Published in final edited form as:

Ann Surg. 2018 August ; 268(2): 340–347. doi:10.1097/SLA.0000000000002421.

A Multi-Institutional Validation Study of Pancreatic Cyst Fluid Protein Analysis for Prediction of High-Risk Intraductal Papillary Mucinous Neoplasms (IPMN) of the Pancreas

Mohammad A. Al Efshat, MD¹, Marc F. Attiyeh, MD¹, Anne A. Eaton, MS², Mithat Gönen, PhD², Denise Prosser, BS³, Anna E. Lokshin, PhD³, Carlos Fernández-del Castillo, MD⁴, Keith D. Lillemoe, MD⁴, Cristina R. Ferrone, MD⁴, Ilaria Pergolini, MD⁴, Mari Mino-Kenudson, MD⁵, Neda Rezaee, MD⁶, Marco Dal Molin, MD⁶, Matthew J. Weiss, MD⁶, John L. Cameron, MD⁶, Ralph H. Hruban, MD⁷, Michael I. D'Angelica, MD¹, T. Peter Kingham, MD¹, Ronald P. DeMatteo, MD¹, William R. Jarnagin, MD¹, Christopher L. Wolfgang, MD, PhD⁶, and Peter J. Allen, MD¹

¹Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

²Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

³Department of Pathology, University of Pittsburgh Cancer Institute, Pittsburgh, PA 15261, USA

⁴Department of Surgery, Massachusetts General Hospital, Boston, MA 02114, USA

⁵Department of Pathology, Massachusetts General Hospital, Boston, MA 02114, USA

⁶Department of Surgery, Johns Hopkins Hospital, Baltimore, MD 21287, USA

⁷Sol Goldman Pancreatic Cancer Research Center, Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA

Corresponding Author: Peter J. Allen, M.D., Professor of Surgery, Memorial Sloan-Kettering Cancer Center 1275 York Ave- Office C896, New York, NY 10021, Phone: 212-639-5132, Facsimile: 212-717-3645, allenp@mskcc.org.

Conflicts of Interest: All authors declare no potential conflicts of interest.

Relevant Notes: This work has been presented as an oral presentation at the 2016 American College of Surgery (ACS) Clinical Congress, Washington, DC, October 16–20, 2016.

Authors' contributions are as follows:

- **Conception and Design:** Al Efshat, Attiyeh, Fernández-del Castillo, Gönen, Lokshin, Lillemoe, Weiss, Cameron, Hruban, D'Angelica, Kingham, DeMatteo, Jarnagin, Wolfgang, Allen.
- **Development of Methodology:** Al Efshat, Gonen, Fernández-del Castillo, Lokshin, Wolfgang, Allen.
- **Acquisition of Data:** Al Efshat, Attiyeh, Pergolini, Rezaee, Dal Molin, Mino-Kenudson, Prosser.
- **Analysis and Interpretation of Data:** Al Efshat, Attiyeh, Gonen, Eaton, Allen.
- **Drafting of Manuscript:** Al Efshat, Allen.
- **Review and Critical Revision of the Manuscript:** Al Efshat, Attiyeh, Eaton, Gonen, Prosser, Lokshin, Lomakin, Fernández-del Castillo, Lillemoe, Ferrone, Mino-Kenudson, Pergolini, Rezaee, Dal Molin, Weiss, Cameron, Hruban, D'Angelica, Kingham, DeMatteo, Jarnagin, Wolfgang, Allen.
- **Administrative, Technical, and Material Support:** Al Efshat, Attiyeh, Pergolini, Rezaee, Ferrone, Dal Molin, Mino-Kenudson, Prosser, Lomakin, Allen.
- **Study Supervision:** Allen.

Abstract

OBJECTIVE—Preliminary work by our group suggested that proteins within the pancreatic cyst fluid (CF) may discriminate degree of IPMN dysplasia. We sought to externally validate these markers, and determine if their inclusion in a pre-operative clinical nomogram could increase diagnostic accuracy.

SUMMARY BACKGROUND DATA—IPMN is the most common radiographically identifiable precursor to pancreatic cancer, however, the timing and frequency of its malignant progression is unknown, and there are currently no reliable pre-operative tests that can determine the grade of dysplasia in IPMN.

METHODS—Clinical and radiographic data, as well as CF samples, were obtained from 149 patients who underwent resection for IPMN at one of three institutions. High-risk disease was defined as the presence of high-grade dysplasia or invasive carcinoma. Multianalyte bead array analysis (Luminex) of CF was performed for four protein markers that were previously associated with high-risk disease. Logistic regression models were fit on training data, with and without adjustment for a previously developed clinical nomogram, and validated with an external testing set. The models incorporating clinical risk score were presented graphically as nomograms.

RESULTS—Within the group of 149 resected patients, 89 (60%) had low-risk disease, and 60 (40%) had high-risk disease. All four CF markers (MMP9, CA72-4, sFASL and IL-4) were over-expressed in patients with high-risk IPMN ($p < 0.05$). Two predictive models based on pre-selected combinations of CF markers had concordance indices of 0.76 (model-1) and 0.80 (model-2). Integration of each CF marker model into a previously described clinical nomogram lead to increased discrimination compared to either the CF models or nomogram alone (c-indices of 0.84 and 0.83, respectively).

CONCLUSIONS—This multi-institutional study validated two CF protein marker models for pre-operative identification of high-risk IPMN. When combined with a clinical nomogram, the ability to predict high-grade dysplasia was even stronger.

Keywords

Intraductal papillary mucinous neoplasms; Pancreas; Cyst fluid; Dysplasia; Biomarkers

INTRODUCTION

Intraductal papillary mucinous neoplasms (IPMN) of the pancreas are a heterogeneous group of lesions that grow within the ductal system of the pancreas and may involve the main pancreatic duct (MD-IPMN), branch ducts (BD-IPMN), or both (Mixed-IPMN).[1] Since the recognition of this entity over three decades ago [2][3], IPMN have gained increased attention because of their frequent identification on routine cross-sectional imaging, and because they are the most common radiographically identifiable precursors of pancreatic cancer. [4, 5]

The pathway of IPMN progression from non-invasive to invasive disease is believed to be responsible for approximately 20%–30% of pancreatic cancer cases.[5] These cystic adenomatous lesions are believed to progress from low-grade dysplasia, to high-grade

dysplasia, to invasive cancer.[6] Once invasive disease develops, survival outcomes are similar to conventional pancreatic cancer when controlled for stage.[7, 8] Therefore, most clinicians believe that resection of high-grade dysplasia presents an opportunity to cure a lesion prior to the development of an incurable invasive disease. However, previous recommendations for the routine resection of IPMN have led to some overtreatment, with many reports identifying high-risk pathology in less than half of all patients who have undergone pancreatectomy.[9, 10]

The ability of current laboratory, radiographic, and endoscopic tests to distinguish between low-risk (low and moderate dysplasia) and high-risk (high-grade dysplasia and invasive) IPMN is limited.[11, 12] The presence of a dilated main pancreatic duct (MPD) on preoperative imaging remains the most commonly implemented criterion for prediction of high-grade dysplasia or invasive disease, as it has been shown that approximately 60% of patients with resected MD-IPMN harbor high-grade dysplasia or invasive disease. [13] When the main duct is not dilated (BD-IPMN), only 10–15% of resected patients will be found to have high-grade disease [14]. The most recent consensus guidelines for management of IPMN have recommended resection for all IPMN patients with a dilated MPD>1cm or an enhancing solid component.[15] Reports have shown however, that in this setting, 60% of resected patients will have high-risk disease. If only 60% of these patients have high-risk disease, then 40% of patients will undergo a major operation for low-risk disease.[16] Hence, improved markers of high-risk disease are needed.

Over the past several years, our group has investigated several serum and cyst fluid protein markers in patients who have undergone resection for IPMN. These studies have demonstrated consistent differences in the patterns of cyst fluid protein expression between low-risk and high-risk patients.[17, 18] In the most recent study [19], we analyzed the cyst fluid of 78 patients with resected IPMN utilizing antibody bead array (Luminex) with 87 different protein markers. High-risk disease was found to have a positive correlation with 29 of the measured proteins, of which 23 (79%) were inflammatory markers. In this previous study, none of the pro-inflammatory or neoplastic markers were found to be significantly overexpressed in the low-risk group. Multivariate modeling was performed on these data, and resulted in two different mathematical prediction models with two different sets of markers (MMP9 + CA72-4, and sFASL + IL-4) that each had a discrimination index (c-index) of 86% for predicting high-risk disease. In these 78 patients, if MPD dilation was used as the criterion for high-risk disease (as recommended by the current guidelines) then the c-index would have been only 73%.

In this current study, we sought to validate, on an independent multi-institutional patient set, these two predictive models using antibody bead array (Luminex). We employ a larger set of cyst fluid samples from three high-volume academic hospitals (Johns Hopkins Hospital, Massachusetts General Hospital, Memorial Sloan Kettering). We also aimed to incorporate these protein-level models into a recently developed clinical and radiographic prediction model (nomogram) to investigate whether a combined model could result in a highly accurate prediction tool for high-risk IPMN.

METHODS

Patients

Following IRB approval, and using prospectively maintained pancreatic databases from the Pancreatic Surgery Consortium, each one of the three participating institutions (Memorial Sloan Kettering (MSK), Massachusetts General Hospital (MGH) and Johns Hopkins Hospital (JHH)) randomly selected 50 patients who had undergone resection for a pathologically-proven IPMN, between January 2004 and September 2015. Eligible patients were required to have had adequate banked cyst fluid (at least 250µl) available for the study. Patients with concurrent malignancies (e.g., cholangiocarcinoma, neuroendocrine tumor) were excluded.

All patients had previously signed one of the corresponding institutional review board-approved tissue banking protocols (MSK IRB# 00-032, JHH IRB# AM00028936, MGH IRB# 2003p001289), and a waiver of authorization was obtained from each institutional IRB prior to accessing the electronic medical records. Demographic data, clinical data, laboratory, radiographic and pathologic features were extracted from the clinical databases. Main duct IPMN (MD-IPMN) was recorded if the main pancreatic duct (MPD) was dilated (>0.5cm) on radiographic evaluation or main-duct involvement was reported on pathologic assessment, branch duct (BD-IPMN) was defined if there was cystic disease in the absence of main duct dilation (< 0.5cm) or involvement. A diagnosis of mixed-type (a combination of both main duct dilation and cystic disease) was classified as main duct disease for the purposed of this study.

Histopathologic assessment of resected specimens was performed as per institutional protocol by a dedicated gastrointestinal pathologist. Grade of dysplasia was defined as the highest degree of dysplasia identified within the examined specimen.[6] Based on the outcome of interest, patients were classified as “high-risk” if their pathology report showed high-grade dysplasia, or invasive carcinoma. Lesions with low-grade or intermediate-grade dysplasia were defined as “low-risk”.

Cyst Fluid Samples

Cyst fluid samples were aspirated with an 18 – 21-gauge needle either intraoperatively by the surgeon or upon arrival to the surgical pathology suite by a pathologist. Samples were then aliquoted, and stored at –80°C. Based on the recorded time between resection and freezing, only samples that had been refrigerated within 60 minutes of resection and underwent no prior freeze-thaw cycles were eligible. All the samples, de-identified to study number only, were shipped overnight on dry ice to University of Pittsburgh Cancer Institute (UPCI) Luminex core facility, where all multiplex assays were performed.

Multiplex Biomarker Analysis

Multianalyte analysis was performed using commercially available plates that included the four pre-specified proteins.[19] Luminex Multiplex Bead Immunoassays were performed in 96-well microplate format as per standard protocol (See supplementary files SF1 and SF2 for the detailed steps and protocols of all assays).[19–21] Briefly, dilutions were followed

according to each Vendor's protocol recommendations, unless otherwise noted. Millipore research human cytokine/chemokine Panel 1 (5-Plex) (Catalog # HCYTOMAG-60K-05) was diluted two-fold in order to increase bead recovery. Additional Millipore assay conducted was Milli Research Human Cancer Biomarker Panel 1 (1-Plex) (Catalog # HCCBP1MAG-58K-01) at 6-fold. R&D systems assays conducted were LMPM000 and LMPM911 at 50-fold. And finally, the UPCI Core-developed assay conducted (5plex) at 5-fold. (See Supplementary table ST1 for a summary of these multiplex assays and dilutions). Quality control data for each core-developed assay, including correlation with commercial ELISA, can be found on the UPCI Luminex Core Facility website (<http://upci.upmc.edu/cpf/luminex.cfm>).

Upon completion, all Luminex assay plates were then read on a BioRad Bioplex[®] 100 or 200 instrument utilizing BioPlex Manager[®] 4.1.1. Quantification of markers was either exported directly from this software or via XML files analyzed by the scaler program. All the resulting concentrations were normalized according to a scaling procedure, previously developed by UPCI group, to account for variations across different experiments (batch effect).[20]

Statistical Analysis

Our sample size was selected based on the precision with which we could estimate the sensitivity for the grade of dysplasia, and thus we did not use traditional sample size paradigms. With a sample size of 150, we could estimate sensitivity using a 95% confidence interval to within $\pm 9.2\%$ if true sensitivity was 80%. The precision would increase to $\pm 6.2\%$ if true sensitivity was 90%. Of note, we chose sensitivity over specificity because the clinical emphasis was felt to be weighted towards avoiding non-operative management of high-risk lesions that should ideally be resected to prevent progression to cancer. This translates into a need to decrease false negative results and improve both negative predictive value and sensitivity.

Protein concentrations were log-transformed before analysis to produce more normally-distributed data. Continuous variables, including protein concentrations, were summarized using median and range and compared between the low-risk and high-risk groups using the Mann-Whitney test. Categorical variables were summarized using frequency and percentage and compared using Fisher's exact test. Dependence among proteins was evaluated using Pearson's correlation coefficient, accepting correlation coefficients of less than 0.7 to represent independence between the selected markers.

The data were randomly split into a training (n=104) and testing dataset (n=45). Multivariable logistic regression was used to assess relationships between pre-selected combinations of protein markers and high-risk disease in the training dataset. The markers to be included in the models were selected based on multivariate modeling from our previous work [19]. Model-1 was based on MMP9 and CA72-4 log concentrations, while model-2 was a three-leaf decision tree based on sFASL, IL-4, and the IPMN duct type (decision tree classification details are shown in Table 3). For model-2, sFASL and IL-4 were expressed as percentiles in their respective datasets.

Predictions from our clinical nomogram were added as a predictor in each model to determine if the markers remained independently associated with outcome.[22] Model performance was assessed on the test dataset using concordance indices (c-indices), with values ranging from 0.5 (as good as chance) to 1.0 (perfect discrimination). The final multivariable models were visually represented as nomograms and validated using the testing dataset. All statistical analysis was done in R 3.1.1 using the *rms*, *Hmisc*, *pROC*, and *readxl* packages, and *p*-values less than 0.05 were considered significant.

RESULTS

Patient Characteristics

In total, 154 patients who underwent resection of IPMN, and had cyst fluid banked per protocol, were initially included from the three participating institutions (MSK, MGH and JHH). However, during the multiplex assays, five samples were found to be too viscous for the assay (all five samples were high-risk), and thus had to be excluded due to inadequate protein measurement. The remaining 149 patients constituted our study cohort. (Figure 1)

The clinical, pathologic and radiographic characteristics of the cohort (stratified by disease risk) are summarized in Table 1. High-risk disease was identified in 60 patients (40%), and 89 (60%) had low-risk disease (low and intermediate-grade dysplasia). Median age at resection was similar in both groups (71 years). Males represented 58% (n=86) of the cohort, and constituted 72% of the high-risk group. Female patients were significantly more likely to harbor low-risk disease. BD-IPMN was present in 87 patients (58%) and the majority of BD-IPMN patients (n=72, 83%) had low-risk disease. Within the group of 62 patients (42%) with MD-IPMN, the majority (n=45, 73%) had high-risk disease on final pathological analysis. Main-duct IPMN was significantly associated with high-risk disease ($p<.001$).

Weight loss, jaundice, a history of smoking, an abrupt change in MPD diameter and the presence of solid component were more common in the high-risk group ($p = 0.05$). The training dataset (n=104) had 65 low-risk and 39 high-risk patients (63% high-risk), while the validation dataset (n=45) had 21 high-risk and 24 low-risk patients (47% high-risk).

Multiplex Assays and Univariate Analysis of Cyst Fluid Biomarker Levels between the High-risk and Low-risk Groups

Correlation between the four markers that were included in the predictive models (MMP9, CA72-4, sFASL and IL-4) was low to moderate, with correlation coefficients of 0.25–0.68 (supplementary table ST2). Univariate analysis showed differential expression ($P = 0.05$) of all four protein markers between the high-risk and low-risk groups (Table 2, Figure 2).

We then performed further analysis of each IPMN duct subtype (MD-IPMN and BD-IPMN) separately to see whether these biomarkers had different associations with high-risk disease in MD-IPMN and BD-IPMN. However, given the relatively small number in each subgroup, and the known low power of interaction tests, these tests were not precise, and the ultimate decision not to stratify the cohort based on duct type was based on these considerations.

Validation of Predictive Cyst Fluid Biologic Models

The two pre-selected multivariate models, maintained high predictive ability in this multi-institutional cohort. Model-1, which included the log concentration of cyst fluid MMP9 and CA-72-4 as continuous predictors, had a c-index of 0.76 in the training set and 0.80 in the validation (testing) set. Model-2 had a c-index of 0.80 in the training set and 0.79 in the validation set. Odds ratios and p-values for the different components in each model are summarized in Table 3.

Cyst fluid protein markers remained independently associated with the degree of dysplasia when predictions from our previous clinical nomogram (c-index of 0.80 in the training set and 0.77 in the validation set) were added as a predictor in each model (all p-value<0.05). [22] Furthermore, diagnostic accuracy was further improved when the two cyst fluid models were merged with the clinical nomogram compared to either of the cyst fluid models or the nomogram alone (Table 3). Combining model-1 with the clinical nomogram (tested by receiver operating characteristic curve analysis) yielded the highest discriminative potential in both the training and validation set with c-indices of 0.88 and 0.84, respectively. Similarly, adding the clinical nomogram to model-2 resulted in an improved predictive performance in the training and validations groups (c-indices 0.82 and 0.83, respectively). The combined clinical and protein-level nomograms are visually presented in Figure 3.

DISCUSSION

The distinction between high-risk and low-risk IPMN is of utmost importance; as resection of IPMN in the setting of high-risk disease is considered indicated and appropriate. Similarly, if low-risk disease could be reliably identified, these patients would be radiographically monitored, and could avoid the morbidity of operation until high-risk disease developed.[23] In the current study, we developed and independently validated two separate cyst fluid biomarker models of dysplasia from 149 patients, who underwent resection at three different institutions, using antibody bead array (Luminex). Additionally, we demonstrated that incorporation of these two predictive cyst fluid marker models into a recently developed clinical and radiographic nomogram resulted in an even higher diagnostic discrimination for high-risk IPMN.[22] We believe that this is the first study to demonstrate the benefit of combining clinical, radiographic, and cyst fluid protein expression in defining high-risk IPMN.

Because of the limitations of imaging and cytology in defining high-risk IPMN, the diagnostic utility of cyst fluid analysis has been extensively evaluated. [11] A variety of protein markers including CEA, CA19-9, CA15-3, M1 mucin and amylase; as well as DNA and miRNA markers have been studied as potential diagnostic markers of pancreatic cyst sub-type. Although some of these markers have shown the ability to identify mucinous lesions and differentiation of histopathologic sub-type, their ability to discriminate grade of dysplasia in IPMN has not been encouraging.[24–29] For example, cyst fluid CEA level has been consistently shown to be a useful predictor for identification of mucinous lesions with up to 79% accuracy and 85% accuracy when combined with the presence of extracellular mucin.[28, 30] However, the degree of elevation of cyst fluid CEA has not been found to be predictive of the grade of dysplasia in these patients.[31–33] In a recent retrospective study

from Japan, multivariate analysis showed a weak association between CEA level in the pancreatic fluid and invasive carcinoma in mixed and MD-IPMN (OR, 1.002; 95% CI, 1.000–1.003; $P = .048$). The area under the curve (AUC) value for pancreatic fluid CEA level in mixed IPMN was 0.796 ($P < .001$), while the AUC for pancreatic fluid CEA MD-IPMN was 0.877 ($P < .001$). [34] Multiple other studies developed clinical and radiographic nomograms to predict invasive disease in MD-IPMN and BD-IPMN with variable diagnostic accuracy. [35–37]

Our group has focused on cyst fluid protein expression as a tool for discrimination of grade of dysplasia in IPMN. Previous experiments have demonstrated consistent patterns of protein expression between low-risk and high-risk lesions. [17–19] For example, we previously evaluated the cyst fluid for inflammatory markers as a surrogate for tumor-associated neutrophils (TAN), as tissue studies from our group have identified an increased number of tumor associated neutrophils (TAN) in high-risk lesions. [38] These initial evaluations identified overexpression of inflammatory markers such as IL-1b, IL-5, and IL-8 in high-risk lesions; cyst fluid IL-1b remained a significant predictor of high-risk disease on multivariate analysis that included IPMN subtype; and cyst fluid inflammatory markers were associated with the degree of TAN. [17] High-mobility group (HMG) A2 protein was also found to be significantly higher in cyst fluid of high-grade compared with low-grade IPMN. [39] Other groups found an association between neutrophil-to-lymphocyte ratio and risk of invasive disease in IPMN. [40–42]

More recently [19], we evaluated the association of TAN with malignant progression in 78 patients with resected IPMN, and performed a multiplexed assay for 87 different cyst fluid proteins, including cyst fluid inflammatory markers (CFIM), as possible surrogate markers for parenchymal inflammation. The majority (96%) of the low-risk lesions demonstrated no TAN, whereas 89% of invasive lesions expressed high levels of TAN. The grade of dysplasia was also found to have positive correlation with 29 of the measured proteins, of which 23 (79%) were CFIM. These findings suggested that CFIM may be an excellent surrogate marker for the identification of the TAN-dysplasia association. Because many of the individual markers were overexpressed in high-risk lesions, we used multivariate modeling to develop predictive models and chose the two most promising models to validate in the current study.

This current study expands on and validates this previous work and, to our knowledge, is the first to describe a highly predictive biologic model for identification of high-risk IPMN based on cyst fluid protein expression analysis and clinical and radiographic factors. Our multi-institutional cohort of resected IPMN patients was typical of what is reported in the literature (31). The minority of patients had high-risk disease ($n=60$, 40%), and degree of dysplasia was associated with main duct involvement. The majority of patients with BD-IPMN ($n=72$, 83%) had low-risk disease, while most of MD-IPMN ($n=45$, 75%) were found to have high-risk disease on final pathological analysis. Other factors associated with high-risk disease included weight loss, jaundice, history of smoking, and the presence of solid component ($p = 0.05$). These are all well-known risk factors for high-risk disease and coincide with our findings in a clinical nomogram study as well as what others have previously published literature. [22, 43, 44]

The performance of the two multivariate models in the training and validation sets was encouraging as both model-1 and model-2 demonstrated a high predictive ability for identifying high-risk disease (c-indices 0.76, 0.80 and 0.80, 0.79, respectively for the training and validations sets). Additionally, all four markers were independently associated with the degree of dysplasia when variables from our clinical nomogram (which accounts for duct subtype) were added as a predictor in each model. Importantly, combining these protein-level models with our clinical-radiographic nomogram resulted in a higher predictive ability than any of the models or nomogram alone (Table 3).

The most recent consensus guidelines (2012 International Consensus Guidelines) for the management of IPMN rely on clinical and radiographic features, rather than biologic markers, as the primary predictive tools for identifying high-risk disease, and are typically referenced as a factor in the treatment decision making for patients with IPMN.[15] These guidelines consider the presence of pancreatitis, cyst diameter >3cm in BD-IPMN, thickened or enhancing cyst walls, main duct diameter 5–9 mm, non-enhancing mural nodule, or abrupt change in caliber of pancreatic duct with distal pancreatic atrophy as “worrisome features”. Factors considered to be “high-risk stigmata” include obstructive jaundice, enhancing solid component within cyst, and main duct dilation ≥10 mm. Resection is recommended for all patients with dilated MPD ≥10mm, i.e. MD-IPMN, and for any BD-IPMN >3cm with “high-risk stigmata”. Multiple reports however have shown that patients who undergo resection for MD-IPMN by these guidelines will have only a 50% – 60% chance of having high-grade dysplasia at the time of resection.[45] Conversely, high-grade dysplasia is present in approximately 10% – 20% of patients who undergo resection in the absence of a dilated pancreatic duct (branch duct IPMN).[12, 14] Additionally, Aso et al. investigated the associations between the malignant grade of IPMN and the number of high-risk stigmata and worrisome features (as outlined by the consensus guidelines), as well as the statistical significance of each factor for predicting malignancy, and found a high prevalence of malignancy in patients with MD-IPMN (64 %) regardless of the presence or absence of high-risk stigmata.[44] The overall sensitivity and specificity of high-risk stigmata were 57% and 90% for predicting malignant IPMN (high-grade dysplasia or invasion) and 69% and 83% for predicting invasive carcinoma, respectively.[44] Similarly, a recent study from Japan found that the accuracy of the 2012 consensus guidelines for predicting malignancy in IPMN is 45%.[46] These findings highlight the shortcomings of the current consensus guidelines and suggest a persistent need for a better tool to identify those with high-risk IPMN. Given the findings of the current study, we believe that the combination of cyst fluid protein assessment and clinical and radiographic features may allow for a robust prediction model that can decrease the risk of “undertreatment” for high-risk IPMN lesions which need to be resected to prevent progression into malignancy.

The most critical flaw in this study is that the samples were not obtained pre-operatively. Most cyst fluid samples are obtained clinically using endoscopic ultrasound guidance. These fluid samples often contain contaminants from the gastrointestinal tract that may alter protein levels.[47] In addition, this study includes only patients with pathologically proven IPMN. The ability to preoperatively identify the histopathologic sub-type of the given lesion (serous cystadenoma vs. IPMN vs. mucinous cystic neoplasm) is an additional dilemma that is only partially solved with cyst fluid CEA measurement. Previous work by our group has

not found these markers to be elevated in other histopathologic sub-types, however prospective assessment on all patients being considered for resection is needed, and a prospective study is ongoing.[26] Although prospective evaluation with EUS obtained samples will allow us to validate this approach in those undergoing resection, its applicability to patients who are selected for radiographic surveillance will remain unknown. The defined problem is that we cannot determine grade of dysplasia without resection, and therefore those who are selected for radiographic surveillance cannot have their grade (or even histopathologic sub-type) determined. Upcoming prospective studies will define a 3 to 5-year period of surveillance without progression as “success” however even this prospective evaluation will fall short and require large numbers of patients for validation.

CONCLUSIONS

This multi-institutional study validated two models, with high objective predictive ability, for the identification of high-risk IPMN based on cyst fluid protein expression. When combined with our clinical nomogram, the selected markers provided a stronger ability to predict high-risk disease than either the nomogram or cyst fluid markers alone. This study was relatively large in size, multi-institutional in nature, and included an independent dataset for validation. The latter served to decrease the risk of over-fitting the models to our data and provided a fair and unbiased test of the models. Measurement of cyst fluid proteins can provide an accurate assessment of the degree of dysplasia, and with further development may be helpful for surgical decision-making in these patients. Given the consistent findings of elevated inflammatory markers in the cyst fluid of high-risk lesions in multiple studies, an anti-inflammatory strategy may be a reasonable approach to prevent IPMN progression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work would not have been possible without the generous and unconditional collaboration between the four contributing institutions (Memorial Sloan Kettering, Johns Hopkins Hospital, Massachusetts General Hospital and University of Pittsburgh Cancer Institute).

Source of Funding: Supported in part by NIH/NCI research project (R01) grant, project 5R01CA182076-02, and the NCI SPORE grant CA62924.

References

1. Hruban RH, Takaori K, Klimstra DS, et al. An illustrated consensus on the classification of pancreatic intraepithelial neoplasia and intraductal papillary mucinous neoplasms. *Am J Surg Pathol.* 2004; 28:977–987. [PubMed: 15252303]
2. Sessa F, Solcia E, Capella C, et al. Intraductal papillary-mucinous tumours represent a distinct group of pancreatic neoplasms: an investigation of tumour cell differentiation and K-ras, p53 and c-erbB-2 abnormalities in 26 patients. *Virchows Arch.* 1994; 425:357–367. [PubMed: 7820300]
3. Ohhashi T, Saitoh Y, Maguchi H, et al. Four cases of “mucin producing” cancer of the pancreas on specific findings of the papilla of Vater [abstract]. *Prog Diag Endosc.* 1982; 20:348–51.
4. Matthaei H, Dal Molin M, Maitra A. Identification and analysis of precursors to invasive pancreatic cancer. *Methods Mol Biol.* 2013; 980:1–12. [PubMed: 23359146]

5. Maitra A, Fukushima N, Takaori K, Hruban RH. Precursors to invasive pancreatic cancer. *Adv Anat Pathol.* 2005; 12:81–91. [PubMed: 15731576]
6. Basturk O, Hong SM, Wood LD, et al. A Revised Classification System and Recommendations From the Baltimore Consensus Meeting for Neoplastic Precursor Lesions in the Pancreas. *Am J Surg Pathol.* 2015; 39:1730–1741. [PubMed: 26559377]
7. Yopp AC, Katabi N, Janakos M, et al. Invasive carcinoma arising in intraductal papillary mucinous neoplasms of the pancreas: a matched control study with conventional pancreatic ductal adenocarcinoma. *Ann Surg.* 2011; 253:968–974. [PubMed: 21422912]
8. Poultides GA, Reddy S, Cameron JL, et al. Histopathologic basis for the favorable survival after resection of intraductal papillary mucinous neoplasm-associated invasive adenocarcinoma of the pancreas. *Ann Surg.* 2010; 251:470–476. [PubMed: 20142731]
9. Sohn TA, Yeo CJ, Cameron JL, et al. Intraductal papillary mucinous neoplasms of the pancreas: an updated experience. *Ann Surg.* 2004; 239:788–797. discussion 797–789. [PubMed: 15166958]
10. Kang MJ, Jang JY, Lee KB, et al. Long-term prospective cohort study of patients undergoing pancreatotomy for intraductal papillary mucinous neoplasm of the pancreas: implications for postoperative surveillance. *Ann Surg.* 2014; 260:356–363. [PubMed: 24378847]
11. Maker AV, Lee LS, Raut CP, et al. Cytology from pancreatic cysts has marginal utility in surgical decision-making. *Ann Surg Oncol.* 2008; 15:3187–3192. [PubMed: 18766406]
12. Heckler M, Michalski CW, Schaeffe S, et al. The Sendai and Fukuoka consensus criteria for the management of branch duct IPMN - A meta-analysis on their accuracy. *Pancreatol.* 2017; 17:255–262. [PubMed: 28189431]
13. Marchegiani G, Mino-Kenudson M, Sahora K, et al. IPMN involving the main pancreatic duct: biology, epidemiology, and long-term outcomes following resection. *Ann Surg.* 2015; 261:976–983. [PubMed: 24979607]
14. Allen PJ. The management of intraductal papillary mucinous neoplasms of the pancreas. *Surg Oncol Clin N Am.* 2010; 19:297–310. [PubMed: 20159516]
15. Tanaka M, Fernandez-del Castillo C, Adsay V, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. *Pancreatol.* 2012; 12:183–197. [PubMed: 22687371]
16. Lafemina J, Katabi N, Klimstra D, et al. Malignant progression in IPMN: a cohort analysis of patients initially selected for resection or observation. *Ann Surg Oncol.* 2013; 20:440–447. [PubMed: 23111706]
17. Maker AV, Katabi N, Qin LX, et al. Cyst fluid interleukin-1beta (IL1beta) levels predict the risk of carcinoma in intraductal papillary mucinous neoplasms of the pancreas. *Clin Cancer Res.* 2011; 17:1502–1508. [PubMed: 21266527]
18. Maker AV, Katabi N, Gonen M, et al. Pancreatic cyst fluid and serum mucin levels predict dysplasia in intraductal papillary mucinous neoplasms of the pancreas. *Ann Surg Oncol.* 2011; 18:199–206. [PubMed: 20717734]
19. Sadot E, Basturk O, Klimstra DS, et al. Tumor-associated Neutrophils and Malignant Progression in Intraductal Papillary Mucinous Neoplasms: An Opportunity for Identification of High-risk Disease. *Ann Surg.* 2015; 262:1102–1107. [PubMed: 25563865]
20. Brand RE, Nolen BM, Zeh HJ, et al. Serum biomarker panels for the detection of pancreatic cancer. *Clin Cancer Res.* 2011; 17:805–816. [PubMed: 21325298]
21. Yurkovetsky Z, Skates S, Lomakin A, et al. Development of a multimarker assay for early detection of ovarian cancer. *J Clin Oncol.* 2010; 28:2159–2166. [PubMed: 20368574]
22. Attiyeh MA, Fernandez-Del Castillo C, Al Efishat M, et al. Development and Validation of a Multi-Institutional Preoperative Nomogram for Predicting Grade of Dysplasia in Intraductal Papillary Mucinous Neoplasms (IPMNs) of the Pancreas: A Report from The Pancreatic Surgery Consortium. *Ann Surg.* 2016
23. Kneuert PJ, Pitt HA, Bilimoria KY, et al. Risk of morbidity and mortality following hepato-pancreato-biliary surgery. *J Gastrointest Surg.* 2012; 16:1727–1735. [PubMed: 22760965]
24. Wu J, Jiao Y, Dal Molin M, et al. Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways. *Proc Natl Acad Sci U S A.* 2011; 108:21188–21193. [PubMed: 22158988]

25. Wu J, Matthaei H, Maitra A, et al. Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development. *Sci Transl Med*. 2011; 3:92ra66.
26. Allen PJ, Qin LX, Tang L, et al. Pancreatic cyst fluid protein expression profiling for discriminating between serous cystadenoma and intraductal papillary mucinous neoplasm. *Ann Surg*. 2009; 250:754–760. [PubMed: 19806054]
27. Lubezky N, Loewenstein S, Ben-Haim M, et al. MicroRNA expression signatures in intraductal papillary mucinous neoplasm of the pancreas. *Surgery*. 2013; 153:663–672. [PubMed: 23305591]
28. Brugge WR, Lewandrowski K, Lee-Lewandrowski E, et al. Diagnosis of pancreatic cystic neoplasms: a report of the cooperative pancreatic cyst study. *Gastroenterology*. 2004; 126:1330–1336. [PubMed: 15131794]
29. Utomo WK, Looijenga LH, Bruno MJ, et al. A MicroRNA Panel in Pancreatic Cyst Fluid for the Risk Stratification of Pancreatic Cysts in a Prospective Cohort. *Mol Ther Nucleic Acids*. 2016; 5:e350. [PubMed: 28131248]
30. Reid MD, Lewis MM, Willingham FF, Adsay NV. The Evolving Role of Pathology in New Developments, Classification, Terminology, and Diagnosis of Pancreatobiliary Neoplasms. *Arch Pathol Lab Med*. 2017; 141:366–380. [PubMed: 28055239]
31. Nagula S, Kennedy T, Schattner MA, et al. Evaluation of cyst fluid CEA analysis in the diagnosis of mucinous cysts of the pancreas. *J Gastrointest Surg*. 2010; 14:1997–2003. [PubMed: 20658204]
32. Pais SA, Attasaranya S, Leblanc JK, et al. Role of endoscopic ultrasound in the diagnosis of intraductal papillary mucinous neoplasms: correlation with surgical histopathology. *Clin Gastroenterol Hepatol*. 2007; 5:489–495. [PubMed: 17350894]
33. Kucera S, Centeno BA, Springett G, et al. Cyst fluid carcinoembryonic antigen level is not predictive of invasive cancer in patients with intraductal papillary mucinous neoplasm of the pancreas. *JOP*. 2012; 13:409–413. [PubMed: 22797397]
34. Hirono S, Kawai M, Okada KI, et al. Factors Associated With Invasive Intraductal Papillary Mucinous Carcinoma of the Pancreas. *JAMA Surg*. 2017; 152:e165054. [PubMed: 28122068]
35. Jang JY, Park T, Lee S, et al. Proposed Nomogram Predicting the Individual Risk of Malignancy in the Patients With Branch Duct Type Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Ann Surg*. 2016
36. Shimizu Y, Yamaue H, Maguchi H, et al. Validation of a nomogram for predicting the probability of carcinoma in patients with intraductal papillary mucinous neoplasm in 180 pancreatic resection patients at 3 high-volume centers. *Pancreas*. 2015; 44:459–464. [PubMed: 25423557]
37. Correa-Gallego C, Do R, Lafemina J, et al. Predicting dysplasia and invasive carcinoma in intraductal papillary mucinous neoplasms of the pancreas: development of a preoperative nomogram. *Ann Surg Oncol*. 2013; 20:4348–4355. [PubMed: 24046103]
38. Reid MDBO, Thirabanasak D, Hruban RH, Klimstra DS, Bagci P, Altinel D, Adsay V. Tumor-infiltrating neutrophils in pancreatic neoplasia. *Mod Pathol*. 2011:1612–1619. [PubMed: 21822201]
39. DiMaio CJ, Weis-Garcia F, Bagiella E, et al. Pancreatic cyst fluid concentration of high-mobility group A2 protein acts as a differential biomarker of dysplasia in intraductal papillary mucinous neoplasm. *Gastrointest Endosc*. 2016; 83:1205–1209. [PubMed: 26408423]
40. Gemenetzi G, Bagante F, Griffin JF, et al. Neutrophil-to-lymphocyte Ratio is a Predictive Marker for Invasive Malignancy in Intraductal Papillary Mucinous Neoplasms of the Pancreas. *Ann Surg*. 2016
41. Arima K, Okabe H, Hashimoto D, et al. The Neutrophil-to-Lymphocyte Ratio Predicts Malignant Potential in Intraductal Papillary Mucinous Neoplasms. *J Gastrointest Surg*. 2015; 19:2171–2177. [PubMed: 26443528]
42. Goh BK, Tan DM, Chan CY, et al. Are preoperative blood neutrophil-to-lymphocyte and platelet-to-lymphocyte ratios useful in predicting malignancy in surgically-treated mucin-producing pancreatic cystic neoplasms? *J Surg Oncol*. 2015; 112:366–371. [PubMed: 26280242]
43. Roch AM, Ceppa EP, Al-Haddad MA, et al. The natural history of main duct-involved, mixed-type intraductal papillary mucinous neoplasm: parameters predictive of progression. *Ann Surg*. 2014; 260:680–688. discussion 688–690. [PubMed: 25203885]

44. Aso T, Ohtsuka T, Matsunaga T, et al. “High-risk stigmata” of the 2012 international consensus guidelines correlate with the malignant grade of branch duct intraductal papillary mucinous neoplasms of the pancreas. *Pancreas*. 2014; 43:1239–1243. [PubMed: 25036910]
45. Salvia R, Fernandez-del Castillo C, Bassi C, et al. Main-duct intraductal papillary mucinous neoplasms of the pancreas: clinical predictors of malignancy and long-term survival following resection. *Ann Surg*. 2004; 239:678–685. discussion 685–677. [PubMed: 15082972]
46. Yamada S, Fujii T, Murotani K, et al. Comparison of the international consensus guidelines for predicting malignancy in intraductal papillary mucinous neoplasms. *Surgery*. 2016; 159:878–884. [PubMed: 26506564]
47. Belsley NA, Pitman MB, Lauwers GY, et al. Serous cystadenoma of the pancreas: limitations and pitfalls of endoscopic ultrasound-guided fine-needle aspiration biopsy. *Cancer*. 2008; 114:102–110. [PubMed: 18260088]

FIGURE 1. Study Cohort

Study Cohort. Number of patients included from each institution, excluded patients, and the breakdown of numbers between the low-risk and high-risk groups.

FIGURE 2. Representative Biomarkers with Differential Expression between Low-risk and High-risk Groups

Differential Expression of Protein Markers between Low-risk and High-risk Groups. Log scales of standardized concentrations are presented.

FIGURE 3. Combined Clinical and Molecular Nomogram for Predicting High-Risk IPMN
 Combined Clinical and Protein-level Nomogram for Predicting High-Risk IPMN. (A) Based on Model-1. (B) Based on Model-2. Nomogram score is calculated per the method described in our previous work[22], log concentrations of the CF protein markers (CA72-4 and MMP9 in Model1, sFASL and IL-4 in Model 2) are recorded and given points according to the model's point scale bar, and the total points are added and translated into a probability of high-risk disease.

Table 1

Clinical, Biologic, Radiographic, and Pathologic Characteristics with Univariate Analysis, Stratified by Low-risk and High-risk Groups (N=149). Median (low, high) or N (%).

Characteristics	Total (n=149)	High-risk* (n=60)	Low-risk (n=89)	p-value
Age at operation	71 (32, 88)	70.5 (44, 88)	71 (32, 88)	0.434
Gender				0.007
Female	63 (42)	17 (28)	46 (52)	
Male	86 (58)	43 (72)	43 (48)	
Institution				0.041
JH	47 (32)	18 (30)	29 (32)	
MGH	48 (32)	26 (43)	22 (25)	
MSK	54 (36)	16 (27)	38 (43)	
BMI	26 (18, 45)	27.76 (18, 40)	26.07 (18, 45)	0.067
Symptoms	50 (34)	26 (43)	24 (27)	0.051
Weight loss	28 (19)	17 (28)	11 (12)	0.019
Diabetes	40 (27)	21 (35)	19 (21)	0.089
Smoking	71 (47.7)	39 (65)	32 (36)	<0.001
Jaundice	6 (4)	5 (8)	1 (1)	0.039
Largest cyst size				0.409
<=3cm	56 (38.1)	21 (35.6)	35 (39.8)	
>3cm	83 (56.5)	33 (55.9)	50 (56.8)	
None seen	8 (5.4)	5 (8.5)	3 (3.4)	
Serum CA 19-9(units/mL)	22 (1, 6618)	29 (1, 6618)	16 (1, 113)	0.058
Cyst fluid CEA (ng/ml)	150 (7.2, 100800)	155.5 (7.2, 50000)	150 (10, 100800)	0.487
MPD size (cm)	0.5 (0, 2.1)	1 (0, 2.1)	0.3 (0, 1.8)	<0.001
MPD size range				<0.001
0.5cm<MPD<=1.0cm	41 (28)	26 (43)	15 (17)	
MPD<=0.5cm	87 (58)	15 (25)	72 (81)	
MPD>1.0cm	21 (14)	19 (32)	2 (2)	
Radiographic duct type				<0.001
Branch duct	87 (58)	15 (25)	72 (81)	
Main duct	62 (42)	45 (75)	17 (19)	
Abrupt MPD change	12 (9)	10 (18)	2 (3)	0.005
Solid component	22 (15)	18 (30)	4 (5)	<0.001

* High risk disease included High-grade dysplasia alone (n=34) and High-grade dysplasia with invasion (n=26).

Continuous variables are summarized using median and range (p-values from rank sum test), rounded to the nearest one.

Categorical variables are summarized using frequency and percentage (p-values from Fisher's exact test).

Abbreviations: MSK=Memorial Sloan Kettering Cancer Center, JH=Johns Hopkins Hospital, MGH=Massachusetts General Hospital, MPD=Main pancreatic duct, CEA=Carcinoembryonic antigen.

TABLE 2

Differentially Expressed Cyst Fluid Biomarkers between the Low-risk and High-risk Groups

Protein marker*	Overall Concentration	Low-risk	High-risk	p-value
Log sFASL (pg/ml)	3.408 (2.766, 9.334)	3.258 (2.766, 7.66)	3.77 (2.874, 9.334)	<0.001
Log MMP-9 (pg/ml)	9.255 (5.72, 15.56)	7.898 (5.72, 13.93)	10.4 (5.743, 15.56)	<0.001
Log IL-4 (pg/ml)	0.6419 (0.1398, 8.006)	0.5878 (0.1398, 8.006)	0.8064 (0.1398, 6.385)	0.02
Log CA72-4 (U/ml)	1.135 (−0.837, 11.93)	0.5735 (−0.837, 4.635)	3.144 (−0.3567, 11.93)	<0.001

* All are Cyst fluid inflammatory markers except: sFASL.

Analysis of cyst fluid proteins was performed using Mann-Whitney test.

Log concentrations presented as median; interquartile range.

TABLE 3

Performance Metrics for the Training and Validation Sets in Single and Combined Models.

Predictive Model	Variable	Odds ratio	p-value	Training set c-index (n=104)	Validation set c-index (n = 45 patients)
Clinical Nomogram alone	Nomogram LP	2.43 (1.50–3.93)	0.0003	0.80	0.77
Model-1[†]	Log CA72-4	1.46 (1.12–1.91)	0.0056	0.76	0.80
	Log MMP9	1.28 (1.05–1.57)	0.0134		
Clinical Nomogram Plus Model-1	Nomogram LP	1.96 (1.23–3.13)	0.0050	0.88	0.84
	Log CA72-4	1.55 (1.11–2.17)	0.0097		
	Log MMP9	1.34 (1.05–1.73)	0.0204		
Model-2[§]	Int-risk [*]	16.07 (5.10–50.68)	<.0001	0.80	0.79
	High-risk [*]	13.27 (3.94–44.59)	<.0001		
Clinical Nomogram Plus Model-2	Nomogram LP	1.64 (0.98–2.76)	0.0604	0.82	0.83
	Int-risk [*]	4.33 (1.03–18.28)	0.0462		
	High-risk [*]	4.80 (1.16–19.82)	0.0300		

LP = Linear predictor, Int-risk= Intermediate risk (Main duct IPMN, sFASL and IL4 not elevated), High-risk= sFASL above the 90th per centile or IL4 above the 84th percentile.

[†]Log (MMP9) and log (CA-72-4) as continuous predictors.[§]Three-level categorical predictor from a decision tree based on IL-4, sFASL and Main/Branch duct.^{*}Versus low risk (Branch duct IPMN, sFASL and IL4 not elevated).