Reviewer A

The authors report risk factors associated with insufficient wedge resection margin for lung cancer, suggesting a predictive model and the derived nomogram to estimate the risk of the margin. Although this manuscript is well written, there are several crucial problems that should be clarified. My comments and questions are given below.

Comment 1: Major concern: I reckon that the nomogram suggested in this article should be evaluated in a validation set. A nomogram made in a training set is usually validated in a different set because how effectively it works remains unknown. In addition, some of the risk factors included in the nomogram were subjective and meaningless. For example, tumor's max distance to pleura was included although it was not significantly associated with the quality of wedge resection in the multivariable analysis. Furthermore, there were no good reasons why diabetes mellitus was a risk factor for insufficient margin. The authors insisted that this factor would be related to poorer quality of life, but the cardiovascular disease was not a significant factor. Therefore, I doubt the nomogram suggested in this article could work well in the real world, and the authors should evaluate it in the paper.

Reply 1: We greatly appreciate your insightful comments and suggestions. We fully agree that evaluating the nomogram in a validation cohort is superior to bootstrap resampling in the model development data set. Following the initial submission, we commenced the collection of clinical data, including a cohort of 180 patients who underwent pulmonary wedge resection from March 2023 to January 2024 in our center for internal validation. The model maintains good performance in the internal validation dataset, evidenced by a C-index of 0.761. In the future, an external validation will still be needed to further confirm the generalization of our model in a multicenter setting.

Both you and the third reviewer raised doubts about the inclusion of diabetes and other quality of life-related factors in the final model and suggested that these factors be excluded from the predictive model. Despite the statistical correlation between quality of life-related factors and wedge resection margins, we concur that there is no definitive technical correlation with whether the wedge surgical margin is guideline concordant or not. The explanation that factors like diabetes may inadvertently lead surgeons to reduce the resection range is neither sufficient nor persuasive. Consequently, we have omitted these factors from the multivariable regression analysis and developed a revised predictive model. The revised model attained satisfactory predictive performance, with a slight decrease in the C-index from 0.754 to 0.720.

The updated model still demonstrates a marginally significant correlation (OR=1.337, P=0.06) between the tumor's max distance to pleura and guideline discordant resection margin. Including the tumor's max distance to pleura in the final model, as opposed to its exclusion, appears to enhance the model's predictive performance to a certain extent. Meanwhile, max distance to pleura is clinically correlated with wedge resection. An increased max distance to pleura is generally understood to augment the range of resection and

surgical complexity, potentially compromising the attainment of guideline concordant surgical margin. Therefore, we will still retain this variable in the final model.

Changes in the text: We have revised the predictive model and validated its predictive performance in an internal validation cohort (see Table 3-5, Figure 1-2).

Comment 2: Minor concerns: Line 111: How did you measure tumor's max and minimal distance to pleura? Please clarify the definition.

Reply 2: We appreciate this suggestion. Tumor's max and minimal distance to the pleura, which were used for risk evaluation and predictive model construction, were evaluated based on the preoperative CT examination. Experienced thoracic surgeons, with more than five years of practice, analyzed the preoperative 3D reconstructed chest CT images and measured the distance from tumor to the nearest pleura for each patient. Meanwhile, we also verified our measurements with the CT reports when available.

Changes in the text: We have improved the description about the measurement of tumor's max and minimal distance to pleura in the Methods section (Page 8, Line 125-129).

Comment 3: Line 113: Some patients underwent wedge resection combined with segmentectomy or lobectomy. In these patients, was the purpose of wedge resection to confirm the pathological malignancy in frozen specimen? If so, these patients should not be included in this study.

Reply 3: We thank you for your comment. Patients undergoing combined lobectomy or segmentectomy with wedge resection presented with multiple primary lung tumors distributed in different lobes or segments. Lobectomy or segmentectomy was performed on the main lesions, while wedge resection was applied to the remaining tumors simultaneously. The previous phrasing could lead to misunderstandings. We have amended the text accordingly in the Results section (Page , Line).

Changes in the text: We have amended the expression in the text (Page 11, Line 172-174).

Comment 4: Line 121: How was the margin distance evaluated? It is a crucial point that should be clarified. Please confirm it according to the paper from Goldstein (Am J Clin Pathol 2003;120:720-4.).

Reply 4: We appreciate your comment. In this study, the wedge resection margin distance was measured using the gross cut-surface margin distance method described by Goldstein et al. During margin distance measurement, the tumor was transected, and its perimeter was delineated. Subsequently, the distance from the tumor to the cutting edge was measured. The description of margin distance evaluation has been included in the Methods section (Page 8, Line 122-125).

Changes in the text: We have added the methods for margin distance evaluation in the Methods section (Page 8, Line 122-125).

Comment 5: Line 165: You mentioned that 37% of the surgeries were guideline discordant. Additional resections were performed for these patients? If so, how did you calculate the margin for the additional specimens?

Reply 5: Thank you for your comment. In our department, intraoperative macroscopic evaluation of tumor margin was regularly conducted by surgeons after the resection of pulmonary tumor and the associated lung tissue. If margin distance was considered inadequate by surgeons during surgery, a re-resection would be carried out to ensure a more optimal margin. Precise measurement of resection margin distance, however, was conducted postoperatively, while intraoperative frozen section examination solely confirmed tumor malignancy and assessed the pathological status of margins. Consequently, most guideline discordant

margins were identified postoperatively, precluding the possibility of immediate re-resection. For patients receiving additional resection intraoperatively, pathologists would attempt to put the two excised specimens together and roughly estimate the combined margin distance. For those patients with additionally resected specimens, whose closest cutting edge was unable to identify, due to the lack of accurate margin distance data, they were excluded from this study.

Comment 6: Line 187: "Univariate" should be changed into "univariable".

Reply 6: Thanks for pointing this out. We have corrected the word in the text.

Changes in the text: We have corrected the word in the text (Page 12, Line 195).

Authors have performed a retrospective study to evaluate clinical features associated with NCCN-guideline discordant resection margin distance defined a ratio of <1 for resection margin distance to tumor size. Their findings support an admixture of poor clinical status parameters (age, BMI, poor pulmonary function, and diabetes) and surgical factors (preoperative localization, tumor size, distance from pleura) as being predictive of discordant (insufficient) margin. They combine these factors into a nomogram for preoperative prediction.

Comment 1: While the study size (530 patients) is adequate, the findings might be strengthened if authors also were able to show that guideline discordant margins were also associated with pathologic positive final margin or with recurrence. Micropapillary pattern ≥5% (PMID: 23926067), STAS (PMID: 25629637), and vascular invasion (PMID: 38204657) have further been associated with recurrence after wedge resection, and if this data is available these features might also be valuable to test as a function of guideline discordant margin status and/or nomogram prediction, strengthening the assertion that these clinical parameters are predictive of not just guideline discordance, but real risk of recurrence and poor outcome.

Reply 1: We greatly appreciate your comments and suggestions. Prior studies have demonstrated a significant association between inadequate or guideline discordant resection margins and the presence of malignant positive margins (Sawabata N et al. Ann Thorac Surg. 2004; El-Sherif A et al. Ann Surg Oncol. 2007; Mohiuddin K et al. J Thorac Cardiovasc Surg. 2014). Regrettably, our lung cancer database does not record sufficient postoperative survival and recurrence data, therefore, we are unable to analyze the correlation between insufficient margins and recurrence or patient survival. Concerning the assessment of malignant positive margins, we routinely send resected specimens for frozen section examination during wedge resection surgery to confirm the malignancy of tumors and to evaluate whether the margins are pathologically positive or not. Therefore, when performing radical surgery for early-stage lung cancer, very few patients have positive wedge resection margins.

We fully agree that pathological features including micropapillary pattern ≥ 5%, spread through air spaces (STAS), and vascular invasion are significantly associated with postoperative recurrence following wedge resection. These factors serve as robust indicators of local recurrence and are instrumental in enhancing or assessing our predictive model's performance. Given the extensive time span of patients included in this study (2014-2023), certain pathological features, particularly STAS, were not consistently documented by pathologists. The retrospective nature of this study complicates the re-evaluation of these pathological factors

for individual patients. Additionally, the present study's objective was to develop a preoperative predictive model; thus, postoperative pathological features were not the primary focus. Future research exploring the correlation between these pathological features and guideline-discordant wedge resection margins would be both interesting and valuable.

Changes in the text: We have discussed the lack of long-term survival or recurrence data as one of the study's limitations in the Discussion section (Page 18, Line 309-310).

Comment 2: The intro suggests that lobectomy is still the preferred surgical option for NSCLC, however based the two recent RCT's, sublobar resections are now being considered by many as standard of care (PMID 36780674 & 35461558).

Reply 2: Thanks very much for pointing this out. We have modified the expression about lobectomy and sublobar resection in the Introduction section (Page 6, Line 74-78).

Changes in the text: We have modified the expression about lobectomy and sublobar resection in the Introduction section (Page 6, Line 74-78).

Comment 3: Minor: Table 1: "(" in the wrong direction for minimal distance to pleural column.

Reply 3: Thanks for pointing out this mistake. We have corrected it in Table 1.

Changes in the text: We have corrected this error in Table 1.

Comment 4: Line 288 Discussion: states methanol fixation when it should be formalin fixation.

Reply 4: Thank you for pointing out this mistake. We have corrected the word in the Discussion section.

Changes in the text: We have corrected the word in the Discussion section (Page 17, Line 302).

<mark>**Reviewer C**</mark>

The study evaluates the possibility to build a predicting nomogram for insufficient margins in wedge resections. The authors performed the analysis on a retrospective series of 932 patients. Results show a predicting model which includes diabetes, age, BMI, longer max distance to the pleura and larger tumor size.

Comment 1: An insufficient resection margin is well defined by the authors; however, this is a result of the anatomy, the surgical technique, and the final decision of the surgeon. If the surgeon doesn't want to remove too much parenchyma and cut too much close to the lesion, this is an intentional insufficient margin, and you cannot build a predictive model for this situation. In your study you included diabetes, age, and low respiratory function in order to do this; I fell this part completely unsensed.

You should concentrate your model, if you want do rewrite the paper, on the technical issues able to reduce the "radicality" of a wedge, I mean with the intention to guide the surgeon between atypical and typical resections.

Reply 1: We very much appreciate your suggestions. Both you and the first reviewer have expressed concerns regarding the inclusion of diabetes and other quality of life-related factors in the final model, recommending their exclusion from the predictive model. As we replied to the first reviewer, we concurred

that these factors lack a direct technical correlation with guideline discordant wedge surgical margin. The rationale that quality of life factors might lead surgeons to inadvertently limit the extent of resection was deemed insufficient and unconvincing. Consequently, we revised our predictive model by omitting these factors from the multivariable regression analysis and the ultimate predictive model. The newly developed model maintained satisfactory predictive performance, with a modest decrease in the C-index from 0.754 to 0.720.

Changes in the text: We have revised the predictive model (see Table 3, Figure 1-2).