

Identification of a novel signature based on mTORC1 pathway for hepatocellular carcinoma.

Zhuomao Mo

Department of Traditional Chinese Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong province, 510080, China

Shaoju Luo

Department of Traditional Chinese Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong province, 510080, China

Zhirui Cao

Department of Traditional Chinese Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong province, 510080, China

Hao Hu

Department of Traditional Chinese Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong province, 510080, China

Ling Yu

Department of Traditional Chinese Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, Guangdong province, 510080, China

Shijun Zhang (✉ 2806973376@qq.com)

First Affiliated Hospital of Sun Yat-sen University <https://orcid.org/0000-0001-5251-0115>

Research

Keywords: hepatocellular carcinoma, mTORC1, signature, TCGA.

Posted Date: February 17th, 2020

DOI: <https://doi.org/10.21203/rs.2.23745/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background mTORC1 signal pathway play a role in the initiation and progression of hepatocellular carcinoma (HCC), but no relevant gene signature was developed. This research aimed to explore the potential correlation between mTORC1 signal pathway and HCC and establish the related genes signature.

Methods HCC cases were retrieved from The Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC) and Gene Expression Omnibus (GEO) databases. The genes to be included in mTORC1-associated signature were selected by performing univariate, multivariate Cox regression analysis and lasso regression analysis. Then, the signature was verified by survival analysis and multiple receiver operating characteristic (ROC) curve. Furthermore, a nomogram was established and evaluated by C-index, calibration plot and ROC curve.

Results The signature was established with the six genes (ETF1, GSR, SKAP2, HSPD1, CACYBP and PNP). Under the grouping from signature, patients in the high- risk group showed worse survival than those in the low-risk group in both three datasets. The univariate and multivariate Cox regression analysis indicated that mTORC1 related signature can be the potential independent prognostic factor in HCC.

Conclusion The mTORC1 associated gene signature established and validated in our research could be used as a potential prognostic factor in HCC.

1. Background

Hepatocellular carcinoma (HCC) is one of the most popular cancer around the world, becoming the second leading causes of tumor related death [1]. Due to the high rate of metastasis, HCC patients with advanced stage are usually with a poor prognosis [2]. As the treatment and biomarkers of HCC have developed, the clinical outcomes of HCC patients are still unsatisfactory [3]. The occurrence and development of HCC involved interactions between genetics, epigenetics and transcriptomic alterations [4]. Many studies have verified that different genes have an effect on predicting prognosis of HCC. For example, Gu et al found that *CCL14* was a potential prognostic biomarker that correlated with tumor immune cell infiltration in HCC [5]. Another study found the strong correlations between *PRPF3* expression and prognosis in HCC [6]. However, as a biomarker, a single gene usually has a lower prognostic performance than multigene prognostic signature. Therefore, many gene related signature have been developed for predicting prognosis of HCC. For instance, Zhang et al established a gene signature associated with HCC microenvironment and successfully verified it [7]. Other predictive signatures based on immune [8] and glycolysis [9] also play an important role in prognosis of HCC.

Generally, the gene researches usually focus on comparing the gene expression between two groups, or pay attention to the highly upregulated and downregulated genes. Nevertheless, some genes which showed no significant difference but had important biological function and characteristics were omitted. In view of this, a computational method, Gene Set Enrichment Analysis (GSEA) determined whether a

prior defined set of genes shows statistically significant differences between two biological states [10]. The advantage of GSEA is that it can identify the genes which expression is based on the trend of overall level. Consequently, in this research, we identified the pathway and gene with GSEA. Then, we constructed the signature based on related genes and verified it, providing the more comprehensive and accurate prognostic model for clinic.

2. Materials And Methods

2.1 Data collection

The gene expression data with the type of level 3 RNA-seq FPKM dataset and the clinical messages in TCGA website (<https://portal.gdc.cancer.gov/>) were retrieved. A total of 377 HCC cases have been downloaded and analyzed, which included 50 normal samples and 327 tumor samples. As the cohorts of validation, GSE76427 datasets were retrieved from Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>), while ICGC-LIRI dataset was retrieved from International Cancer Genome Consortium (ICGC) database (<https://icgc.org/>).

2.2 Identification of pathways and related genes.

After data collection, we extracted the clinical details and generated the expression matrix of all genes in HCC of TCGA dataset. Then we divided the patients into two groups according to the survival status and employed GSEA to choose the most relevant pathway. The pathways were considered for further analyses if a normalized P-value <0.01 . After that, we collected the related genes of involved pathway from Molecular Signatures Database (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>).

2.3 Construction of signature

First, we performed the differentially expressed analysis to select the related genes. The “limma” package under R studio software was employed and a P-value <0.05 was considered statistically significant. Second, we employed the univariate and multivariate Cox regression analysis to choose the prognostic genes among the differentially expressed genes. The genes in this section were eligible for further selection if a P-value <0.05 . Third, the lasso regression analysis was executed for checking selected genes. In this analysis, a lasso penalty was applied, to simultaneously accounting for shrinkage and variable selection. The optimal value of the lambda penalty parameter was defined by performing 10 cross-validations. Using the “glmnet” package, the coefficient of each included gene and risk score of each case were calculated. The calculation formula of risk score was following: . The cases were divided into two groups (high risk or low risk), according to the risk score median. To explore the time-dependent prognostic value of our gene signature, the survival analysis was performed using the “survival” package in the R studio software. The relationship between signature and other clinical messages were also evaluated and visualized with a heatmap. Besides, the multiple Receiver Operating Characteristic (ROC) curve was performed for checking the accuracy. Moreover, the univariate and multivariate Cox regression analysis were performed to verify whether the risk score is an independent prognostic factor.

2.4 Predictive nomogram design

Using the “rms” package and Cox regression model, a predictive nomogram based on age, gender, stage and risk score was constructed to predict the overall survival (OS) at 3 years and 5 years for HCC patients. Then, we used the Harrell’s concordance index (C-index), calibration plot and ROC curve to evaluate the nomogram.

3. Results

The clinical data details of the patients used in this study are shown in Table 1. Figure 1 show the screening process and validation of our study. The result of GSEA in Figure 2 showed that 6 pathways were eligible and the details of pathways were summarized in Table 2. Considering the highest normalized enrichment score of mTORC1 pathways, we chose the 199 relevant genes in this pathway for further analysis. The differentially expressed analysis showed that 160 genes were significantly different in HCC (Appendix S1). After that, the univariate and multivariate Cox regression analysis demonstrated that 15 genes (*ETF1*, *CTSC*, *GSR*, *HSPE1*, *SKAP2*, *HSPD1*, *TES*, *TFRC*, *ASNS*, *EPRS*, *CANX*, *CACYBP*, *UNG*, *TBK1* and *PNP*) were included (Appendix S2 and S3). As illustrated in Figure 3, the results of lasso regression analysis further confirmed the signature composed of 6 genes (*ETF1*, *GSR*, *SKAP2*, *HSPD1*, *CACYBP* and *PNP*). And the coefficients of *ETF1*, *GSR*, *SKAP2*, *HSPD1*, *CACYBP* and *PNP* were 0.03402, 0.00670 0.02556, 0.00181, 0.02034 and 0.00916, respectively. The survival analysis of TCGA, GSE76427 and ICGC datasets both indicated that significant difference was found between two groups in Figure 4 ($P < 0.05$).

Besides, the correlation between risk score and clinical data were shown in Figure 5. It illustrated that significant difference was found in tumor size and stage between two groups (Figure 5a). All the genes included in the signature were highly expressed in the high-risk group and lowly expressed in the low-risk group. And the multiple ROC curve demonstrated that the risk score got the highest predictability (ROC=0.802) among analyzed factors (Figure 5b). Furthermore, the results of univariate and multivariate Cox regression analysis (Figure 5c and 5d) revealed that the signature can be the independent prognostic factor in HCC ($P < 0.01$).

Finally, we built the nomogram based on established signature to predict 3-years and 5-years OS for HCC patients (Figure 6a). The C-index of nomogram was 0.73, and the calibration plot for the probability of survival at 3 years (Figure 6b) and 5 years (Figure 6c) showed good agreement between the prediction by nomogram and real observation. Moreover, the ROC curve of 3-year (AUC=0.735) and 5-year OS (AUC=0.742) showed the moderate accuracy in Figure 6d.

4. Discussions

In the initiation and progression of hepatocellular carcinoma, genetic factors usually play an important role. Meanwhile, mRNA gene signature based on a certain characteristic like glycolysis [11] and immune

[12] have been developed for prognosis of cancers. In this research, we explored specific function to identify genes by comparing survival group and death group in GSEA. According to our results, six signal pathways was found to be highly related to survival and we established the gene signature with mTORC1 signal pathway. As we all known, mTOR pathway is a serine/threonine protein kinase belonging to the PI3K-related kinase family [13], which comprised of two distinct complexes (mTORC1 and mTORC2). With Raptor as its unique and key protein component, mTORC1 plays an important role in cell survival, autophagy, and metabolism [14]. Concerning for the mTOR signal pathway in HCC, it has been found that aberrant mTOR signaling was present in half of the HCC cases [15]. Meanwhile, an intact mTORC1 axis [16] and mTORC2-Akt1 cascade [17] were required for c-Myc-driven hepatocarcinogenesis. Moreover, some researches [15, 18] provided the theoretical basis of mTOR signaling pathway-oriented targeting treatment for HCC in clinic. Overall, mTOR signal pathway plays an important role in the development and progression of HCC.

In this study, we identified six genes in signature by performing the differentially expressed analysis, univariate Cox regression analysis and lasso regression analysis. Among our included genes, five genes (*ETF1*, *GSR*, *HSPD1*, *CACYBP* and *PNP*) have been found to be related to HCC from previous studies. Singh et al found that *ETF1*, *CNOT6* and *XRN1* gene in HepG2 cell led to significant alteration in stability of specific mRNAs and this mechanism may hold novel cancer therapeutic targets [19]. In another research, McLoughlin concluded that *GSR*, *TRXR1*, *NRF2* and oxidative stress determined hepatocellular carcinoma malignancy [20]. Lee's study [21] found that *HSPD1* was down-regulated during early apoptosis of the hepatoma cell mediated by *Paeoniae Radix*. In terms of *CACYBP*, it have been verified that *CACYBP* can promote hepatocellular carcinoma progression in the absence of *RNF41* mediated degradation [22]. Moreover, a study [23] found that *PNP*/fludarabine suicide gene system induced HCC cell apoptosis and inhibited the growth of HCC cells. Although we found no evidence supporting the correlation between *SKAP2* and HCC, it has been verified that *SKAP2* promotes podosome formation to facilitates tumor-associated macrophage infiltration and metastatic progression [24].

Being different from the previous prognostic studies in HCC, our predictive model firstly concentrated on mTORC1 signal pathway. More importantly, mTORC1 signal pathway was identified by GSEA, which indicated the underlying mechanism between survival of HCC and mTORC1 pathway. Moreover, the validation from three different datasets and a rigorous screening process enabled the identification of a reliable signature. However, our study has some limitations. First, prognostic signature showed a relatively low diagnostic performance, which may be attributed to the inadequate genes (199 related genes) in the initiation of screening process. More genes associated with mTORC1 signal pathways need to be confirmed in the future. Second, using a single characteristic (mTORC1 signal pathway) to establish the predictive model is an intrinsic weakness. Indeed, many other mechanisms, such as metabolism [25] and immune [8], have an effect on the progression of HCC. Furthermore, it's necessary to perform more independent trials and functional experiments to shed light on the mechanism linking mTORC1 signal pathway and HCC.

5. Conclusion

Our study is the first to identify a novel gene signature related to mTORC1 signal pathway that could be used as a potential prognostic factor in hepatocellular carcinoma.

Abbreviations

HCC: Hepatocellular carcinoma

TCGA: The Cancer Genome Atlas

ICGC: International Cancer Genome Consortium

GEO: Gene Expression Omnibus

ROC: Receiver operating characteristic

GSEA: Gene set enrichment analysis

OS: overall survival

C-index: Concordance index

Declarations

Ethics approval and consent to participate:

Not applicable.

Consent for publication:

Not applicable.

Availability of data and materials:

The datasets generated and/or analyzed during the current study are available in the TCGA (<https://portal.gdc.cancer.gov/>), GEO (<https://www.ncbi.nlm.nih.gov/geo/>), and ICGC (<https://icgc.org/>).

Competing interests:

The authors declare that they have no competing interests.

Funding:

This study was funded by the National Natural Science Foundation of China (NO. 81873248, NO. 81673903).

Authors' contributions:

ZM and SZ designed the manuscript. ZM and SL wrote and completed the manuscript. HH, LY and ZC completed the data download and analysis. All the authors approved the final manuscript.

Acknowledgements:

Not applicable.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer*. 2015; 136:E359-386.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010; 127:2893-2917.
3. El-Serag HB, Marrero JA, Rudolph L, Reddy KR. Diagnosis and treatment of hepatocellular carcinoma. *Gastroenterology*. 2008; 134:1752-1763.
4. Aravalli RN, Cressman EN, Steer CJ. Cellular and molecular mechanisms of hepatocellular carcinoma: an update. *Arch Toxicol*. 2013; 87:227-247.
5. Gu Y, Li X, Bi Y, Zheng Y, Wang J, Li X, Huang Z, Chen L, Huang Y, Huang Y. CCL14 is a prognostic biomarker and correlates with immune infiltrates in hepatocellular carcinoma. *Aging (Albany NY)*. 2020; 12:784-807.
6. Liu Y, Yang Y, Luo Y, Wang J, Lu X, Yang Z, Yang J. Prognostic potential of PRPF3 in hepatocellular carcinoma. *Aging (Albany NY)*. 2020; 12:912-930.
7. Zhang FP, Huang YP, Luo WX, Deng WY, Liu CQ, Xu LB, Liu C. Construction of a risk score prognosis model based on hepatocellular carcinoma microenvironment. *World J Gastroenterol*. 2020; 26:134-153.
8. He Y, Dang Q, Li J, Zhang Q, Yu X, Xue M, Guo W. Prediction of hepatocellular carcinoma prognosis based on expression of an immune-related gene set. *Aging (Albany NY)*. 2020; 12:965-977.
9. Jiang L, Zhao L, Bi J, Guan Q, Qi A, Wei Q, He M, Wei M, Zhao L. Glycolysis gene expression profilings screen for prognostic risk signature of hepatocellular carcinoma. *Aging (Albany NY)*. 2019; 11:10861-10882.
10. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005; 102:15545-15550.

11. Zhang L, Zhang Z, Yu Z. Identification of a novel glycolysis-related gene signature for predicting metastasis and survival in patients with lung adenocarcinoma. *Journal of Translational Medicine*. 2019; 17:423.
12. She Y, Kong X, Ge Y, Yin P, Liu Z, Chen J, Gao F, Fang S. Immune-related gene signature for predicting the prognosis of head and neck squamous cell carcinoma. *Cancer Cell International*. 2020; 20:22.
13. Raught B, Gingras AC, Sonenberg N. The target of rapamycin (TOR) proteins. *Proc Natl Acad Sci U S A*. 2001; 98:7037-7044.
14. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. *Cell*. 2012; 149:274-293.
15. Villanueva A, Chiang DY, Newell P, Peix J, Thung S, Alsinet C, Tovar V, Roayaie S, Minguez B, Sole M, et al. Pivotal role of mTOR signaling in hepatocellular carcinoma. *Gastroenterology*. 2008; 135:1972-1983, 1983.e1971-1911.
16. Liu P, Ge M, Hu J, Li X, Che L, Sun K, Cheng L, Huang Y, Pilo MG, Cigliano A, et al. A functional mammalian target of rapamycin complex 1 signaling is indispensable for c-Myc-driven hepatocarcinogenesis. *Hepatology*. 2017; 66:167-181.
17. Xu Z, Xu M, Liu P, Zhang S, Shang R, Qiao Y, Che L, Ribback S, Cigliano A, Evert K, et al. The mTORC2-Akt1 Cascade Is Crucial for c-Myc to Promote Hepatocarcinogenesis in Mice and Humans. *Hepatology*. 2019; 70:1600-1613.
18. Lu C, Ning Z, Wang A, Chen D, Liu X, Xia T, Tekcham DS, Wang W, Li T, Liu X, et al. USP10 suppresses tumor progression by inhibiting mTOR activation in hepatocellular carcinoma. *Cancer Lett*. 2018; 436:139-148.
19. Singh P, James RS, Mee CJ, Morozov IY. mRNA levels are buffered upon knockdown of RNA decay and translation factors via adjustment of transcription rates in human HepG2 cells. *RNA Biology*. 2019; 16:1147-1155.
20. McLoughlin MR, Orlicky DJ, Prigge JR, Krishna P, Talago EA, Cavigli IR, Eriksson S, Miller CG, Kundert JA, Sayin VI, et al. TrxR1, Gsr, and oxidative stress determine hepatocellular carcinoma malignancy. *Proceedings of the National Academy of Sciences*. 2019; 116:11408-11417.
21. Lee SMY, Li MLY, Tse YC, Leung SCL, Lee MMS, Tsui SKW, Fung KP, Lee CY, Waye MMY. Paeoniae Radix, a Chinese herbal extract, inhibit hepatoma cells growth by inducing apoptosis in a p53 independent pathway. *Life Sciences*. 2002; 71:2267-2277.
22. Lian Y-F, Huang Y-L, Zhang Y-J, Chen D-M, Wang J-L, Wei H, Bi Y-H, Jiang Z-W, Li P, Chen M-S, Huang Y-H. CACYBP Enhances Cytoplasmic Retention of P27^{Kip1} to Promote Hepatocellular Carcinoma Progression in the Absence of RNF41 Mediated Degradation. *Theranostics*. 2019; 9:8392-8408.
23. Zhang B, Chen M, Zhang Y, Chen W, Zhang L, Chen L. An ultrasonic nanobubble-mediated PNP/fludarabine suicide gene system: A new approach for the treatment of hepatocellular carcinoma. *PLOS ONE*. 2018; 13:e0196686.
24. Tanaka M, Shimamura S, Kuriyama S, Maeda D, Goto A, Aiba N. SKAP2 Promotes Podosome Formation to Facilitate Tumor-Associated Macrophage Infiltration and Metastatic Progression.

25. Liu G-M, Xie W-X, Zhang C-Y, Xu J-W. Identification of a four-gene metabolic signature predicting overall survival for hepatocellular carcinoma. *Journal of Cellular Physiology*. 2020; 235:1624-1636.

Tables

Table 1. Baseline patient characteristic in TCGA, GEO and ICGC cohorts.

Clinical characteristics		Number	percent (%)
TCGA-LIHC (n=377)			
Survival status	Survival	249	66
	Death	128	34
Age (1 patient missing)	≤65 years	235	62.5
	>65 years	141	37.5
Gender	Female	122	68
	Male	255	32
Stage (24 patients missing)	I	175	50
	II	87	24.6
	III	86	24.4
	IV	5	1
Grade (5 patients missing)	G1	55	14
	G2	180	48
	G3	124	33
	G4	13	5
T classification (3 patients missing)	T1	185	49
	T2	95	26
	T3	81	22
	T4	13	3
GSE76427 (n=115)			
Survival status	Survival	92	80
	Death	23	20
Age	≤65 years	65	56.5
	>65 years	50	43.5
Gender	Female	22	19.1
	Male	93	80.9
Stage	I	55	47.8
	II	35	30.4
	III	21	18.3
	IV	4	3.5
ICGC-LIRI (n=260)			
Survival status	Survival	214	82.4
	Death	46	17.6
Age	≤65 years	98	37.7
	>65 years	162	62.3
Stage	I	40	15.4
	II	117	45
	III	80	30.8
	IV	23	8.8

Table 2. Details of signal pathways selected by GSEA.

Name	Size	ES	NES	Normalized p-value
MTORC1 SIGNALING	200	0.581694	1.982214	0
UV_RESPONSE_UP	158	0.506378	1.955258	0
GLYCOLYSIS	200	0.506672	1.88337	0
G2M_CHECKPOINT	199	0.73077	1.881815	0.002088
MYC_TARGETS_V1	199	0.685811	1.860498	0.004141
E2F_TARGETS	200	0.745006	1.856894	0.002079

Note: ES=Enrichment Score, NES=Normalized Enrichment Score.

Figures

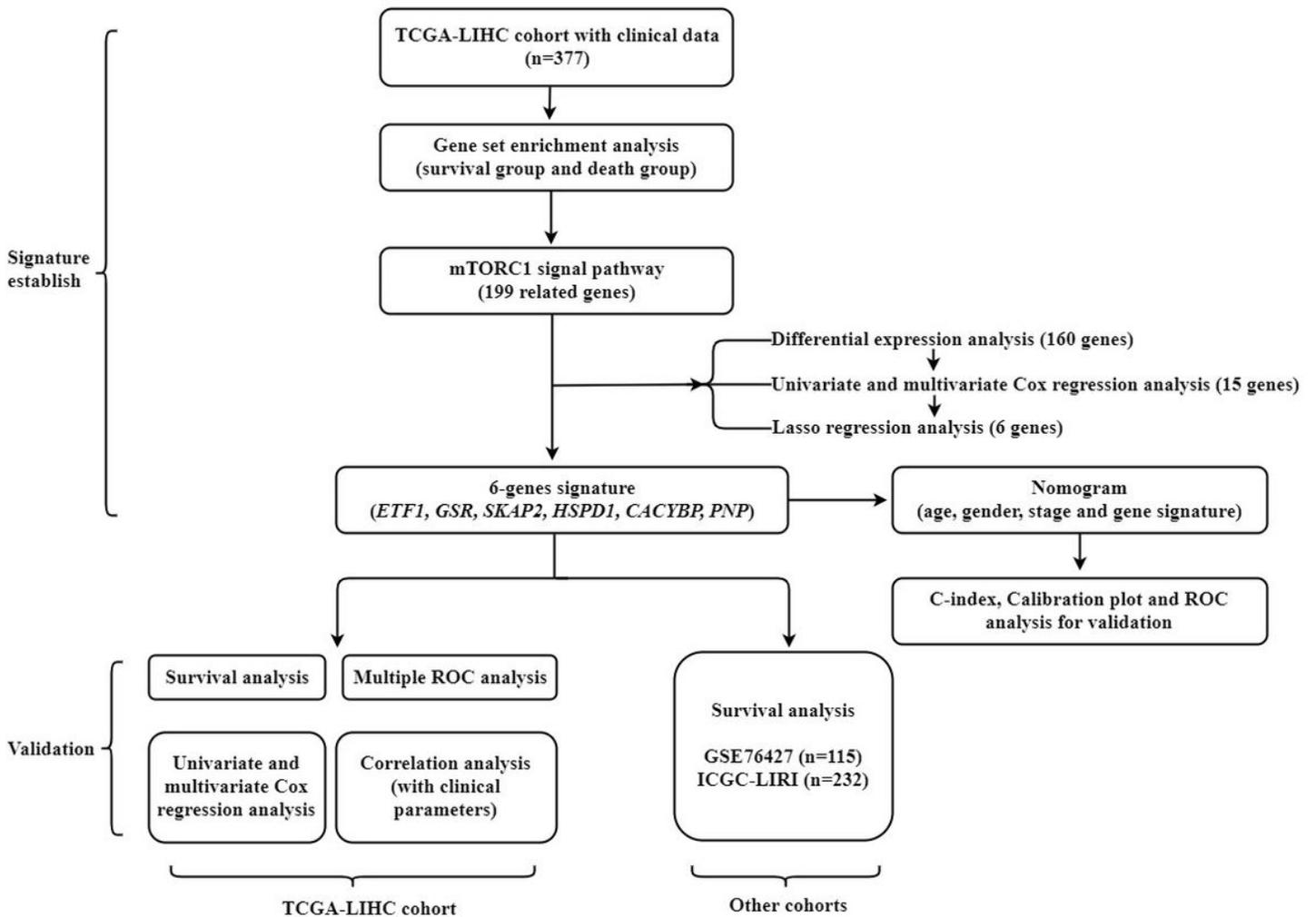


Figure 1

The flowchart of this study.

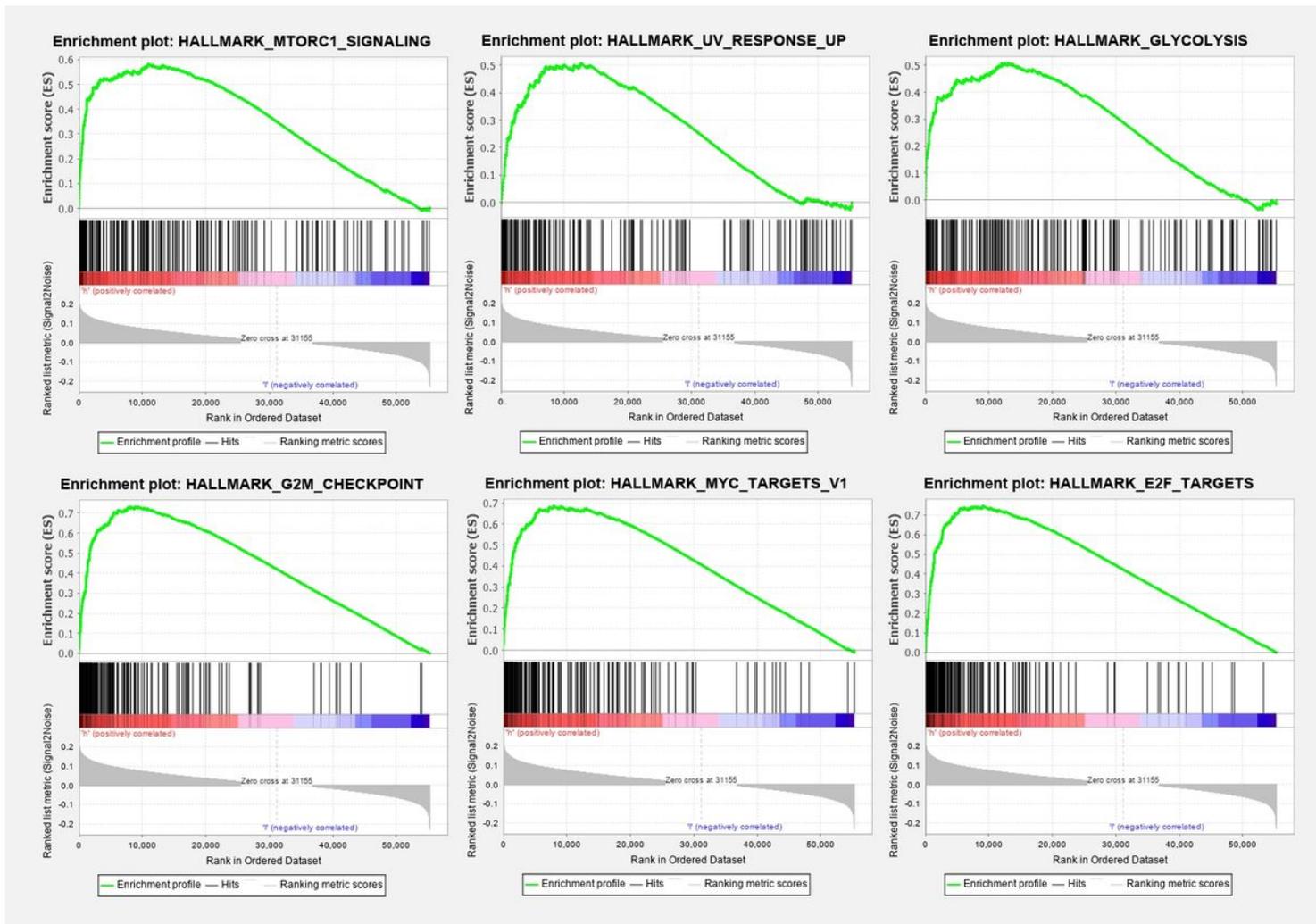


Figure 2

Enrichment plots of six pathways which importantly differentiated between survival and death group.

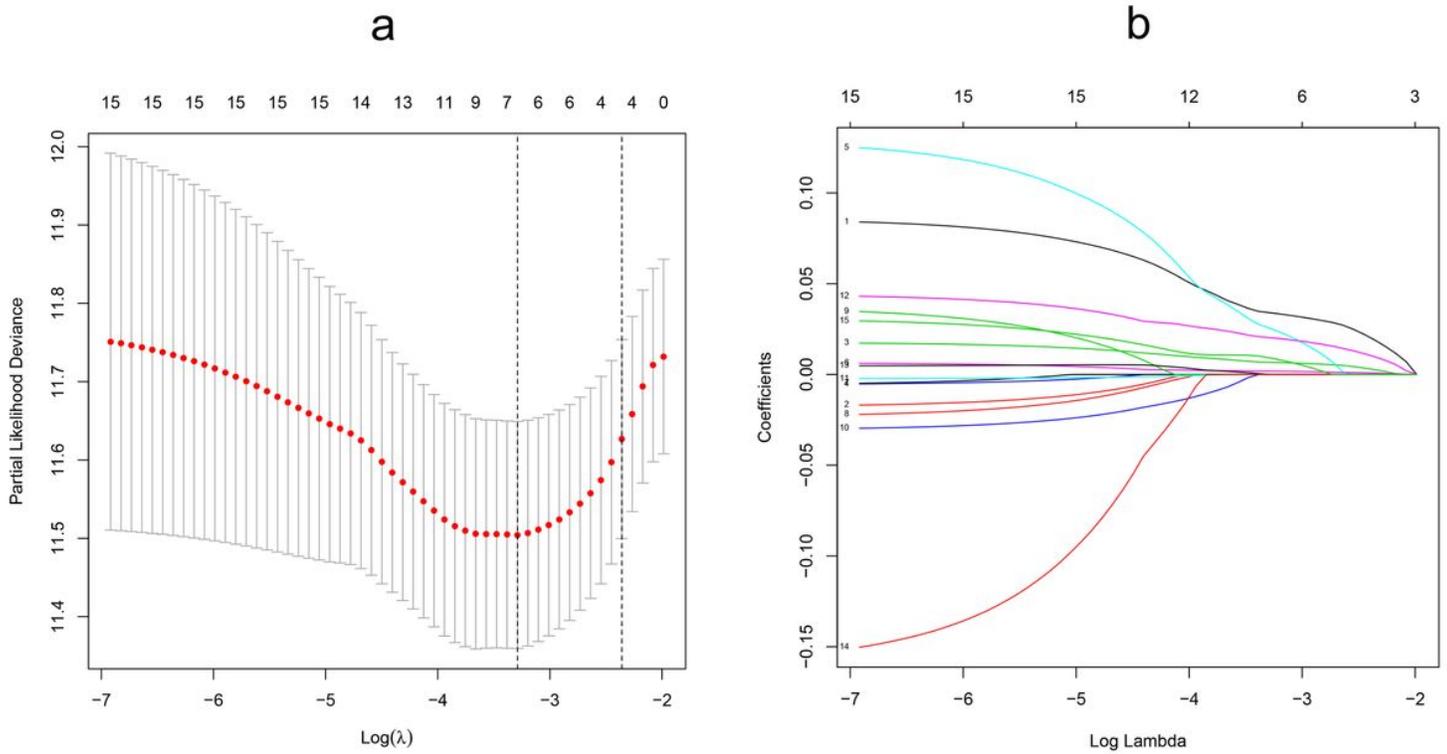


Figure 3

Lasso regression analysis results. Notes: Panels (a) show the partial likelihood deviance for the lasso regression, panels (b) show the lasso regression analysis.

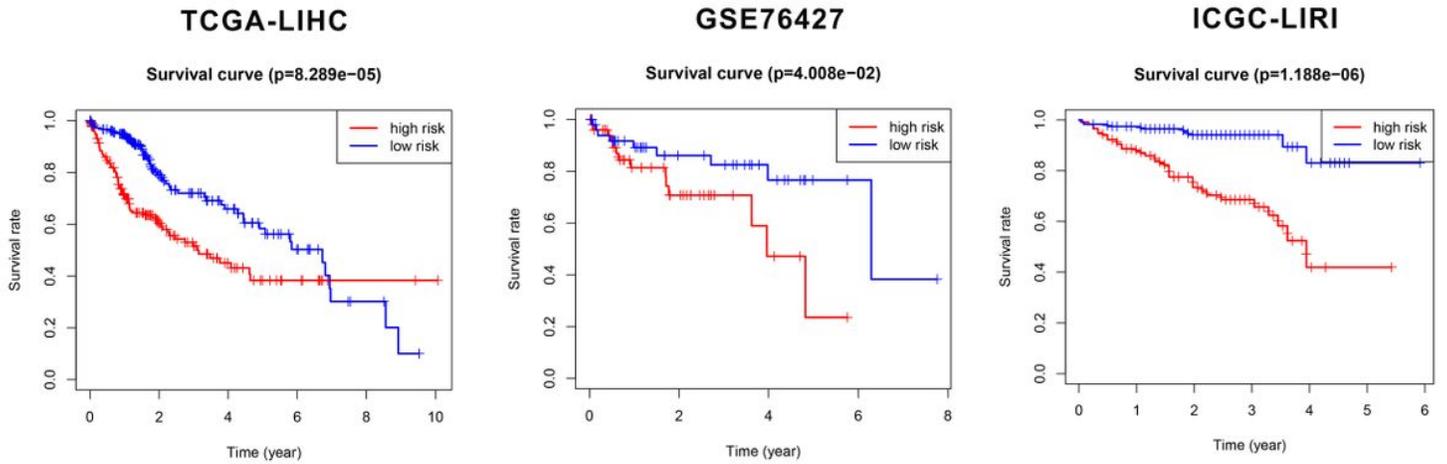


Figure 4

Survival analysis results of three datasets.

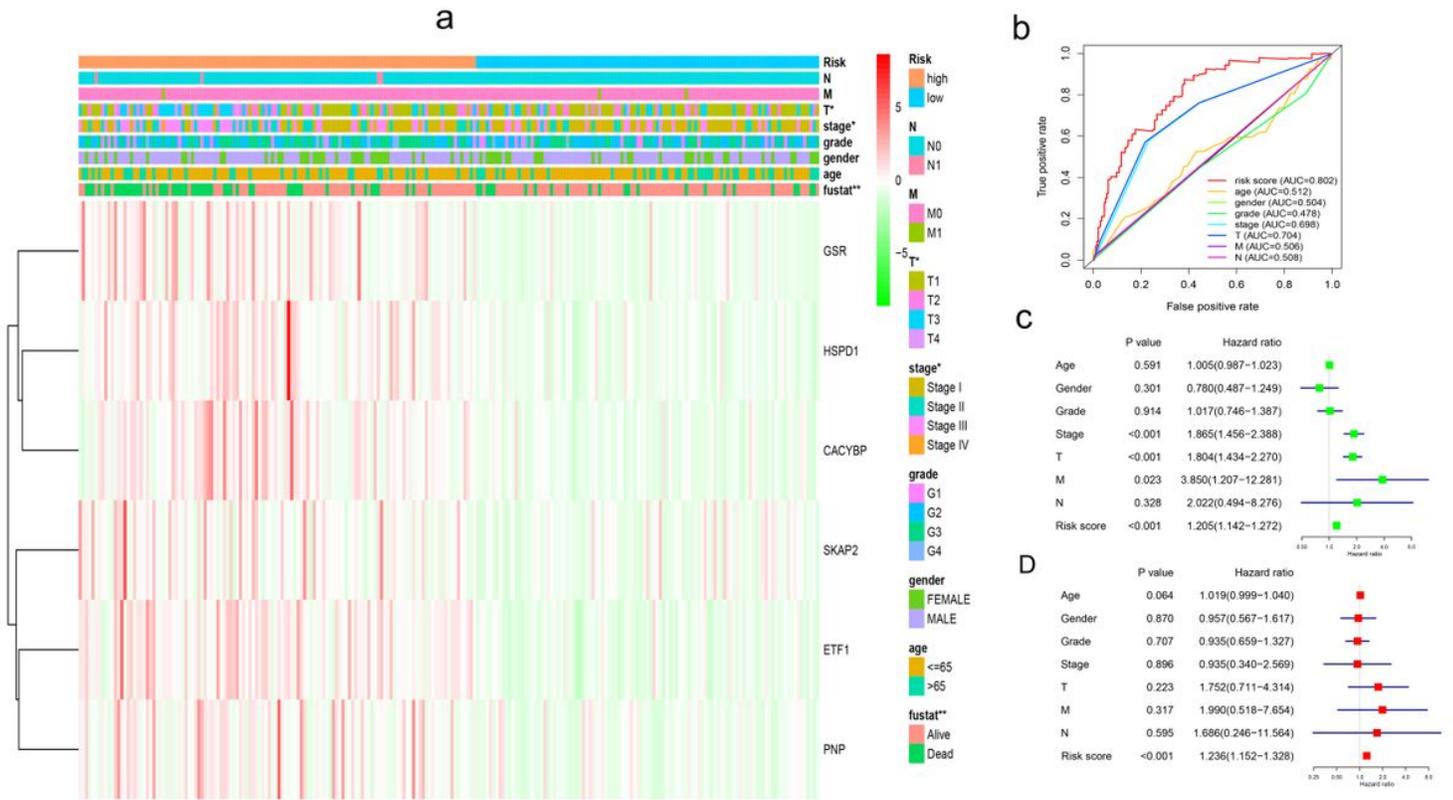


Figure 5

Figure 5. Heatmap, multiple ROC curve and Cox regression analysis results. Notes: Panels (a) show the heatmap of the established signature with other clinical parameters, panels (b) show the multiple ROC curve, panels (c) show the result of univariate Cox regression analysis, and panels (d) show the result of multivariate Cox regression analysis. In the heatmap, symbol “*” in means $P < 0.05$, “**” means $P < 0.01$, “T” represents tumor size, “M” represents distant metastasis, and “N” represents regional lymph node metastasis.

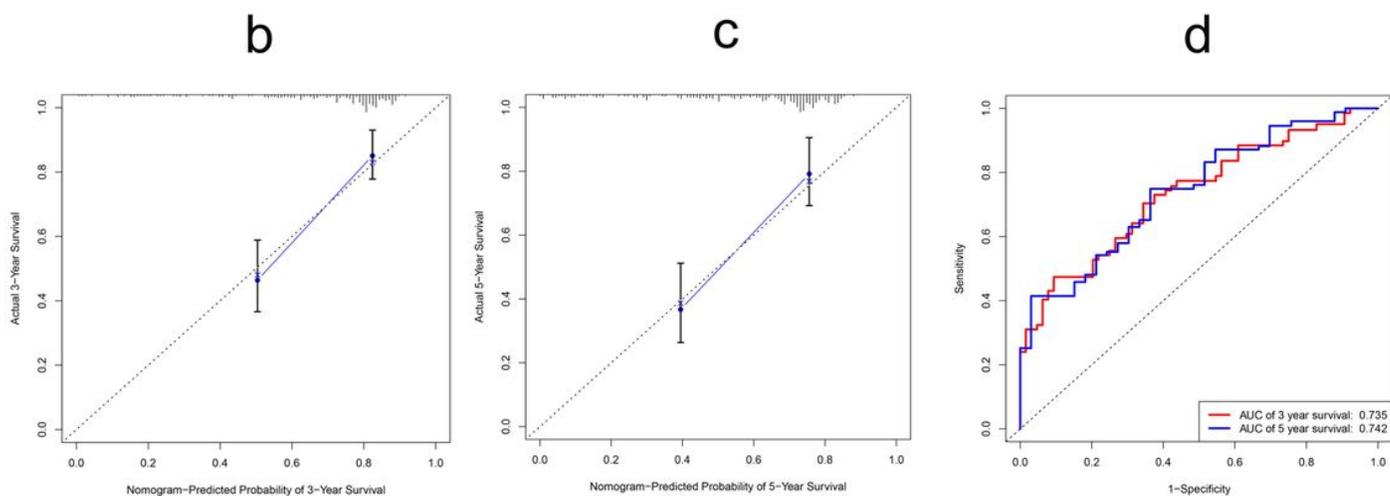
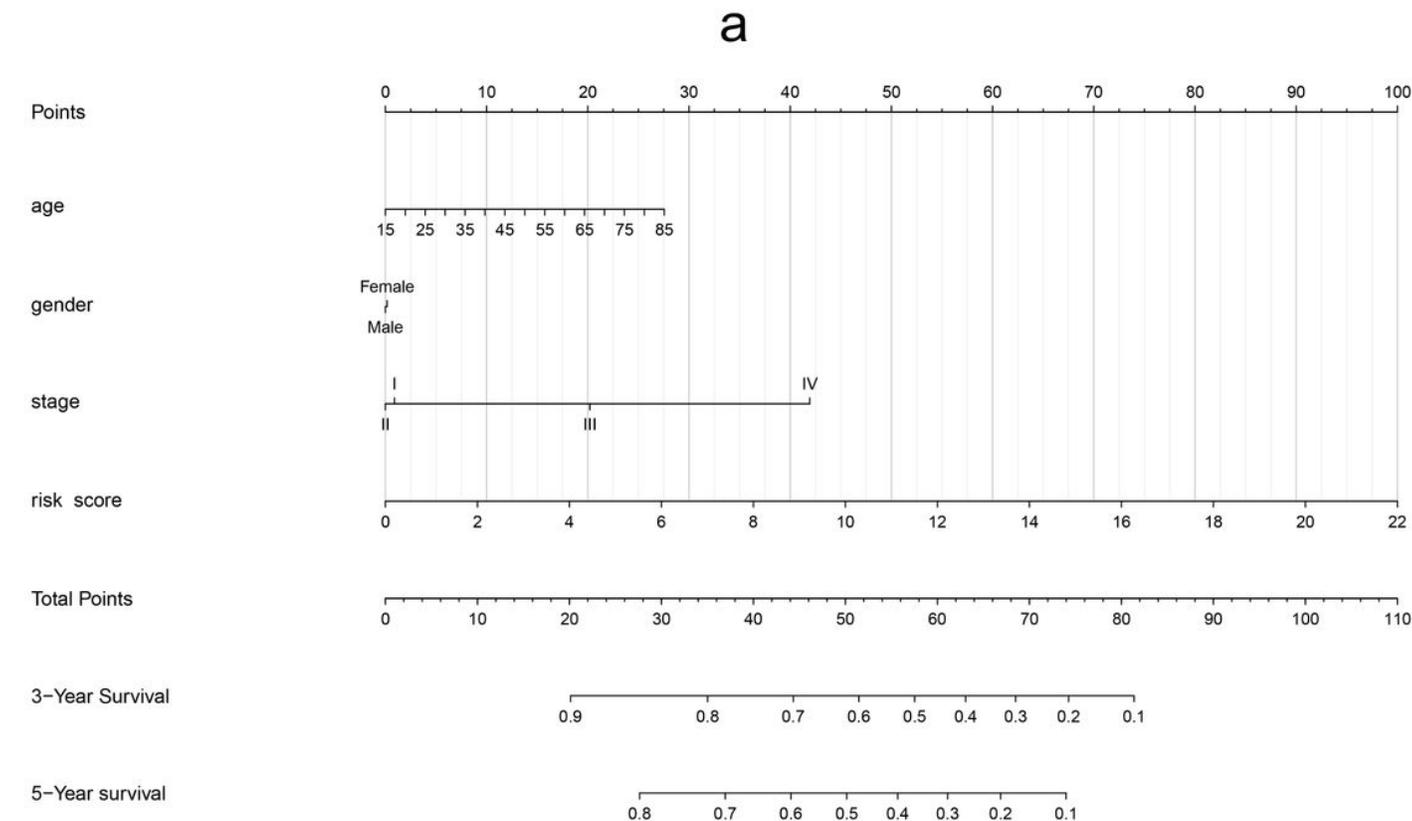


Figure 6

Construction and validation of Nomogram. Notes: Panels (a) show the nomogram based on age, gender, stage and gene signature, panels (b) (c) show the calibration plot, and panels (d) show the ROC curve of 3-years and 5-years OS.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- S3.docx
- S2.docx
- S1.pdf