Peer Review File

Article Information: http://dx.doi.org/10.21037/jgo-20-436

Review Comments:

Comment 1: English language of the paper is poor. Please have the paper polished by English-speaking professionals. For example, line 56 "inadequate patients", line 67 should be "were significantly associated with AL".

Reply 1: Thanks for your suggestions. The manuscript has been polished by English-speaking professionals.

Changes in the text: We have changed our text as advised (see Page 3, line 55 and 66). Proof of English language editing has been added as an attachment.

Comment 2: Abstract. Methods, please briefly introduce subjects, measures, and outcomes of this study by using PICOS criteria. The collection of major predictive factors is necessary here. Results, please specify the factors included in to the predictive model.

Reply 2: Thanks for suggestion. We didn't use PICOS criteria because we didn't give any intervention. Instead, we wrote the article following TRIPOD checklist which is professional for prediction model development and validation study (see Page 6, line 104-105). A total of 20 variables were included and it's prolix to elucidate them in abstract so we specified them in methods of text part.

Changes in the text: We modified the methods to include the outcome of the study---"***Patients diagnosed with AL within 6 months' follow-up were recorded.***" (see Page 3, line 58-59) We added the number of candidate factors in methods of abstract (see Page 3, line 59-60). All of 8 factors in the predictive model were elaborated in results--- "***sex, distance of tumor from the anal verge, bowel stenosis or obstruction, preoperative hemoglobin, surgeon volume, diabetes, neoadjuvant chemoradiotherapy, and surgical approach***" (see Page 3, line 64-66).

Comment 3: Conclusion should be made with cautions due to potential sample selection bias and lack of external validation samples.

Reply 3: Thanks for your advice. Just as explained in discussion, this was a retrospective study in a single center, which might have caused selection bias for patients and lacked external validation of other centers. We are now conducting a prospective multicenter study to further verify our model through website. (http://www.changhai-rc-al-prediction.org)

Changes in the text: Potential sample selection bias was explained in Page 15, line 318. We have modified the lack of external validation samples in conclusion as advised (see Page 15, line 320-322).

Comment 4: Introduction. Line 93-94, it would be helpful to briefly review existing knowledge on factors associated with AL, as well as predictive studies of AL. Line 94-95, please provide specific data to support "small sample sizes and inconsistent findings". Importantly, why a predictive model is necessary should be specified. In the last paragraph, the authors should compare the limitations and strengths between traditional approaches of predictive model development and machine-learning based random forest.

Reply 4:

Thanks for your advice. We have mentioned existing knowledge on factors and studies associated

with AL in discussion that Koyama M et al believed a temporary stoma could indirectly accelerate anastomotic healing, thus reducing the incidence of AL (1). Yun JA et al discovered that a temporary stoma could not reduce the incidence of AL (2). Shinji S identified that distance of tumor from the anal verge and sex are risk factors for AL.(3) Xiao C et al reported that anemia or massive blood loss during operation and nCRT were considered as predictors(4). Zhang W et al added that diabetes was related to AL (5). Various factors have been proved to be associated with AL.

Just as explained in introduction that we aimed to analyze a large number of rectal cancer patients after AR to illustrate the risk factors of AL, and to create a random forest classifier to better predict the incidence of AL and give an advice on whether to do a temporary stoma. So we constructed the model.

We have compared the limitations and strengths between traditional approaches of predictive model development and machine-learning based random forest. First of all, we reviewed widely used machine learning in introduction that the performance of the established models predicting the incidence of AL remains unsatisfactory, while random forest, a new and highly flexible machine learning algorithm, has wide application prospects, and has been demonstrated to have better performance in disease prediction.(6,7) To further emphasize the preciseness, we illustrated the mechanism of random forest in methods that the pre-processed data set was split into training set and validation set, grid-search cross validation technique and 5-fold cross validation were applied to prevent overfitting. Besides, the differences between 2 models were illuminated in discussion that the principle of the nomogram to predict AL is based on logistic regression that has limitations in the fitting of model creation, while machine learning based random forest can partly overcome the limitations of the regression models and has shown better predictive value than the traditional prediction model.

Changes in the text: We have briefly introduced the current knowledge of anastomotic healing---***"According to the current knowledge, the healing of the anastomotic site depends on the tension and blood supply around the anastomotic site"*** (see Page 5, line 94-96). Some articles about the previous prediction studies or supporting "small sample sizes and inconsistent findings" have been cited (see Page 5, line 96).

Comment 5: Methodology. The validation sample is small and validation sample from the same hospital can not guarantee the external validity of the predictive model developed by the authors. Please briefly describe how the patients were followed up in this part. For candidate predictors, the authors need to specify why these variables were selected. For predicting, surgeon-related factors are also important, such as the years of clinical services of treating physicians, and training experiences but the authors did not include such factors.

Reply 5:

Thanks for your suggestion. We admitted that the study lacked external validation of other centers which would be further validated in our following studies by adopting the website (http://www.changhai-rc-al-prediction.org) and we have modified the expression in the article. The overall incidence of AL was 6.2% in our center and the number of AL patients in test set was only 45. However, a total of 836 patients validating the model were a relatively large member as far as we know, because we are not only interested in patients with AL. We want to identify the patients with or without AL in a more precise way to avoid unnecessary ileostomy.

AL was diagnosed via digital rectal examination (DRE), endoscopy, or imaging examination within 6 months. The follow-up was conducted via outpatient or telephone. Variables were selected as

candidate predictors because of previous reports (8-11) and clinical experiences.

According to the surgeon-related factors, we initially found the incidence of AL was associated with different surgeons. Since every surgeon in our center had had strict training program before operating independently, we focused on the changes of surgery volumes. A total of 8 surgeons were well-trained and surgery volumes are increasing year by year. Considering that the data spanning 10 years, we classified surgeons by surgery volumes per year rather than surgeons themselves. It is concluded that low volume surgeons and doing a laparoscopic surgery in the surgeon's first operating year are more likely to cause AL, which means that a learning curve were also exist even when surgeons were well-trained. In addition, we added Table S as the proof of different surgeons' qualifications.

Changes in the text: The reason of variable selection was explained at the end of Variables (see Page 8, line 147-148). A total of 8 surgeons were well-trained. Considering that the data spanning 10 years, we classified surgeons by surgery volumes per year rather than surgeons themselves. The qualifications of surgeons have been illustrated (see Page 7, line 144-145 and Table S).

Comment 6: Statistics. Line 148-151, it remains unclear the objectives of these comparisons and the comparisons were made between which groups. I did not agree with PSM approach in this study. The focus of this part is to identify predictors, not risk factors. The authors should avoid to use the term "risk factors". It is unnecessary to ascertain the role of temporary stoma. In this part, the primary aim is to identify predictors via univariate and multivariate analyses. To further illustrate the strengths of machine-learning based random forest over traditional logistic regression, In suggest the authors to also present the predictive model based on logistic regression. In addition to AUC, sensitivity and specificity of the new model should be also calculated and reported. A further question is the small sample size of validation sample, resulting only 45 patients with AL. This may result in unstable findings on the performance assessment of the new predictive model. I suggest the authors to include more sample for validation.

Reply 6:

The comparisons were made between patients with AL and without AL (Table 1). A temporary stoma is widely used in AR of lower RC. The role of temporary stoma remained controversial since some researchers believed stoma could lower the incidence of AL (1) while others didn't think so (2). In addition, a temporary stoma could have great impact on patients such as ordinary inconvenience and second operation. We further verified the effect of stoma on AL by utilizing the data of our center so as to provide some clinical advice.

This is a retrospective study, and PSM was adopted to balance selection bias so as to identify whether it could reduce the incidence of AL. Nomogram based on logistic regression was also constructed (Figure 3) to compare 2 models (Figure 2a,2b,2c; Figure 4a,4b), thus illustrate the strengths of machine-learning based random forest.

We have mentioned in results that the sensitivity of training and validation set was 0.827 and 0.818, respectively, and the specificity was 0.739 and 0.67, respectively. According to the data in the test set, the AUC was 0.87, while the sensitivity and specificity were 0.844 and 0.697, respectively.

A total of 5220 patients were included to train the model and divided into training set and validation set by 5-fold cross validation. In addition, a total of 836 patients were included in the test set, which means our study were double validated, so the model is relatively rigorous. We admit that only 45 patients with AL in test set are relatively small. Nevertheless, the aim of our study is identifying not

only AL patients but also non-AL patients, thus avoid unnecessary ileostomy. And our sample size might be the biggest one among the current single center studies. Actually, the study lacked external validation of other centers and we are conducting a prospective study to further verify it by the website (http://www.changhai-rc-al-prediction.org) which can be seen in our following article.

Changes in the text: We have changed some “risk factors” into “predictive factors” or “predictors” as suggested.

Reference

1. Koyama M, Murata A, Sakamoto Y, et al. Risk Factors for Anastomotic Leakage After Intersphincteric Resection Without a Protective Defunctioning Stoma for Lower Rectal Cancer. Ann Surg Oncol 2016;23 Suppl 2:S249-56.

2. Yun JA, Cho YB, Park YA, et al. Clinical manifestations and risk factors of anastomotic leakage after low anterior resection for rectal cancer. ANZ J Surg 2017;87:908-14.

3. Shinji S, Ueda Y, Yamada T, et al. Male sex and history of ischemic heart disease are major risk factors for anastomotic leakage after laparoscopic anterior resection in patients with rectal cancer. BMC Gastroenterol 2018;18:117.

4. Xiao C, Zhou M, Yang X, et al. Novel nomogram with microvascular density in the surgical margins can accurately predict the risk for anastomotic leakage after anterior resection for rectal cancer. 2019;120:1412-9.

5. Zhang W, Lou Z, Liu Q, et al. Multicenter analysis of risk factors for anastomotic leakage after middle and low rectal cancer resection without diverting stoma: a retrospective study of 319 consecutive patients. Int J Colorectal Dis 2017;32:1431-7.

6. Mogensen UB, Ishwaran H, Gerds TA. Evaluating Random Forests for Survival Analysis using Prediction Error Curves. J Stat Softw 2012;50:1-23.

7. Charoentong P, Luedde T, Herpel E, et al.

8. Rencuzogullari A, Benlice C, Valente M, et al. Predictors of Anastomotic Leak in Elderly Patients After Colectomy: Nomogram-Based Assessment From the American College of Surgeons National Surgical Quality Program Procedure-Targeted Cohort. Dis Colon Rectum 2017;60:527-36.

9. Yao HH, Shao F, Huang Q, et al. Nomogram to predict anastomotic leakage after laparoscopic anterior resection with intracorporeal rectal transection and double-stapling technique anastomosis for rectal cancer. Hepatogastroenterology 2014;61:1257-61.

10. Park JS, Huh JW, Park YA, et al. Risk Factors of Anastomotic Leakage and Long-Term Survival After Colorectal Surgery. Medicine (Baltimore) 2016;95:e2890.

11. Zheng H, Wu Z, Wu Y, et al. Laparoscopic surgery may decrease the risk of clinical anastomotic leakage and a nomogram to predict anastomotic leakage after anterior resection for rectal cancer. Int J Colorectal Dis 2019;34:319-28.