Annals of Surgery Publish Ahead of Print
DOI:10.1097/SLA.0000000000006702

OPEN

Personalized surveillance intervals for intraductal papillary mucinous neoplasm (IPMN): multicenter study using parametric models

Claudio Ricci MD, PhD[1-2] Stefano Crippa, MD, [3], Johnathan Hee, MD[4-5], Hyesol Jung, MD [6], Gabriele Capurso, MD [7], Marco Ferronato, MD [1-8], José Lariño Noia, MD [9], Myrte Gorris, MD [10], Paula Ghaneh, MD [11], Ihsan Ekin Demir, MD [12], Nuzhat Ahmad, MD [13], Max Heckler MD [14], Giulio Belfiori MD [3], Francesca Aleotti, MD [3], Youngmin Han, MD [6], Wooil Kwon, MD [6], Gaetano Lauri, MD [7], Matteo Tacelli, MD [7], Olivier Busch, MD [10], Kulbir Mann, MD [11], Marina Migliori, MD [8], Paolo Giorgio Arcidiacono, MD [7], Helmut Friess, MD [12], Charles M Vollmer, MD [15], Thilo Hackert, MD [14-15] Marc Besselink, MD[10], Riccardo Casadei, MD[1,2], Jin-Young Jang, MD [6], Brian Kim-Poh Goh, MD [4-17], Massimo Falconi, MD[3], Giovanni Marchegiani, MD [18]

[1] Department of Internal Medicine and Surgery (DIMEC); Alma Mater Studiorum, University of Bologna

[2] Division of Pancreatic Surgery, IRCCS Azienda Ospedaliero-Universitaria Di Bologna, via Albertoni 15- Italia

[3] Division of Pancreatic Surgery, Pancreas Translational and Clinical Research Center, San Raffaele Scientific Institute, Milan, Italy

[4] Department of Hepatopancreatobiliary and Transplant Surgery, Singapore General Hospital and National Cancer Center Singapore, Singapore

[5] Ministry of Health Holdings Singapore

[6] Department of Surgery and Cancer Research Institute, Seoul National University Hospital, Seoul National University College of Medicine, Seoul, Korea.

[7] Pancreato-Biliary Endoscopy and Endoscopic Ultrasound, Pancreas Translational and Clinical Research Center, San Raffaele Scientific Institute, Milan, Italy.

[8] IRCCS Azienda Ospedaliero-Universitaria di Bologna, Internal Medicine Unit, Italy

[9] Endoscopy and Pancreatic Unit, University Hospital of Santiago de Compostela, Santiago de Compostela, Spain.

[10] Department of Surgery, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands; Cancer Center Amsterdam, Amsterdam, The Netherlands.

[11] Royal Liverpool University Hospital, Liverpool, United Kingdom

[12] Department of Gastroenterology, Klinikum rechts der Isar, School of Medicine, Technical University Munich, Munich, Germany

[13] Division of Gastroenterology and Hepatology, University of Pennsylvania, Philadelphia, Pennsylvania

[14] Department of Surgery, Division of General Surgery, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania

[15] Department of General, Visceral and Transplantation Surgery, Heidelberg University Hospital, Heidelberg, Germany.

[16] Department of General, Visceral and Thoracic Surgery, University Hospital Hamburg-Eppendorf, Hamburg, Germany

[17] Surgery Academic Clinical Program, Duke-NUS Medical School, Singapore

[18] Hepatopancreatobiliary and Liver Transplant Surgery, Department of Surgical, Oncological and Gastroenterological Sciences (DiSCOG), University of Padua, Padua, Italy.

Correspondence should be addressed to:

Claudio Ricci, Department of Internal Medicine and Surgery (DIMEC); Alma Mater Studiorum, University of Bologna

Pancreas Unit, S.Orsola-Malpighi Hospital, Italy, Via Massarenti 9, Bologna, Italy, ZIP code 40138

claudio.ricci6@unibo.it

Disclosure: All authors disclosed any support from the National Institutes of Health (NIH), Wellcome Trust, Howard Hughes Medical Institute (HHMI), and others. The authors did not receive funds or grants for the manuscript.

Running Head: Parametric analysis for branch-duct intraductal papillary mucinous neoplasm

evolution

Keywords: branch-duct intraductal papillary mucinous neoplasm

Word count: 3242

Abstract

Objective: The aim was to build a calculator for personalized surveillance of BD-IPMNs.

Summary background Data: The interval time for surveillance of low-risk branch duct intraductal papillary mucinous neoplasms (BD-IPMNs) has not been established yet.

Methods: The study included an international cohort of BD-IPMNs without worrisome features (WFs) or high-risk stigmata (HRS). IPMN evolution was defined as the occurrence of HRS or WFs. The derivation cohort comprised 60% of patients. The validation group comprised the remaining patients. A parametric survival model was developed in the derivation cohort using Akaike (AIC) and Bayesian (BIC) information criteria and c-index. A "k-fold" validation was used to measure the covariate effect on the accelerated failure time. Two models ("standard" and "conservative") were built and validated using the second cohort.

Results: The derivation and validation cohorts included 1,992 and 1,119 BD-IPMNs. The lognormal distribution best fitted the derivation cohort (AIC=2673; BIC=2718). The pooled c-index was 0.689 (0.668 to 0.718, 95%CI). The factors reducing the time needed for IPMN evolution were age [- 2% (-1% to -3%) for each year] and cyst size [-2% (0% to -3%); for each mm]. The "conservative" model, called PANORAMA, was the only one that correctly classified the validation cohort (c-index 0.712 vs 0.696; P=0.072).

Conclusion and Relevance : The development of WF and HRS is influenced by the patient's age and cyst size. After a prudential first control at six months, repeating a semestral/annual follow-up in this time frame could be too tight.

Introduction

In recent years, several studies[1-4] have demonstrated that branch-duct intraductal papillary mucinous neoplasms (BD-IPMN) without worrisome features (WFs) and high-risk stigmata (HRS) are at very low risk of progression toward malignancy. Nonetheless, the small subset of low-risk BD-IPMNs developing WFs or HRs during surveillance deserves special attention as it shapes the follow-up duration and timing.[5] Available guidelines[6-8], in their latest iterations, still recommend multiple and repeated evaluations over time to capture radiological changes requiring further diagnostic evaluation or therapeutic interventions, like endoscopic ultrasonography (EUS) or surgery. The overall length of surveillance and its intervals are based on the risk of the IPMN evolving over time. This evolution is measured with Hazard ratio (HR) using Cox regression, which distills the risk into a constant baseline risk, shared across all patients, and relative risk, which describes how individual covariates influence the risk.[9] However, the most obvious and rational approach should be to calculate the interval time of follow-up, estimating the time needed to observe relevant changes such as WFs or HRS occurrence.[10] This approach reflects the knowledge that BD-IPMNs progress to malignancy through an established chain of events depicted by given clinical or radiologic features (WF and HRS) and can be explored using the "Accelerated Failure Time" (AFT) model. [11] The AFT model assumes that the impact of covariates is to speed up or slow down the disease progression. In contrast, traditional nonparametric or semiparametric models, such as the Kaplan-Meier or Cox Regression, assume that the disease progression is constant for the entire population and observation time.

AFT can be calculated only using a parametric survival method[11,] which permits violating the assumption by exploring how the risk changes over time.

Through this approach, the present article aims to build a prediction model of BD-IPMN evolution over time. This model could serve as the backbone of a personalized surveillance strategy based on the evolution vs. stability of the cysts.

Methods

Patients and eligibility criteria

This study was approved by the local review board of each participating center and followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. Informed consent policy varied across centers, and in most cases, it was waived, given the use of historic de-identified data.[12] Patients with presumed BD-IPMNs (radiologically or cytologically/histologically diagnosed) at low risk of degeneration, lacking WFs or HRS, were enrolled. The presumptive diagnosis was based on the presence of one or more dilated branch duct(s) communicating with a normal (5 mm or smaller) main pancreatic duct (MPD) on high-resolution axial imaging or EUS. The exclusion criteria were: i) a follow-up less than 12 months from diagnosis; ii) a previous history of pancreatic resection; iii) cysts highly suspicious for an alternative diagnosis such as mucinous or serous cystic neoplasms, pseudocysts, solid pseudopapillary tumors, and cystic neuroendocrine tumors; iv) absence of relevant data for the study. The overall cohort of patients was divided into derivation and validation cohorts with a ratio of approximately 60/40. The derivation cohort was used to develop and internally validate the model. This cohort included all patients with presumed BD-IPMNs observed between January 2000 and December 2022 at nine tertiary centers in Western countries (Supplementary Table 1, Supplemental Digital Content 1, http://links.lww.com/SLA/F438). The external validation cohort comprised all patients with presumed BD-IPMNs observed between January 2006 and December 2017 at two tertiary centers in Eastern countries (Supplementary Table 2, Supplemental Digital Content 1, http://links.lww.com/SLA/F438).

Data Collection, Definitions, Endpoints

Age, gender, and presence of high-risk individuals for familial pancreatic cancer (FPC)[13] were registered. Clinical data included pancreas-related symptoms such as jaundice, acute pancreatitis,

new onset, or worsening diabetes. Radiological and endoscopic characteristics of the cyst were collected, including location and cyst size. The largest size was considered for the analysis when multiple cysts were present. Duration of surveillance was defined as the interval from diagnosis to the date of last follow-up, surgery, or death. Surveillance was performed according to each center's policies in line with the International Association of Pancreatology (IAP) or European guidelines. [6-8, 14-15] The target event of the study was any relevant changes (new onset of WFs or HRS) occurring during follow-up. WFs and HRS were defined according to the latest update of the IAP guidelines.[6] However, because of its retrospective nature, mural nodule size was not always available for analysis. For this reason, non-enhancing mural nodules were considered WF, and enhancing mural nodules were considered HRS.[15] The final endpoint was to build a model to estimate the time needed to observe a relevant change.

Statistical analysis

The descriptive characteristics of patients were reported as number and percentage or median and interquartile (IQR) or range. The analysis was performed in three steps: i) the model development using the derivation cohort; ii) the internal validation splitting of the derivation cohort into two groups, "in-sample" and "out-of-sample"; iii) the external validation using the model developed in derivation cohort to fit the survival data of training cohort. In the first step, the time-to-event model was built using the Western cohort, comparing Weibull, Exponential, Log-normal, and Log-logistic hazard distribution (Supplementary Figure 1, Supplemental Digital Content 1, http://links.lww.com/SLA/F438 and Supplementary Methods, Supplemental Digital Content 1, http://links.lww.com/SLA/F438). The effect of the covariates was reported using the accelerated failure time (AFT) in percentage:[16.] The covariates accelerate or decelerate the time to the event of interest, shortening or extending the time for clinical or radiological relevant change of BD-IPMN. The best model distribution was selected, measuring the goodness of fit: the smallest value of Akaike (AIC) and Bayesian (BIC) information criteria indicates the best model [17]. Moreover, the

ability to discriminate patients with and without progression was measured with Harrel's c-index.[18] The discrimination power was considered sufficient, good, very good, strong, and perfect for the c-index around 0.6, 0.7, 0.8, 0.9, and 1, respectively. [18] In the second step, the model was internally validated through a 5-fold cross-validation approach. Five times, the model was developed on nearly 80% of the derivation cohort and tested in the remaining 20% so that all patients were included in the test group once.[19] Two different prediction models were obtained using the pooled values of coefficients after the 5-fold cross-validation. In the standard prediction model, the weight of the covariates was assumed to be equal to the mean effect observed in the derivation cohort. The 95% CI upper effect of covariate was used in the conservative model. Finally, the prediction model was externally validated in the Eastern cohort, measuring calibration[21] and discrimination power.[20] For this purpose, the prediction models were used to fit the survival data of the external cohort in a parametric survival function. All statistical analysis was performed using STATA (Statistical software for data science, release 18, Stata Corp, College Station, TX) and R-project (R version 4.2.2; R Foundation for Statistical Computing, Vienna, Austria). A P value <0.05 was considered statistically significant.

Results

Overall population

Table 1 shows the differences between derivation (n=1,992) and validation (n=1,119) cohorts. The median follow-up was 59 months (33 to 86). Considering the presence of several significant differences, the search for heterogeneity in demographics, clinical, and radiological features between validation and derivation cohorts was satisfied.

Model selection using the derivation cohort

Table 2 reports the parametric model performances, while **Figure 1 and Supplementary Figure 2,** Supplemental Digital Content 1, http://links.lww.com/SLA/F438 show the risk curve distribution

of parametric models. The model that best fitted the derivation cohort data was the one with the lognormal distribution of the risk change. The lognormal model had the smallest AIC and BIC (2673 and 2718, respectively) and discrimination power of good quality (0.691). The σ value was 1.3 (1.2 to 1.4), suggesting that the BD-IPMN cohort had a high risk of progression during the first five years, but the risk decreased after without returning to zero. Sex, site of lesions, familial pancreatic cancer, and length of the follow-up did not influence the AFT. Factors significantly influencing the AFT were the patient's age and the cyst size. With the increase in age, the time for change occurrence was reduced by 1.7% each year. With the increase in cyst size, the time for change occurrence was reduced by 2% for each year.

Model development and internal validation using the derivation cohort

In **Table 3,** the results of k-fold cross-validation are reported. In each of the five samples, the patient's age at diagnosis and cyst size were always confirmed as the only factors influencing the AFT. Moreover, in each of the five samples, C-statistics were similar in both "in-sample" and "out-of-sample" cohorts, indicating a low risk of overfitting and good discrimination quality. The pooled data from the in-sample cohort suggested that the patient's age and cyst size accelerated the time needed for change. As age increased, the time for change was reduced by 2% (1% to 3%) for each year; as cyst size increased, the time for change was reduced by 2% (0% to 3%) for each mm. The logarithmic transformation of these values and the use of the coefficient permitted the development of the two individual case-predicting algorithms. The standard model calculates the predicted median time for a change (in months) by solving the following equation: $time = e^{(6.2+(size\ in\ mm*-0.02)+(age\ in\ yrs*-0.02))}$. In the conservative model, the effect of age and diameter was emphasized using the 95 CI superior coefficient for the two covariates: $time = e^{(6.2+(size\ in\ mm*-0.03)+(age\ in\ yrs*-0.03))}$. **Figure 2** summarizes conservative prediction models with a contour plot. **Table 4** reports the predicted times for change divided into three categories based on time values in terciles: Q1, individuals with a high risk of change; Q2, subjects with intermediate

risk; Q3, individuals with low risk. In the "Standard" model, patients at high risk of change did so after a median of 80 months (range 46 to 91), while those at intermediate and low risk after a median of 101 months (92 to 114) and 134 months (122 to 217), respectively. Using the "Prudential" model, patients at a high risk of change did so after a median of 33 months (range 14 to 40). Those at intermediate and low risk did so after a median of 46 (40 to 55) and 70 (55 to 200) months, respectively.

External Validation

Figure 3 shows the calibration curves, with the derivation and training models being similar and very close to the reference line, indicating that the observed and predicted risks are very similar. **Table 5** shows the external validation process. First, a naïve model was developed using the training cohort population and regressing a parametric survival model based on lognormal distribution. In this model, the patient's age and cyst size predicted the AFT similarly to that of the derivation cohort: i) increasing age, for each year, the AFT was increased to 2% (1% to 4%); ii) increasing cyst size, for each mm, the AFT was increased of 5% (3% to 7%). The external validity was confirmed by comparing the naïve model with the standard and conservative models. Using the standard model, the patients of the training cohort were classified into three different categories (namely high, intermediate, and low risk). However, discrimination of these categories was not accurate as both intermediate and low-risk patients had AFTs that were not significantly different from those of high-risk patients. The suboptimal performance of the standard model was confirmed by a statistically significant difference with the naïve model two C-statistics (0.712 vs. 0.640; $P<0.001$).

On the contrary, using the conservative algorithm to classify the patients of the training cohort, the three categories well discriminated patients at different progression times. In detail, the time needed to change in patients classified as intermediate and low risk significantly decelerated by 60% and 118% compared to high risk. The conservative algorithm developed on the validation cohort has

the same discrimination power as the model directly created using the original data of the training cohort (0.712 vs. 0.696; P=0.072) and, for this reason, can be considered externally validated. The individual case-predicting algorithms are available on the Excel spreadsheet in beta version (**supplementary material,** Supplemental Digital Content 1, http://links.lww.com/SLA/F438) and at the following website: https://dpcg.nl/pancreascalculator/ after the publication of the manuscript.

Discussion

This international multicenter study provides the first online available model for personalized surveillance intervals in patients with low-risk BD-IPMN. This model determines the optimal surveillance interval based on two straightforward parameters, namely, the patient's age and cyst size. For this reason, the calculator was called PANORAMA ("Personalized Age and Size-based Surveillance for Low-Risk BD-IPMN of the Pancreas"). It was developed and internally validated in a large cohort of tertiary referral Western centers and externally validated in a large cohort of Eastern centers with good discrimination power. The heterogeneity of the two populations regarding demographic, clinical, and radiological factors confirmed the robustness of such a prediction model. However, the model was validated, excluding those BD-IPMNs undergoing rapid evolution. Therefore, it should be used in clinical practice only at centers of expertise for the management of pancreatic diseases and applied after the first six months of prudential observation, as recommended by the Kyoto guidelines[6]. Of note, the PANORAMA use calculator is unsuitable for any BD-IPMN with WFs or HRS.

Due to the high prevalence of BD-IPMN in the general population (nearly 2%)[22], surveillance in patients without WFs or HRS represents a contemporary "hot topic" with considerable implications for healthcare system organization and costs.[23] In the last decade, several studies[1,5] have demonstrated the indolent behavior of most of these cysts. Nonetheless, all available guidelines and related updates [6-8,14] still recommend long-term follow-up in patients fit for surgery because the risk

of progression is not null. These policies contrast the concept of the cost-effectiveness of oncologic screening, as they do not depict the actual risk of degeneration of IPMNs once considered in a dynamic fashion.

The current study uses a parametric survival approach to build a model that predicts the time needed to observe the relevant changes of low-risk BD-IPMNs, namely the development of WF and HRS. The first remarkable observation was that the change risk was not constant over time but decreased after five years, though never returning to zero and having a lognormal distribution of hazard function. This data is coherent with many observations associated with human cancers [24] and clinical observations regarding BD-IPMN progression.[5] Moreover, it confirmed that most low-risk BD-IPMNs did not significantly change within the first five years of observation, and therefore, repeating a six-month / annual follow-up in this time frame could be too tight. In developing the prediction model in the derivation cohort, two main factors were associated with BD-IPMN evolution, namely cyst size and patient's age. The evolution of cyst size over time is a well-known risk factor for malignancy in surgical series of IPMNs. [25] Moreover, Han Y et al. [25] recently confirmed that a large cyst diameter at the baseline increased the risk of WF development during surveillance. In the present study, the time for WFs/HRS occurrence was shorter by 2% (1% to 3%, months) for each mm in cyst size. This observation seems logical because BD-IPMNs can be assimilated to a hollow spheroid internally covered by an epithelium "at risk of change"; therefore, the larger the size, the larger the number of cells at risk.[26-27] Moreover, the size at diagnosis estimates the time already passed from the initial occurrence, even if the cyst was undiagnosed. In other words, the larger the diameter, the older the cyst, thus increasing the risk of change. Also, the effect of age on WFs/HRS development seems to be aligned with epidemiological, clinical, and biological knowledge about BD-IPMNs. Prevalence and incidence studies [28-31] showed that the median size of pancreatic cysts increases with age. Another prevalence study [22] demonstrated that the standardized prevalence of WFs and HRS increases with age. In a molecular study, Omori et al.[32] reported that BD-IPMN and PDAC in the same patient shared several gene mutations,

demonstrating a time-dependent progression from adenoma to carcinoma driven by somatic mutations. Finally, in a simulation model developed by Koopmann et al.[33], the overall risk of PDAC in patients with BD-IPMNs increased with age. In the present effort, the PANORAMA calculator was built by solving the parametric survival equation after k-fold internal validation. In the conservative version, the effects of age and size were emphasized to 95% CI, assuming the worst scenario as the target of surveillance. In other words, the conservative PANORAMA model was based on the worst-case scenario to obtain an interval time that could capture the changes of evolving BD-IPMNs. Predictably, the conservative version of the calculator fits the data of the external validation cohort better than the standard one. The surveillance interval times obtained were sound and in line with the IPMN progression time reported by Koopmann et al.[31.] For instance, the conservative PANORAMA calculator suggests that a 75-year-old patient with a BD-IPMN of 30 mm in size could be followed every two years rather than every 6-12 months as recommended by the current guidelines.[6,14] Moreover, if the same patient has a BD-IPMN of 15 mm, the PANORAMA calculator recommends an observation every three years rather than after 12-18 months [6,14]. These recommendations should, however, be applied with caution and only by clinicians with high expertise in BD-IPMN management. Thus, it is always advisable to maintain a prudential first follow-up six months after the initial diagnosis in all presumed BD-IPMN to allow for an adequate test of time. Thereafter, in the continuing absence of WF and HRS, the application of the PANORAMA calculator might have a beneficial impact on healthcare costs without impacting patient safety.

The present study has several limitations. First, its retrospective nature could produce intrinsic bias. Second, although all involved centers are tertiary hubs for pancreatic cyst observation, detecting some BD-IPMN progressions could be suboptimal. Third, the study covered a long time for enrollment, during which the management of IPMNs evolved, resulting in a potential bias for the analysis. However, the management of low-risk BD-IPMN has only been slightly modified through different updates and versions of the IAP guidelines.[6-8] Fourth, the risk of incorrect stratification of

BD-IPMNs at baseline could generate a bias in interpreting their natural history. All patients with very early progression (6 months) were excluded from the analysis to reduce this phenomenon, prudentially limiting the use of the calculator only after the first follow-up at six months, as recommended by the International guidelines.[6] Another limitation is that the model suggested the interval time for BD-IPMN surveillance at the first diagnosis. However, it did not solve the dilemma of the follow-up modification when the changes were observed. In other words, the interval time should be modified when a WF occurs and surgery is not performed. It should be remembered that the interval time suggested by PANORAMA was valid only for non-evolutive and low-risk cysts. Further studies applying the same AFT model in a population with WFs could clarify this key aspect. Fifth, the analysis did not include as a specific endpoint the occurrence of concomitant PDAC as such entity is not depictable without surgery. However, as previously described,[34] the overall risk of cancer (both concomitant PDAC and malignant IPMN) was very low (<1%) once WF and HRS were lacking. The main strength of the study was that it is based on pragmatical cornerstones: i) few low-risk BD-IPMNs will change, developing WFs or HRS and becoming at higher risk for cancer; ii) an ideal approach to timely capture patients at higher risk of cancer development should be designed.

In conclusion, the conservative PANORAMA calculator accurately predicted the time needed for low-risk BD-IPMNs to evolve. This information allowed to build a novel surveillance schedule for BD-IPMN without WF and HRS at diagnosis based on two pillars, namely age and cyst size. Further multicentric prospective studies are needed to confirm these data and reinforce the safety of this novel tailored surveillance protocol.

Acknowledgements: We thank Carlos Fernández-Del Castillo for providing us with the patient data of the Department of Surgery, Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts.

Ethics statement: This study was approved by the local review board of each participating center and followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. Informed consent policy varied across centers, and in most cases, it was waived, given the use of historic de-identified data.

References

1- Marchegiani G, Pollini T, Andrianello S, et al. Progression vs Cyst Stability of Branch-Duct Intraductal Papillary Mucinous Neoplasms After Observation and Surgery. JAMA Surg. 2021; 156:654-661.

2- Crippa S, Pezzilli R, Bissolati M, et al. Active Surveillance Beyond 5 Years Is Required for Presumed Branch-Duct Intraductal Papillary Mucinous Neoplasms Undergoing Non-Operative Management. Am J Gastroenterol. 2017; 112:1153-1161.

3- Pergolini I, Sahora K, Ferrone CR, et al. Long-term Risk of Pancreatic Malignancy in Patients With Branch Duct Intraductal Papillary Mucinous Neoplasm in a Referral Center. Gastroenterology. 2017;153:1284-1294.e1.

4- Lee BS, Nguyen AK, Tekeste TF, et al. Long-term follow-up of branch-duct intraductal papillary mucinous neoplasms with No change in first 5 Years of diagnosis. Pancreatology. 2021;21:144-154.

5- Marchegiani G, Andrianello S, Pollini T, et al. "Trivial" Cysts Redefine the Risk of Cancer in Presumed Branch-Duct Intraductal Papillary Mucinous Neoplasms of the Pancreas: A Potential Target for Follow-Up Discontinuation? Am J Gastroenterol. 2019; 114:1678-1684

6- Ohtsuka T, Fernandez-Del Castillo C, Furukawa T, et al. International evidence-based Kyoto guidelines for the management of intraductal papillary mucinous neoplasm of the pancreas. Pancreatology. 2023;28:S1424-3903(23)01883-5.

7- Vege SS, Ziring B, Jain R, Moayyedi P; Clinical Guidelines Committee; American Gastroenterology Association. American gastroenterological association institute guideline on the diagnosis and management of asymptomatic neoplastic pancreatic cysts. Gastroenterology. 2015 Apr;148(4):819-823

8- European Study Group on Cystic Tumours of the Pancreas. European evidence-based guidelines on pancreatic cystic neoplasms. Gut. 2018 May;67(5):789-804.

9- Adelian R, Jamali J, Zare N, Ayatollahi SM, Pooladfar GR, Roustaei N. Comparison of Cox's Regression Model and Parametric Models in Evaluating the Prognostic Factors for Survival after Liver Transplantation in Shiraz during 2000-2012. Int J Organ Transplant Med. 2015; 6:119-125.

10- Hosseini Teshnizi S, Taghi Ayatollahi SM. Comparison of Cox Regression and Parametric Models: Application for Assessment of Survival of Pediatric Cases of Acute Leukemia in Southern Iran. Asian Pac J Cancer Prev. 2017; 18:981-985.

11- Gamel JW, Vogel RL. Comparison of parametric and non-parametric survival methods using simulated clinical data. Stat Med. 1997; 16:1629-1643.

12- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. Int J Surg. 2014; 12:1495-1499.

13- Goggins M, Overbeek KA, Brand R, et al. Management of patients with increased risk for familial pancreatic cancer: updated recommendations from the International Cancer of the Pancreas Screening (CAPS) Consortium. Gut. 2020;69:7-17.

14- European Study Group on Cystic Tumours of the Pancreas. European evidence-based guidelines on pancreatic cystic neoplasms. Gut. 2018; 67:789-804.

15- Tanaka M, Fernández-Del Castillo C, Kamisawa T, et al. Revisions of international consensus Fukuoka guidelines for the management of IPMN of the pancreas. Pancreatology. 2017;17(5):738-753.

16- Crowther MJ, Royston P, Clements M. A flexible parametric accelerated failure time model and the extension to time-dependent acceleration factors. Biostatistics. 2023;14;24:811-831.

17- Wigfield PC, Heeg B, Ouwens M. Improved estimation of overall survival and progression-free survival for state transition modeling. J Comp Eff Res. 2024;13(1):e230031.

18- Harrell Jr FE, Kerry LEE, Daniel MB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15: 361-387.

19- Steyerberg EW , Harrell FEJr , Borsboom GJ et al. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001; 54: 774–781

20- Ramspek CL, Jager KJ, Dekker FW, Zoccali C, van Diepen M. External validation of prognostic models: what, why, how, when and where? Clin Kidney J. 2020 Nov 24;14(1):49-58.

21- Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013; 13: 33

22- Ricci C, Migliori M, Imbrogno A, et al. Prevalence of Asymptomatic Intraductal Papillary Mucinous Neoplasms in Healthy and Ill Populations Detected by Ultrasonography: A Single-Center Study of 6353 Outpatients. Pancreas. 2019;48:113-120.

23- Budde C, Beyer G, Kühn JP, Lerch MM, Mayerle J. The Clinical and Socio-Economic Relevance of Increased IPMN Detection Rates and Management Choices. Viszeralmedizin. 2015;31(1):47-52

24- Abadi A, Amanpour F, Bajdik C, Yavari P. Breast Cancer Survival Analysis: Applying the Generalized Gamma Distribution under Different Conditions of the Proportional Hazards and Accelerated Failure Time Assumptions. Int J Prev Med. 2012;3:644-651.

25- Han Y, Kwon W, Lee M, et al. Optimal Surveillance Interval of Branch Duct Intraductal Papillary Mucinous Neoplasm of the Pancreas. JAMA Surg. 2024 Jan 17:e237010.

26- Han Y, Lee H, Kang JS, et al.. Progression of Pancreatic Branch Duct Intraductal Papillary Mucinous Neoplasm Associates With Cyst Size. Gastroenterology. 2018;154:576-584.

27- Han Y, Jang JY, Oh MY, Kim HS, Byun Y, Kang JS, Kim SH, Lee KB, Kim H, Kwon W. Natural history and optimal treatment strategy of intraductal papillary mucinous neoplasm of the pancreas: Analysis using a nomogram and Markov decision model. J Hepatobiliary Pancreat Sci. 2021;28:131-142.

28- Ricci C, Casadei R, Taffurelli G, et al. Risk Factors for Malignancy of Branch-Duct Intraductal Papillary Mucinous Neoplasms: A Critical Evaluation of the Fukuoka Guidelines With a Systematic Review and Meta-analysis. Pancreas. 2016;45:1243-1254.

29- Chang YR, Park JK, Jang JY, Kwon W, Yoon JH, Kim SW. Incidental pancreatic cystic neoplasms in an asymptomatic
healthy population of 21,745 individuals: large-scale, single center cohort study. Medicine (Baltimore). 2016;95(51):e5535.

30- Chang YR, Park JK, Jang JY, Kwon W, Yoon JH, Kim SW. Incidental pancreatic cystic neoplasms in an asymptomatic healthy population of 21,745 individuals: large-scale, singlecenter cohort study. Medicine (Baltimore). 2016;95(51):e5535

31- Schweber AB, Agarunov E, Brooks C, Hur C, Gonda TA. Prevalence, incidence, and risk of progression of asymptomatic pancreatic cysts in large sample real-world data. Pancreas. 2021;50(9):1287–92.

32- Omori Y, Ono Y, Tanino M, et al. Pathways of Progression From Intraductal Papillary Mucinous Neoplasm to Pancreatic Ductal Adenocarcinoma Based on Molecular Features. Gastroenterology. 2019;156(3):647-661.e2.

33- Koopmann BDM, Dunnewind N, van Duuren LA, et al. The Natural Disease Course of Pancreatic Cyst-Associated Neoplasia, Dysplasia, and Ductal Adenocarcinoma: Results of a Microsimulation Model. Gastroenterology. 2023;165(6):1522-1532.

34- Marchegiani G, Pollini T, Burelli A, Han Y, Jung HS, Kwon W, Rocha Castellanos DM, Crippa S, Belfiori G, Arcidiacono PG, Capurso G, Apadula L, Zaccari P, Noia JL, Gorris M, Busch O, Ponweera A, Mann K, Demir IE, Phillip V, Ahmad N, Hackert T, Heckler M,

Lennon AM, Afghani E, Vallicella D, Dall'Olio T, Nepi A, Vollmer CM, Friess H, Ghaneh P, Besselink M, Falconi M, Bassi C, Goh BK, Jang JY, Fernández-Del Castillo C, Salvia R. Surveillance for Presumed BD-IPMN of the Pancreas: Stability, Size, and Age Identify Targets for Discontinuation. Gastroenterology. 2023 Oct;165(4):1016-1024.e5.

Figure 1-: Hazard function

Legend: red = lognormal distribution; blue= logistic distribution; green = exponential distribution; orange= Weibull distribution.

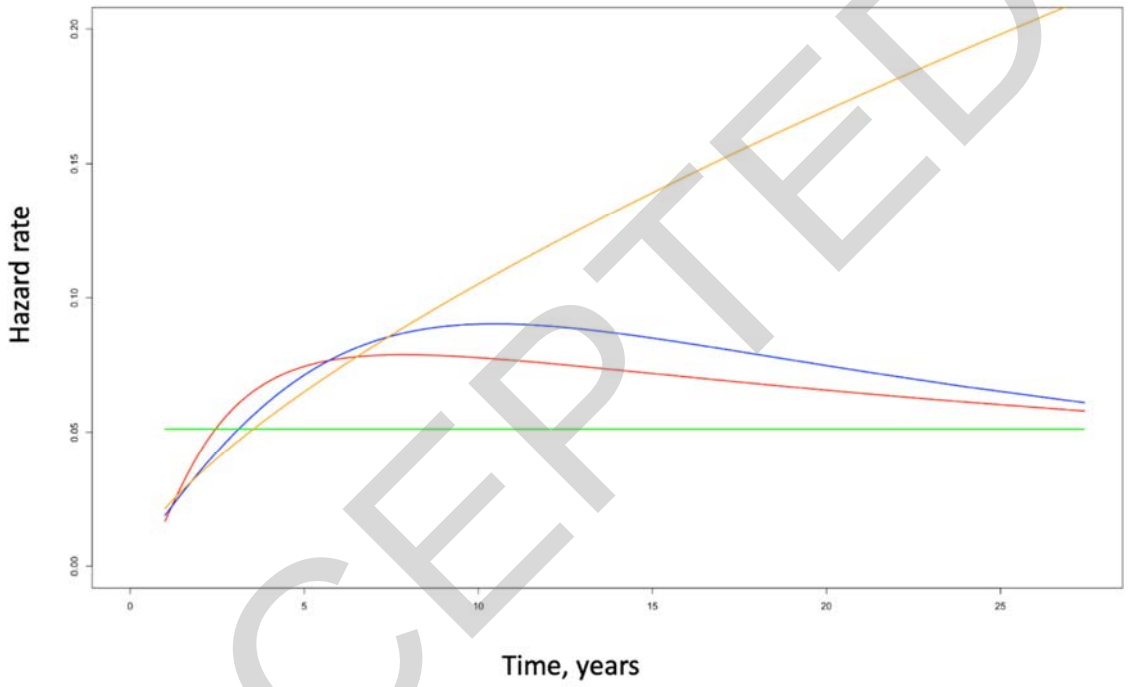


Figure 2 –Contour plot of the conservative model.

Legend: Time needed to observe a relevant change, according to variations in patients 'age and cyst's diameter. Individual time is accessible for any given patient at https://dpcg.nl/pancreascalculator/. The contour plots are derived from the following equations:

$$time = e^{(6.2+(size\ in\ mm*-0.03)+(age\ in\ yrs*-0.03))}$$

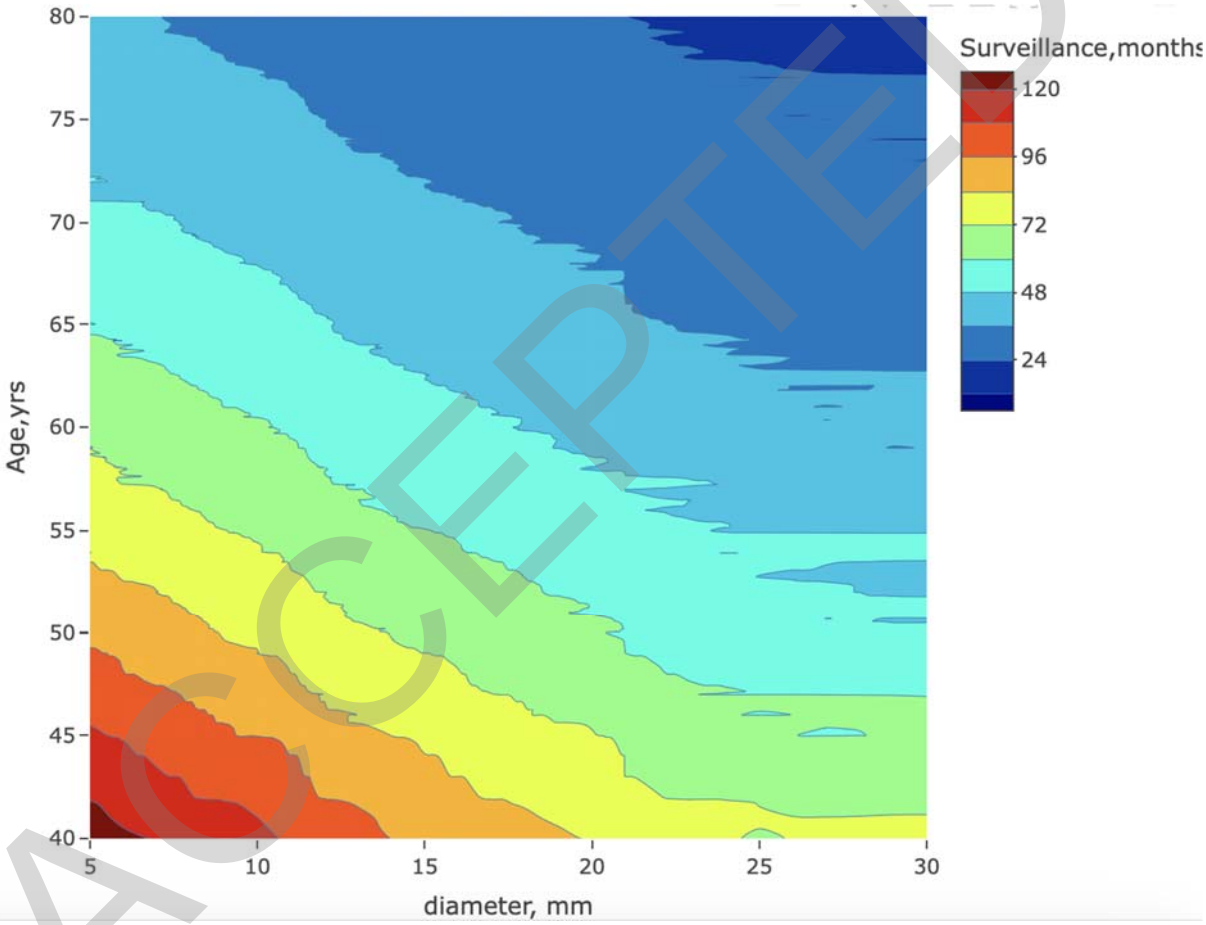
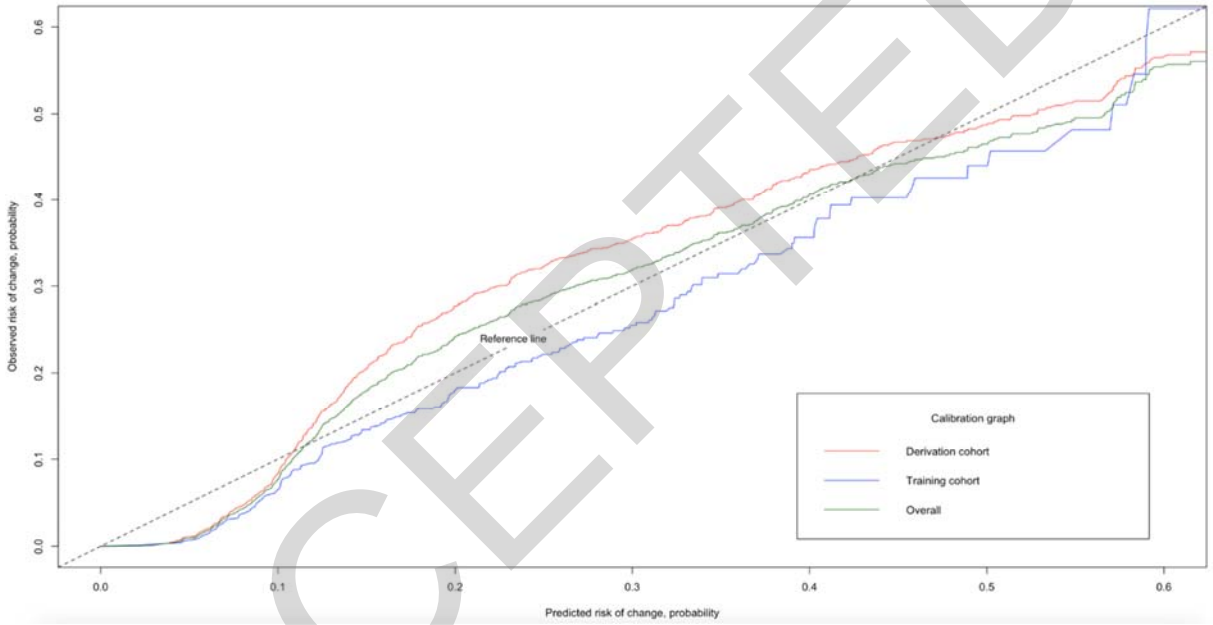


Figure 3 Calibration curve

Legend: The Y-axis represents the observed risk of relevant change; the X-axis represents the predicted pertinent change risk; the dotted line at 45 degrees indicates perfect calibration, as expected and observed probabilities are equal.



Characteristics of patients	Derivation cohort§ n=1,992	Validation cohort° N=1,119	P value
	n (%) or Median (IQR)		
***Age**, years*	66 (58-73)	66 (59-73)	0.177
Sex			
Male	809 (40.6)	507 (45.3)	0.011
Female	1,183 (59.3)	612 (54.7)	
FPC			
No	1,869 (93.8)	1,110 (99.2)	<0.001
Yes	123 (6.2)	9 (0.8)	
Size of largest cyst, mm	12 (8 to 17)	11 (7 to 15)	<0.001
Site of lesion			
Head	507 (25.5)	398 (35.6)	<0.001
Body	390 (19.6)	425 (38)	
Tail	126 (6.3)	90 (8)	
Multiple	969 (48.6)	206 (18.4)	
Observation time, months	55 (30-93)	62 (36-79)	0.007
Relevant change during surveillance^			
No	1,434 (72)	926 (82.7)	<0.001
Yes	558 (28)	193 (17.3)	
HRs during surveillance			
No	1,974 (99.1)	1,093 (97.7)	0.002
Yes	18 (0.9)	26 (2.3)	
WFs during surveillance			
No	1,452 (72.9)	952 (85)	<0.001
Yes	540 (27.1)	167 (15)	
Surgical resection			
No	1,896 (95.2)	1,090 (97.7)	0.001
Yes	95 (4.8)	26 (2.3)	
PDAC occurrence			
No	1,972 (99)	1,114 (99.6)	0.141
Yes	20 (1)	5 (0.4)	

Table 1- Characteristics and outcomes of 3,111 patients with BD-IPMN without worrisome features nor high-risk stigmata included in the study.

Legend: IQR=Interquartile rage; FPC= familiar Pancreatic Cancer; ^= worrisome features or high-risk stigmata; WF= Worrisome features; HR= High-risk stigmata. §= Western centers; °= Eastern centers.

Table 2- Parametric distributions of failure curve for BD-IPMNs in the derivation cohort

Parameters	Distributions °			
	Weibull (AFT)	**Exponential (AFT)**	**Lognormal (AFT)**	**Loglogistic (AFT)**
Age (for each year)	1.7% (1.0% to 2.5%)	1.9% (1% to 2.6%)	1.7% (0.9% to 2.5%)	1.7% (1% to 2.6%)
Cyst size (for each mm)	2% (0.7% to 3.1%)	2.2% (0.9% to 3.4%)	2% (1% to 3.5%)	1.9% (0.8% to 3.2%)
Ancillary parameter^	$p = 1.1$ (1.0 to 1.2)	$p = 1$	$\sigma = 1.3$ (1.2 to 1.4)	$\gamma = 0.7$ (0.7 to 0.8)
Goodness of fit (AIC, BIC)	2705, 2750	2712, 2751	2673,2718	2688,2733
C statistic	0.691 (0.669 to 0.714)	0.692 (0.669 to 0.714)	0.691 (0.688 to 0.713)	0.692 (0.670 to 0.715)

Legend: °= sex, site of lesions, familial pancreatic cancer, and length of the follow-up did not influence the scale of failure function. AFT = Accelerated failure time reported as a percentage. The exponential transformation of the coefficient permits the obtaining of this value. ^= parameter which determines the shape of the risk curve; AIC= Akaike information criterion; Bayesian information criterion.

Table 3- Modelling for Accelerated Failure Time (AFT) in BD-IPMNs using the derivation cohort.

	k-fold cross validation					
	1	**2**	**3**	**4**	**5**	**Pooled (95% CI)**
Training set, n	1547	1574	1582	1537	1590	
Age (for each year)	2% (1% to 3%)	1% (0% to 2%)	2% (1% to 2%)	2% (1% to 3%)	2% (1% to 3%)	2% (1% to 3%)
Larger cyst diameter (for each mm)	2% (1% to 4%)	2% (1% to 4%)	2% (0% to 3%)	2% (0% to 3%)	2% (0.1% to 3%)	2% (0% to 3%)
Constant	6.3	5.9	6.3	6.2	6.1	6.2
C statistics (95% CI)	0.699 (0.674 to 0.724)	0.674 (0.649 to 0.701)	0.689 (0.664 to 0.713)	0.691 (0.665 to 0.717)	0.684 (0.659 to 0.710)	0.689 (0.668 to 0.718)
Test set, n	444	418	409	455	402	
C statistics (95% CI)	0.653 (0.602 to 0.704)	0.741 (0.697 to 0.786)	0.694 (0.644 to 0.744)	0.694 (0.649 to 0.740)	0.702 (0.654 to 0.752)	0.702 (0.654 to 0.752)

Legend: The time for change can be obtained by solving the following equation: time= EXP (6.2+ (size in mm *-0,02) + (age in years* -0,02)); in the conservative scenario, the 95 CI superior CI can be used time= EXP (6.2+ (size in mm *-0,03) + (age in years* -0,03)) emphasizing the effect of age and cyst diameter.

Table 4- Predicted interval time calculated in the derivation cohort reported in three risk categories.

	Q1 – High risk (n=702)	**Q2 – Intermediate risk (n=626)**	**Q3 – Low risk (n=624)**
PANORAMA "Standard" Model	80 (46 to 91) months	101 (92 to 114) months	134 (122 to 217) months
PANORAMA "Conservative" Model	33 (14 to 40) months	46 (40 to 55) months	70 (55 to 200) months

Legend: PANORAMA= Personalized Age and Size-based Surveillance for Low-Risk BD-IPMN of the Pancreas; The interval risk categories were obtained by solving the two-equations and dividing the predicted interval time in terciles (Q1 <33%; Q2=33-66%; Q3>66%). Within each category, the time for change was reported with range.

Table 5- Algorithms for case-predictions interval time fitted in the training cohort.

Parameters	Naive model (AFT)*	Standard PANORAMA model (AFT)^	Conservative PANORAMA model (AFT)^
High risk category	-	Referent	Referent
Intermediate risk category	-	-16% (-213% to 56%)	-60% (-12% to -129%)
Low risk category	-	-90% (-858% to 62%)	-118% (-49% to -314%)
Age	2% (1% to 4%)	-	-
Diameter	5% (3% to 7%)	-	-
Goodness of fit (AIC, BIC)	1152, 1177	1179, 1204	1162, 1187
C statistic; P-value §	0.712 (0.678 to 0.747); referent	0.640 (0.622 to 659); P<0.001	0.696 (0.664 to 0.728); P=0.072

Legend: AFT = Accelerated failure time reported in percentage; *= The naive prediction model was built using the age and diameter of the training cohort to calculate AFT; PANORAMA= Personalized Age and Size-based Surveillance for Low-Risk BD-IPMN of the Pancreas; ^= Standard and Prudential were developed using the equations developed in the derivation cohort, and these equations were used to classify the patients of the training cohort in high, intermediate, and low risk of progression; AIC= Akaike information criterion; Bayesian information criterion; §= Harrel C statistic indicated the discrimination power and the referent was the naïve model.