



# The milestone for preventing post-ERCP pancreatitis using novel simplified predictive scoring system: a propensity score analysis

Masafumi Chiba<sup>1</sup> · Masayuki Kato<sup>1</sup> · Yuji Kinoshita<sup>2</sup> · Nana Shimamoto<sup>1</sup> · Youichi Tomita<sup>2</sup> · Takahiro Abe<sup>1</sup> · Keisuke Kanazawa<sup>1</sup> · Shintaro Tsukinaga<sup>1</sup> · Masanori Nakano<sup>2</sup> · Yuichi Torisu<sup>2</sup> · Hirobumi Toyoizumi<sup>1</sup> · Kazuki Sumiyama<sup>1</sup>

Received: 7 August 2020 / Accepted: 15 November 2020 / Published online: 30 November 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

**Background** Post-ERCP pancreatitis (PEP) with trans-papillary approach remains a major issue, and the multi-factorial etiology can lead to the development of unpredictable PEP. Therefore, the early identification of PEP is highly desirable to assist with the health cost containment, the reduction in unnecessary admissions, earlier appropriate primary care, and intensive care for preventing progression of severe pancreatitis. This study aimed to establish a simplified predictive scoring system for PEP.

**Methods** Between January 1, 2012, and December 31, 2019, 3362 consecutive trans-papillary ERCP procedures were retrospectively analyzed. Significant risk factors were extracted by univariate, multivariate, and propensity score analyses, and the probability of PEP in the combinations of each factor were quantified using propensity score analysis. The results were internally validated using bootstrapping resampling.

**Results** In the scoring system with four stratifications using combinations of only five extracted risk factors, the very high-risk group showed 28.79% (95% confidence interval [CI], 18.30%–41.25%;  $P < 0.001$ ) in the predicted incidence rate of PEP, and 9.09% (95% CI, 3.41%–18.74%;  $P < 0.001$ ) in that of severe PEP; although the adjusted prevalence revealed 3.74% in PEP and 0.90% in severe PEP, respectively. The prediction model had an area under the curve of 0.86 (95% CI, 0.82–0.89) and the optimism-corrected model as an internal validation had an area under the curve of 0.81 (95% CI, 0.77–0.86).

**Conclusions** We established and validated a simplified predictive scoring system for PEP using five risk factors immediately after ERCP to assist with the early identification of PEP.

**Keywords** Endoscopic retrograde cholangiopancreatography (ERCP) · Post-ERCP pancreatitis (PEP) · Predictive scoring system · Propensity score analysis · Internal validation

Post-endoscopic retrograde cholangiopancreatography (ERCP) pancreatitis (PEP) remains a major issue with the trans-papillary approach, in which the incidence rate of PEP

has been shown to be 3.5%–9.7% and the mortality rate has been shown to be 0.1%–0.7% [1].

To date, several patient- or procedure-related risk factors of PEP have been demonstrated, while the use of pancreatic stents and rectal nonsteroidal anti-inflammatory drugs (NSAIDs) have been shown to prevent PEP [2, 3]. Although these prophylactic methods are helpful for the prevention of PEP, they cannot completely prevent PEP due to the complexities associated with trans-papillary ERCP (e.g., selection of cannulation methods, unintentional pancreatic injections, or guidewire pass) [3, 4]. Invariable factors, such as young age and difficult cannulation arising from personal anatomy, are other unavoidable issues [2–4]. As a result, this multi-factorial etiology makes the prediction of PEP difficult [4, 5].

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00464-020-08173-4>) contains supplementary material, which is available to authorized users.

✉ Masafumi Chiba  
cc10972009720@gmail.com

<sup>1</sup> Department of Endoscopy, The Jikei University School of Medicine, 3-25-8, Nishi-Shimbashi, Minato-ku, Tokyo 105-8461, Japan

<sup>2</sup> Division of Gastroenterology and Hepatology, Department of Internal Medicine, The Jikei University School of Medicine, Tokyo, Japan

Therefore, if a simple prediction model for PEP is to be feasible immediately after ERCP, earlier distinction between low- and high-risk patients with PEP may assist with the health cost containment and the reduction in unnecessary admissions [4, 6, 7]. Moreover, earlier appropriate primary care and intensive care can prevent progression of severe pancreatitis [8–10]. However, a simplified and practical prediction system for PEP is not yet available, and early identification of PEP remains a challenge in this field because only few studies on prediction models for PEP immediately after ERCP have been reported even though there are many reports on the risk factors associated with PEP [3, 5, 11–13]. Therefore, the aim of this study was to establish a simple predictive scoring system for PEP immediately after ERCP.

## Methods

### Study design and patients

The present study was a retrospective single-center cohort study including consecutive patients with suspected hepatobiliary-pancreatic disorders who underwent trans-papillary ERCP attempts between January 1, 2012 and December 31, 2019. Data on ERCP procedures were retrieved from the Jikei University ERCP database and medical records. The database was updated immediately after each procedure and contained data of > 3 months of follow up. Written informed consent to undergo ERCP was obtained from all patients prior to participation. The opportunity to opt out of this study participation was also provided, and the requirement for informed consent was waived due to the anonymous retrospective observational study (opt-out method of informed consent).

All participants who underwent trans-papillary ERCP were enrolled in this trial. Patients with no papilla (e.g., pancreaticoduodenectomy, Roux-en-Y with hepaticojejunostomy), and those with ongoing acute pancreatitis were excluded. Patients in whom the target papilla could not be reached and discontinued (e.g., gastrointestinal tract stenosis and food residue at endoscopy), and patients who only underwent stent removal without ERCP were also excluded. This study complied with the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) statement [14], and was approved by the Human Subjects Committee of Jikei University School of Medicine [ID no. 32-027 (10102)]. This study was also subsequently registered with the University Hospital Medical Information Network Clinical Trials Registry (identification no. UMIN000040471), and was conducted in accordance with the ethical principles of the Declaration of Helsinki (Fortaleza revision).

### ERCP procedures

ERCP was performed under fluoroscopic view by seven experts who had experience of > 300 ERCPs with naïve papilla, or by 16 trainees who had experience of < 300 with experts' interference, depending on the situation [15, 16]. Side-viewing duodenoscopes (JF-260, TJF-260V; Olympus Medical Systems, Tokyo, Japan) were used in most cases (e.g. normal anatomy, Billroth I). Forward-viewing scopes, including balloon assisted endoscope and/or cholangioscope, were used in some cases (e.g. Billroth II, Roux-en-Y with gastrectomy) depending on the gastrointestinal anatomy, or when both diagnosis and treatment of pancreaticobiliary conditions were required (Supplementary Tables 1, 2). Selective pancreatic- or biliary-cannulation methods of contrast-assisted, single-guidewire assisted (SGW), pancreatic guidewire assisted (PGW), double-guidewire assisted (DGW), rendezvous, or crossover was chosen depending on the situation. PGW technique is defined as the placement of a pancreatic guidewire to facilitate biliary cannulation using only contrast-assisted method [17]. DGW technique is defined as the combination of both pancreatic duct guidewire placement and SGW to achieve selective biliary cannulation without contrast injection [18]. Pancreatic stents of 5Fr diameter, or occasional endoscopic nasopancreatic drainage of that diameter, were also used in cases of easy pancreatic stenting.

Patients > 13 years old who underwent ERCP were conscious but sedated with intravenous midazolam and pethidine administration, and those < 12 years old were administered general anesthesia during the ERCP procedure.

### Definitions of post-ERCP pancreatitis and severity

Diagnosis of PEP complied with revision of the Atlanta classification [1], while severity of PEP complied with the Japanese Guidelines for the Management of Acute Pancreatitis 2015 [19]. We used the Japanese Guidelines for the following reasons. First, the majority of patients who undergo ERCP are admitted to hospital in Japan. Second, initial endoscopic biliary drainage was usually performed for patients with acute cholangitis or obstructive jaundice. Third, the subsequent endoscopic treatments (e.g., stone extraction with endoscopic sphincterotomy and/or endoscopic papillary balloon dilatation) or surgeries for etiology were continuously performed, thus, maintaining the admission in Japan [20]. Thus, the number of prolonged hospitalizations ( $\geq 10$  days) because of these processes facilitate severe PEP in patients with mild PEP, if cotton criteria or other classification are applied [21, 22]. Therefore, we used the Japanese Guidelines.

## Selection of PEP risk factors

Univariable analysis was performed to select 35 potential predictor variables for PEP based on previous research, including ASGE and ESGE risk factors; however, our database did not include the total number of attempts on the papilla (Supplementary Table 1, 2) [2, 3]. The significant factors ( $P < 0.05$ ) from univariate analysis were adopted into multivariate analysis in order to narrow the selected variables. In the multivariate analysis, to avoid imbalances when there were seven or fewer events per confounder, the number of dependent variables was adopted within 15.4 (108/7) items [23]. Variables with a  $P$ -value  $< 0.10$  were included but those with a  $P$ -value  $> 0.10$  were removed. To adjust for selection bias and to reduce potential confounders, we used three types of propensity score analysis: The regression adjustment method, the inverse probability of treatment weighting (IPTW), and one-to-one propensity score matching. These three methods were used to verify the significance and robustness of each risk factor with the same direction [24]. In the three propensity score analyses, one outcome was defined as PEP incidence based on revision of the Atlanta classification [1], and the other was defined as PEP score for the severity based on prognosis factors (0 to 9 pts) plus CT grade (Grade 1 = 1 pts, Grade 2 = 2 pts, Grade 3 = 3 pts) from the Japanese Guidelines for the Management of Acute Pancreatitis 2015 [19]. Finally, when each exposure had one more significance ( $P < 0.05$ ) among the six propensity score analyses, the significant factors were employed in the final predictive scoring system.

## Development of predictive scoring system for PEP and internal validation

To establish a predictive scoring system, the probability of PEP was quantified using a propensity score analysis in combinations of the employed factors. The probability of PEP was also calculated using the ASGE's 12 risk factors and the ESGE definite 7 risk factors adding the risk factor of "Absence of pancreatic stent" (Table 1), because each guideline recommends use of pancreatic stents to prevent PEP [2, 3]. As for discrimination, the area under the receiver operating characteristic (AUC) curves was calculated and compared for PEP prediction in each model. Subsequently, four thresholds were set for stratification based on the range of propensity score, as follows: low-, moderate-, high-, and very high-risk groups, which were unrelated to total number of risk factors. Calibration of the predicted scoring model was verified for both PEP and severe PEP by calculating the incidence rate and odds ratio. Pearson goodness-of-fit test ( $P > 0.05$  acceptable) was also performed. Furthermore, bootstrap resampling analysis (i.e., 1000 bootstrap resamples) was performed as an internal validation to evaluate

the optimism of the present model [25]. The calibration was calculated by an optimism adjusted frequency of PEP, and the discrimination was calculated by the receiver operating characteristic curve and AUC.

## Statistical analysis

In multivariate analysis, variables with multicollinearity (variance inflation factor  $> 10$ ), non-observation in PEP, and a perfect prediction arising from few numbers were excluded. One-way analysis of variance with a Bonferroni correction was used for multiple comparisons of the AUCs for PEP. Missing values were excluded for complete case analysis. The standardized mean differences ( $< 0.1$ ) were assessed to balance baseline characteristics in the regression adjustment by propensity score analysis and in the IPTW analysis [24]. To adjust for differences in patient-related and procedure-related covariate, one-to-one propensity score matching was employed using calipers with a width of 0.2 standard deviations. The covariates were all employed from patient-related and procedure-related risk factors (Supplementary Tables 1 and 2), which were known as risk factors of PEP in previous studies [2, 3]. However, non-observation covariates in the PEP group, covariates after ERCP allocation (e.g., prevention drug for PEP), and multicollinearity variables (variance inflation factor  $> 10$ ) were excluded. Data are presented as mean (standard deviation, SD) or frequencies (%), as appropriate. The proportions of categorical variables were compared using the Chi-square test or Fisher's exact test (e.g., analysis of PEP ratio after propensity score matching). The mean values of the continuous variables were compared using the Mann–Whitney  $U$ -test (e.g., propensity score matching PEP score). Two-sided  $P$ -values  $< 0.05$  were considered significant. All analyses were performed using Stata version 15 (StataCorp LP; College Station, TX, USA).

## Results

### Patients, procedures, and post-ERCP pancreatitis

Following exclusion of non-papillary ERCP procedures, 3,362 eligible patients who underwent trans-papillary ERCP procedures were enrolled in this study (Fig. 1). The raw incidence rates of PEP and severe PEP were 3.2% (108/3362) and 0.77% (26/3362), respectively (Table 1). In terms of the papilla condition, PEP occurred in 96 cases (88.9%) with naïve papilla, and in 12 cases (11.1%) without naïve papilla (Table 1, Supplementary Table 1). The overall success rate was 96.9% (3160/3261) in selective CBD cannulation and 97.0% (98/101) in intentional selective pancreatic ductal cannulation. The final diagnosis and indications of ERCP

**Table 1** Patient's characteristics and univariate analysis to select the risk factors for PEP

Independent variable	PEP group (n = 108)	Non-PEP group (n = 3254)			OR (95%CI)	p value
<b>Patient-related risk factors for PEP</b>						
Mean (range) age (years)	66.7 (24–91)	69.7 (5–99)			–	–
No. age < 60 <sup>a</sup>	31 (33.3)	678 (20.8)			1.67 (1.01–2.77)	0.045
No. age < 40 <sup>a,b</sup>	5 (4.6)	116 (3.6)			0.86 (0.21–3.53)	0.84
No. female <sup>b,c</sup>	39 (36.1)	1143 (35.1)			1.04 (0.70–1.56)	0.83
Absence of chronic pancreatitis <sup>b</sup>	105 (97.2)	3154 (96.9)			1.11 (0.35–3.56)	0.86
Previous pancreatitis <sup>b,c</sup>	2 (1.9)	184 (5.7)			0.32 (0.77–1.29)	0.11
Previous PEP <sup>b,c</sup>	6 (5.6)	166 (5.1)			1.09 (0.47–2.53)	0.83
Normal serum bilirubin (< 2.0 mg/dL) <sup>b,d</sup>	66 (61.1)	2067 (63.5)			0.90 (0.61–1.34)	0.61
Clinical suspicious of SOD <sup>b,c</sup>	2 (1.9)	8 (0.3)			7.66 (1.61–36.49)	0.011
Naïve papilla <sup>a</sup>	96 (88.9)	1402 (43.1)			12.37 (5.72–26.75)	< 0.001
<b>Procedure-related risk factors</b>						
SIF H290S/single balloon short type <sup>a</sup>	1 (0.9)	1 (0.03)			29.26 (1.82–471.8)	0.017
EST Knife for cannulation only <sup>a</sup>	25 (23.2)	260 (8.0)			3.34 (2.09–5.34)	< 0.001
SGW <sup>a</sup>	20 (18.5)	234 (7.2)			2.34 (1.37–3.99)	0.002
PGW <sup>a</sup>	42 (38.9)	211 (6.5)			7.53 (4.70–12.06)	< 0.001
Rendezvous method <sup>a</sup>	1 (0.9)	5 (0.2)			10.17 (1.17–88.41)	0.036
Precut sphincterotomy <sup>a,b</sup>	10 (9.3)	34 (1.0)			1.70 (0.74–3.87)	0.21
Repetitive pancreatic guide wire pass <sup>b,c</sup>	1 (0.9)	25 (0.8)			1.21 (0.16–8.99)	0.85
EPLBD(> 12 mm) <sup>b</sup>	0 (0.0)	11 (0.3)			N.A	N0.A
Fully-covered SEMS across papilla <sup>a</sup>	2 (1.9)	4 (0.1)			14.71 (2.66–81.25)	0.002
Expert only (Naïve papilla ≥ 300)	15 (13.9)	478 (14.7)			Reference	–
Trainee only (Naïve papilla < 300)	41 (38.0)	2167 (66.6)			0.54 (0.29–0.99)	0.045
Substitute for expert	52 (51.0)	609 (19.3)			2.71 (1.51–4.86)	0.001
Difficult cannulation (> 5 min) <sup>c</sup>	87 (80.6)	1087 (34.9)			Reference	–
Difficult cannulation (> 10 min) <sup>b</sup>	69 (63.9)	519 (16.7)			2.54 (1.17–5.51)	0.018
Difficult cannulation (> 15 min)	61 (56.5)	310 (10.0)			5.14 (2.41–10.97)	< 0.001
Total procedure time (> 30 min)	61 (57.0)	826 (25.6)			3.86 (2.61–5.70)	< 0.001
Pancreatic injections (≥ 1) <sup>b,c</sup>	95 (88.0)	1027 (35.8)			13.13 (7.32–23.56)	< 0.001
Failure of selective CBD cannulation	11 (10.3)	90 (2.9)			3.90 (2.02–7.53)	< 0.001
<b>Prophylaxis-related factors</b>						
Absence of pancreatic stent	88 (81.5)	3087 (94.9)			0.24 (0.14–0.37)	< 0.001
<b>Other information</b>						
30-days mortality	1 (0.9)				52 (1.6)	
Post ERCP hyperamylasemia	108 (100)				213 (6.6)	
Incidence rate of all PEP	108 (3.21)				–	
Incidence rate of severe PEP	26 (0.77)				–	
<b>Prognosis factors of PEP</b>						
	0 pt	1 pt	2 pt	3 pt	4 pt	≥ 5pt
	43 (39.8)	43 (39.8)	15 (13.9)	6 (5.6)	1 (0.9)	0 (0)
<b>CT grade of PEP</b>						
	Grade 1	Grade 2			Grade 3	
	85 (78.7)	23 (21.3)			0 (0.0)	

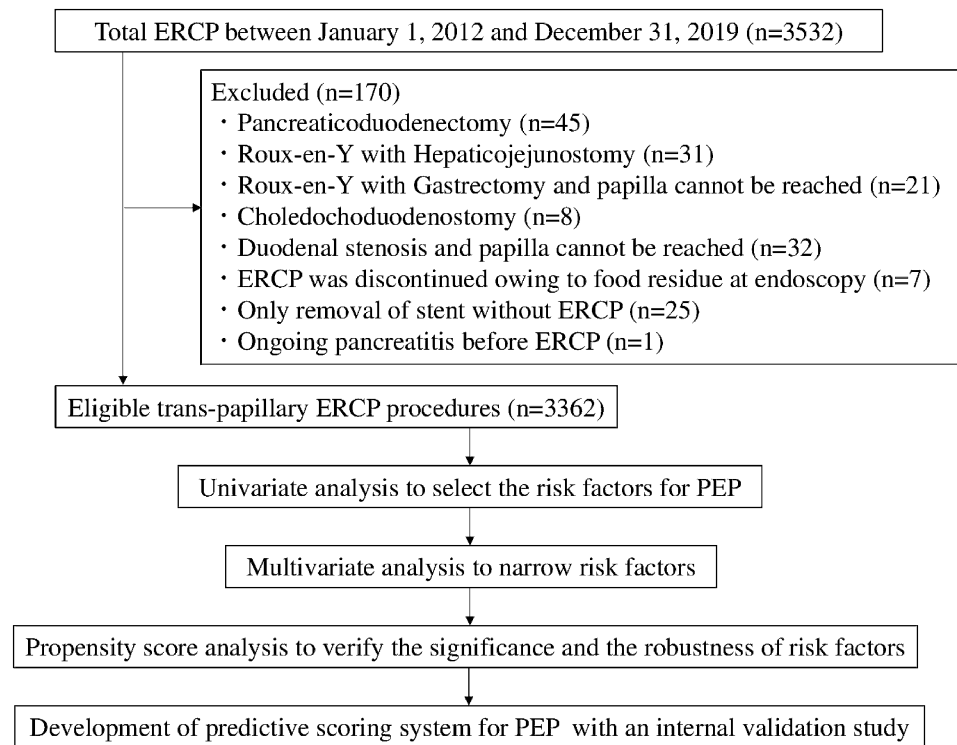
Unless indicated otherwise, data are presented as n (%). Of note, percentages may not add up to 100% because of rounding or missing values

<sup>a</sup>Complete version of detail including each reference variable is available in Supplemental Tables 1 and 2

<sup>b</sup>Twelve risk factors of ASGE guideline 2017 for PEP

<sup>c</sup>Seven definite risk factors of ESGE guideline 2020 for PEP

**Fig. 1** Patient information and the flow of study



were bile duct stone (42.9%), malignant biliary stricture (33.0%), and benign biliary non-stricture (12.5%) (Supplementary Table 1). The frequency of hyperamylasemia without PEP was 6.6% (213/3254) in the non-PEP group (Table 1), and no PEP-related 30 day-mortality occurred in this series.

### Univariate analysis to select risk factors for PEP

Univariate analysis revealed significant odds ratios in patient- and procedure-related risk factors (Table 1, Supplementary Tables 1 and 2). However, more than half of the 12 risk factors of ASGE showed no statistical significance [Age < 40, female, absence of chronic pancreatitis, previous pancreatitis, previous PEP, normal serum bilirubin, precut sphincterotomy, repetitive pancreatic guide wire pass, and EPLBD (> 12 mm)]. Moreover, 4 of the 7 definitive risk factors of ESGE showed no statistical difference (female, previous pancreatitis, previous PEP, and repetitive pancreatic guide wire pass) (Table 1, Supplementary Tables 1 and 2) [2, 3]. In particular, pancreatic guidewire-assisted cannulation (PGW) showed a significant association with PEP, as did single-guidewire assisted cannulation (SGW) and the rendezvous method (Table 1, Supplementary Table 2). However, there was no significant difference in number of PEP between PGW with pancreatic stent ( $n = 18$ ) and PGW

without pancreatic stent ( $n = 24$ ) ( $P = 0.25$ ) (Supplementary Table 2).

With regards to preventative drugs for PEP, no drug revealed significance at the stage of univariate analysis. A small number of nafamostat mesilate and aggressive hydration with lactated Ringer's solution was observed in the PEP and/or non-PEP groups, respectively (Supplementary Table 2).

### Multivariate analysis to narrow risk factors

Three factors showed significance for PEP: Naïve papilla, difficult cannulation (> 15 min), and pancreatic injections ( $\geq 1$ ) (Table 2). PGW (OR, 1.83; 95% CI, 0.99–3.38;  $P = 0.052$ ) and absence of pancreatic stent (OR, 1.74; 95% CI, 0.90–3.35;  $P = 0.098$ ), showed odds ratios with borderline significance ( $P < 0.10$ ) (Table 2).

### Propensity score analysis to verify the significance of risk factors

Of five candidate risk factors for the incidence of PEP, all exposures without absence of pancreatic stent showed statistical significance in adjusted logistic regression (Table 3). Furthermore, exposures of PGW and pancreatic injections ( $\geq 1$ ) showed statistical significance in the IPTW and in one-to-one propensity score matching analysis (Table 3).

**Table 2** Multiple logistic regression to narrow risk factors

Variables	Adjusted OR	95% CI	<i>p</i> Value
Patient-related risk factors			
Age < 60	1.41	0.87–2.27	0.16
Clinical suspicious of SOD	2.75	0.47–15.86	0.26
<b>Sig1. Naïve papilla</b>	<b>3.12</b>	<b>1.58–6.18</b>	<b>0.001</b>
Procedure-related risk factors			
EST Knife for cannulation only	1.20	0.69–2.11	0.52
SGW	1.62	0.88–2.97	0.12
<b>*PGW</b>	<b>1.83</b>	<b>0.99–3.38</b>	<b>0.052</b>
Rendezvous method	8.60	0.63–117.83	0.11
Biliary brushing cytology	1.31	0.66–2.62	0.44
Pancreatic-duct brushing cytology	2.21	0.43–11.29	0.34
Trainee only (Naïve papilla < 300)	1.20	0.60–2.41	0.61
Substitute for expert	1.10	0.57–2.11	0.78
<b>Sig2. difficult cannulation (&gt; 15 min)</b>	<b>2.77</b>	<b>1.50–5.12</b>	<b>0.001</b>
Total procedure time (> 30 min)	1.47	0.85–2.55	0.17
<b>Sig3. pancreatic injections (≥ 1)</b>	<b>4.81</b>	<b>2.38–9.72</b>	<b>&lt; 0.001</b>
Prophylaxis-related factors			
<b>*Absence of Pancreatic stent</b>	<b>1.74</b>	<b>0.90–3.35</b>	<b>0.098</b>

After the initial univariate analysis (Table 1), significant variables were adopted into the multiple logistic regression without multicollinearity variables (variance inflation factor > 10), non-observation variables in PEP and perfect prediction variables

Sig1–3: significant factor

\*Borderline significant factor ( $P < 0.10$ )

Bold indicates significance ( $< 0.05$ )

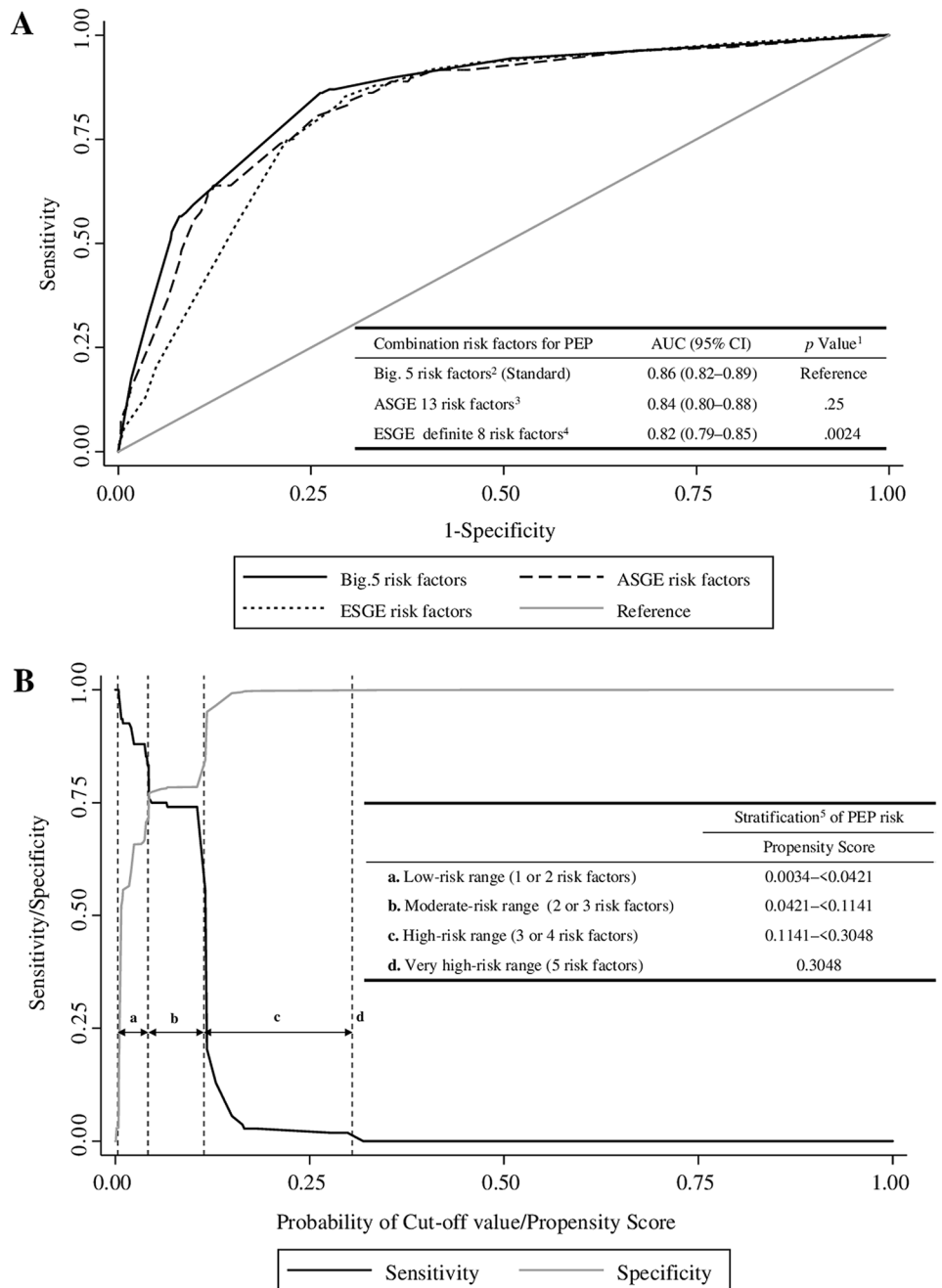
**Table 3** Three types of propensity score analyses to verify the significance and the robustness of “Big. 5” risk factors for PEP

Outcome = PEP	Adjusted logistic regression (adjusted, $n = 2683$ )		IPTW (adjusted, $n = 2683$ )		One-to-one matching analysis Exposure vs non-exposure, (n:n)	
	OR (95% CI)	<i>p</i> value	OR (95% CI)	<i>p</i> value	PEP ratio, %	<i>p</i> value
Exposure						
1. Naïve papilla	3.43 (1.66–7.13)	0.001	2.77 (0.71–10.87)	0.14	3.55 vs 0.89 (563:563)	0.004
2. PGW	3.72 (1.68–8.22)	0.001	6.18 (2.78–13.75)	< 0.001	20.78 vs 16.88 (77:77)	0.54
3. Difficult cannulation (> 15 min)	3.04 (1.47–6.30)	0.003	1.92 (0.86–4.31)	0.11	10.53 vs 4.61 (152:152)	0.05
4. Pancreatic injections (≥ 1)	6.33 (3.11–12.89)	< 0.001	3.94 (1.09–14.30)	0.037	4.90 vs 0.75 (531:531)	< 0.001
5. Absence of pancreatic stent	1.53 (0.69–3.41)	0.30	0.42 (0.14–1.29)	0.13	17.86 vs 11.90 (84:84)	0.28
Outcome = PEP score <sup>a</sup> (severity)	Adjusted linear regression (adjusted, $n = 2683$ )		IPTW (adjusted, $n = 2683$ )		One-to-one matching analysis Exposure vs non-exposure, (n:n)	
	$\beta$ (95% CI)	<i>p</i> value	$\beta$ (95% CI)	<i>p</i> value	PEP score <sup>C</sup> , pts	<i>p</i> value
Exposure						
1. Naïve papilla	0.07 (0.03–0.12)	0.001	0.07 (0.03–0.11)	0.001	0.09 vs 0.01 (563:563)	0.0023
2. PGW	0.19 (0.10–0.29)	< 0.001	0.27 (0.06–0.49)	0.012	0.46 vs 0.38 (77:77)	0.57
3. Difficult cannulation (> 15 min)	0.12 (0.05–0.20)	0.001	0.02 (– 0.02–0.06)	0.37	0.21 vs 0.11 (152:152)	0.06
4. Pancreatic injections (≥ 1)	0.08 (0.04–0.12)	< 0.001	0.07 (– 0.003–0.13)	0.06	0.10 vs 0.03 (531:531)	< 0.001
5. Absence of pancreatic stent	0.13 (0.02–0.23)	0.023	– 0.04 (– 0.14–0.06)	0.43	0.36 vs 0.23 (84:84)	0.30

All covariate (Table 1 and supplemental Tables 1 and 2) were used to adjust in propensity score analysis, without non-observation covariate in PEP group and variables after therapy allocation (e.g., prevention drug for PEP). Multicollinearity variables were also dropped (variance inflation factor > 10)

<sup>a</sup>Prognosis factors (0–9 pts)+CT grade (Grade1 = 1 pts, Grade2 = 2 pts, Grade3 = 3 pts): The Japanese guidelines for the management of acute pancreatitis 2015

**Fig. 2** Comparison of area under the curve for PEP (A) and stratification based on propensity score (B) in predictive scoring system for PEP. 1. One-way ANOVA with Bonferroni correction. 2. Five risk factors for PEP in the present model (Table 3). 3. Twelve risk factors of ASGE for PEP plus absence of pancreatic stent (Table 1). 4. Definite 7 risk factors of ESGE for PEP plus absence of pancreatic stent (Table 1). 5. Each threshold was set based on the range of propensity score



In PEP score for severity, all exposures showed statistical significance in the adjusted linear regression. Exposures of naïve papilla and PGW showed statistical significance in the IPTW, while those of naïve papilla and pancreatic injections ( $\geq 1$ ) also showed statistical significance in one-to-one propensity score matching analysis. Finally, five exposures revealed significance in the six propensity score analyses, as follows: 1. Naïve papilla, 2. PGW, 3. difficult cannulation ( $> 15$  min), 4. pancreatic injections ( $\geq 1$ ), and 5. absence of pancreatic stent. Thus, the five exposures were named “Big.

5”, and were employed for the subsequent predictive scoring system for PEP (Table 3).

**Predictive scoring system for PEP and internal validation**

The prediction model consisting of “Big. 5” revealed an AUC of 0.86 as a predictive probability of PEP (Fig. 2A). Inclusive of the absence of pancreatic stent, a total of 13 risk factors of ASGEs and 8 of ESGEs revealed AUC of 0.84 and 0.82, respectively. The AUC of the “Big. 5”

**Table 4** Lookup table for propensity score in the prediction model

Propensity score	Combinations
Very high-risk combinations for PEP	
0.3048	1. Naïve papilla, 2. PGW, 3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
High-risk combinations for PEP	
0.2009	1. Naïve papilla, 2. PGW, 3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ )
0.1905	1. Naïve papilla, 3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
0.1510	2. PGW, 3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
<b>0.1143</b>	<b>1. Naïve papilla, 3. Difficult Cannulation (&gt; 15 min), 4. Pancreatic injections (<math>\geq 1</math>)</b>
0.1141	1. Naïve papilla, 2. PGW, 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
Moderate risk combinations for PEP	
0.0869	2. PGW, 3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ )
0.0811	3. Difficult Cannulation (> 15 min), 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
0.0684	1. Naïve papilla, 3. Difficult Cannulation (> 15 min), 5. Absence of Pancreatic stent
0.0627	1. Naïve papilla, 2. PGW, 4. Pancreatic injections ( $\geq 1$ )
0.0582	1. Naïve papilla, 3. Difficult Cannulation (> 15 min), 5. Absence of Pancreatic stent
<b>0.0422</b>	<b>3. Difficult Cannulation (&gt; 15 min), 4. Pancreatic injections (<math>\geq 1</math>)</b>
0.0421	2. PGW, 4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
Low risk combination for PEP	
0.0288	1. Naïve papilla, 4. Pancreatic injections ( $\geq 1$ )
0.0223	3. Difficult Cannulation (> 15 min), 5. Absence of Pancreatic stent
0.0200	2. PGW, 4. Pancreatic injections ( $\geq 1$ )
0.0182	4. Pancreatic injections ( $\geq 1$ ), 5. Absence of Pancreatic stent
0.0146	1. Naïve papilla, 5. Absence of Pancreatic stent
0.0078	4. Pancreatic injections ( $\geq 1$ )
0.0034	5. Absence of Pancreatic stent

prediction model was significantly higher than that of ESGE ( $P=0.0024$ ), but not that of ASGE ( $P=0.25$ ) (Fig. 2A).

Four thresholds were set based on the range of propensity scores for stratifications from the low-risk group to the very high-risk group (Fig. 2B, Table 4). The diagnostic performance at different predicted probability cut-offs is shown in Supplementary Table 3.

The moderate-risk group (range: 0.0421–0.0869) usually had 3 items, which contained either one or both of “difficult cannulation (> 15 min)” and “pancreatic injections ( $\geq 1$ )” (Table 4). Likewise, the high-risk group (range: 0.1141–0.2009) usually had 4 items, which contained at least 2 of “difficult cannulation (> 15 min)”, “pancreatic injections ( $\geq 1$ )” and/or “naïve papilla” (Table 4).

The distribution ratio of four stratifications was 71.4% in the low-risk group, 18.9% in the moderate-risk group, 7.5% in the high-risk group, and 2.3% in the very high-risk group (Table 5). As for calibration, the Pearson goodness-of-fit test indicated goodness-of-fit for the model ( $P=0.993$ , Group = 20). Although the adjusted prevalence of PEP was 3.74% in the present model, the predicted incidence rate of PEP in the very high-risk group rose to 28.79% (95% CI,

18.30–41.25%) (Table 5). Conversely, predicted incidence rate of PEP in the low-risk group decreased in 0.73% (95% CI, 0.41–1.20%). Moreover, the very high-risk group was approximately 60-fold more likely to have PEP than the low-risk group (OR, 55.11; 95% CI, 26.40–115.07;  $P<0.001$ ) (Table 5). Likewise, although the adjusted prevalence of severe PEP was 0.90% in this model, the predicted incidence rate of severe PEP in the very high-risk group rose to 9.09% (95% CI, 3.41–18.74%). Conversely, that of severe PEP in low-risk group decreased to 0.15% (95% CI, 0.03–0.43%). The very high-risk group was approximately 70-fold more likely to have severe PEP than the low-risk group (OR, 68.57; 95% CI, 16.75–280.70;  $P<0.001$ ) (Table 5).

With regards to internal validation, the optimism-corrected frequency of PEP was acceptable in the very high-risk group at 29.82% (95% CI, 19.58–42.59%) and in the low-risk group at 0.73% (95% CI, 0.45–1.18%) compared to the predicted incidence rate (Table 5). The optimism-corrected AUC also revealed good performance at 0.81 (Fig. 3). However, a validation study could not be performed owing to the small number of severe PEP (26 events), which requires at least 100 events [26].



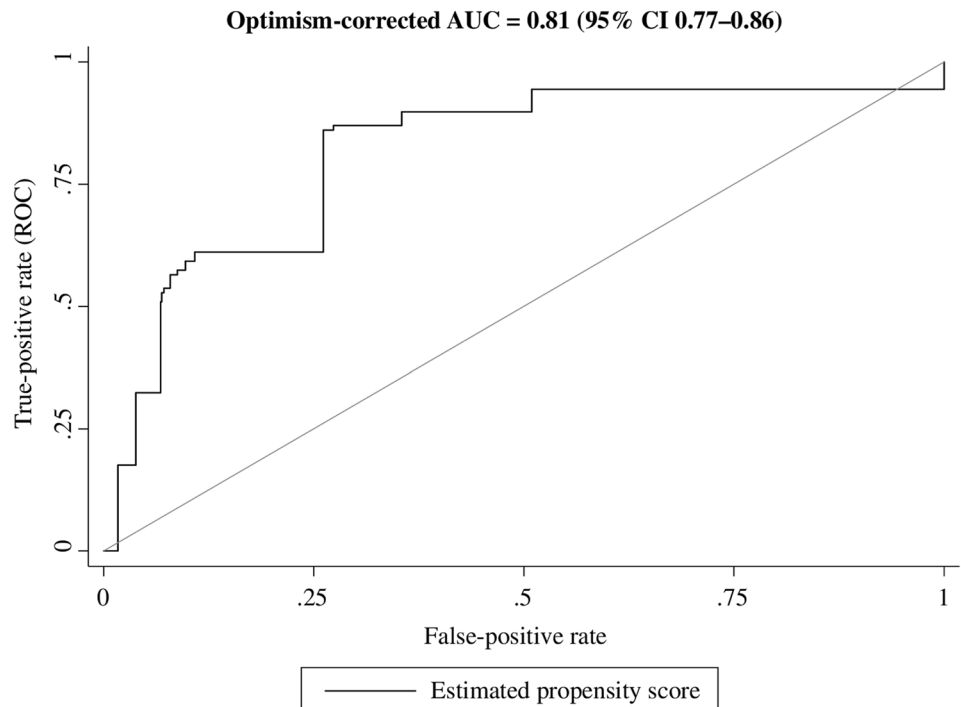
**Table 5** Predicted and observed incidence rates of PEP and odds ratios using an internally validated risk-score model

Risk category	Low-risk group	Moderate-risk group	High-risk group	Very high-risk group
Propensity score	0.0034–<0.0421	0.0421–<0.1141	0.1141–<0.3048	0.3048
Frequency, <i>n</i> (%)	2060 (71.4)	544 (18.9)	215 (7.5)	66 (2.3)
Incidence rate of PEP, <i>n</i> (%) [PEP prevalence = 3.74% (108/2885)]				
Predicted, % (95% CI)	0.73% (0.41–1.20)	5.88% (4.06–8.20)	19.53% (14.46–25.47)	28.79% (18.30–41.25)
Observed <sup>a</sup> , % (95% CI)	0.73% (0.45–1.18)	4.94% (3.44–7.04)	23.88% (18.62–30.07)	29.82% (19.58–42.59)
Odds of PEP				
Predicted OR (95% CI)	Reference	8.52 (4.58–15.85)	33.10 (17.99–60.89)	55.11 (26.40–115.07)
<i>p</i> value	–	<0.001	<0.001	<0.001
Incidence rate of severe PEP <sup>b</sup> , <i>n</i> (%) [severe PEP prevalence = 0.90% (26/2885)]				
Predicted, % (95% CI)	0.15% (0.03–0.43)	1.29% (0.52–2.63)	4.65% (2.25–8.38)	9.09% (3.41–18.74)
Odds of severe PEP				
Predicted OR (95% CI)	Reference	8.94 (2.30–34.68)	33.45 (9.13–122.50)	68.57 (16.75–280.70)
<i>p</i> value	–	0.002	<0.001	<0.001

<sup>a</sup>Resampling the model 1000 times

<sup>b</sup>In the setting of severe PEP, internal validation test was waived because validation studies should have at least 100 events to be meaningful (*n* = 26)

**Fig. 3** Optimism-corrected ROC curve and AUC in internal validation for PEP resampling the model 1000 times



## Discussion

In the present study, we established a simple predictive scoring system for PEP immediately after ERCP. This simplified clinical scoring system for PEP was derived from the combination of five risk factors as follows: 1. Naïve papilla, 2. PGW, 3. difficult cannulation (> 15 min), 4. pancreatic injections ( $\geq 1$ ), and 5. absence of pancreatic stent, which

we named the “Big. 5”. These factors were strengthened using three types of propensity score analysis. In the scoring system with four stratifications, the very high-risk group had a 28.79% predicted incidence rate of PEP, while that of severe PEP had a 9.09% predicted incidence rate of PEP, although the adjusted prevalence was only 3.74% in PEP and 0.90% in severe PEP. Conversely, the low-risk group had a 0.73% predicted incidence rate of PEP, while the severe PEP group had a 0.15% predicted incidence rate of PEP. The

very high-risk group was approximately 60-fold more likely to have PEP and 70-fold more likely to have severe PEP compared to the low-risk group. The present model showed a good calibration of the observed PEP frequency and a good discrimination of optimism-corrected AUC of 0.81 in the internal validation using 1000 times bootstrap resampling.

In view of the predictive scoring system for PEP, a small number of cohort studies used combinations of patient- and procedure-related risk factors to indicate odds ratios or frequencies [12, 13]. However, these were only univariate and multivariate logistic analysis methods, and the total integral number employed for the scoring system used no validation or split sample method. For reason, we employed both propensity score analysis and the bootstrap method to provide stable estimates with low bias, which can avoid the overfitting bias, overly pessimistic estimates of performance with large variability, and confounding factors [23, 24, 26, 27]. Thus, we established the simplified model using the Big 5, which consisted of five risk factors; this is comparable to the 13 factors in ASGE and 8 factors in ESGE (Fig. 2A).

In the patient-related risk factors, naïve papilla accounted for 88.9% in the PEP group. However, the other risk factors were flawed by univariate or multivariate regression analysis (Table 1, 2). There are several possible reasons for this, including the small number in this series (e.g., suspicious SOD, previous PEP, absence of chronic pancreatitis) may have decreased statistic power, or other patient-related factors (e.g., normal serum bilirubin) may have only had weak effects compared to naïve papilla and the significant procedure-related factors. The small number of suspicious SOD may reflect the daily medical practice because only 1.5% of patients with a functional gastrointestinal symptom were reported to have suspicious of SOD [28]. In addition, our result may support that the hypothesis that younger patients have no significant relationship with PEP [29]. However, it is also important to consider that other potential factors without naïve papilla had an incidence of 10.9% in the PEP group, and these factors may also have influenced the incidence of PEP.

In terms of ERCP procedures, PGW was a significant independent risk factor for PEP, because accomplishment of PGW may require both pancreatic injections and pancreatic guidewire passes, which are risk factors of PEP [2, 3]. This finding supports the risk of PGW in that there was no significance in the PEP group between PGW with pancreatic stent in 18 cases and PGW without pancreatic stent in 24 cases ( $P=0.25$ , Supplementary Table 2). However, the pancreatic stent should be installed when PGW is performed due to difficult cannulation with the easy pancreatic stent [2, 3]. In addition, non-use of the pancreatic stent was a significant risk factor of severe PEP in our adjusted linear regression of propensity score analysis (Table 3).

Concerning the proposal applied method for the present model: the moderate-risk group contained either one or both of “difficult cannulation ( $> 15$  min)” and “pancreatic injections ( $\geq 1$ )”. Likewise, the high-risk group contained at least two of the three items “difficult cannulation ( $> 15$  min)”, “pancreatic injections ( $\geq 1$ )” and/or “naïve papilla” (Table 4). The results were highly suggestive to prevent PEP since “difficult cannulation ( $> 15$  min)” is the only element to judge continuation or cessation of the ERCP procedure. For example, if the propensity score was 0.0627, containing “naïve papilla”, “PGW”, and “pancreatic injections ( $\geq 1$ )” (Table 4), the addition of “difficult cannulation ( $> 15$  min)” may lead to progression from the moderate-risk group with a PEP incidence of 4.94% to the high-risk group with a PEP incidence of 23.88% (Table 4, 5). This result may indicate the necessity of switching from an ERCP approach to an alternative endoscopic treatment, such as an endoscopic ultrasound-guided biliary drainage [30], if selective cannulation failed within 15 min.

The present model can stratify PEP candidates with low-risk or high-risk by score. Our stratifications suggest that it is possible to choose high-risk and low-risk groups from cases immediately after ERCP. These refinements may aid the early prediction of PEP, and allow selection of patients who require earlier therapeutic interventions, such as rectal NSAIDs, aggressive hydration, and nafamostat mesilate, possibly preventing the transfer of severe PEP [2–4, 9]. Indeed, in the current study, we demonstrated that the high-risk group and the very high-risk group were approximately 30-fold and 70-fold more likely incidence rate of severe PEP compared to the low-risk group, respectively (Table 5). In contrast, the low-risk group held the most common frequency of 71.8% in adjusted ERCP cases and had an observed incidence rate of only 0.73% (Table 5). These results suggest that the low- and moderate-risk group could be discharged early in conjunction with 2–6 h post-ERCP monitoring with serum amylase level [31–33]. Based on these considerations, this model could solve outpatient issues by introducing discharge or admission on the day of surgery in order to reduce medical expenses not only in the US but also in other countries including Japan [5–7]. In summary, the low-risk group may be able to undergo ERCP as an outpatient treatment and the moderate-risk group may need a temporary hospital stay with 2–6 h post-ERCP monitoring with serum amylase in the US. The high- and very high-risk group may need early and appropriate primary and intensive care to prevent progression of severe pancreatitis both in Japan and in the US.

Our study has several limitations. First, this was a retrospective study at a single center. Second, although propensity score analyses were used because of the difficulty to control multi-factorial etiology assignment in conjunction with ethical considerations, unobserved selection biases and

potential confounding factors may still remain. Third, the present model could not indicate the predictive ability for severe PEP owing to the small number of events ( $n=26$ ). Finally, a small number of risk factors were included in some cases, such as suspicious SOD, because the result may influence the reduced statistical power. Therefore, further investigations with a prospective multicenter study may be necessary to overcome these limitations.

In conclusion, we established and validated a simplified predictive scoring system for PEP using five minimal risk factors immediately after ERCP to help early identification of PEP and earlier therapeutic interventions. Thus, this scoring system is simple and easy to use and also achieves a high AUC.

**Acknowledgements** The authors thank the Department of Surgery, and the Division of Gastroenterology and Hepatology for patient recruitment. The authors also thank Enago for the English language editing of this paper.

**Author contributions** Conception and design: MC and MK Performing endoscopic examinations: MC, MK, YK, NS, YT, TA, KK, ST, YT and HT Acquisition of data: MC, MK, YK, NS, YT and TA Statistical analysis: MC and MK Analysis and interpretation of the data: MC, MK, YK, NS, YT, TA, KK, ST, MN and YT Writing of the original paper: MC, MK and KS Revision of the paper: MC and MK Approval of the final draft submitted: All authors.

**Funding** The authors declare that no funding support was received for this study.

### Compliance with ethical standards

**Disclosures** Drs. Masafumi Chiba, Masayuki Kato, Yuji Kinoshita, Nana Shimamoto, Youichi Tomita, Takahiro Abe, Keisuke Kanazawa, Shintaro Tsukinaga, Masanori Nakano, Yuichi Torisu, Hirobumi Toyozumi, and Kazuki Sumiyama have no conflicts of interest or financial ties to disclose.

### References

- Banks PA, Bollen TL, Dervenis C, Gooszen HG, Johnson CD, Sarr MG, Tsiotis GG, Vege SS, Acute Pancreatitis Classification Working Group (2013) Classification of acute pancreatitis–2012: revision of the Atlanta classification and definitions by international consensus. *Gut* 62:102–111
- ASGE Standards of Practice Committee, Chandrasekhara V, Khashab MA, Muthusamy VR, Acosta RD, Agrawal D, Bruining DH, Eloubeidi MA, Fanelli RD, Faulx AL, Gurudu SR, Kothari S, Lightdale JR, Qumseya BJ, Shaikat A, Wang A, Wani SB, Yang J, DeWitt JM (2017) Adverse events associated with ERCP. *Gastrointest Endosc* 85:32–47
- Dumonceau JM, Kapral C, Aabakken L, Papanikolaou IS, Tringali A, Vanbiervliet G, Beyna T, Dinis-Ribeiro M, Hritz I, Mariani A, Paspatis G, Radaelli F, Lakhtakia S, Veitch AM, van Hooff JE (2020) ERCP-related adverse events: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 52:127–149
- Leerhoy B, Elmunzer BJ (2018) How to avoid post-endoscopic retrograde cholangiopancreatography pancreatitis. *Gastrointest Endosc Clin N Am* 28:439–454
- Tryliskyy Y, Bryce GJ (2018) Post-ERCP pancreatitis: Pathophysiology, early identification and risk stratification. *Adv Clin Exp Med* 27:149–154
- Neermark S, Rasmussen D, Rysgaard S, Gluud LL, Novovic S, Schmidt PN (2019) The cost of endoscopic treatment for walled-off pancreatic necrosis. *Pancreatol* 19:828–833
- Nicolas-Perez D, Castilla-Rodriguez I, Gimeno-Garcia AZ, Romero-Garcia R, Nunez-Diaz V, Quintero E (2015) Prevention of post-endoscopic retrograde cholangiopancreatography pancreatitis: a cost-effectiveness analysis. *Pancreas* 44:204–210
- Zhang ZF, Duan ZJ, Wang LX, Zhao G, Deng WG (2017) Aggressive hydration with lactated Ringer solution in prevention of post-endoscopic retrograde cholangiopancreatography pancreatitis: a meta-analysis of randomized controlled trials. *J Clin Gastroenterol* 51:e17–e26
- Yu G, Li S, Wan R, Wang X, Hu G (2015) Nafamostat mesilate for prevention of post-ERCP pancreatitis: a meta-analysis of prospective, randomized, controlled trials. *Pancreas* 44:561–569
- Kim SJ, Kang DH, Kim HW, Choi CW, Park SB, Song BJ, Nam HS (2016) A randomized comparative study of 24- and 6-hour infusion of nafamostat mesilate for the prevention of post-endoscopic retrograde cholangiopancreatography pancreatitis: a prospective randomized comparison trial. *Pancreas* 45:1179–1183
- Cote GA, Lynch S, Easler JJ, Keen A, Vassell PA, Sherman S, Hui S, Xu H (2015) Development and validation of a prediction model for admission after endoscopic retrograde cholangiopancreatography. *Clin Gastroenterol Hepatol* 13:2323–32
- Jeurnink SM, Siersema PD, Steyerberg EW, Dees J, Poley JW, Haringsma J, Kuipers EJ (2011) Predictors of complications after endoscopic retrograde cholangiopancreatography: a prognostic model for early discharge. *Surg Endosc* 25:2892–2900
- Freeman ML, DiSario JA, Nelson DB, Fennerty MB, Lee JG, Bjorkman DJ, Overby CS, Aas J, Ryan ME, Bochna GS, Shaw MJ, Snady HW, Erickson RV, Moore JP, Roel JP (2001) Risk factors for post-ERCP pancreatitis: a prospective, multicenter study. *Gastrointest Endosc* 54:425–434
- Collins GS, Reitsma JB, Altman DG, Moons KG (2015) Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Br J Surg* 102:148–158
- Siau K, Dunckley P, Feeney M, Johnson G, Joint Advisory Group on Gastrointestinal Endoscopy (2019) ERCP assessment tool: evidence of validity and competency development during training. *Endoscopy* 51:1017–1026
- Wani S, Han S, Simon V, Hall M, Early D, Aagaard E, Abidi WM, Banerjee S, Baron TH, Bartel M, Bowman E, Brauer BC, Buscaglia JM, Carlin L, Chak A, Chatrath H, Choudhary A, Confer B, Cote GA, Das KK, DiMaio CJ, Dries AM, Edmundowicz SA, El Chafic AH, El Hajj I, Ellert S, Ferreira J, Gamboa A, Gan IS, Gangarosa L, Gannavarapu B, Gordon SR, Guda NM, Hammad HT, Harris C, Jalaj S, Jowell P, Kenshil S, Klapman J, Kochman ML, Komanduri S, Lang G, Lee LS, Loren DE, Lukens FJ, Mullady D, Muthusamy RV, Nett AS, Olyae MS, Pakseresht K, Perera P, Pfau P, Piraka C, Ponerros JM, Rastogi A, Razzak A, Riff B, Saligram S, Scheiman JM, Schuster I, Shah RJ, Sharma R, Spaete JP, Singh A, Sohail M, Sreenarasimhaiah J, Stevens T, Tabibian JH, Tzimas D, Uppal DS, Urayama S, Vitterbo D, Wang AY, Wassef W, Yachimski P, Zepeda-Gomez S, Zuchelli T, Keswani RN (2019) Setting minimum standards for training in EUS and ERCP: results from a prospective multicenter study evaluating learning curves and competence among advanced endoscopy trainees. *Gastrointest Endosc* 89:1160–1168.e9

17. Freeman ML, Guda NM (2005) ERCP cannulation: a review of reported techniques. *Gastrointest Endosc* 61:112–125
18. Tanaka R, Itoi T, Sofuni A, Itokawa F, Kurihara T, Tsuchiya T, Tsuji S, Ishii K, Ikeuchi N, Umeda J, Tonozuka R, Honjo M, Mukai S, Moriyasu F (2013) Is the double-guidewire technique superior to the pancreatic duct guidewire technique in cases of pancreatic duct opacification? *J Gastroenterol Hepatol* 28:1787–1793
19. Yokoe M, Takada T, Mayumi T, Yoshida M, Isaji S, Wada K, Itoi T, Sata N, Gabata T, Igarashi H, Kataoka K, Hirota M, Kadoya M, Kitamura N, Kimura Y, Kiriyama S, Shirai K, Hattori T, Takeda K, Takeyama Y, Hirota M, Sekimoto M, Shikata S, Arata S, Hirata K (2015) Japanese guidelines for the management of acute pancreatitis: Japanese guidelines 2015. *J Hepatobiliary Pancreat Sci* 22:405–432
20. Miura F, Okamoto K, Takada T, Strasberg SM, Asbun HJ, Pitt HA, Gomi H, Solomkin JS, Schlossberg D, Han HS, Kim MH, Hwang TL, Chen MF, Huang WS, Kiriyama S, Itoi T, Garden OJ, Liau KH, Horiguchi A, Liu KH, Su CH, Gouma DJ, Belli G, Dervenis C, Jagannath P, Chan ACW, Lau WY, Endo I, Suzuki K, Yoon YS, de Santibanes E, Gimenez ME, Jonas E, Singh H, Honda G, Asai K, Mori Y, Wada K, Higuchi R, Watanabe M, Rikiyama T, Sata N, Kano N, Umezawa A, Mukai S, Tokumura H, Hata J, Kozaka K, Iwashita Y, Hibi T, Yokoe M, Kimura T, Kitano S, Inomata M, Hirata K, Sumiyama Y, Inui K, Yamamoto M (2018) Tokyo Guidelines 2018: initial management of acute biliary infection and flowchart for acute cholangitis. *J Hepatobiliary Pancreat Sci* 25:31–40
21. Cotton PB, Lehman G, Vennes J, Geenen JE, Russell RC, Meyers WC, Liguory C, Nickl N (1991) Endoscopic sphincterotomy complications and their management: an attempt at consensus. *Gastrointest Endosc* 37:383–393
22. Mine T, Morizane T, Kawaguchi Y, Akashi R, Hanada K, Ito T, Kanno A, Kida M, Miyagawa H, Yamaguchi T, Mayumi T, Takeyama Y, Shimosegawa T (2017) Clinical practice guideline for post-ERCP pancreatitis. *J Gastroenterol* 52:1013–1022
23. Cepeda MS, Boston R, Farrar JT, Strom BL (2003) Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol* 158:280–287
24. Austin PC (2011) An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res* 46:399–424
25. Efron B (1979) Bootstrap methods: another look at the Jackknife. *Ann Statist* 7:1–26
26. Steyerberg EW (2018) Validation in prediction research: the waste by data splitting. *J Clin Epidemiol* 103:131–133
27. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD (2001) Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* 54:774–781
28. Drossman DA, Li Z, Andruzzi E, Temple RD, Talley NJ, Thompson WG, Whitehead WE, Janssens J, Funch-Jensen P, Corazziari E (1993) U.S. householder survey of functional gastrointestinal disorders. Prevalence, sociodemography, and health impact. *Dig Dis Sci* 38:1569–1580
29. Ding X, Zhang F, Wang Y (2015) Risk factors for post-ERCP pancreatitis: a systematic review and meta-analysis. *Surgeon* 13:218–229
30. Sharaiha RZ, Khan MA, Kamal F, Tyberg A, Tombazzi CR, Ali B, Tombazzi C, Kahaleh M (2017) Efficacy and safety of EUS-guided biliary drainage in comparison with percutaneous biliary drainage when ERCP fails: a systematic review and meta-analysis. *Gastrointest Endosc* 85:904–914
31. Ito K, Fujita N, Noda Y, Kobayashi G, Horaguchi J, Takasawa O, Obana T (2007) Relationship between post-ERCP pancreatitis and the change of serum amylase level after the procedure. *World J Gastroenterol* 13:3855–3860
32. Jeurnink SM, Poley JW, Steyerberg EW, Kuipers EJ, Siersema PD (2008) ERCP as an outpatient treatment: a review. *Gastrointest Endosc* 68:118–123
33. Rabago L, Guerra I, Moran M, Quintanilla E, Collado D, Chico I, Olivares A, Castro JL, Gea F (2010) Is outpatient ERCP suitable, feasible, and safe? The experience of a Spanish community hospital. *Surg Endosc* 24:1701–1706

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.