nature portfolio

Peer Review File

Perioperative Tislelizumab Plus Lenvatinib Treatment for Resectable Hepatocellular Carcinoma at High Risk of Recurrence: Single-arm Phase II Trial

Corresponding Author: Professor TianQiang Song

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The authors report the results of a phase 2 trial testing neoadjuvant tislelizumab and lenvatinib in patients with resectable HCC. They report that this regimen is relatively well tolerated with low high grade AE rates, feasible, and has potential efficacy with an ORR of 55.6%. Pathologic responses were also quite high among resected patients.

I consider the novelty of this study to be low, as 4 prior neoadjuvant trials of similar sample sizes have already been reported and published in 2021 and 2022, including one testing a TKI plus anti-PD-1 combination. The authors include some of these prior studies as references, but do not incorporate them in the introduction or discussion. In my view, the feasibility of neoadjuvant therapy in HCC was also previously established. What is needed is a proof of efficacy with improved long-term outcomes which requires at least a randomized study. As the authors are also testing another anti-PD-1 ab with an established anti-VEGF TKI, I'm not sure of the scientific advancement that is to be made here.

While the high ORR reported in this study may be interesting, it is contrast to prior data from large randomized studies particularly of lenvatinib plus immunotherapy which may suggest limited generalizability of the trial results. In fact, the reported ORR rate of tislelizumab and lenvatinib is more than double the ORR rate of pembrolizumab plus lenvatinib. Are the authors suggesting that one anti-PD1 antibody is differentiated from another enough to be responsible for such a dramatic improvement in tumor response rates? If so, some biological explanation is needed.

The TabPFN classifier is interesting, but very preliminary. I'm not sure if this adds anything to the paper as there is no clinical validation or new mechanistic understanding.

Reviewer #2

(Remarks to the Author)
Chen et report a safety and efficacy study of tislelizumab plus lenvatinib as perioperative treatment in patients with resectable hepatocellular carcinoma (HCC) at high risk of recurrence. Overall, the manuscript addresses an important clinical question and presents encouraging early results (compared to previous perioperative studies). The study's limitations include its single-center design, single-arm approach, and small sample size. I have several suggestions for strengthening the clinical relevance of this study.

Specific comments:

1. Due to differences in clinical backgrounds of the enrolled patients and regional variations across different clinical trials, it is difficult to directly compare the efficacy among clinical trials. Since this is a single-center study involving initially resectable patients, would it be possible to retrospectively analyze the prognostic outcomes of patients with matched clinicopathological factors who underwent upfront surgery, as a reference?

2. The authors should clarify why only 9 of 20 surgical patients received adjuvant tislelizumab–lenvatinib. For example, was this due to patient preference, performance status, toxicity, or other selection factors? Given the DFS difference between

adjuvant and non-adjuvant groups, detailing these reasons and baseline characteristics would help assess potential confounding factors.

3. Given the high MPR rate in this study (60%, which is 2-3 times higher than that reported in other perioperative HCC trials), the authors should report the absolute number and percentage of MPR patients who experienced recurrence, and the pattern of recurrence. This would clarify whether the high pathological response rate translated into lower recurrence risk and aligns with standard perioperative trial reporting.

4. Regarding the feasibility of the study, the definition of a delay in surgery of 49 days or more appears to be rather extensive. The authors should provide a thorough basis for this definition. Moreover, 70.4% (n=19) of patients who underwent surgery within 49 days appears to be a low rate. A thorough discussion of this result is necessary.

5. The impact of lower dose administration on the therapeutic effect warrants further investigation. The findings, which indicated that patients who achieved CR exhibited a more favorable prognosis compared to those who underwent surgical intervention, suggest that enhancing the therapeutic efficacy of preoperative treatment may lead to improvements in patient outcomes. If the dosage was reduced due to safety concerns, it would have been preferable to verify the appropriate dosage in advance. The authors should provide a rationale for their decision to forego preliminary testing.

6. The predictive model (based solely on hematologic factors) performed well in a single small cohort, but comparing it with standard clinical factors (e.g., tumor size, number, vascular invasion, stage) or integrating both would help demonstrate added value beyond established criteria.

7. As stated in the Limitations, it is imperative to ascertain the impact of markers of responsiveness derived from machine learning models on long-term prognosis. Concurrently, the impact on short-term outcomes should be examined. The finding that patients who achieved MPR exhibited a high rate of early recurrence is counterintuitive.

Minor comments:

1. Please verify the annotations in Figure 1B. The main text describes that PD patients did not undergo surgery. However, Figure 1B shows that one PD patient had surgery. Additionally, Figure 1B indicates that 22 out of 27 patients underwent surgery, which is inconsistent with the 20 cases described in the Results section.

2. The results and notation in Extended Data Fig. 1 appear to be reversed. In the K-M curve, the group with “Adjuvant therapy: No” appears to have a DFS of 18.8 months and a 1YDFS rate of 77.8%.

3. The authors stated that "only one of them experienced tumor recurrence at the data cutoff point, which indicated that the necessity of surgery for this population needs further discussion, as previously reported." Yet no recurrence can be confirmed from Figure 1. The figure must be revised.

4. In Fig.2D, the vertical axis should be labeled "Overall survival" rather than "DFS."

Reviewer #3

(Remarks to the Author)
The authors present a small prospective study about the use of perioperative immunotherapy plus TKI for resection HCC. The authors developed a machine learning model to predict preoperative response after combination therapy and validated is in an external test cohort. The single-arm phase II study is well designed and the manuscript is well written, with appropriate review of existing literature. The methodology for the design and validation of the machine leanring model is comprehensive and well described. However there are some areas which need to be considered and addressed:

Major comments

-The results demonstrate a mDFS of 18.8 months for 20 patients who received perioperative therapy and surgery. However there is no control arm for the study and thus comments on the effectiveness of this regimen (as stated at the end of the abstract -line 38) cannot be fully made based on this analyses.

-There is heterogeneity in the perioperative regimens patients recieved, with only 9/20 receiving adjuvant therapy after resection. DFS rates were comparable between those who did and did not have adjuvant therapy. The implications on this finding for the optimal perioperative regimen needs to be addressed and the exact regimen proposed by the authors must be made clear.

-For the prediction model, the sample size of both the training and testing cohorts are small. There is a class imbalance between responders and non-responders in both training (84% vs 16%) and testing (67% vs 33%) cohorts. This can impact the model performance in larger datasets. Therefore the results of the predictive model must be interpretated cautiously. Much larger cohorts are needed for robust testing of this model.

-The authors suggest in the conclusion the TabPFN model identifies populations unlikely to respond to perioperative treatment. This appears to contradict the confusion matrix in Figure 3D, which suggests the model has a high sensitivity for detecting responders (100% in both cohorts) but low specificity (33.0%) in the testing set. The model only detected 2 out of 6

non-responders and therefore this conclusion cannot be made based on the data presented. This low specificity means a large proportion of non-responders may receive perioperative therapy and risk disease progression prior to surgery.

Introduction

-Define perioperative intervention timing (i.e. pre- and post- resection)

line 78 - based on BCLC guidelines lenvatinib is not first-line for advanced HCC (only if ICI not feasible). Would suggest rephrasing to "is a recommended treatment for advanced HCC"

-Introduction is very well written but quite long. I would suggest moving some of the text to the Discussion. For example:

line 72-77: "In a phase II study, preoperative..... 1-year RFS of 53.85%"

and

line 79-87 : "Yi et al. reported .... antitumour activity in advanced HCC"

and

line 99-107: "Currently, research on prognostic ... Cox regerssion is often challenging"

Methods
For the predictive model please state clarify the target variable. In line 429 it is unclear if the Response and Non-Resposne is referring to pre-operative status only.

Results
-On line 142 the authors state one patient experienced TRAE and refused surgery but on line 150 they state one patient experienced nerve injury leading to surgical cancellation. Could the authors clarify if this is the same patient and whether the cancellation of surgery was due to patient choice or a clinical decision?

-For safety results please outline how many TRAEs led to treatment discontinuation.

-For the efficacy in line 155 please state this is referring specifically to preoperative ORR and DCR.

-The results for DFS and survival are a little confusing to read, particurlarly the subgroup. Consider reporting (1st) the DFS and OS for the 20 patients who had perioperative therapy and surgery (2nd) DFS/OS for those who did and did not have adjuvant therapy (3rd) DFS/OS for MPR patients (4th) DFS/OS for surgical vs non-surgical

-Consider using statistical comparision of characteristics between the training and validation set (Extended Data Table 3) and commenting on any noteable differences in the results text.

-Please outline the six features selected for the final model in the main text.

Discussion
-The authors outline in line 279 four patients did not undergo surgery due to achieving perioperative CR, with one of four experiencing tumour recurrence. This is a really important discussion point. However in the Results line 137 the authors state two of the three patients who refused surgery due to CR had disease progression. This needs clarification. Could the authors also state whether these patients had operable disease at the time of tumour recurrence/progression?

-Please tone down the use of "excellent" for the TabPFN model performance in line 297.

-The high sensitivity of the TabPFN in both cohorts is mentioned in the discussion but not implicitly stated in the results or figures (though can be determined from Figure 3D confusion matrix). Please state the sensitivity (recall) in the results (100% in training and testing set). I would also suggest stating specificity, especially in the testing set as this is only 33%. This low specificity means a large proportion of non-responders may receive perioperative therapy and risk disease progression prior to surgery. The authors should comment on this.

Reviewer #4

(Remarks to the Author)
This is a prospective single-arm Phase II trial investigating the efficacy and safety of peri-operative tislelizumab in combination with lenvatinib in patients with resectable hepatocellular carcinoma at high-risk of recurrence. Of the 20 patients who underwent surgery, 60% of patients achieved major pathological response, whereby 20% achieved pathological complete response. In the cohort of 27 patients, ORR were 55.6% and 81.5% per RECIST v1.1 and mRECIST respectively, and Grade 3 TRAEs were reported in 7.4% of patients. The median DFS was 18.8 months with a 1-year DFS of 74.0%. The

finding of this trial is clinical meaningfully and the authors proposed a predictive model for recurrence using this treatment combination based on this trial data, for better patient selection.

Comments
1. Line 170-172, could the authors please confirm if you mean to report the 1-year DFS or the median DFS as the 1-year DFS cannot be “not estimable”.
2. Could the authors please check the Extended Data Figure 3 and 2-year OS rates stated in the manuscript (Line 173-176). Which is the correct data as the 2-year OS rate for the no surgery cohort appears to be around or below the 75% mark on the y-axis.
3. The authors acknowledged that no hypothesis-based sample size computation was performed at the trial design stage.
4. Could the authors please include the method used to compute the median follow-up time?
5. It will be good if the authors can include the test used to compare the survival between groups (as reported in the manuscript, eg. Line 172) in the statistical section.
6. Line 412 to 414, could the authors explain how censoring was done if the patient is alive and recurrence-free at time of data cut-off? How often is the follow up for disease recurrence?

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
The author responses to my concerns do not address the underlying issues that I see regarding the lack of new clinical or scientific insights. This is another small non-randomized study which is largely affirming of prior similarly sized trials regarding the feasibility of neoadjuvant therapy in HCC. In particular, the previously conducted cabo nivo trial also already included high-risk HCC. The observed ORR is also extraordinary, but without clear explanation as to why drug effects in this study seem superior to other studies.

Reviewer #2

(Remarks to the Author)
The authors addressed all my comments. The only minor comment I still have is the use of the lower dose of lenvatinib in this study (8 mg/day). This may have contributed to lower toxicity, but also may have had different biological effects on the tumor, facilitating immune responses via effects on the vessels, as seem for lower doses of regorafenib, for example (see PMID: 33234602 and 38374347). This deserves some discussion.

Reviewer #3

(Remarks to the Author)
The authors have fully engaged with the comments and have answered them thoroughly.

Reviewer #4

(Remarks to the Author)
The authors have satisfactorily addressed all my previous comments; however, I have a comment on the revised version.

Comments
1. Line 319-321 Given the small sample size of 9 vs 11 patients in the groups of patients with and without adjuvant therapy, the study lacks power to detect any significant difference. Hence, the absence of a significant p-value cannot be used to implied that there is no significant difference. Would suggest the authors to rephrase the sentence and remove the p-value.

Comments:

Reviewer #1, Comments 1:

I consider the novelty of this study to be low, as 4 prior neoadjuvant trials of similar sample sizes have already been reported and published in 2021 and 2022, including one testing a TKI plus anti-PD-1 combination. The authors include some of these prior studies as references, but do not incorporate them in the introduction or discussion. In my view, the feasibility of neoadjuvant therapy in HCC was also previously established. What is needed is a proof of efficacy with improved long-term outcomes which requires at least a randomized study. As the authors are also testing another anti-PD-1 ab with an established anti-VEGF TKI, I'm not sure of the scientific advancement that is to be made here.

Response:

We sincerely appreciate your insightful comments regarding the novelty of our study. In response, we have added a detailed discussion in the manuscript (Discussion section, lines233-240) regarding the four representative neoadjuvant clinical trials published internationally between 2021 and 2022[1-4], including PD-1 monotherapy, PD-1 ± CTLA-4 combinations, and PD-1 plus TKI regimens.

Our study specifically focused on patients with resectable, high-recurrence-risk HCC, evaluating the feasibility and preliminary efficacy of neoadjuvant tislelizumab plus lenvatinib. Compared with the four previously published trials, our regimen demonstrates several advantages:

Most existing neoadjuvant studies: including cemiplimab (2022), nivolumab ipilimumab (PD-1 plus CTLA-4) (2022), cabozantinib plus nivolumab (2021), and camrelizumab plus apatinib (2022): had inclusion criteria that differed from ours in terms of etiological background, enrolling populations with a lower proportion of viral hepatitis-related HCC, and treatment efficacy may therefore vary across populations. Cabozantinib plus nivolumab primarily targeted conversion of locally advanced tumors to resectable disease, while camrelizumab plus apatinib, although also a PD-1 plus TKI regimen, was associated with a higher incidence of treatment-related adverse events (TRAEs), with 88.9% of patients experiencing any-grade TRAEs and 16.7% experiencing grade ≥3 TRAEs; in contrast, tislelizumab plus lenvatinib in our study was associated with 74.1% any-grade and 7.4% grade ≥3 TRAEs. Furthermore, the camrelizumab plus apatinib trial enrolled a considerable proportion of patients with BCLC stage B (n=5) and C (n=13), highlighting the

heterogeneity in patient selection across neoadjuvant studies and the lack of consensus regarding the optimal target population, and indicating that conducting phase III trials in this setting remains relatively premature. Due to these differences and patient heterogeneity, results from prior studies cannot be directly extrapolated to patients with primary resectable, high-recurrence-risk HCC. In contrast, our study focuses on a more homogeneous, clinically relevant population and provides a feasible and safe neoadjuvant approach, highlighting its potential value for patients at high risk of recurrence.

- Novelty and Future Value

Therefore, our study is not merely a repetition of previous "feasibility" studies. It: Focuses on resectable, high-recurrence-risk patients , demonstrating the potential and feasibility of neoadjuvant therapy;Evaluates a PD-1 plus TKI combination that appears both safer and more effective in the neoadjuvant setting;Provides a foundation for future large-scale randomized controlled trials.Similar to advanced HCC treatment , where multiple PD-1 plus TKI combinations continue to be validated, neoadjuvant therapy also requires multiple early-phase exploratory trials to identify optimal regimens. Thus, the existence of the four prior studies does not imply that this field is fully mature.

- Summary

We fully agree that establishing standards for neoadjuvant therapy requires large phase III randomized trials. Nevertheless, our study provides meaningful contributions by focusing on resectable, high-recurrence-risk patients, evaluating a PD-1 plus TKI regimen with promising efficacy and acceptable safety, and offering early-phase evidence in a clinically relevant population. We hope that these findings can help guide the design of future large-scale studies and support further exploration of neoadjuvant therapy in HCC.

References:

1.Kaseb, A.O., et al. Perioperative nivolumab monotherapy versus nivolumab plus ipilimumab in resectable hepatocellular carcinoma: a randomised, open-label, phase 2 trial. Lancet Gastroenterol Hepatol 7, 208-218 (2022).

2.Marron, T.U., et al. Neoadjuvant cemiplimab for resectable hepatocellular carcinoma: a single-arm, open-label, phase 2 trial. Lancet Gastroenterol Hepatol 7, 219-229 (2022).

3.Ho, W.J., et al. Neoadjuvant Cabozantinib and Nivolumab Converts Locally

Advanced HCC into Resectable Disease with Enhanced Antitumor Immunity. Nat Cancer 2, 891-903 (2021).
4.Xia, Y., et al. Efficacy and safety of camrelizumab plus apatinib during the perioperative period in resectable hepatocellular carcinoma: a single-arm, open label, phase II clinical trial. J Immunother Cancer 10 (2022).

Reviewer #1, Comments 2:

While the high ORR reported in this study may be interesting, it is contrast to prior data from large randomized studies particularly of lenvatinib plus immunotherapy which may suggest limited generalizability of the trial results. In fact, the reported ORR rate of tislelizumab and lenvatinib is more than double the ORR rate of pembrolizumab plus lenvatinib. Are the authors suggesting that one anti-PD1 antibody is differentiated from another enough to be responsible for such a dramatic improvement in tumor response rates? If so, some biological explanation is needed.

Response:

We appreciate the reviewer's insightful comment regarding the high ORR observed in our study. Motivated by this suggestion, and considering our interest in exploring potential biological factors underlying this response, we examined the role of the tumor immune microenvironment. Previous studies have shown that effective antitumor immune responses, particularly T cell-mediated immunity, are closely associated with favorable responses to immunotherapy[1]. Additionally, the presence and maturation of tertiary lymphoid structures (TLSs) within tumors have been reported to provide a supportive microenvironment for T cell priming and activation, thereby enhancing antitumor immune responses, improved immune responses, and better clinical outcomes in HCC and other solid tumors[2-4].

Inspired by these findings, and considering the relatively high ORR in our cohort, we conducted an exploratory analysis of TLSs in patients with available tissue specimens (n=23, including 20 postoperative and 3 preoperative samples). Histological evaluation revealed abundant and well-developed TLSs in 13 patients (56.5%). Notably, among these 13 patients, 11 achieved an objective response (ORR, 84.6% RECIST v1.1). Consistently, of the 12 patients who achieved ORR in this cohort, 11 had TLSs present (91.7%).

For illustrative purposes, representative images of TLSs are provided in this response letter. Fig.R1A shows a representative tissue specimen with abundant,

well-formed TLSs located in the post-treatment necrotic area and at the tumor margin (TLSs-positive), while Fig.R1B shows a specimen lacking TLSs (TLSs-negative). These observations provide preliminary biological evidence supporting the association between TLSs and enhanced antitumor immune activity, suggesting the potential of TLSs as a biomarker for predicting therapeutic benefit in the neoadjuvant setting. These findings indicate that the high prevalence of TLSs in a considerable proportion of the samples from this cohort may be one factor contributing to the observed high ORR.

We also note that response rates may depend on patient characteristics, disease stage, and study context. Even for the same drug combination (Lenvatinib plus PD-1 inhibitor)[5], reported ORRs in previous studies differ substantially: 116-KEYNOTE-524 (Phase Ib, advanced HCC, lenvatinib plus pembrolizumab) reported an ORR of 36% (RECIST v1.1), LEAP-002 (Phase III, advanced HCC, lenvatinib plus pembrolizumab) reported an ORR of 26.1% (RECIST v1.1)[6], and BGB-A317-211 (Phase II, advanced HCC, lenvatinib plus tislelizumab) reported an IRC-RECIST-assessed ORR of 38.7% (RECIST v1.1)[7]. Our study enrolled resectable, high-recurrence-risk HCC patients across BCLC stages A-C (A: 48.1%, B: 44.4%, C: 7.4%), which may contribute to the observed response outcomes.

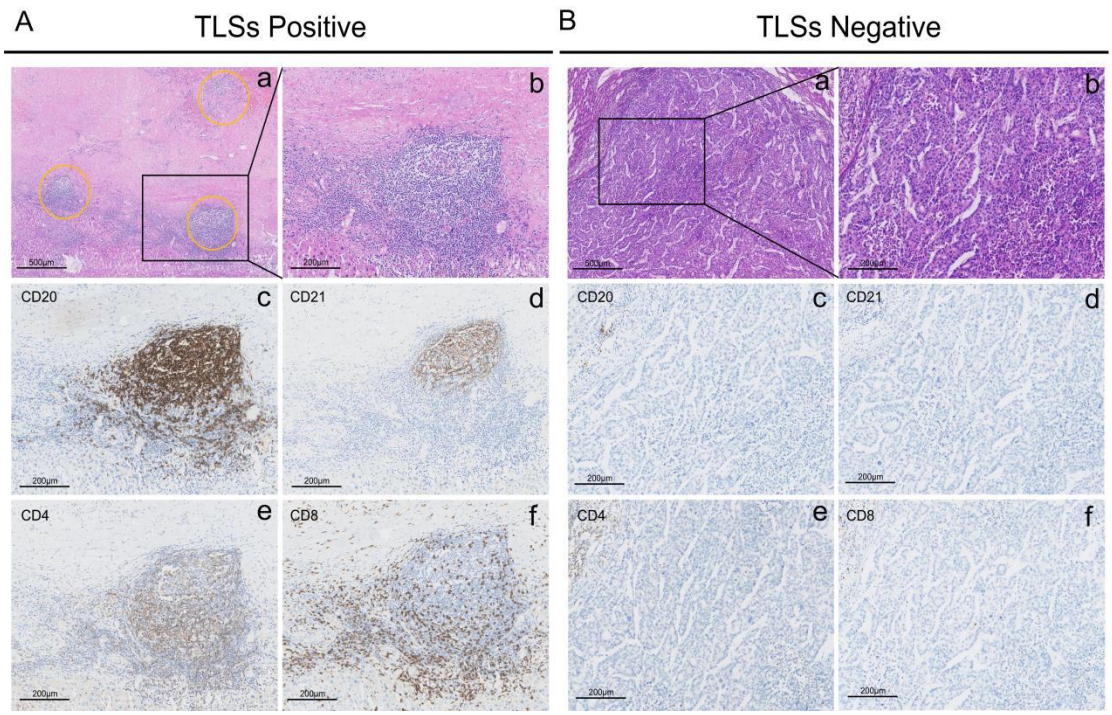


Fig.R1 Representative images of TLSs in HCC tissue specimens. (Aa,b) Representative images of the TLSs located in the post-treatment necrotic area and at

the tumor margin in H&E-stained sections. TLSs are highlighted by the yellow circles. Scale bar: 500μm, 200 μm. (Ba,b) Representative H&E-stained image of HCC tissue lacking TLSs. Scale bar: 500μm, 200 μm. (Ac, Bc) CD20 staining showed the number of B cells. Scale bar: 200 μm. (Ad, Bd) The FDC meshworks were labelled by CD21. Scale bar: 200 μm. (Ae,f; Be,f) CD4 and CD8 staining showed the number of T cells. Scale bar: 200 μm.

References:

1.Waldman, A. D., Fritz, J. M. & Lenardo, M. J. A guide to cancer immunotherapy: from T cell basic science to clinical practice. Nat. Rev. Immunol. 20, 651-668 (2020).

2.Su, J.Y., et al. Tertiary lymphoid structures in HCC: Influence on immune cell profiles in tumors and on efficacy of adjuvant PD-1 inhibitor therapy after hepatectomy. Hepatology (2025).

3.Teillaud, J.L., Houel, A., Panouillot, M., Riffard, C. & Dieu-Nosjean, M.C. Tertiary lymphoid structures in anticancer immunity. Nat Rev Cancer 24, 629-646 (2024).

4.Shu, D.H., et al. Immunotherapy response induces divergent tertiary lymphoid structure morphologies in hepatocellular carcinoma. Nat Immunol 25, 2110-2123 (2024).

5.Finn, R.S., et al. Phase Ib Study of Lenvatinib Plus Pembrolizumab in Patients With Unresectable Hepatocellular Carcinoma. J Clin Oncol 38, 2960-2970 (2020).

6.Llovet, J.M., et al. Lenvatinib plus pembrolizumab versus lenvatinib plus placebo for advanced hepatocellular carcinoma (LEAP-002): a randomised, double-blind, phase 3 trial. Lancet Oncol 24, 1399-1410 (2023).

7.Xu, L., et al. Efficacy and safety of tislelizumab plus lenvatinib as first-linetreatment in patients with unresectable hepatocellular carcinoma: a multicenter, single-arm, phase 2 trial. BMC Med 22, 172 (2024).

Reviewer #1, Comments 3:

The TabPFN classifier is interesting, but very preliminary. I'm not sure if this adds anything to the paper as there is no clinical validation or new mechanistic understanding.

Response:

We sincerely thank the reviewers for their comments. We are flattered to know that you also think that the TabPFN classifier we use is something of interest. But we

must clarify with you the fact that although the work we are doing is not complicated, it is a meaningful thing, and it is worth exploring and it must be explored by someone.

Although neoadjuvant therapy may bring potential long-term treatment benefits[1-2], its potential impact on the implementation of surgery is of greater concern[2-3]. For patients who could have been operated on, if they are insensitive to neoadjuvant therapy, the timing of operation is delayed or even the operation cannot be performed, which is of great clinical concern[4].

Therefore, most clinicians, not only in liver cancer, have begun to conduct preliminary exploration around the sensitivity of patients to neoadjuvant therapy[1,5-6], most of which focus on the discovery of potential predictive biomarkers. Limited by the sample size, although a number of studies have explored the value of biomarkers in predicting the efficacy of neoadjuvant therapy for liver cancer[7-9], there is still no appropriate and consensus model to guide the precise implementation of perioperative treatment for liver cancer. The answer to this question is still worth exploring.

At the same time, we found that TabPFN has good prediction performance on small samples, can deal with class imbalance, and shows good robustness on multiple datasets[10]. Although it has just been released for training on tabular data, we think it is very suitable for the neoadjuvant field, because small samples and class imbalance are common in the current neoadjuvant treatment of liver cancer[7-9], and there are many tabular data worth using in clinical practice, such as hematological indicators.

Can the TabPFN model constructed by hematological indicators provide a universal and simple option for identifying patients who are not sensitive to neoadjuvant therapy? It was a meaningful attempt. In fact, through our research exploration, it did show a high specificity (100%) for patients who are insensitive to neoadjuvant therapy as we expected, which indicates that it can accurately identify some patients who are insensitive to neoadjuvant therapy, thereby avoiding the delay of possible surgery for these patients. In addition, our model can provide a preliminary framework for identifying potential predictive biomarkers in HCC neoadjuvant therapy, which will guide future basic research and clinical validation efforts.

References:

1.Xia, Y., et al. Efficacy and safety of camrelizumab plus apatinib during the perioperative period in resectable hepatocellular carcinoma: a single-arm, open label, phase II clinical trial. J Immunother Cancer 10, (2022).

2.Llovet, JM., et al. Adjuvant and neoadjuvant immunotherapies in hepatocellular carcinoma. Nat Rev Clin Oncol 21, 294-311 (2024).
3.Hong, WX., et al. Neoadjuvant Intratumoral Immunotherapy with TLR9 Activation and Anti-OX40 Antibody Eradicates Metastatic Cancer. Cancer Res 82, 1396-1408 (2022).
4.Vogel, A., et al. Adjuvant and neoadjuvant therapies for hepatocellular carcinoma. Hepatology 82, 777-793 (2025).
5.Knochelmann, HM., et al. Neoadjuvant presurgical PD-1 inhibition in oral cavity squamous cell carcinoma. Cell Rep Med 2, 100426 (2021).
6.Laza-Briviesca, R., et al. Blood biomarkers associated to complete pathological response on NSCLC patients treated with neoadjuvant chemoimmunotherapy included in NADIM clinical trial. Clin Transl Med 11, e491 (2021).
7.Marron, T.U., et al. Neoadjuvant cemiplimab for resectable hepatocellular carcinoma: a single-arm, open-label, phase 2 trial. Lancet Gastroenterol Hepatol 7, 219-229 (2022).
8.Kaseb, A.O., et al. Perioperative nivolumab monotherapy versus nivolumab plus ipilimumab in resectable hepatocellular carcinoma: a randomised, open-label, phase 2 trial. Lancet Gastroenterol Hepatol 7, 208-218 (2022).
9.Li, Z., et al. Neoadjuvant tislelizumab plus stereotactic body radiotherapy and adjuvant tislelizumab in early-stage resectable hepatocellular carcinoma: the Notable-HCC phase 1b trial. Nat Commun 15, 3260 (2024).
10.Hollmann, N., et al. Accurate predictions on small data with a tabular foundation model. Nature 637, 319-326 (2025).

Reviewer #2, Comments 1:

Due to differences in clinical backgrounds of the enrolled patients and regional variations across different clinical trials, it is difficult to directly compare the efficacy among clinical trials. Since this is a single-center study involving initially resectable patients, would it be possible to retrospectively analyze the prognostic outcomes of patients with matched clinicopathological factors who underwent upfront surgery, as a reference?

Response:

We appreciate the reviewer's insightful comment regarding our study design. This study focused on the proof-of-concept of perioperative immunotherapy for HCC. Therefore, we chose a single-center, single-arm study design at the beginning of the study design, and no statistical hypotheses were made due to the lack of relevant reference for efficacy expectations. Consequently, we have avoided any comparative descriptions of efficacy and safety data as much as possible. Meanwhile, to what extent perioperative immunotherapy can improve the postoperative prognosis of patients with HCC is very curious and is the ultimate goal of research on perioperative treatment of HCC.

According to your advice, we have attempted to retrospectively collected clinical records of patients who underwent surgery alone with similar pathological features in our center. Unfortunately, because our center was conducting many clinical trials on perioperative treatment of HCC at the same time (e.g. NCT04701060, NCT05807776, NCT05621499, NCT05519410, NCT04521153, ChiCRT-2400080080), most patients were enrolled in each clinical trial, and only a few eligible patients meet such requirements. Such that we could not eliminate bias between groups by means of, for example, propensity-score matching. Further, it is not feasible to compare the positioning of effectiveness data obtained in this study with retrospective controls.

However, we think the value of our results can be viewed in the context of some previous data. Xia et al. previously reported a retrospective analysis involving 10,966 patients with HCC in China. Their results showed that the 1-year recurrence rates of CNLC stage 1b, 2a, 2b, 3a (BCLC stage B to C) HCC patients after surgery were 32.4%, 45.7%, 55.9%and 59.9%, respectively[1]. The reported population has similar etiologic (mainly HBV), racial, and geographic characteristics to the population included in our study. The 1-year DFS rate of 74.0% observed in this study is numerically at a good level, suggesting that the result of this study is a positive signal,

which can provide a reference for the design of subsequent similar studies and is worth to be verified by larger randomized controlled trials. We have also added relevant discussion in the revised manuscript in order to provide readers with a better understanding of our results. (Discussion section, lines281-291)

In conclusion, the objective of this study was to provide an early exploration of perioperative immunotherapy for HCC. Therefore, from the perspective of study design, we could not provide rigorous comparative conclusions. However, the value of the results of this study is proof of concept for this perioperative treatment strategy and provides data for the design of similar clinical studies. As we have introduced in this manuscript, perioperative treatment of hepatocellular carcinoma (HCC) is an important clinical problem of how to reduce postoperative recurrence and prolong survival, and it is a current research hotspot. This also implies the importance of our research data.

References:

1.Xia, Y.X., et al. Surgical treatment of primary liver cancer: a report of 10 966 cases. Zhonghua Wai Ke Za Zhi 59, 6-17 (2021).

Reviewer #2, Comments 2:

The authors should clarify why only 9 of 20 surgical patients received adjuvant tislelizumab-lenvatinib. For example, was this due to patient preference, performance status, toxicity, or other selection factors? Given the DFS difference between adjuvant and non-adjuvant groups, detailing these reasons and baselinecharacteristics would help assess potential confounding factors.

Response:

We greatly appreciate the reviewer's thoughtful question and helpful suggestions. There were 11 of 20 patients undergoing surgery did not receive adjuvant therapy in this study. The main reason for not receiving adjuvant therapy was the inability or unwillingness of patients due to COVID-19 to return to the site for treatment according to the protocol-defined algorithm. To clarify this issue, relevant description has been added in Result section, lines134-137. Besides, the baselinecharacteristics, tumor evaluation, and pathological evaluation of the adjuvant therapy group (n=9) and the non-adjuvant therapy group (n=11) are shown in Table R1, and no significant difference is observed between the two groups. During the preoperative period, the incidence of grade 1/2 TRAE in non-adjuvant and adjuvant group is 81.8% and 77.8%

and there is no grade 3 TRAE observed. These indicate that there was no substantial selective bias in the use of adjuvant therapy.

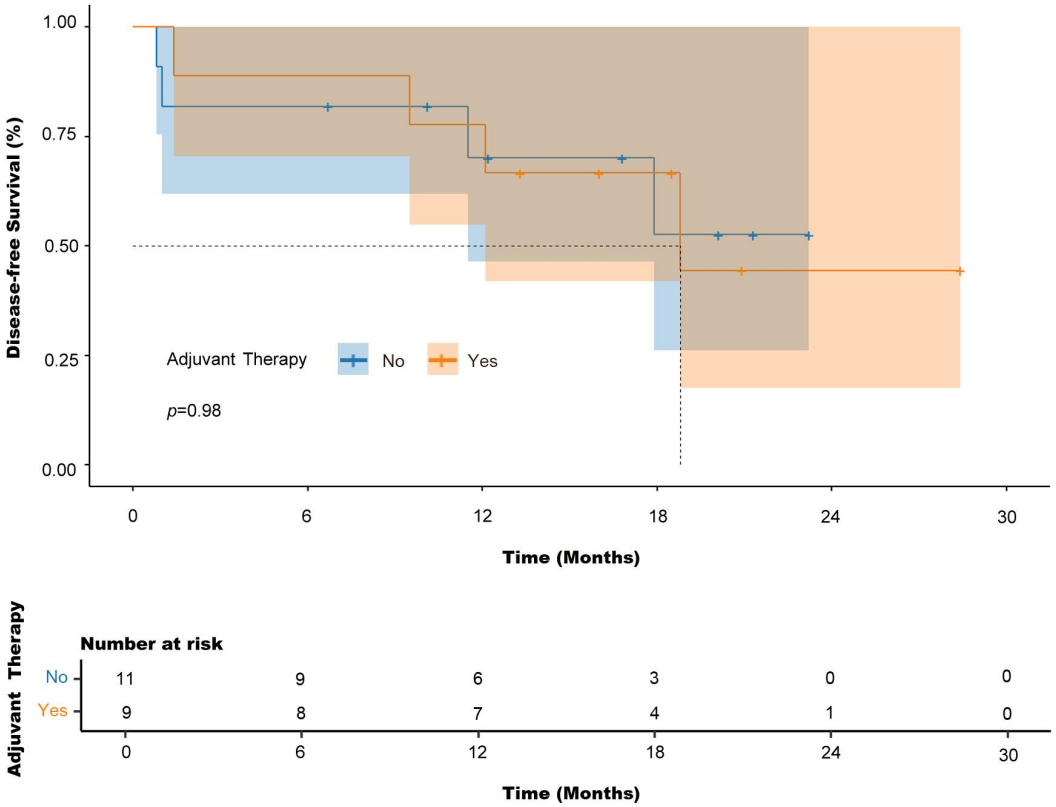
Table R1 Baselinecharacteristics of adjuvant and non-adjuvant group

	Non-adjuvant group (N=11)	Adjuvant group (N=9)	*P* value
Age			
Median (Min-Max)	59 (35-67)	57(44-69)	0.928
Maximum tumor size			
Median (Min-Max)	6.5 (3.7-12.3)	6.3 (3.3-8.9)	0.449
AFP			
Median (Min-Max)	1074 (10.90-65228)	42 (1.94-7833)	0.095
Gender			
Male	10 (90.91 %)	9 (100 %)	1.000
Female	1 (9.09 %)	0 (0 %)	
ECOG PS score			
0	11 (100 %)	9 (100 %)	1.000
Child Pugh class			
A	11 (100%)	9 (100%)	1.000
HBV			
No	1 (9.09 %)	1 (11.11 %)	1.000
Yes	10 (90.91 %)	8 (88.89 %)	
BCLC stage			
A	6 (54.55 %)	3 (33.33 %)	0.406
B	5 (45.45 %)	6 (66.67 %)	
CNLC stage			
Ib	6 (54.55 %)	3 (33.33 %)	0.406
IIa	5 (45.45 %)	6 (66.67 %)	
Tumor number			
1	6 (54.55 %)	3 (33.33 %)	0.271
2	4 (36.36 %)	6 (66.67 %)	
3	1 (9.09 %)	0 (0%)	
ALBI grade			
1	8 (72.73 %)	6 (66.67 %)	1.000
2	3 (27.27 %)	3 (33.33 %)	

Tumor evaluation per mRECIST			
CR	4 (36.36 %)	1 (11.11 %)	0.301
PR	5 (45.45 %)	7 (77.78 %)	
SD	2 (18.18 %)	1 (11.11 %)	
Pathological evaluation			
Non-response	3 (27.27 %)	4 (44.44 %)	0.577
MPR	5 (45.45 %)	4 (44.44 %)	
PCR	3 (27.27 %)	1 (11.11 %)	

We also attempted to explore whether it has influences on the DFS of the two groups. As shown in Extended Data Fig. 1 (Supplementary Fig.1) , there is no significant difference observed in DFS between patients who received adjuvant therapy or not (p=0.98). This is really an interesting phenomenon. As we described in the manuscript, the negative results of the IMbrave 050 study have led to a heated discussion about the value of immunotherapy as adjuvant therapy for HCC[1]. Similarly, we incidentally found that the use of adjuvant therapy after preoperative therapy and surgery did not appear to have a significant effect on DFS, implying that preoperative therapy may play a more important role than adjuvant therapy in the perioperative intervention of resectable HCC at high risk of recurrence. Limited by the study design, these results should be interpreted with caution. And the relationship between the use of adjuvant therapy and prognosis may need to be explored in a dedicated randomized, controlled trial. Nevertheless, it also indicates that there are still many problems to be solved in the exploration of perioperative treatment of HCC, and emphasizes the importance of disclosing more research experience and data for the reference of treatment optimization.

Finally, we have added part of the necessary discussion and stated this limitation in the revised manuscript in order to provide readers with a better understanding of our data and a certain reference for the optimization of future perioperative treatment regimens for HCC. (Discussion section, lines316-325)



Extended Data Fig. 1 (Supplementary Fig.1) Kaplan-Meier curve of disease-free survival for patients receiving/not receiving adjuvant therapy.

References:

1. Fu Y, et.al. Where is the future of adjuvant therapy for hepatocellular carcinoma? J Clin Oncol. 2025, 43(14):1625-1630.

Reviewer #2, Comments 3:

Given the high MPR rate in this study (60%, which is 2-3 times higher than that reported in other perioperative HCC trials), the authors should report the absolute number and percentage of MPR patients who experienced recurrence, and the pattern of recurrence. This would clarify whether the high pathological response rate translated into lower recurrence risk and aligns with standard perioperative trial reporting.

Response:

We sincerely appreciate the reviewer's valuable comments regarding recurrence outcomes. During postoperative follow-up, a total of eight patients developed recurrence. None of the four patients who achieved complete response (CR)

experienced recurrence. Among the eight patients with major pathological response (MPR), three recurred (two extrahepatic and one intrahepatic), whereas among the eight non-MPR patients, five recurred (four intrahepatic and one extrahepatic). Overall, CR patients showed no recurrence, MPR patients had a lower recurrence rate with a tendency toward extrahepatic relapse, and non-MPR patients had a higher recurrence rate, predominantly intrahepatic. Although these differences did not reach statistical significance, the descriptive findings suggest potential variations in recurrence patterns across response groups. Fig.R2 has been added to visually illustrate the distribution of recurrence events.

Similar observations have been reported in other studies. Some studies have found that patients receiving neoadjuvant immune checkpoint inhibitor therapy who achieved a major pathological response (MPR) were associated with longer relapse-free survival (RFS)[1].Even patients who achieve a major pathological response (MPR) or complete pathological response (pCR) after neoadjuvant immunotherapy may still experience recurrence. Studies in high-risk resectable melanoma have shown that, although these patients generally have favorable outcomes and respond to salvage therapy, late recurrences can still occur, often at distant sites rather than at the primary tumor bed[2].

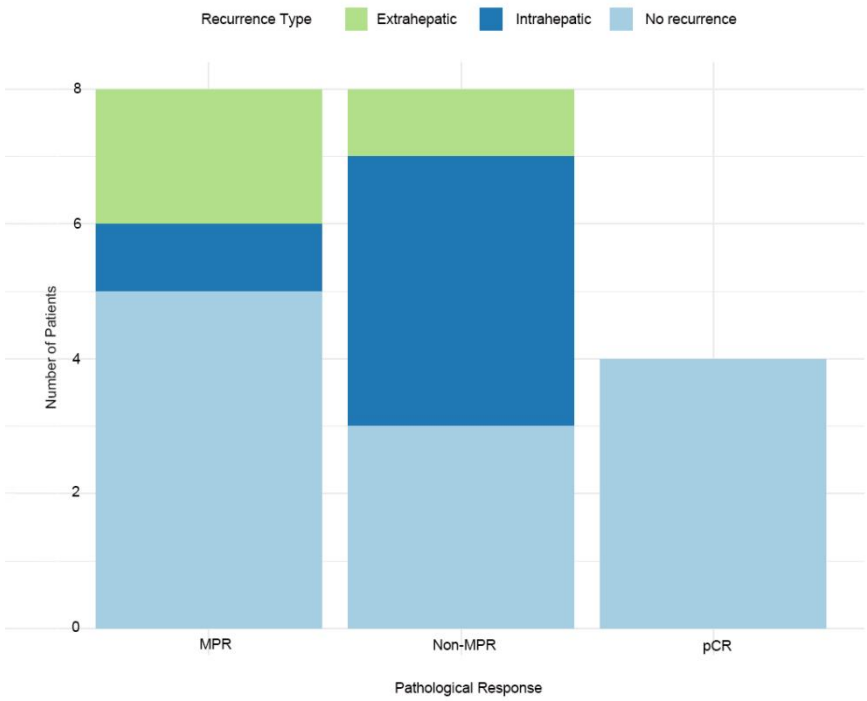


Fig.R2 Recurrence Incidence and Sites by Pathological Response.

References:

1.D'Alessio, A., Stefanini, B., Blanter, J. et al. Pathological response following neoadjuvant immune checkpoint inhibitors in patients with hepatocellular carcinoma: a cross-trial, patient-level analysis. Lancet Oncol. 25, 1465-1475 (2024).

2.Sharon CE, Tortorello GN, Ma KL, et al. Long-term outcomes of neoadjuvant pembrolizumab based on pathological response in patients with resectable stage III/IV cutaneous melanoma. Nat. Rev. Oncol. 34: 806-812 (2023).

Reviewer #2, Comments 4:

Regarding the feasibility of the study, the definition of a delay in surgery of 49 days or more appears to be rather extensive. The authors should provide a thorough basis for this definition. Moreover, 70.4% (n=19) of patients who underwent surgery within 49 days appears to be a low rate. A thorough discussion of this result is necessary.

Response:

Thank you very much for your valuable advice. As the perioperative therapy of HCC is still in the early period of exploration, there is no consensus on the critical surgical window. In reported studies about the perioperative treatment of HCC, the definition of feasibility ranges from 50 to 89 days after treatment 50 days[1]; 60 days[2]; 89 days[3]. It is important to consider the following aspects the the feasibility definition:

First, the potential impact of adverse events from treatment on the surgery must be considered. In the design of most perioperative studies, safety is often selected as the primary or key secondary endpoint, reflecting concerns about the potential for adverse events to affect the procedure. Safety endpoints are typically measured by the incidence and severity of adverse events. However, this approach may overlook the fact that the outcomes of adverse events can also affect the implementation of surgery. The emerging definition of feasibility serves as an effective supplement to the safety endpoint. It should be noted that while some serious adverse events may cause patients to forgo surgery, there are still patients who can quickly recover from adverse events and continue to receive radical treatment opportunities. For this latter group, it is possible to achieve long-term DFS and even OS benefits through active preoperative intervention while preserving the opportunity for radical treatment. Therefore, we still believe that perioperative therapy is feasible, focusing on these

situations. According to the outcomes of adverse events related to immunotherapy, the time from onset to recovery for most adverse events ranges from 2 to 5.9 weeks As shown in Fig.R3[4], Thus, we think that defining the surgical window as 6 weeks (42 days) after treatment, plus an additional 7 days for scheduling surgery, can preserve the opportunity for a subset of patients to receive radical treatment.

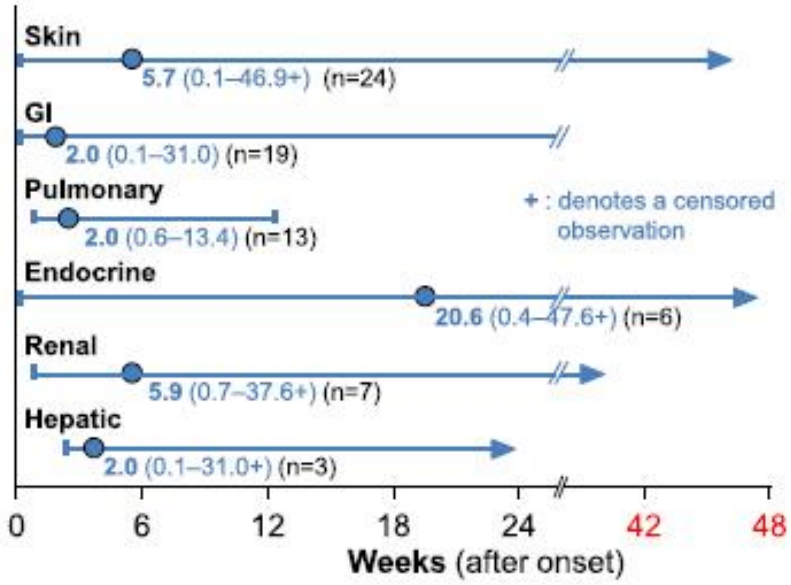


Fig.R3 Time patterns of resolution of irAE of anti-PD-1 antibody[4].

Secondly, when defining the surgical window period, attention should also be paid to its potential impact on prognosis. In other malignant tumors with more established concepts of neoadjuvant therapy, a study by Suleman et al. showed that breast cancer patients who underwent surgery within 4 to 7 weeks after neoadjuvant therapy had a trend of better pCR rates and DFS/OS compared to patients who underwent surgery after 8 weeks [5]. Similarly, related clinical studies in thoracic cancer patients have reported similar findings (6-8 weeks)[6].

Based on these considerations, we chose 49 days as the criterion for evaluating the feasibility of this treatment strategy at the beginning of our study design. This is also similar to the design of a study (defined as 50 days) of radiotherapy combined with immunotherapy in the neoadjuvant treatment of hepatocellular carcinoma reported by Li et al [1].

It is important to emphasize that we also support the completion of surgery as soon as possible after the completion of preoperative treatment. However, we believe that an appropriate surgical window under reasonable circumstances would be more conducive to a thoughtful evaluation of the feasibility of a perioperative treatment strategy.

Finally, we appreciate the opportunity to explain why 70.4% of the patients in this study underwent surgery 49 days after preoperative treatment. As mentioned in the original manuscript, of the 27 patients enrolled, 4 who achieved CR did not undergo surgery due to active refusal or other non-safety reasons. This indicates that preoperative treatment did not negatively affect the performance of surgery. It is also important to fully respect patients' wishes, as this is a key aspect of medical ethics. Among those who underwent surgery, only 1 out of 20 (5%) experienced surgical delays as defined in this study. And the median time from enrollment to surgery was 99.5 (range,56-156) days. It suggested the feasibility of this perioperative treatment strategy.

To facilitate the understanding of this result by subsequent potential readers, we have added relevant descriptions in the revised manuscript. (Result section, lines123-124;Discussion section, lines245-250)

References:

1.Li, Z., et al. Neoadjuvant tislelizumab plus stereotactic body radiotherapy and adjuvant tislelizumab in early-stage resectable hepatocellular carcinoma: the Notable-HCC phase 1b trial. Nat Commun 15, 3260 (2024).

2.Ho, W.J., et al. Neoadjuvant Cabozantinib and Nivolumab Converts Locally Advanced HCC into Resectable Disease with Enhanced Antitumor Immunity. Nat Cancer 2, 891-903 (2021).

3.Pinato, D.J., et al. PRIME-HCC: phase Ib study of neoadjuvant ipilimumab and nivolumab prior to liver resection for hepatocellular carcinoma. BMC Cancer 21, 301 (2021).

4.Eigentler, T.K., et al. Diagnosis, monitoring and management of immune-related adverse drug reactions of anti-PD-1 antibody therapy. Cancer Treat Rev 45, 7-18 (2016).

5.Suleman, K., et al. Does the Timing of Surgery after Neoadjuvant Therapy in Breast Cancer Patients Affect the Outcome? Oncology 98, 168-173 (2020).

6.Fligor, S.C., et al. Time to surgery in thoracic cancers and prioritization during COVID-19: a systematic review. J Thorac Dis 12, 6640-6654 (2020).

Reviewer #2, Comments 5:

The impact of lower dose administration on the therapeutic effect warrants further investigation. The findings, which indicated that patients who achieved CR

exhibited a more favorable prognosis compared to those who underwent surgical intervention, suggest that enhancing the therapeutic efficacy of preoperative treatment may lead to improvements in patient outcomes. If the dosage was reduced due to safety concerns, it would have been preferable to verify the appropriate dosage in advance. The authors should provide a rationale for their decision to forego preliminary testing.

Response:

We sincerely thank the reviewer for the valuable comment. In the neoadjuvant setting, balancing efficacy and safety is particularly important, as reducing adverse events can ensure that patients proceed to surgery without delay[1]. In the LEAP-002 study lenvatinib combined with a PD-1 inhibitor was associated with a high rate of grade ≥3 adverse events (63%)[2]. To improve tolerability, we administered lenvatinib at a fixed dose of 8 mg in this study. Notably, with this regimen, the incidence of grade ≥3 adverse events in our cohort was only 7.4%. Data from the ZSAB-TransGOLP study and several other HCC studies[3-5] indicate that this regimen is feasible and safe in similar clinical settings. This approach may help reduce the incidence of grade 3-4 adverse events, improve patient compliance, and thereby enhance the overall feasibility of neoadjuvant therapy.

References:

1.Llovet, J.M., et al. Adjuvant and neoadjuvant immunotherapies in hepatocellular carcinoma. Nat Rev Clin Oncol 21, 294-311 (2024).

2.Llovet, J.M., et al. Lenvatinib plus pembrolizumab versus lenvatinib plus placebo for advanced hepatocellular carcinoma (LEAP-002): a randomised, double-blind, phase 3 trial. Lancet Oncol 24, 1399-1410 (2023).

3.Shi, G., et al. Conversion therapy of tislelizumab plus lenvatinib and GEMOX in unresectable locally advanced biliary tract cancer (ZSAB-TransGOLP): a multicentre, prospective, phase 2 study. Lancet Oncol (2025).

4.Huang, C., et al. Organ specific responses to first-linelenvatinib plus anti-PD-1 antibodies in patients with unresectable hepatocellular carcinoma: a retrospective analysis. Biomark Res 9, 19 (2021).

5.Han, R.Y., et al. Lenvatinib, sintilimab combined interventional treatment vs bevacizumab, sintilimab combined interventional treatment for intermediate-advanced unresectable hepatocellular carcinoma. World J Gastroenterol 30, 4620-4635 (2024).

Reviewer #2, Comments 6:

The predictive model (based solely on hematologic factors) performed well in a single small cohort, but comparing it with standard clinical factors (e.g., tumor size, number, vascular invasion, stage) or integrating both would help demonstrate added value beyond established criteria.

Response:

This commenter raises an important point, and we appreciate the opportunity to explain it. In fact, these points have been carefully considered in the process of building our prediction model.

First of all, in machine learning modeling we generally do not want collinearity between input features. In the clinical logic and definition, there is an inherent high correlation between clinicopathological factors. Specifically, the BCLC and CNLC staging systems themselves are comprehensive packages of basic indicators such as tumor number, maximum size and tumor thrombus status, and collinearity between them is likely to exist, which has also been mentioned in some studies[1-2]. This is one of the reasons why we do not want to use these indicators for modeling analysis, which is also the case. As shown in Fig.R4 A-B, we found that there was a collinearity problem among the pathological indicators by calculating the variance inflation factor (VIF) and correlation coefficient[3].

In addition, we favored the inclusion of laboratory measures as continuous features rather than selecting pathological factors, which are mostly discrete features. Firstly, TabPFN is specially designed for small data sets, with automatic feature processing capability, and is able to directly accept continuous feature input[4]. Second, continuous features are able to retain more original information in small sample class imbalance problems. Wasikowski's classic work on IEEE TKDE directly proved that continuous data provide richer information than discrete features in small sample class imbalanced scenarios[5]. For example, continuous data from 1 to 5 may be defined as ≤ 5 categories in discrete features, but this actually ignores the data variation between 1 and 5. Continuous data retain this information. Finally, considering the particularity of small data sets, Zantvoort et al. 's study in Nature Digital Medicine confirmed that the information of each feature is extremely valuable in the case of limited sample size ($N \leq 300$)[6].

Therefore, we believe that it is more reasonable to use only continuous hematologic measures to keep the included features clean than to use

clinicopathological factors that may have information loss and serious collinearity problems. Of course, we really appreciate your questions. However, due to the above concerns and the good phenotype of our current model, we chose to try to re-model with traditional indicators to compare with our model instead of combining with traditional indicators. We exhaustive all combinations with more than 2 features. Unsurprisingly, our results are not good, even on the best combination test set (Figure.R4), which indicates that there is overfitting. I hope this answers your questions.

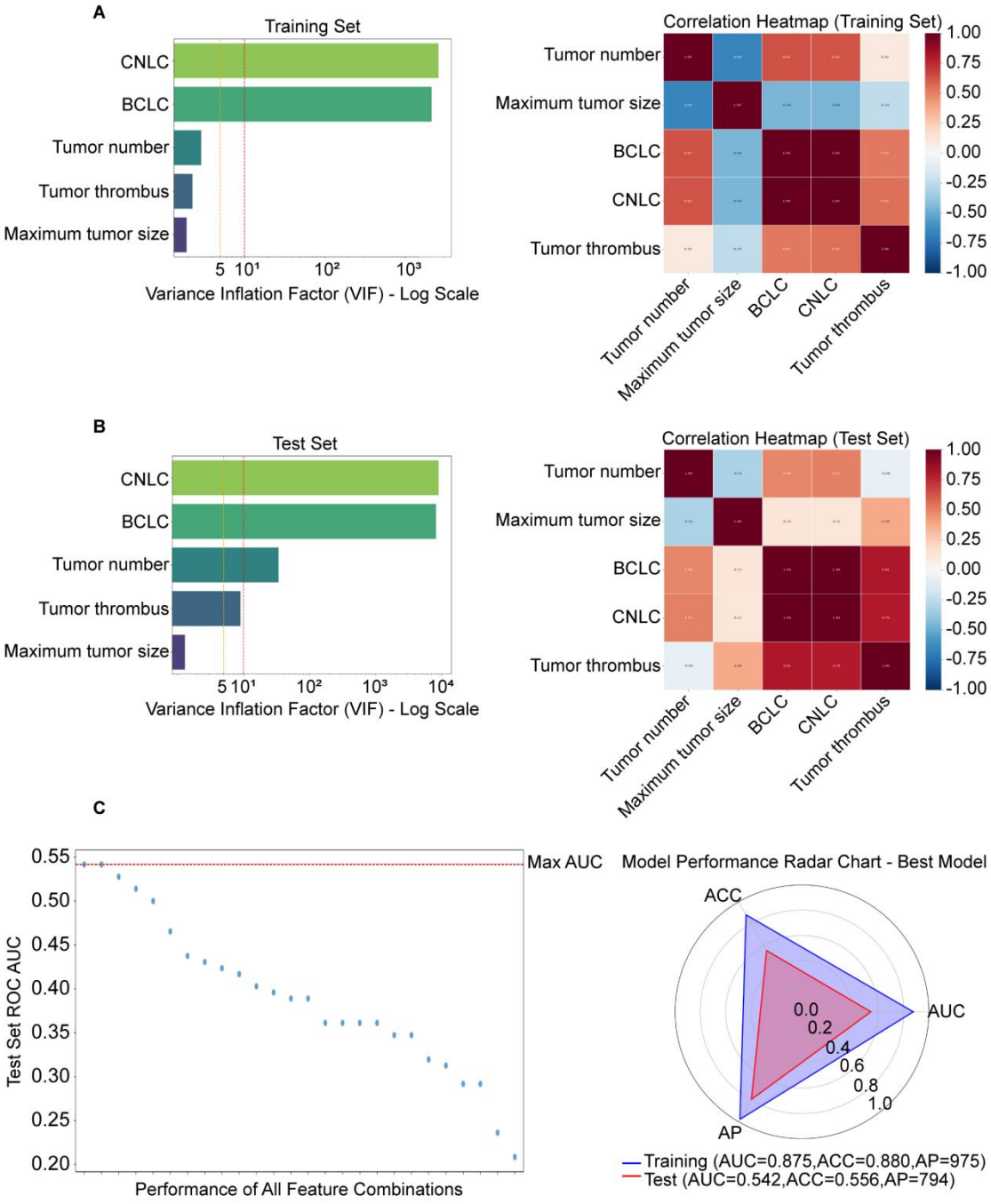


Fig.R4 Model performance training and testing results. A The bar and heat map

matrices demonstrate the VIF and correlation coefficients of the traditional pathological measures in the training set. **B** The bar and heat map matrices present the VIF and correlation coefficients of the traditional pathological measures in the test set. **C** Scatter plots and radar plots show our process of finding the best model for traditional pathological indicators and the performance of the best model in AUC, ACC and AP on the training and test sets.

References:

1.Liu, X., et al. Early presence of anti-angiogenesis-related adverse events as a potential biomarker of antitumor efficacy in metastatic gastric cancer patients treated with apatinib: a cohort study. J Hematol Oncol (2017).

2.Ebadi, M., et al. Visceral Adipose Tissue Radiodensity Is Linked to Prognosis in Hepatocellular Carcinoma Patients Treated with Selective Internal Radiation Therapy. Cancers (Basel) (2020).

3.Osredkar, J., et al. Association of Zn and Cu Levels in Cord Blood and Maternal Milk with Pregnancy Outcomes among the Slovenian Population. Nutrients 14, e4667 (2022).

4.Hollmann, N., et al. Accurate predictions on small data with a tabular foundation model. Nature (2025).

5.Wasikowski, M., et al. Combating the Small Sample Class Imbalance Problem Using Feature Selection. IEEE Transactions on Knowledge and Data Engineering (2010).

6.Zantvoort, K., et al. Estimation of minimal data sets sizes for machine learning predictions in digital mental health interventions. npj Digit. Med. (2024).

Reviewer #2, Comments 7:

As stated in the Limitations, it is imperative to ascertain the impact of markers of responsiveness derived from machine learning models on long-term prognosis. Concurrently, the impact on short-term outcomes should be examined. The finding that patients who achieved MPR exhibited a high rate of early recurrence is counterintuitive.

Response:

We thank the reviewers for emphasizing the importance of evaluating both long-term prognosis and short-term outcomes. Before formally responding to your comments, I would like to present a fact to you: Due to the possibility of delaying

curative surgeries, the implementation of neoadjuvant therapy in clinical practice is fraught with limitations and concerns, which to some extent has led to the issue of a small sample size for neoadjuvant therapy[1]. Clinicians all expect and seek methods to predict the efficacy of neoadjuvant therapy before treatment[2-3]. Therefore, the model constructed in this study and the indicators used are all for predicting the efficacy of neoadjuvant therapy. We do not deny the importance of long-term prognosis and short-term results, but at least at this stage, this is not within the scope of our research.

At the same time, we are very grateful for your raising this question. This indicates that there is an issue with the expression in our original text, which may cause readers to confuse the function of our model and lead to its misuse. Therefore, we have carefully revised the original expression, emphasizing the function of our model. In the Discussion section, lines229-232, the following statement has been added: The model's primary function lies in its ability to identify, before treatment, a subset of patients who are unlikely to respond, thereby preventing a delay in their opportunity for surgery. We hope this explanation can clarify the focus of the study and its clinical relevance.

We sincerely thank the reviewer for raising this point. In our cohort, MPR patients exhibited a lower overall recurrence rate compared with non-MPR patients (3 of 8 vs. 5 of 8). No pCR patients recurred. these results suggest that achieving MPR generally corresponds to a favorable recurrence outcome.

References:

1.Llovet, JM., et al. Adjuvant and neoadjuvant immunotherapies in hepatocellular carcinoma. Nat Rev Clin Oncol 21, 294-311 (2024).

2.Marron, TU., et al. Neoadjuvant cemiplimab for resectable hepatocellular carcinoma: a single-arm, open-label, phase 2 trial. Lancet Gastroenterol Hepatol 7, 219-229 (2022).

3.Xia, Y., et al. Efficacy and safety of camrelizumab plus apatinib during the perioperative period in resectable hepatocellular carcinoma: a single-arm, open label, phase II clinical trial. J Immunother Cancer 10 (2022).

Reviewer #2, Minor Comments 1:

Please verify the annotations in Figure 1B. The main text describes that PD patients did not undergo surgery. However, Figure 1B shows that one PD patient had surgery. Additionally, Figure 1B indicates that 22 out of 27 patients underwent surgery,

which is inconsistent with the 20 cases described in the Results section.

Response:

We sincerely thank the reviewer for carefully examining Figure 1B. The original version of Figure 1B depicted the actual surgical procedures performed among all 27 enrolled patients, including two patients who were not included in the per-protocol (PP) analysis. Specifically, for the patient with progressive disease (PD), subsequent treatment involved a change in systemic therapy combined with local interventions, which achieved better tumor control and allowed for later surgical resection. For the patient with complete response (CR), disease progression occurred after initial remission, and the patient subsequently underwent surgery. According to the trial protocol, these two patients were not included in the main efficacy analysis, which explains why the results section reports 20 per-protocol surgical cases. To prevent any potential confusion, we have revised Figure 1B and updated the figure legend to clearly distinguish between "per-protocol surgical cases" and "actual surgical procedures.

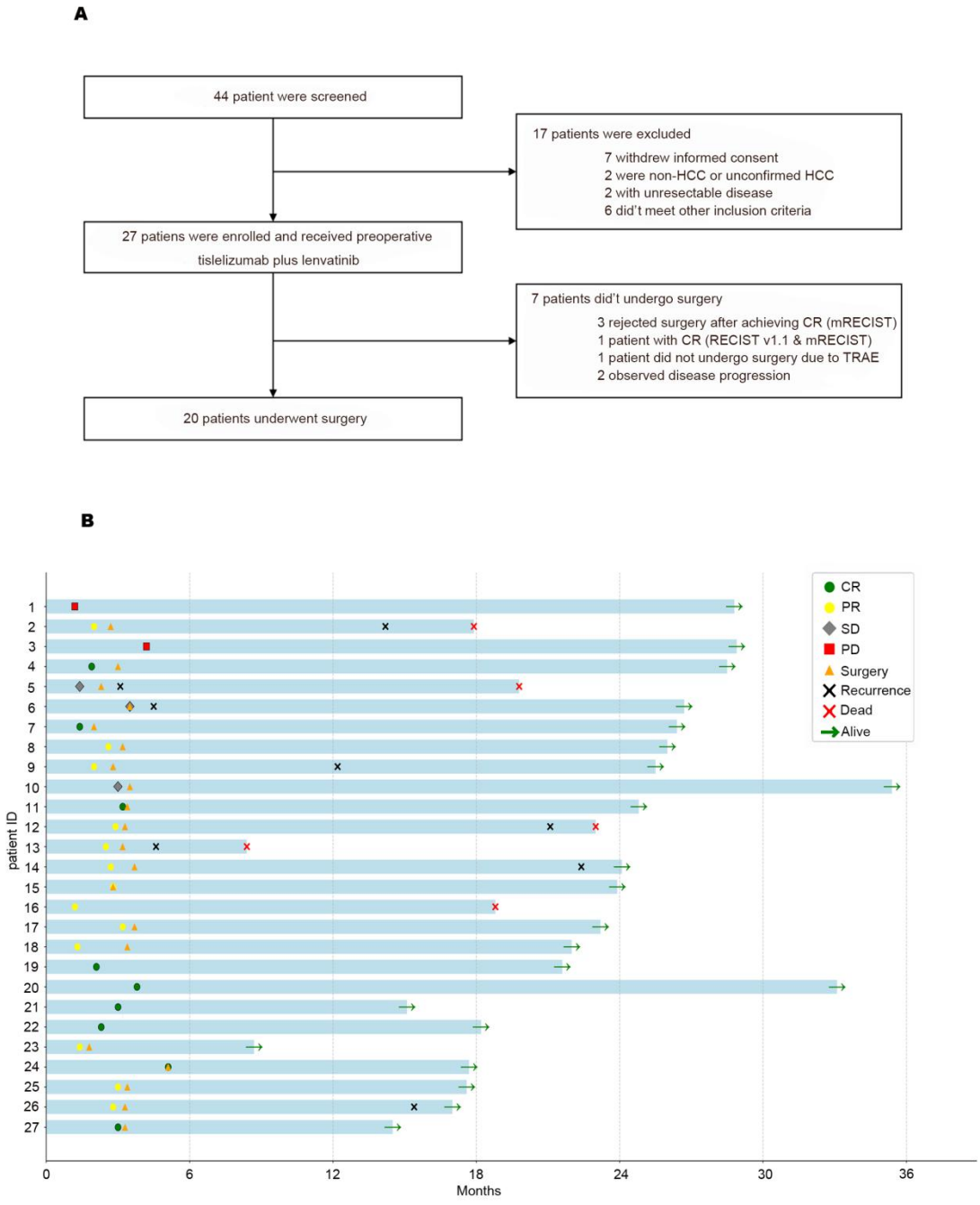


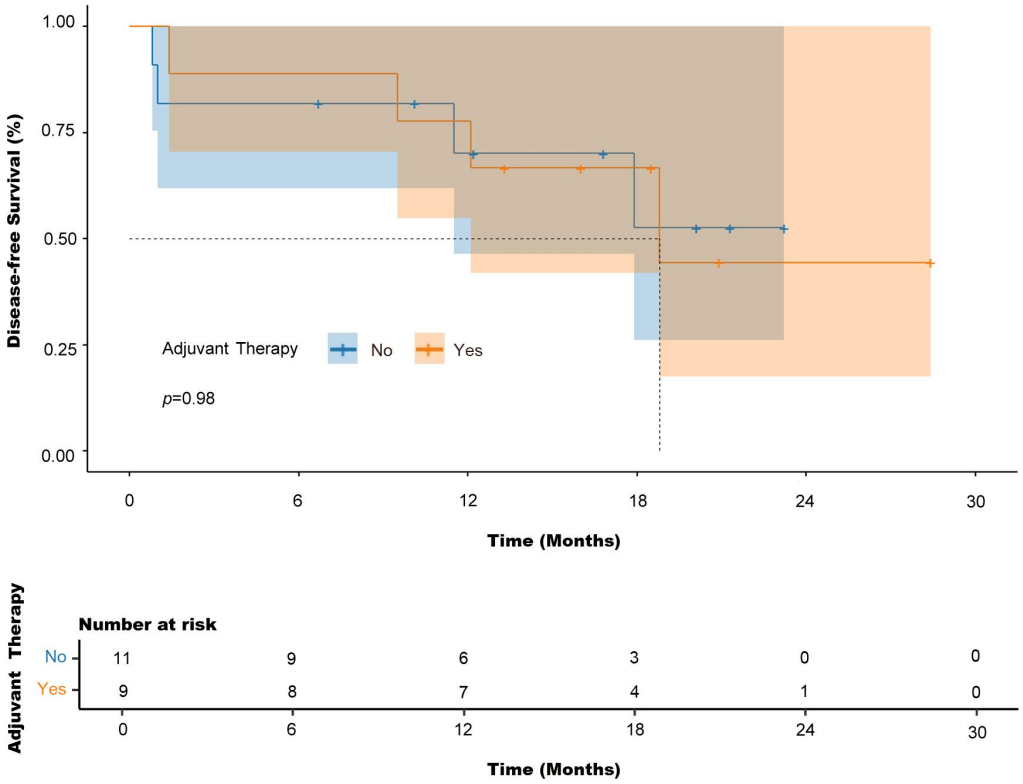
Fig. 1 Trial profiles. A study flow. **B** swimmer plot showing the treatment durations and assessments per mRECIST.

Reviewer #2, Minor Comments 2:

The results and notation in Extended Data Fig. 1 appear to be reversed. In the K-M curve, the group with "Adjuvant therapy: No" appears to have a DFS of 18.8 months and a 1YDFS rate of 77.8%.

Response:

We sincerely appreciate your valuable suggestion and apologize for any confusion caused. We have carefully reviewed the Extended Data Fig.1 (Supplementary Fig.1) and corrected the labeling accordingly.



Extended Data Fig. 1 (Supplementary Fig.1) Kaplan-Meier curve of disease-free survival for patients receiving/not receiving adjuvant therapy.

Reviewer #2, Minor comments 3:

The authors stated that "only one of them experienced tumor recurrence at the data cutoff point, which indicated that the necessity of surgery for this population needs further discussion, as previously reported." Yet no recurrence can be confirmed from Figure 1. The figure must be revised.

Response:

We thank the reviewer for raising this point. To clarify, the patient achieved complete response (CR) according to both RECIST v1.1 and mRECIST. Surgery was not initially performed because the tumor lesion could not be located per the study protocol. The patient subsequently experienced recurrence 12 months later, after which surgical resection was carried out. It should be noted that this recurrence and surgery were off-protocol events and therefore are not shown in Figure 1. We hope

this clarification provides a clear understanding of the patient's clinical course.

Reviewer #2, Minor comments 4:

In Fig.2D, the vertical axis should be labeled "Overall survival" rather than "DFS."

Response:

We thank the reviewer for pointing out this mistake. The label of the vertical axis in Figure. 2D has been corrected to "Overall survival" in the revised version.

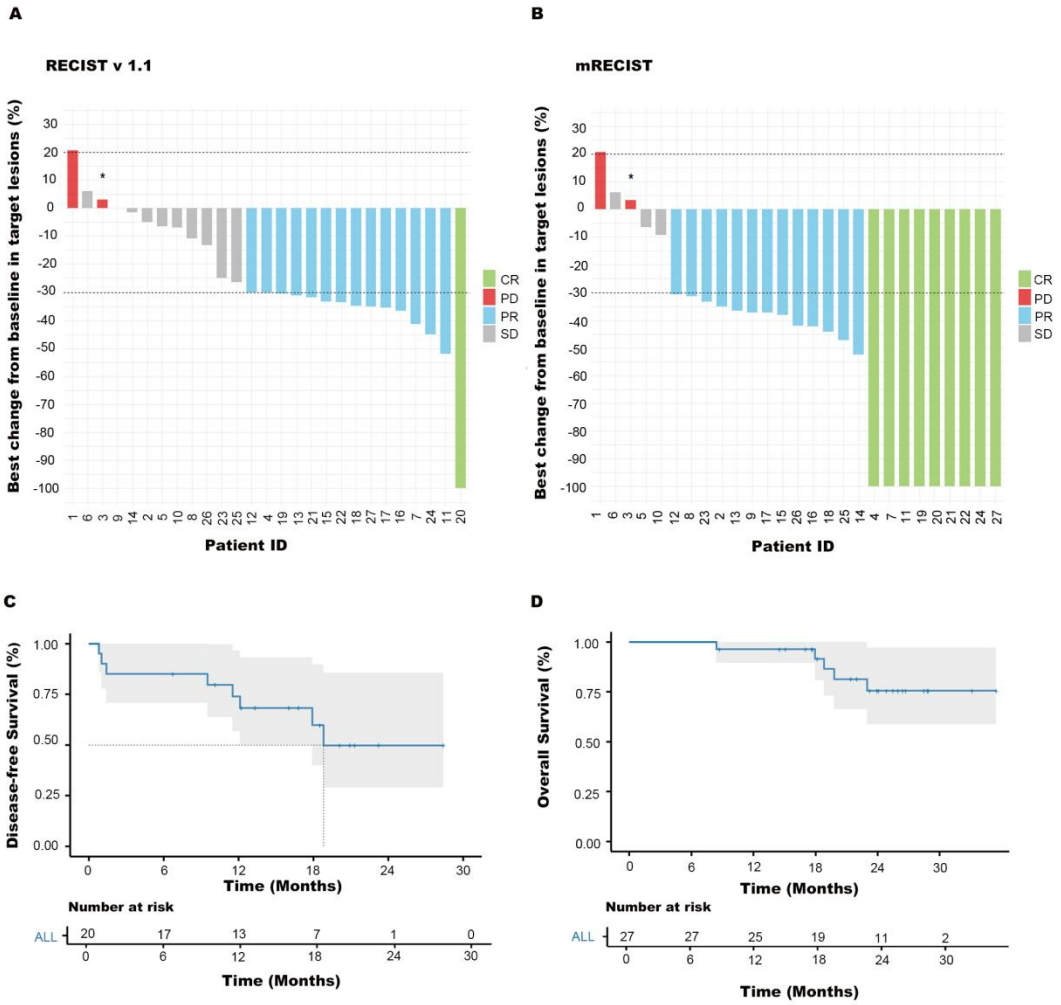


Fig. 2 Clinical outcomes of perioperative tislelizumab plus lenvatinib. Waterfall plot showing the change in the sum of target lesion diameters **A** per RECIST v1.1 and **B** per mRECIST. **C** Kaplan-Meier curve of DFS for patients undergoing surgery. **D** Kaplan-Meier curve of OS for all enrolled patients. *1 patient was assessed as PD due

to the appearance of a new lesion.

Reviewer #3, Comments 1:

The results demonstrate a mDFS of 18.8 months for 20 patients who received perioperative therapy and surgery. However there is no control arm for the study and thus comments on the effectiveness of this regimen (as stated at the end of the abstract -line38) cannot be fully made based on this analyses.

Response:

Thank you very much for your valuable advice. This study focused on the proof-of-concept of perioperative immunotherapy for HCC. Therefore, we initially chose a single-center, single-arm study design and did not formulate statistical hypotheses due to the lack of relevant references for efficacy expectations. As a result, we have avoided any comparative descriptions of efficacy and safety data as much as possible in the revised manuscript. The details are as follows: "In conclusion, tislelizumab-lenvatinib is a potential perioperative regimen for resectable HCC at high recurrence risk and exhibits preliminary efficacy and safety signals."(Abstract section lines40-42).

Additionally, we believe the value of our results can be contextualized with some previous data. Xia et al[1]. previously reported a retrospective analysis involving 10,966 patients with HCC in China. Their results showed that the 1-year recurrence rates for CNLC stage 1b, 2a, 2b, and 3a (BCLC stage B to C) HCC patients after surgery were 32.4%, 45.7%, 55.9%, and 59.9%, respectively. The reported population has similar etiologic (mainly HBV), racial, and geographic characteristics to the population enrolled in our study. The 1-year DFS rate of 74.0% observed in our study is numerically favorable, suggesting that our results are a positive signal. This can provide a reference for the design of subsequent similar studies and is worth verifying through larger randomized controlled trials. We have also added relevant discussion in the revised manuscript in order to provide readers with a better understanding of our results. (lines281-291)

References:

1.Xia, Y.X., et al. Surgical treatment of primary liver cancer: a report of 10 966 cases. Zhonghua Wai Ke Za Zhi 59, 6-17 (2021).

Reviewer #3, Comments 2:

There is heterogeneity in the perioperative regimens patients recieved, with only 9/20 receiving adjuvant therapy after resection. DFS rates were comparable between

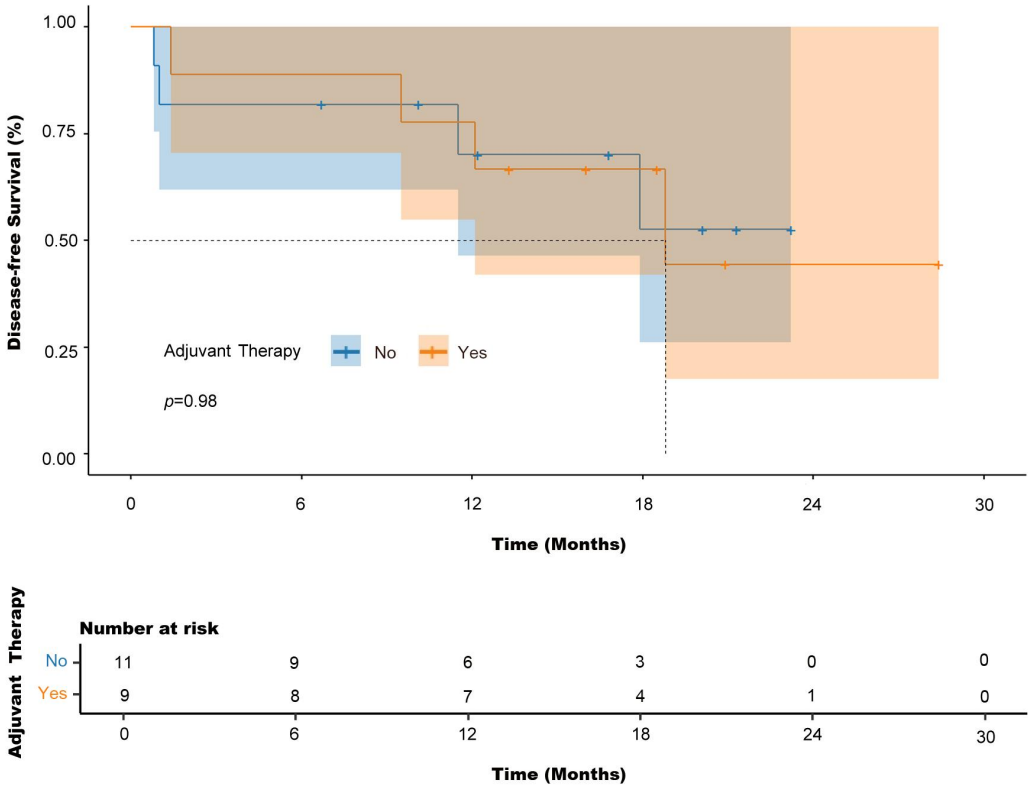
those who did and did not have adjuvant therapy. The implications on this finding for the optimal perioperative regimen needs to be addressed and the exact regimen proposed by the authors must be made clear.

Response:

We are grateful for your thoughtful suggestions. The reason for not receiving adjuvant therapy was the inability or unwillingness of patients due to COVID-19 to return to the site for treatment according to the protocol-defined algorithm. To clarify this issue, relevant description has been added in Result section, lines134-137. Interestingly, when we tried to explore whether this would have an impact on DFS between the two groups, we found no significant difference in DFS (p=0.98, as shown in Extended Data Figure 1 (Supplementary Fig.1)). As we described in the manuscript, the negative results of the IMbrave 050 study have led to a confusion about the effect of immunotherapy as adjuvant therapy for HCC. Our incidental finding followed a similar trend, implying that preoperative therapy may play a more important role than adjuvant therapy in the perioperative intervention of resectable HCC at high risk of recurrence. And it also suggests a direction to explore in the optimization of perioperative treatment.

Returning to the primary objective of the study, we observed the safety and feasibility of tislelizumab plus lenvatinib as a perioperative treatment, particularly in the preoperative period, along with preliminary signals of efficacy. Although patients were heterogeneous with respect to the use of adjuvant therapy, the results regarding DFS in the non-adjuvant and adjuvant groups should be interpreted with caution due to the small sample sizes and are not sufficient to dismiss the value of postoperative therapy in this regimen.

According to your advice, we have added part of the necessary discussion and stated this limitation in the revised manuscript in order to provide readers with a better understanding of our data and a certain reference for the optimization of future perioperative treatment regimens for HCC. (lines316-325)



Extended Data Fig. 1 (Supplementary Fig.1) Kaplan-Meier curve of disease-free survival for patients receiving/not receiving adjuvant therapy.

Reviewer #3, Comments 3:

For the prediction model, the sample size of both the training and testing cohorts are small. There is a class imbalance between responders and non-responders in both training (84% vs 16%) and testing (67% vs 33%) cohorts. This can impact the model performance in larger datasets. Therefore the results of the predictive model must be interpretated cautiously. Much larger cohorts are needed for robust testing of this model.

Response:

We understand the concerns expressed by the reviewers regarding the small sample sizes of the training group and the test group, as well as the imbalance in the distribution of categories (training group: 84% vs 16%; test group: 67% vs 33%). This situation is quite common in prospective neoadjuvant treatment studies for hepatocellular carcinoma. For instance, a phase II trial of simiplicatib neoadjuvant therapy by Marron et al., which only included 21 patients, also exhibited significant category imbalance (response group vs non-response group=15%: 85%)[1]. Similarly, a

randomized controlled trial of nivolumab monotherapy combined with ipilimumab by Kaseb et al., with a total sample size of only 27 cases, had 20 cases undergoing surgery (response group vs non-response group=33%: 67%)[2]. Moreover, the Notable-HCC trial by Li et al. only enrolled 20 cases (response group vs non-response group=63%: 37%)[3]. This is mainly due to the fact that the current neoadjuvant treatment for liver cancer is still in the initial exploration stage. Although we fully agree with the value of testing the robustness of the model in a larger cohort, due to the varying schemes used by researchers in current liver cancer neoadjuvant treatment studies and the inconsistent indicators included, we cannot achieve this within a short period of time. However, we highly appreciate your suggestion and we should adopt a cautious attitude when interpreting the model results.

Therefore, in the Discussion section, lines408-412, we further clarify this issue: Finally, the model in this study has the risk of favoring the majority category. Although we used an independent test set for internal validation, the results of this study are still at an initial stage and require external clinical validation in larger-scale, multi-center cohorts to confirm their broader clinical applicability We sincerely hope that this explanation will dispel the reviewers' doubts and clarify the limitations of this study.

References:

1.Marron, T. U., et al. Neoadjuvant cemiplimab for resectable hepatocellular carcinoma: a single-arm, open-label, phase 2 trial. Lancet Gastroenterol Hepatol (2022).

2.Kaseb, A. O., et al. Perioperative nivolumab monotherapy versus nivolumab plus ipilimumab in resectable hepatocellular carcinoma: a randomised, open-label, phase 2 trial. Lancet Gastroenterol Hepatol (2022).

3.Li, Z., et al. Neoadjuvant tislelizumab plus stereotactic body radiotherapy and adjuvant tislelizumab in early-stage resectable hepatocellular carcinoma: the Notable-HCC phase 1b trial. Nat Commun (2024).

Reviewer #3, Comments 4:

The authors suggest in the conclusion the TabPFN model identifies populations unlikely to respond to perioperative treatment. This appears to contradict the confusion matrix in Figure 3D, which suggests the model has a high sensitivity for detecting responders (100% in both cohorts) but low specificity (33.0%) in the testing

set. The model only detected 2 out of 6 non-responders and therefore this conclusion cannot be made based on the data presented. This low specificity means a large proportion of non-responders may receive perioperative therapy and risk disease progression prior to surgery.

Response:

We sincerely appreciate the important insight provided by the reviewers. It might be due to the unclear expression in the main text that led to your misunderstanding. First, what we need to add in the methods is that we are more concerned about the non-response group, so we regard the non-response group as positive. Secondly, we would like to clarify that the function of this model is to identify some patients who may not respond to perioperative treatment, rather than all. This is based on the model's high specificity (100%). As you mentioned, there are still some patients who may not respond to perioperative treatment being identified as the effective group. Therefore, to avoid the model being misused, we have made modifications and clarifications in the relevant results and discussion sections.

The clarification in the Result section, lines200-205 is: This indicates that although the model successfully identified some of the true non-responders, it did mistakenly classify a portion of the non-responders as responders. Therefore, we must explicitly state here that the function of this model is to identify a certain group of patients who are ineffective for perioperative treatment, but it cannot be used to predict patients who are effective for perioperative treatment. The statement in the Discussion section, lines229-232 is: The model's primary function lies in its ability to identify, before treatment, a subset of patients who are unlikely to respond, thereby preventing a delay in their opportunity for surgery. We hope this can clarify the confusion.

Reviewer #3, Minor comments 1 in Introduction:

Define perioperative intervention timing (i.e. pre- and post- resection)

Response:

We are deeply grateful to the reviewer for this insightful suggestion. We have indeed defined the timing of perioperative intervention (i.e., both pre- and post-resection) in the Methods section lines448-455 of the manuscript.

Reviewer #3, Minor comments 2 in Introduction:

line78 - based on BCLC guidelineslenvatinib is not first-linefor advanced HCC (only if ICI not feasible). Would suggest rephrasing to "is a recommended treatment for advanced HCC"

Response:

We appreciate the reviewer's comment. The sentence lines78-79 has been revised to: "Based on the positive results of the REFLECT study, lenvatinib is a recommended treatment for patients with advanced HCC."

Reviewer #3, Minor comments 3 in Introduction:

Introduction is very well written but quite long. I would suggest moving some of the text to the Discussion. For example:

line72-77: "In a phase II study, preoperative..... 1-year RFS of 53.85%"and line79-87: "Yi et al. reported .... antitumour activity in advanced HCC"and line99-107: "Currently, research on prognostic ... Cox regerssion is often challenging"

Response:

We sincerely thank the reviewer for the thoughtful suggestion regarding the length of the Introduction. We have carefully considered this recommendation and have moved certain parts of the text to the Discussion section to improve the manuscript's flow and readability.

The text from line72-77 ("In a phase II study, preoperative… 1-year RFS of 53.85%") has been moved from the Introduction to the Discussion, and is now located in lines236-240.

The text from line79-87 ("Yi et al. reported .... antitumour activity in advanced HCC") has been moved from the Introduction to the Discussion, and is now located in lines256-263.

The text from line99-107 ("Currently, research on prognostic ... Cox regerssion is often challenging") has been moved from the Introduction to the Discussion, and is now located in lines338-341 and 348-353.

Reviewer #3, Minor comments in Methods:

For the predictive model please state clarify the target variable. In line429 it is unclear if the Response and Non-Resposne is referring to pre-operative status only.

Response:

We are grateful to the reviewers for pointing out this ambiguity. To address this

issue, we have revised the lines484-491 in the "Methods" section to clearly and precisely define the target variable. We confirm that this target variable refers only to the preoperative response status evaluated after neoadjuvant therapy.

Reviewer #3, Minor comments 1 in Results:

On line142 the authors state one patient experienced TRAE and refused surgery but on line150 they state one patient experienced nerve injury leading to surgical cancellation. Could the authors clarify if this is the same patient and whether the cancellation of surgery was due to patient choice or a clinical decision?

Response:

We sincerely appreciate the reviewer's insightful comment. The patient who experienced a treatment-related adverse event (TRAE) and did not undergo surgery is the same patient who developed peripheral nerve injury leading to surgical cancellation. The cancellation of surgery was due to this clinical adverse event. We have clarified and corrected this in the Results section (lines132-133 and 143-145).

Reviewer #3, Minor comments 2 in Results :

For safety results please outlinehow many TRAEs led to treatment discontinuation.

Response:

We sincerely thank the reviewer for highlighting this important detail. In our study, two patients discontinued treatment due to treatment-related adverse events (TRAEs), one during the neoadjuvant phase and the other during the postoperative adjuvant phase. This information has been clearly added to the Results section (lines143-146).

Reviewer #3, Minor comments 3 in Results:

For the efficacy in line155 please state this is referring specifically to preoperative ORR and DCR.

Response:

We sincerely appreciate the reviewer's insightful suggestion. We have now clarified in the Results section (line149) that the efficacy outcomes reported here specifically refer to preoperative ORR and DCR.

Reviewer #3, Minor comments 4 in Results :

The results for DFS and survival are a little confusing to read, particurlarly the subgroup. Consider reporting (1st) the DFS and OS for the 20 patients who had perioperative therapy and surgery (2nd) DFS/OS for those who did and did not have adjuvant therapy (3rd) DFS/OS for MPR patients (4th) DFS/OS for surgical vs non-surgical.

Response:

We sincerely thank the reviewer for their valuable suggestion. Regarding the presentation of DFS and OS results, we have revised the Results section according to the reviewer's recommendation to facilitate clearer understanding for readers. The adjusted presentation order can be found in lines163-174 of the Results section.

Reviewer #3, Minor comments 5 in Results:

Consider using statistical comparision of characteristics between the training and validation set (Extended Data Table 3) and commenting on any noteable differences in the results text.

Response:

We are very grateful to the reviewers for their valuable comments. We have revised Extended Data Table 3 by adding *P* values and noted the test methods used for each indicator in the notes. We have also described the indicators with significant differences in the Results section lines178-180.

Reviewer #3, Minor comments 6 in Results:

Please outlinethe six features selected for the final model in the main text.

Response:

We are grateful for the valuable suggestion made by the reviewer. To enhance readability, in the lines187-190 of the Results section, we have clearly listed the six features selected for the final model.

Reviewer #3, Minor comments in 1 Discussion :

The authors outlinein line279 four patients did not undergo surgery due to achieving perioperative CR, with one of four experiencing tumour recurrence. This is a really important discussion point. However in the Results line137 the authors state two of the three patients who refused surgery due to CR had disease progression. This

needs clarification. Could the authors also state whether these patients had operable disease at the time of tumour recurrence/progression?

Response:

We thank the reviewer for pointing out this important inconsistency. We have carefully revised the text to clarify the descriptions. Specifically, in the Results section (lines125-133), the statement has been modified to avoid the misunderstanding that two of the three patients who refused surgery after achieving CR experienced disease progression. The revised description now reads:

"25.9% (n=7) did not undergo surgery and detailed reasons are as follows: three patients refused surgery after achieving complete response (CR) following treatment with tislelizumab plus lenvatinib; two patients experienced disease progression and the disease became unresectable after re-evaluation; one patient did not undergo surgery because CR was achieved according to both the Response Evaluation Criteria in Solid Tumors guideline, version 1.1 (RECIST v1.1) and modified Response Evaluation Criteria in Solid Tumors (mRECIST); one patient was unable to undergo surgery owing to treatment-related adverse event (TRAE)."

In addition, in the Discussion section (lines329-331), we revised the sentence regarding the four patients who did not receive radical treatment as follows:

"Four patients who did not receive radical treatment owing to some reason. Interestingly, only one of themexperienced subsequently developed disease recurrence and still underwent surgical resection."

We believe these revisions more accurately reflect the data and avoid potential misunderstanding.

Reviewer #3, Minor comments 2 in Discussion:

Please tone down the use of "excellent" for the TabPFN model performance in line297.

Response:

We are grateful to the reviewers for pointing out those overly optimistic statements. We agree that using the word "excellent" to describe the performance of the TabPFN model is subjective, and we have carefully revised the manuscript to adopt a more objective and neutral tone. The revised text in the Discussion section (lines351-353) of the manuscript is as follows: TabPFN is a novel machine-learning model designed for efficient training and capable of generating reliable predictions in

datasets with limited sample sizes.

Reviewer #3, Minor comments 3 in Discussion:

The high sensitivity of the TabPFN in both cohorts is mentioned in the discussion but not implicitly stated in the results or figures (though can be determined from Figure 3D confusion matrix). Please state the sensitivity (recall) in the results (100% in training and testing set). I would also suggest stating specificity, especially in the testing set as this is only 33%. This low specificity means a large proportion of non-responders may receive perioperative therapy and risk disease progression prior to surgery. The authors should comment on this.

Response:

We sincerely appreciate the important suggestion made by the reviewers, which is to enhance the clarity and transparency of our research results. To this end, in the "Results" section, we clearly indicated the sensitivity and specificity values of the training group and the test group. However, what we additionally stated in the methods is that we are more concerned about the situation of the non-response group, so we regarded the non-response group as positive. The "Results" section now includes the following detailed performance description: A detailed performance analysis based on the confusion matrix (Fig. 3D) revealed that the model achieved a sensitivity of 100% and a specificity of 100% in the training set. In the testing cohort, the model achieved a high specificity of 100%, correctly identifying all responders, but demonstrated a limited sensitivity of 33.3%. Similarly, in order to avoid the potential risks of the model being misused, we also further stated in the Result section (lines202-205) that the function of the model is to identify patients who are partially ineffective to perioperative treatment, and it cannot be used to predict the population that is effective to perioperative treatment.

Reviewer #4, Comments 1:

line170-172, could the authors please confirm if you mean to report the 1-year DFS or the median DFS as the 1-year DFS cannot be "not estimable".

Response:

Thank you for your kind reminder. There was indeed a clerical error regarding the median DFS. The corresponding content has been corrected in the revised manuscript.

Reviewer #4, Comments 2:

Could the authors please check the Extended Data Figure 3 and 2-year OS rates stated in the manuscript (line173-176). Which is the correct data as the 2-year OS rate for the no surgery cohort appears to be around or below the 75% mark on the y-axis.

Response:

Thank you for your kind reminder. We have corrected the OS data and the corresponding charts. It is important to note that the revised OS data do not affect our main conclusions. Additionally, we have rechecked all other data involved in this paper, and the corresponding corrections have been marked in the revised manuscript.

Reviewer #4, Comments 3:

The authors acknowledged that no hypothesis-based sample size computation was performed at the trial design stage.

Response:

Thank you for your kind reminder. We also recognize the critical role of hypothesis formulation. However, at the time this study was designed in 2021, there was a significant data deficit regarding the perioperative treatment with anti-PD-1 monoclonal antibody plus tyrosine kinase inhibitor in patients with resectable HCC. This introduced uncertainty regarding their efficacy and safety, making it difficult to perform hypothesis-based sample size computation for this study. Therefore, the current sample size was determined with consideration of the enrollment capacity of the participating site, as part of the initial exploration for safety and efficacy signals for this novel regimen. Additionally, we have avoided any comparative descriptions of efficacy and safety data as much as possible. Furthermore, we have also emphasized this limitation in the original manuscript.

Reviewer #4, Comments 4:

Could the authors please include the method used to compute the median follow-up time?

Response:

Thank you for your question. As it tended to appear in a similar context to DFS, the relevant data reported in the manuscript was the median follow-up time for DFS. The reverse Kaplan-Meier estimation method was employed for this calculation.

Reviewer #4, Comments 5:

It will be good if the authors can include the test used to compare the survival between groups (as reported in the manuscript, eg. line172) in the statistical section.

Response:

Thank you for your advice. The survival between sub-groups was compared using the log-rank test, and the relevant description has been added to the revised manuscript (Methods section, lines512-513). The details are as follows: "The log-rank test was used to assess survival differences between different sub-groups."

Reviewer #4, Comments 6:

line412 to 414, could the authors explain how censoring was done if the patient is alive and recurrence-free at time of data cut-off? How often is the follow up for disease recurrence?

Response:

Thank you for your question. In this study, DFS was defined as the time from surgery to the first diagnosis of recurrent tumor or death. For patients who were alive and recurrence-free at the time of data cut-off, DFS was computed as the time from surgery to the date of the last radiological examination. This censoring approach was relatively conservative to avoid overly optimistic estimates of the results.

For patients who underwent surgery, radiological examinations were performed every 3 months according to the protocol, which was consistent with the recommendations of relevant guidelines(e.g., NCCN). Additionally, unplanned radiological examinations were permitted if patients reported discomfort caused by suspected disease recurrence or if the investigator suspected disease recurrence based on other examination results.

Reviewer #1, Comments:

The author responses to my concerns do not address the underlying issues that I see regarding the lack of new clinical or scientific insights. This is another small non-randomized study which is largely affirming of prior similarly sized trials regarding the feasibility of neoadjuvant therapy in HCC. In particular, the previously conducted cabo nivo trial also already included high-risk HCC. The observed ORR is also extraordinary, but without clear explanation as to why drug effects in this study seem superior to other studies.

Response:

We sincerely appreciate the reviewer's valuable comments. First, we would like to note that neoadjuvant therapy for hepatocellular carcinoma (HCC) is not yet recommended by current clinical guidelines, and the field remains exploratory. In this context, we believe that well-designed exploratory studies still hold meaningful value and potential. Second, regarding studies of cabo nivo trial, we note that their enrolled patient populations differ from those in our study, and the incidence of treatment-related adverse events (TRAEs) was relatively high. This raises certain concerns in the neoadjuvant setting, particularly with respect to maintaining surgical feasibility and perioperative safety. Given the relatively high ORR observed in our study and to further enhance the reliability and transparency of the results, we have made corresponding revisions based on the reviewer's suggestions. In the Methods section (lines 474-479), we added a detailed description of the imaging assessment workflow. Specifically, three board-certified radiologists independently participated in tumor response evaluation based on RECIST v1.1 and mRECIST criteria. Two radiologists performed the initial evaluations independently, and in cases of disagreement, a third senior radiologist made the final assessment to ensure the reliability and consistency of ORR evaluation. In addition, we have incorporated recent comparable studies into the Discussion (lines 271-273) to provide context for our findings. Notably, D'Alessio et al. reported durable disease-free survival (DFS) outcomes1 , offering additional reference for perioperative management strategies and suggesting that the DFS and relatively high ORR observed in our study are not isolated findings. As mentioned in our previous response letter,We conducted an exploratory analysis of tertiary lymphoid structures (TLS) in

available tissue specimens to investigate potential biological factors underlying the high ORR observed in our study. Among the 23 evaluable specimens (20 postoperative and 3 preoperative), 13 (56.5%) showed abundant and well-developed TLSs. Notably, 11 of these 13 patients achieved an objective response (ORR, 84.6% by RECIST v1.1), and conversely, 11 of the 12 patients who achieved ORR in this cohort had TLSs present (91.7%). These observations suggest a potential association between the presence of TLSs in the tumor microenvironment and the relatively high ORR observed in our neoadjuvant cohort, providing complementary biological insight into the antitumor activity. We hope that these revisions and additions adequately address the reviewer's concerns and help readers better understand the relevance and potential value of our study in exploring neoadjuvant therapy for HCC. References: 1.D'Alessio, A. Pathological response following neoadjuvant immune checkpoint inhibitors in patients with hepatocellular carcinoma: a cross-trial, patient-level analysis. Lancet Oncol 25, 1465-1475

Reviewer #2, Comments:

The authors addressed all my comments. The only minor comment I still have is the use of the lower dose of lenvatinib in this study (8 mg/day). This may have contributed to lower toxicity, but also may have had different biological effects on the tumor, facilitating immune responses via effects on the vessels, as seem for lower doses of regorafenib, for example (see PMID: 33234602 and 38374347). This deserves some discussion.

Response:

Thank you for your suggestions. We have added an elaboration on this issue in the discussion section from line 248 to line 254. The specific content is as follows: Kim et al.Consistently, Kim et al. reported in the phase 2 RENOBATE trial that regorafenib, a TKI, combined with nivolumab as first-line therapy for unresectable HCC demonstrated clinical activity with manageable safety. Notably, a reduced dose of regorafenib (80 mg/day) was used, which may have contributed to lower toxicity while maintaining efficacy[1], and preclinical evidence suggested that this reduced dose of regorafenib could normalize tumor vasculature and enhance CD8+ T cell infiltration via CXCL10 expression[2].

References:

1. Kim, H.D., et al. Regorafenib plus nivolumab in unresectable hepatocellular carcinoma: the phase 2 RENOBATE trial. Nat Med 30, 699-707 (2024).
2. Shigeta, K., et al. Regorafenib combined with PD1 blockade increases CD8 T-cell infiltration by inducing CXCL10 expression in hepatocellular carcinoma. J Immunother Cancer 8(2020).

Reviewer #4, Comments:

Line 319-321 Given the small sample size of 9 vs 11 patients in the groups of patients with and without adjuvant therapy, the study lacks power to detect any significant difference. Hence, the absence of a significant p-value cannot be used to implied that there is no significant difference. Would suggest the authors to rephrase the sentence and remove the p-value.

Response:

Thank you for your suggestions. We have revised the description from line 325 to line 327. Interestingly, disease-free survival showed a similar trend between patients who received adjuvant therapy and those who did not.Nonetheless, owing to the limited sample size, we were unable to perform an independent validation.