

Dynamic prognostication and treatment planning for hepatocellular carcinoma: A machine learning-enhanced survival study using multi-centric data

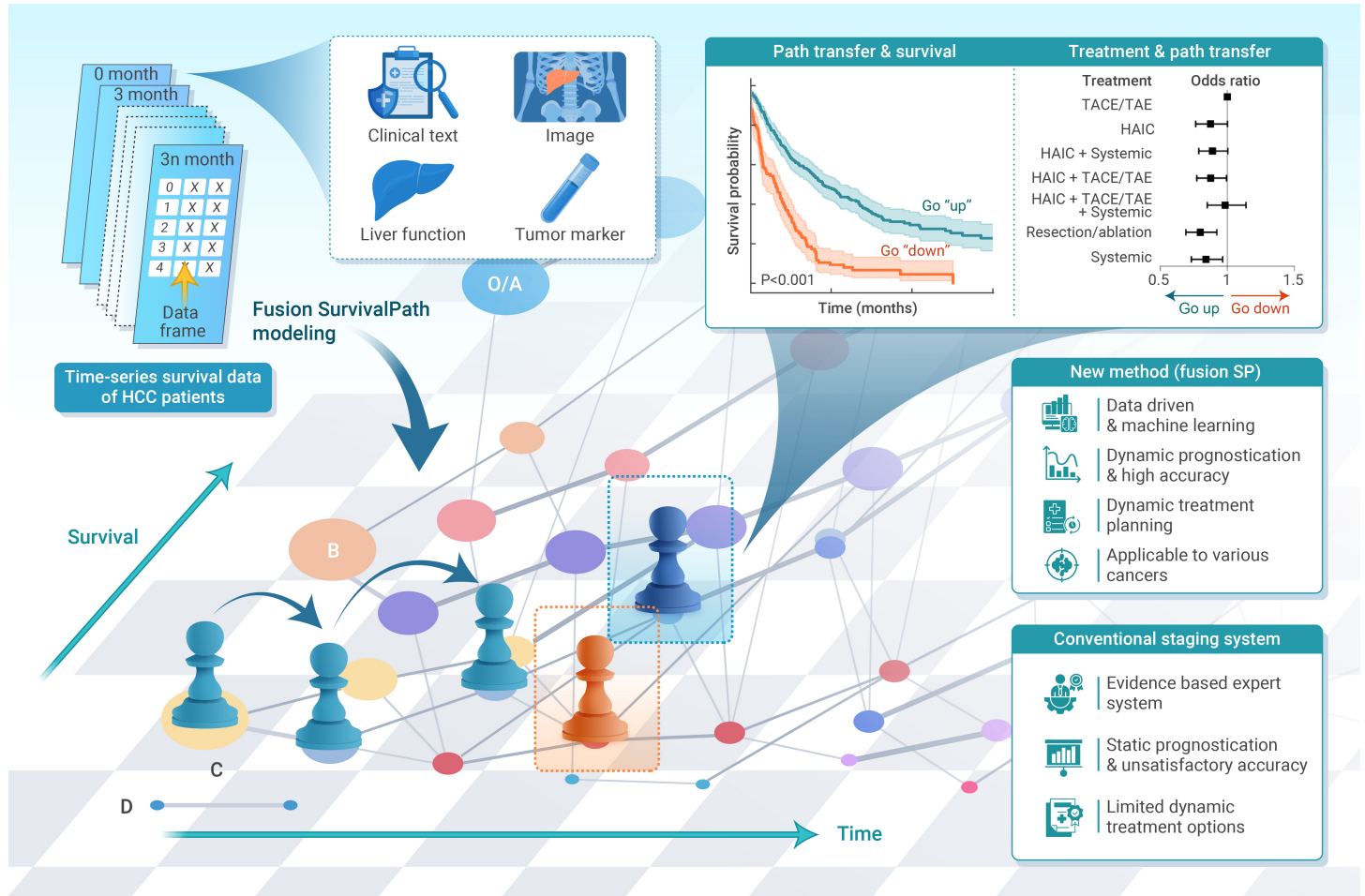
Lujun Shen,^{1,2,16} Yiquan Jiang,^{1,2,16} Linbin Lu,^{3,16} Mengxuan Cui,^{4,16} Jiming Xu,^{4,5,16} Chen Li,^{1,2} Ruizhi Tang,^{1,2} Qi Zeng,⁶ Kai Li,⁷ Juan Nie,⁸ Jingjun Huang,⁹ Boyang Chang,¹⁰ Nan Wu,¹¹ Feng Shi,¹² Ge Ren,¹³ Yao Wang,^{4,5} Zhimei Huang,^{1,2} Chao An,^{1,2} Zhongguo Zhou,¹⁴ Chaofeng Li,¹⁵ Xiong Chen,³ Letao Lin,^{1,2} Peihong Wu,^{1,2} Linfeng Li,^{4,*} Jinhua Huang,^{1,2,*} and Weijun Fan^{1,2,*}

*Correspondence: fanwj@sysucc.org.cn (W.F.); huangjh@sysucc.org.cn (J.H.); linfeng.li@yiducloud.cn (L.L.)

Received: December 3, 2024; Accepted: April 8, 2025; Published Online: April 9, 2025; <https://doi.org/10.59717/j.xinn-med.2025.100125>

© 2025 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Hepatocellular carcinoma (HCC) patients require advanced tools for dynamic prognostication and treatment planning.
- The fusion survival path model (fusion-SP) revolutionizes HCC prognosis tracking using longitudinal data over time.
- Fusion-SP boosts dynamic HCC monitoring and optimizes personalized care.
- The adaptable framework holds potential for broader application across multiple cancer types.

Dynamic prognostication and treatment planning for hepatocellular carcinoma: A machine learning-enhanced survival study using multi-centric data

Lujun Shen,^{1,2,16} Yiquan Jiang,^{1,2,16} Linbin Lu,^{3,16} Mengxuan Cui,^{4,16} Jiming Xu,^{4,5,16} Chen Li,^{1,2} Ruizhi Tang,^{1,2} Qi Zeng,⁶ Kai Li,⁷ Juan Nie,⁸ Jingjun Huang,⁹ Boyang Chang,¹⁰ Nan Wu,¹¹ Feng Shi,¹² Ge Ren,¹³ Yao Wang,^{4,5} Zhimei Huang,^{1,2} Chao An,^{1,2} Zhongguo Zhou,¹⁴ Chaofeng Li,¹⁵ Xiong Chen,³ Letao Lin,^{1,2} Peihong Wu,^{1,2} Linfeng Li,^{4,*} Jinhua Huang,^{1,2,*} and Weijun Fan^{1,2,*}

¹Department of Minimally Invasive therapy, Sun Yat-sen University Cancer Center, Guangzhou 510080, China

²State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou 510060, China

³Department of Oncology, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou 350025, China

⁴Yidu Cloud Technology Inc, Beijing 100101, China

⁵Department of Automation, Tsinghua University, Beijing 100084, China

⁶Cancer center, The Fifth Affiliated Hospital of Sun Yat-sen University, Zhuhai 519099, China

⁷Department of Ultrasound, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou 510620, China

⁸Department of Research & Education, Guangzhou Concord Cancer Center, Guangzhou 510054, China

⁹Department of Minimally Invasive Interventional Radiology, Second Affiliated Hospital of Guangzhou Medical University, Guangzhou 510260, China

¹⁰Department of Interventional Radiology, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou 510620, China

¹¹Department of Family Medicine, Memorial University of Newfoundland, St. John's NL A1B 3V6, Canada

¹²Department of Interventional Radiology, Guangdong Provincial People's Hospital, Guangzhou 510060, China

¹³Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR 999077, China

¹⁴Department of Liver Surgery, Sun Yat-sen University Cancer Center, Guangzhou 510080, China

¹⁵Information Center, Sun Yat-sen University Cancer Center, Guangzhou 510080, China

¹⁶These authors contributed equally

*Correspondence: fanwj@sysucc.org.cn (W.F.); huangjh@sysucc.org.cn (J.H.); linfeng.li@yiduccloud.cn (L.L.)

Received: December 3, 2024; Accepted: April 8, 2025; Published Online: April 9, 2025; <https://doi.org/10.59717/j.xinn-med.2025.100125>

© 2025 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Shen L., Jiang Y., Lu L., et al. (2025). Dynamic prognostication and treatment planning for hepatocellular carcinoma: A machine learning-enhanced survival study using multi-centric data. *The Innovation Medicine* 3:100125.

A reliable system for dynamic prognostication and management of hepatocellular carcinoma (HCC) is urgently needed but currently unavailable. In our previous work, we developed a machine learning algorithm termed "Survival Path" (raw-SP) to enhance prognostication with longitudinal survival data. However, the previous model was limited to intermediate stage HCC patients, and it faced the risk of overfitting due to path proliferation. In this study, we developed a novel framework incorporating nodal fusion techniques to mitigate the risk of overfitting, and expanded the model's applicability to all stages of HCC patients. A post-fusion survival map (fusion-SP) containing 14 different paths was built, which demonstrated superior or non-inferior accuracy in dynamic prognosis prediction for HCC patients compared with raw-SP, as well as traditional staging systems within the first 15 months since initial diagnosis in large-scale derivation, internal and external validation cohorts. Subgroup analysis showed the fusion-SP demonstrated superior performance in dynamic prognostication compared to other models among patients with BCLC stage C disease and initial tumor burden above up-to-seven criteria. Under the framework of fusion-SP, novel and distinct optimal combination treatment strategies for advanced-stage HCC patients at different key nodes were uncovered, where traditional staging frameworks fall short. The fusion-SP framework could serve as a robust tool for facilitating dynamic prognosis prediction and treatment planning for HCC. Moreover, our streamlined methodology holds the potential to be applied across various types of cancers.

INTRODUCTION

Hepatocellular carcinoma (HCC) ranks as the sixth most prevalent cancer and the third leading cause of cancer-related mortality globally. It is characterized by multicentric origins, high risk of intrahepatic recurrence, and poor prognosis.¹ The Barcelona Clinic Liver Cancer (BCLC) staging and treatment protocol is currently the most widely adopted system worldwide.²

Traditional staging methodologies, including the BCLC staging system, offer a static snapshot of HCC status at a single time point.³ However, HCC exhibits marked interpatient heterogeneity in progression patterns, which substantially limits the efficacy of static classification systems for dynamic

disease monitoring and prognosis prediction.^{4,5} The 2022 update of BCLC strategy introduced the concepts of treatment stage migration (TSM) and untreatable progression to facilitate dynamic treatment adjustments in cases where recommended therapies prove suboptimal.² In 2023, Alessandro Vitale et al. proposed an evidence-based framework for HCC based on the novel concept of "multiparametric therapeutic hierarchy", enabling dynamic adaptation of staging-based algorithms.⁶ Despite this, there remains a gap in dynamic prognostication and treatment planning.

Machine learning presents a compelling solution by analyzing longitudinal clinical data, encompassing laboratory tests, imaging results, and treatment histories.⁷⁻¹⁰ This enables machine learning models to capture the dynamic nature of HCC progression and provide more accurate and personalized prognostications compared to traditional staging systems. In our previous work,⁵ we introduced a machine learning-based model called "survival paths" (SP), to dynamically predict the prognosis of BCLC stage B HCC patients based on manually-curated datasets.¹¹ While this model demonstrated superior discriminative capability over BCLC and AJCC staging systems, our prior model encountered limitations. Derived from a relatively small and incomplete dataset (around 1600/400 patients in internal/external cohort respectively, all at BCLC stage B at baseline), its generalizability was limited. Moreover, the algorithm's exponential path proliferation led to overfitting risks, data exhaustion, and insufficient data utilization for distant time period due to the early stop of node bifurcation.¹²

To address these limitations, this study implemented several key enhancements.¹³ We expanded the study scope to encompass patients across all BCLC stages at baseline. We also refined the SP algorithm through advanced state fusion techniques. These enhancements yielded significant findings across multiple levels. Firstly, a brand-new survival path model consisting of 14 distinct paths using fusion techniques (fusion-SP model) was built, which demonstrated improved and externally validated dynamic prognostic accuracy compared to conventional SP model. Secondly, a series of key features were identified that contributed to the bifurcation of paths, which provides important hints in the causes of disease progression. Thirdly, the increased sample size at the nodal level after fusion facilitated uncovering novel treatment combinations that might bring survival benefits to patients.

These advancements hold promise for developing a more robust and clinical

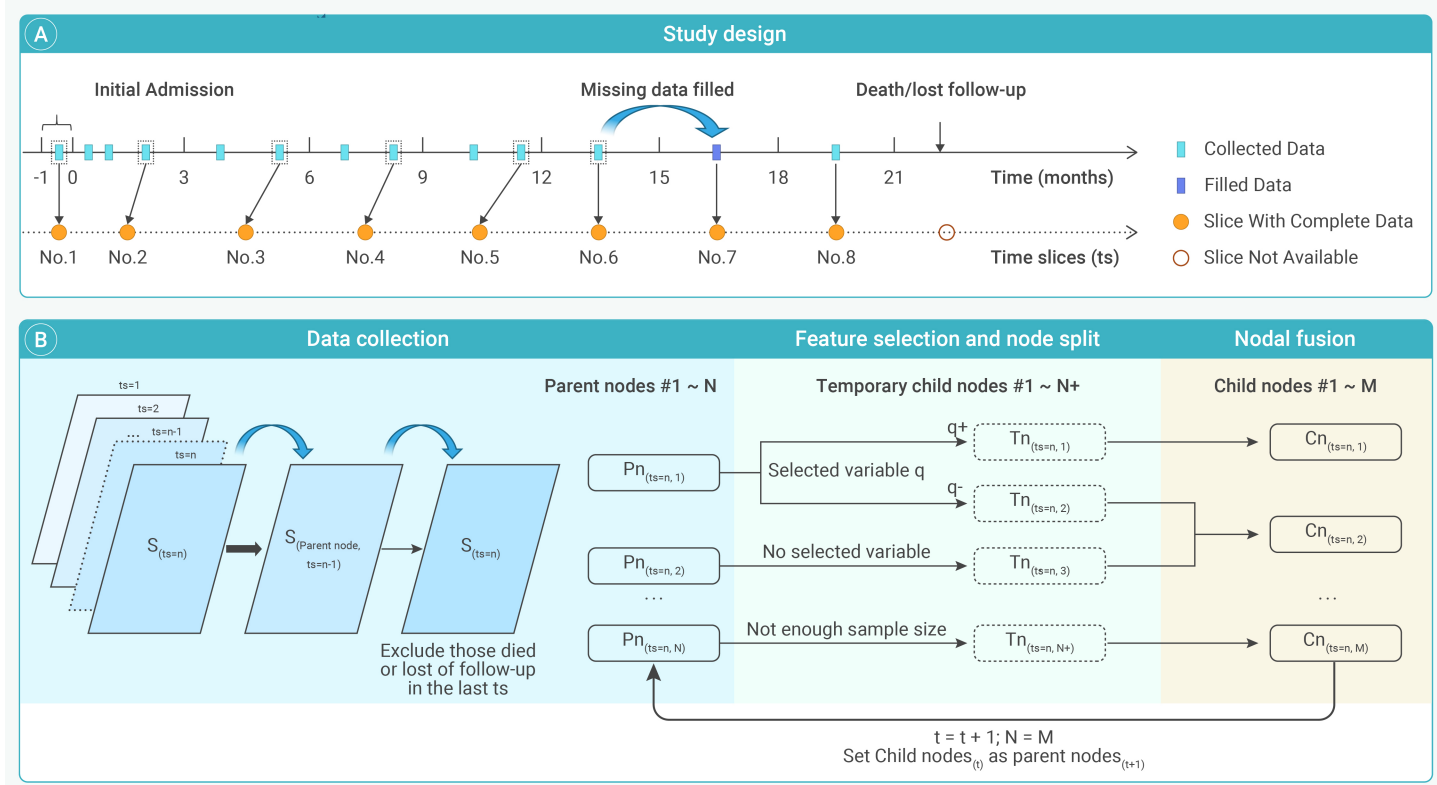


Figure 1. Flowchart of study design (A) The time-series data of HCC patients were converted into data of time slices with constant time interval, then undergo workflow of survival path modeling. (B) The whole population were initially divided into four subgroups based on BCLC classification and each subgroup in the first time slice sequentially undergo the procedure of *Feature Selection (FS)*, *Node Splitting (NS)* and *Nodal Fusion (NF)*. The computed child nodes serve as parent nodes in the next time slice and the data that represent patients of each node in the next time slice undergo next *FS*, *NS* and *NF* process; the analytical cycles continue until the completion of analysis of data of the last time slice.

cally applicable tool for dynamic prognosis prediction and personalized treatment planning in HCC patients.

MATERIALS AND METHODS

Study design and participants

A retrospective review was conducted on 33,285 patients diagnosed with HCC at Sun Yat-sen University Cancer Center (SYSUCC) between January 2007 and December 2022. Inclusion criteria were as follows: (1) patients who were newly diagnosed with HCC during their first visit; (2) with complete initial diagnostic data including abdominal computed tomography (CT) or magnetic resonance imaging (MRI), chest radiography or CT, routine bloodwork (biochemistry tests, serum AFP level, and coagulation indices); (3) without a history of other malignancies. A total of 19,485 HCC patients were included as the internal dataset, which were randomly split into derivation set ($N=15,588$, 80%) and internal validation set ($N=3,897$, 20%). For multicenter validation, a total of 1,337 HCC patients from five medical centers in southern China were reviewed between January 2010 and June 2023 (421 patients from The Fifth Affiliated Hospital of Sun Yat-sen University; 492 patients from The Third Affiliated Hospital of Sun Yat-sen University; 268 patients from Mengchao Hepatobiliary Hospital of Fujian Medical University, 116 patients from The Second Affiliated Hospital of Guangdong Medical University and 40 patients from Guangdong Provincial People's Hospital), with 1,282 meeting the inclusion criteria. The study is registered on Chinese Clinical Trial Registry (ChiCTR2100041874).

Transformation of time-series data for analysis

The longitudinal data of patients on demographics, imaging exam, serum tumor markers, and biochemical and coagulation test results, were collected and organized into standard time intervals. For each patient, the starting point was defined as the time of their initial diagnosis of HCC. A 3-month interval was applied, thus transforming the time periods of [-1, 1], (1, 3], (3, 6], (6, 9], (9, 12], (12, 15], (15, 18], (18, 21], (21, 24], (24, 27], (27, 30], (30, 33] months into 12 consecutive time slices, indexed from time slice no.1 to 12 (Figure 1A).

The completeness of the original dataset at baseline was reported in Table S1. Within each time slice, the latest available data were selected for variables with multiple entries. For surviving patients, missing values were imputed using the preceding time slice data (Figure S1). If a patient died or was lost to follow-up, subsequent time slice data were considered non-existent.

Based on past literatures on HCC,^{14,15} a total of eleven variables were designed, covering imaging results and laboratory tests; all the variables were dichotomized (Table S2). The primary outcome was overall survival (OS). For patients in the first time slice, length of OS was defined as the duration from diagnosis to death due to any causes. In subsequent time slices, lengths of OS were defined as the time from the image examination of that particular time slice to death due to any causes.

The work flow of survival path mapping

The SP was constructed based on derivation data following a specific workflow (Figure 1B). Patients were initially grouped into four subgroups based on their BCLC stages at diagnosis (0/A, B, C, or D), considered as parent-nodes at time slice no.1. For conventional SP model, a feature selection algorithm was utilized to identify the most discriminative feature at parent node and bifurcated each parent-node into two temporary child-nodes, which served as parent-nodes for the bifurcation at next time slice. In terms of the survival path with fusion techniques (fusion-SP), an additional merging algorithm combined similar temporary child-nodes from different parent-nodes was applied following nodal split of parent nodes at each time slice; the merged nodes served as parent-nodes for the next time slice. The building of SP models starts from the bifurcation of parent nodes at time slice no.1 and ends with the completion of the machine learning process described above in the last time slice.

Feature selection algorithm and nodal splitting. Step 1: All collected variables prepared for feature selection were discretized into binary values, either 0 or 1 (Table S2). In each parent node, "non-outlier" features were selected, whose properties were exhibited by more than 5% and less than 95% of the

patients.

Step 2: We performed univariate analysis with Kaplan-Meier method to identify features with statistical significance ($P < 0.05$) in predicting OS, which were further included in a multivariate random survival forest model¹⁶⁻¹⁸ to identify the most discriminative feature at each parent node. Random survival forest model was built by "Log-rank splitting" rules as follows,

$$L(X) = \frac{\sum_{j=1}^m (d_{j,L} - Y_{j,L} \frac{d_j}{Y_j})}{\sqrt{\sum_{j=1}^m \frac{Y_{j,L}}{Y_j} \left(1 - \frac{Y_{j,L}}{Y_j}\right) \left(\frac{Y_j - d_j}{Y_j - 1}\right) d_j}}$$

Where $d_{j,L}$, $d_{j,R}$ and $Y_{j,L}$, $Y_{j,R}$ refers to the number of deaths and individuals at risk at time t_j in daughter nodes left (L) and right (R). And Y_j and d_j were defined as follows,

$$Y_j = Y_{j,L} + Y_{j,R}$$

$$d_j = d_{j,L} + d_{j,R}$$

The value $|L(X)|$ is a measure of node separation. The larger the value, the greater the survival difference between daughter nodes left (L) and right (R), and the better the split is. Importance was determined by the Breiman-Cutler variable importance.¹⁹

Parent-nodes with sample size less than 100, or lacking any significant feature identified, were not subjected to further division. The resulting nodes, whether bifurcated or undivided, were regarded as temporary child-nodes for that specific time slice.

Nodal fusion algorithm. To consolidate temporary child-nodes with comparable BCLC stages and OS profiles into a single entity, a merging algorithm was developed as follows:

Step 1: Clustering of child nodes

We initiated by conducting a clustering analysis based on the distribution of BCLC stages (OA, B, and C) among all temporary child-nodes. Clustering analysis were performed by Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm²⁰ in which the loss optimizing criterion was as follows,

$$C \in C, d_{db}(p, q) \leq \min \varepsilon \forall p, q \in C \cap C_i \in C^C$$

Where $d_{db}(p, q)$ gives the smallest distance of ε so that the two points p and q in the same cluster are density-connected. The minimum number of points required within the epsilon neighborhood for core points was set at 1, which was empirical as at least 1 "neighbor" is required to constitute a cluster in SP models. The epsilon neighborhood size was set at 50 based on the identification of the sudden increase of the kNN distance (knee) that can bestly distinguish clusters from outliers in k-nearest neighbor (k-NN) distplot (Figure S2).

Step 2: Screening of the clustered nodes

Within each identified cluster, Cox proportional hazard models were constructed utilizing the clustered nodes as categorical features as follows,

$$h(t, X) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m)$$

where X_i represents each cluster as a categorical variable and β_i donates its corresponding coefficient.

Employing a backward elimination approach, we systematically removed nodes with the highest median survival until the P -value of the survival model exceeded 0.05.

Step 3: Finalization of fused nodes

The nodes remaining in the survival models were amalgamated into a final child node, representing a combined node with similar survival patterns. Meanwhile, the uncombined nodes remained as independent child nodes. Both the combined and uncombined child-nodes were designated as parent-nodes for the subsequent time slice.

Comparative analysis of prognostic models

A comparative analysis was conducted to assess the effectiveness of the fusion-SP model, raw-SP model, the BCLC staging system and the AJCC staging system in dynamic prognosis prediction. To evaluate the perfor-

mance of model for patients at each specific time slice, the time slice-specific C-index was calculated based on patients with valid data at each time slice. To assess the cumulative performance of model over time, we calculated the patient-specific C-index based on the newest update of each patient's node by the time of a specific time slice, with all patients considered. Moreover, we further stratified the cohorts into subgroups by BCLC stage, tumor burden, and liver function to identify subgroups where the fusion-SP model demonstrated superior prognostic accuracy in HCC.

Treatment strategies based on survival path transfer in fusion-SP model

Within the fusion-SP framework, as adopted in our previous study,⁵ the core of the treatment recommendation was to identify the optimal treatment that led to patients' transferring to "upper" nodes with improved survival at the subsequent time slice. In this study, four key nodes at paths VI and VII were selected as demonstration. The odds ratios of the different treatment options comparing to embolization therapy (TACE/TAE) were calculated using multivariate logistic regression, adjusting for tumor size, number of intrahepatic tumors, vascular invasion, distant metastasis, and AFP levels.

Statistical analysis

Baseline characteristics of categorical variables were reported as counts (percentages). Chi-squared test was used to compare categorical variables between groups. The multivariate random survival forest model was constructed using the "rfsrc" function from the "randomForestSRC" package.¹⁶⁻¹⁸ Clustering analysis was performed using "dbscan" function from the "dbscan" package.²⁰ The Cox proportional hazards regression model was built using the "coxph" function from the "survival" package.²¹ The DeepHit survival neural network was constructed using the "deephit" function from the "survivalmodels" package.²² The C-index was calculated using the "Cindex" function from the "survival" package,²³ and compared using the "compareC" function from the "compareC" package.²⁴ Two-sided P values less than 0.05 were considered significant. All analyses were performed using R software (version 4.1.2, R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Baseline characteristics of patients

The derivation cohort consisted of 15,588 HCC patients, the internal validation cohort consisted of 3,897 HCC patients, while the multicenter validation cohort included 1,282 HCC patients. The median follow-up durations were 23.5 months for the derivation cohort, 24.5 months for the internal validation cohort, and 41.3 months for the multicenter validation cohort, respectively. The baseline characteristics of patients were displayed in Table 1. Compared to the derivation cohort, the external validation cohort had a significantly higher proportion of patients with BCLC B stage ($P < 0.001$), HBV infection ($P < 0.001$), high initial ALBI grade ($P < 0.001$) and high initial Child-Pugh class ($P < 0.001$). No significant difference in baseline characteristics were observed between internal validation cohort and derivation cohort. The first-line treatments for the derivation cohort and internal validation cohort are summarized in Table S3.

The fusion-SP developed by the derivation cohort

Using three months as the interval of each time slice, the time-series data of the derivation cohort was converted into data of time slices (Figure 1A), which included 15588, 10320, 7938, 6281, 5284, 4482, 3818, 3278, 2839, 2381, 1966 and 1561 patient with valid data, respectively. In accordance with the BCLC staging system (0/A, B, C, and D) at time slice no.1, we established four parent nodes for the development of raw-SP and fusion-SP models. The count of children nodes and those with sufficient sample size at each time slice for both SP models is detailed in Table S4.

The raw-SP model presented an exponential increase in the number of nodes since the first 4 time slices, which resulted in insufficient sample sizes for bifurcation of the majority nodes (71.9%) since time slice no. 4. The number of nodes were 32, 57, and 70 at time slices no. 4, 8, and 12, respectively. By contrast, the fusion-SP model (Figure 2 and Tables S4-5) effectively reduced number of nodes, with only 7, 5, and 6 child nodes at time slices no. 4, 8, and 12, respectively. Moreover, except for the time slice no.12, the majority of nodes (>80%) in fusion-SP had sample size large enough

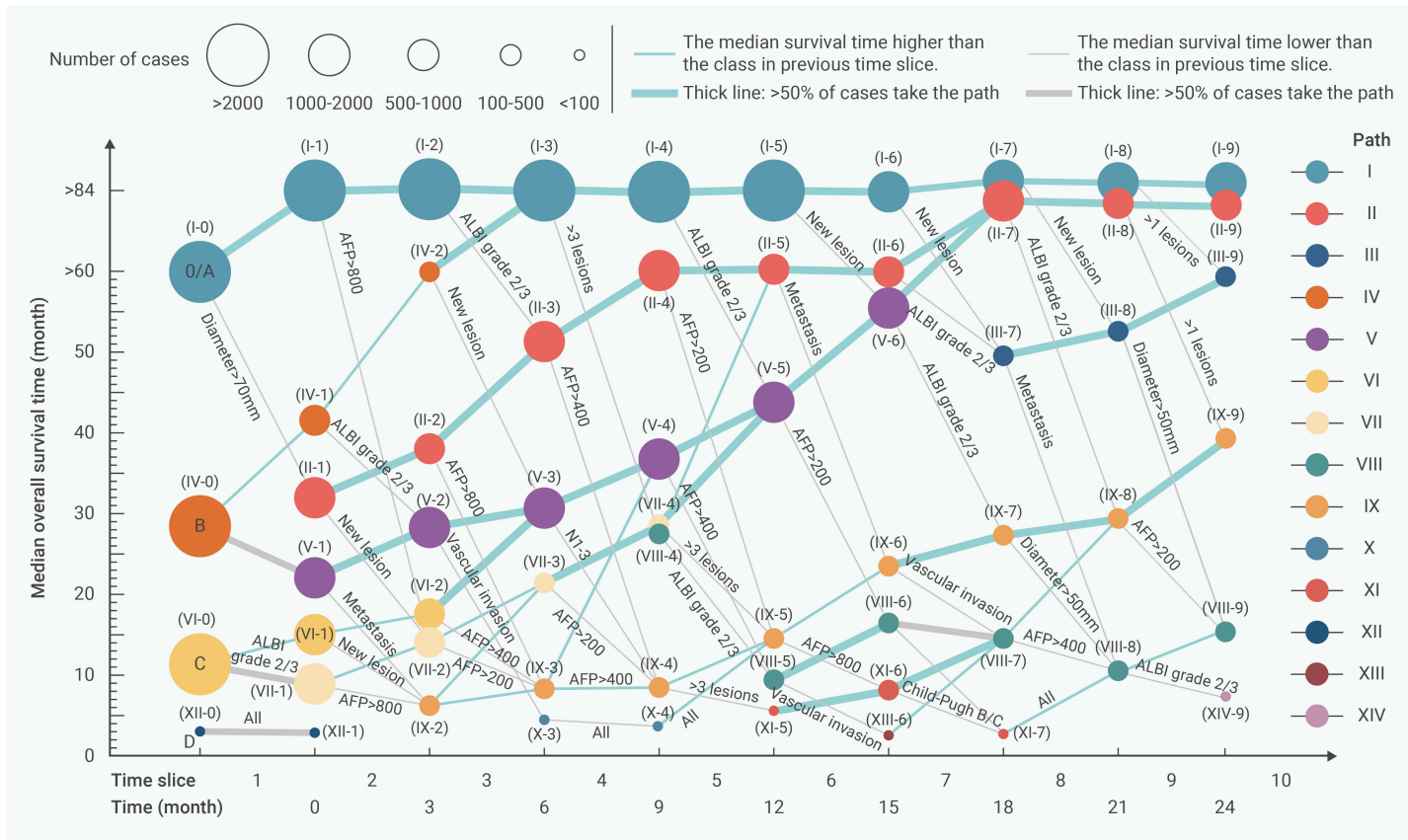


Figure 2. The survival path with nodal fusion technique for HCC patients Using patients' BCLC stage (0 or A, B, C, and D) at their initial diagnosis as starting nodes, the population was divided into cascades of subgroups, which was further visualized by two-dimensional graph, with the time slices on x-axis and median OS time on y-axis. A total of 14 different paths were constructed, which reflected the status change of patients in the first two years during dynamic treatment and monitoring. The nomenclature of nodes was based on the exact "name of path" and the "time slice" at which they are located.

(>100 cases) for robust survival estimation and further bifurcation.

The characteristics of patients in each node in the fusion-SP model were displayed in Figure 3A and Tables S6-8. Of the 14 paths constructed in fusion-SP model, four paths (Paths I, II, IV, and V) led to long-term survival (>60 months). The sustained improvement in clinical parameters among patients following these paths indicated a cure-directed disease evolution. It's notably that approximately 20% of patients in path I still carry tumors, which was well controlled after regular monitoring and timely treatment. However, disease progression could occur in a small proportion of patients on these paths due to occurrence of new lesions or deterioration of liver function (ALBI grade 2/3). The nodes in two paths (Paths XII and XIV) presented poor prognosis (median OS <10 months) and absence of upward transitional bifurcations in subsequent time slices. Therefore, these nodes were defined as incurable nodes. The tumor burden and tumor marker for patients on Paths XIII and XI gradually worsened as time progressed, indicating there lacked effective treatments to control the disease under the current medical care. The results of sensitivity analyses confirmed the robustness of our model (Table S9).

Enhanced dynamic prognosis prediction in fusion-SP

We compared the accuracy in dynamic prognostic prediction between the fusion-SP, raw-SP, BCLC staging system, and AJCC staging system in the internal and multicenter validation cohorts (Figure 3 and Tables S10-11). The predictive capabilities of the fusion-SP were consistently superior to BCLC or AJCC staging systems throughout all time slices (Figure 3B). In comparison with raw-SP model, the advantages of the fusion-SP started at time slice no.6 in the internal validation cohort, and no. 4 in the multicenter validation cohort. Notably, while no significant differences in the time slice-specific C-index between the raw-SP and fusion-SP models within the derivation cohort from time slice no. 6 and onwards, the fusion-SP model demonstrated significantly higher C-index values when compared to the raw-SP in time slices no.10 and no.12 in the internal validation cohort, and no.6 in the multicenter

validation cohort. The C-index values for the fusion-SP were 0.705 (95% CI: 0.702-0.709), 0.722 (0.717-0.727), and 0.662 (0.654-0.669), respectively, which were significantly higher than the values of 0.664 0.660-0.667), 0.655 (0.649-0.660), and 0.554 (0.547-0.562) of the raw-SP, respectively. The Kaplan-Meier curves (Figures 4 & S3) showed the fusion-SP significantly reduced the complexity of raw-SP models. Subgroup analysis based on BCLC stages, tumor burden and liver function (Figures 3C-D and Tables S12-14) revealed fusion-SP's superiority in patients with BCLC stage C disease and initial tumor burden above up-to-seven criteria. A performance comparison of between the fusion-SP model and state-of-the-art (SOTA) machine learning models - including Cox proportional hazards regression model, DeepHit survival neural network, and Random Forest for survival model - is presented in Table S15.

Individualized and dynamic treatment strategies based on fusion-SP

The recommended treatments for patients with advanced stage HCC after failure of first-line treatment follows fixed principles based on the BCLC guideline. As the fusion-SP has advantages in dynamic prognostication for this group of patients, we tentatively explored the optimal treatment under fusion-SP (Figure 5). The nodes VI-1 and VII-1 stemmed from the node of initial BCLC stage C disease (Node VI-0). The trajectory of Node initial C - Node VI-1 - Node VI-2 and trajectory of Node initial C - Node VII-1 - Node VII-2) went upward, while the trajectory of Node initial C - Node VI-1 - Node IX-2 and Node initial C - Node VII-1 - Node IX-2 went downward. Survival analysis revealed that all downward transfers significantly correlated with unfavorable prognosis ($P < 0.001$).

During treatment strategy analysis, the top seven treatment combinations employed among patients in Nodes VI-1 and VII-1 were analyzed for their value in facilitating transfer to upper node in the next time slice. In Node VII-1, patients had initial ALBI grade 1 and all prevalent regimens, compared to TACE/TAE, were linked to improved outcomes. By contrast, in Node VI-1,

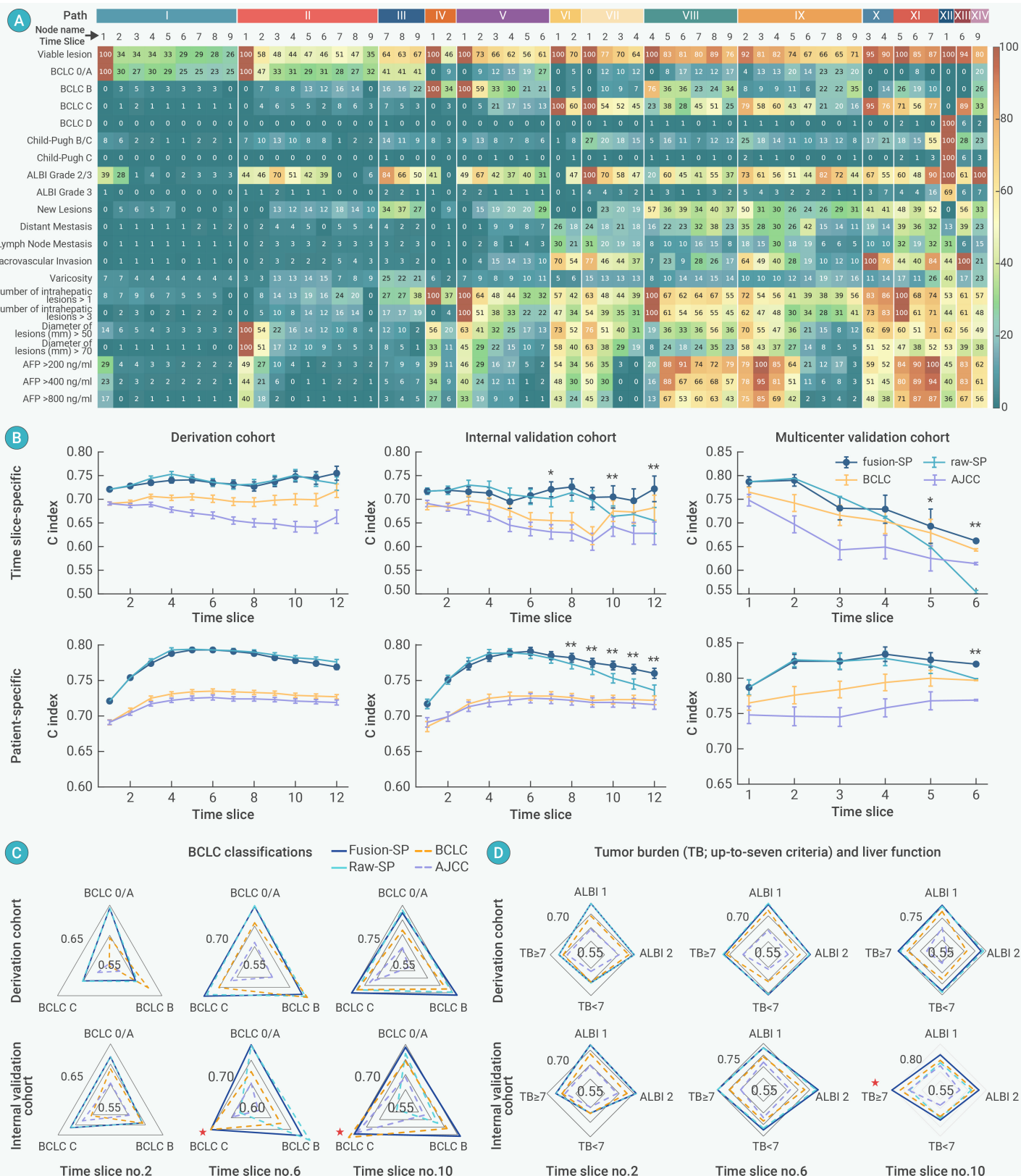


Figure 3. In-depth analysis of nodes in fusion-SP model and their discriminative ability for dynamic prognostication (A) A heatmap describing the ratio of the key characteristics of patients in each node was created, which reflects the trend of characteristics changes in each path of the fusion-SP model. (B) The line chart presents the c-index of different prediction system across distinct time slices. The comparison of C-indexes is depicted for the survival path with fusion (fusion-SP, blue solid lines), survival path without fusion (raw-SP, green dashed lines), the Barcelona Clinic Liver Cancer staging system (BCLC, yellow dashed lines), and the American Joint Committee on Cancer staging system (AJCC, purple dashed lines) in the derivation cohort, internal validation cohort, and multicenter validation cohort. Statistical significance between the fusion-SP and raw-SP models is denoted by asterisks, with a single asterisk (*) signifying P values less than 0.1 (marginal significance) and a double asterisk (**) indicating P values less than 0.05 (significance). (C, D) These radar charts present C-index of different prediction systems in the derivation and internal validation cohort by (C) BCLC stage, (D) tumor burden and liver function. Statistical significance between the fusion-SP and raw-SP models is denoted by asterisks, with an asterisk signifying P values less than 0.05.

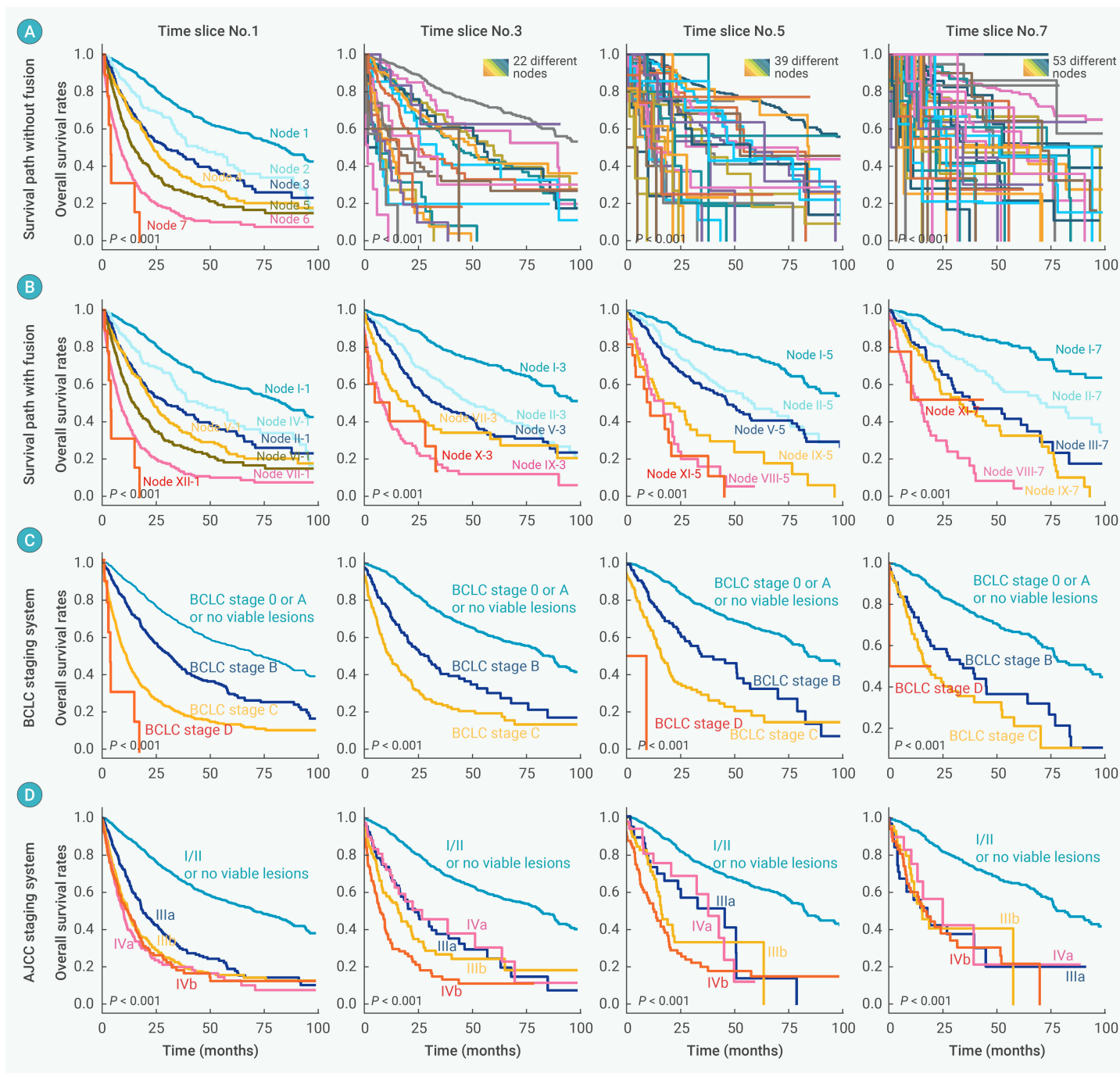


Figure 4. Kaplan-Meier Curves of the internal validation cohort Kaplan-Meier curves stratified by (A) nodes of Raw-SP model, (B) nodes of fusion-SP model, (C) BCLC staging system, and (D) AJCC staging system at time slice no.1, 3, 5 and 7 were displayed. The raw-SP model had more than twenty nodes since time slice no.3, which increased the difficulty in reading and application of the model. By contrast, the fusion technique greatly simplified the SP model for clinical application, while maintained the superiority in dynamic prognosis prediction compared to conventional staging systems.

patients had initial ALBI grade 2/3 and neither hepatic artery infusion chemotherapy (HAIC) alone, HAIC combined with systemic treatment, nor the triplet combination of HAIC, TACE/TAE and systemic treatment was associated with a higher chance of transferring to the upper node. The discrepancy in the therapeutic impact of different regimens was also observed among patients in Nodes VI-2 and VII-2 at time slice no.2. Although there were no significant differences in baseline characteristics between the two nodes (Figure 3A); in Node VII-2, all other prevalent regimens, except for triplet combination of HAIC, TACE/TAE and systemic treatment, were linked to improved outcomes compared to TACE/TAE. By contrast, in Node VI-2, no significant advantages for these prevalent regimens in transferring to subsequent upper node were found. Figure 6 showcased the fusion-SP model's capability to dynamic predict prognosis and optimize treatment strategies in

advanced HCC.

DISCUSSION

In this study, we utilized longitudinal data from HCC patients to construct SP models spanning all clinical stages of liver cancer. Through developing a novel nodal fusion algorithm for the survival model, we systemically evaluated its performance in dynamic prognosis prediction. Our findings demonstrated that nodal fusion algorithm can significantly enhance the prognosis predictive ability of the SP model, mitigate the risk of overfitting, and offer a more intuitive solution for visualizing and tracking patients' statuses. These advancements render the fusion-SP model a robust, powerful tool for dynamic prognosis prediction and personalized treatment planning in HCC patients.

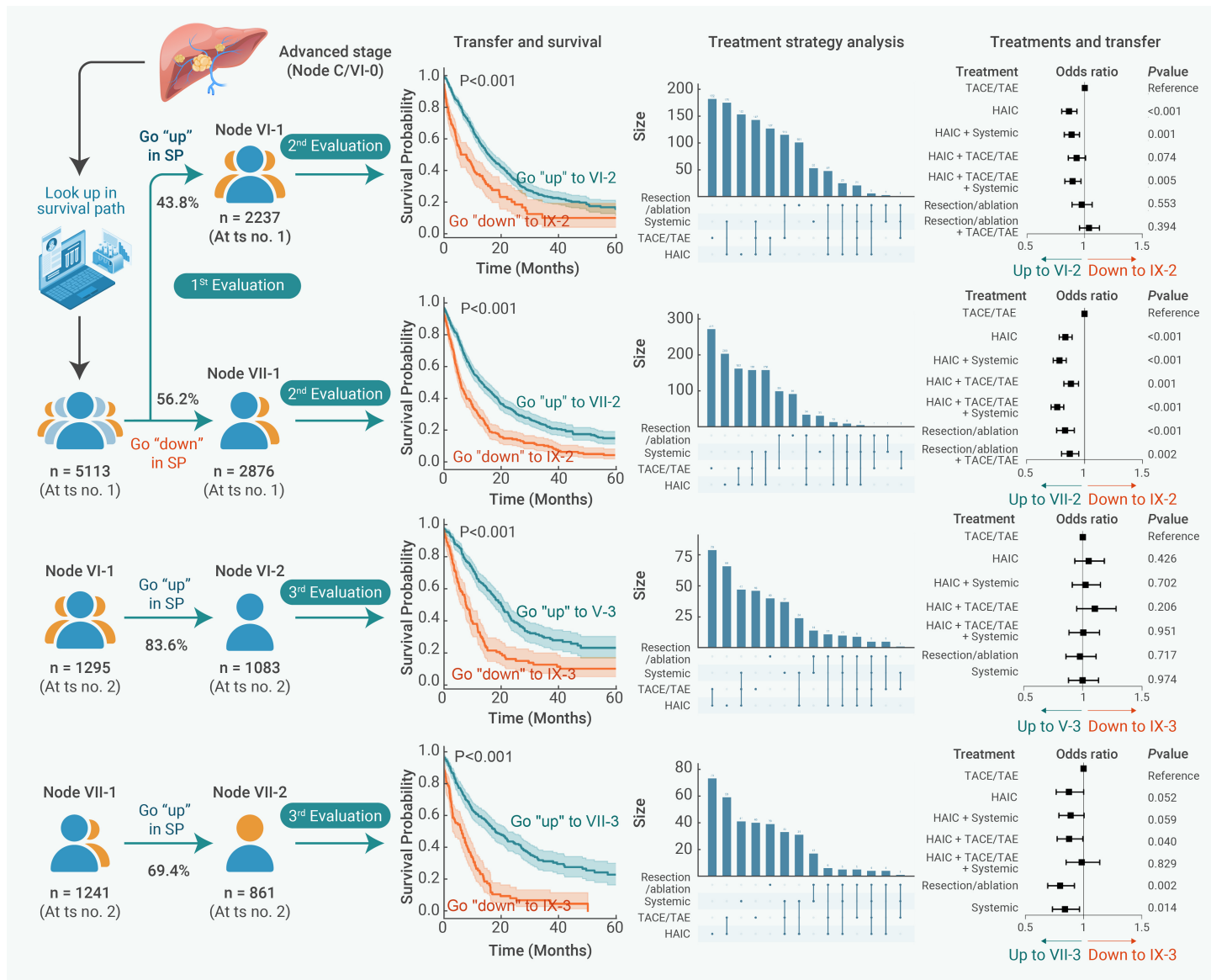


Figure 5. Dynamic treatment recommendations for advanced stage HCC patients based on fusion-SP model built by derivation cohort This figure presented a demonstration on dynamic treatment planning based on fusion-SP model. Patients with newly diagnosed advanced stage HCC underwent first evaluation and were categorized into two subgroups (node VI-1 and VII-1). The associations between subsequent transfers and overall survival were analyzed by Kaplan-Meier curves. Treatment strategy analysis based on multivariate logistic model was conducted to identify the optimal treatment that contribute to transfer to the upper node in the next time slice, as visualized by the forest plot. For patients went up to VI-2 from VI-1 and VII-2 from VII-1, additional round evaluation for the value of subsequent transfers and treatment strategy analysis was conducted. Based on the fusion-SP model, although VI-1 and VII-1 refers to advanced stage HCC patients at time slice no.1, the odds ratio for transferring to the "upper" node in the next time slice by different treatment arms process differs. Similar phenomenon is also observed among patients in next time slices (node VI-2 and VII-2).

While previous SP models have demonstrated superior accuracy in dynamic prognosis over static staging systems within specific timeframes, their sample size requirements increased exponentially with extended time horizons, resulting in deteriorating accuracy in long-term prognosis prediction, and a heightened risk of overfitting.⁵ Our study validates the efficacy of the fusion technique in addressing these issues. We observed a significant reduction in the number of nodes in fusion-SP as compared to the raw-SP, which made the model more robust in estimating the OS time for each fused node. The model's predictive accuracy on OS time estimation was reliably improved in remote time slices, evidenced by the results of time slices no.10 and 12 in the internal validation, as well as time slice no. 6 in external validation.

One notable advantage of utilization of the survival map is its ability to plainly demonstrate the potential transferring of disease status throughout the entire management cycle. In fact, the BCLC classification system can be seen as a set of such states, enabling prognosis prediction at any point during the dynamic treatment process.²⁵ Our previous experience in dynamic modeling for HCC found patients with specific BCLC classification at specific

time slice could be allocated in different nodes and share distinct survival duration and transition dynamics. The transitional complexity captured by SP models highlights the limitations of monitoring HCC development and progression solely through static, time-independent staging systems, such as BCLC or AJCC.

Beyond prognosis prediction, fusion-SP also has the potential to provide evidence-based treatment recommendations. Currently, robust evidence remains scarce for managing HCC patients after progression on first- and second-line therapies.^{26,27} By continuously monitoring patient trajectories within the SP framework, we can yield Level 2 evidence (per Oxford CEBM criteria) for dynamic prognosis prediction. Additionally, by comparing the efficacy of interventions at selected nodes using case-control designs, we can generate level 2-3 evidence for treatment planning.²⁸ In cases of poor prognosis and limited treatment options based on guidelines, the fusion-SP model could supplement clinical decision-making by offering data-driven therapeutic strategies. On the other hand, for patients in conditions with available mature treatments and expected good therapeutic effects, if the recommended treatment by fusion-SP model is not in line with the current stan-

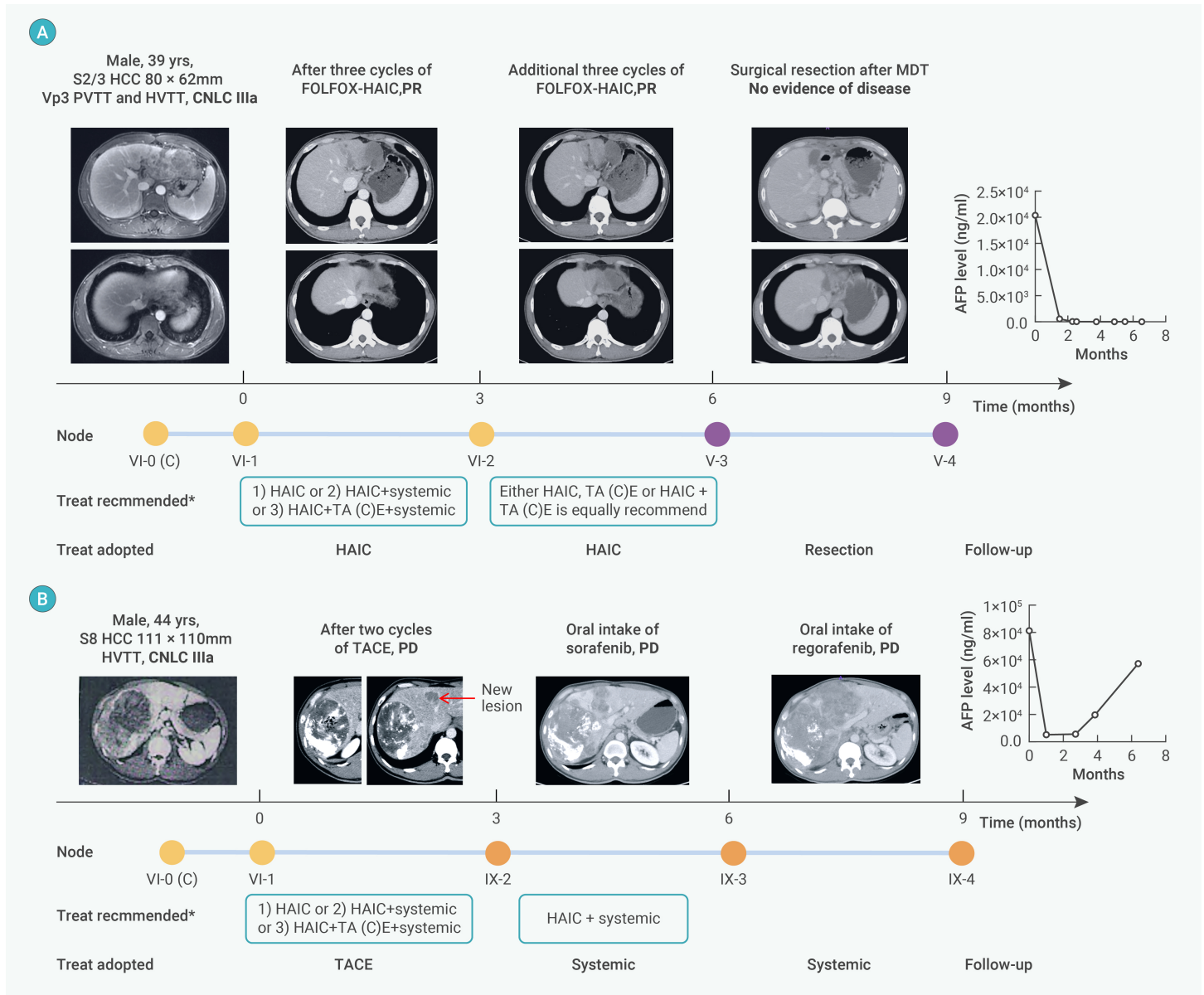


Figure 6. Representative cases of HCC in real world in demonstration the capability of fusion-SP model for dynamic prognosis prediction and treatment planning (A) A 39-year-old male patient with CNLC stage IIIa HCC was located at node VI-1 at initial diagnosis based on fusion-SP model. The recommended treatment could be either of the following: 1) HAIC, 2) HAIC+ systemic, 3) HAIC+TA(C)E+systemic. In real world, the patient received FOLFOX-HAIC and transferred to Node VI-2 at next time slice; he continued the HAIC treatment and received surgical resection in time slice no.3, with NED achieved and serum AFP level return to normal. (B) A 44-year-old male patient with CNLC stage IIIa HCC was also located at node VI-1 at initial diagnosis. He received conventional TACE as first-line treatment. Although the index tumor was well controlled, new lesions were identified at the end of the no.1 time slice, indicating the patient was transferred to Node IX-2. The patient received systemic therapies at subsequent time slices while the disease continued progressing. * refers to treatments recommended based on Fusion-SP model.

dard treatment, clinical discordance may emerge. In principle, guideline-endorsed therapies should remain the primary option. If the treatment recommended by the survival path model is also reasonable with growing evidence, prospective studies are needed to demonstrate its value.

Based on the fusion-SP model built in this study, three distinct trajectories (Paths I, II, and V) lead to prolonged survival (>60 months). However, disease progression could occur in a small proportion of patients on these trajectories due to occurrence of new lesions or deterioration of liver function (ALBI grade 2/3). This finding is consistent with our previous study in intermediate stage HCC patients,⁵ which provides important hints that adjuvant therapies may be warranted to mitigate the occurrence of new lesions and improve the liver function for this group of patients.²⁹

In Nodes VI-1, VII-1, and VII-2, we discovered that HAIC alone provided better survival outcome compared to TACE. Whereas TACE traditionally serves as a fundamental treatment for intermediate and advanced stage HCC, and is recommended as the first-line or second-line treatment in

CNLC,³⁰ BCLC,² and NCCN³¹ guidelines; on the other hand, HAIC is gaining popularity,^{32,33} particularly for patients with macrovascular invasion. These findings suggest that HAIC could take precedence over traditional TACE treatment when managing disease progression in patients within these nodes. Prospective trials are warranted to validate the value of HAIC among patients who failed TACE.

Although local treatments, such as resection or tumor ablation is conventionally contraindicated for BCLC stage B/C patients,² our model surprisingly identified local treatment might offer higher chance for transferring to subsequent upper nodes from nodes VII-1, VII-2. In fact, there are growing evidence that local treatments could bring survival benefit for highly selective advanced stage HCC patients after successful conversion therapy.^{34,35} More studies are needed to identify the advanced stage HCC patients who can benefit from potential curative treatment.

Theoretically, the fusion-SP model enables the identification of suitable treatment plans by analyzing the efficacy of different treatment methods for

Table 1. Baseline characteristics of derivation, internal validation, and multicenter external validation cohort at initial diagnosis.

N (%)	Derivation (N=15588)	Internal validation (N=3897)	<i>P</i> _{derivation vs. internal validation}	External validation (N=1282)	<i>P</i> _{derivation vs. external validation}
Demographics					
Male	13521 (86.7)	3386 (86.9)	0.828	1094 (85.3)	0.168
Age			0.561		
<50	5373 (34.5)	1342 (34.4)		466 (36.3)	0.242
[50, 60)	4931 (31.6)	1203 (30.9)		379 (29.6)	
≥ 60	5284 (33.9)	1352 (34.7)		437 (34.1)	
Etiology					
HBV	12726 (81.6)	3218 (82.6)	0.183	1156 (90.2)	<0.001
HCV	230 (1.5)	59 (1.5)	0.917	18 (1.4)	0.933
Staging and Liver function					
BCLC			0.832		<0.001
0/A	7709 (49.5)	1911 (49.0)		640 (49.9)	
B	2711 (17.4)	690 (17.7)		497 (38.8)	
C	5113 (32.8)	1279 (32.8)		126 (9.8)	
D	55 (0.4)	17 (0.4)		19 (1.5)	
AJCC			0.587		<0.001
I/II	8796 (56.4)	2174 (55.8)		846 (66.0)	
IIIa	1644 (10.5)	434 (11.1)		302 (23.6)	
IIIb	2689 (17.3)	688 (17.7)		52 (4.1)	
IVa	1170 (7.5)	300 (7.7)		49 (3.8)	
IVb	1289 (8.3)	301 (7.7)		33 (2.6)	
ALBI grade			0.31		<0.001
Grade 1	8326 (53.4)	2126 (54.6)		416 (32.4)	
Grade 2	6985 (44.8)	1711 (43.9)		711 (55.5)	
Grade 3	277 (1.8)	60 (1.5)		155 (12.1)	
ChildPugh score			0.69		
Class A	13921 (89.3)	3469 (89.0)		739 (57.6)	<0.001
Class B	1612 (10.3)	411 (10.5)		524 (40.9)	
Class C	55 (0.4)	17 (0.4)		19 (1.5)	
Imaging exam					
Distant Metastasis	1289 (8.3)	301 (7.7)	0.28	33 (2.6)	<0.001
Local Lymph Node Metastasis	1575 (10.1)	391 (10.0)	0.919	64 (5.0)	<0.001
Macrovascular Invasion	3784 (24.3)	950 (24.4)	0.91	81 (6.3)	<0.001
Varicosity	1236 (7.9)	285 (7.3)	0.212	232 (18.1)	<0.001
Amount of intrahepatic tumors			0.056		
= 1	9298 (59.6)	2315 (59.4)		582 (45.4)	<0.001
2 or 3	1955 (12.5)	541 (13.9)		315 (24.6)	
> 3	4335 (27.8)	1041 (26.7)		385 (30.0)	
Largest diameter of tumor (mm)			0.195		<0.001
≤ 50	7797 (50.0)	1978 (50.8)		809 (63.1)	
(50, 70]	2098 (13.5)	552 (14.2)		149 (11.6)	
> 70	5693 (36.5)	1367 (35.1)		324 (25.3)	
Tumor marker					
AFP (ng/ml)			0.599		<0.001
≤ 200	9028 (57.9)	2299 (59.0)		814 (63.5)	
(200, 400]	941 (6.0)	220 (5.6)		100 (7.8)	
(400, 800]	873 (5.6)	211 (5.4)		67 (5.2)	
> 800	4746 (30.4)	1167 (29.9)		301 (23.5)	

similar patients in each node. However, implementation of this strategy faces challenges as patients within the same node often received diverse treatment modalities, and treatment effects were influenced by various confounding factors. Despite the employment of multivariate adjusting could help reducing confounding effects,³⁶ current approaches still fall short in providing high-quality evidence for personalized treatment. Therefore, it is an important future direction to design treatment recommendation algorithms that optimally leverage large data and provide high-quality evidence. One potential solution lies in obtaining individualized treatment effects for each intervention through causal inference.^{37,38} Recent researches indicate that causal machine learning holds promise in deriving novel insights for estimating average treatment effects,^{39,40} and its integration into the fusion-SP model may enhance the quantity and quality of treatment related evidences. Proof-of-concept trials are essential to ascertain its value.

While our study has brought significant improvement to the conventional SP methodology, there remains enormous room for future development. Firstly, our model lacked the incorporation of multimodal features. Given the rapid advancements in radiomics⁴¹ and molecular diagnostics⁴² in HCC, integrating key features from multimodal sources may potentially bolster the model's performance. Secondly, recent studies showed the PIVKA-II is also an important biomarker for HCC and is of prognostic value, especially for AFP negative HCC patients.⁴³⁻⁴⁵ Incorporating this biomarker into our model could further improve its prognostic capabilities. However, due to the limited availability of PIVKA-II data in our cancer center before 2018, it was not included in the model. As data on this biomarker continues to accumulate, future iterations of our model will incorporate this biomarker. Thirdly, due to the limitation of sample size in both validation cohorts and accessibility of treatment information in external validation cohort, we were currently unable to validate so many correlations between interventions and path transfer in the fusion-SP model. Prospective studies are needed in the future to validate the therapeutic effect of the identified optimal intervention for patients in specific node. Also, despite our study's inclusion of a large cohort for modeling and multicenter cohort for testing, it remains a retrospective study; meanwhile, there existed significant sample size disparity and differences in baseline characteristics between derivation cohort and multicenter validation cohort, which might introduce bias in validation. Prospective validation studies are necessary to confirm the model's performance. Lastly, although fusion-SP, as a general framework, had the potential to be applied for dynamic prognostication across different cancer types, at present, we do not have access to data other than HCC. We are planning to expand the application of fusion-SP model in other malignant tumors in the future.

CONCLUSION

The application of fusion algorithms in the SP system has demonstrated the potential to significantly enhance the accuracy of dynamic prognosis prediction for HCC patients compared to raw-SP. Fusion-SP could serve as a new tool to facilitate dynamic prognostication and treatment planning for cancer patients in the future.

REFERENCES

- Brown Z. J., Tsilimigras D. I., Ruff S. M., et al. (2023). Management of Hepatocellular Carcinoma: A Review. *JAMA Surg.* **158**:410–420. DOI:10.1001/jamasurg.2022.7989
- Reig M., Forner A., Rimola J., et al. (2022). BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J. Hepatol.* **76**:681–693. DOI:10.1016/j.jhep.2021.11.018
- Tsilimigras D. I., Aziz H. and Pawlik T. M. (2022). Critical analysis of the updated Barcelona Clinic Liver Cancer (BCLC) group guidelines. *Ann. Surg. Oncol.* **29**:7231–7234. DOI:10.1245/s10434-022-12242-4
- Dhanasekaran R. (2021). Deciphering tumor heterogeneity in hepatocellular carcinoma (HCC)-multi-omic and singulomic approaches. *Semin. Liver Dis.* **41**:9–18. DOI:10.1055/s-0040-1722261
- Shen L., Zeng Q., Guo P., et al. (2018). Dynamically prognosticating patients with hepatocellular carcinoma through survival paths mapping based on time-series data. *Nat. Commun.* **9**:2230. DOI:10.1038/s41467-018-04633-7
- Vitale A., Cabibbo G., Iavarone M., et al. (2023). Personalised management of patients with hepatocellular carcinoma: A multiparametric therapeutic hierarchy concept. *Lancet Oncol.* **24**:e312–e322. DOI:10.1016/s1470-2045(23)00186-9
- Forrest I. S., Petrazzini B. O., Duffy Á., et al. (2023). Machine learning-based marker for coronary artery disease: Derivation and validation in two longitudinal cohorts. *Lancet* **401**:215–225. DOI:10.1016/s0140-6736(22)02079-7
- Martí-Juan G., Sanroma-Guell G. and Piella G. (2020). A survey on machine and statistical learning for longitudinal analysis of neuroimaging data in Alzheimer's disease. *Comput. Methods Programs Biomed.* **189**:105348. DOI:10.1016/j.cmpb.2020.105348
- Dai J., Xu H., Chen T., et al. (2025). Artificial intelligence for medicine 2025: Navigating the endless frontier. *Innov. Med.* **3**:100120. DOI:10.59717/j.xinn-med.2025.100120
- Liu Y., Chen Y. and Han L. (2023). Bioinformatics: Advancing biomedical discovery and innovation in the era of big data and artificial intelligence. *Innov. Med.* **1**:100012. DOI:10.59717/j.xinn-med.2023.100012
- Shen L., Mo J., Yang C., et al. (2023). SurvivalPath: A R package for conducting personalized survival path mapping based on time-series survival data. *PLoS Comput. Biol.* **19**:e1010830. DOI:10.1371/journal.pcbi.1010830
- Shen L., Jiang Y., Zhang T., et al. (2024). Machine learning for dynamic prognostication of patients with hepatocellular carcinoma using time-series data: Survival path versus dynamic-DeepHit HCC model. *Cancer Inform.* **23**:11769351241289719. DOI:10.1177/11769351241289719
- Feng G., Xu H., Wan S., et al. (2024). Twelve practical recommendations for developing and applying clinical predictive models. *Innov. Med.* **2**:100105. DOI:10.59717/j.xinn-med.2024.100105
- Alim A. and Karataş C. (2021). Prognostic factors of liver transplantation for HCC: Comparative literature review. *J. Gastrointest. Cancer* **52**:1223–1231. DOI:10.1007/s12029-021-00730-x
- Lu X., Meng J., Wang H., et al. (2023). DNA replication stress stratifies prognosis and enables exploitable therapeutic vulnerabilities of HBV-associated hepatocellular carcinoma: An *in-silico* precision oncology strategy. *Innov. Med.* **1**:100014. DOI:10.59717/j.xinn-med.2023.100014
- Kogalur H. I. a. U. B. (2007). Random survival forests for R. *R News* **7**:841–860.
- Kogalur H. I. a. U. B. (2023). Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC).
- Lauer H. I. a. U. B. K. a. E. H. B. a. M. S. (2008). Random survival forests. *Ann. Appl. Statist.* **2**:841–860.
- Breiman L. (2001). Random Forests. *Machine Learning* **45**:5–32. DOI:10.1023/A:1010933404324
- Hahsler M., Piekenbrock M. and Doran D. (2019). dbscan: Fast density-based clustering with R. *J. Stat. Softw.* **91**:1–30. DOI:10.18637/jss.v091.i01
- Andersen P. K. and Gill R. D. (1982). Cox's regression model for counting processes: A large sample study. *Ann. Stat.* **10**:1100–1120,1121.
- Lee C., Yoon J. and Schaar M. V. (2020). Dynamic-DeepHit: A deep learning approach for dynamic survival analysis with competing risks based on longitudinal data. *IEEE Trans. Biomed. Eng.* **67**:122–133. DOI:10.1109/tbme.2019.2909027
- Therneau T. M. (2024). A Package for Survival Analysis in R. <https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf>.
- Kang L., Chen W., Petrick N. A., et al. (2015). Comparing two correlated C indices with right-censored survival outcome: A one-shot nonparametric approach. *Stat. Med.* **34**:685–703. DOI:10.1002/sim.6370
- Ayuso C., Rimola J., Vilana R., et al. (2018). Diagnosis and staging of hepatocellular carcinoma (HCC): Current guidelines. *Eur. J. Radiol.* **101**:72–81. DOI:10.1016/j.ejrad.2018.01.025
- Villarruel-Melquiades F., Mendoza-Garrido M. E., García-Cuellar C. M., et al. (2023). Current and novel approaches in the pharmacological treatment of hepatocellular carcinoma. *World J. Gastroenterol.* **29**:2571–2599. DOI:10.3748/wjg.v29.i17.2571
- Huang T., Xu H., Wang H., et al. (2023). Artificial intelligence for medicine: Progress, challenges, and perspectives. *Innov. Med.* **1**:100030. DOI:10.59717/j.xinn-med.2023.100030
- Durieux N., Vandenput S. and Pasleau F. (2013). [OCEBM levels of evidence system]. *Rev. Med. Liege.* **68**:644–649.
- Llovet J. M., Pinyol R., Yarchoan M., et al. (2024). Adjuvant and neoadjuvant immunotherapies in hepatocellular carcinoma. *Nat. Rev. Clin. Oncol.* **21**:294–311. DOI:10.1038/s41571-024-00868-0
- Xie D. Y., Zhu K., Ren Z. G., et al. (2023). A review of 2022 Chinese clinical guidelines on the management of hepatocellular carcinoma: Updates and insights. *Hepatobiliary Surg. Nutr.* **12**:216–228. DOI:10.21037/hbsn-22-469
- Benson A. B., D'Angelica M. I., Abbott D. E., et al. (2021). Hepatobiliary cancers, version 2.2021, NCCN clinical practice guidelines in oncology. *J. Natl. Compr. Canc. Netw.* **19**:541–565. DOI:10.6004/jnccn.2021.0022
- Lyu N., Wang X., Li J. B., et al. (2022). Arterial chemotherapy of oxaliplatin plus fluorouracil versus sorafenib in advanced hepatocellular carcinoma: A biomolecular exploratory, randomized, phase III trial (FOHAIC-1). *J. Clin. Oncol.* **40**:468–480. DOI:10.1200/jco.21.01963
- Villanueva A. (2019). Hepatocellular carcinoma. *N. Engl. J. Med.* **380**:1450–1462. DOI:10.1056/NEJMr1713263
- Liang C., He Z., Tao Q., et al. (2023). From conversion to resection for unresectable hepatocellular carcinoma: A review of the latest strategies. *J. Clin. Med.* **12**. DOI:10.3390/jcm12247665
- Sun H. C., Zhou J., Wang Z., et al. (2022). Chinese expert consensus on conversion therapy for hepatocellular carcinoma (2021 edition). *Hepatobiliary Surg. Nutr.* **11**:227–252. DOI:10.21037/hbsn-21-328

36. Balzer L. B., Petersen M. L. and van der Laan M. J. (2016). Targeted estimation and inference for the sample average treatment effect in trials with and without pair-matching. *Stat. Med.* **35**:3717–3732. DOI:10.1002/sim.6965
37. Hernán M. A. and Robins J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *Am. J. Epidemiol.* **183**:758–764. DOI:10.1093/aje/kwv254
38. Yarmolinsky J., Wade K. H., Richmond R. C., et al. (2018). Causal inference in cancer epidemiology: What is the role of mendelian randomization. *Cancer Epidemiol. Biomarkers Prev.* **27**:995–1010. DOI:10.1158/1055-9965.Epi-17-1177
39. Feuerriegel S., Frauen D., Melnychuk V., et al. (2024). Causal machine learning for predicting treatment outcomes. *Nat. Med.* **30**:958–968. DOI:10.1038/s41591-024-02902-1
40. Liu R., Hunold K. M., Caterino J. M., et al. (2023). Estimating treatment effects for time-to-treatment antibiotic stewardship in sepsis. *Nat. Mach. Intell.* **5**:421–431. DOI:10.1038/s42256-023-00638-0
41. Xia T. Y., Zhou Z. H., Meng X. P., et al. (2023). Predicting microvascular invasion in hepatocellular carcinoma using CT-based radiomics model. *Radiology* **307**:e222729. DOI:10.1148/radiol.222729
42. Chidambaranathan-Reghupaty S., Fisher P. B. and Sarkar D. (2021). Hepatocellular carcinoma (HCC): Epidemiology, etiology and molecular classification. *Adv. Cancer Res.* **149**:1–61. DOI:10.1016/bs.acr.2020.10.001
43. Feng H., Li B., Li Z., et al. (2021). PIVKA-II serves as a potential biomarker that complements AFP for the diagnosis of hepatocellular carcinoma. *BMC Cancer* **21**:401. DOI:10.1186/s12885-021-08138-3
44. Su T. H., Wu C. H., Liu T. H., et al. (2023). Clinical practice guidelines and real-life practice in hepatocellular carcinoma: A Taiwan perspective. *Clin. Mol. Hepatol.* **29**:230–241. DOI:10.3350/cmh.2022.0421
45. Tian S., Chen Y., Zhang Y., et al. (2023). Clinical value of serum AFP and PIVKA-II for diagnosis, treatment and prognosis of hepatocellular carcinoma. *J. Clin. Lab. Anal.* **37**:e24823. DOI:10.1002/jcla.24823

FUNDING AND ACKNOWLEDGMENTS

This study was supported by grants from National Key R&D Program of China (2023YFC2414000) and National Natural Science Foundation of China (No. 81801804). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

We would like to thank Ms. Meizhen Zhu and Mr. Xueting Shen for giving their best love to their son, Dr. Lujun Shen and his family when he pursues the excellence in medicine. The authors sincerely thank Yidu Cloud (Beijing) Technology Ltd. for the establishment of big data intelligence platform at SYSUCC, Ms. Cong-hui Lu, Mr. Jun Wu and Ms. Shu-Yi Zheng from Yidu Cloud for assistance during data collection and analyses.

AUTHOR CONTRIBUTIONS

Conception and design: LJS, YQJ, MXC, WJF; Data analysis and interpretation: LJS, YQJ, LBL, MXC, JMX, QZ, KL, JN, JHH, BYC, QFC, NW; Resources: LFL, JHH, WJF; Funding acquisition: WJF; Writing-original draft: LJS, YQJ, LBL, MXC, WJF; Writing-review & editing: All authors. Final approval of manuscript: All authors. LJS (shenlj@sysucc.org.cn) is responsible for the concept of nodal fusion of survival paths. YQJ and SLJ have accessed and verified the data. All authors contributed to and approved the manuscript

DECLARATION OF INTERESTS

The authors declare no competing interests.

ETHICAL STATEMENT AND PATIENT CONSENT

The study protocol was approved by the Department of Clinical Research of SYSUCC (2023-FXY-213), with ethical approval obtained from the Hospital Ethics Committee of SYSUCC (B2023-634-01). Written informed consent was waived due to the retrospective nature of the study.

DATA AND CODE AVAILABILITY

Data are available from the corresponding author upon reasonable request. The code and demonstration data are publicly available on GitHub at <https://github.com/CMXyiduyun/Fusion-survival-path>. A web-based interface of fusion survival path model at www.fusionbot.cn was established to facilitate constructing models online.

SUPPLEMENTAL INFORMATION

It can be found online at <https://doi.org/10.59717/j.xinn-med.2025.100125>