Peer review file
Article information: http://dx.doi.org/10.21037/jtd-21-636

Reviewer A

Comment 1:1. The definition of postoperative complications should be based on Clavien-Dindo classification. Authors included dyspnea in complications, however, dyspnea is not subjective factor for postoperative complications.
Reply 1: Thank you for the reviewer's valuable advice. The definition of our postoperative complications has been based on Clavien-Dindo classification. According to the definition of Clavien-Dindo classification, the dyspnea can be classified into the Grade II. Since patients with postoperative dyspnea usually need expectorant and antiasthmatic drugs to promote expectoration and reduce airway spasm.
Changes in the text: We have modified our text as advised (see Page 7, line 106-110).

Comment 2: The postoperative complications should be classified as minor and major complications.
Reply 2: Thank you for the reviewer's valuable advice. According to the advice of the reviewer, we have classified the postoperative complications (PC) into minor and major complications. The minor PC included: pleural effusion, dyspnea, arrhythmia, air leakage, fever, while the major PC included: reoperation, chylothorax, mechanical ventilation, cardio-dysfunction, pulmonary embolism, and death.
Changes in the text: We have modified our text as advised (see Page 7, line 103-105).

Reviewer B

Comment 1: To me, the main problem in this paper is the heterogeneity of the population under study. Your series includes from young people undergoing surgery to treat spontaneous pneumothorax to elderly patients undergoing NSCLC anatomical resection. Most published risk scoring systems are intended to help decisions on surgical indication in homogeneous populations, usually patients proposed for anatomical lung resection. Your results stratify several groups of cases but do not help to decide in risky cases undergoing major lung resection.
Reply 1: Thank you for the reviewer's valuable advice. The heterogeneity of the population will certainly affect the feasibility of the risk scoring systems. And here, we re-validated the model by different age groups and ASA levels, and we also found that different age groups and ASA levels have no effect on the risk model, which indicates that the model is still applicable for different patients.

Subgroup	Development data set			Validation data set		
	AUC	Lower confidence limit	Upper confidence limit	AUC	Lower confidence limit	upper confidence limit
Age						
< 65	0.7123	0.6484	0.7762	0.6515	0.5629	0.74
≥65	0.6267	0.5096	0.7438	0.8901	0.7454	1
ASA level						
I	0.6804	0.6126	0.7483	0.6946	0.5912	0.7981

II	0.7680	0.6858	0.8501	0.7056	0.5663	0.8450
II	0.7063	0.2965	1.0000	0.7222	0.5093	0.9351

Changes in the text: We have supplemented the data in Supplemental-Table S4.

Comment 2: According to pre-previous comment, I'm suggesting including in the analysis just cases undergoing scheduled anatomical lung resection. You are decreasing case number in the analysis, but the precision and relevance of the model will be higher.
Reply 2: Thank you for the reviewer's valuable advice. First, considering the risk of reducing the case number, we think that we can prove the feasibility of this risk model by analyzing the subgroup of different types of surgery. The results of subgroup analysis are as follows, and show that AUC >0.668. In addition, the type of surgery was included as a candidate variable to establish the risk model at first, and was not selected to be included in the model at last, which indicates that this risk model is applicable to different types of surgery.

Subgroup	Development data set			Validation data set		
	AUC	Lower confidence limit	Upper confidence limit	AUC	Lower confidence limit	Upper confidence limit
Surgery type						
Non-anatomical lung surgery	0.6716	0.5929	0.7503	0.6680	0.5436	0.7924
Anatomical lung surgery	0.6682	0.5784	0.7580	0.6780	0.5512	0.8049
Mediastinal mass resection	0.9042	0.8246	0.9838	0.8554	0.6764	1.0000
Bilateral sympathectomy	0.9265	0.8400	1.0000	1.0000	1.0000	1.0000
Other procedures	0.9688	0.9167	1.0000	0.8125	0.4319	1.0000

Changes in the text: We have supplemented the data in Supplemental-Table S4.

Comment 3: One of the variables in the model is LVEF. I'm surprised that young people undergoing easy procedures had their LVEF measured. What are the rates of missing values in the series?
Reply 3: Thank you for the reviewer's valuable advice. LVEF is not necessary in young patients with simple surgery, such as hernia or appendix. But our institute has always insisted on to carry out cardiac ultrasound examination for most patients when undergoing thoracic surgery. The reasons are that, first, hidden heart disease can be ruled out, second, the evaluation of cardiopulmonary function before operation can be more perfect. The missing values in the series was about 707 in total 3023 cases, thus the rate of missing values was 23.4%.
Changes in the text: No change in the text.

Comment 4: In most risk models, predicted postoperative values are included instead of measured preoperative ones. Could you include that parameter in your model; if not, why not commenting on that in the discussion section?

Reply 4: Thank you for the reviewer's valuable advice. The original intention of our model is to predict postoperative complications through preoperative characteristics and intraoperative variables, which should not included the postoperative variables, otherwise there is no predictability.
Changes in the text: We have explained in our text as advised (see Page 11, line 169-172)

Comment 5: DLCO was not considered in the analysis. Any comments on that?
Reply 5: Thank you for the reviewer's valuable advice. DLCO is really an important examination, but it is not a routine examination in our institute. If the patient has poor pulmonary function, old age, or pulmonary interstitial disease, DLCO examination must be performed. In addition, if DLCO is included in the analysis of risk model, many cases without DLCO values may be excluded, which results in bias of results. We have given comments in discussion.
Changes in the text: We have modified our text as advised (see Page 13, line 234-237)

Comment 6: I understand the term "neutrophil rate" in the text as the percent of neutrophils in preoperative blood test. High percent could mean acute lung or pleural infections and, consequently, emergency surgery for that. Emergency surgery usually encompasses higher risk for the patient and shouldn't be included in the analysis.
Reply 6: Thank you for the reviewer's valuable advice. Emergency surgery shouldn't be included in the analysis, and we should give more detailed exclusion criteria.
Changes in the text: We have modified our text as advised (see Page 5, line 79)

Comment 7: The score RCRI should be referenced. I'm suggesting the modified using thoracic "ThRCRI" (doi: 10.1016/j.athoracsur.2010.03.042.) instead as a better parameter for thoracic patients; its use has been validated in several papers.
Reply 7: Thank you for the reviewer's valuable advice. The "ThRCRI" is a better parameter for referenced in thoracic patients and has been validated in several papers, but Wotton's observation has revealed that both RCRI and ThRCRI scores failed to accuratedly predict the risk of cardiac complications in patients undergoing elective resection of lung cancer (1). In addition, the RCRI is also used for evaluation in thoracoscopic surgery (2). Therefore, we continue to use RCRI to evaluate the cardiac function.

1.Robin Wotton, Andrea Marshall, Amy Kerr, et al. Does the revised cardiac risk index predict cardiac complications following elective lung resection? J Cardiothorac Surg 2013, 1;8:220.

2.Woo Sik Yu, Hee Suk Jung, Jin Gu Lee, et al. Safety of thoracoscopic surgery for lung cancer without interruption of anti-platelet agents. J Thorac Dis 2015, 7 (11):2024-32.

Changes in the text: No change in the text.