frontiers | Frontiers in Oncology

TYPE Original Research
PUBLISHED 29 February 2024
DOI 10.3389/fonc.2023.1327147

OPEN ACCESS

EDITED BY
Francisco Tustumi,
University of São Paulo, Brazil

REVIEWED BY
Marina Alessandra Pereira,
Universidade de São Paulo, Brazil
Eric Nakamura,
University of São Paulo, Brazil
Fan Feng,
The 302th Hospital of PLA, China

*CORRESPONDENCE
Weidong Pei
peiwd@aliyun.com

RECEIVED 24 October 2023
ACCEPTED 08 December 2023
PUBLISHED 29 February 2024

CITATION
Zhang Z-H, Du Y, Wei S and Pei W (2024) Multilayered insights: a machine learning approach for personalized prognostic assessment in hepatocellular carcinoma. *Front. Oncol.* 13:1327147. doi: 10.3389/fonc.2023.1327147

Multilayered insights: a machine learning approach for personalized prognostic assessment in hepatocellular carcinoma

Zhao-Han Zhang[1], Yunxiang Du[2], Shuzhen Wei[2] and Weidong Pei[3]*

[1]Shenyang No.20 High School, Shenyang, China, [2]Department of Oncology, Huai'an 82 Hospital, China RongTong Medical Healthcare Group Co., Ltd., Chengdu, China, [3]Department of Discipline Development, China RongTong Medical Healthcare Group Co., Ltd., Chengdu, China

Background: Hepatocellular carcinoma (HCC) is a complex malignancy, and precise prognosis assessment is vital for personalized treatment decisions.

Objective: This study aimed to develop a multi-level prognostic risk model for HCC, offering individualized prognosis assessment and treatment guidance.

Methods: By utilizing data from The Cancer Genome Atlas (TCGA) and the Surveillance, Epidemiology, and End Results (SEER) database, we performed differential gene expression analysis to identify genes associated with survival in HCC patients. The HCC Differential Gene Prognostic Model (HCC-DGPM) was developed through multivariate Cox regression. Clinical indicators were incorporated into the HCC-DGPM using Cox regression, leading to the creation of the HCC Multilevel Prognostic Model (HCC-MLPM). Immune function was evaluated using single-sample Gene Set Enrichment Analysis (ssGSEA), and immune cell infiltration was assessed. Patient responsiveness to immunotherapy was evaluated using the Immunophenoscore (IPS). Clinical drug responsiveness was investigated using drug-related information from the TCGA database. Cox regression, Kaplan-Meier analysis, and trend association tests were conducted.

Results: Seven differentially expressed genes from the TCGA database were used to construct the HCC-DGPM. Additionally, four clinical indicators associated with survival were identified from the SEER database for model adjustment. The adjusted HCC-MLPM showed significantly improved discriminative capacity (*AUC*=0.819 *vs.* 0.724). External validation involving 153 HCC patients from the International Cancer Genome Consortium (ICGC) database verified the performance of the HCC-MLPM (*AUC*=0.776). Significantly, the HCC-MLPM exhibited predictive capacity for patient response to immunotherapy and clinical drug efficacy ($P < 0.05$).

Conclusion: This study offers comprehensive insights into HCC prognosis and develops predictive models to enhance patient outcomes. The

evaluation of immune function, immune cell infiltration, and clinical drug responsiveness enhances our comprehension and management of HCC.

KEYWORDS

hepatocellular carcinoma, prognosis risk model, machine learning, immune function, drug responsiveness

1 Introduction

Primary liver cancer, a prevalent malignancy of the digestive system, ranks as the sixth most frequently occurring tumor globally and is the second leading cause of mortality (1, 2). Hepatocellular carcinoma (HCC) is the prevailing pathological subtype of primary liver cancer, accounting for 75%-85% of all cases (3). The poor prognosis of HCC arises from its early propensity for metastasis, often involving dissemination to the portal vein or distant organs (4). Patients with early-stage HCC have access to a potentially curative treatment option with a long-term survival rate exceeding 5% at 60 years, while patients with advanced-stage tumors experience a median survival period ranging from 1 to 2 years (5–7). Therefore, timely identification, early intervention, and the implementation of rational and effective treatment strategies are crucial for patients diagnosed with HCC (8).

Surgical resection is considered represents the primary therapeutic approach for patients with early-stage HCC and often leads to favorable outcomes (9, 10). However, for individuals diagnosed with intermediate or advanced-stage HCC, surgical resection is no longer a feasible option due to tumor progression and metastasis. Local regional therapies, such as ablation, arterial-directed therapies, or external beam radiation therapy, are the preferred treatment modalities for patients with localized liver disease that cannot be surgically removed or are not suitable for surgery. Systemic therapies are recommended for patients who undergo disease progression after local regional therapies or those with metastases outside the liver (11). This focus on systemic therapies highlights the importance of considering the tumor microenvironment, drug responsiveness, and immunotherapy as crucial factors (12–14). Targeted therapies are particularly relevant for patients diagnosed with intermediate or advanced-stage HCC (15, 16). Sorafenib, initially approved for advanced HCC treatment, is hindered by the development of resistance (17, 18). Subsequently, other multi-kinase inhibitors, such as Lenvatinib, Regorafenib, Cabozantinib, and the VEGFR2 inhibitor ramucirumab, have been approved as second-line targeted treatment options (19). With a deeper understanding of the interplay between the immune system and cancer, immune checkpoint inhibitors (ICI) have been integrated into the therapeutic arsenal for patients with advanced HCC. Nivolumab, an ICI agent, has been FDA approval for the management of advanced HCC (20, 21). Given the expanding range of treatment methods, the selection of the most suitable treatment plan for patients has become critical.

Therefore, to provide optimal treatment approaches for different stages of HCC progression, conventional methods often assign patients to specific stages based on the Barcelona Clinic Liver Cancer (BCLC) classification (6, 22). BCLC staging, widely utilized in HCC, categorizes patients into different stages based on factors such as tumor size, number, liver function, and symptoms (23). Treatment strategies, such as surgical resection, liver transplantation, radiofrequency ablation, radiation therapy, and targeted therapy, are determined for patients according to their corresponding stages (6). However, significant heterogeneity exists among patients, including genetic variations, immune environments, and tumor heterogeneity. Relying solely on conventional staging methods may insufficiently consider individual patient characteristics and the complexities of the disease, potentially leading to inaccurate prognosis assessments and suboptimal personalized treatments.

Recent advancements in tumor genomics research have facilitated the utilization of extensive tumor genomic data to gain insights into the complexity and individual variations of tumors. The Cancer Genome Atlas (TCGA) database, as a comprehensive repository of diverse cancer-related data, offers new opportunities for exploring prognostic risk assessment in HCC (24, 25). Therefore, in contrast to conventional approaches, we consider incorporating factors such as molecular biology information and the tumor microenvironment based on clinical indicators. By leveraging the potential of powerful machine learning and big data analysis techniques, we can extract valuable insights from extensive tumor genomic data to construct prognostic risk assessment models.

Through a comprehensive analysis of clinical indicators, molecular biology information, tumor microenvironment, and other multi-level factors, o our objective is to establish a

Abbreviations: AUC, Area under the curve; BCLC, Barcelona Clinic Liver Cancer; DEG, Differentially expressed gene; GSEA, Gene Set Enrichment Analysis; HCC, Hepatocellular carcinoma; HCC-DGPM, Hepatocellular Carcinoma Differential Gene Prognostic Model; HCC-MLPM, Hepatocellular Carcinoma Multilevel Prognostic Model; HR, Hazard ratios; ICGC, International Cancer Genome Consortium; IPS, Immunophenoscore; KM, Kaplan-Meier; NCCN, National Comprehensive Cancer Network; OS, Overall survival; ROC, Receiver operating characteristic; SEER, Surveillance, Epidemiology, and End Results; ssGSEA, Single-sample Gene Set Enrichment Analysis; TCGA, The Cancer Genome Atlas.

comprehensive and accurate HCC prognostic risk model and explore its association with drug responsiveness and immunotherapy (26, 27). By improving the accurate assessment of prognostic risk in HCC patients, we can provide essential evidence to inform the development of personalized treatment plans, thus improving prognosis assessment and treatment outcomes for patients. Moreover, by leveraging the extensive resources of databases such as TCGA and SEER, this study has the potential to make significant breakthroughs and advancements in the field of HCC prognostic assessment and personalized treatment.

2 Research design and methods

2.1 Research workflow

The research workflow (depicted in Figure 1) comprised the subsequent steps: 1) Identification of differentially expressed genes (DEGs) associated with survival through gene expression analysis; 2) Development of the HCC Differential Gene Prognostic Model (HCC-DGPM) by integrating significant DEGs; 3) Selection of clinical indices linked to survival; 4) Model adjustment and validation, culminating in the HCC Multilevel Prognostic Model (HCC-MLPM); 5) Evaluation of immune function and analysis of clinical drug responsiveness.

2.2 Data collection and preparation

The dataset for liver cancer was obtained from the TCGA database through the website, and the gene expression matrices of adjacent non-cancerous and cancerous tissues were used for the analysis. The dataset consisted of 50 samples of normal liver tissue and 370 samples of liver cancer. Additionally, data from 30,684 patients diagnosed with primary liver cancer between 1988 and 2015 were extracted from the SEER database using SEER*stat software 8.4.0 (https://seer.cancer.gov/). After data cleansing, a total of 3,017 patient records were available for further analysis.

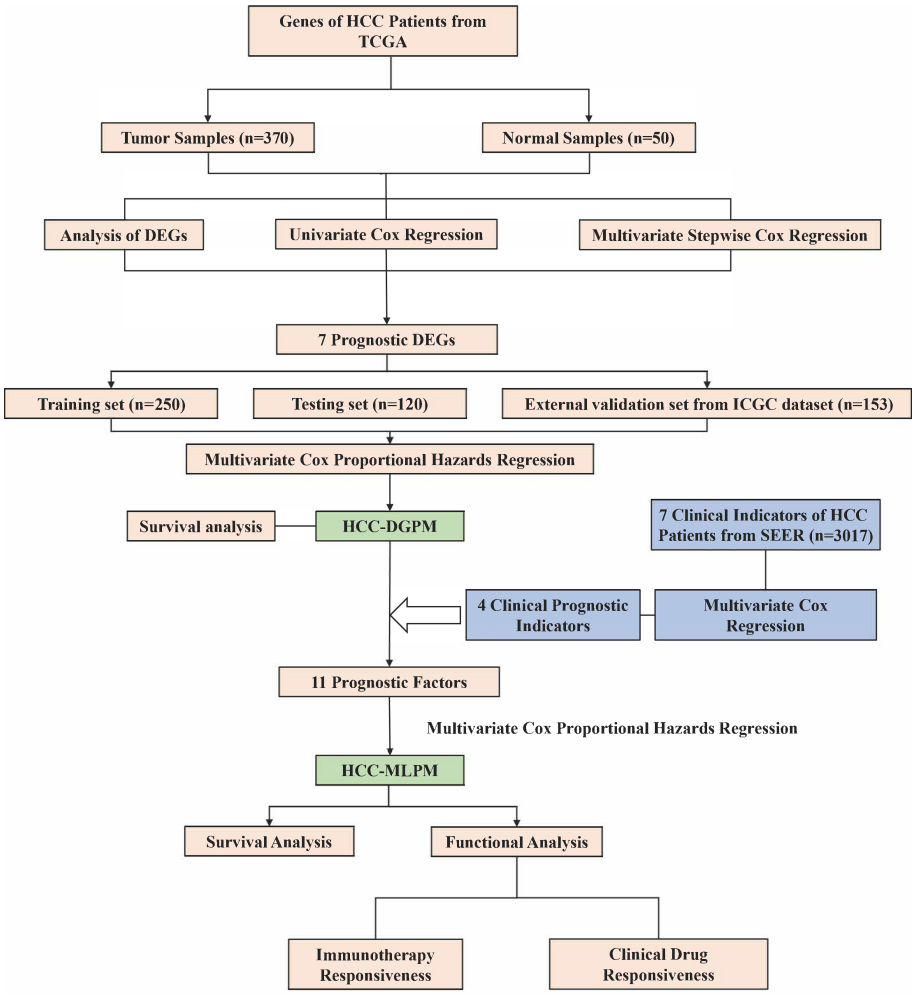

FIGURE 1
Research workflow for the construction of hepatocellular carcinoma prognostic model. HCC, Hepatocellular carcinoma; TCGA, The Cancer Genome Atlas; SEER, Surveillance, Epidemiology, and End Results; DEGs, Differentially expressed genes; HCC-DGPM, HCC Differential Gene Prognostic Model; HCC-MLPM, Hepatocellular Carcinoma Multilevel Prognostic Model.

The International Cancer Genome Consortium (ICGC) database provided data from 369 HCC patients for the study. After further data cleaning, cases with incomplete clinical information were excluded, resulting in a final sample size of 153 cases.

It is important to note that all the included patients were diagnosed with primary liver cancer. The year of the initial diagnosis was categorized into 5-year intervals and considered as an ordinal variable. Age 45 was chosen as the threshold to classify cases with early-onset HCC, and age was divided into 10-year intervals.

2.3 Selection of HCC prognostic DEGs

The "limma" package in R was used to identify the DEGs between cancerous and adjacent non-cancerous tissues (28). The DEG threshold was set as an absolute log2-fold change (FC) ≥ 1 and an adjusted $P < 0.05$. Volcano plots illustrating the DEGs were generated using the "ggplot2" package (https://ggplot2.tidyverse.org/). Subsequent screening involved conducting both univariate and stepwise multivariate Cox regression analyses. In the univariate Cox regression analysis, each DEG was evaluated individually to assess its association with the survival outcome. In this context, the survival outcomes were solely considered as "death". This analysis facilitated the identification of genes that exhibited a significant correlation with patient survival. Subsequently, a stepwise multivariate Cox regression analysis was performed.

2.4 Construction and validation of HCC-DGPM

The patients from the TCGA dataset were randomly assigned to training (n = 250) and testing (n = 120) sets in a 7:3 ratio, facilitated by the "caret" R package for random assignment (29). The training set was used to train the model, while the testing set was used to assess the predictive performance of the model. HCC-DGPM was constructed using the multivariate Cox regression method. External validation set was performed using the ICGC dataset (n = 153). The performance of HCC-DGPM was evaluated using receiver operating characteristic (ROC) curves, with a higher area under the curve (AUC) indicating improved predictive accuracy. To enhance the precision of our prognostic model, calibration curves were utilized, employing the "rms", "survival", and "ResourceSelection" R packages (https://CRAN.R-project.org/package). These curves are a measure of how closely the model's predictions align with actual outcomes. The closer these curves lie to the 45-degree line, the more accurate the model is, indicating a high degree of concordance between predicted and observed results. The HCC-DGPM was validated with Kaplan-Meier (KM) curves, and the methodology for establishing the cutoff value for risk groups was not initially specified. The cutoff value used for delineating high and low risk groups was determined by the method that maximizes (sensitivity + specificity - 1). The established cutoff value for the risk score was 1.65.

2.5 Model adjustment and validation

To improve the model's performance, we performed multivariate Cox regression analysis to identify clinical indicators associated with HCC patient survival using the SEER database. SSubsequently, these indicators were used to refine the HCC-DGPM, resulting in the HCC-MLPM. ROC curves and KM curves were generated to assess the performance of the HCC-MLPM and provide additional validation of its effectiveness.

2.6 Immune evaluation of the model

Gene Set Enrichment Analysis (GSEA) was performed to investigate the impact of the risk score on the biological function of HCC patients. The annotated gene setlist was selected using a significance threshold of $P < 0.05$. Moreover, the "ssGSEA" R package was used to estimate the infiltration levels of 28 distinct immune cell types in HCC patients (30), taking into account their risk scores. The IPS was used to assess patient responsiveness to immunotherapy (31).

2.7 Clinical drug responsiveness evaluation

To explore the variations in clinical drug responsiveness among patients, we analyzed the clinical drug information and patients' drug responsiveness data retrieved from the TCGA database. We evaluated the effects of chemotherapy drugs such as Gemcitabine, Cisplatin, Doxorubicin, 5-fluorouracil (5-FU), Oxaliplatin, Adriamycin, and Cytoxan, alongside targeted therapy agents including Sorafenib, Everolimus, Sunitinib, and Temsirolimus. This comprehensive assessment was crucial as it is well-recognized that therapeutic efficacy varies significantly between treatments, independent of other evaluated variables.

Based on the risk scores generated by our model, patients were categorized into high-risk and low-risk groups. Subsequently, a proportional stacked bar chart was used to visually represent and analyze the disease progression in these two patient groups.

2.8 Statistical analysis

Cox regression models were employed to calculate hazard ratios (HR) and assess the relationship between gene expression and survival outcomes. The KM method was utilized to generate survival curves, and the log-rank test was applied to compare these curves. The Jonckheere-Terpstra test and Cochran-Mantel-Haenszel test were performed to evaluate the trend association between the diagnosis year and patient characteristics for numerical and categorical data, respectively. Cox proportional hazards regression models were used to HRs and their corresponding 95% confidence intervals for prognostic factors related to overall survival (OS). In the multivariate Cox regression analyses, a stepwise procedure was conducted with an entry criterion of $P < 0.05$ to identify the most statistically significant

prognostic factors. The significance level for all statistical tests was set at $P < 0.05$. Statistical analyses were conducted using SAS 13.2 (SAS Institute, Cary, NC, USA), and the KM curves were plotted using R Software.

2.9 Availability of data

The data utilized in this study can be obtained by contacting the authors due to restrictions imposed by the data providers, namely TCGA, SEER, and GSEA databases. Access to these databases is available via their dedicated websites: TCGA (https://portal.gdc.cancer.gov/), SEER (https://seer.cancer.gov/), GSEA (http://software.broadinstitute.org/gsea/index.jsp), and ICGC (https://icgcportal.genomics.cn/). Researchers interested in accessing the data may reach out to the authors for additional information and support in acquiring the required permissions and data access.

3 Results

3.1 Analysis of differential gene expression

To establish a prognostic model, 370 HCC samples and 50 samples of normal liver tissue from TCGA were involved. Compared to normal group, a total of 1761 DEGs were identified, comprising 1,091 upregulated genes and 670 downregulated genes in HCC group. The expression patterns of these DEGs are illustrated in Figure 2A through volcano plots.

3.2 Identification of genes associated with patient prognosis

To further identify the prognostic genes, a univariate Cox regression was performed on the 1,761 genes to identify genes significantly associated with patient prognosis. This analysis yielded a subset of 89 genes that showed a significant association at a significance level of $P < 0.001$. Further analysis using multivariate Cox stepwise regression identified seven genes significantly associated with the survival of HCC patients: MYBL2, SF3B4, CDCA8, NUF2, HMMR, PON1, and PAGE1 (Table 1).

KM analysis was performed to assess the impact of the seven identified genes on patient survival. The results revealed a significant correlation between these genes and patient prognosis, as evidenced by distinct survival patterns observed in patient groups with high and low expression levels of these genes. Moreover, the survival analysis demonstrated that patients with higher gene expression levels had a significantly poorer prognosis compared to those with lower expression levels (Figures 2B–H).

3.3 Construction and validation of the HCC-DGPM

Therefore, HCC-DGPM was constructed through univariate Cox proportional hazards regression analysis using the expression levels of the seven identified genes (MYBL2, SF3B4, CDCA8, NUF2, HMMR, PON1, and PAGE1) as covariates. The regression coefficients were used to assign weights to each gene, allowing for the development of a risk score formula to calculate the individual risk score for each patient The risk score formula is defined as follows: Risk score = (Expression level of MYBL2 × 0.4820) + (Expression level of SF3B4 × 3.0446) + (Expression level of CDCA8 × 3.1851) + (Expression level of NUF2 × 0.1932) + (Expression level of HMMR × 3.1205) + (Expression level of P0N1 × 0.7698) + (Expression level of PAGE1 × 1.2691).

ROC curve was performed to assess the predictive ability of the HCC-DGPM in determining patient outcomes. The training dataset exhibited an AUC of 0.723 (Figure 3A) for the HCC-DGPM, while the testing and external validation sets showed AUC values of 0.724 & 0.719 (Figures 3B, C). These results suggest that the HCC-DGPM has a moderate predictive ability to distinguish between high-risk and low-risk patients. To enhance the credibility of our model's accuracy, we performed calibration curve analyses following the ROC assessments (Figures 3D–F). The results from these calibration curves lend further credence to the model's predictive acumen, highlighting its prospective value in a clinical setting.

In addition, survival analysis was conducted based on the newly calculated risk score, allowing for the classification of patients into high-risk and low-risk groups for model validation purposes. The KM curves demonstrated significant differences in survival between the high-risk and low-risk groups (Figures 3G–H).

3.4 Model adjustment

Clinical indicators, including Age, Race, Sex, tumor size (T), node involvement (N), metastasis (M), and stage, were screened from the SEER database due to their potential correlation with patient survival in HCC. Four indicators were identified as significantly associated with survival outcomes (Table 2). Multivariate Cox regression analysis was conducted to determine the clinical factors significantly associated with patient survival. Four indicators were found to be significantly correlated with survival outcomes (Table 2). Next, the identified clinical indicators from the SEER database were integrated with the risk scores obtained from the 7 DEGs, resulting in the development of a novel predictive model (HCC-MLPM). The adjusted predictive HCC-MLPM is represented by the following formula: Risk Score = (Expression level of MYBL2 × 0.4820) + (Expression level of SF3B4 × 3.0446) + (Expression level of CDCA8 × 3.1851) + (Expression level of NUF2 × 0.1932) + (Expression level of HMMR × 3.1205) + (Expression level of PON1 × 0.7698) + (Expression level of PAGE1 × 1.2691) + (Age × 1.5079) + (T × 2.9376) + (N × 0.8721) + (M × 3.0453).

3.5 Evaluation of HCC-MLPM

The performance of the HCC-MLPM was evaluated using both the training and testing datasets. In the training dataset, the HCC-MLPM demonstrated improved predictive ability with an AUC of 0.826 (Figure 4A). In the testing dataset, the HCC-MLPM achieved

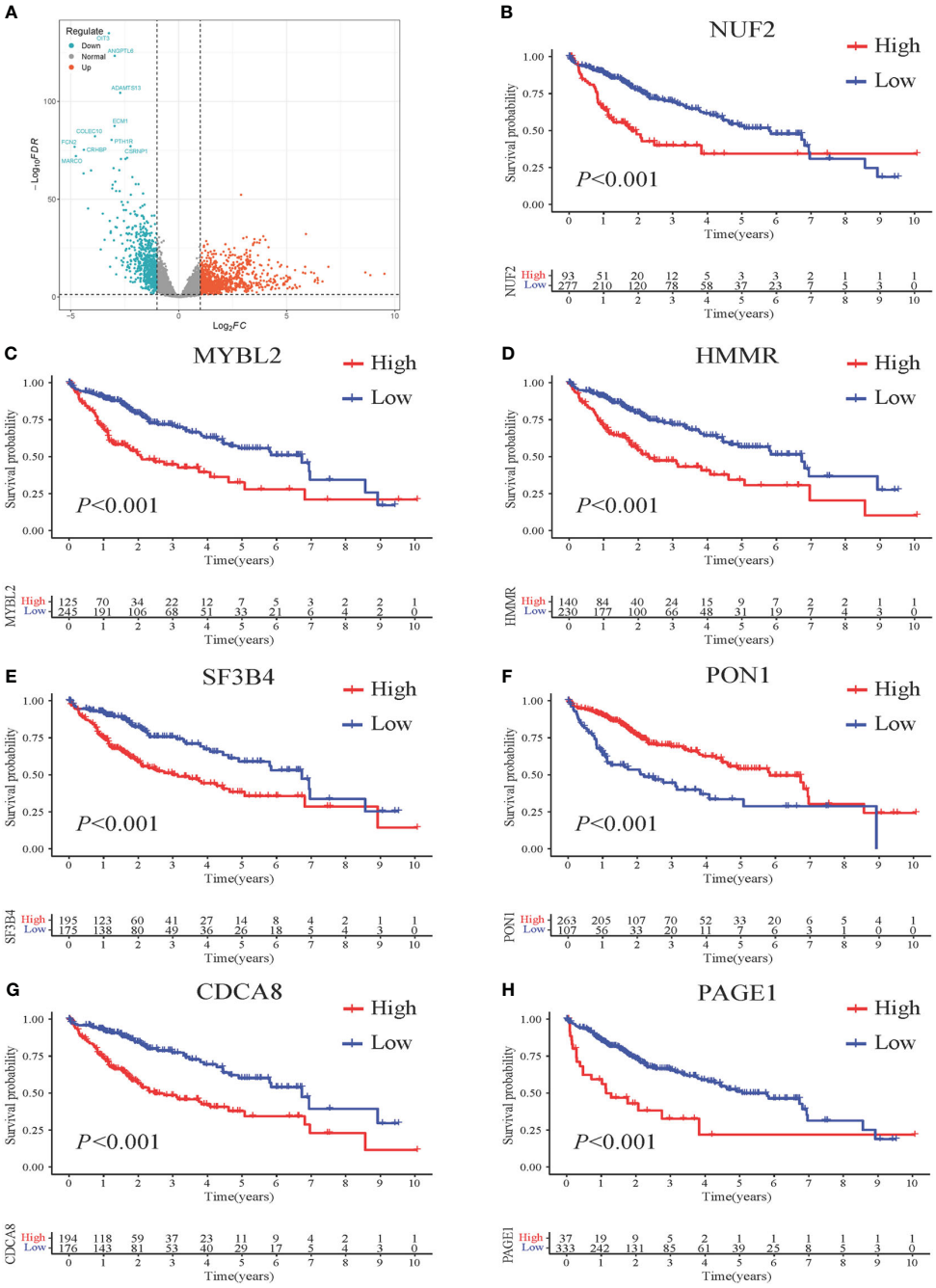

FIGURE 2
Survival impact of 7 DEGs screened from HCC patients. **(A)** Volcano plots illustrating the DEGs from HCC *vs*. normal. Genes upregulated in HCC are represented in red, while genes downregulated in HCC are shown in blue. The x-axis represents the log2-fold change in gene expression, indicating the magnitude of change, and the y-axis represents the statistical significance (-log10 p-value) of the differential expression. **(B–H)** Kaplan-Meier (KM) curves demonstrating the association between gene expression levels and patient survival. High-risk patients are depicted in red, while low-risk patients are represented in blue. The x-axis represents the survival time, and the y-axis represents the survival probability.

an AUC of 0.819 (Figure 4B), further validating its enhanced predictive capacity. Similarly, the HCC-MLPM exhibited satisfactory predictive capability in the external validation dataset, achieving an AUC of 0.776 (Figure 4C). These results indicate that the HCC-MLPM effectively discriminates between high-risk and low-risk patients in this external validation set. Calibration curves were subsequently integrated, serving as an additional verification stratum for the model's validity (Figures 4D–F).

Furthermore, by utilizing the scoring system derived from the adjusted model, patients were categorized into high-risk and low-risk groups. KM curves showed that patients in the low-risk group had superior overall survival outcomes compared to those in the high-risk group (Figures 4G–H). These findings indicate that the integration of risk scores based on differential gene expression, along with the selected clinical indicators, significantly improved the predictive performance of the HCC-MLPM.

TABLE 1 Differential genes associated with OS in HCC patients.

Name	HR	HR.95L	HR.95H	*P value*
MYBL2	0.4820	0.31979	0.7264	0.000488
SF3B4	3.0446	1.90062	4.8770	3.63e-06
CDCA8	3.1851	1.65058	6.1463	0.000552
NUF2	0.1932	0.08874	0.4205	3.43e-05
HMMR	3.1205	1.73930	5.5985	0.000136
PON1	0.7698	0.67562	0.8771	8.47e-05
PAGE1	1.2691	1.11467	1.4448	0.000318

3.6 Stratified survival analysis based on clinical indicators

This section delves into a detailed survival analysis of HCC patients within the HCC-MLPM framework, stratified according to key clinical indicators. The KM curves display distinct survival probabilities over time for groups stratified by key clinical indicators: Age (Figure 5A), T (Figure 5B), N (Figure 5C), and M (Figure 5D). These curves reveal considerable variation in survival outcomes across these different clinical stratifications (P<0.001), underscoring the significant impact of each indicator on survival. They highlight the potential utility of these clinical indicators in refining the HCC-MLPM.

3.7 Gene set enrichment analysis

To assess the immune function associated with the HCC-MLPM, Gene Set Enrichment Analysis (GSEA) was performed. The analysis revealed that high-risk patients showed a stronger association with

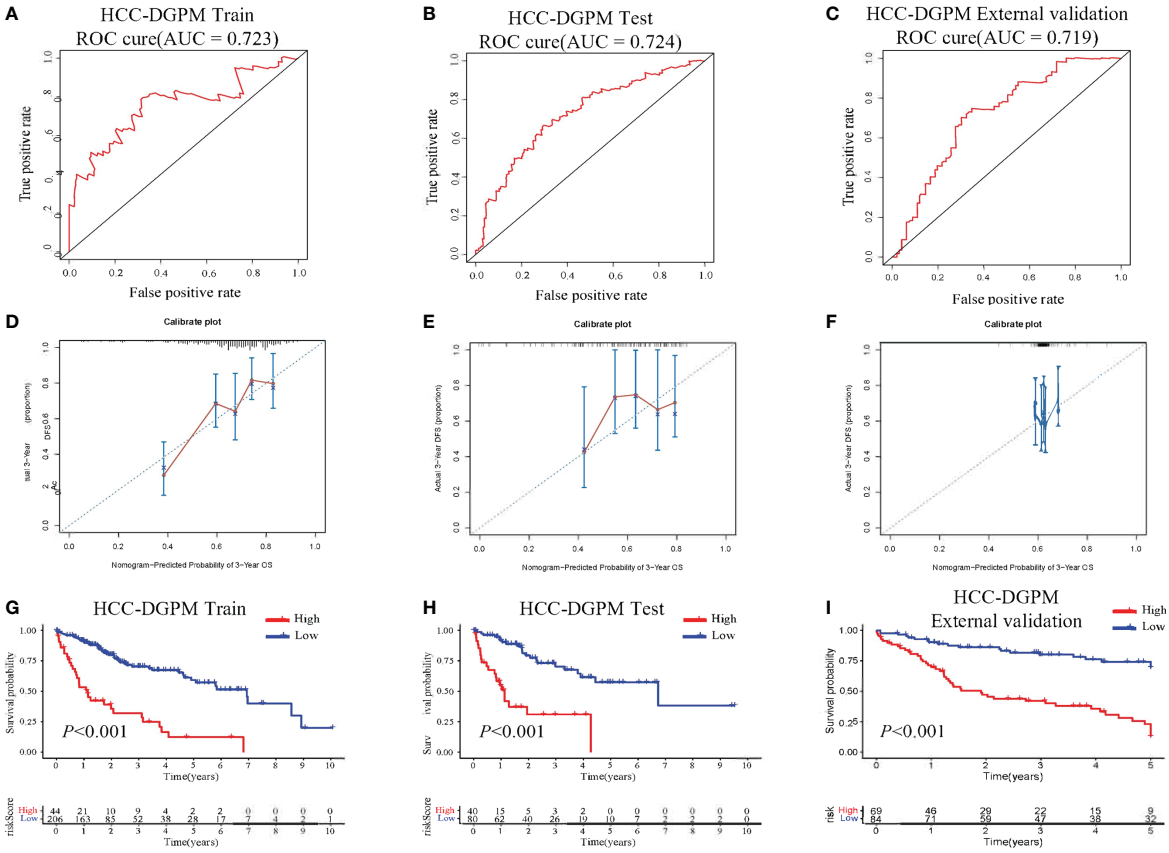

FIGURE 3
Performance Evaluation of the HCC-DGPM. **(A)** ROC curve of the training dataset. **(B)** ROC curve of the testing set. **(C)** ROC curve of the external validation set. The x-axis represents the false-positive rate, while the y-axis represents the true-positive rate. **(D–F)** Calibration curves for the training, testing, and external validation sets, respectively. **(G–I)** KM curves for overall survival (OS) in the training, testing and external validation set sets respectively (red: High risk; blue: Low risk).

TABLE 2 Risk factors in the SEER database.

risk factors	HR	HR.95L	HR.95H	*P value*
Age	1.5079	1.2375	1.8374	4.63e-05
T	2.9376	2.5259	3.4165	2e-16
N	0.8721	0.7641	0.9954	0.0425
M	3.0453	2.4834	3.7343	2e-16

cellular processes related to the cell cycle and DNA replication, indicating a more aggressive tumor phenotype compared to low-risk patients (Figures 6A–C). Moreover, the high-risk group exhibited a closer association with immune response compared to the low-risk group. This was evident from the enrichment of gene sets related to the toll-like receptor signaling pathway, cytokine-cytokine receptor interaction, and chemokine signaling pathway. These findings highlight a significant correlation between the risk score and the immune status of HCC (Figures 6D–F).

3.8 Immune assessment of the HCC-MLPM

Using the single-sample Gene Set Enrichment Analysis (ssGSEA) method, we conducted an analysis of immune cell infiltration in HCC patients, comparing the high-risk and low-risk groups. Violin plots further illustrated significantly lower infiltration levels of activated B cells, activated CD8+ T cells, natural killer cells, immature B cells, mast cells, and memory CD4+ T cells in the high-risk group. Conversely, the infiltration level of activated CD4+ T cells was significantly higher in the high-risk group (Figure 7A). Additionally, we examined the expression changes of immune checkpoint markers between the high-risk and low-risk groups. Remarkably, the high-risk group exhibited a significant upregulation in the expression levels of most immune checkpoint markers (Figure 7B). These findings indicate that the high-risk group of HCC patients displays lower levels of immune cell infiltration, particularly in specific immune cell subsets, along with higher expression of immune checkpoint markers. These observations suggest the presence of a potentially

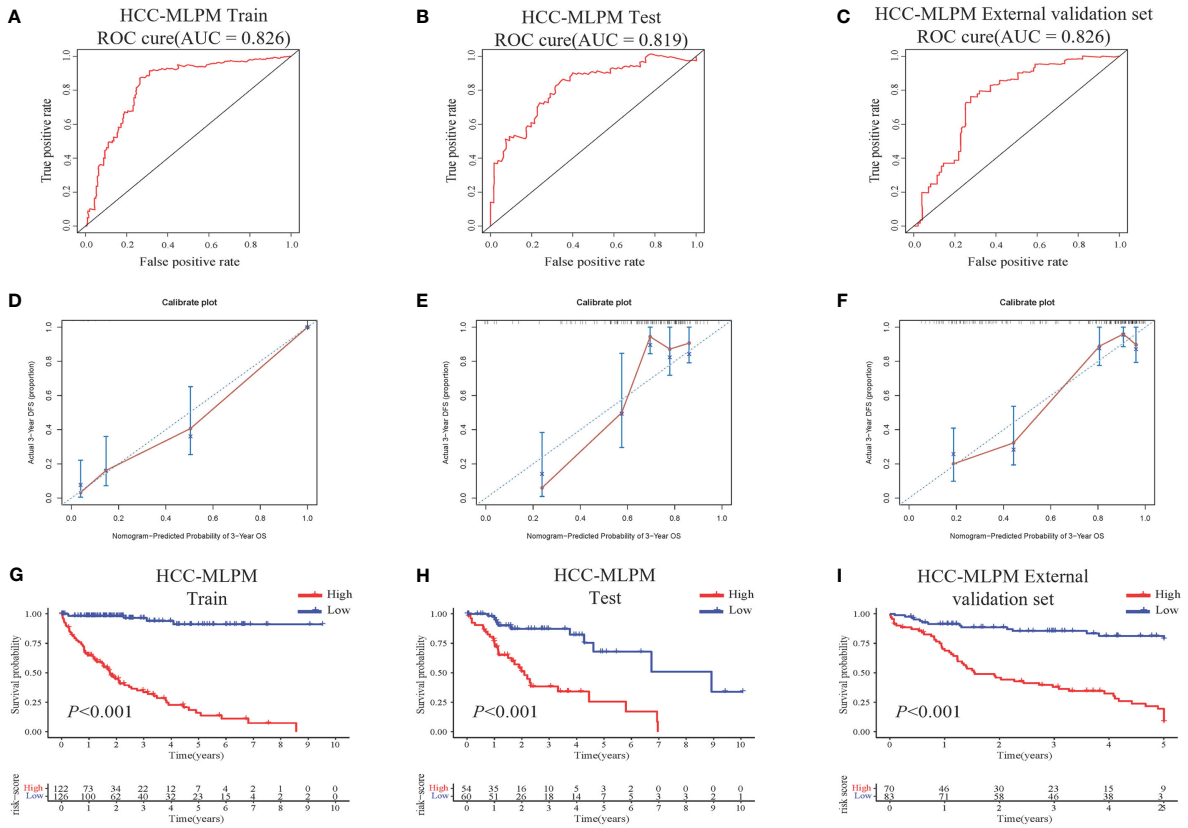

FIGURE 4
Performance Evaluation of the HCC-MLPM. **(A)** ROC curve of the training dataset. **(B)** ROC curve of the testing set. **(C)** ROC curve of the external validation set. **(D–F)** Calibration curves for the training, testing, and external validation sets, respectively. **(G–I)** KM survival curves for OS of the training, testing and external dataset respectively (red: High risk; bule: Low risk).

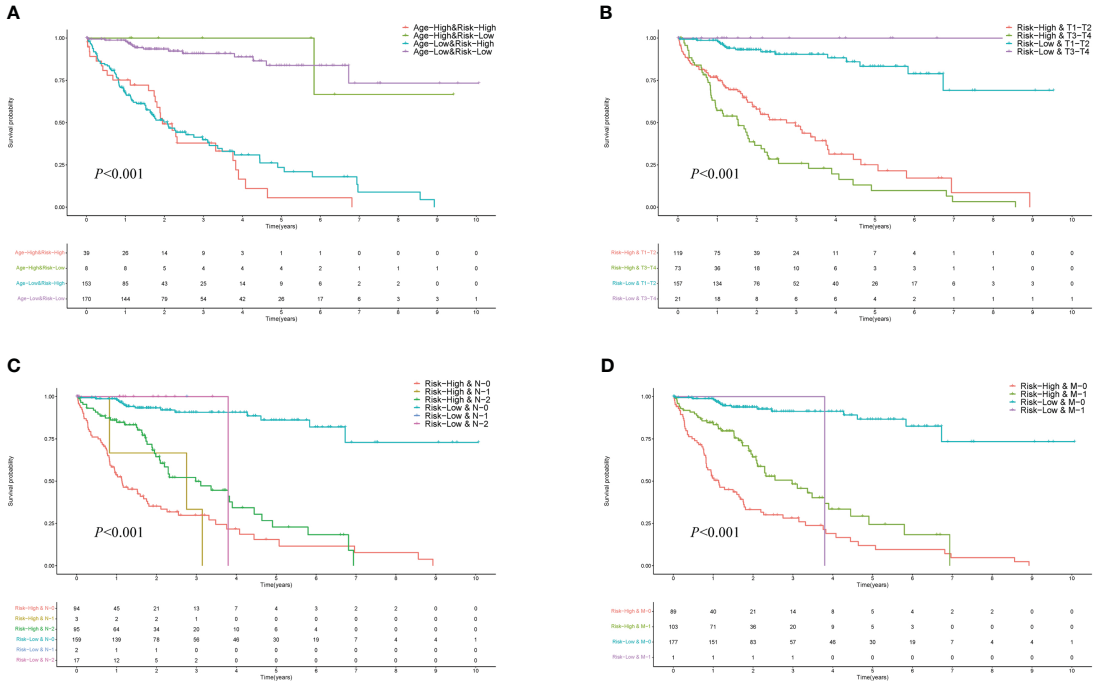

FIGURE 5
Performance for Different Clinical Indicators in HCC-MLPM. **(A)** KM curves for Age. **(B)** KM curves for tumor size (T). **(C)** KM curves for node involvement (N). **(D)** KM curves for metastasis (M).

immunosuppressive microenvironment in the high-risk group, which may contribute to disease progression and poorer prognosis.

To assess the responsiveness of patients in both groups to immunotherapy, we utilized an unsupervised clustering approach based on the characteristics of the tumor microenvironment, specifically the IPS. The patients were categorized into four groups: immune-enriched/fibrotic (IE/F), immune-enriched (IE), fibrotic (F), and immune-depleted (D). Among these groups, IE/F and IE demonstrated a more favorable response to immunotherapy, while F and D were associated with relatively poorer responses. In our analysis, we observed a higher proportion of patients in the low-risk group with an IE/F microenvironment. However, the proportions of IE and F were comparable between the two patient groups (IPS) (Figure 7C).

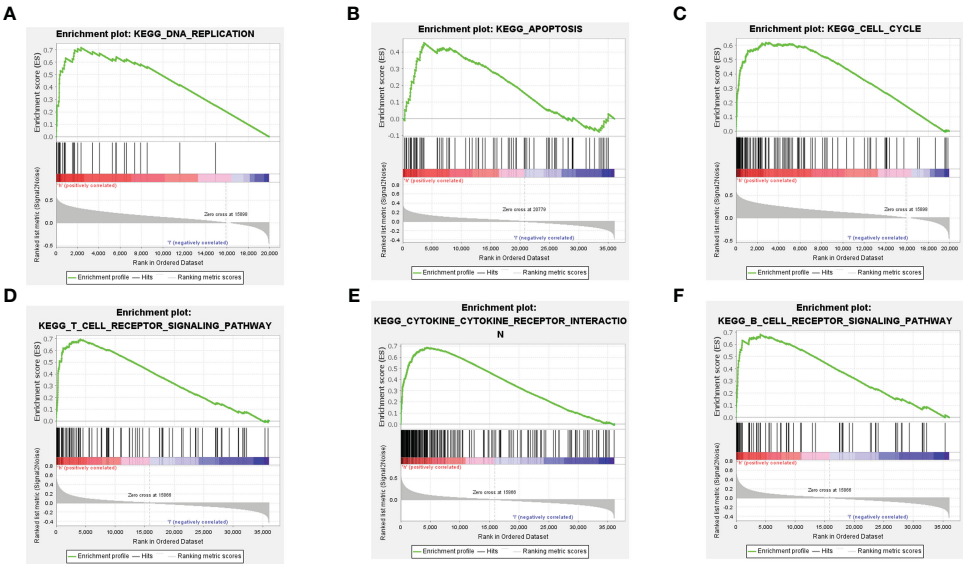

FIGURE 6
Gene Set Enrichment Analysis (GSEA) of the HCC-MLPM. **(A–C)** DNA replication, apoptosis and cell cycle enrichment analysis by GSEA. **(D-F)** T and B cell receptor pathway, cytokine receptor interaction pathway enrichment analysis by GSEA.

10.3389/fonc.2023.1327147

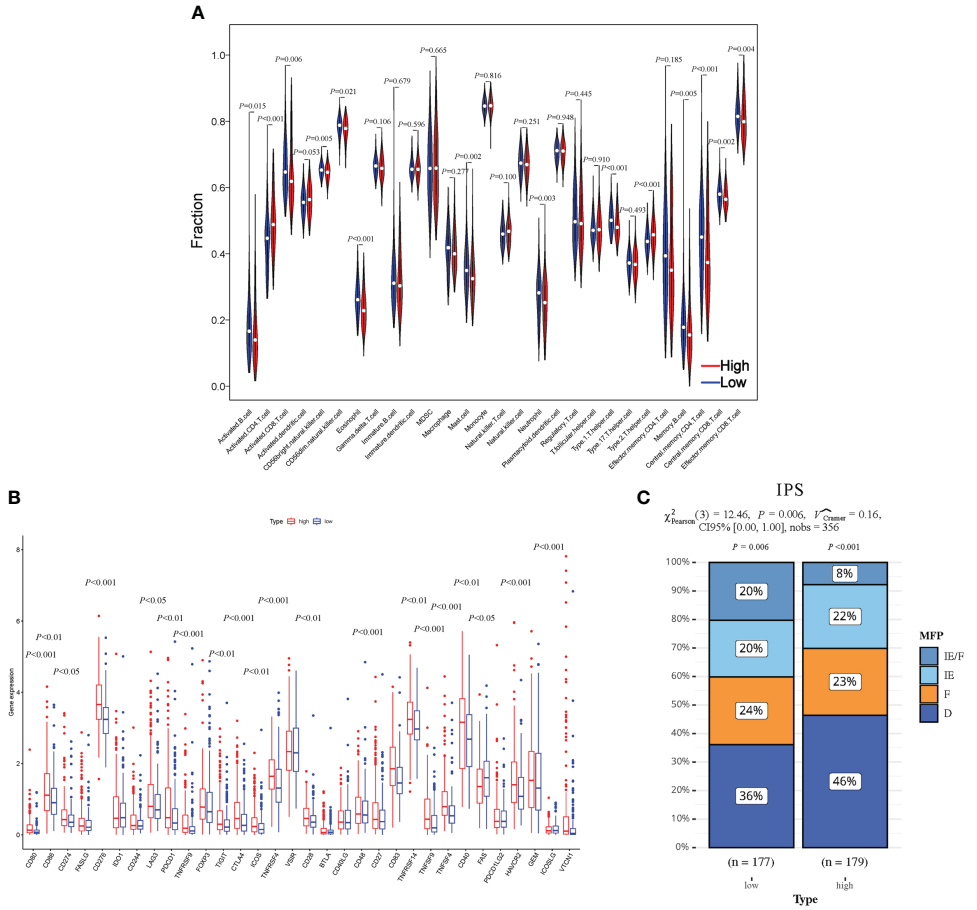

FIGURE 7
Evaluation of the HCC-MLPM on tumor immunity. **(A)** The Violin plots for infiltration levels of specific immune cell types. **(B)** Boxplot for infiltration levels of immune checkpoint markers. **(C)** The performance for immunotherapy. IE/F, Immune-enriched/Fibrotic; IE, Immune-enriched; F, Fibrotic; D, Depleted; High, high risk; Low, low risk. (red: High risk; bule: Low risk).

3.9 Assessment of clinical drug responsiveness

To investigate the clinical drug responsiveness predicted by HCC-MLPM, we utilized drug information obtained from the TCGA database. We examined the correlation between drug targets (Sorafenib) and cytotoxic drugs with the risk scores derived from our model. The results revealed significant differences in drug responsiveness between the high-risk and low-risk groups. Notably, a higher proportion of patients in the low-risk group demonstrated disease stability when treated with both targeted therapy and chemotherapy regimens (Figures 8A, B). This finding suggests that these treatment approaches were more likely to be effective in the low-risk group, providing potential therapeutic options for this particular subgroup of patients.

4 Discussion

We have successfully developed a machine learning-based prognostic risk model specifically designed for patients with HCC. This study integrates diverse indicators from a multicenter dataset, providing a comprehensive tool to aid in personalized treatment decisions for HCC patients. Remarkably, this model accurately predicts patient survival outcomes and offers insights into the effectiveness of immunotherapy and other clinical drug treatments in HCC patients.

The molecular pathogenesis of HCC is highly complex and heterogeneous. Currently, clinical treatment decisions for HCC patients are primarily based on limited clinical and pathological indicators (32). However, the development and progression of HCC are influenced by various factors, such as genetic variations, aberrant cell signaling pathways, and alterations in the immune environment (33, 34). Therefore, there is an urgent need to develop a dependable prognostic risk assessment model to enable personalized treatment for HCC. Previous studies have predominantly focused on individual tumor indicators, including age, stage, pathological type, tumor size, and other clinical factors, which have been extensively utilized in clinical assessments (35, 36). With advancements in genomics and immunology, there is a growing emphasis on the molecular characteristics of HCC and the impact of the immune environment on patient prognosis. Researchers have made efforts to predict patient prognosis from a biological perspective by integrating diverse information sources, such as gene expression data, protein expression

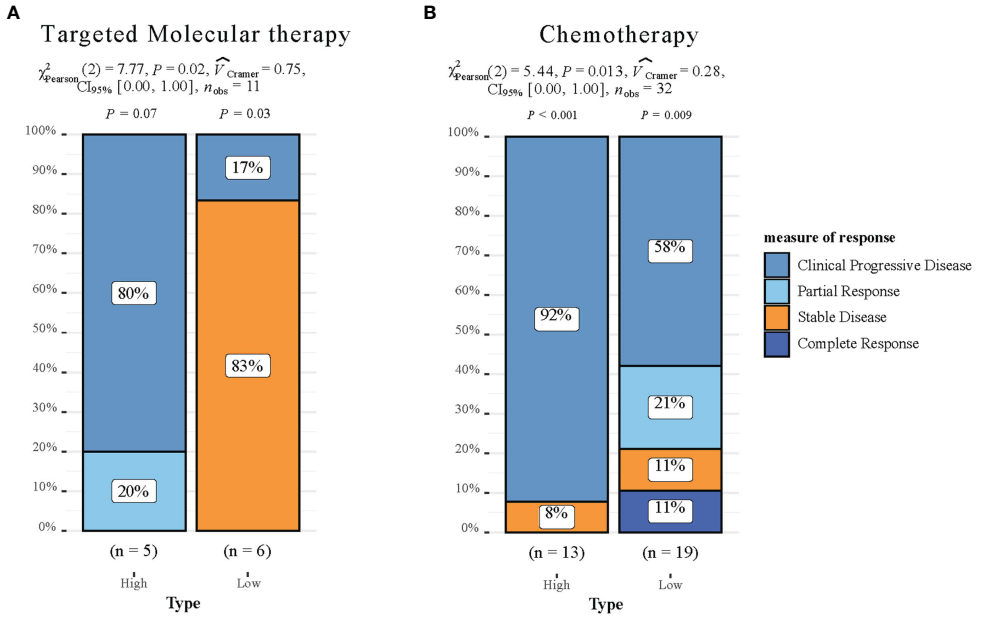

FIGURE 8
Drug responsiveness Evaluation of the HCC-MLPM. **(A)** Performance for predicting targeted molecular therapy of HCC-MLPM. **(B)** Performance for chemotherapy of HCC-MLPM. (High: high risk; Low: low risk).

data, and immune cell infiltration (37, 38). However, these evaluation methods provide a limited perspective on patient prognosis, resulting in inherent limitations. In contrast, our study not only takes into account clinical and pathological indicators that reflect the overall patient condition and disease severity but also places significant emphasis on the biological characteristics of liver cell tumors. This approach involves the identification of differentially expressed genes in HCC and the unveiling of potential biological mechanisms. By utilizing a modeling approach that incorporates comprehensive multi-level indicators, our model can offer a more comprehensive and dependable prognostic risk assessment for HCC patients (39, 40). Similar methodologies have demonstrated favorable outcomes in studies focusing on diverse cancer types, underscoring their potential utility in personalized medicine. Specifically, researchers investigating breast cancer, bladder cancer, and colorectal cancer have achieved robust predictive results by integrating comprehensive models with diverse datasets encompassing clinical, gene expression, and proteomic information (41–43). These investigations additionally validate the feasibility of our modeling approach.

In terms of methodology, our research has made significant advancements. Cox regression, an advanced machine learning technique, has provided strong technical support. Machine learning, in comparison to traditional statistical methods, excels in managing complex data structures and relationships, facilitating the extraction of potential features and patterns from extensive clinical and gene expression data (41). By employing the feature selection and optimization process of Cox regression, we have identified the most relevant indicators for prognosticating HCC patients. This approach effectively reduces the dimensionality of the feature space and enhances the predictive performance of the model (AUC = 0.724 *vs.* 0.819). Importantly, our research leverages the generalizability of machine learning, enabling the evaluation of the model's predictive performance and reliability across diverse datasets from multiple centers, including TCGA, SEER, and ICGC. The comprehensive integration of biological factors, clinical and pathological features, and multi-level indicators in our model substantially enhances its capacity to capture the intricacies of patient survival in HCC.

Our research findings include 11 risk factors, including 7 identified from the TCGA dataset (MYBL2, SF3B4, CDCA8, NUF2, HMMR, P0N1, and PAGE1), and an additional 4 acquired from the SEER database (Age, T, N, M). Several studies have recognized the substantial impact of certain factors on patient outcomes in HCC (44, 45). Our stratified survival analysis revealed that age plays a critical role in determining survival rates, aligning with previous findings. Furthermore, the correlation observed between smaller tumor size and better prognoses in our study underscores the importance of early detection and diagnosis in HCC, as reflected in the Kaplan-Meier curves for T. N and M, which indicate the aggressive progression of HCC, were also found to significantly affect survival rates in our analysis. These findings advocate for a nuanced understanding of HCC, indicating the inadequacy of generic treatment strategies. Moreover, our study validated these risk factors against the National Comprehensive Cancer Network (NCCN) guidelines (11), confirming the model's emphasis on tumor staging and its reliability.

Apart from established clinical factors, our research innovatively identified 7 genes that influence prognosis. These genes play critical roles in regulating cell-cell interactions, extracellular matrix remodeling, angiogenesis, and inflammatory responses within the tumor microenvironment. For example,

MYBL2 plays an important role in regulating the cell cycle, as its high expression is correlated with the staging and grading of various cancers (46). SF3B4 is involved in regulating the cell cycle, cell differentiation, and immune deficiency. Mutations in SF3B4 can lead to abnormal cell growth and contribute to disease development (47). CDCA8 controls the process of cell mitosis and has been identified as an unfavorable prognostic predictor in liver cancer (48). NUF2 participates in chromosome segregation and has been positively correlated with differential immune cell infiltration and various immune biomarkers (49). HMMR is associated with the infiltration levels of neutrophils, CD8+ T cells, and CD4+ T cells in the immune system, as well as the prognosis of patients with cancer (50). PON1 plays a role in cell adhesion and migration, contributing to the regulation of tumor development, oxidative stress, and inflammatory responses (51). PAGE1 is involved in cell apoptosis and immune regulation (52). These gene abnormalities play a role in altering the tumor microenvironment, which impacts the growth, infiltration, and metastasis of HCC. By integrating these differential gene factors into the prognostic risk assessment model, we capture the intricacies of patient survival in HCC.

In addition to assessing the model's performance in predicting patient survival, our research closely integrates with clinical treatment through the evaluation of patients' immune infiltration status and their responses to clinical drugs. This enhances our comprehensive understanding of HCC. In recent years, immunotherapy has emerged as a significant breakthrough in HCC treatment, utilizing the patient's immune system to target tumor cells (53, 54). We examined the association between the model's predictive results and the immune status by employing GSEA analysis and assessing immune cell infiltration. The results indicated a significant correlation ($P < 0.05$) between high-risk patients and malignant tumor phenotypes, particularly in terms of cell cycle, DNA replication, and immune responses. We have discerned a significant elevation in the infiltration levels of Type2 T helper (Th2) cells within the cohort of high-risk HCC patients ($P < 0.001$), indicating a Th2-dominated immune microenvironment. The cytokines secreted by Th2 cells, such as IL-4 and IL-10, may facilitate tumor growth and assist in the tumor's evasion of immune surveillance. Therapeutic interventions targeting the Th2 cell pathway, such as PD-1/PD-L1 and CTLA-4 inhibitors, have demonstrated potential in the treatment of other cancers (55, 56). This observation underscores the importance of considering the immune microenvironment when devising therapeutic strategies for HCC.

Additionally, we introduced the novel IPS to assess both the immune system's activity level and the extent of immune cell infiltration in the tumor microenvironment. This was done with the aim of identifying potential variations in patient response to immunotherapy. The IPS quantifies patients' potential responsiveness to immunotherapy based on the analysis of expression patterns in immune-related genes. Higher IPS scores generally reflect a more active immune system and an increased likelihood of positive response to immunotherapy (31). We computed IPS scores for patients in both the high-risk and low-risk groups, facilitating a comparison of their immunotherapy responsiveness. The findings showed that the low-risk group exhibited significantly higher responsiveness to immunotherapy ($P < 0.05$), providing theoretical support for the application of immunotherapy in low-risk patients.

Although our research has shown promising results, it is important to acknowledge its limitations. First, the development and prognosis of HCC are influenced by various biological and environmental factors. While we thoroughly considered clinical data and genetic information, it is conceivable that other factors, not accounted for in the model, may also contribute. This underscores the necessity for continual improvement and refinement. Secondly, as our model lacks support from Supplementary Databases, it is advisable to conduct further prospective studies to validate and refine it in relation to immunotherapy and clinical drug responsiveness.

In conclusion, we have successfully developed a machine learning-based prognostic risk model for HCC, providing robust support for personalized treatment strategies in HCC patients. Furthermore, this study highlights the potential importance of utilizing multi-level modeling approaches in the realm of personalized medicine.

Data availability statement

The data utilized in this study can be obtained by contacting the authors due to restrictions imposed by the data providers, namely the TCGA, SEER, and GSEA databases. Access to these databases is available via their dedicated websites: TCGA (https://portal.gdc.cancer.gov/), SEER (https://seer.cancer.gov/), GSEA (http://software.broadinstitute.org/gsea/index.jsp), and ICGC (https://icgcportal.genomics.cn/). Researchers interested in accessing the data may reach out to the authors for additional information and support in acquiring the required permissions and data access.

Author contributions

Z-HZ: Data curation, Methodology, Software, Writing – original draft. YD: Supervision, Validation, Writing – review & editing. SW: Investigation, Writing – original draft. WP: Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

Authors YD, SW and WP were employed by the company China RongTong Medical Healthcare Group Co., Ltd.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of hepatocellular carcinoma. *Hepatology* (2021) 73 (Suppl 1):4–13. doi: 10.1002/hep.31288

2. Samant H, Amiri HS, Zibari GB. Addressing the worldwide hepatocellular carcinoma: epidemiology, prevention and management. *J Gastrointest Oncol* (2021) 12 Suppl 2:S361–73. doi: 10.21037/jgo.2020.02.08

3. Yao J, Liang X, Liu Y, Li S, Zheng M. Trends in incidence and prognostic factors of two subtypes of primary liver cancers: A surveillance, epidemiology, and end results-based population study. *Cancer Control.* (2022) 29:10732748211051548. doi: 10.1177/10732748211051548

4. Moon AM, Singal AG, Tapper EB. Contemporary epidemiology of chronic liver disease and cirrhosis. *Clin Gastroenterol Hepatol* (2020) 18:2650–66. doi: 10.1016/j.cgh.2019.07.060

5. European Association for the Study of the Liver. Electronic address: easloffice@easloffice.eu and European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *J Hepatol* (2018) 69:182–236. doi: 10.1016/j.jhep.2018.03.019

6. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers.* (2021) 7:6. doi: 10.1038/s41572-020-00240-3

7. Marrero JA, Kulik LM, Sirlin CB, Zhu AX, Finn RS, Abecassis MM, et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the american association for the study of liver diseases. *Hepatology* (2018) 68:723–50. doi: 10.1002/hep.29913

8. Ayuso C, Rimola J, Vilana R, Burrel M, Darnell A, García-Criado Á, et al. Diagnosis and staging of hepatocellular carcinoma (HCC): current guidelines. *Eur J Radiol* (2018) 101:72–81. doi: 10.1016/j.ejrad.2018.01.025

9. Allaire M, Goumard C, Lim C, Le Cleach A, Wagner M, Scatton O. New frontiers in liver resection for hepatocellular carcinoma. *JHEP Rep* (2020) 2:100134. doi: 10.1016/j.jhepr.2020.100134

10. Kabir T, Tan ZZ, Syn NL, Wu E, Lin JD, Zhao JJ, et al. Laparoscopic versus open resection of hepatocellular carcinoma in patients with cirrhosis: meta-analysis. *Br J Surg* (2021) 109:21–9. doi: 10.1093/bjs/znab376

11. Benson AB, D'Angelica MI, Abbott DE, Anaya DA, Anders R, Are C, et al. Hepatobiliary cancers, version 2.2021, NCCN clinical practice guidelines in oncology. *J Natl Compr Cancer Network* (2021) 19:541–65. doi: 10.6004/jnccn.2021.0022

12. Shimose S, Iwamoto H, Tanaka M, Niizeki T, Shirono T, Kajiwara A, et al. Multimolecular-targeted agents for intermediate-stage hepatocellular carcinoma influence time to stage progression and overall survival. *Oncology* (2021) 99:756–65. doi: 10.1159/000518612

13. D'Angelo S, Secondulfo M, De Cristofano R, Sorrentino P. Sorafenib and entecavir: the dioscuri of treatment for advanced hepatocellular carcinoma? *World J Gastroenterol* (2013) 19:2141–3. doi: 10.3748/wjg.v19.i14.2141

14. Shen X, Li N, Li H, Zhang T, Wang F, Li Q. Increased prevalence of regulatory T cells in the tumor microenvironment and its correlation with TNM stage of hepatocellular carcinoma. *J Cancer Res Clin Oncol* (2010) 136:1745–54. doi: 10.1007/s00432-010-0833-8

15. Lim H, Ramjeesingh R, Liu D, Tam VC, Knox JJ, Card PB, et al. Optimizing survival and the changing landscape of targeted therapy for intermediate and advanced hepatocellular carcinoma: A systematic review. *J Natl Cancer Inst* (2021) 113:123–36. doi: 10.1093/jnci/djaa119

16. Foerster F, Gairing SJ, Müller L, Galle PR. NAFLD-driven HCC: Safety and efficacy of current and emerging treatment options. *J Hepatol* (2022) 76:446–57. doi: 10.1016/j.jhep.2021.09.007

17. Couri T, Pillai A. Goals and targets for personalized therapy for HCC. *Hepatol Int* (2019) 13:125–37. doi: 10.1007/s12072-018-9919-1

18. Lee S-W, Lee T-Y, Peng Y-C, Yang S-S, Yeh H-Z, Chang C-S. Sorafenib treatment on Chinese patients with advanced hepatocellular carcinoma: A study on prognostic factors of the viral and tumor status. *Med (Baltimore).* (2019) 98:e17692. doi: 10.1097/MD.0000000000017692

19. Faivre S, Rimassa L, Finn RS. Molecular therapies for HCC: Looking outside the box. *J Hepatol* (2020) 72:342–52. doi: 10.1016/j.jhep.2019.09.010

20. Finkelmeier F, Waidmann O, Trojan J. Nivolumab for the treatment of hepatocellular carcinoma. *Expert Rev Anticancer Ther* (2018) 18:1169–75. doi: 10.1080/14737140.2018.1535315

21. Paik J. Nivolumab plus relatlimab: first approval. *Drugs* (2022) 82:925–31. doi: 10.1007/s40265-022-01723-1

22. Sidali S, Trépo E, Sutter O, Nault J-C. New concepts in the treatment of hepatocellular carcinoma. *United Eur Gastroenterol J* (2022) 10:765–74. doi: 10.1002/ueg2.12286

23. Guiu B, Garin E, Allimant C, Edeline J, Salem R. TARE in hepatocellular carcinoma: from the right to the left of BCLC. *Cardiovasc Intervent Radiol* (2022) 45:1599–607. doi: 10.1007/s00270-022-03072-8

24. Liñares-Blanco J, Pazos A, Fernandez-Lozano C. Machine learning analysis of TCGA cancer data. *PeerJ Comput Sci* (2021) 7:e584. doi: 10.7717/peerj-cs.584

25. Liu J, Sun G, Pan S, Qin M, Ouyang R, Li Z, et al. The Cancer Genome Atlas (TCGA) based m6A methylation-related genes predict prognosis in hepatocellular carcinoma. *Bioengineered* (2020) 11:759–68. doi: 10.1080/21655979.2020.1787764

26. Donisi C, Puzzoni M, Ziranu P, Lai E, Mariani S, Saba G, et al. Immune checkpoint inhibitors in the treatment of HCC. *Front Oncol* (2020) 10:601240. doi: 10.3389/fonc.2020.601240

27. Llovet JM, Montal R, Sia D, Finn RS. Molecular therapies and precision medicine for hepatocellular carcinoma. *Nat Rev Clin Oncol* (2018) 15:599–616. doi: 10.1038/s41571-018-0073-4

28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* (2015) 43:e47. doi: 10.1093/nar/gkv007

29. Kuhn M. Building predictive models in R using the caret package. *J Stat Software* (2008) 28:1–26. doi: 10.18637/jss.v028.i05

30. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7

31. Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, Isaeva O, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* (2021) 39:845–865.e7. doi: 10.1016/j.ccell.2021.04.014

32. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado Á, et al. BCLC strategy for prognosis prediction and treatment recommendation: The 2022 update. *J Hepatol* (2022) 76:681–93. doi: 10.1016/j.jhep.2021.11.018

33. Dhanasekaran R, Bandoh S, Roberts LR. Molecular pathogenesis of hepatocellular carcinoma and impact of therapeutic advances. *F1000Res* (2016) 5:F1000 Faculty Rev–879. doi: 10.12688/f1000research.6946.1

34. Macdonald GA. Pathogenesis of hepatocellular carcinoma. *Clin Liver Dis* (2001) 5:69–85. doi: 10.1016/S1089-3261(05)70154-9

35. Ma L, Deng K, Zhang C, Li H, Luo Y, Yang Y, et al. Nomograms for predicting hepatocellular carcinoma recurrence and overall postoperative patient survival. *Front Oncol* (2022) 12:843589. doi: 10.3389/fonc.2022.843589

36. Zhang H, Du X, Dong H, Xu W, Zhou P, Liu S, et al. Risk factors and predictive nomograms for early death of patients with advanced hepatocellular carcinoma: a large retrospective study based on the SEER database. *BMC Gastroenterol* (2022) 22:348. doi: 10.1186/s12876-022-02424-5

37. Hu B, Yang X-B, Sang X-T. Molecular subtypes based on immune-related genes predict the prognosis for hepatocellular carcinoma patients. *Int Immunopharmacol.* (2021) 90:107164. doi: 10.1016/j.intimp.2020.107164

38. Song X, Du R, Gui H, Zhou M, Zhong W, Mao C, et al. Identification of potential hub genes related to the progression and prognosis of hepatocellular carcinoma through integrated bioinformatics analysis. *Oncol Rep* (2020) 43:133–46. doi: 10.3892/or.2019.7400

39. Yang B, Zhang Y, Pang S, Shang X, Zhao X, Han M. Integrating multi-omic data with deep subspace fusion clustering for cancer subtype prediction. *IEEE/ACM Trans Comput Biol Bioinform* (2021) 18:216–26. doi: 10.1109/TCBB.2019.2951413

40. Vangimalla RR, Sreevalsan-Nair J. HCNM: heterogeneous correlation network model for multi-level integrative study of multi-omics data for cancer subtype prediction. *Annu Int Conf IEEE Eng Med Biol Soc* (2021) 2021:1880–6. doi: 10.1109/EMBC46164.2021.9630781

41. Pou SA, Díaz M del P, Osella AR. Applying multilevel model to the relationship of dietary patterns and colorectal cancer: an ongoing case-control study in Córdoba, Argentina. *Eur J Nutr* (2012) 51:755–64. doi: 10.1007/s00394-011-0255-7

42. Hiatt RA, Porco TC, Liu F, Balke K, Balmain A, Barlow J, et al. A multilevel model of postmenopausal breast cancer incidence. *Cancer Epidemiol Biomarkers Prev* (2014) 23:2078–92. doi: 10.1158/1055-9965.EPI-14-0403

43. Peng C, Li A, Wang M. Discovery of bladder cancer-related genes using integrative heterogeneous network modeling of multi-omics data. *Sci Rep* (2017) 7:15639. doi: 10.1038/s41598-017-15890-9

44. Liu M, Xu M, Tang T. Association between chemotherapy and prognostic factors of survival in hepatocellular carcinoma: a SEER population-based cohort study. *Sci Rep* (2021) 11:23754. doi: 10.1038/s41598-021-02698-x

45. Ding J, Wen Z. Survival improvement and prognosis for hepatocellular carcinoma: analysis of the SEER database. *BMC Cancer*. (2021) 21:1157. doi: 10.1186/s12885-021-08904-3

46. Chen X, Lu Y, Yu H, Du K, Zhang Y, Nan Y, et al. Pan-cancer analysis indicates that MYBL2 is associated with the prognosis and immunotherapy of multiple cancers as an oncogene. *Cell Cycle* (2021) 20:2291–308. doi: 10.1080/15384101.2021.1982494

47. Yan L, Yang X, Yang X, Yuan X, Wei L, Si Y, et al. The role of splicing factor SF3B4 in congenital diseases and tumors. *Discovery Med* (2021) 32:123–32.

48. Shuai Y, Fan E, Zhong Q, Chen Q, Feng G, Gou X, et al. CDCA8 as an independent predictor for a poor prognosis in liver cancer. *Cancer Cell Int* (2021) 21:159. doi: 10.1186/s12935-021-01850-x

49. Xie X, Jiang S, Li X. Nuf2 is a prognostic-related biomarker and correlated with immune infiltrates in hepatocellular carcinoma. *Front Oncol* (2021) 11:621373. doi: 10.3389/fonc.2021.621373

50. Ma X, Xie M, Xue Z, Yao J, Wang Y, Xue X, et al. HMMR associates with immune infiltrates and acts as a prognostic biomaker in lung adenocarcinoma. *Comput Biol Med* (2022) 151 Pt A:106213. doi: 10.1016/j.compbiomed.2022.106213

51. Cai D, Zhao Z, Hu J, Dai X, Zhong G, Gong J, et al. Identification of the tumor immune microenvironment and therapeutic biomarkers by a novel molecular subtype based on aging-related genes in hepatocellular carcinoma. *Front Surg* (2022) 9:836080. doi: 10.3389/fsurg.2022.836080

52. Cui Y, Jiang N. Identification of a seven-gene signature predicting clinical outcome of liver cancer based on tumor mutational burden. *Hum Cell* (2022) 35:1192–206. doi: 10.1007/s13577-022-00708-2

53. Rodríguez Pérez Á, Campillo-Davo D, Van Tendeloo VFI, Benítez-Ribas D. Cellular immunotherapy: a clinical state-of-the-art of a new paradigm for cancer treatment. *Clin Transl Oncol* (2020) 22:1923–37. doi: 10.1007/s12094-020-02344-4

54. Abbott M, Ustoyev Y. Cancer and the immune system: the history and background of immunotherapy. *Semin Oncol Nurs*. (2019) 35:150923. doi: 10.1016/j.soncn.2019.08.002

55. Zhang H, Dai Z, Wu W, Wang Z, Zhang N, Zhang L, et al. Regulatory mechanisms of immune checkpoints PD-L1 and CTLA-4 in cancer. *J Exp Clin Cancer Res* (2021) 40:184. doi: 10.1186/s13046-021-01987-7

56. Pandey P, Khan F, Qari HA, Upadhyay TK, Alkhateeb AF, Oves M. Revolutionization in cancer therapeutics *via* targeting major immune checkpoints PD-1, PD-L1 and CTLA-4. *Pharmaceuticals* (2022) 15:335. doi: 10.3390/ph15030335