29.4)	35 (29.2)	18 (30)	0.996
	2	74 (41.1)	50 (41.7)	24 (40)	
	3	30 (16.7)	20 (16.7)	10 (16.7)	
	4	23 (12.8)	15 (12.5)	8 (13.3)	
T stage	1	22 (12.2)	13 (10.8)	9 (15.0)	0.711
	2	94 (52.2)	66 (55.0)	28 (46.7)	
	3	38 (21.1)	26 (21.7)	12 (20.0)	
	4	10 (5.6)	6 (5.0)	4 (6.7)	
	NA	16 (8.9)	9 (7.5)	7 (11.7)	
N stage	0	69 (38.3)	47 (39.2)	22 (36.7)	0.779
	1	73 (40.6)	47 (39.2)	26 (43.3)	
	2	26 (14.4)	19 (15.8)	7 (11.7)	
	NA	12 (6.7)	7 (5.8)	5 (8.3)	
Differentiation	WD	14 (7.8)	9 (7.5)	5 (8.3)	0.862
	MD	118 (65.6)	80 (66.7)	38 (63.3)	
	PD	26 (14.4)	18 (15.0)	8 (13.3)	
	NA	22 (12.2)	13 (10.8)	9 (15.0)	

		Total		Training set	Validation set	p-value
Lymphatic	No	78 (43.3)		52 (43.3)	26 (43.3)	0.948
invasion	Yes	77 (42.8)		52 (43.3)	25 (41.7)	
	NA	25 (13.9)		16 (13.3)	9 (15.0)	
Venous	No	61 (33.9)		41 (34.2)	20 (33.3)	0.575
invasion	Yes	85 (47.2)		54 (45.0)	31 (51.7)	
	NA	34 (18.9)		25 (20.8)	9 (15.0)	
Perineural	No	19 (10.6)		14 (11.7)	5 (8.3)	0.547
invasion	Yes	145 (80.6)		97 (80.8)	48 (80.0)	
	NA	16 (8.9)		9 (7.5)	7 (11.7)	

Healthy control

		Total N=573		Training set n=382	Validation set n= 191	p-value
Age	mean (range)	56.9 (38-79)		56.6 (38-79)	57.5 (40-78)	0.250
Sex	M	334 (58.3)		218 (57.1)	116 (60.7)	0.420
	F	239 (41.7)		164 (42.9)	75 (39.3)	

PPPD, pylorus preserving pancreaticoduodenectomy; PD, pancreaticoduodenectomy; DP, distal pancreatectomy; TP, total pancreatectomy; NA, not available; WD, well differentiated; MD, moderate differentiated; PD, poorly differentiated; *Others, bypass surgery and open biopsy

Table 2. Diagnostic performance of new biomarker panel

	Training and test set				Validation set			
	AUC	Accuracy (%)	Specificity (%)	Sensitivity (%)	AUC	Accuracy (%)	Specificity (%)	Sensitivity (%)
RF								
Panel	0.992	95	95	96	0.993	95	96	93
CA19-9 + CEA	0.924	90	95	76	0.960	92	94	83
CA19-9	0.921	90	95	74	0.960	90	94	78
CEA	0.666	77	95	20	0.797	78	95	25
GLM								
Panel	0.983	94	95	92	0.983	94	95	92
CA19-9 + CEA	0.852	87	95	62	0.928	91	94	80
CA19-9	0.848	88	95	66	0.923	92	94	83
CEA	0.732	78	95	24	0.814	80	96	28
GLM + RF								
Panel	0.984	94	95	92	0.986	95	96	92
CA19-9 + CEA	0.934	91	95	78	0.962	91	93	87
CA19-9	0.933	90	95	75	0.964	90	94	80
CEA	0.732	78	95	24	0.814	80	96	28
RIDGE								
Panel	0.987	95	95	93	0.985	95	96	92
CA19-9 + CEA	0.852	87	95	62	0.928	91	94	80
CA19-9	0.848	88	95	67	0.924	92	94	83
CEA	0.732	78	95	24	0.816	80	96	28
SVM								
Panel	0.990	95	95	95	0.991	97	98	92
CA19-9 + CEA	0.900	89	95	71	0.964	92	97	77
CA19-9	0.912	88	95	68	0.967	92	96	77
CEA	0.627	78	95	25	0.692	78	95	27

AUC, area under the curve; RF, random forest; CA19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; GLM, generalized linear model; SVM, support vector machine

Table 3. Result of comparing 2 to 5 markers combinations with 6 marker panels among markers

	Training and test set				Validation set			
	AUC	Accuracy (%)	Specificity (%)	Sensitivity (%)	AUC	Accuracy (%)	Specificity (%)	Sensitivity (%)
Panel vs Combination of 2 markers (RF)								
Panel	0.992	95	95	96	0.993	95	96	93
Max. CA19-9 + TTR	0.989	95	95	94	0.993	96	96	97
Min. B2M + CEA	0.686	76	95	18	0.8	81	97	30
Panel vs Combination of 3 markers (RF)								
Panel	0.992	95	95	96	0.993	95	96	93
Max. CA19-9 + TTR + D-Dimer	0.991	95	95	95	0.991	96	95	98
Min. CA125 + B2M + CEA	0.726	79	95	29	0.815	80	93	40
Panel vs Combination of 4 markers (RF)								
Panel	0.992	95	95	96	0.993	95	96	93
Max. CA19-9 + TTR + CRP + ApoA2	0.993	95	95	94	0.997	96	96	97
Min. CA125 + B2M + CEA + LRG1	0.774	80	95	33	0.841	81	95	37
Panel vs Combination of 4 markers (GLM)								
Panel	0.983	94	95	92	0.983	94	95	92
CA19-9 + TTR + CRP + ApoA2	0.982	95	95	94	0.986	95	96	93
Panel vs Combination of 4 markers (GLM + RF)								
Panel	0.984	94	95	92	0.986	95	96	92
CA19-9 + TTR + CRP + ApoA2	0.986	95	95	93	0.995	97	97	95
Panel vs Combination of 4 markers (RIDGE)								
Panel	0.987	95	95	93	0.985	95	96	92
CA19-9 + TTR + CRP + ApoA2	0.986	95	95	94	0.986	95	96	93

Panel vs Combination of 4 markers (SVM)							
Panel	0.990	95	95	95	0.991	97	98	92
CA19-9 + TTR + CRP + ApoA2	0.989	95	95	94	0.992	97	97	95
Panel vs Combination of 5 markers (RF)							
Panel	0.992	95	95	96	0.993	95	96	93
Max. CA19-9 + TTR + CEA + ApoA2 + CRP	0.993	95	95	95	0.997	96	95	97
Min. CA125 + B2M + CEA + LRG1 + CYFRA21.1	0.845	82	95	41	0.873	94	98	40
Panel vs Combination of 5 markers (GLM)							
Panel	0.983	94	95	92	0.983	94	95	92
CA19-9 + TTR + CEA + ApoA2 +CRP	0.981	95	95	94	0.987	95	95	93
Panel vs Combination of 5 markers (GLM + RF)							
Panel	0.984	94	95	92	0.986	95	96	92
CA19-9 + TTR + CEA + ApoA2 +CRP	0.984	95	95	94	0.997	97	97	97
Panel vs Combination of 5 markers (RIDGE)							
Panel	0.987	95	95	93	0.985	95	96	92
CA19-9 + TTR + CEA + ApoA2 +CRP	0.986	95	95	93	0.987	95	96	93
Panel vs Combination of 5 markers (SVM)							
Panel	0.990	95	95	95	0.991	97	98	92
CA19-9 + TTR + CEA + ApoA2 +CRP	0.988	95	95	95	0.992	96	98	92

AUC, area under the curve; RF, random forest; CA19-9, carbohydrate antigen 19-9; TTR, transthyretin; CEA, carcinoembryonic antigen; B2M, beta-2 microglobulin; CA125, cancer antigen 125; ApoA2, Apolipoprotein A2; CRP, C-reactive protein; LRG1, leucine rich alpha-2-glycoprotein 1; GLM, generalized linear model; SVM, support vector machine; CYRFA12.1, cytokeratin 19 fragment 21.1

Figures

Figure 1. Study schematic flow chart. The marker panels used in the model generation consist of 11 candidate biomarkers and 2,047 combinations. After combining age and gender variables to each panel, the combination is applied to the five classification algorithms. Selection criteria for optimal biomarker combinations are as follows; (1) the top 10% out of the 2047 sets, (2) sets containing CEA and CA19-9, (3) minimal difference between training set and validation set, (4) excellent performance independent of the linear and nonlinear methods

Figure content (flowchart labels):

180 samples from pancreatic cancer patients (PC)
573 samples from normal control (NL)

Candidate 19 markers
from 6 cancer patients
and normal control

Training set
120 PC
382 NL

Validation set
60 PC
191 NL

+10-Fold CV

Marker selection   Model development   Validation

GLM, RIDGE, SVM, RF, GLM+RF

11 marker
+Age +Sex

Selection criteria
•Top 10% set
•Including CEA, CA19-9
•Minimum difference between training set and validation set
•Excellent accuracy in 5 statistical models

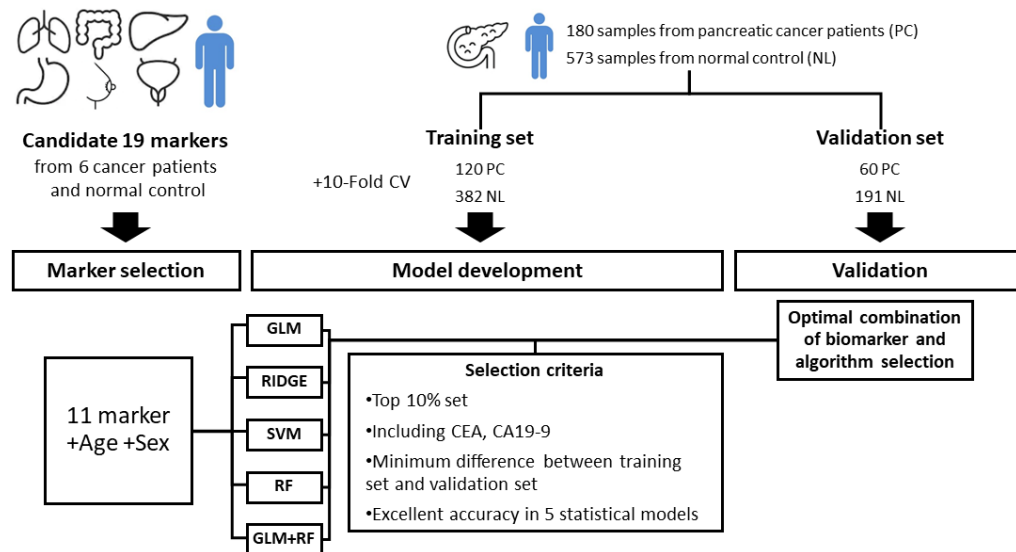Optimal combination of biomarker and algorithm selection

Figure 2. Comparison of 11 candidate markers concentration between PDAC samples and normal control samples. Among the 11 markers (ApoA1, CA125, CA19-9, CRP, CYFRA21.1, LRG1, CEA, ApoA2, TTR, B2M, and D.Dimer), all except B2M showed statistical differences between PDAC and normal controls.
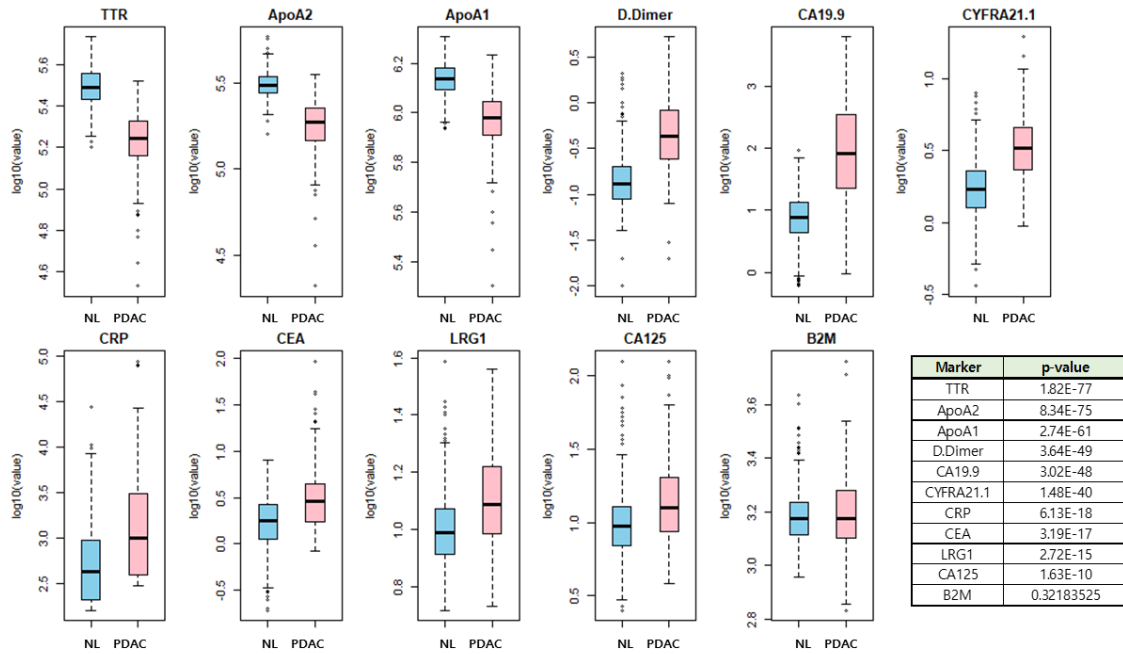

Marker	p-value
TTR	1.82E-77
ApoA2	8.34E-75
ApoA1	2.74E-61
D.Dimer	3.64E-49
CA19.9	3.02E-48
CYFRA21.1	1.48E-40
CRP	6.13E-18
CEA	3.19E-17
LRG1	2.72E-15
CA125	1.63E-10
B2M	0.32183525

Figure 3. Diagnostic performance of new model in training set and validation set.The AUC was 0.992 in the training set and 0.993 in the validation set. Comparing CA19-9, CEA and the combination of CA19-9 and CEA, the AUC of the new model was 0.993 in the validation set, and the AUC of CEA+ CA19-9 was 0.960 and CA19-9 alone was 0.960.
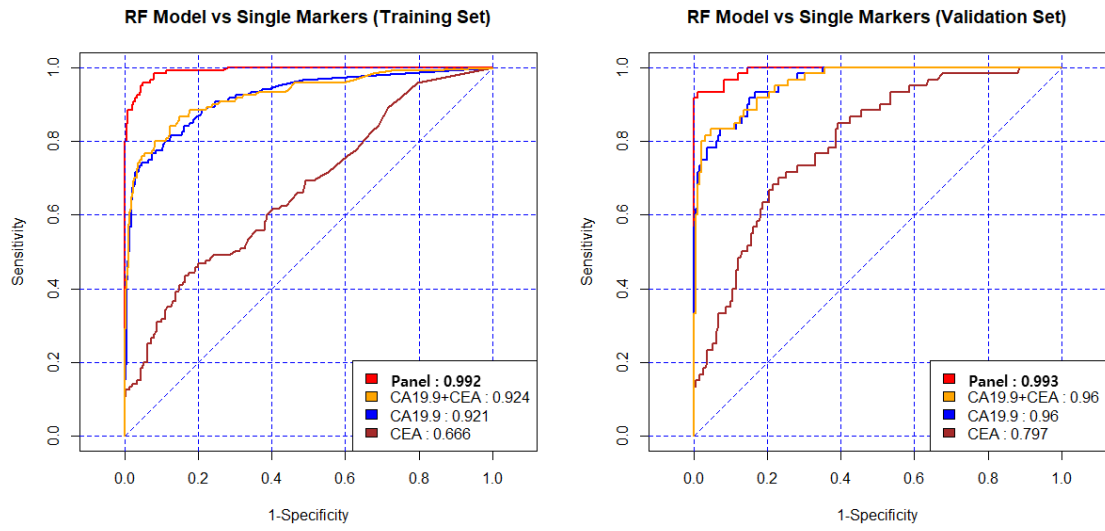
Figure 4. Diagnostic performance according to the PDAC stage. In the validation set, the diagnostic accuracy according to the cancer stage was 89%, 92%, 100%, and 100 % in stage 1, 2, 3, and 4, respectively. Especially, in stages 1 and 2, the new model improved the diagnostic ability compared to CA19-9 alone.
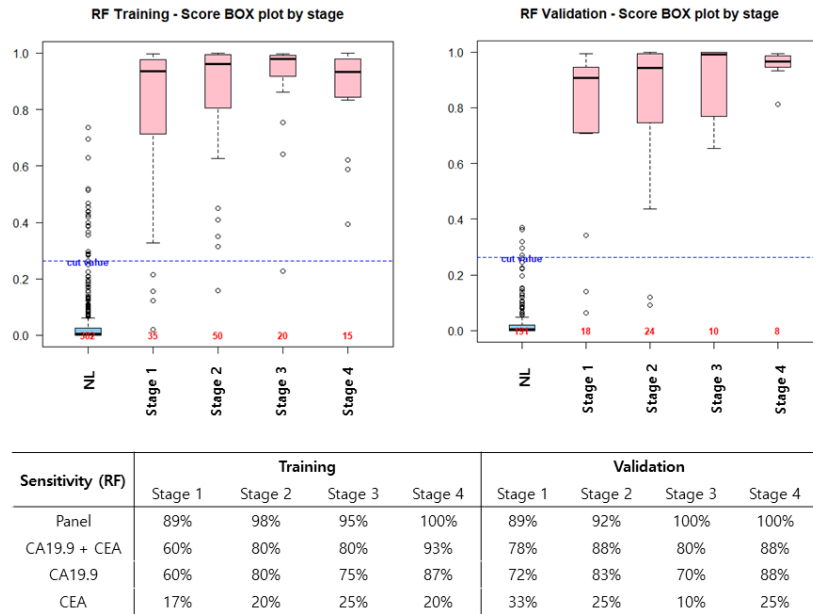

Sensitivity (RF)	Training				Validation			
	Stage 1	Stage 2	Stage 3	Stage 4	Stage 1	Stage 2	Stage 3	Stage 4
Panel	89%	98%	95%	100%	89%	92%	100%	100%
CA19.9 + CEA	60%	80%	80%	93%	78%	88%	80%	88%
CA19.9	60%	80%	75%	87%	72%	83%	70%	88%
CEA	17%	20%	25%	20%	33%	25%	10%	25%

Figure 5. Discriminate between PDAC and benign pancreatic disease.

The concentrations of 6 markers measured in the IPMN sample were between the normal controls and the PDAC group. The AUC, specificity and sensitivity were 0.9638, 91.1% and 81.4% in the training set, and 0.9403, 91.7% and 86.7% in the test set.
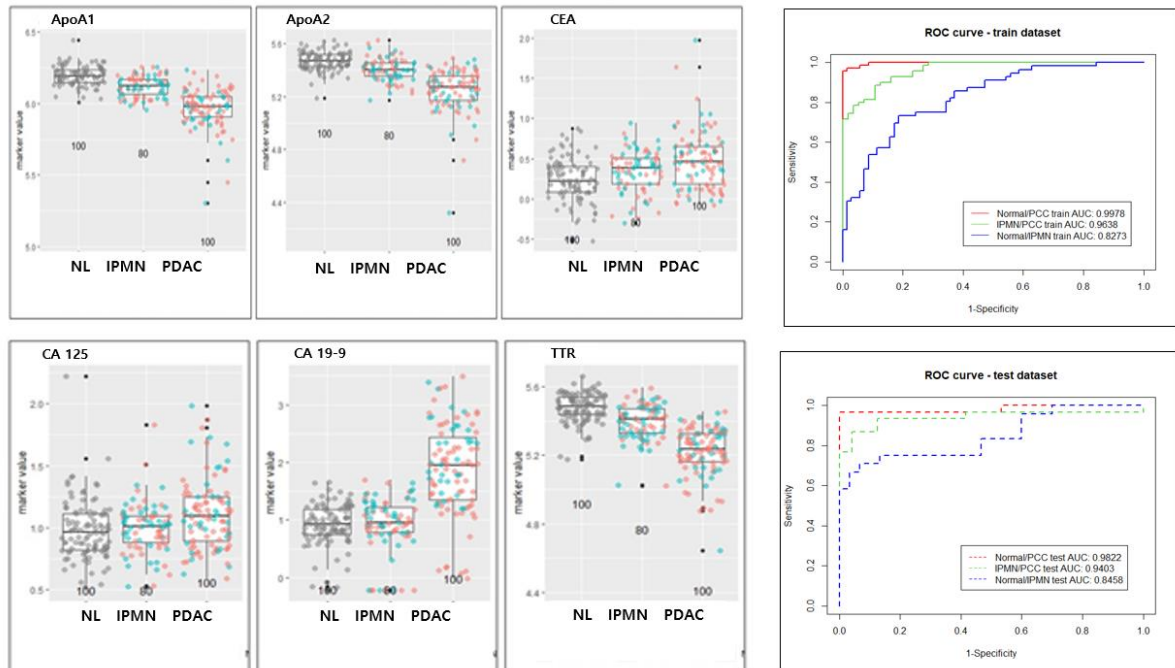
국문초록

췌장암 조기 진단을 위한 복합 바이오마커 개발

서울대학교 대학원

의학과 외과학 전공

김 홍 범

배경: 췌장암에서 단일 종양 표지자는 진단 가치가 높지 않다. 이러한 한계는 여러개의 바이오 마커 조합을 고유 분석 알고리즘에 적용 하여 진단 능력을 향상시킬 수 있다. 본 연구는 췌관선암에 진단에서 다중 바이오 마커 조합과 진단 알고리즘을 개발하고 성능을 검증하는 것이다.

방법: 본 연구는 180 명의 췌관선암 환자와 573 명의 건강한 대조군의 혈액 샘플을 사용하였다. 췌장 양성질환과 감별하기 위하여 췌관내유두상종양 샘플을 사용하였다. 후보 마커는 다양한 암에서 일반적으로 발현되고 췌장암에서도 발현이

1

입증 된 11 개의 마커로 구성하였다. 샘플은 개발, 훈련 세트 및 검증 세트로 구분하였다. 최적의 모델을 결정하기 위해 다섯 가지 선형 또는 비선형 분류 방법을 사용하였다.

결과: 11 개 마커 중 10 개의 마커가 대조군과 췌관선암 사이에서 유의한 차이를 보였다. 5 개의 개별 알고리즘에 모두 적용되는 2047 개의 조합을 확인하였고 6 개로 (ApoA1, CA125, CA19-9, CEA, ApoA2 및 TTR) 구성된 새로운 바이오 마커 조합을 찾았다. 곡선 아래 영역 (Area under the curve; AUC), 특이도 및 민감도는 훈련 세트에서 각각 0.992, 95 % 및 96 % 였고 검증 세트에서 0.993, 96 % 및 93 %였다. 췌관내유두상 종양과 비교하였을 때 AUC는 0.964였고, 특이도 및 민감도는 91.1%, 81.4%였다.

결론: 본 연구를 통하여 췌관선암의 진단에서 6개의 바이오 마커로 구성된 진단 알고리즘이 개발되고 검증 되었다. 이러한 알고리즘은 췌관선암의 조기 진단에 도움이 될 것이다.