

Peer Review File

Article information: <https://dx.doi.org/10.21037/gS-21-550>

Reviewer A

Yin et al. attempted to develop prediction models for the occurrence of POPF and drainage management based on a large collective of PD, performed at a single Chinese center over a 5-year period. Statistical methods are appropriate and English language is acceptable. The topic is interesting, however not original and plenty of high-quality literature already exists. The manuscript has certain potential if some important issues be considered and considerably revised. My major issues of concern are focused on the methods of the study:

1. the retrospective character of the study should be compensated by a stringent and standardized surgical management of the patients. However, the authors don't provide any information on the extent of standardisation in their institution concerning PD: how many surgeons performed these procedures and how was the number of PDs distributed b/n surgeons. Was there a single standard technique of PD and what were the variations: PJ vs PG, type of PJ -Peng or Blumgart or invagination vs PG, stents or no stents, external or internal, how many drainages, what type, reconstruction mode (single vs double loop), minimally invasive vs open, criteria for different techniques, criteria for pp vs classic etc. All these technical issues play a significant, sometimes decisive role for the development of POPF, so the authors should give a detailed description in the methods.

Reply 1: There were five surgeons performed PDs in our center. Among 993 patients, Dr Miao performed 423 PDs, Dr Dai performed 262 PDs, Dr Jiang performed 171 PDs, Dr Wu performed 75 PDs and Dr Gao performed 62 PDs. In our center, we actually have relative routines to perform PD or PPPD. We always choose PJ in our PDs. PJ was performed using two methods: two-layer invagination and modified one-layer duct-to-mucosa. The latter method is a new approach to PJ that has been utilized since 2011. An internal stent was routinely placed unless if the diameter was dilated larger than 5mm. No external stent was used in all cases. During the study period, intraoperative drainage tubes were routinely placed in situ in all cases. All the digestive reconstruction were performed by Child method. In the 993 patients, only

18 patients received laparoscopic pancreaticoduodenectomy. Classic or pylorus-preserving pancreaticoduodenectomy was performed according to the indication as well as preference of the individual surgeon.

Changes in the text: We have added modified our text as advised (see Page 4, line 98-106).

2. some contradictory or implausible statements should be cleared: the minimal time for PD was given as 90 minutes - did this patient died intraoperatively -sorry, I haven't seen a 90 minutes PD for 20 years of pancreatic surgery, having visited some of the largest centers worldwide. Also intraoperative infusion of only 500 ml is quite weird. Line 77 says drainage amylase was measured on POD1 and POD 3, however line 98 and the tables declare drainage amylase was determined only on/after POD3? Also, high drainage amylase on POD3 is already a POPF according to definition, and not a risk factor as used by the authors.

Reply 2: As the 90 min PD, we confirmed that is true. This procedure was performed by Dr Miao Yi in 2015/11/25 and the patient was discharged on POD 9 without any complication.

Changes in the text: Almost all the patients were received less than 3 hours PPPDs with only 500 ml of intraoperative infusion. POD1 drain amylase was not routinely measured in all cases. POD 3 drain amylase was used to evaluate probability of the early removal of drainage, not pancreatic fistula. That may make some misunderstanding of our study.

Changes in the text: No changes.

3. line 56 in the introduction claims that FRS/aFRS were not validated - this is not true - these have been validated in a number of well-designed and large studies. The authors mention first in the discussion, that they tested the accuracy of FRS/aFRS for their collective and found it had low AUC. This is a problem - first the raw data of this calculation are missing, second the authors declared no pancreatic texture was recorded for their patients - how did they calculated FRS then? On the other side they say in the conclusion that their own models were not better than the FRS.

Reply 3: In our context, we actually acknowledged that FRS or a-FRS were well validated, but these models still had some drawbacks in our mind. The AUC of our data based on FRS was 0.683. The raw data will be added in the revised manuscript. Although we had the raw data about pancreatic texture, the low AUC of FRS made us rethink our evaluation of parenchymal consistency was objective or not, so we did not

include this factor. Our model I was not better FRS, but the AUC of model II was elevated.

Changes in the text: We added raw data about AUC of FRS (see Page 6, line 169, Figure 6)

4. how low ALT increases the risk for POPF - literature?

Reply 4: Relatively normal ALT levels were mostly observed in patients with a normal pancreas and no history of jaundice, which has been reported as a risk factor [Yu L, Huang Q, Xie F, Lin X, Liu C. Risk factors of postoperative complications of pancreatoduodenectomy. *Hepatogastroenterology*. 2014, 61(135):2091-5. Ausania F, Snowden CP, Prentis JM, Holmes LR, Jaques BC, White SA, French JJ, Manas DM, Charnley RM. Effects of low cardiopulmonary reserve on pancreatic leak following pancreaticoduodenectomy. *Br J Surg* 2012, 99(9): 1290-1294.]. It was interesting to note that ALT, but not preoperative bilirubin, was significantly related to the incidence of POPF in our study. However, it is improper to simply conclude that preoperative improvement in liver function will increase POPF rate; on the contrary, conditions including cirrhosis and portal hypertension may compromise surgical outcomes [El Nakeeb A, Sultan AM, Salah T, El Hemaly M, Hamdy E, Salem A, Moneer A, Said R, AbuEleneen A, Abu Zeid M, Abdallah T, Abdel Wahab M. Impact of cirrhosis on surgical outcome after pancreaticoduodenectomy. *World J Gastroenterol* 2013, 19(41): 7129-7137.]

Changes in the text: We have added modified our text as advised (see Page 8, line 226-230).

5. Lines 228-231 state that fibrosis of pancreas was not used in the model, since it is impossible to analyse it intraoperatively. This is not true - plenty of studies show that all histomorphologic features aof pancreas may be quantitatively and rapidly determined on frozen section intraoperatively (e.g. Belyaev O, Munding J, Herzog T, Suelberg D, Tannapfel A, Schmidt WE, Mueller CA, Uhl W. Histomorphological Features of the Pancreatic Remnant as Independent Risk Factors for Postoperative Pancreatic Fistula: A Matched-Pairs Analysis. *Pancreatology* 2011;11(5):516-524.)

Reply 5: Thank you for your kind remind.

Changes in the text: We have revised our manuscript (see Page 9, line 243-245).

6. the parameters used for both models are quite uncommon - in the intraoperative model 1 some of the parameters included are actually preoperative, such as ALAT in

serum, and the pancreatic duct diameter may be determined preop on imaging. Pathology itself is not a pure intraop parameter since it is determined at final pathology in most cases. Postop Model 2 includes pancreatic duct diameter and intraoperative colloid fluid volume?!?

Reply 6: Our intraoperative model indeed contains some preoperative parameters, actually we used the pre/intraoperative as the title before. But we want to emphasize that model I was calculated during the operation to decide whether place drains or not, rather than before the operation. Pathology could be predicted by surgeon's experience or frozen section during the operation. Model II was calculated in order to decide whether remove the drain on POD3 or not. In a word, these two models were named by the decision-making time of drain management.

Change in the text: No change.

7. a major problem is that the most important risk factor for POPF - soft pancreatic texture is not recorded at all. Techniques for quantitative measurement of pancreatic tissue softness exist and have been validated in many studies - durometry is being currently applied as a standard in many german and italian institutions (e.g. Belyaev O, Herden H, Meier JJ, Muller CA, Seelig MH, Herzog T, Tannapfel A, Schmidt WE, Uhl W. Assessment of pancreatic hardness – surgeon versus durometer. J Surg Res 2010;158(1): 53-60.)

Reply 7: Unfortunately, we did not use the quantitative method for measurement of pancreatic texture so far. Although we have the raw data about pancreatic texture, we still decided exclude this relative subjective parameter in this study.

Change in the text: No change.

8. the authors declare their model predicted well the risk for POPF in low risk patients - unfortunately nobody needs POPF prediction in low risk patients - it is a really rare event in this group. The models failed to predict POPF exactly there where it is most needed - in the high risk group.

Reply 8: The original aim of our study is to help us to decide whether place drains and early drain removal. The patient safety is our most concern. The advantage is that we can set the different predictive risk value to expand or narrow the benefit populations.

Change in the text: No change.

Reviewer B

It seems like a good study to compensate for the weakness of RFS in the US.

Reply: Thank you for your kind appreciation.

Change in the text: No change.

Reviewer C

1. Were the patients included in this study all, consecutive cases? Were there any cases excluded, such as combined resection of liver? Please comment or describe the point more in detail.

Reply 1: Our study included all consecutive patients in the study period. Some cases were excluded in the model validation process because missing data. There were no patients received HPD.

Change in the text: No change.

2. The entire cohort was randomly divided into 2 groups in a scale of 2:1. The 2 groups were well balanced in each characteristic. Did the author use any stratification factor?

Reply 2: We did not use any stratification factor for randomization.

Change in the text: No change.

3. Preoperative ALT has not been well evaluated for the risk factor of POPF previously. Why did the author choice the factor for analysis? Please comment or cite the appropriate references.

Reply 3: Relatively normal ALT levels were mostly observed in patients with a normal pancreas and no history of jaundice, which has been reported as a risk factor [Yu L, Huang Q, Xie F, Lin X, Liu C. Risk factors of postoperative complications of pancreatoduodenectomy. *Hepatogastroenterology*. 2014, 61(135):2091-5. Ausania F, Snowden CP, Prentis JM, Holmes LR, Jaques BC, White SA, French JJ, Manas DM, Charnley RM. Effects of low cardiopulmonary reserve on pancreatic leak following pancreaticoduodenectomy. *Br J Surg* 2012, 99(9): 1290-1294.]. It was interesting to note that ALT, but not preoperative bilirubin, was significantly related to the incidence of POPF in our study. However, it is improper to simply conclude that preoperative improvement in liver function will increase POPF rate; on the contrary, conditions including cirrhosis and portal hypertension may compromise surgical outcomes [El Nakeeb A, Sultan AM, Salah T, El Hemaly M, Hamdy E, Salem A, Moneer A, Said R, AbuEleneen A, Abu Zeid M, Abdallah T, Abdel Wahab M. Impact of cirrhosis on

surgical outcome after pancreaticoduodenectomy. World J Gastroenterol 2013, 19(41): 7129-7137.]

Change in the text: We have added modified our text as advised (see Page 8, line 226-230).

4. Did the authors use any drug for the prevention or for the treatment of POPF, such as octreotide? Use or no-Use of prophylactic agents for POPF should be shown in Method section.

Reply 4: Octreotide was routinely used until POD 5 in all cases.

Change in the text: We have revised our manuscript (see Page 4, line 107).

5. Please state more in detail about the anastomosis procedure of pancreas. This might affect the results and should be included in discussion.

Reply 5: Pancreaticojejunostomy (PJ) was performed using two different methods: two-layer invagination and modified one-layer duct-to-mucosa which was reported in our previous study. There was no significant difference between the two methods in the occurrence of POPF (18.2% vs. 15.3, P=0.245). During PJ, either interrupted or running suturing was adopted, depending on the surgeon's preference.

Change in the text: We have revised our manuscript (see Page 4, line 101-104; Table 1).

6. Although the author demonstrated the results performing a validation by FRS (Fistula Risk Score) in Discussion section, this should be placed in Results section, to compare the model they used.

Reply 6: That is right. We added the AUC of FRS in Results section.

Change in the text: We added raw data about AUC of FRS (see Page 6, line 169, Figure 6)

Reviewer D

This study is interesting in that it presents nomograms at two time points using separate data before and after surgery. However, it seems that major revision are needed in several points.

* Major points

1. Testing by dividing into two groups by one institution appears to be internal validation, not external validation, by definition. External validation refers to validating the nomogram developed with patient groups from other institutions.

Reply 1: According to the Tripod explanation, prediction model development studies with validation may be done in participant data collected by the same investigators, commonly using the same predictor and outcome definitions and measurements, but sampled from a later time period (so-called "temporal" or "narrow" validation); by other investigators in another hospital or country, sometimes using different definitions and measurements (geographic or broad validation). We acknowledged that our study had some drawbacks in separating training and validation cohort by randomization, which was not recommended in Tripod statement.

Change in the text: No change.

2. There are too many missing values to validate of Model II. (Internal validation (n=440, excluding 221 cases with miss data) and external validation (n=205, excluding 172 cases with miss data)

Reply 1: The missing data were mostly attributed to drain amylase on POD 3. DFA levels were routinely measured on or after POD3, in the early study period, we did not test DFA on POD 3 for every patient because we would not remove the drainage in such early time.

Change in the text: No change.

* Minor points

1. Table 1

- add the data for open/ laparoscopic/ Robotic sugery
- duct diameter 3.0 (1.0–5.0), 3.0 (1.0–20.0), $p < 0.001$?
- check the mean value of EBL, Intraoperative colloid infusion, Intraoperative crystal infusion. All of the value are same.

Reply 1: In the 993 patients, only 18 patients received minimally invasive (laparoscopic) pancreaticoduodenectomy. As for duct diameter, the P value < 0.001 was correct when we used Mann-Whitney *U* test to compare POPF group and no-POPF group. The mean value of EBL and intraoperative infusion were also correct.

Change in the text: No change.

2. Table 2

- multivariable analysis: intraoperative colloid infusion (OR 1.00, 95%CI 1.00–1.00,

P=0.001)?

Reply 2: Intraoperative colloid infusion as continuous data was put into multivariable analysis. The OR and P value were correct according to our analysis.

Change in the text: No change.

3. Line 26 intrao-perative

Reply 3: Thank you for pointing this typo.

Change in the text: We have revised our manuscript (see Page 1, line 26).

4. Line 74/line 127/line 145 combined portal-superior vein resection

Reply 4: Thank you for pointing this typo. We have revised it as portal-superior mesenteric vein resection

Change in the text: We have corrected all the terms in our context.

5. Line 205 eraly

Reply 5: Thank you for pointing this typo. We had revised it as early.