Dear Editor:

Thank you for your letter and for your comments concerning our manuscript. Those comments are all valuable and very helpful for revising and improving our paper, as well as the important guiding significance to our researches. We have studied those comments carefully and have made revision in the paper. Revised portion can be found in the paper. The main corrections in the paper and the responds to the comments are as follows:

Reviewer #1

Comment 1: The first thing I need is to reduce the abstract, as it contains 388 words. I want to reduce it to fewer than 300 words. The keywords should be (Keywords, not Key words) and must align with the Mesh Keywords. There should be approximately five keywords, listed in alphabetical order

Reply: Thank you very much for your constructive comments. We have carefully revised the abstract and reduced its length from 388 words to fewer than 300 words as requested (298 words). In addition, we have corrected the formatting of the keywords section and ensured that the selected five keywords are consistent with MeSH terms and listed in alphabetical order. We greatly appreciate your constructive suggestions, which have helped us to further improve the clarity and quality of our manuscript.

Comment 2: Additionally, you initially selected 128 people, but some were excluded, resulting in a final count of 109. I hope you can explain the reasons for the exclusions and the reason for choosing this number in the method section

Reply: We thank the reviewer for the helpful comment. We have clarified the reasons for patient exclusion in the *Methods* section. Specifically, among the 128 adult recipients who underwent right trilobe SLT, 19 patients were excluded: re-transplantation (n=7), because their surgical background and prognosis differ from primary transplants; combined multi-organ

transplantation (n=10), because these cases involve additional factors beyond isolated SLT; and loss to follow-up (n=2), because their outcomes could not be assessed. Therefore, 109 patients were finally included for analysis. Modify the Methods section as follows:

"During the study period, 128 adult recipients underwent right trilobe split liver transplantation (SLT). Of these, 109 patients met the study criteria and were included in the final analysis. Patients were excluded if they underwent re-transplantation (n=7, different surgical background and prognosis), combined multi-organ transplantation (n=10, additional confounding factors), or were lost to follow-up (n=2, outcomes not available)."

Comment 3: Define "early postoperative complications" more explicitly (are these within 30 days?)

Reply: We sincerely thank you for this valuable comment. In our study, *early postoperative complications* were defined as those occurring within 30 days after surgery. We have now clarified this definition in the **Methods section** (*Data sources* section) to avoid any ambiguity. The content is provided below and has been integrated into the revised manuscript:

"Early postoperative complications were defined as complications occurring within 30 days after surgery, including intra-abdominal hemorrhage, vascular complications, biliary complications, ascites, and infections."

Comment 4: Provide rationale for excluding Child–Pugh score in multivariate model, beyond collinearity.

Reply: We sincerely thank the reviewer for this insightful comment. The Child–Pugh score was not included in our multivariate model for the following reasons, beyond the concern of potential collinearity. First, the Child–Pugh score is a composite variable that incorporates several clinical and laboratory parameters (bilirubin, albumin, INR, ascites, and encephalopathy), some of which were already included as covariates in our

analysis. Including the score alongside its individual components could therefore result in redundancy and obscure the contributions of these parameters. Second, the Child–Pugh classification, while useful for broad stratification of liver function, is less granular and may introduce subjectivity due to its semi-quantitative nature (e.g., assessment of ascites and encephalopathy). By focusing on objective, continuous laboratory parameters, our model allows for a more precise evaluation of prognostic factors. Finally, exclusion of the Child–Pugh score also avoids over-adjustment, ensuring that the model remains parsimonious and interpretable.

Comment 5: Include an example of how the nomogram could be applied clinically.

Reply: We sincerely appreciate the reviewer's valuable suggestion. To address this point, we have now included a representative example to demonstrate how the nomogram could be applied in clinical practice. Specifically, we described a hypothetical patient with defined values of SII, MELD score, intraoperative blood loss, and IV segment resection status, and then illustrated how the total score is calculated and translated into an individualized risk prediction. This example has been added to the revised manuscript (Results section, *To develop a diagnostic nomogram for predicting EPC*), which we believe makes the clinical applicability of our nomogram clearer and more intuitive for readers. Thank you again for this constructive comment, which has helped us improve the clarity and clinical relevance of our study. The content is provided below and has been integrated into the revised manuscript:

"To illustrate the clinical applicability of the nomogram, we present a representative case. A patient with an SII value of 5.87, intraoperative blood loss of 2500 mL, no resection of segment IV, and a MELD score of 11 achieved a total nomogram score of 159. According to the model, this score corresponded to a 44.6% predicted probability of developing early

postoperative complications. This example highlights how the nomogram can be used to translate individual patient characteristics into a quantitative risk estimate, thereby supporting perioperative risk stratification and tailored management strategies."

Comment 6: Add calibration plots and decision curves in the main text instead of supplementary for a stronger impact.

Reply: We sincerely thank the reviewer for this valuable suggestion. We would like to clarify that both the calibration plots and the decision curve analyses have already been presented in Figure 3 of the main text, rather than in the supplementary materials. We fully agree with the reviewer's point on the importance of these analyses, which is why we had included them in the main figures from the initial submission. To further highlight their impact, we have revised the figure legend and relevant description in the Results section to explicitly emphasize the role of calibration and decision curve analyses in supporting our findings. We hope this clarification and revision adequately address your concern.

Comment 7: Simplify SHAP explanations for clinical readers (e.g., a plain-language description).

Reply: We sincerely thank the reviewer for this valuable suggestion. We agree that a plain-language description of SHAP would make our manuscript more accessible to clinical readers. In the revised version, we have simplified the explanation of SHAP values by removing overly technical details and providing a concise, reader-friendly description of how SHAP highlights the relative contribution of each feature to the prediction (Results section, *Risk Biomarkers for EPC Predicted by Machine Learning Models*). This adjustment aims to enhance clarity for clinicians without compromising the scientific rigor of the analysis. The content is provided below and has been integrated into the revised manuscript:

"To enhance interpretability, we applied SHAP analysis, a method that helps explain how each factor influences the model's predictions. This approach showed not only which variables were most important overall but also how they affected risk for individual patients. The six strongest predictors of early postoperative complications were higher LnSII, AFR, and MELD scores, greater intraoperative blood loss, longer operative time, and IV WPL"

Comment 8: Table 1 and 2 are dense—consider highlighting statistically significant results in bold for clarity.

Reply: We sincerely thank the reviewer for this valuable suggestion. To improve clarity and readability, we have revised Tables 1 and 2 by highlighting the statistically significant results in bold.

Comment 9: Expand more on how the findings could guide future clinical trials or perioperative management protocols.

Reply: We appreciate the reviewer's valuable suggestion to further elaborate on how our findings may inform future clinical trials and perioperative management. We added this description to the Discussion:

"Our findings offer several tangible implications for the design of future clinical trials and the refinement of perioperative management protocols in split liver transplantation (SLT). The predictive model, incorporating variables such as segment IV resection, MELD score, SII (systemic immune-inflammation index), and intraoperative blood loss, provides a framework for risk stratification that could be prospectively validated. Specifically, the nomogram's ability to estimate the probability of early postoperative complications enables the identification of high-risk recipients at the time of transplant evaluation. This risk quantification could guide the development of adaptive trial designs, where patients with higher predicted risk scores are allocated to more intensive monitoring arms or novel protective interventions—such as tailored immunosuppression regimens or enhanced

hemodynamic support—while standard management is maintained for lower-risk patients. Furthermore, the strong predictive value of intraoperative blood loss and the SII underscores the importance of targeting modifiable factors; for instance, future trials could test standardized surgical techniques to minimize blood loss or prehabilitation protocols to optimize the recipient's inflammatory and immune status before surgery. In routine clinical practice, the nomogram could be integrated into perioperative checklists to assist multidisciplinary teams in making data-driven decisions regarding postoperative care intensity, such as the timing of ICU discharge or the threshold for diagnostic imaging, thereby fostering a more personalized and proactive approach to managing SLT recipients."

Comment 10: Acknowledge potential overfitting due to small sample size when using ML.

Reply: We acknowledge your important point regarding the potential for model overfitting, a legitimate concern given our sample size when employing multiple machine learning algorithms (Random Forest, XGBoost, Support Vector Machine, and Logistic Regression). To proactively address this, we implemented a rigorous validation strategy centered on repeated ten-fold cross-validation, which provides a more reliable estimate of model performance on unseen data compared to a single train-test split. Furthermore, within the training folds of each cross-validation cycle, we performed hyperparameter tuning using a randomized or grid search to prevent the models from fitting the noise in the training data. The consistent performance of our final model, as evidenced by the cross-validated metrics, and the fact that the most influential variables identified (MELD, blood loss) have strong biological and clinical plausibility in the context of liver transplantation, provide reassurance that the model has captured a generalizable signal rather than idiosyncratic noise. We also acknowledge that the limited sample size remains an inherent limitation of our study, and

future validation using larger, multi-center cohorts will be essential to further confirm the robustness and clinical applicability of our prediction model. We added this description to the limitation:

“Fourth, the relatively small sample size may limit the generalizability of our findings. Future studies with larger cohorts are warranted to further validate the performance of the model. Additionally, we plan to incorporate multiple statistical approaches to minimize the risk of overfitting associated with limited sample size.”

Comment 11: Grammer corrections 1. in the abstract-Result "intraoperative blood loss, and operation time as key contributors. "should be changed with "intraoperative blood loss, and operation time as key contributors to the model. " 2. "Kaplan–Meier analysis indicated that SII and MELD score" should be changed to "Kaplan–Meier analysis indicated that SII and MELD scores" 3. Core Tip: This study applied an interpretable machine learning framework to evaluate risk factors for early postoperative complications (EPC) in adult recipients of right tri-segment split liver transplantation (SLT). should be changed with "Core Tip: This study employed an interpretable machine learning framework to assess risk factors for early postoperative complications (EPCs) in adult recipients of right tri-segment split liver transplantation (SLT). "

Reply: We thank the reviewer for carefully pointing out the grammatical issues in our manuscript. We have thoroughly revised the text according to the suggestions, and the corrected version has been incorporated into the revised manuscript. We sincerely appreciate the reviewer’s attention to detail, which has helped us improve the clarity and readability of the paper.

Reviewer #2 This manuscript by Wang et al. presents a timely and clinically valuable study on the development of a model for predicting early

complications (ECP) after split liver transplantation. The ability to stratify patients based on their risk for adverse events such as bile leakage, hemorrhage, and impaired graft function is a significant challenge in clinical practice. Therefore, a reliable predictive model has the potential to directly impact post-operative management, guide more intensive monitoring for high-risk individuals, and ultimately improve patient outcomes. The authors are to be commended for tackling this complex problem and for presenting their data in a clear and well-structured manuscript. The results appear robust and hold considerable promise for clinical translation.

Comment 1: Structural Clarification of the Background Section: The authors currently introduce specific examples of ECP (e.g., bile leakage, hemorrhage) within the introduction. While this provides immediate context, I suggest that this detailed clinical information would be more impactful if integrated into a dedicated background section. The introduction could be sharpened to briefly state the clinical problem of ECP in broad terms, while a subsequent background section could more thoroughly review the incidence, clinical consequences, and economic burden of specific complications like bile leakage and hemorrhage. This structural change would allow the introduction to focus on establishing the knowledge gap—the lack of a reliable predictive model—while the background section would build a stronger, more compelling rationale for why filling this gap is so critically important. It would help the reader fully appreciate the clinical significance of the study before delving into the methodology.

Reply: Thank you for this insightful suggestion. Following your recommendation, we have revised the manuscript to better distinguish the Introduction and Background sections:

"Split liver transplantation (SLT) is associated with an increased risk of early post-transplant complications (EPC), which present substantial clinical challenges. SLT involves greater technical complexity and introduces unique

surgical and physiological stresses. The extensive transection surface of partial grafts, coupled with resultant hemodynamic alterations, predisposes recipients to EPCs such as bile leakage, hemorrhage, and impaired graft function. These complications can significantly worsen patient outcomes by elevating morbidity and mortality. Bile leakage, one of the most common EPCs, is influenced by both surgical technique and graft quality and is linked to prolonged hospitalization, higher infection risk, and potential graft loss. Hemorrhage represents another critical complication, directly contributing to increased morbidity and mortality. Impaired graft function, ranging from transient dysfunction to primary non-function, further complicates post-transplant recovery, with incidence varying according to donor and recipient characteristics. Collectively, these complications impose a heavy clinical burden, resulting in extended intensive care unit stays, higher healthcare costs, and reduced patient quality of life. The unpredictability of EPC highlights a crucial knowledge gap: the lack of reliable predictive tools to identify high-risk patients and inform preventive strategies. Developing such a model is essential to optimize outcomes, improve resource allocation, and enhance the long-term success of SLT. The present study seeks to address this gap by identifying predictive factors for EPC, with the ultimate goal of refining clinical decision-making and perioperative management."

Comment 2: Enhanced Accessibility of the Machine Learning Methodology: In the section titled "Risk Biomarkers for ECP Predicted by Machine Learning Models," the authors present their findings effectively. However, to make the manuscript more accessible to its broad intended audience, which undoubtedly includes clinicians and surgeons, I recommend adding a brief primer on the core principles of the specific machine learning algorithms employed. For instance, a short paragraph explaining the fundamental logic of a Random Forest (as an ensemble of decision trees) or a Support Vector Machine (as finding an optimal hyperplane for classification) would be

immensely helpful. This does not require a deep dive into mathematical complexities but should focus on the intuitive logic behind why each algorithm was chosen and how it generally operates to identify patterns. This addition will demystify the methodology for readers less familiar with computational techniques, thereby enhancing the manuscript's educational value, fostering interdisciplinary dialogue, and strengthening the perceived validity of the approach.

Reply: We sincerely thank the reviewer for their insightful suggestion to enhance the accessibility of the machine learning methodology for a broader audience, including clinicians and surgeons. We agree that providing an intuitive explanation of the core principles behind the machine learning algorithms used in our study will improve the manuscript's educational value and foster interdisciplinary dialogue. To address this, we have added a concise primer in the "*Machine learning*" section, which explains the fundamental logic, their rationale for selection, and how they identify patterns relevant to predicting early complications post-liver transplantation (ECP). The primer avoids mathematical complexities and is tailored to readers less familiar with computational techniques. The content is provided below and has been integrated into the revised manuscript:

"Comprehensive perioperative donor and recipient variables were evaluated using four machine learning algorithms (Random Forest, Support Vector Machine, XGBoost, and Logistic Regression). RF is an ensemble method that builds multiple decision trees, each acting like a flowchart that makes predictions based on patient features (e.g., age, or biomarker levels). RF combines the predictions of all trees through voting to produce a robust result. Its ability to capture complex, non-linear relationships in data makes it well-suited for identifying ECP risk factors in our diverse dataset. SVM identifies an optimal boundary (or hyperplane) that separates patients into groups, such as those likely to develop ECP versus those who are not, based on clinical features. It maximizes the separation between groups to ensure reliable

classification. SVM's strength in handling high-dimensional data makes it effective for pinpointing key ECP predictors. XGBoost is an advanced ensemble method that builds decision trees sequentially, with each tree correcting errors from the previous ones. This iterative approach enhances prediction accuracy, particularly for complex datasets. XGBoost has high performance in identifying subtle patterns in clinical data. LR is a statistical method that predicts the probability of an outcome (e.g., ECP occurrence) by analyzing relationships between patient features and the outcome. It is straightforward and interpretable, making it valuable for identifying linear relationships between clinical variables and ECP risk."