

A riskscore-based nomogram for the prediction of overall survival in cervical squamous cell carcinoma

Lingling Li

School of Medicine, Nankai University, Tianjin, China

Sujie Zhang

Department of Oncology, The First Medical Center of Chinese PLA General Hospital, Beijing, China

Lijie Wang

Department of Oncology, The First Medical Center of Chinese PLA General Hospital, Beijing, China

Haitao Tao

Department of Oncology, The First Medical Center of Chinese PLA General Hospital, Beijing, China

Yaping Long

School of Medicine, Nankai University, Tianjin, China

Yiming Gao

Department of Oncology, The First Medical Center of Chinese PLA General Hospital, Beijing, China

Yi Hu (✉ huyi301zlx@ sina.com)

Chinese PLA General Hospital <https://orcid.org/0000-0001-9319-5692>

Yi Sun

Shanghai Institute of Microsystem and Information Technology Chinese Academy of Sciences,
Shanghai, China

Research

Keywords: cervical squamous cell carcinoma, nomogram, overall survival, riskscore, immune score, tumor microenvironment.

Posted Date: October 22nd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-94429/v1>

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

[Read Full License](#)

Abstract

Background:

Cervical cancer is still the major cause of cancer-related death among women. However, the prognosis of cervical cancer varies even in the same stage. Thus, exploring prognostic biomarkers that could reflect its biological heterogeneity may contribute to identify patients with a poor prognosis.

Methods:

Based on the ESTIMATE algorithm, we acquired the immune/stromal scores of cervical squamous cell carcinoma (CSCC) patients collected from The Cancer Genome Atlas (TCGA) dataset. Subsequently, we analyzed the DEGs between high- and low immune score groups using R package edgeR and performed K-M analysis to illustrate the relationship between differentially expressed genes (DEGs) and the overall survival to select survival-related DEGs. Then the LASSO regression model was constructed with the package “glmnet” in R to evaluate the riskscore of each patient. Finally, we developed a nomogram composing riskscore and clinicopathological characteristics to predict the overall survival (OS) of CSCC patients. The R software v3.6.1 was used for statistical analyses. All statistical tests were two-tailed.

Results:

We established a riskscore model composed of two genes including FOXP3 and ZAP70. The receiver operating characteristic (ROC) curve demonstrated a good potency of the riskscore model. Ultimately, we constructed a nomogram composing riskscore, age and stage to predict the overall survival (OS) of CSCC patients. The area under the ROC curve (AUC) of the nomogram for OS was 0.805, 0.723 and 0.748 for the first, third, and fifth years, respectively. The concordance index (C-index) was 0.746. The calibration curves also showed optimal accuracy of the nomogram for survival prediction.

Conclusion:

The nomogram based on riskscore could predict overall survival in CSCC and may benefit those patients through individualized immunotherapy.

Background

Cervical cancer represents one of the most common solid tumors in females worldwide, with 530,000 new cases and 270,000 deaths annually around the world[1]. The overwhelming majority of these new cases arise in developing countries, including China[1]. With the improvement of early diagnosis rate of cervical cancer, the incidence has been decreased due to application of various screening methods and effective vaccines against HPV. Nonetheless, on account of the lack of related knowledge education, healthcare resources and economic factors, the incidence rate of cervical cancer in some underdeveloped and remote areas remains high, with a younger trend[2]. There are multiple methods to treat cervical cancer, including surgical resection, radiotherapy, chemotherapy, immunotherapy and so on. However, the

incidence of patients with stage I-III experiencing metastatic disease ranges from 15–61%[3]. Moreover, the median overall survival for patients with disease progression varied a lot, with a range of 7 to 53 months[4], indicating that it is crucial to identify prognostic factors for cervical cancer. To our knowledge, lymph node status and the International Federation of Gynecology and Obstetrics (FIGO) stage are important prognostic factors for cervical cancer, but it seems that the prediction efficiency of these factors is limited[5].

Recently, Substantial progress has been made in immunotherapy for cervical cancer, and immune checkpoint inhibitors have been playing its role as a new therapeutic option for these patients[6]. However, clinical studies have shown the response rate of cervical cancer to checkpoint immunotherapy to be in the region to 10–25%[7, 8], so it is necessary to narrow down the patients who could be expected to have satisfactory effects using ideal biomarkers[9]. The tumor microenvironment (TME) consists of immune cells, stromal cells, extracellular matrix molecules and inflammatory mediators[10]. In cervical cancer, the composition of the TME has an impact on survival, with a higher ratio of CD8 TILs to CD4 Tregs being associated with less extensive disease (lymph node negativity) and improved survival rates[11, 12], suggesting that the understanding of the TME may help identify optimal biomarker to predict overall survival in cervical cancer and improve the efficacy of cancer treatments in the future. To our knowledge, there has not been a prognostic nomogram based on immune-related factors in cervical cancer up to now.

ESTIMATE was a R package designed by Yoshihara et al[13] to obtain immune and stromal scores based on the specific gene expression signatures. Several reports have documented the application of ESTIMATE in multiple tumors[14–17]. However, whether the prognosis in CSCC, which is the most common histologic subtype of cervical cancer, could be investigated using ESTIMATE algorithm is still unclear. In the present study, the TME and the gene expression profile of CSCC were investigated to establish a riskscore prognostic model for CSCC based on The Cancer Genome Atlas (TCGA) database.

Materials And Methods

Data collection

The transcriptional profiles and clinical information of CSCC patients were retrieved from the TCGA website (<https://cancergenome.nih.gov/>). The inclusive criteria were as follows: (1) only patients with primary CSCC were enrolled, (2) no history of other malignancies, (3) only samples with RNA-sequencing data, (4) complete clinicopathological materials including age and stage, which are found to be correlated with overall survival in our study, (5) complete survival data with a survival time of more than 1 mo, (6) minimum follow-up of 60 days. Ultimately, 210 CSCC samples with complete RNA-seq data and the corresponding clinical information were subjected to consequent analyses. This study meets TCGA's publication guidelines. Immune and stromal scores were calculated by applying the ESTIMATE algorithm[13] to the mRNA expression data.

Overall survival curve

Kaplan-Meier (K-M) plots were performed to illustrate the correlation of immune/stromal scores with patients' overall survival (OS). The relationship was tested by the log-rank test.

Differentially expressed genes analysis

To screen for differentially expressed genes (DEGs), 210 CSCC patients obtained from the TCGA dataset were divided into high and low immune score groups according to the ESTIMATE results. The differentially expressed genes (DEGs) were identified using the package edgeR[40] in R software (Version 3.6.1; <https://www.r-project.org/>). The genes that satisfied the inclusion criteria of adj.P-value < 0.05 and $|log FC| \geq 1$ were identified as statistically significant DEGs and included in this study. K-M analysis was performed to illustrate the relationship between DEGs and patients' OS. The relationship was tested by the log-rank test and $P < 0.01$ was considered as significant. Volcano plot and heatmaps were generated using the "gplots" and "pheatmap" package in R software, respectively.

GO and KEGG enrichment analysis of DEGs

The Gene Ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway analyses were performed by "clusterProfiler" package[41] in R to identify potential biological processes (BP), molecular functions (MF), cellular components (CC) and signaling pathways related to the DEGs.

Construction of the riskscore prognostic model

To prevent overfitting[42], the least absolute shrinkage and selection operator (LASSO) regression with L1-penalty was applied, which is an appropriate solution to establish signatures if there are numerous correlated covariates with small sample size[43]. In our study, the LASSO regression model was constructed with the package "glmnet" in R and the penalty parameter "lambda" was used to choose the best model based on leave-one-out cross-validation[44]. Finally, based on the LASSO analysis, we extracted genes from DEGs which were regarded as significant in K-M analysis and their corresponding coefficients. Combining expression levels of the selected RNAs (RNA_i) with the relative coefficients, a riskscore for each sample was calculated: $Riskscore = \sum_{i=1}^n RNA_i \times Coe_i$, the Coe_i represents the coefficient of each mRNA from the model. CSCC patients were divided into low- and high-risk groups based on the median riskscore. The K-M survival curves for risk were performed. The predictive ability of the riskscore prognosis model was assessed by "timeROC" package in R software.

Estimation of immune infiltration

The CIBERSORT algorithm[45] was applied to estimate immune infiltration cell composition including 22 types of immune cells, based on normalized gene expression profiles. Samples with the CIBERSORT $P < 0.05$ were considered eligible for subsequent analysis. For each sample, all the estimates of immune cell type fractions summed up to 1[46].

Development and validation of the nomogram

Using the package of “rms” in R, a nomogram that can visualize the prognostic ability of various factors in a single figure[47] was established based on the riskscore and clinicopathological data. The concordance index (C-index) and receiver operating characteristic (ROC) analysis were used to determine its predictive accuracy. The calibration curves[48] were used to determine its discriminatory capacity. The internal validation of the nomogram was conducted by bootstraps with 1000 resamples.

Statistical analysis

The R software v3.6.1 was used for statistical analyses. All statistical tests were two-tailed.

Results

Immune scores were significantly related to CSCC prognosis

The flow chart of our study procedure is presented in Fig. 1. In this study, 210 CSCC patients with complete mRNA expression data and the corresponding clinical information were included for subsequent analysis. According to the ESTIMATE algorithm, immune scores ranged from -1199.72 and 3426.59, stromal scores ranged from -2309.11 to 813.34, and ESTIMATE scores ranged from -3241.44 to 4017.39 (Fig. 2A, B and C). To explore the influence of immune/stromal scores on prognosis, we divided 210 CSCC patients into high- and low-score groups and constructed K-M curves. We found that the immune scores were significantly positively correlated with overall survival (OS) ($P=0.005$) (Fig. 2D). On the contrary, stromal scores were not associated with OS ($P=0.364$) (Fig. 2E). Then, we investigated the correlation between immune/stromal scores and clinical characteristics. The results showed that the distributions of immune/stromal scores did not vary with stage (Fig. 3A and B) and age (Fig. 3C and D).

Differentially expressed genes with immune score

As stromal scores were not correlated with overall survival in this study, we only explored the difference of gene expression profiles between high- and low immune score groups to construct the prognostic model. Finally, according to the inclusion criteria, 641 DEGs were identified and exhibited by volcano plot (Fig. 4A). Compared with the low immune scores group, 504 genes were upregulated and 137 genes were downregulated in the high immune score group (Supplementary Table 1). Among these DEGs, 72 DEGs were significantly positively correlated with OS ($P<0.01$) (Supplementary Table 2). An additional file shows this in more detail [see Additional file 1]. The heatmap showed the expression profiles of the 72 genes of prognostic value of cases in low and high immune scores groups (Fig. 4B).

Functional enrichment analysis

The potential functions of the 641 DEGs were evaluated by GO analysis and KEGG pathway. The top 5 GO terms identified included immunoglobulin complex, antigen binding, adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains, lymphocyte mediated immunity and complement activation classical pathway (Fig. 5A). The KEGG

pathway analysis revealed that the DEGs were enriched in allograft rejection, viral protein interaction with cytokine and cytokine receptor, cytokine-cytokine receptor interaction, cell adhesion molecules (CAMs) and staphylococcus aureus infection (Fig. 5B).

Construction of a riskscore model and evaluation of its predictive ability

Seventy-two genes of prognostic value were included in LASSO analysis, in which the contributions of 72 genes were weighted by their relative coefficients (Fig. 6A and B). In consequence, two genes were selected for constructing the riskscore model, and the final riskscore formula was as follows: Risk score = [-0.0115 × expression level of FOXP3] + [-0.0201 × expression level of ZAP70]. Then the riskscore of each sample was calculated according to the formula and the samples were devided into high- or low-risk groups based on the median riskscore. The distribution of the riskscore was indicated in Fig. 7A. As indicated in the K-M curve plot (Fig. 7B), there was a significant difference between high- and low-risk groups in patients' OS. Patients in the low-risk group had significantly longer OS than those in the high-risk group. The area under the ROC curve (AUC) of the prognostic model for OS was 0.766, 0.717 and 0.701 for the first, third, and fifth years, respectively (Fig. 7C), demonstrating that the riskscore model had good performance.

Correlation of the riskscore with immune infiltration

By using the CIBERSORT algorithm, we acquired the relative fractions of 22 immune cell types. 185 cases of CSCC in the TCGA dataset were enrolled for subsequent analysis after the filter criteria with *P* value < 0.05, with 97 cases in low-risk group and 88 cases in high-risk group. As shown in the bar plot, the T cells CD8 and the macrophage M0 were the most significant enrichment of immune cells among 22 immune cell types (Fig. 8A). As shown in the heatmap, the fractions of immune cells in CSCC varied significantly between the low- and high-risk groups (Fig. 8B). The vioplot showed the fraction of T cells CD8, the macrophage M1 and resting Mast cells was significantly higher in low-risk group compared with high-risk group. In contrast, the fraction of resting CD4 memory T cells, the macrophage M0 and activated Dendritic cells were significantly higher in high-risk group patients compared with low-risk group (Fig. 8C). As the expresion T cells CD4 naive in all samples were estimated as zero, we eliminated this cell type in the analysis of the correlation between different tumor-infiltrating immune cells (TIIICs). The corHeatmap indicated a weak to moderate correlation in the proportions of 21 types of TIIICs (Fig. 8D). These results indicated that the different immune cell infiltration in patients with CSCC may be a potential prognostic biomarker and targets for immunotherapy.

A nomogram based on the riskscore model

To establish a clinically applicable method to evaluate the prognosis of CSCC patients, a nomogram was generated for OS prediction using the three prognostic factors including riskscore, age and pathologic stage (Fig. 9A). The concordance index (C-index) was 0.746. The calibration plots for the possibility of 3- and 5-year survival showed an optimal agreement between the prediction by the novel nomogram (Fig. 9B-C). The AUC was 0.805, 0.723 and 0.748 for the 1, 3 and 5-year survival times, respectively

(Fig. 10A). However, the AUC was 0.3, 0.375 and 0.358 for the 1, 3 and 5-year survival times (Fig. 10B) when immunescore was included as a single prognostic factor to predict patients' OS, suggesting that the nomogram has a much better predictive accuracy than the immunescore model and our method to construct a prognostic model is effective. In brief, the nomogram has a good predictive accuracy and aid clinical management.

Discussion

Cervical cancer remains a huge challenge for public health in developing countries. The past decade has witnessed an astonishing outpouring of research on the treatment of cervical cancer. Significant progress has been made in precise immunotherapy, which is taking the place of chemotherapy gradually. Further, it is urgently needed to identify effective biomarkers in tumor immune microenvironment correlated with prognosis.

In our study, we applied the ESTIMATE algorithm to calculate immune and stromal scores of 210 CSCC samples obtained from the TCGA dataset, where the results showed that high immune scores were significantly associated with better OS. As stromal scores were not found to be significantly related to prognosis, we just analyzed the DEGs between high- and low immune score groups. Then 641 DEGs were identified. GO analysis showed that the DEGs were mainly involved in the TME, such as immunoglobulin complex, antigen binding, adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains, lymphocyte mediated immunity and complement activation classical pathway. K-M analysis of 641 DEGs suggested that 72 DEGs were significantly correlated with prognosis. Subsequently, LASSO regression was applied to construct a riskscore model that could identify CSCC patients with unfavorable outcomes. Finally, we selected 2 genes including FOXP3 and ZAP70. Though the difference of expression level of FOXP3 and ZAP70 between two groups were not obvious in the heatmap (Fig. 4B), both FOXP3 and ZAP70 satisfied the criteria of $\log FC > 1$. It may result from the expression of FOXP3 and ZAP70 being in the range of 0–2, which was all green in the heatmap and was hard to distinguish. However, it did not mean the difference of expression level of FOXP3 and ZAP70 between two groups was not obvious, as they satisfied the criteria of $FC > 2$. The heatmap only showed the difference of expression level between low-immune-score group and high-immune-score group, it did not reflect the contribution of predictive ability of genes to the overall survival. The AUC for the riskscore model for predicting the 1-, 3-, and 5-year survival were 0.766, 0.717 and 0.701, respectively, indicating that this model had a good performance for OS prediction and modifying the TME may impede or eliminate tumors. The genes (FOXP3 and ZAP70) that formed our riskscore model may be potential targets and provide better performance in combination.

FOXP3 is a transcription factor belonging to the family of forkhead box (FOX) protein. All the family members have a forkhead (FKH) domain which functions as a transcriptional repressor or activator[18–20] and regulate the transcription of approximately 700 genes[21]. FOXP3 was found to be expressed in both regulatory T cells[22, 23], based on which the detection of the FOXP3 expression in tumor samples was used as an indicator of regulatory T cells[24], and in tumor cells, including breast cancer, lung cancer,

gastric cancer, thyroid cancer, colon cancer, bladder cancer and glioblastoma cell lines[25–29]. However, the role of FOXP3 in cancer is inconsistent and even the opposite. In some types of human tumors, the expression of FOXP3 is upregulated with promotion of the development of tumors, leading to a poor prognosis. While in some other types of tumors, it is a totally different story[30]. It was reported that the expression level of FOXP3 was high in the cytoplasm and nucleus as well as cancer interstitium in cervical cancer, but low in cervical intraepithelial neoplasia (CIN), and the expression of FOXP3 was even not observed in normal cervical epithelium, indicating that FOXP3 may not be involved in the initiation of this malignancy but may facilitate tumor growth[31]. It was also found that the expression of FOXP3 in metastatic lymph nodes was significantly stronger than that in normal lymph nodes, suggesting the correlation between FOXP3 and tumor metastasis[31]. The protein tyrosine kinase zeta-associated protein (ZAP70) is a protein tyrosine kinase acting in the T cell receptor signal transduction system, which has been identified to play a crucial role in T cell activation and other immune response[32–34]. Studies have accumulated pointing to a correlation of ZAP70 expression in CLL cells with IGVH status[35–37]. ZAP70 expression has been considered as an effective prognostic factor in chronic lymphocytic leukemia [37]. However, the expression of ZAP70 and the impact on prognosis in cervical cancer are unclear. For all we know, there has not been a riskscore prognostic model consisting of these two gene signature in cervical cancer.

Then, we constructed a nomogram comprising riskscore, age and the pathologic stage.

The nomogram we developed was the visualization of Cox regression model specifically to provide a more individualized prediction of outcome of each patient. Each variable had a corresponding score on the “points” scale at the top. After summing all of the assigned scores and locating it on the “total points” scale, it was easily able to discover the corresponding probability of 1-, 3- and 5- year OS. The results of both AUC curves and calibration plot showed that the nomogram we established had a good predictive accuracy.

The past decade has witnessed the evolution of the treatment of metastatic cervical cancer from cytotoxic agents to immunotherapy. A high frequency of PD-L1 expression has been reported in up to 80% of CSCC patients, making this disease a prime candidate for PD-1/PD- L1 inhibition[38]. Furthermore, identifying immune mechanisms to overcome immune suppression and understanding the biomarkers of prognosis will contribute to the development of effective individualized targeted approaches. We investigated immune mechanisms and the component of TIICs subpopulation between patients in the low- and high-risk groups. The results have shown a significantly higher fraction of T cells CD8, the macrophage M1 and resting Mast cells in low-risk group patients, while a significantly higher fraction of resting CD4 memory T cells, the macrophage M0 and activated Dendritic cells in high-risk group patients. A recent study has demonstrated the levels of CD8(+) and CD8(+) CD28(+) T cells in the early stage cervical cancer group were markedly lower than those in the CIN group and the control group and the ratio of CD4(+)CD25(+)/CD4(+) in the cervical cancer group was significantly higher than in the control group[39], which is consistent with the results of our study. Based on all these results, we could draw a conclusion that the riskscore model is associated with the immunosuppressive environment.

As far as we know, the prognostic nomogram we developed is the first one based on immune-related factors in CSCC. However, we are obliged to admit the limitations in our study. The riskscore model is necessary to be further validated in multicenter clinical trials and prospective studies.

Conclusion

In conclusion, our research developed a riskscore model that is based on two immune-related genes to predict the OS of CSCC, which provides an immunological perspective to uncover the mechanisms and assists in clinical decision making for personalized treatment. The nomogram we developed based on riskscore could predict overall survival in CSCC and may benefit those patients through individualized immunotherapy.

Abbreviations

cervical squamous cell carcinoma (CSCC)

The Cancer Genome Atlas (TCGA)

Differentially Expressed Genes (DEGs)

receiver operating characteristic (ROC)

overall survival (OS)

area under the ROC curve (AUC)

concordance index (C-index)

International Federation of Gynecology and Obstetrics (FIGO)

tumor microenvironment (TME)

forkhead box (FOX)

Gene Ontology (GO)

Kyoto encyclopedia of genes and genomes (KEGG)

potential biological processes (BP)

molecular functions (MF)

cellular components (CC)

least absolute shrinkage and selection operator (LASSO)

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the The Cancer Genome Atlas (TCGA) dataset[<https://portal.gdc.cancer.gov/>].

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Military Health special research project (16BJZ24).

Author Contributions

LLL and YH contributed conception and design of the study; LLL and YS organized the database; LLL performed the statistical analysis; LLL wrote the whole manuscript; All authors contributed to manuscript revision, read and approved the submitted version.

Acknowledgments

Not applicable

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
2. Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ, He J. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66:115–32.
3. Pfaendler KS, Tewari KS. Changing paradigms in the systemic treatment of advanced cervical cancer. Am J Obstet Gynecol. 2016;214:22–30.
4. Elit L, Fyles AW, Devries MC, Oliver TK, Fung-Kee-Fung M. Follow-up for women after treatment for cervical cancer: a systematic review. Gynecol Oncol. 2009;114:528–35.

5. Benedet JL, Odicino F, Maisonneuve P, Beller U, Creasman WT, Heintz AP, Ngan HY, Pecorelli S. Carcinoma of the cervix uteri. *Int J Gynaecol Obstet.* 2003;83(Suppl 1):41–78.
6. Enwere EK, Kornaga EN, Dean M, Koulis TA, Phan T, Kalantarian M, Köbel M, Ghatare P, Magliocco AM, Lees-Miller SP, Doll CM. Expression of PD-L1 and presence of CD8-positive T cells in pre-treatment specimens of locally advanced cervical cancer. *Mod Pathol.* 2017;30:577–86.
7. Saglam O, Conejo-Garcia J. PD-1/PD-L1 immune checkpoint inhibitors in advanced cervical cancer. *Integr Cancer Sci Ther.* 2018; 5.
8. Naumann RW, Hollebecque A, Meyer T, Devlin MJ, Oaknin A, Kerger J, López-Picazo JM, Machiels JP, Delord JP, Evans TRJ, Boni V, Calvo E, Topalian SL, et al. Safety and Efficacy of Nivolumab Monotherapy in Recurrent or Metastatic Cervical, Vaginal, or Vulvar Carcinoma: Results From the Phase I/II CheckMate 358 Trial. *J Clin Oncol.* 2019;37:2825–34.
9. Kagabu M, Nagasawa T, Fukagawa D, Tomabechi H, Sato S, Shoji T, Baba T. Immunotherapy for Uterine Cervical Cancer. *Healthcare (Basel).* 2019; 7.
10. De Nola R, Menga A, Castegna A, Loizzi V, Ranieri G, Cicinelli E, Cormio G. The Crowded Crosstalk between Cancer Cells and Stromal Microenvironment in Gynecological Malignancies: Biological Pathways and Therapeutic Implication. *Int J Mol Sci.* 2019; 20.
11. Shah W, Yan X, Jing L, Zhou Y, Chen H, Wang Y. A reversed CD4/CD8 ratio of tumor-infiltrating lymphocytes and a high percentage of CD4(+)FOXP3(+) regulatory T cells are significantly associated with clinical outcome in squamous cell carcinoma of the cervix. *Cell Mol Immunol.* 2011;8:59–66.
12. Komdeur FL, Prins TM, van de Wall S, Plat A, Wisman GBA, Hollema H, Daemen T, Church DN, de Bruyn M, Nijman HW. CD103 + tumor-infiltrating lymphocytes are tumor-reactive intraepithelial CD8 + T cells associated with prognostic benefit and therapy response in cervical cancer. *Oncoimmunology.* 2017;6:e1338230.
13. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013;4:2612.
14. Shah N, Wang P, Wongvipat J, Karthaus WR, Abida W, Armenia J, Rockowitz S, Drier Y, Bernstein BE, Long HW, Freedman ML, Arora VK, Zheng D, et al. Regulation of the glucocorticoid receptor via a BET-dependent enhancer drives antiandrogen resistance in prostate cancer. *Elife.* 2017; 6.
15. Priedigkeit N, Watters RJ, Lucas PC, Basudan A, Bhargava R, Horne W, Kolls JK, Fang Z, Rosenzweig MQ, Brufsky AM, Weiss KR, Oesterreich S, Lee AV. Exome-capture RNA sequencing of decade-old breast cancers and matched decalcified bone metastases. *JCI Insight.* 2017; 2.
16. Alonso MH, Aussó S, Lopez-Doriga A, Cordero D, Guinó E, Solé X, Barenys M, de Oca J, Capella G, Salazar R, Sanz-Pamplona R, Moreno V. Comprehensive analysis of copy number aberrations in microsatellite stable colon cancer in view of stromal component. *Br J Cancer.* 2017;117:421–31.
17. Jia D, Li S, Li D, Xue H, Yang D, Liu Y. Mining TCGA database for genes of prognostic value in glioblastoma microenvironment. *Aging.* 2018;10:592–605.

18. Wang L, Liu R, Li W, Chen C, Katoh H, Chen GY, McNally B, Lin L, Zhou P, Zuo T, Cooney KA, Liu Y, Zheng P. Somatic single hits inactivate the X-linked tumor suppressor FOXP3 in the prostate. *Cancer Cell*. 2009;16:336–46.
19. Zuo T, Liu R, Zhang H, Chang X, Liu Y, Wang L, Zheng P, Liu Y. FOXP3 is a novel transcriptional repressor for the breast cancer oncogene SKP2. *J Clin Invest*. 2007;117:3765–73.
20. Li W, Wang L, Katoh H, Liu R, Zheng P, Liu Y. Identification of a tumor suppressor relay between the FOXP3 and the Hippo pathways in breast and prostate cancers. *Cancer Res*. 2011;71:2162–71.
21. Zheng Y, Josefowicz SZ, Kas A, Chu TT, Gavin MA, Rudensky AY. Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature*. 2007;445:936–40.
22. Sakaguchi S, Sakaguchi N, Asano M, Itoh M, Toda M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J Immunol*. 1995;155:1151–64.
23. Takahashi T, Tagami T, Yamazaki S, Uede T, Shimizu J, Sakaguchi N, Mak TW, Sakaguchi S. Immunologic self-tolerance maintained by CD25(+)CD4(+) regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4. *J Exp Med*. 2000;192:303–10.
24. Wolf D, Wolf AM, Rumpold H, Fiegl H, Zeimet AG, Muller-Holzner E, Deibl M, Gastl G, Gunsilius E, Marth C. The expression of the regulatory T cell-specific forkhead box transcription factor FoxP3 is associated with poor prognosis in ovarian cancer. *Clin Cancer Res*. 2005;11:8326–31.
25. Liu R, Liu C, Chen D, Yang WH, Liu X, Liu CG, Dugas CM, Tang F, Zheng P, Liu Y, Wang L. FOXP3 Controls an miR-146/NF-κB Negative Feedback Loop That Inhibits Apoptosis in Breast Cancer Cells. *Cancer Res*. 2015;75:1703–13.
26. Yang S, Liu Y, Li MY, Ng CSH, Yang SL, Wang S, Zou C, Dong Y, Du J, Long X, Liu LZ, Wan IYP, Mok T, et al. FOXP3 promotes tumor growth and metastasis by activating Wnt/β-catenin signaling pathway and EMT in non-small cell lung cancer. *Mol Cancer*. 2017;16:124.
27. Ma GF, Miao Q, Liu YM, Gao H, Lian JJ, Wang YN, Zeng XQ, Luo TC, Ma LL, Shen ZB, Sun YH, Chen SY. High FoxP3 expression in tumour cells predicts better survival in gastric cancer and its role in tumour microenvironment. *Br J Cancer*. 2014;110:1552–60.
28. Chu R, Liu SY, Vlantis AC, van Hasselt CA, Ng EK, Fan MD, Ng SK, Chan AB, Du J, Wei W, Liu X, Liu Z, Chen GG. Inhibition of Foxp3 in cancer cells induces apoptosis of thyroid cancer cells. *Mol Cell Endocrinol*. 2015;399:228–34.
29. Ebert LM, Tan BS, Browning J, Svobodova S, Russell SE, Kirkpatrick N, Gedye C, Moss D, Ng SP, MacGregor D, Davis ID, Cebon J, Chen W. The regulatory T cell-associated transcription factor FoxP3 is expressed by tumor cells. *Cancer Res*. 2008;68:3001–9.
30. Jia H, Qi H, Gong Z, Yang S, Ren J, Liu Y, Li MY, Chen GG. The expression of FOXP3 and its role in human cancers. *Biochim Biophys Acta Rev Cancer*. 2019;1871:170–8.
31. Luo Q, Zhang S, Wei H, Pang X, Zhang H. Roles of Foxp3 in the occurrence and development of cervical cancer. *Int J Clin Exp Pathol*. 2015;8:8717–30.

32. Fu D, Liu B, Zang LE, Jiang H. MiR-631/ZAP70: A novel axis in the migration and invasion of prostate cancer cells. *Biochem Biophys Res Commun.* 2016;469:345–51.
33. Alsadeq A, Fedders H, Vokuhl C, Belau NM, Zimmermann M, Wirbelauer T, Spielberg S, Vossen-Gajcy M, Cario G, Schrappe M, Schewe DM. The role of ZAP70 kinase in acute lymphoblastic leukemia infiltration into the central nervous system. *Haematologica.* 2017;102:346–55.
34. Laufer JM, Lyck R, Legler DF. ZAP70 expression enhances chemokine-driven chronic lymphocytic leukemia cell migration and arrest by valency regulation of integrins. *Faseb J.* 2018;32:4824–35.
35. Wiestner A, Rosenwald A, Barry TS, Wright G, Davis RE, Henrickson SE, Zhao H, Ibbotson RE, Orchard JA, Davis Z, Stetler-Stevenson M, Raffeld M, Arthur DC, et al. ZAP-70 expression identifies a chronic lymphocytic leukemia subtype with unmutated immunoglobulin genes, inferior clinical outcome, and distinct gene expression profile. *Blood.* 2003;101:4944–51.
36. Crespo M, Bosch F, Villamor N, Bellosillo B, Colomer D, Rozman M, Marcé S, López-Guillermo A, Campo E, Montserrat E. ZAP-70 expression as a surrogate for immunoglobulin-variable-region mutations in chronic lymphocytic leukemia. *N Engl J Med.* 2003;348:1764–75.
37. Rassenti LZ, Huynh L, Toy TL, Chen L, Keating MJ, Gribben JG, Neuberg DS, Flinn IW, Rai KR, Byrd JC, Kay NE, Greaves A, Weiss A, et al. ZAP-70 compared with immunoglobulin heavy-chain gene mutation status as a predictor of disease progression in chronic lymphocytic leukemia. *N Engl J Med.* 2004;351:893–901.
38. Mezache L, Paniccia B, Nyinawabera A, Nuovo GJ. Enhanced expression of PD-L1 in cervical intraepithelial neoplasia and cervical cancers. *Mod Pathol.* 2015;28:1594–602.
39. Zhang L, Zhang H, Huang Y, Xi X, Sun Y. Expression of immune cell markers and tumor markers in patients with cervical cancer. *Int J Gynecol Cancer.* 2020;30:969–74.
40. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010;26:139–40.
41. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics.* 2012;16:284–7.
42. Peng D, Wang L, Li H, Cai C, Tan Y, Xu B, Le H. An immune infiltration signature to predict the overall survival of patients with colon cancer. *IUBMB Life.* 2019;71:1760–70.
43. Gui J, Li H. Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data. *Bioinformatics.* 2005;21:3001–8.
44. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33:1–22.
45. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12:453–7.
46. Bense RD, Sotiriou C, Piccart-Gebhart MJ, Haanen J, van Vugt M, de Vries EGE, Schröder CP, Fehrmann RSN. Relevance of Tumor-Infiltrating Immune Cell Composition and Functionality for Disease Outcome in Breast Cancer. *J Natl Cancer Inst.* 2017; 109.

47. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. *J Clin Oncol.* 2008;26:1364–70.
48. Long J, Wang A, Bai Y, Lin J, Yang X, Wang D, Yang X, Jiang Y, Zhao H. Development and validation of a TP53-associated immune prognostic model for hepatocellular carcinoma. *EBioMedicine.* 2019;42:363–74.

Figures

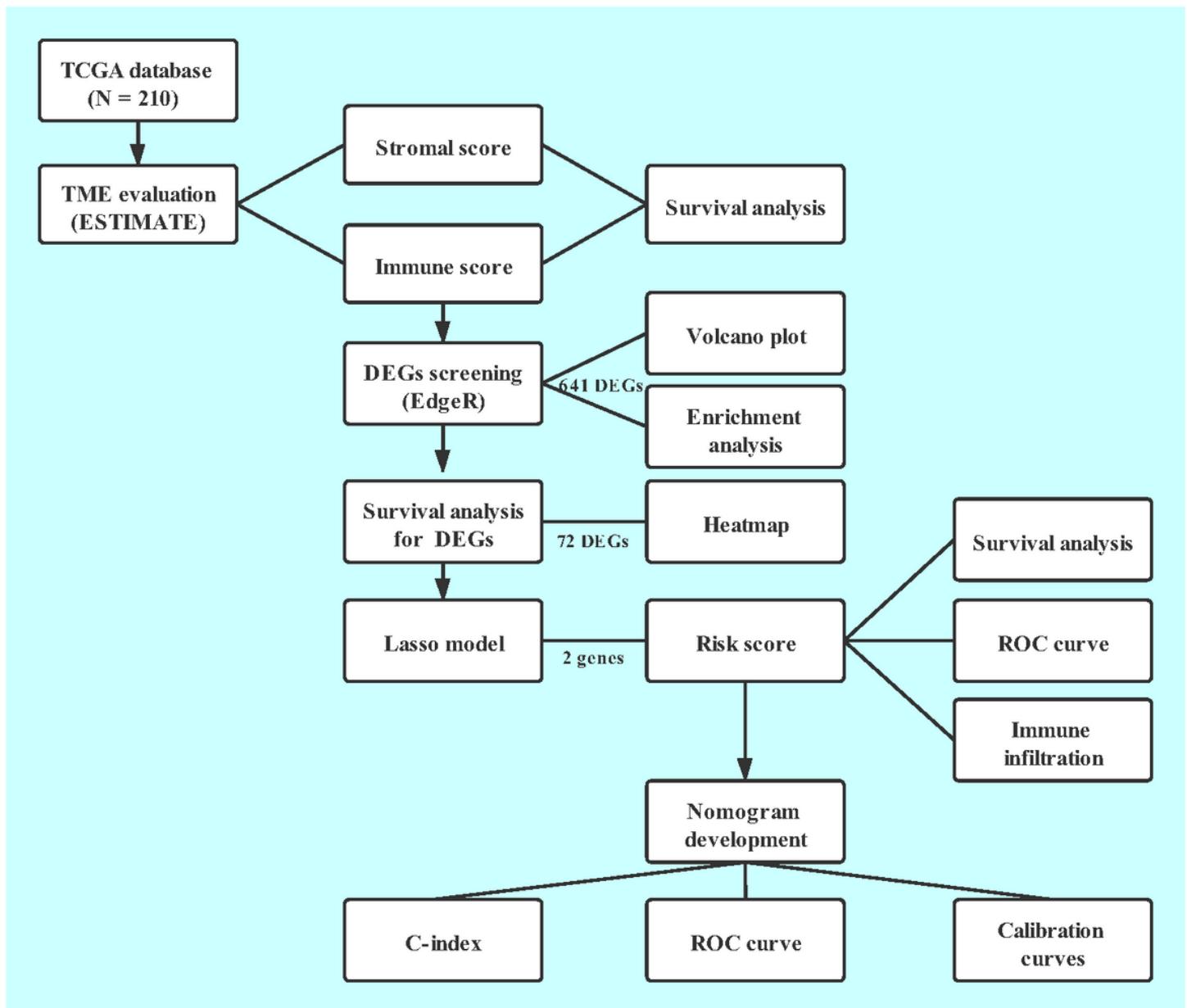


Figure 1

Flow chart of the study.

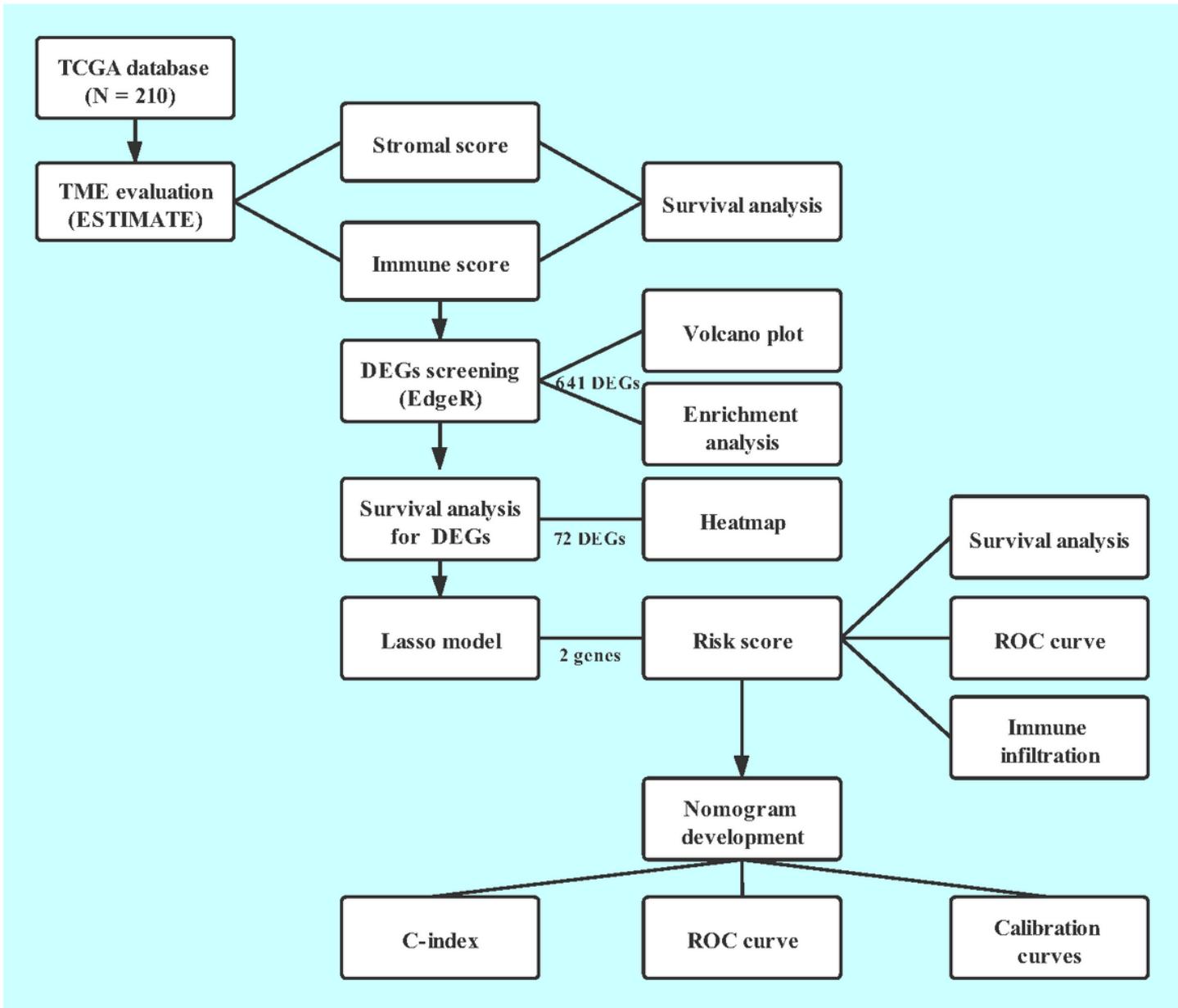


Figure 1

Flow chart of the study.

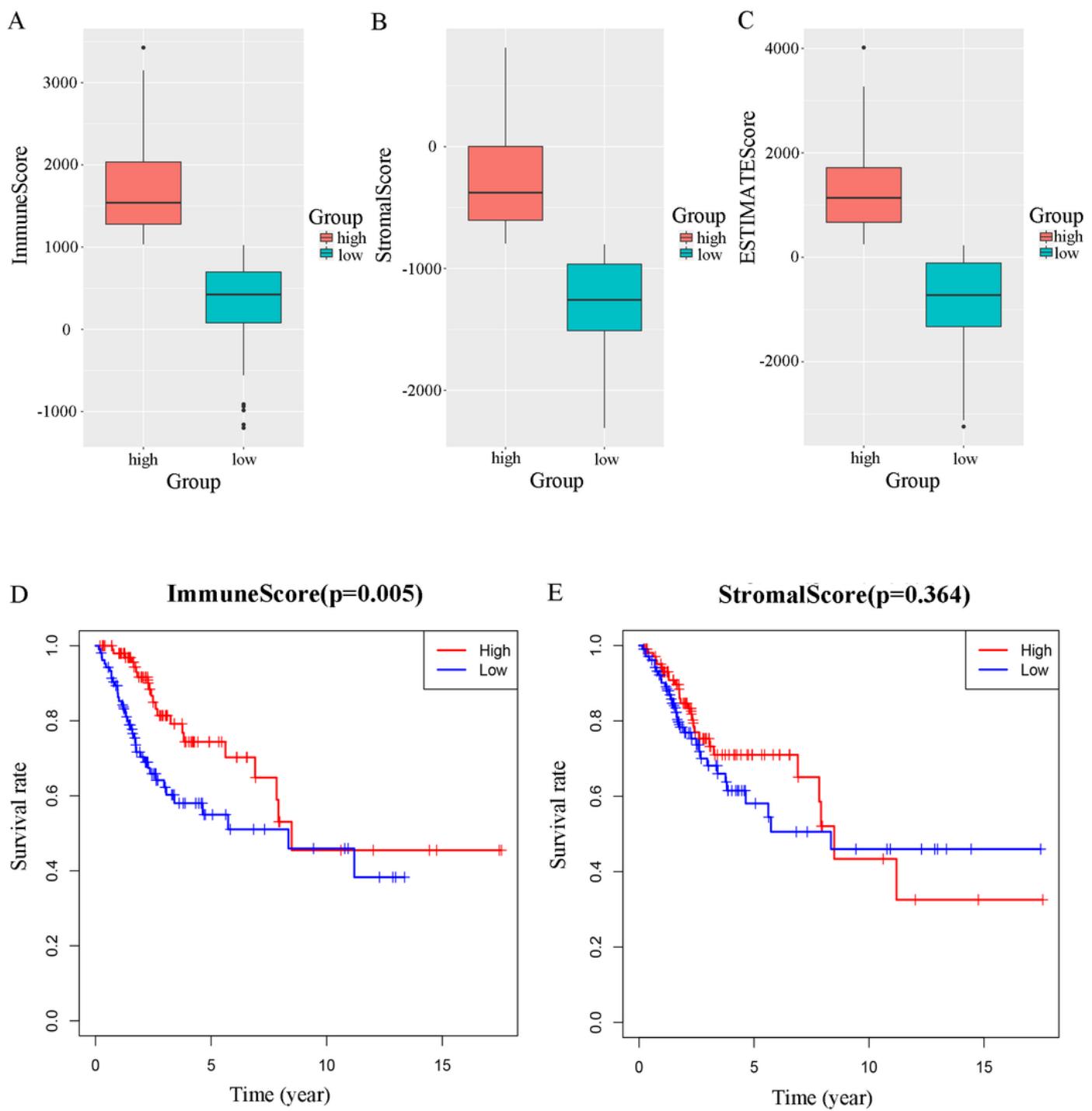


Figure 2

The distribution of the immune/stromal scores and their correlation with overall survival. (A, B, C) Immune/Stromal/Estimate scores calculated by the ESTIMATE algorithm(13). (D, E) The correlation between immune/stromal scores and overall survival in CSCC.

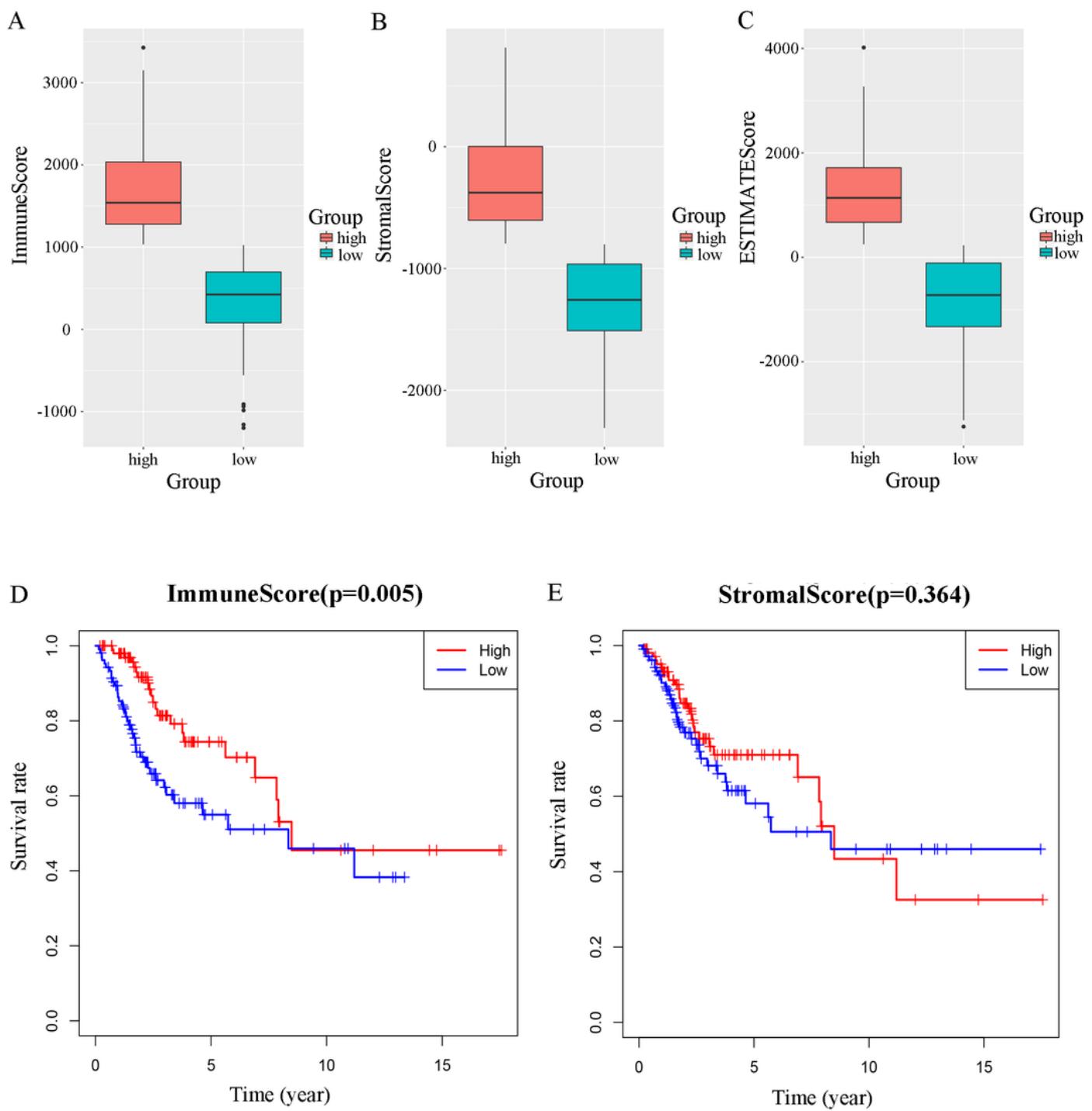


Figure 2

The distribution of the immune/stromal scores and their correlation with overall survival. (A, B, C) Immune/Stromal/Estimate scores calculated by the ESTIMATE algorithm(13). (D, E) The correlation between immune/stromal scores and overall survival in CSCC.

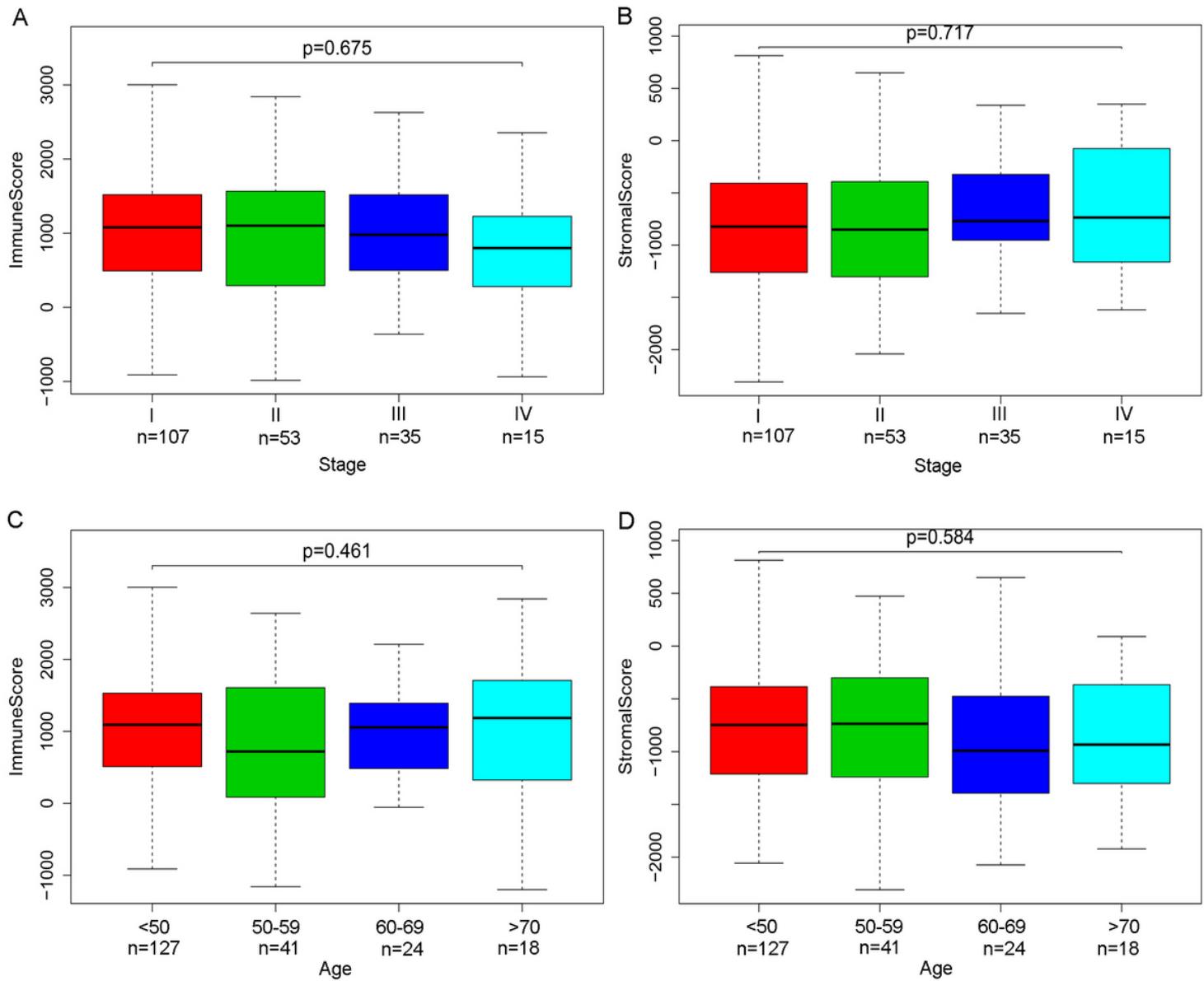


Figure 3

Correlation of the immune/stromal scores with clinical characteristics. (A, B) The distributions of immune/stromal scores did not vary with stage. (C, D) The distributions of immune/stromal scores did not vary with age.

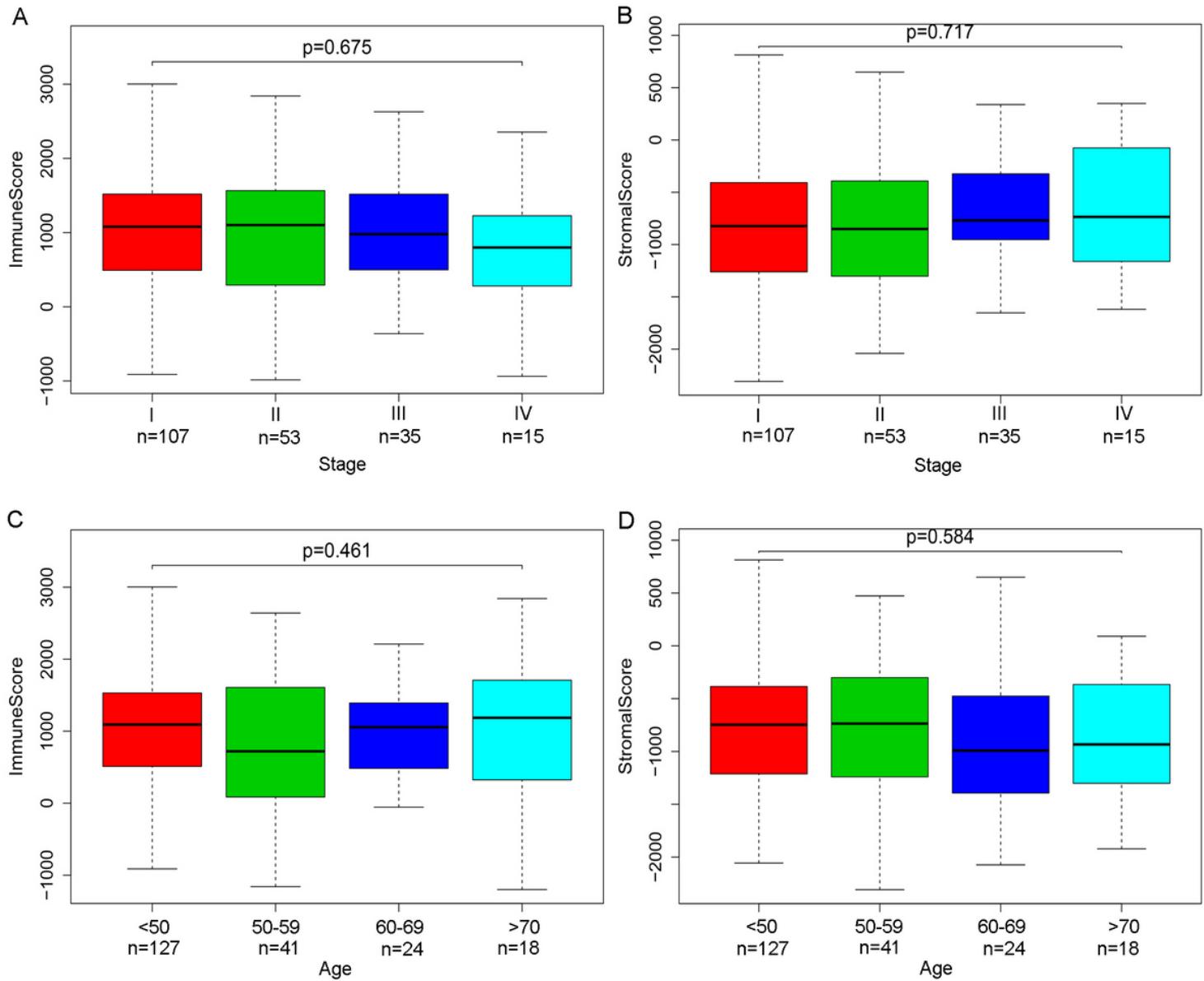


Figure 3

Correlation of the immune/stromal scores with clinical characteristics. (A, B) The distributions of immune/stromal scores did not vary with stage. (C, D) The distributions of immune/stromal scores did not vary with age.

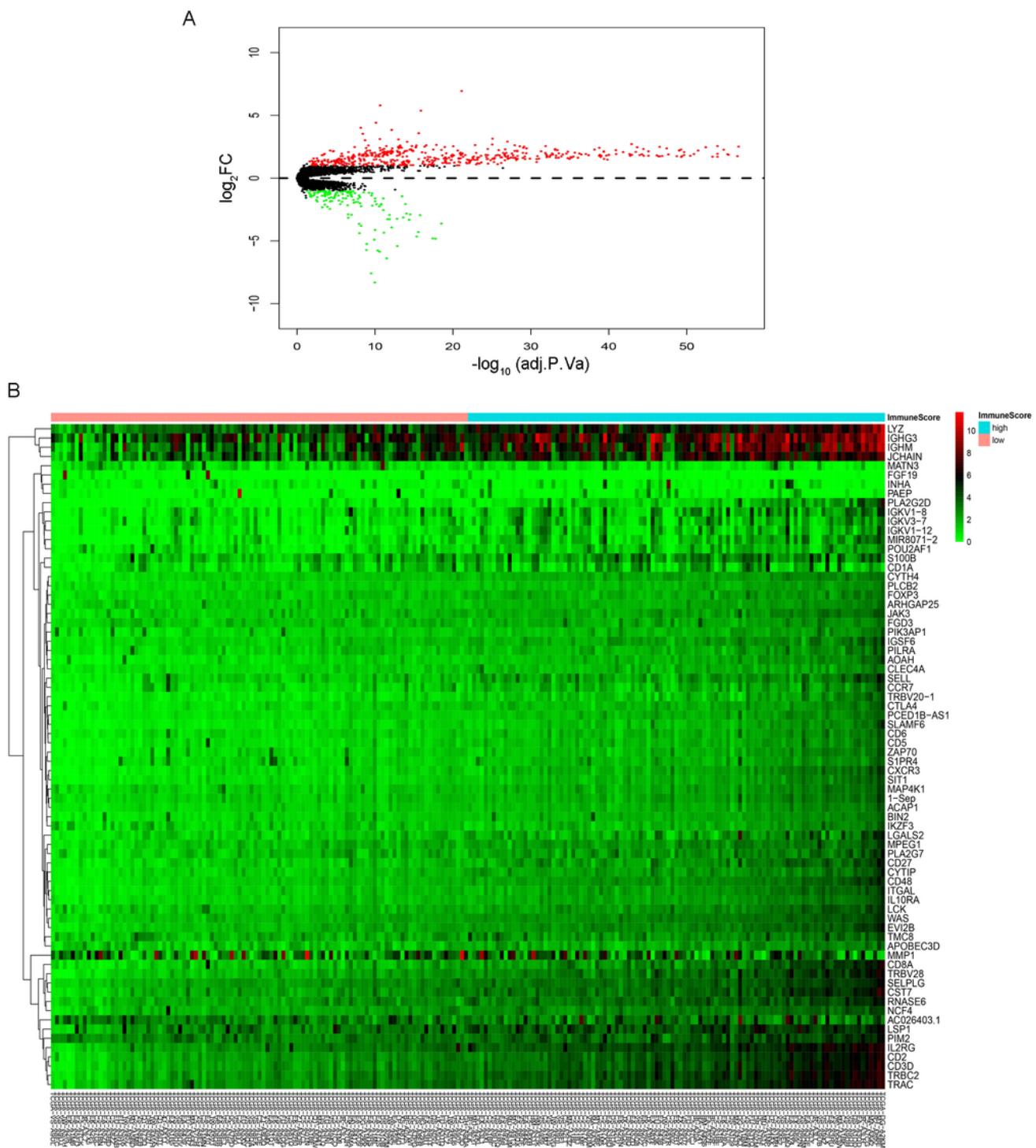


Figure 4

Comparison of gene expression profile in high- and low- immune/stromal scores group. (A) Volcano plot of 641 DEGs based on immune scores. (B) Heatmap of the 72 prognostic DEGs.

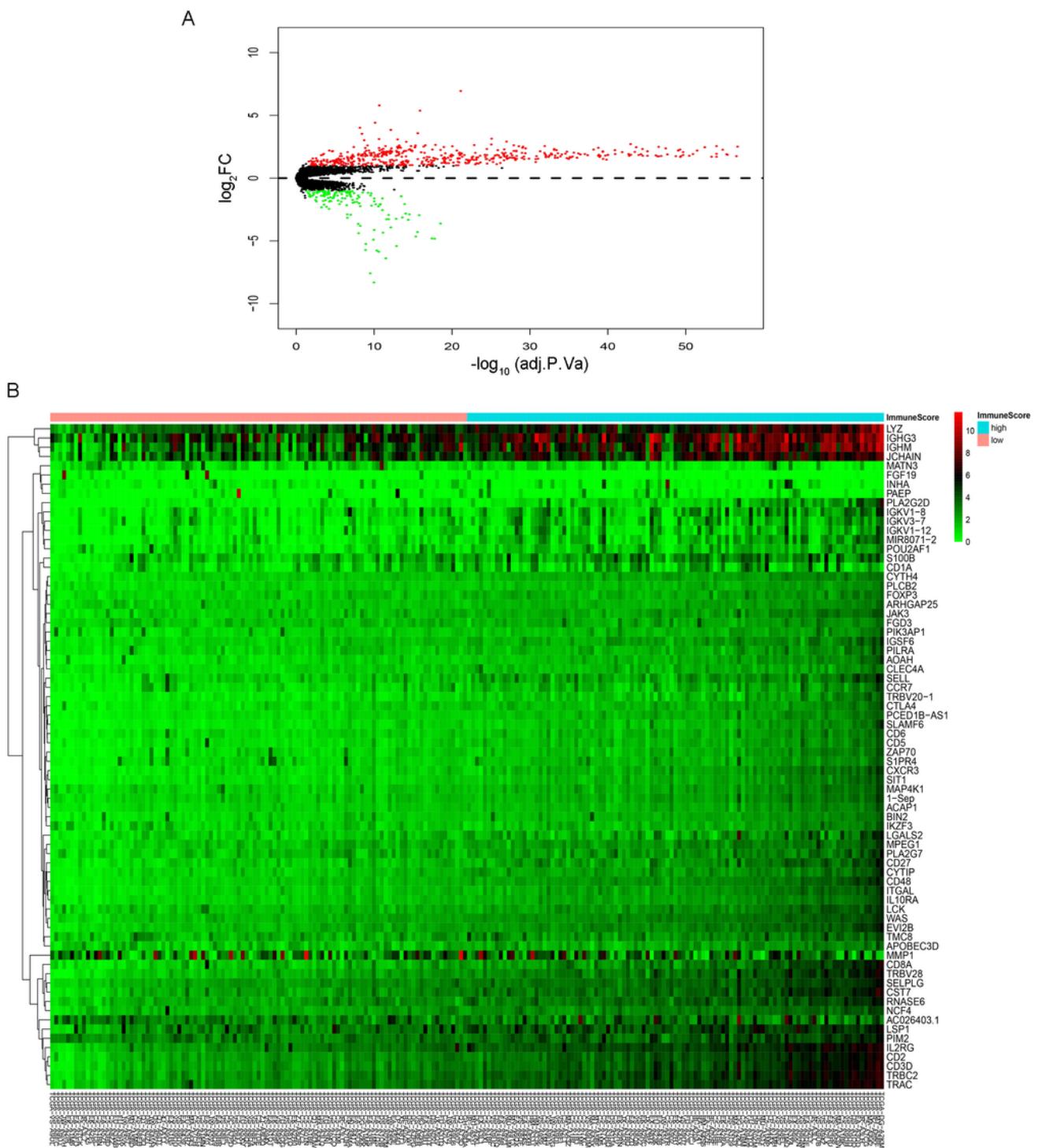
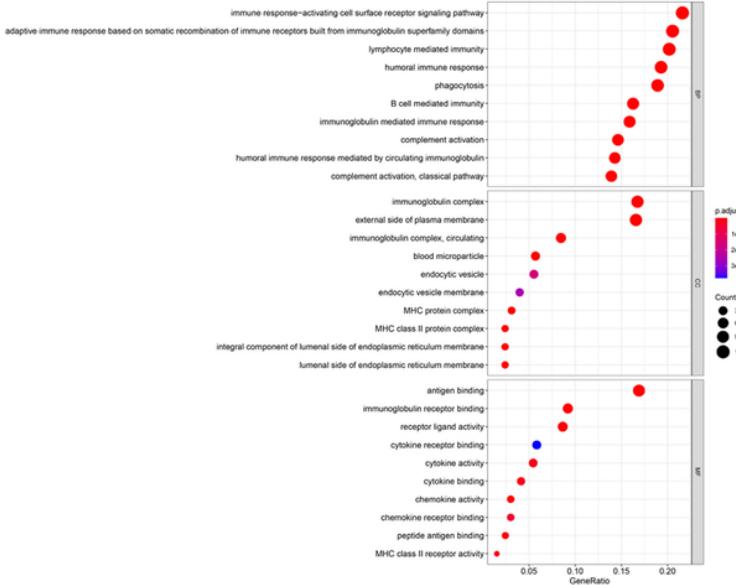


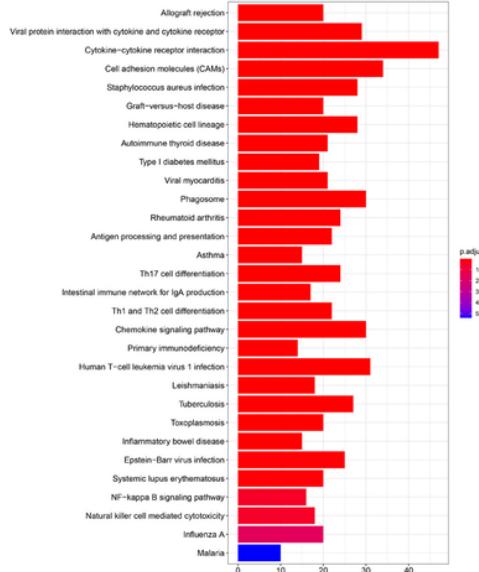
Figure 4

Comparison of gene expression profile in high- and low- immune/stromal scores group. (A) Volcano plot of 641 DEGs based on immune scores. (B) Heatmap of the 72 prognostic DEGs.

A

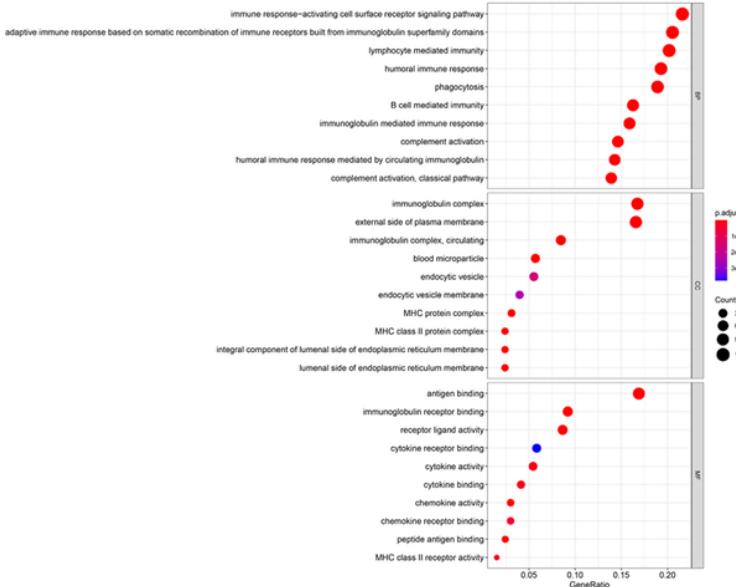


B

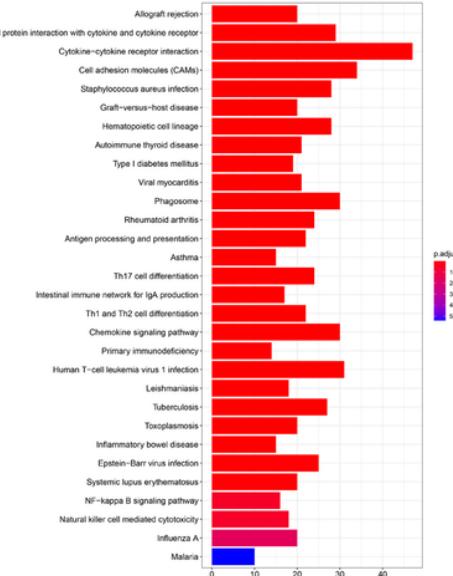
**Figure 5**

GO analysis and KEGG pathway analysis of the 641 DEGs based on immune scores. (A) GO enrichment analyses of the 641 DEGs. (B) KEGG pathway enrichment analysis of the 641 DEGs.

A



B

**Figure 5**

GO analysis and KEGG pathway analysis of the 641 DEGs based on immune scores. (A) GO enrichment analyses of the 641 DEGs. (B) KEGG pathway enrichment analysis of the 641 DEGs.

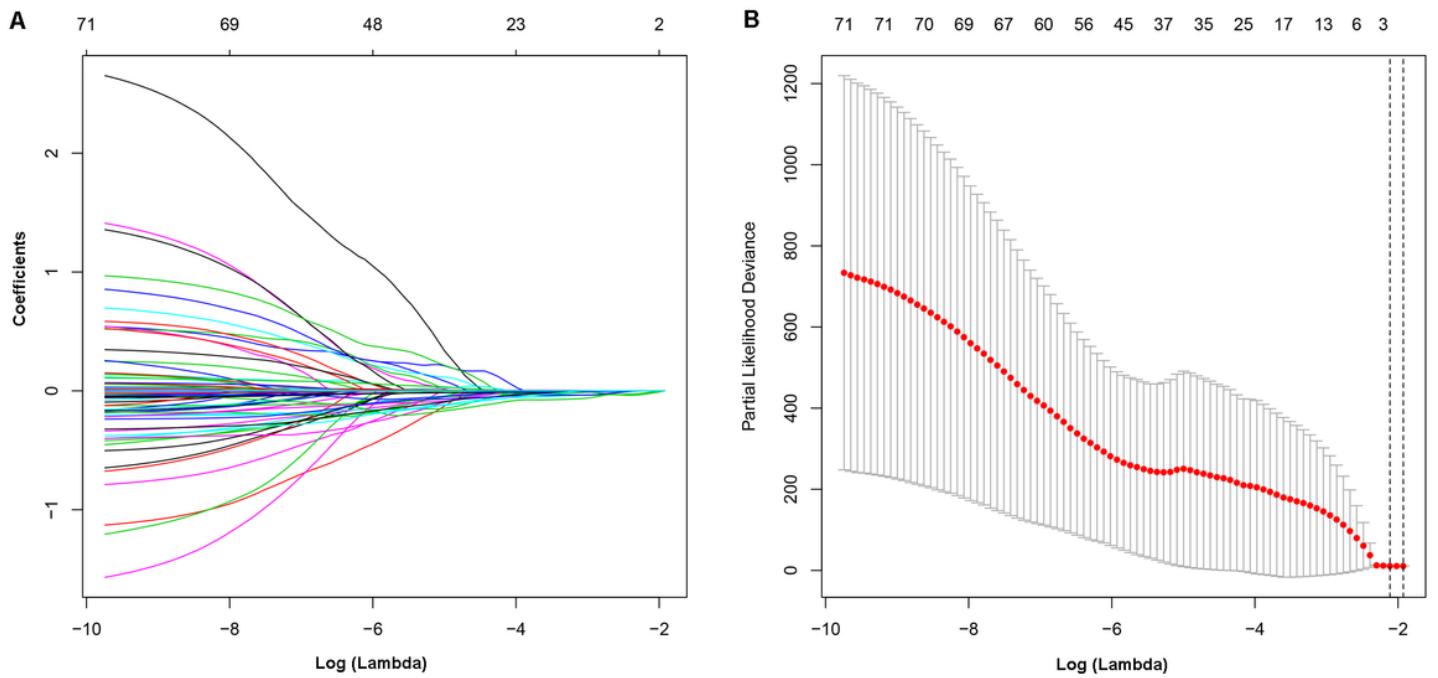


Figure 6

LASSO profiles of the 72 prognostic DEGs in CSCC. (A) LASSO coefficient profiles of the 72 prognostic DEGs in CSCC. (B) Lasso deviance profiles of the 72 prognostic DEGs in CSCC.

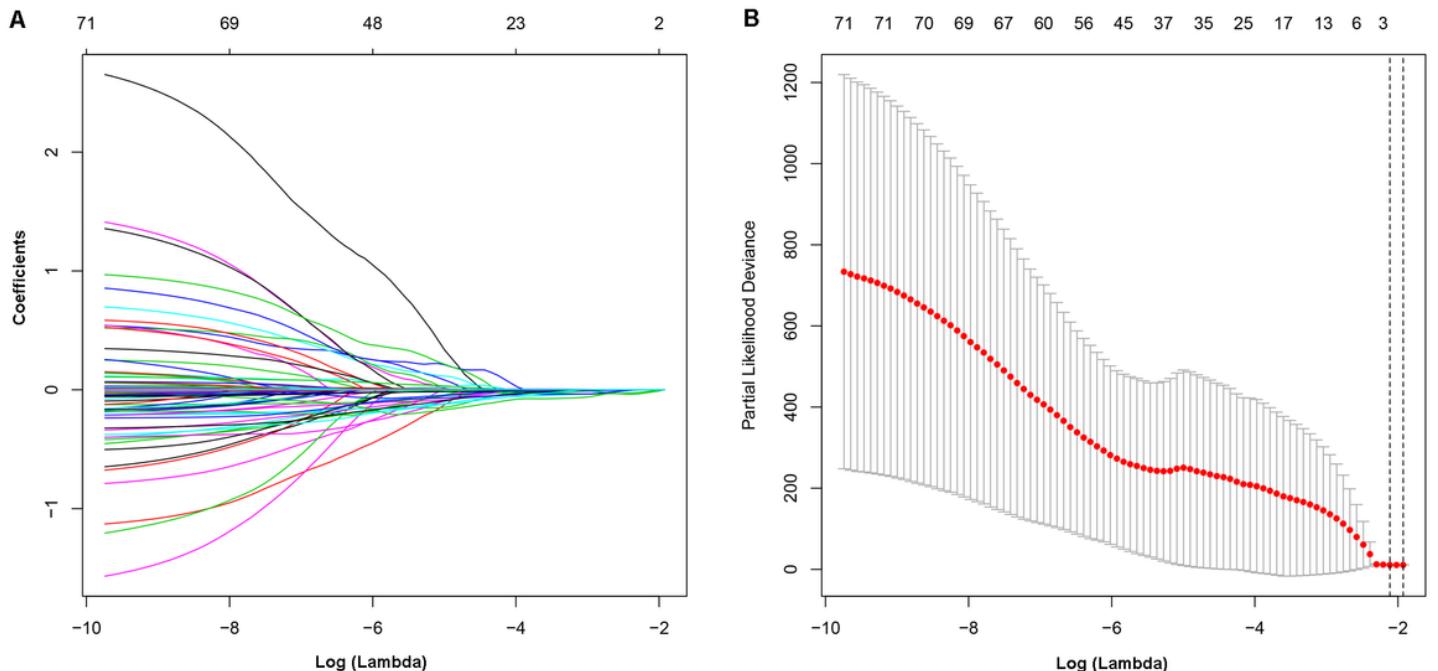


Figure 6

LASSO profiles of the 72 prognostic DEGs in CSCC. (A) LASSO coefficient profiles of the 72 prognostic DEGs in CSCC. (B) Lasso deviance profiles of the 72 prognostic DEGs in CSCC.

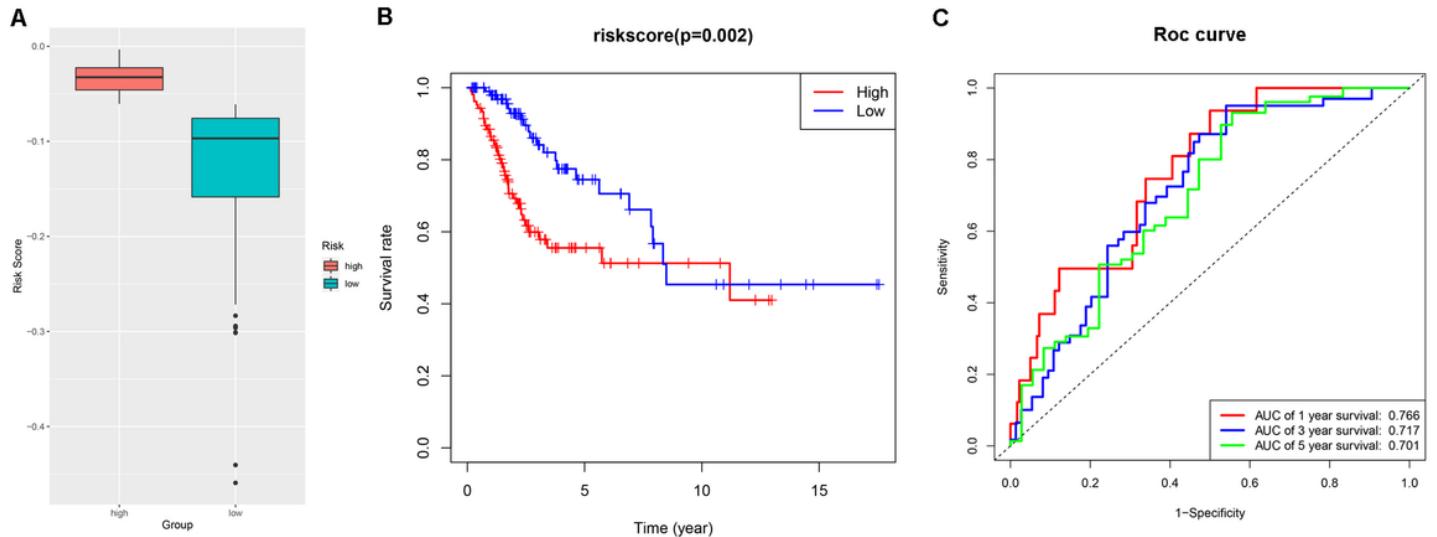


Figure 7

Prognostic analysis of the riskscore model. (A) The distribution of the riskscore. (B) K-M analysis of the riskscore. (C) ROC curves of the riskscore model.

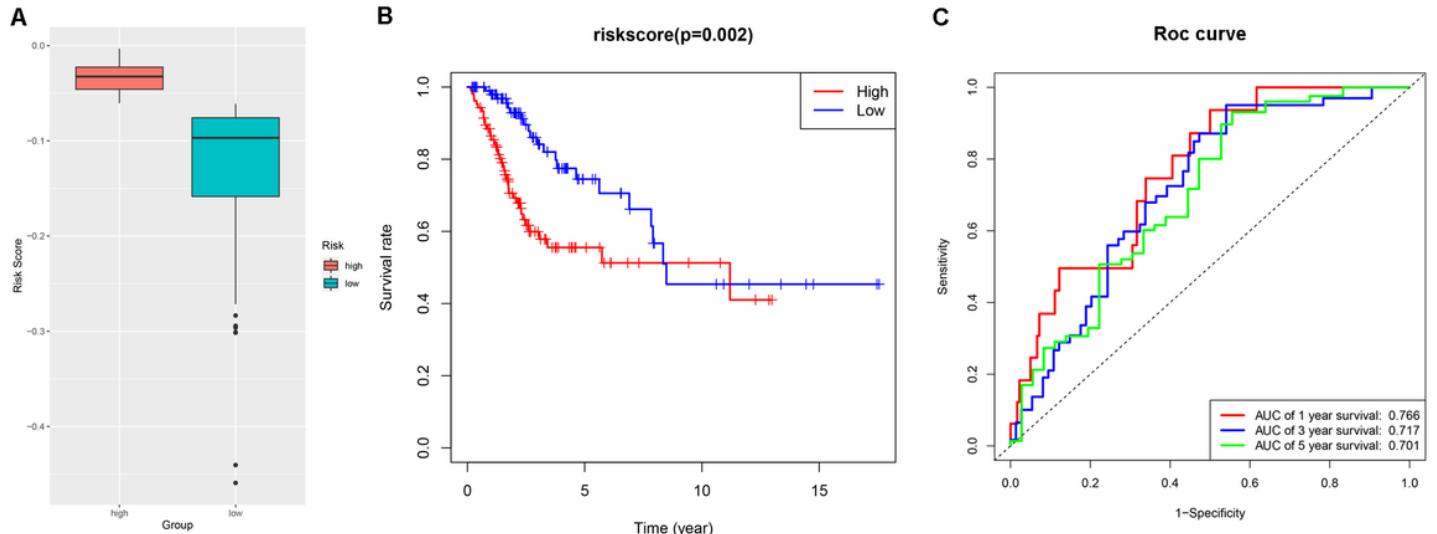


Figure 7

Prognostic analysis of the riskscore model. (A) The distribution of the riskscore. (B) K-M analysis of the riskscore. (C) ROC curves of the riskscore model.

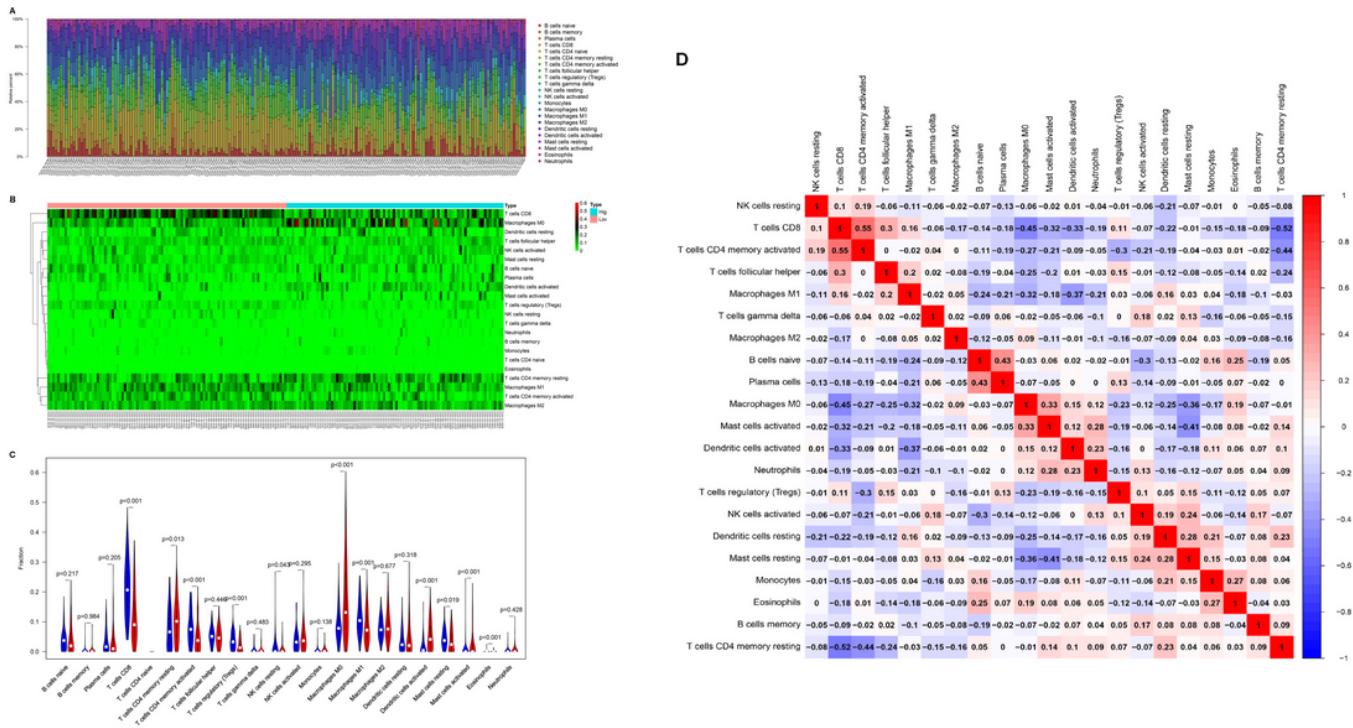


Figure 8

Landscape of immune infiltration in high- and low-risk CSCC patients. (A) The barplot showed the proportion of 22 immune cells infiltration in all patients. (B) Heat map of the 22 immune cell fractions in high- and low-risk patients. (C) Violin plots visualizing significantly different immune cells between high- and low-risk patients. (D) Correlation matrix of immune cells infiltration.

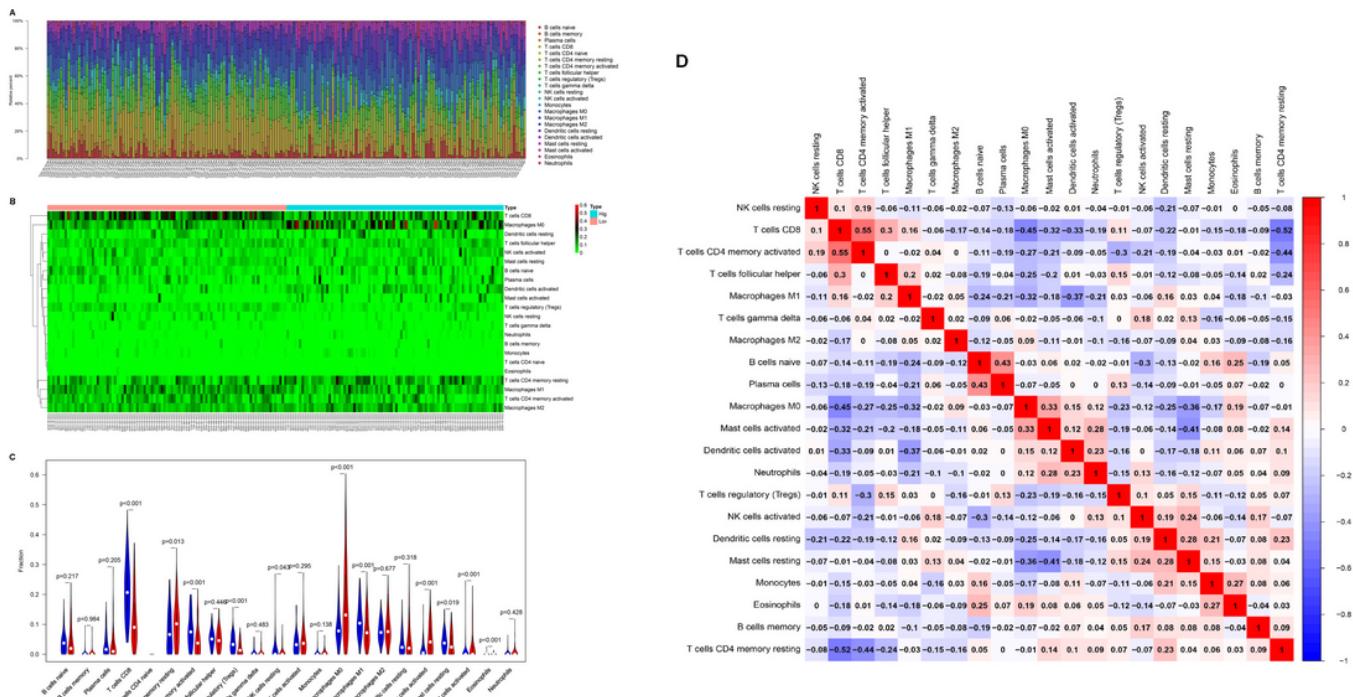


Figure 8

Landscape of immune infiltration in high- and low-risk CSCC patients. (A) The barplot showed the proportion of 22 immune cells infiltration in all patients. (B) Heat map of the 22 immune cell fractions in high- and low-risk patients. (C) Violin plots visualizing significantly different immune cells between high- and low-risk patients. (D) Correlation matrix of immune cells infiltration.

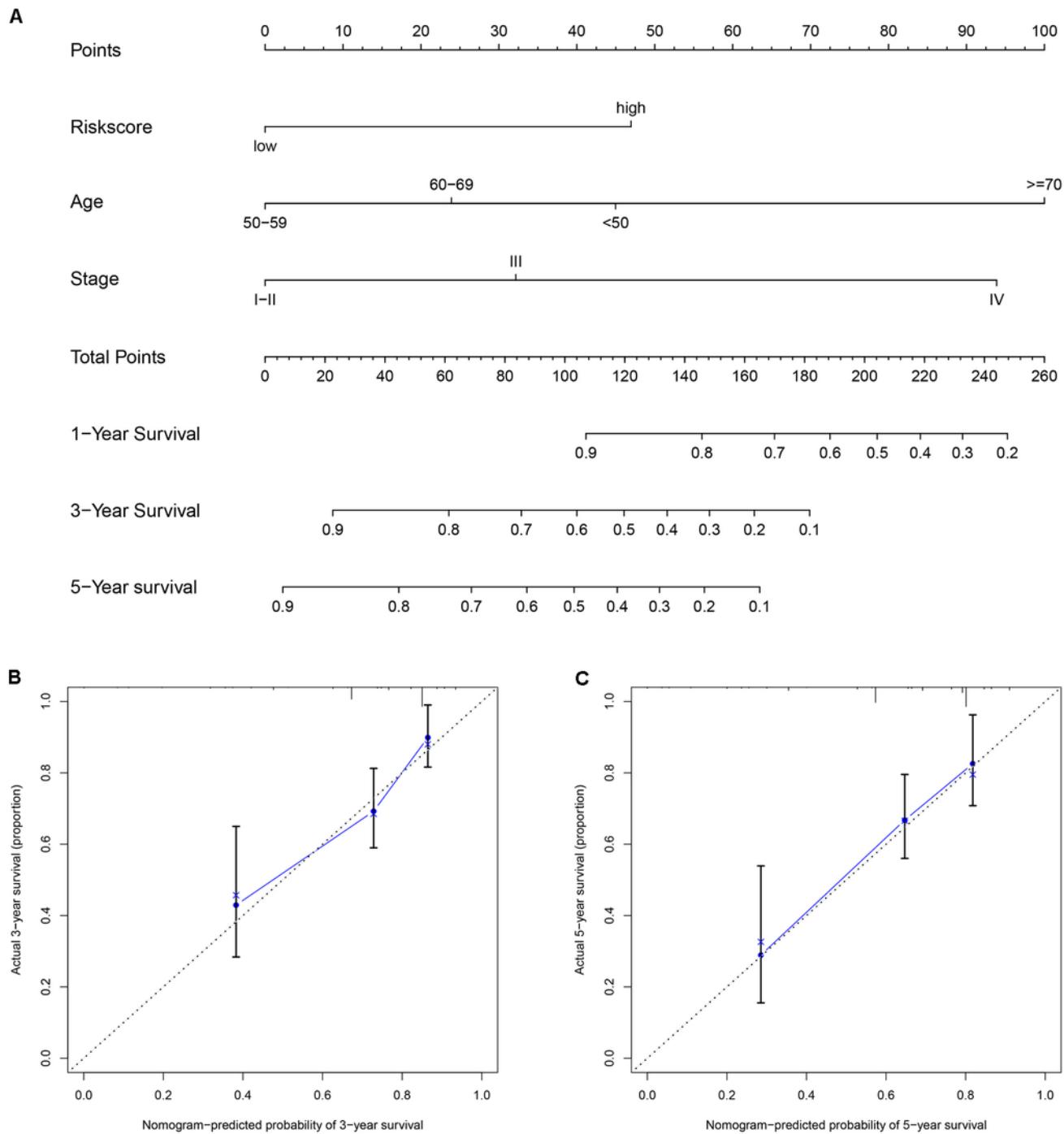


Figure 9

The construction of a nomogram based on riskscore, age and stage. (A) Nomogram for predicting the probability of 1-, 3-, and 5-year overall survival for cervical squamous cell carcinoma. (B,C) Calibration plots of the nomogram for predicting the probability of OS at 3 and 5 year.

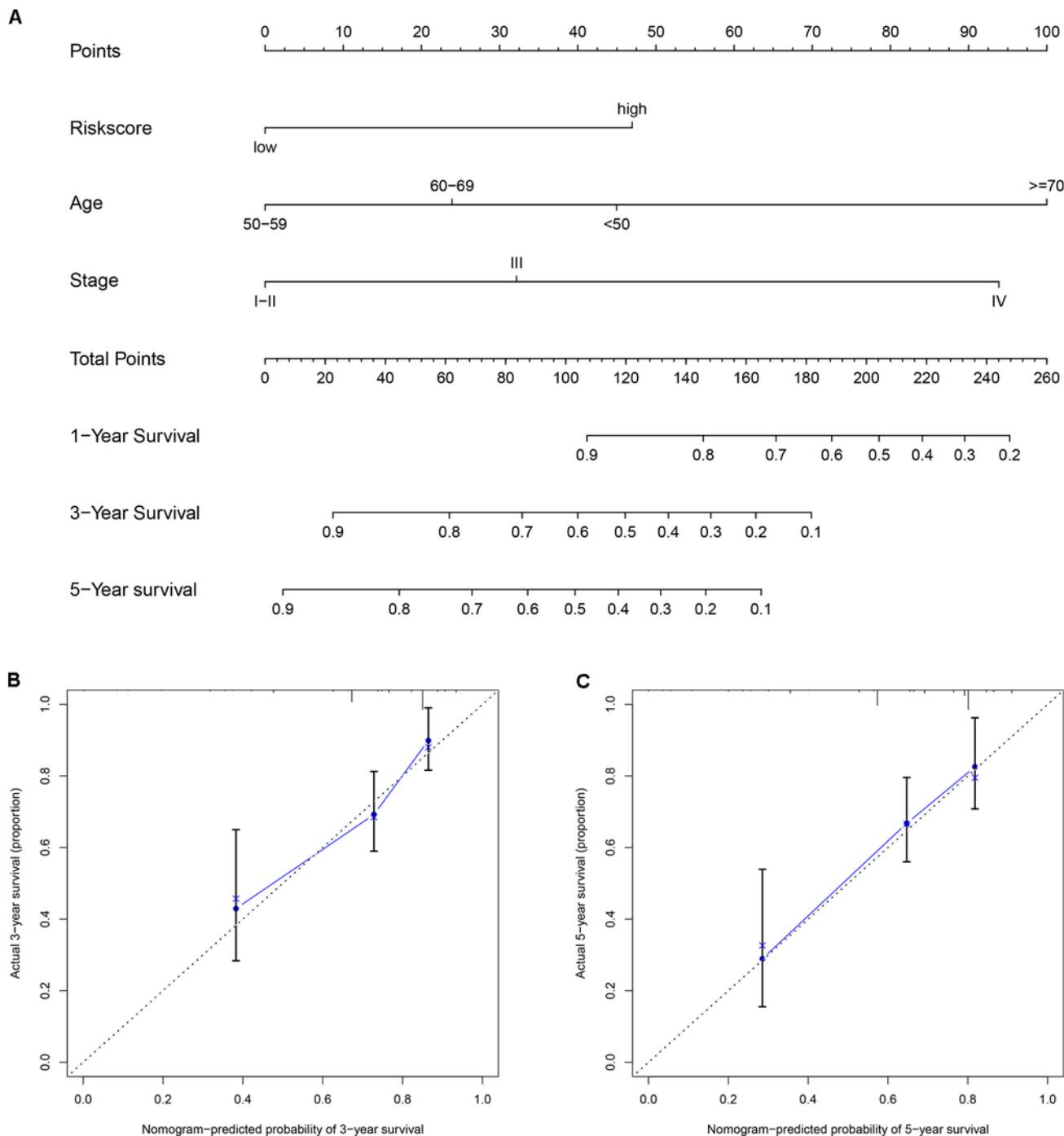


Figure 9

The construction of a nomogram based on riskscore, age and stage. (A) Nomogram for predicting the probability of 1-, 3-, and 5-year overall survival for cervical squamous cell carcinoma. (B,C) Calibration plots of the nomogram for predicting the probability of OS at 3 and 5 year.

plots of the nomogram for predicting the probability of OS at 3 and 5 year.

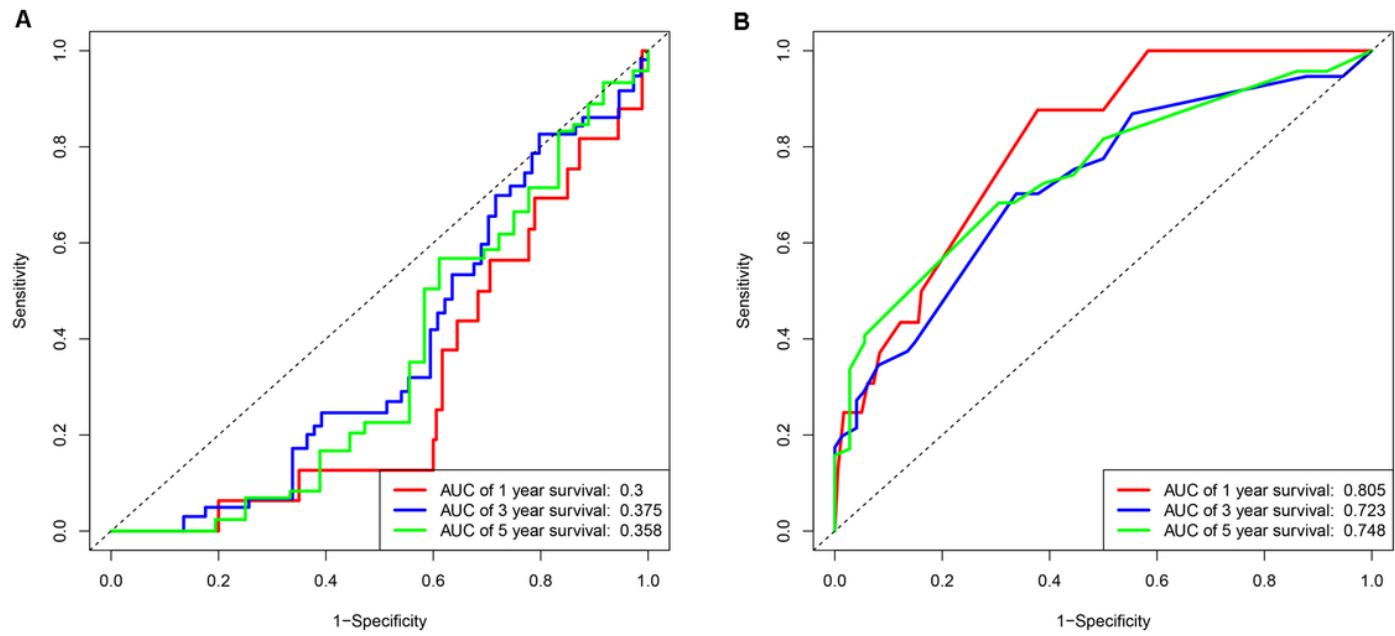


Figure 10

Comparison of ROC curve analyses of nomogram and immune score. (A) ROCcurve analyses of the nomogram. (B) ROC curve analyses of immune score.

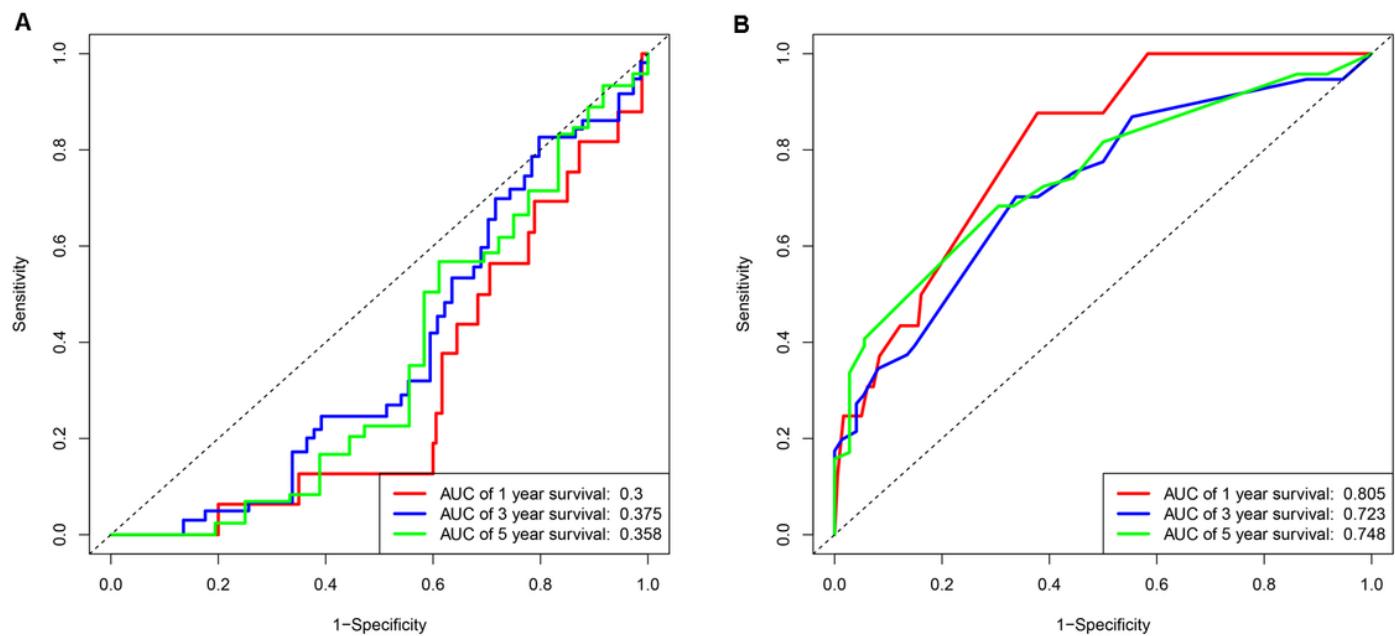


Figure 10

Comparison of ROC curve analyses of nomogram and immune score. (A) ROCcurve analyses of the nomogram. (B) ROC curve analyses of immune score.

Supplementary Files

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

- [Additionalfile1.docx](#)
- [Additionalfile1.docx](#)