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

Marc A. Attiyeh, MD[*], **Carlos Fernández-del Castillo, MD**[†], **Mohammad Al Efishat, MD**[*], **Anne A. Eaton, MS**[‡], **Mithat Gönen, PhD**[‡], **Ruqayyah Batts, BA**[*], **Ilaria Pergolini, MD**[†], **Neda Rezaee, MD**[§], **Keith D. Lillemoe, MD**[†], **Cristina R. Ferrone, MD**[†], **Mari Mino-Kenudson, MD**[‖], **Matthew J. Weiss, MD**[§], **John L. Cameron, MD**[§], **Ralph H. Hruban, MD**[¶], **Michael I. D'Angelica, MD**[*], **Ronald P. DeMatteo, MD**[*], **T. Peter Kingham, MD**[*], **William R. Jarnagin, MD**[*], **Christopher L. Wolfgang, MD**[§], and **Peter J. Allen, MD**[*]

[*]Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[†]Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

[‡]Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[§]Department of Surgery, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

[‖]Department of Pathology, Massachusetts General Hospital, Boston, MA, USA

[¶]Department of Pathology, The Sol Goldman Pancreatic Cancer Research Center, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Abstract

Objective—Previous nomogram models for patients undergoing resection of intraductal papillary mucinous neoplasms (IPMNs) have been relatively small single-institutional series. Our objective was to improve upon these studies by developing and independently validating a new model using a large multi-institutional dataset.

Summary Background Data—IPMNs represent the most common radiographically identifiable precursor lesions of pancreatic cancer. They are a heterogenous group of neoplasms in which more accurate markers of high-grade dysplasia or early invasive carcinoma could help avoid unnecessary surgery in one case and support potentially curative intervention (resection) in another.

Corresponding author: Peter J. Allen, MD, Department of Surgery, C896, Memorial Sloan Kettering Cancer Center, New York, NY 10021, 212-639-5132 phone, 212-717-3645 facsimile, allenp@mskcc.org.
Disclosures: None

Methods—Prospectively maintained databases from three institutions were queried for patients who had undergone resection of IPMNs between 2005 and 2015. Patients were separated into main duct [main and mixed-type (MD)] and branch duct (BD) types based on preoperative imaging. Logistic regression modeling was used on a training subset to develop two independent nomograms (MD and BD) to predict low-risk (low- or intermediate-grade dysplasia) or high-risk (high-grade dysplasia or invasive carcinoma) disease. Model performance was then evaluated using an independent validation set.

Results—We identified 1,028 patients who underwent resection for IPMNs [MD: n = 454 (44%), BD: n = 574 (56%)] during the ten-year study period. High-risk disease was present in 487 patients (47%). Patients with high-risk disease comprised 71% and 29% of MD and BD groups, respectively (p < 0.0001). MD and BD nomograms were developed on the training set [70% of total (n = 720); MD: n = 318, BD: n = 402] and validated on the test set [30% (n = 308); MD: n = 136, BD: n = 172]. The presence of jaundice was almost exclusively associated with high-risk disease (57 of 58 patients, 98%). Cyst size > 3.0 cm, solid component/mural nodule, pain symptoms, and weight loss were significantly associated with high-risk disease. C-indices were 0.82 and 0.81 on training and independent validation sets, respectively; Brier scores were 0.173 and 0.175, respectively.

Conclusion—For patients with suspected IPMNs, we present an independently validated model for the prediction of high-risk disease.

INTRODUCTION

Intraductal papillary mucinous neoplasms (IPMNs) of the pancreas are radiographically identifiable precursors of invasive pancreatic cancer. The incidence of IPMNs is rising mostly due to the increasing use of high-resolution cross-sectional imaging.[1,2] These cystic neoplasms have been shown to evolve from low-grade dysplasia to high-grade dysplasia to invasive carcinoma, and this pathway of progression is believed to account for 20%–30% of pancreatic cancers.[3] The timing and frequency of malignant progression are unknown, and therefore the management of patients with IPMNs is controversial.[4,5] This controversy exists because current laboratory, endoscopic, cytologic, and imaging technologies are unable to reliably distinguish between IPMNs that are at low-risk (low- to intermediate-grade dysplasia) from those that are at high-risk (high-grade dysplasia) of progressing to invasive cancer.

Presently, the most accurate factor associated with high-risk IPMNs is dilation of the main pancreatic duct on preoperative imaging [main duct IPMNs (MD-IPMN)]. Patients who undergo resection for MD-IPMN have a 50%–60% chance of having high-grade dysplasia or invasive carcinoma at the time of resection.[6] Conversely, high-grade dysplasia is present in only 10%–15% of patients who undergo resection in the absence of a dilated pancreatic duct [branch duct IPMNs (BD-IPMN)].[5] The 2012 International Consensus Guidelines (ICG2012) therefore recommend resection for patients with MD-IPMN and observation for the majority of patients with BD-IPMN.[4,7,8] The identification of more accurate markers of high-grade dysplasia could allow for more rational treatment decision-making. Low-risk patients could avoid a potentially morbid and life-threatening operation, and high-risk patients could undergo resection hopefully prior to the development of invasive disease.

Previously published data from Memorial Sloan Kettering described a nomogram-derived objective risk score that could be used to assess the probability of patients with IPMNs having high-risk disease.[9] While only a single-institutional study, the nomograms for MD-IPMN and BD-IPMN each had a relatively strong concordance index of 0.74, demonstrating a significant association between nomogram-predicted and actual risk of having high-risk disease.

The current study sought to build on these and other previous nomograms by expanding our patient population to include a large multi-institutional dataset.[10,11] These data were gathered from three high-volume institutions, and previous factors that were found to be associated with the presence of high-risk IPMNs were included in the analysis.

METHODS

Prospectively maintained databases from three of the institutions of The Pancreatic Surgery Consortium were included in the study. The Consortium is composed of five independent groups from four high-volume institutions [Memorial Sloan Kettering (MSK), Johns Hopkins Hospital (JHH), Massachusetts General Hospital (MGH), and University of Verona (UV)]. The current study was conceived and designed by investigators from MSK, JHH, and MGH. Data from MSK, JHH, and MGH were combined into a cumulative database which was queried for patients who had undergone resection of pathologically proven IPMNs between 2005 and 2015. Patients resected for a recurrent IPMN, pancreatic adenocarcinoma in the absence of an IPMN, and patients who had postoperative pathological findings of concurrent malignancies (e.g., cholangiocarcinoma, neuroendocrine tumor) were excluded. Preoperative imaging reports were reviewed to ensure that all cases were radiographically described as predominantly cystic in nature.

Demographic, clinical, laboratory, radiological, and pathological factors were extracted from the databases. The presence of symptoms was interpreted as any episode of abdominal pain or gastrointestinal (GI) disturbance in the upper abdomen, and the symptoms of weight loss and jaundice were recorded separately. Factors such as alcohol use or smoking were defined as any past or current use. Laboratory results were recorded from those obtained at pre-operative testing. If multiple cysts were seen on imaging, the cyst size and location were recorded as that of the largest cyst. Main duct dilation measurements were stratified across three categories: ≤0.5 cm, >0.5 cm and ≤1.0 cm, and >1.0 cm. IPMN sub-types were assigned based on main duct dilation (≤0.5 cm: BD-IPMN; >0.5 cm: MD-IPMN).[4] Any findings on imaging described as a solid component, thickened or enhanced cyst, concurrent lesion, and/or mural nodule were initially recorded separately but later combined into a single variable ("solid component/mural nodule") as the composite was thought to be more replicable across observers. A concurrent lesion was defined as a concurrent non-cystic finding (e.g., a mass in the head of the pancreas and a cyst in the tail). A radiological diagnosis of mixed-type was classified as main duct.

Pathological analysis was performed by dedicated gastrointestinal pathologists at each of the three institutions, and all pathology had been previously reviewed. The determination of risk was based on the highest grade of dysplasia noted in the resected lesion: low- and

intermediate-grade were classified as "low-risk," while high-grade dysplasia and invasive carcinoma were classified as "high-risk." Any incidence of adenocarcinoma on pathology with a concurrent IPMN was recorded as an invasive IPMN and therefore "high-risk." A breakdown of the study cohort can be found in Figure 1.

Statistical Analysis

The outcome of interest was the level of risk (low- vs. high-risk) determined by the grade of dysplasia on pathologic analysis. The data were split into a training set (70% of patients) and a validation set (30% of patients), stratified by MD- and BD-IPMN. Univariate and multivariate models were built from the training set to predict the probability of high-risk disease in future patients. Based on the significant difference between the levels of high-risk in the two duct type groups, separate nomograms for MD-IPMN and BD-IPMN were created. In addition, a history of jaundice had an extremely high positive predictive value (57/58 patients with jaundice in the training dataset had high-risk disease). We therefore designed our model to assign a predicted probability of high-risk disease of 1 to patients with jaundice, and those patients were consequently excluded from further model building.

Patient characteristics were summarized separately for MD-IPMN and BD-IPMN using median and range for continuous covariates, and frequency and percentage for categorical covariates. Differences between patients with low- and high-risk disease were assessed using the Wilcoxon rank sum test and Fisher's exact test. Variable selection was based on univariate significance, clinical importance, and results from prior studies. Multivariable modeling was done using logistic regression and assessed using concordance indices (c-indices), calibration plots, and Brier scores (mean squared prediction error). The concordance index is a measure of model discrimination and represents the probability that given a pair of patients, the model assigns a higher risk to the patient who is truly high risk compared to the patient who is truly low risk. Calibration plots show the true (observed) rate of high-risk disease in groups of patients defined by model-predicted risk of high-risk disease; in a well-calibrated model, the observed and expected rates are very similar. The final multivariable model was visually represented using nomograms and validated using the test datasets. 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

A total of 1,073 patients underwent pancreatic resection for IPMNs at one of the three institutions between 2005 and 2015. Resection was performed for recurrence in 20 patients, and these patients were excluded. In addition, 25 additional patients were excluded because the IPMNs were identified at the time of resection for a separate pathologically distinct malignancy (e.g., distal cholangiocarcinoma). The remaining 1,028 patients constituted our study group. Gender was equally distributed (49% male; 51% female). Median age at resection was 68 years (IQR 60–75 years). High-risk disease was identified on final pathological analysis in 487 patients (47%). Patients with MD-IPMN had a significantly higher likelihood of having high-risk disease (high-risk disease: 71% MD vs. 29% BD; $p <$ 0.0001). The training and validation sets contained 720 (70%) and 308 (30%) patients,

respectively. The distribution of MD-IPMN and BD-IPMN were comparable between the two groups (44% MD-IPMN and 56% BD-IPMN in each of the training and validation sets).

Univariate analysis identified seven variables that were significantly different between low- and high-risk groups in the MD-IPMN and BD-IPMN subsets (Table 1). Patients with isolated main duct dilation were more likely to have high-risk disease (compared to mixed-type), and mixed-type lesions were more likely to have high-risk disease when the cyst size was greater than 3.0 cm (i.e. mixed-type with large branch duct component). High-risk disease was also associated with a solid component/mural nodule, a history of weight loss, pain and/or GI symptoms, and main pancreatic duct dilatation greater than 1.0 cm. For BD-IPMN, high-risk disease was associated with a cyst size greater than 3.0 cm, solid component/mural nodule, pain and/or GI symptoms, older age, and male gender. Preoperative CA 19-9 levels were only available for approximately 60% of the patients in each of the main and branch duct groups, and despite their significance in univariate analysis, they were excluded from further modeling.

Based on univariate results, a multivariate logistic regression model was built for MD- and BD-IPMN (Table 2), and nomograms were created to predict high-risk disease (high-grade dysplasia or invasive carcinoma) (Figure 2). Patients with jaundice were assigned a high-risk probability of 1, and the rest were assigned a high-risk probability based on the nomogram that matched their radiological diagnosis (MD-IPMN or BD-IPMN). For example, a 70-year-old non-jaundiced asymptomatic male with a 3.5 cm BD-IPMN without high-risk imaging features would have a score of 136, resulting in a probability of high-risk disease of 32%. We initially tested the model using our training set, and the c-index was 0.82 (Brier score 0.173). The validation set was then applied to the model, and the c-index was 0.81 (Brier score 0.175). Calibration plots for training and validation sets can be seen in Figure 3. The data points on the training calibration plot are expected to be close to the equivalence line since they were used to build the model. The strength of the model is displayed in the validation calibration plot as the equivalence line generally falls within the 95% confidence interval of the observed rate of high-risk disease for each group indicating accurate prediction of high-risk disease on new unseen data.

DISCUSSION

Currently, our ability to accurately identify high-risk disease in patients with IPMNs is limited. Resection is generally recommended for patients with MD-IPMN, yet up to 40% of these patients will have low-risk disease at the time of resection.[12] The consequences of these limitations should not be understated as pancreaticoduodenectomy continues to be associated with a 2%–4% risk of mortality and a 20%–25% risk of major morbidity at institutions with the largest operative volumes.[13] A recent report from MSK highlighted the difficulty in identifying those at high-risk for progression to invasive cancer.[12] In this study of 186 patients who underwent resection for IPMNs, there were 75 patients (40%) who proved to have only low- or intermediate-grade dysplasia. The median age of patients in this study was 69 years, the risk of dying from operative complications was 2%, and the major operative complication rate was 37%. Improving our ability to predict high-risk IPMNs would improve clinical care. Patients with low-risk lesions could be monitored and avoid a

life-threatening operation until high-risk disease developed, and patients with high-risk lesions could undergo resection hopefully prior to the development of pancreatic cancer.

In the current study, we developed and independently validated a preoperative clinical model for IPMN that strongly predicts the risk of having high-grade dysplasia or invasive cancer. Analysis of our predictive model suggests that it may be better than the ICG2012 at identifying the presence of high-risk disease. The c-index of the model on validation data was 0.81 which highlights the model's ability to discriminate between low- and high-risk disease in a large group of patients 81% of the time. Currently, the reported rate of high-risk disease in patients with main duct dilation undergoing resection for presumed MD-IPMN is approximately 60%.[6,12] In addition, a separate model that determined high-risk probability based solely on the presence of main vs. branch duct disease was run on our validation dataset and the c-index was 0.74. Therefore, our model is able to predict higher-risk disease better than the presence of main duct dilation alone.

The strengths of this study include the large sample size, the multi-institutional nature of the data, and the use of an independent validation dataset. Prior studies have typically been single-institutional and without independent validation. Validation on an independent dataset decreases the risk of over-fitting the model to an individual dataset, and the similarity between the c-indices of the training (0.82) and validation (0.81) sets suggest that this model is widely applicable. An additional advantage of this model, and nomograms in general, is that they assign risk probabilities on a continuous scale as an individualized risk score rather than splitting patients into two broad risk groups. This allows for additional stratification of risk and for patients and doctors to tailor treatment decisions based on patients' individual risks.

The prevalence of high-risk disease in the present study is in accord with existing literature: 71% of resected MD-IPMN were found to have high-risk disease (defined as having high-grade dysplasia or invasive carcinoma) compared to only 29% of resected BD-IPMN. In patients with MD-IPMN, cyst size greater than 3.0 cm carried an odds ratio (OR) of 2.19 and an even higher OR of 3.61 when there were no cysts seen ($p = 0.002$). These results are similar to previously published reports that have demonstrated a slightly lower risk of high-grade or invasive IPMNs in patients undergoing resection for mixed-type IPMNs when compared to pure main duct disease.[14,15] Interestingly, to our knowledge, the association in mixed-type IPMNs between a larger cyst size and higher-risk disease has not been reported.

The presence of a solid component, mural nodule, concurrent lesion, or thickened or enhancing cyst on imaging was associated with the presence of high-risk disease in both MD- and BD-IPMN. We combined these findings into a single variable ("solid component/mural nodule") as previous studies have documented the difficulty in distinguishing between these features.[16] As part of a study by Do et al. investigating interobserver agreement, four independent radiologists reviewed pancreatic protocol CT studies for 84 patients who had undergone resection for IPMNs. They classified the lesion as main, branch, or mixed and provided their estimation of the presence of malignant features such as a solid component or a mural nodule. The study results showed that while the radiologists' estimations of cyst size and MPD diameter were comparable, their assessment of malignant features was more

variable. The authors suggested that these markers could be better determined as part of a tumor board conference where a consensus can be reached especially for factors that consensus guidelines have established as indications for resection.

Jaundice is considered a marker of high-risk disease in patients with a cystic lesion of the head of the pancreas. The ICG2012 recommend resection in such cases.[4] Similarly, in our study, jaundice was found to be a very strong predictor of high-risk disease. For MD and BD patients who presented with a history of jaundice, 37/37 (100%) and 20/21 (95%) of patients were found to have high-risk disease following resection, respectively. (The single low-risk patient with a history of jaundice also had prior episodes of biliary stricture and cholangitis which we believe to be the cause of his symptom.) Our model automatically assigns a predicted probability of high-risk disease of 1 to these patients and excludes them from the logistic regression analysis, allowing us to more accurately measure associations between the remaining factors and high-risk disease. Therefore, our model agrees with previously published guidelines stating that patients with jaundice should undergo resection.

Currently, standard recommendations for the management of IPMN are based on metaanalyses first published in 2006 and later updated in the ICG2012.[4,7] These guidelines attempt to stratify patients into higher-risk groups and aid surgeons with treatment recommendations. Recommendations include resection for all patients with MD-IPMN and resection for BD-IPMN with 'high-risk stigmata' on imaging (e.g., mural nodules). Observation is generally recommended for BD-IPMN without radiological findings of a solid component or mural nodularity. A review of the 2006 guidelines by Nagai et al. found that while the guidelines had near perfect sensitivity (97%), their low specificity (30%) resulted in many patients with low-risk IPMNs being resected.[17] In 2015, a validity study examined the conclusions drawn from the ICG2012 and found that many mixed-type IPMNs were actually low-risk which further reduced the likelihood of high-risk disease in the MD-IPMN group.[18] Their results for sensitivity and specificity for the ICG2012 guidelines were 88% and 65%, respectively. These findings suggest a continued need for an improved ability to discriminate between patients with low- vs. high-risk IPMNs.

A nomogram is a graphical representation of a complex statistical formula that accepts multiple input variables and provides an easy-to-understand answer to a focused question. As a prognostic tool, nomograms provide an individualized risk score for a given patient. Both preoperative (e.g., estimating risk of severity of disease) and postoperative (e.g., predicting recurrence-free or overall survival) nomograms have been described in the literature on topics such as breast, GI, and prostate cancer. A recent review highlighted the strengths and pitfalls of using clinical nomograms.[19] The authors highlighted four key performance metrics: validation, discrimination, calibration, and clinical usefulness. With respect to validation, our study used an independent dataset to validate the model in order to ensure a more fair and unbiased assessment of the model. With respect to discrimination, nomograms are typically scored using a c-index ranging from 0.5 (as good as chance) to 1.0 (perfect discrimination). The c-index of our model was 0.81 on a validation dataset meaning that 81% of the time, the model assigned a lower probability to a patient with truly low-risk disease than a patient with high-risk disease. With respect to calibration, the accuracy of a nomogram is best depicted by a calibration plot showing the relationship between predicted

risk and actual risk. An ideal plot would show a diagonal line ($y=x$). The calibration plot of our model on the training set (Figure 3a) demonstrated a strong association between the nomograms and the data. The validation plot (Figure 3b) was expectedly weaker (i.e. larger confidence intervals, data points further away from the "ideal line") but was still able to show high accuracy for patients in the lower (~10%–30%) and higher (~90%) risk groups. Finally, with respect to clinical usefulness, our results suggest that our model may be a better predictor of high-risk disease and therefore could be a useful adjunct to clinical decision-making. It provides a risk assessment on a continuous scale, as opposed to the ICG2012's categorical criteria, that is easier to apply to an individual patient in the context of associated co-morbidities and life expectancy. Further work will include a prospective analysis to determine whether it significantly improves patient outcomes when compared to clinician-directed management.

In 2013, researchers from MSK published an IPMN nomogram that sought to predict high-grade dysplasia and invasive carcinoma based on more limited data.[9] The nomograms developed in this previous study contained the same factors presented here, namely, solid component/mural nodule, lesion size, and weight loss. The endpoint in that study was a three-level ordinal outcome: benign, high-grade dysplasia, and invasive carcinoma. In the present study, we elected to use the simpler and more clinically useful endpoint of high-risk disease, a composite of high-grade dysplasia and invasive carcinoma. By expanding our sample size to include two other large pancreatic centers and by including independent validation, our study serves not only to support prior results but also to expand and strengthen the model by identifying other possible markers of high-risk disease. In 2010, researchers in Japan created and later validated a nomogram that could predict the probability of carcinoma in IPMNs.[11,20] While this study had the benefit of being externally validated, the sample size used (n = 81 for the training set; n = 180 for validation) was relatively small. In addition, one factor they found to be significant was cytology grade. As part of our data collection, we attempted to collect data on cytology in a similar manner, but the variability as well as lack of specificity in reports led to the exclusion of this variable from the final model. Practically speaking, the use of a non-standardized cytology grade in the model renders a nomogram difficult to apply to other centers.

The present study has several weaknesses. First, our cohort only included patients who underwent resection resulting in a selection bias in our data; it is unknown whether these factors would remain significant in unresected patients. This makes it difficult to apply the nomograms to patients who carry a diagnosis of an incidental cystic lesion. Additional studies are currently being developed to validate our model on unresected patients undergoing surveillance. Second, as mentioned previously, the clinical utility our model offers has yet to be demonstrated to be significant due to the lack of any prospective analysis. We anticipate future studies will further validate our model by applying it to patients with IPMNs and determining its accuracy and usefulness. However, for this particular type of study, patients determined to likely be low-risk and therefore managed conservatively will not have a pathological diagnosis to support or refute the model's prediction. Assuming the patient does not undergo resection during their follow-up period, one possible solution is to define a length of time (e.g., at least five years) after which it can be safely stated that in the absence of clinical or radiographic evidence of disease

progression, the patient likely had a low-risk lesion. Third, our study found preoperative CA 19-9 to be predictive of high-risk disease on univariate analysis. However, due to its specificity for malignancy and that approximately 40% of our cohort did not have CA 19-9 levels available, it was excluded from subsequent analysis. Finally, our model is not meant to replace a clinician's decision-making with regard to resecting an IPMN. Although nomograms can predict the likelihood of identifying a high-risk lesion, only the surgeon and patient can best balance risks and benefits and decide the threshold for which resection is indicated.

In conclusion, for patients with suspected IPMNs, we present an independently validated model containing two nomograms for predicting high-risk disease. Our study is the largest to date to identify significant factors contributing to high-risk disease in IPMNs, and our model displays strong objective predictive power when validated with independent data. As noted previously, future studies will need to expand these nomograms to include unresected patients so their applicability can go beyond pre-operative patients. Finally, studies investigating the use of cyst fluid characteristics as diagnostic and/or prognostic markers may enhance our model even further.

Acknowledgments

References

1. Gaujoux S, Brennan MF, Gonen M, et al. Cystic lesions of the pancreas: changes in the presentation and management of 1,424 patients at a single institution over a 15-year time period. J Am Coll Surg. 2011; 212(4):590–600. discussion 600-603. [PubMed: 21463795]

2. Fernandez-del Castillo C, Warshaw AL. Current management of cystic neoplasms of the pancreas. Adv Surg. 2000; 34:237–248. [PubMed: 10997221]

3. Maitra A, Fukushima N, Takaori K, et al. Precursors to invasive pancreatic cancer. Adv Anat Pathol. 2005; 12(2):81–91. [PubMed: 15731576]

4. Tanaka M, Fernandez-del Castillo C, Adsay V, et al. International consensus guidelines 2012 for the management of IPMN and MCN of the pancreas. Pancreatology. 2012; 12(3):183–197. [PubMed: 22687371]

5. Allen PJ. The management of intraductal papillary mucinous neoplasms of the pancreas. Surg Oncol Clin N Am. 2010; 19(2):297–310. [PubMed: 20159516]

6. 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(5):678–685. discussion 685-687. [PubMed: 15082972]

7. Tanaka M, Chari S, Adsay V, et al. International consensus guidelines for management of intraductal papillary mucinous neoplasms and mucinous cystic neoplasms of the pancreas. Pancreatology. 2006; 6(1–2):17–32. [PubMed: 16327281]

8. Schmidt CM, White PB, Waters JA, et al. Intraductal papillary mucinous neoplasms: predictors of malignant and invasive pathology. Ann Surg. 2007; 246(4):644–651. discussion 651-654. [PubMed: 17893501]

9. 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(13):4348–4355. [PubMed: 24046103]

10. Hijioka S, Shimizu Y, Mizuno N, et al. Can long-term follow-up strategies be determined using a nomogram-based prediction model of malignancy among intraductal papillary mucinous neoplasms of the pancreas? Pancreas. 2014; 43(3):367–372. [PubMed: 24622065]

11. Shimizu Y, Kanemitsu Y, Sano T, et al. A nomogram for predicting the probability of carcinoma in patients with intraductal papillary-mucinous neoplasm. World J Surg. 2010; 34(12):2932–2938. [PubMed: 20845037]

12. 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(2):440–447. [PubMed: 23111706]

13. Kneuertz PJ, Pitt HA, Bilimoria KY, et al. Risk of morbidity and mortality following hepato-pancreato-biliary surgery. J Gastrointest Surg. 2012; 16(9):1727–1735. [PubMed: 22760965]

14. Crippa S, Fernandez-del Castillo C. Management of intraductal papillary mucinous neoplasms. Curr Gastroenterol Rep. 2008; 10(2):136–143. [PubMed: 18462599]

15. 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(4):680–688. discussion 688-690. [PubMed: 25203885]

16. Do RK, Katz SS, Gollub MJ, et al. Interobserver agreement for detection of malignant features of intraductal papillary mucinous neoplasms of the pancreas on MDCT. AJR Am J Roentgenol. 2014; 203(5):973–979. [PubMed: 25341134]

17. Nagai K, Doi R, Ito T, et al. Single-institution validation of the international consensus guidelines for treatment of branch duct intraductal papillary mucinous neoplasms of the pancreas. J Hepatobiliary Pancreat Surg. 2009; 16(3):353–358. [PubMed: 19280108]

18. Watanabe Y, Nishihara K, Niina Y, et al. Validity of the management strategy for intraductal papillary mucinous neoplasm advocated by the international consensus guidelines 2012: a retrospective review. Surg Today. 2015

19. Balachandran VP, Gonen M, Smith JJ, et al. Nomograms in oncology: more than meets the eye. Lancet Oncol. 2015; 16(4):e173–180. [PubMed: 25846097]

20. 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(3):459–464. [PubMed: 25423557]
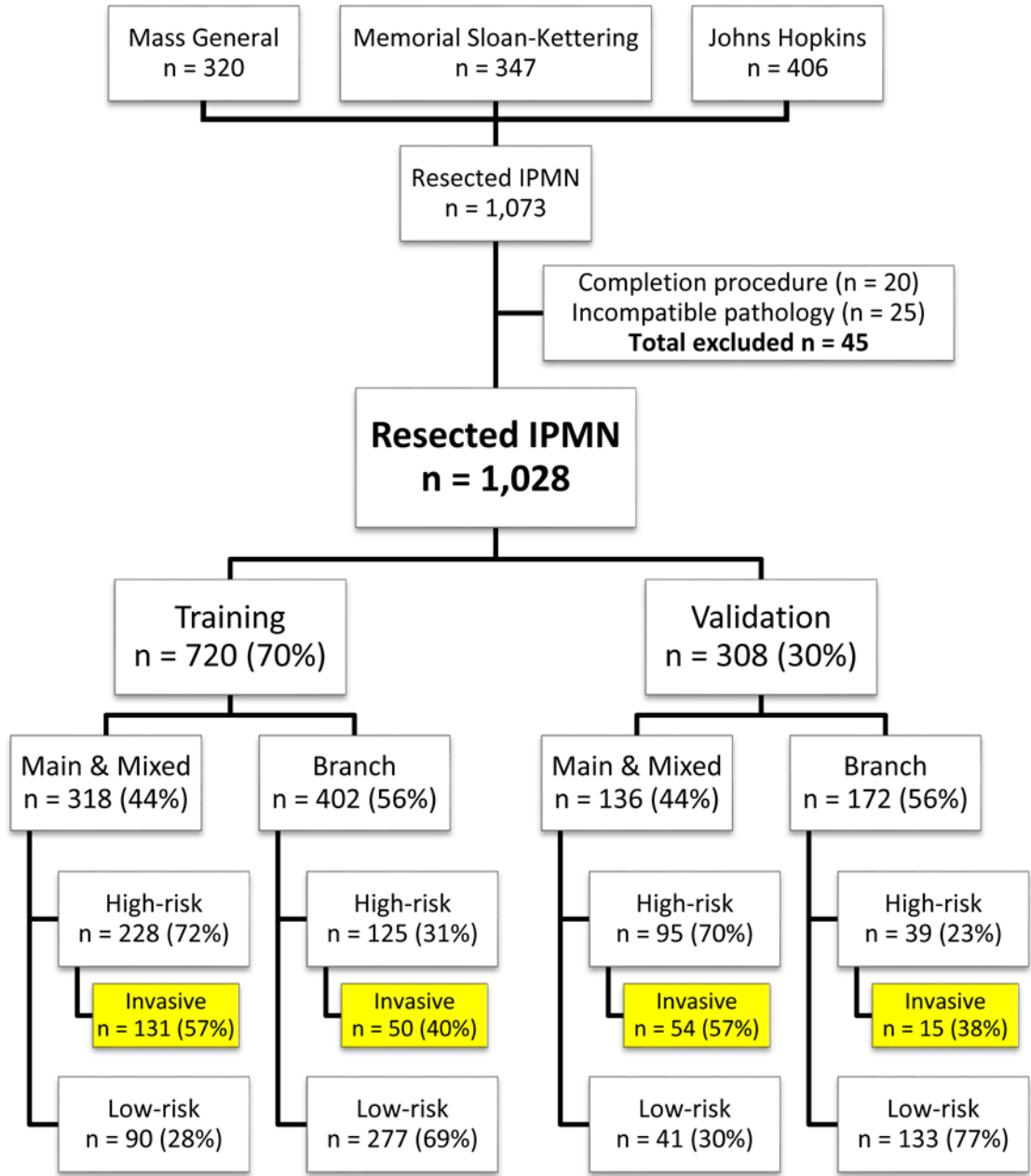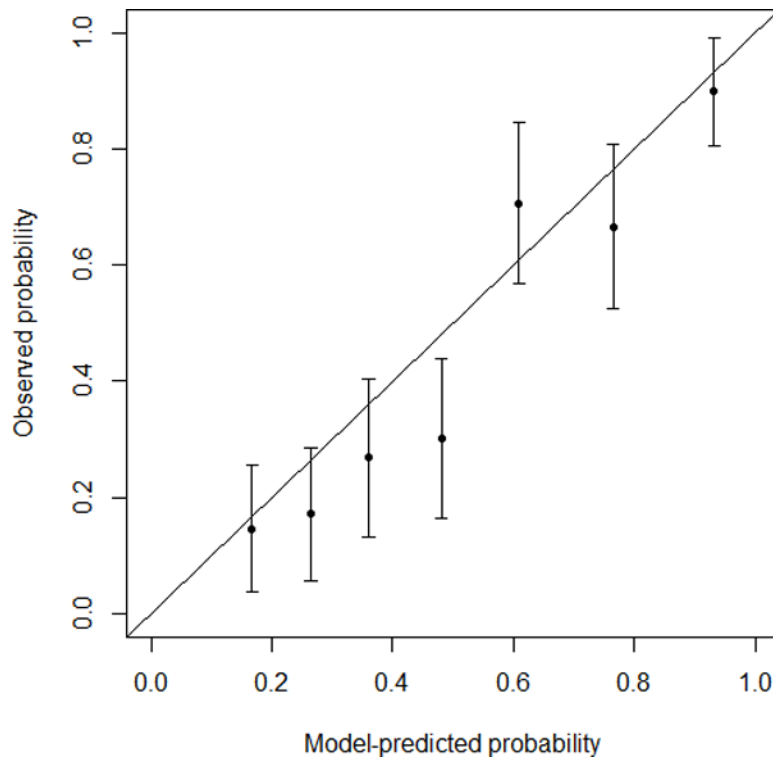
FIGURE 1.
Study cohort
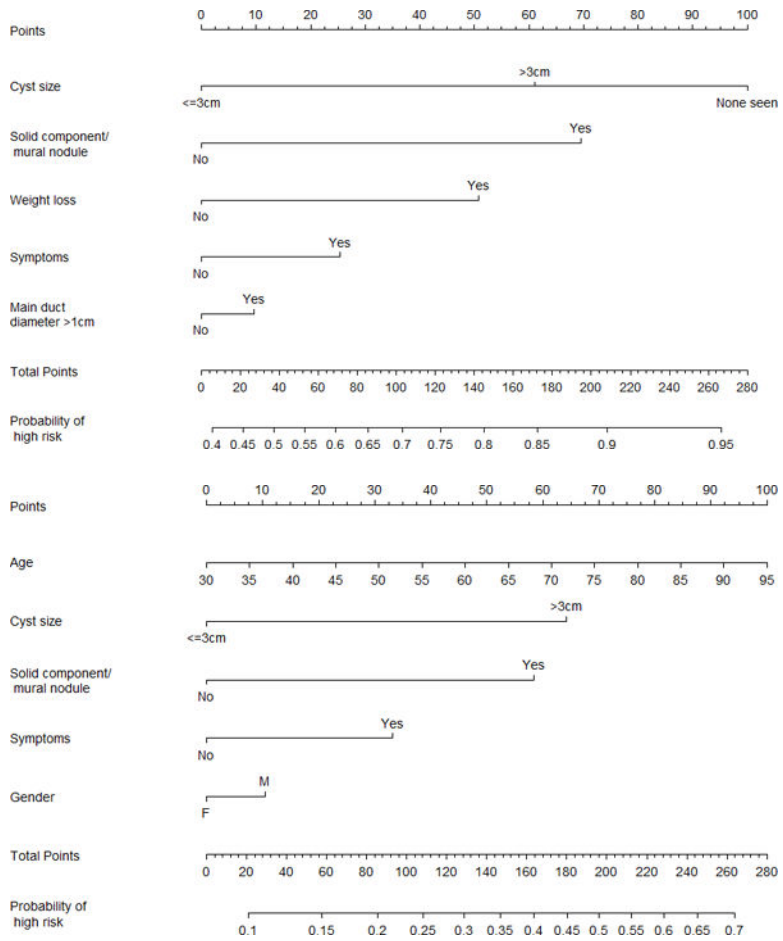
FIGURE 2.

Clinical nomograms for predicting high-risk disease in non-jaundiced patients with MD-IPMN (2a) and BD-IPMN (2b). If a patient has symptoms of jaundice, assign probability = 1.

○ FIGURE 2a – Main duct nomogram.

○ FIGURE 2b – Branch duct nomogram.
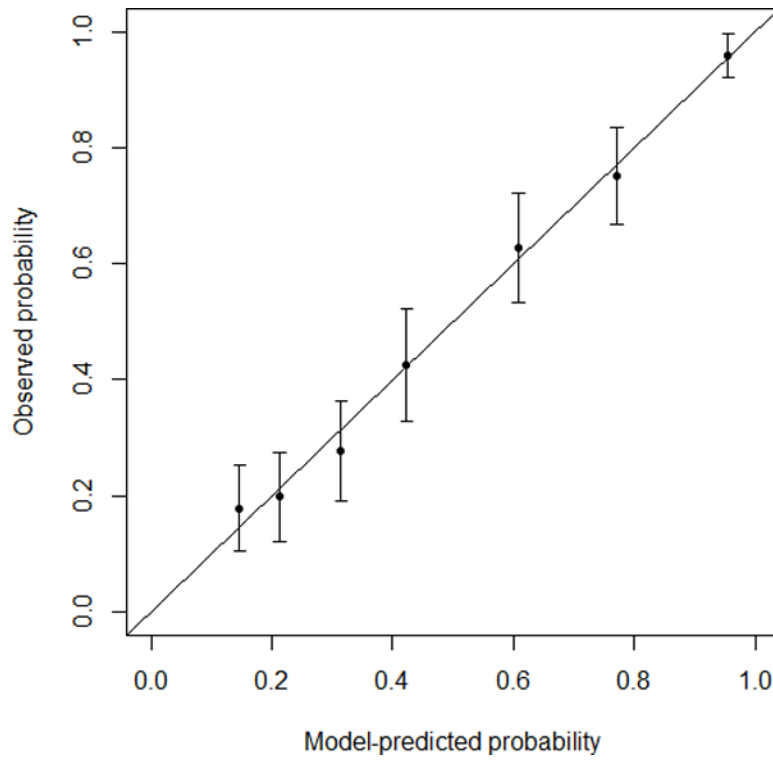
FIGURE 3.

Observed probability of high risk disease by model-predicted probability of high risk disease in training data (3a) and validation data (3b). Error bars represent 95% confidence intervals for observed probabilities.

- ○ FIGURE 3a – Training calibration plot. C-index 0.82. Brier score 0.173.

- ○ FIGURE 3b – Validation calibration plot. C-index 0.81. Brier score 0.175.

TABLE 1

Patient characteristics from training population and univariate analysis (n = 662). Patients with a history of jaundice (n = 58) were excluded from training set. Median (low–high) or N (%).

| | Training set (n = 662) | | | | | | | |
| | Main duct and mixed-type (n = 281) | | | | Branch duct (n = 381) | | | |
	Total (n)	High-risk (n = 191)	Low-risk (n = 90)	P-value	Total (n)	High-risk (n = 105)	Low-risk (n = 276)	P-value
Institution								
MSK	99 (35%)	71 (37%)	28 (31%)		129 (34%)	40 (38%)	89 (32%)	
JHH	93 (33%)	65 (34%)	28 (31%)	0.331	152 (40%)	39 (37%)	113 (41%)	0.569
MGH	89 (32%)	55 (29%)	34 (38%)		100 (26%)	26 (25%)	74 (27%)	
Age	68 (18–92)	67 (18–92)	69 (30–89)	0.244	67 (34–92)	70 (41–92)	66 (34–88)	0.011[a]
Body Mass Index	25.8 (15.5–46.1)	25.9 (15.5–46.1)	25.4 (17.5–38.1)	0.922	26.0 (15.0–47.0)	25.7 (17.6–47.0)	26.3 (15.0–43.2)	0.383
Gender								
Male	154 (55%)	112 (59%)	42 (47%)		160 (42%)	50 (48%)	110 (40%)	
Female	127 (45%)	79 (41%)	48 (53%)	0.072	221 (58%)	55 (52%)	166 (60%)	0.201[a]
Diabetes								
Yes	78 (28%)	56 (29%)	22 (24%)		58 (15%)	15 (14%)	43 (16%)	
No	203 (72%)	135 (71%)	68 (76%)	0.476	323 (85%)	90 (86%)	233 (84%)	0.873
Pancreatitis								
Yes	91 (32%)	62 (32%)	29 (32%)		81 (21%)	25 (24%)	56 (20%)	
No	190 (68%)	129 (68%)	61 (68%)	1.000	300 (79%)	80 (76%)	220 (80%)	0.484
Personal history of cancer								
Yes	64 (23%)	39 (20%)	25 (28%)	0.174	68 (18%)	20 (19%)	48 (17%)	0.765

	Training set (n = 662)							
	Main duct and mixed-type (n = 281)				Branch duct (n = 381)			
	Total (n)	High-risk (n = 191)	Low-risk (n = 90)	P-value	Total (n)	High-risk (n = 105)	Low-risk (n = 276)	P-value
No	217 (77%)	152 (80%)	65 (72%)		313 (82%)	85 (81%)	228 (83%)	
Family history of pancreatic cancer								
Yes	31 (11%)	20 (10%)	11 (12%)	0.686	64 (17%)	7 (7%)	57 (21%)	<0.001
No	250 (89%)	171 (90%)	79 (88%)		317 (83%)	98 (93%)	219 (79%)	
Symptomatic								
Yes	159 (57%)	117 (61%)	42 (47%)	0.028[a]	167 (44%)	53 (50%)	114 (41%)	0.133[a]
No	122 (43%)	74 (39%)	48 (53%)		214 (56%)	52 (50%)	162 (59%)	
Weight loss								
Yes	93 (33%)	73 (38%)	20 (22%)	0.010[a]	51 (13%)	16 (15%)	35 (13%)	0.505
No	188 (67%)	118 (62%)	70 (78%)		330 (87%)	89 (85%)	241 (87%)	
CA 19-9 (serum) > 40[b]								
Yes	42 (23%)	36 (29%)	6 (10%)	0.003	35 (16%)	16 (28%)	19 (12%)	0.007
No	141 (77%)	86 (71%)	55 (90%)		185 (84%)	42 (72%)	143 (88%)	
Solid component/ mural nodule[c]								
Yes	122 (43%)	94 (49%)	28 (31%)	0.005[a]	102 (27%)	41 (39%)	61 (22%)	0.001[a]
No	159 (57%)	97 (51%)	62 (69%)		279 (73%)	64 (61%)	215 (78%)	
Number of cysts								
0	52 (19%)	42 (22%)	10 (11%)	0.001[a]	N/A	N/A	N/A	0.381
1	158 (56%)	110 (58%)	48 (53%)		223 (60%)	69 (66%)	154 (58%)	
2	29 (10%)	21 (11%)	8 (9%)		53 (14%)	14 (13%)	39 (15%)	

	Training set (n = 662)								
	Main duct and mixed-type (n = 281)					Branch duct (n = 381)			
	Total (n)	High-risk (n = 191)	Low-risk (n = 90)	P-value	Total (n)	High-risk (n = 105)	Low-risk (n = 276)	P-value
3 +	42 (15%)	18 (9%)	24 (27%)		94 (26%)	22 (21%)	72 (27%)	
Largest cyst size[d]								
3.0 cm	113 (42%)	63 (34%)	50 (58%)		253 (67%)	54 (52%)	199 (72%)	
> 3.0 cm	106 (39%)	80 (43%)	26 (30%)	<0.001[a]	127 (33%)	50 (48%)	77 (28%)	<0.001[a]
None seen	52 (19%)	42 (23%)	10 (12%)		N/A	N/A	N/A	N/A
MPD size								
0.5 cm < and 1.0 cm	220 (78%)	144 (75%)	76 (84%)		N/A	N/A	N/A	N/A
> 1.0 cm	61 (22%)	47 (25%)	14 (16%)	0.091[a]	N/A	N/A	N/A	

[a] Variables used in subsequent multivariate analysis

[b] Preoperative CA 19-9 was only available for 183/281 (65%) MD-IPMN and 220/381 (58%) BD-IPMN patients

[c] Solid component, thickened or enhanced cyst, mural nodule, or concurrent lesion

[d] Ten MD-IPMN patients and one BD-IPMN patient did not have cyst size information available.

MSK=Memorial Sloan Kettering, JHH=Johns Hopkins, MGH=Massachusetts General Hospital

MPD=Main pancreatic duct

IQR=Interquartile range

TABLE 2

Main and branch duct nomogram models (multivariable logistic regression models fit on the training data, excluding patients with jaundice). Odds ratios refer to the odds of having high-risk disease (vs. low-risk).

	Main duct and mixed-type (n = 271)		Branch duct (n = 380)	
	Odds ratio (95% CI)	*P*-value	Odds ratio (95% CI)	*P*-value
Largest cyst size[a]				
> 3.0 cm	2.19 (1.20–4.05)	0.002	2.24 (1.37–3.65)	0.001
None seen	3.61 (1.65–8.53)		N/A	N/A
Solid component/mural nodule	2.44 (1.39–4.39)	0.002	2.08 (1.25–3.45)	0.005
Weight loss	1.92 (0.92–4.12)	0.086	N/A	N/A
Symptomatic	1.39 (0.73–2.64)	0.316	1.51 (0.94–2.44)	0.087
Main duct > 1.0 cm	1.13 (0.55–2.40)	0.742	N/A	N/A
Age[b]	N/A	N/A	1.02 (1.00–1.05)	0.119
Gender (male)	N/A	N/A	1.14 (0.70–1.85)	0.593

[a]Odds ratio compared to reference category, 3.0cm

[b]Odds ratio per one year increase in age

CI=Confidence Interval