Check for updates

A Five-MicroRNA Signature Predicts the Prognosis in Nasopharyngeal Carcinoma

Shixiong Wu, Cen Zhang, Jing Xie, Shuang Li* and Shuo Huang*

Department of Otolaryngology, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: There is no effective prognostic signature that could predict the prognosis of nasopharyngeal carcinoma (NPC).

Methods: We constructed a prognostic signature based on five microRNAs using random forest and Least Absolute Shrinkage And Selection Operator (LASSO) algorithm on the GSE32960 cohort ($N = 213$). We verified its prognostic value using three independent external validation cohorts (GSE36682, $N = 62$; GSE70970, $N = 246$; and TCGA-HNSC, $N = 523$). Through principal component analysis, receiver operating characteristic curve analysis, and C-index calculation, we confirmed the predictive accuracy of this prognostic signature.

Results: We calculated the risk score based on the LASSO algorithm and divided the patients into high- and low-risk groups according to the calculated optimal cutoff value. The patients in the high-risk group tended to have a worse prognosis outcome and chemotherapy response. The time-dependent receiver operating characteristic curve showed that the 1-year overall survival rate of the five-microRNA signature had an area under the curve of more than 0.83. A functional annotation analysis of the five-microRNA signature showed that the patients in the high-risk group were usually accompanied by activation of DNA repair and MYC-target pathways, while the patients in the low-risk group had higher immune-related pathway signals.

Conclusions: We constructed a five-microRNA prognostic signature, which could accurately predict the prognosis of nasopharyngeal carcinoma, and constructed a nomogram that could conveniently predict the overall survival of patients.

Keywords: nasopharyngeal carcinoma, microRNA, prognosis, chemotherapy response, predicts

INTRODUCTION

Nasopharyngeal cancer (NPC) is a rare tumor, with the highest incidence in southeast Asia, such as southeast China, India, and Thailand [1]. The main cause of treatment failure in NPC patients was the occurrence of local metastatic events, in which more than 70% of patients presented local progression [2]. An accurate prognosis prediction can guide the selection of treatment options for patients with NPC. Currently, the tumor lymph node metastasis (TNM) system is still a widely recognized system for predicting prognosis [3]. However, previous studies have found that, among

NPC patients with the same tumor stage, even after receiving similar treatment regimens, the treatment outcomes are quite different (4). The current TNM staging system only focuses on the anatomical features of the tumor and ignores the biological heterogeneity of the tumor so that it cannot accurately predict the prognosis of individual patients (5). Therefore, it is urgent to find a new biomarker that could predict the prognosis independently or in combination with the TNM staging system so as to improve the accuracy of the prediction.

MicroRNA plays a significant role in the development of cancer and its therapeutic response. Recent studies have demonstrated the value of some microRNA in predicting the prognosis of NPC, and some microRNA signatures have been constructed to predict the prognosis of NPC (6–8). Studies have reported that Epstein–Barr virus (EBV)-related microRNA have potentially important effects on NPC. The in-depth study of microRNA encoded by EBV has potential significance and importance for the discovery of prognostic markers of NPC (9).

At present, Least Absolute Shrinkage and Selection Operator (LASSO) has been widely used to construct the prognostic model (10–12). Here we developed and validated a prognostic signature based on five microRNAs in four large cohorts (GSE32960, GSE36682, GSE70970, and TCGA-HNSC)—a total of 1,143 patients with NPC. The random forest and LASSO Cox regression algorithm were used to select the characteristics and quantify the Cox regression coefficient, to build a scoring model, which could not only accurately predict the prognosis and progress of NPC. An accurate nomogram was provided for clinicians to predict the prognosis of NPC patients.

MATERIALS AND METHODS

Cohort Information and Data Preprocessing

We selected four large microRNA microarray/sequencing cohorts (GSE32960: $N = 312$, GSE36682: $N = 62$, GSE70976 (13): $N = 246$, and TCGA-HNSC: $N = 523$) from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) and The Cancer Genome Atlas (TCGA; https://cancergenome. nih.gov/) database, all of which contained detailed prognostic follow-up data. The cohort (GSE32960) with the largest number of patients in the GEO database was selected as the training cohort, and the other three cohorts (GSE36682, GSE70976, and TCGA-HNSC) were selected as independent external validation cohorts. The GSE32960 cohort contained 312 non-metastatic NPC tissues and 18 non-cancerous nasopharyngitis tissues. The GSE36682 cohort contained 62 cases of NPC and six cases of normal tissues. The GSE70976 cohort contained 246 cases of NPC and 17 cases of normal tissues. We retained only NPC samples with prognostic follow-up data. The level 3 microRNA sequencing data of head and neck squamous cell carcinomas were downloaded from the UCSC Xena website (https://gdc. xenahubs.net/download/TCGA-HNSC.mirna.tsv.gz) based on the TCGA database. The cohort contained 523 head and neck squamous cell carcinoma samples (including NPC). The data

normalization method adopted reads per million mapped reads. All data were $\log2(x + 1)$ transformed. The baseline information for all cohorts used in this study is listed in **Table 1**, and the flow chart of the experiment is shown in **Figure 1**.

Construct and Validate the Five-MicroRNA Signature

We first used the random forest algorithm to feature select all microRNA and found the microRNA which had the confirmed ability to distinguish the overall survival outcome (dead vs. alive). Next, we performed the LASSO Cox regression analysis using the R package "glmnet" (14) on training cohort GSE32960. The LASSO algorithm can reduce the dimensionality of high-latitude data and explain the characteristics of the data with a model with fewer variables (15). Tenfold cross-validation was used to avoid overfitting the model built with the training cohort. Finally, we

TABLE 1 | Baseline information for the cohort that we used.

Series accession numbers	GSE32960	GSE36682	GSE70970	TCGA-HNSC
Platform	MicroRNA array	Human miRNA 1K	nCounter® Human miRNA assay (v1.0, Nanostring)	Illumina miRNAseq
Number of patients	312	62	246	523
T				
Tx	–	–	2	12
T1	66	–	74	38
T2	89	–	50	149
T3	71	–	52	140
T4	86	–	67	184
N				
Nx	–	–	1	19
N0	44	–	49	245
N1	148	–	83	85
N2	72	–	90	165
N3	48	–	23	9
M				
Mx	–	–	–	20
M0	–	–	–	497
M1	–	–	–	6
Stage				
I	12	–	–	27
II	86	–	–	84
III	91	–	–	92
IV	123	–	–	320
Age				
≥65	–	26	35	196
<65	–	36	211	326
Chemotherapy				
Yes	268	–	126	169
No	44	–	120	342
Unknown/ not available	–	–	–	17
Survival outcome	Overall survival (OS), metastasis-free survival (MFS), disease-free survival (DFS), relapse-free survival (RFS)	OS	OS, MFS, DFS, RFS, DSS	OS, MFS

FIGURE 1 | Flow chart of the study. PCA, principal component analysis; OS, overall survival; DFS, disease-free survival; MFS, metastasis-free survival; DSS, disease-free survival; RFS, relapse-free survival.
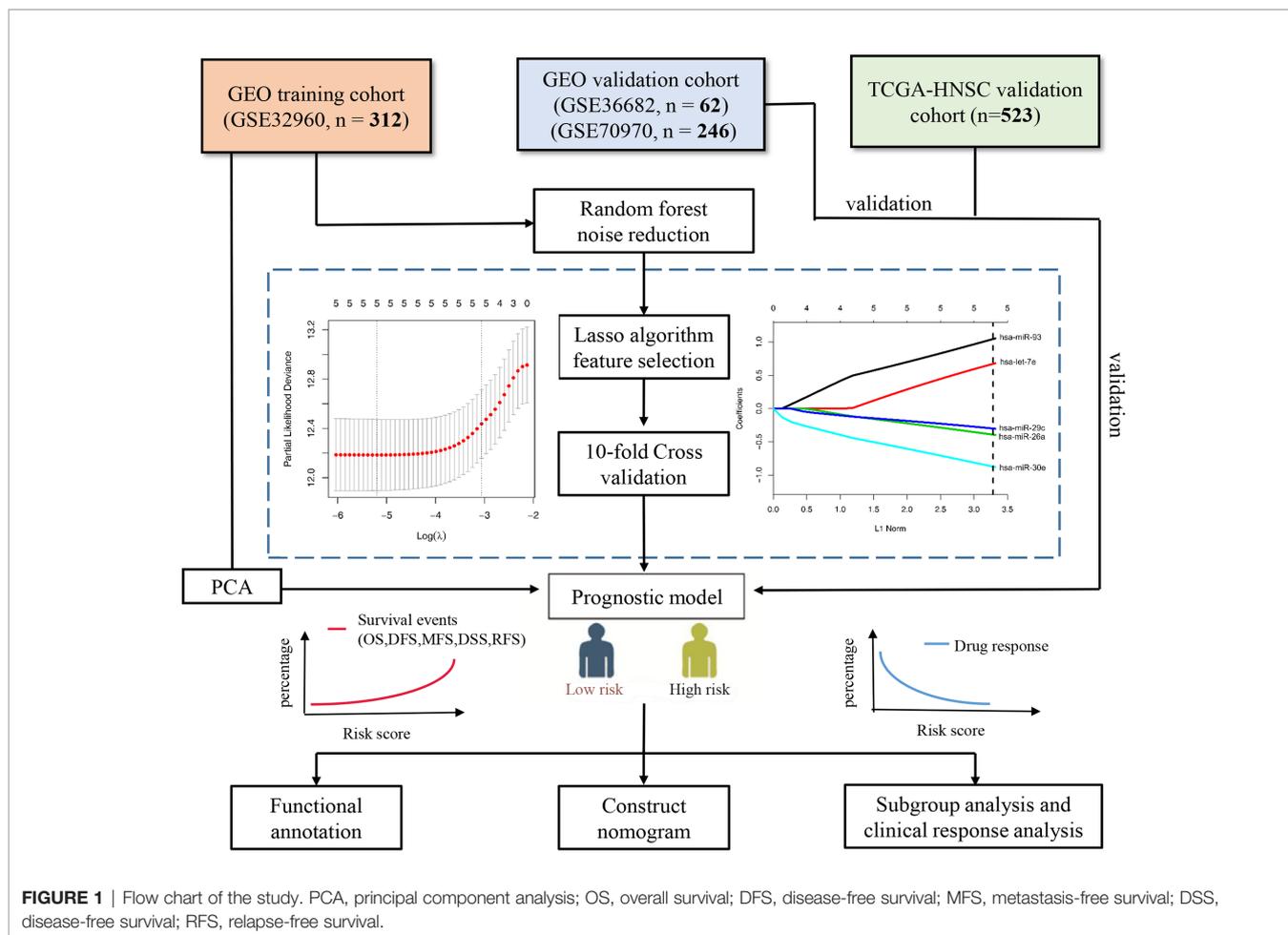
constructed a scoring system based on the regression coefficients calculated by the LASSO cox regression analysis. We divided all NPC patients into high- and low-risk groups according to the cutoff values given by R package "survminer". The Kaplan–Meier survival curve was used to explore the prognostic value of the five-microRNA signature, the time-dependent receiver operating characteristic (ROC) curve was used to measure the prediction accuracy, and principal component analysis (PCA) was used to observe the ability of the model to distinguish the overall survival outcome events. The prognostic value of the five microRNAs was validated by three independent external validation cohorts (GSE36682, GSE70970, and TCGA-HNSC).

Functional Annotation of the Five-MicroRNA Signature

To carry out functional annotation with the five-microRNA signature, we downloaded RNA-seq data from the TCGA database. By calculating the correlation coefficient between risk score and all genes, the top 500 genes with the highest correlation were extracted for Gene Ontology, Kyoto Encyclopedia of Genes and Genomes, and gene set enrichment analysis (GSEA). Enrichment analysis and visualization were performed by the R package "ClusterProfiler" (16).

Statistical Analysis

For the statistical test of the two groups of data, if the data characteristics meet the normal distribution, the t-test was used; if the data characteristics do not meet the normal distribution, the non-parametric test (Wilcoxon test) was chosen. The R package "rms" was used to construct a nomogram. The R package "rmda" was used to perform decision curve analysis (DCA) to evaluate the net benefits of the prediction signature (17). P-value < 0.05 was considered statistically significant.

RESULTS

Feature Selection and Prognostic Signature Construction

We used random forest algorithm to select the features of all microRNAs in the training cohort GSE32960. Random forest is an algorithm that can effectively reduce the dimension and noise of high-latitude data (18). After selection, only five microRNAs (has-miR-93, has-let-7e, has-miR29c, has-miR-26a, and has-miR-30e) were selected as candidate features. Subsequently, we used the LASSO Cox regression algorithm to construct a prognostic signature based on the five microRNAs (**Figure 1**).

The calculation method of the scoring model is as follows: risk score = 1.088 * hsa-miR-93 + 0.673 * hsa-let-7e + (-0.299) * hsa-miR-29c + (-0.395) * hsa-miR-26a + (-0.927) * hsa-miR-30e. The gene symbols of microRNA in the formula represent the transcriptional expression of the microRNA.

The Prognostic Value of the Five-MicroRNA Signature Was Validated in Multiple Cohorts

According to the cutoff value given by R package "survminer" (19), the NPC patients in all the cohorts were divided into high-risk and low-risk groups. In the training cohort GSE32960, our results indicate that the five-microRNA signature has a high predictive value for overall survival (OS; hazard ratio, HR = 5.54; 95% confidence interval, CI: 2.58–11.89, $p < 0.0001$), metastasis-free survival (MFS; HR = 4.91, 95%CI: 2.23–10.81, $p < 0.0001$), disease-free survival (DFS; HR = 4.24, 95%CI: 2.16–8.33, $p < 0.0001$), and relapse-free survival (RFS; HR = 2.52, 95%CI: 0.96–6.61, $p = 0.0077$) of NPC (**Figures 2A–D**). In the external validation cohort GSE36682, we found that the five-microRNA signature has a high predictive value for the OS (HR = 2.87, 95% CI: 1.24–6.63, $p = 0.0299$) of NPC (**Supplementary Figure S1A**). In the external validation cohort GSE70970, we found that the five-microRNA signature has a high predictive value for OS (HR = 2.32, 95%CI: 1.33–4.03, $p = 0.0003$), MFS (HR = 2.05, 95% CI: 0.95–4.42, $p = 0.0337$), DFS (HR = 1.90, 95%CI: 1.16–3.10,
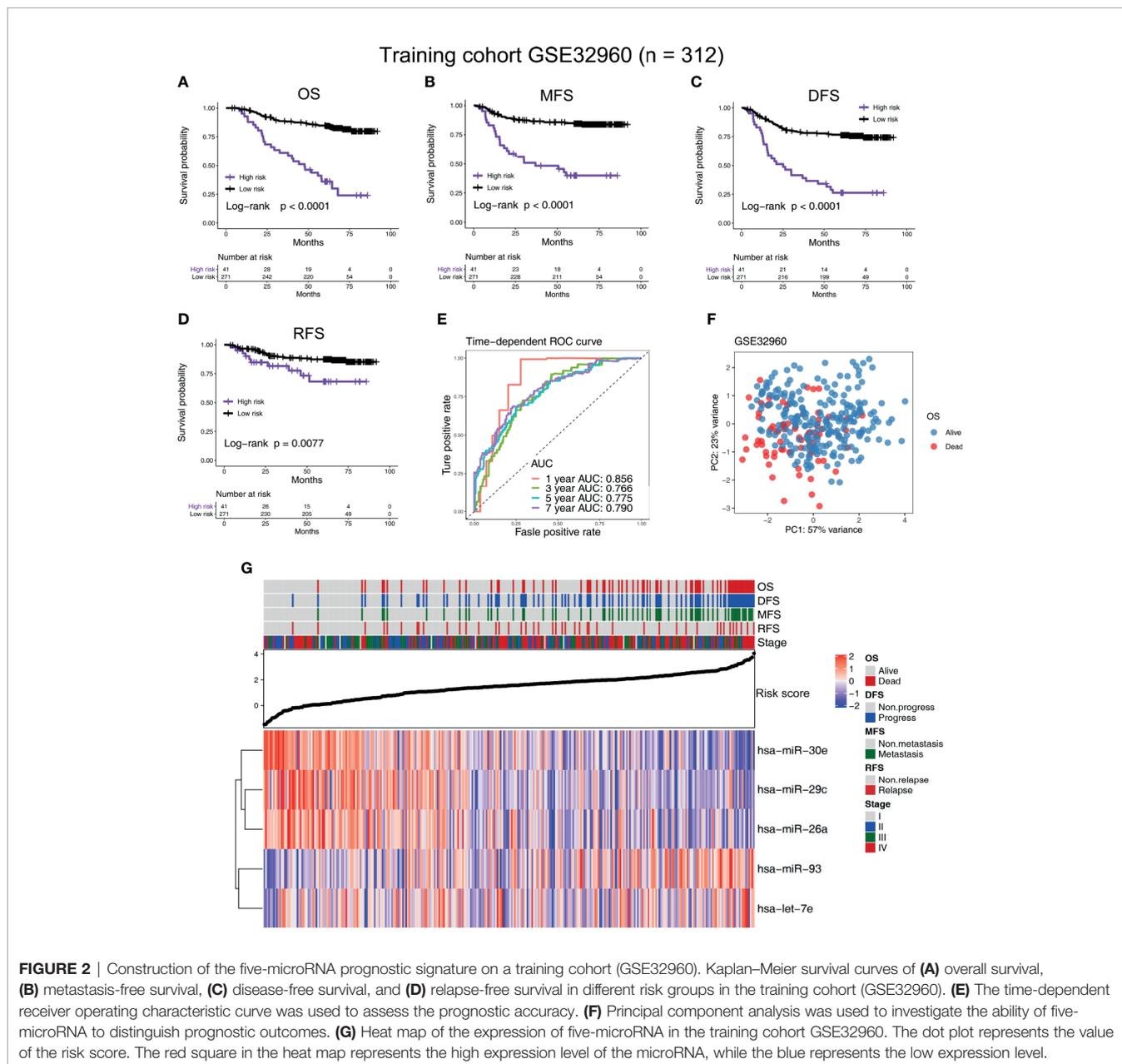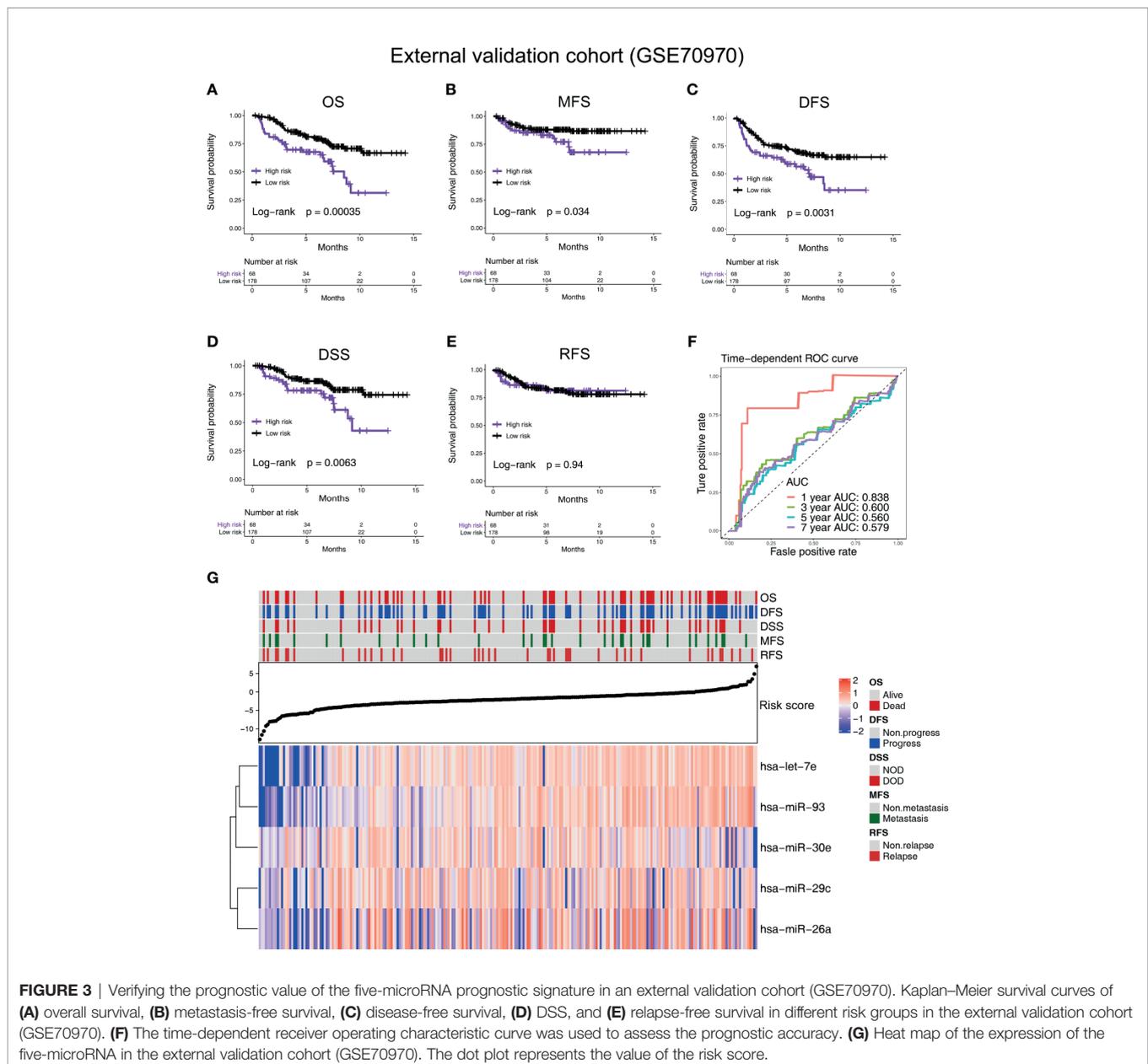

FIGURE 2 | Construction of the five-microRNA prognostic signature on a training cohort (GSE32960). Kaplan–Meier survival curves of **(A)** overall survival, **(B)** metastasis-free survival, **(C)** disease-free survival, and **(D)** relapse-free survival in different risk groups in the training cohort (GSE32960). **(E)** The time-dependent receiver operating characteristic curve was used to assess the prognostic accuracy. **(F)** Principal component analysis was used to investigate the ability of five-microRNA to distinguish prognostic outcomes. **(G)** Heat map of the expression of five-microRNA in the training cohort GSE32960. The dot plot represents the value of the risk score. The red square in the heat map represents the high expression level of the microRNA, while the blue represents the low expression level.

p = 0.0031), disease-specific survival (DSS, HR = 2.19, 95%CI: 1.12–4.28, p = 0.0063), and RFS (HR = 2.19, 95%CI: 1.10–4.36, p = 0.9449; **Figures 3A–E**). The TCGA-HNSC cohort showed that the prognostic signature could predict the OS (HR = 1.57, 95%CI: 1.15–2.14, p = 0.0013) and MFS (HR = 1.68, 95%CI: 1.15–2.43, p = 0.0024) of NPC (**Supplementary Figures S2A, B**). The time-dependent ROC curves in the GSE32960, GSE36682, GSE70970, and TCGA-HNSC cohorts showed that the five-microRNA signature had a high prediction accuracy. In particular, the area under the curve of prediction of 1-year OS was greater than 0.83 in three cohorts (GSE32960: 0.856, **Figure 2E**; GSE70970: 0.838, **Figure 3F**; and GSE36682: 0.898, **Figure S1B**). The PCA showed that the dominant features of the five microRNAs could significantly distinguish the prognostic

outcome of OS (**Figure 2F**). The heat map of the expression of the five microRNAs indicated that, with the increase of risk score, each microRNA has a significant trend of change, and the higher the risk score is, the more survival outcome events will occur (**Figures 2G**, **3G**, **Supplementary Figures S1C**, and **S2**). We calculated the C-index of the prognostic model in all the cohorts and found that the C-index in all of the three cohorts of GEO was greater than 0.77, indicating the high consistency of the prediction model (**Supplementary Table S3**).

The Nomogram Was Constructed by Cox Regression Analysis

We used univariate and multivariate Cox regression analysis to analyze some prognostic factors, including risk score, in the

FIGURE 3 | Verifying the prognostic value of the five-microRNA prognostic signature in an external validation cohort (GSE70970). Kaplan–Meier survival curves of **(A)** overall survival, **(B)** metastasis-free survival, **(C)** disease-free survival, **(D)** DSS, and **(E)** relapse-free survival in different risk groups in the external validation cohort (GSE70970). **(F)** The time-dependent receiver operating characteristic curve was used to assess the prognostic accuracy. **(G)** Heat map of the expression of the five-microRNA in the external validation cohort (GSE70970). The dot plot represents the value of the risk score.

training cohort GSE32960 and found that risk score (HR = 2.46, 95%CI: 1.90–3.19, $p < 0.001$) and N stage (HR = 1.54, 95%CI: 1.09–2.04, $p = 0.0144$) could predict OS in NPC patients independently of other factors (**Figures 4A, B**). Based on this result, we constructed a nomogram that conveniently predicted the OS of each NPC patient for 1, 3, and 5 years (**Figure 4C**). We found through DCA and clinical impact curve that the risk score combined with N stage could improve the clinical net benefits (**Figures 4D–F**).

Subgroup Survival Analysis and Functional Annotation of the Five-MicroRNA Signature

We found that, in NPC patients with different tumor stages (III/IV *vs.* I/II), T stages (III/IV *vs.* I/II), gender (male *vs*, female), age (>=65 *vs*. <65), whether or not to undergo radiotherapy and whether or not to undergo chemotherapy, the five-microRNA signature showed a remarkable ability to predict OS (**Figure 5**). In the TCGA head and neck squamous cell carcinoma cohort, a

subgroup analysis based on anatomical site subdivision showed a little difference in the expression of the risk score in each site (**Supplementary Figure S3**).

To understand the signaling pathways and biological functions associated with the five-microRNA signature, we conducted an enrichment analysis of the 500 genes most associated with the risk score, and the results indicated that the five-microRNA signature was associated with T cell activation, B cell activation, regulation of T cell activation, regulation of lymphocyte activation, immune response, and other biological functions (**Figure 6D** and **Supplementary Table S1**) as well as associated with MAPK, ether lipid metabolism, arachidonic acid metabolism, and ABC transporter, and other signaling pathways (**Figure 6D** and **Supplementary Table S2**). The GSEA analysis results indicated that patients with a high-risk score were usually activated by DNA repair, MYC targets, E2F targets, mTORC1, oxidative phosphorylation, and other signaling pathways (**Figure 6E**, **Table 2**). The NPC patients in the low-risk group were usually accompanied by a higher activation of INF-γ, INF-α,
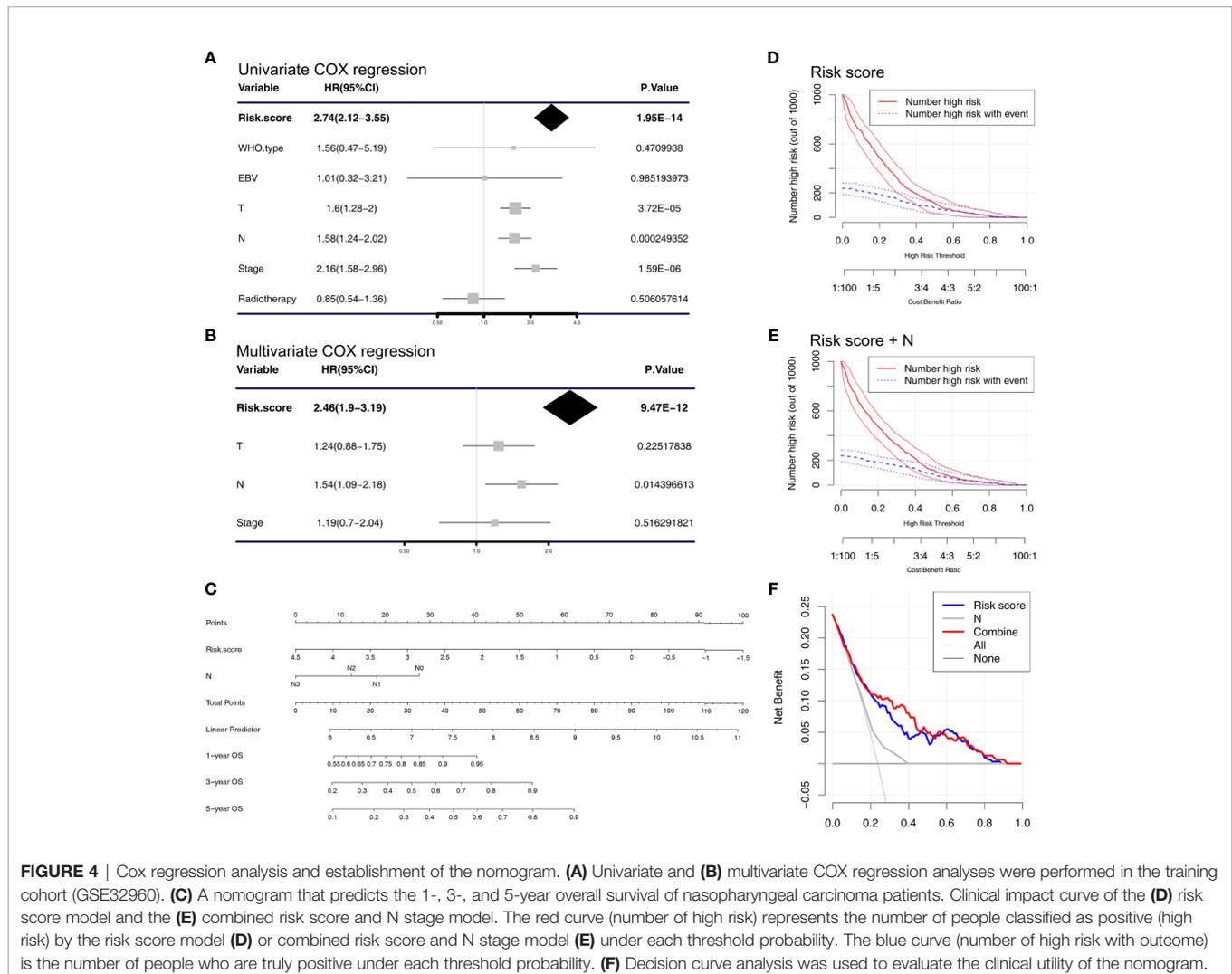

FIGURE 4 | Cox regression analysis and establishment of the nomogram. **(A)** Univariate and **(B)** multivariate COX regression analyses were performed in the training cohort (GSE32960). **(C)** A nomogram that predicts the 1-, 3-, and 5-year overall survival of nasopharyngeal carcinoma patients. Clinical impact curve of the **(D)** risk score model and the **(E)** combined risk score and N stage model. The red curve (number of high risk) represents the number of people classified as positive (high risk) by the risk score model **(D)** or combined risk score and N stage model **(E)** under each threshold probability. The blue curve (number of high risk with outcome) is the number of people who are truly positive under each threshold probability. **(F)** Decision curve analysis was used to evaluate the clinical utility of the nomogram.
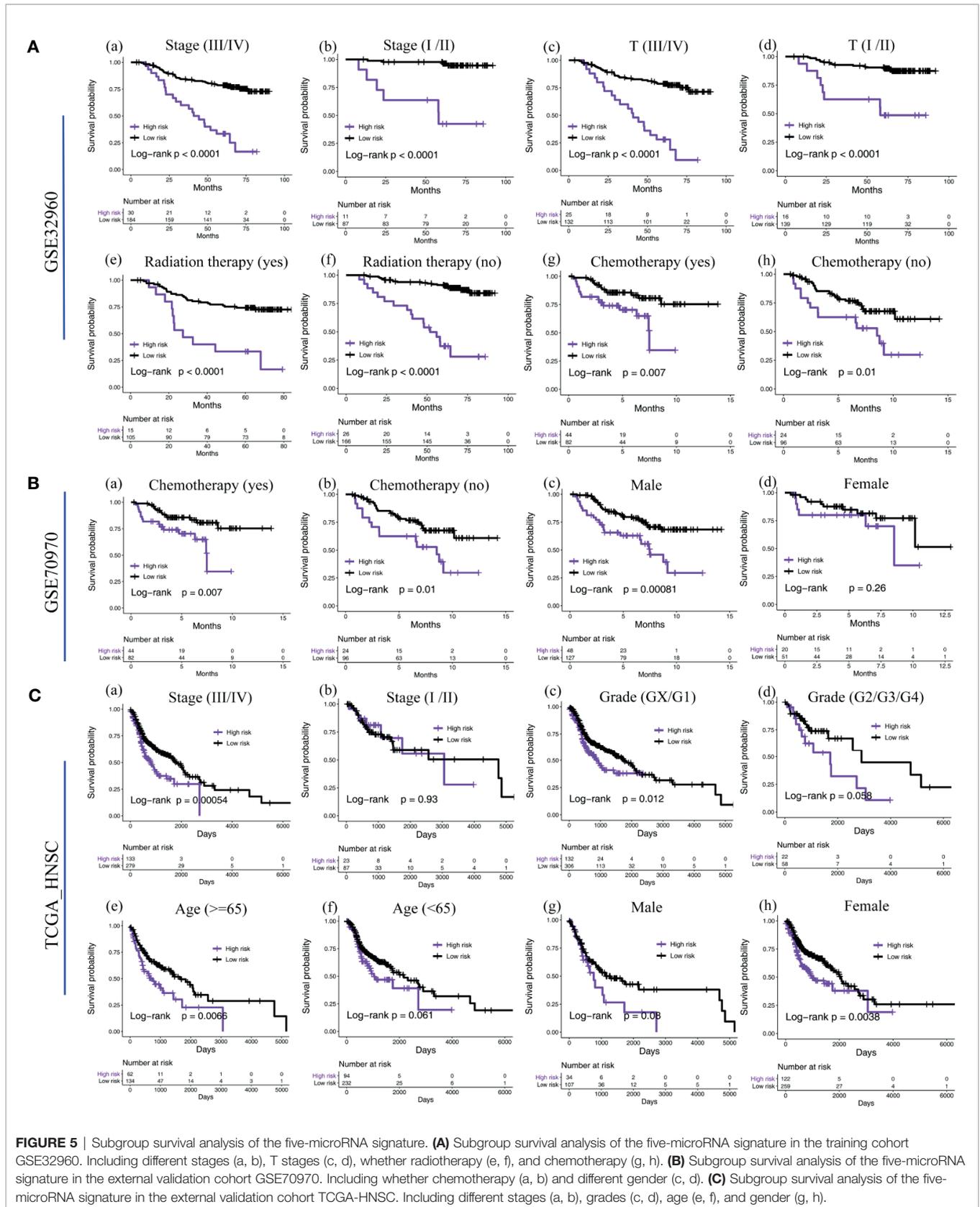
FIGURE 5 | Subgroup survival analysis of the five-microRNA signature. **(A)** Subgroup survival analysis of the five-microRNA signature in the training cohort GSE32960. Including different stages (a, b), T stages (c, d), whether radiotherapy (e, f), and chemotherapy (g, h). **(B)** Subgroup survival analysis of the five-microRNA signature in the external validation cohort GSE70970. Including whether chemotherapy (a, b) and different gender (c, d). **(C)** Subgroup survival analysis of the five-microRNA signature in the external validation cohort TCGA-HNSC. Including different stages (a, b), grades (c, d), age (e, f), and gender (g, h).
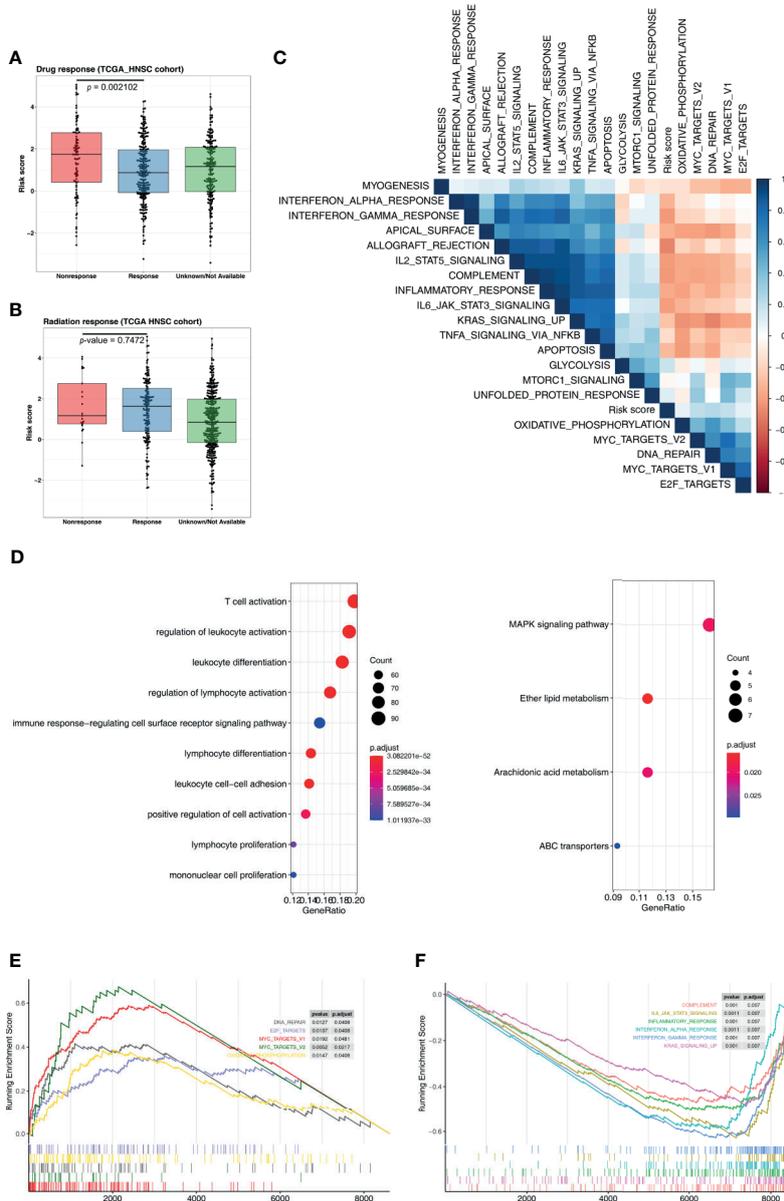
FIGURE 6 | Five-microRNA signature predicts the chemotherapy response and functional annotation. Distribution of risk scores in different **(A)** chemotherapy and **(B)** radiotherapy responses. **(C)** Heat map showing the correlation between the risk score and the signaling pathways. The blue square represents a positive correlation, and the red square represents a negative correlation. **(D)** Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analysis for the five-microRNA signature. Enrichment plots of gene set enrichment analysis based on the risk score, the pathway activated by the **(E)** high and **(F)** low risk score.

KRAS, inflammatory response, apoptosis, and other related signaling pathways (**Figure 6F** and **Table 2**). The risk score was significantly correlated with the signal values of these signaling pathways (**Figure 6C**).

Next, the R package "miRNAtap" was used for functional enrichment analysis of each microRNA separately (has-miR-93, has-let-7e, has-miR29c, has-miR-26a, and has-miR-30e). It was found that each microRNA participated in different major biological functions, such as hsa-let-7e—which was mainly

involved in cell cycle-related signaling pathways (cell cycle process, DNA replication, regulation of cell cycle process, cell cycle phase transition, *etc.*), hsa-miR-29c—which was mainly involved in the regulation of gene expression-related signaling pathways, and hsa-miR-30e—which was mainly involved in the regulation of metabolic-related biological events (carboxylic acid metabolic process, coenzyme metabolic process, ribonucleotide metabolic process, *etc.*; **Supplementary Table S4**).

TABLE 2 | Gene set enrichment analysis of the risk score.

Term	Set size	Enrichment score	NES	P-value	Rank
MYC_TARGETS_V2	28	0.67	3.95	0.00	2,139
DNA_REPAIR	76	0.42	3.40	0.01	1,096
OXIDATIVE_PHOSPHORYLATION	94	0.39	3.30	0.02	2,006
E2F_TARGETS	86	0.36	3.05	0.01	3,234
GLYCOLYSIS	85	0.23	1.99	0.01	3,083
MTORC1_SIGNALING	95	0.23	1.95	0.02	2,118
UNFOLDED_PROTEIN_RESPONSE	54	0.25	1.84	0.01	460
MYOGENESIS	125	-0.30	-1.73	0.00	4,863
TNFA_SIGNALING_VIA_NFKB	106	-0.33	-1.81	0.00	3,089
APICAL_SURFACE	21	-0.49	-1.89	0.00	2,934
APOPTOSIS	79	-0.36	-1.90	0.00	2,500
IL2_STAT5_SIGNALING	116	-0.46	-2.55	0.00	1,478
COMPLEMENT	122	-0.47	-2.66	0.00	2,436
KRAS_SIGNALING_UP	120	-0.48	-2.71	0.00	1,238
INFLAMMATORY_RESPONSE	134	-0.50	-2.89	0.00	2,565
IL6_JAK_STAT3_SIGNALING	60	-0.62	-3.13	0.00	1,459
INTERFERON_ALPHA_RESPONSE	85	-0.59	-3.14	0.00	1,682
INTERFERON_GAMMA_RESPONSE	163	-0.62	-3.68	0.00	1,964
ALLOGRAFT_REJECTION	147	-0.70	-4.08	0.00	987

NES, normalized enrichment score.

Five-MicroRNA Signature Is Associated With Chemotherapy Response and Immune Checkpoint

An analysis of the clinical efficacy of the TCGA-HNSC cohort exhibited that, among NPC patients undergoing chemotherapy, the risk score of nonresponse patients was significantly higher than that of patients with response ($p = 0.0021$), indicating that the five-microRNA signature is a potential prediction target of chemotherapy (**Figure 6A**). At the same time, we found that the risk score had no predictive value for the radiotherapy response ($p = 0.7472$, **Figure 6B**). By calculating correlations between risk scores and major immune checkpoints, we found that the risk scores were significantly negatively correlated with PD-L1 and PD-1, but not with TMB (**Supplementary Figure S4**). Moreover, the risk score was also correlated with tumor stage and T stage, while N stage was not (**Supplementary Figure S4**).

DISCUSSION

NPC is a kind of malignant head and neck cancer (20, 21). At present, the development of comprehensive treatment methods of radiotherapy and chemotherapy has significantly improved the prognosis and outcome of NPC patients. However, NPC is still considered a malignant and invasive tumor due to its high aggressiveness, delayed diagnosis, and relatively poor prognosis (22).

MicroRNAs are endogenous non-coding RNA families of approximately 22 nucleotides in length (23). miRNAs are believed to have important effects on the development, progression, and response to treatment of tumors (24). Currently, researchers have found that microRNA plays a crucial role in the development of NPC. In recent years, studies on the

pathogenesis and development of microRNA in NPC and its prognosis have been developing rapidly (9, 25–28). A meta-analysis identified 65 microRNAs as potential prognostic markers for NPC, providing new targets for future studies (29).

Currently, there are some prognostic signatures of microRNA in NPC (6, 7, 30, 31). Two miRNAs in the five-miRNA signature, has-miR29c and has-miR-30e, have been studied in nasopharyngeal carcinoma (32–34). It is possible that these two miRNAs have a particularly significant predictive role in the prognosis of nasopharyngeal carcinoma, so they are also used in other signatures. However, most of these studies ignore the analysis of subgroups and cannot prove that the prognostic signature still has a predictive value in certain subgroups of NPC patients. Our study conducted a subgroup analysis on the prognostic signature of five microRNAs and found that the five-microRNA signature showed a good prognostic value in each subgroup (different stages, grades, whether chemotherapy or radiotherapy, *etc.*), indicating that the five-microRNA signature could accurately predict the prognosis of various groups of NPC patients. At the same time, our study added a large number of cohorts to verify the results so as to avoid the overfitting effect and ensure that the model can exert its efficiency stably.

Resistance to chemotherapy is currently a major concern in the treatment of NPC. Recently, many studies have found that microRNAs are key molecules in the production of resistance to chemotherapy (35, 36). Our results indicated that the risk score based on the five-microRNA signature was negatively correlated with the efficacy of chemotherapy. We could predict the chemotherapy sensitivity of NPC patients according to the risk score so as to propose specific treatment plans for each NPC patient. However, due to the small number of cohorts with documented chemotherapy responses, our results were not validated in more cohorts. We will collect our chemotherapy response data in the future to verify our conclusions.

Through the analysis of the function annotation, we found that the higher risk score tends to the activation of cell cycle-related signal pathways, and the low-risk group of patients with significant immune response related to activating signaling pathways; it also provides a new way of thinking for the treatment of NPC patients, namely, the high-risk group of patients can try drugs with a targeted cell cycle, and the low patriarch of patients can try immune checkpoint inhibitors. This part of the conjecture needs further verification, which is our future research direction.

In addition, according to the univariate and multivariate Cox regression analyses, we found that the risk score and N stage could be independent of other prognostic factors to predict the prognosis of patients with NPC. On this basis, we constructed a nomogram, which could be convenient to predict the 1, 3, and 5-year overall survival in patients with NPC. The DCA analysis found that the joint risk score and N stage could have a higher clinical net benefit.

In summary, through four big miRNA data cohorts, with a total of more than 1,143 NPC patients, we develop and validate a five-microRNA prognostic signature. It could accurately predict the prognosis of patients with NPC and progress by constructing

a nomogram, combined with risk score and N stage, to accurately predict the prognosis of patients with NPC. Besides this, based on the risk score, a potential layered treatment scheme for NPC was proposed.

DATA AVAILABILITY STATEMENT

The data of the TCGA database can be downloaded from UCSC Xena website (https://gdc.xenahubs.net). All the microarrays data (GSE32960, GSE36682, GSE70970) can be downloaded from the Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/).

AUTHOR CONTRIBUTIONS

SW and SL conceived and designed the study. SH involved in most of the revision of the manuscript. CZ and SW performed the analysis procedures. SW, SH, CZ, JX, and SL analyzed the results. SW, SH, JX, and SL contributed the analysis tools. SW, SH, CZ, JX, and SL contributed to the writing of the manuscript.

All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.723362/full#supplementary-material

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A, et al. Global Cancer Statistics, 2012. *CA Cancer J Clin* (2015) 65(2):87–108. doi: 10.3322/caac.21262

2. Pan JJ, Ng WT, Zong JF, Lee SW, Choi HC, Chan LL, et al. Prognostic Nomogram for Refining the Prognostication of the Proposed 8th Edition of the AJCC/UICC Staging System for Nasopharyngeal Cancer in the Era of Intensity-Modulated Radiotherapy. *Cancer* (2016) 122(21):3307–15. doi: 10.1002/cncr.30198

3. Moertel CG, Fleming TR, Macdonald JS, Haller DG, Laurie JA, Tangen CM, et al. Fluorouracil Plus Levamisole as Effective Adjuvant Therapy After Resection of Stage III Colon Carcinoma: A Final Report. *Ann Intern Med* (1995) 122(5):321–6. doi: 10.7326/0003-4819-122-5-199503010-00001

4. Mao YP, Xie FY, Liu LZ, Sun Y, Li L, Tang LL, et al. Re-Evaluation of 6th Edition of AJCC Staging System for Nasopharyngeal Carcinoma and Proposed Improvement Based on Magnetic Resonance Imaging. *Int J Radiat Oncol Biol Phys* (2009) 73(5):1326–34. doi: 10.1016/j.ijrobp.2008.07.062

5. Edge SB, Compton CC. The American Joint Committee on Cancer: The 7th Edition of the AJCC Cancer Staging Manual and the Future of TNM. *Ann Surg Oncol* (2010) 17(6):1471–4. doi: 10.1245/s10434-010-0985-4

6. Liu N, Cui RX, Sun Y, Guo R, Mao YP, Tang LL, et al. A four-miRNA Signature Identified From Genome-Wide Serum miRNA Profiling Predicts Survival in Patients With Nasopharyngeal Carcinoma. *Int J Cancer* (2014) 134(6):1359–68. doi: 10.1002/ijc.28468

7. Zhang S, Yue W, Xie Y, Liu L, Li S, Dang W, et al. The fourmicroRNA Signature Identified by Bioinformatics Analysis Predicts the Prognosis of Nasopharyngeal Carcinoma Patients. *Oncol Rep* (2019) 42(5):1767–80. doi: 10.3892/or.2019.7316

8. Li G, Qiu Y, Su Z, Ren S, Liu C, Tian Y, et al. Genome-Wide Analyses of Radioresistance-Associated miRNA Expression Profile in Nasopharyngeal Carcinoma Using Next Generation Deep Sequencing. *PloS One* (2013) 8(12):e84486. doi: 10.1371/journal.pone.0084486

9. Spence T, Bruce J, Yip KW, Liu FF. MicroRNAs in Nasopharyngeal Carcinoma. *Chin Clin Oncol* (2016) 5(2):17. doi: 10.21037/cco.2016.03.09

10. Jiang Y, Chen C, Xie J, Wang W, Zha X, Lv W, et al. Radiomics Signature of Computed Tomography Imaging for Prediction of Survival and

Chemotherapeutic Benefits in Gastric Cancer. *EBioMedicine* (2018) 36:171–82. doi: 10.1016/j.ebiom.2018.09.007

11. Li H, Liu J, Chen J, Wang H, Yang L, Chen F, et al. A Serum microRNA Signature Predicts Trastuzumab Benefit in HER2-Positive Metastatic Breast Cancer Patients. *Nat Commun* (2018) 9(1):1614. doi: 10.1038/s41467-018-03537-w

12. Zhu GQ, Zhou YJ, Qiu LX, Wang B, Yang Y, Liao WT, et al. Prognostic Alternative mRNA Splicing Signature in Hepatocellular Carcinoma: A Study Based on Large-Scale Sequencing Data. *Carcinogenesis* (2019) 40(9):1077–85. doi: 10.1093/carcin/bgz073

13. Zheng N, Wang K, He J, Qiu Y, Xie G, Su M, et al. Effects of ADMA on Gene Expression and Metabolism in Serum-Starved LoVo Cells. *Sci Rep* (2016) 6:25892. doi: 10.1038/srep25892

14. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw* (2010) 33(1):1–22. doi: 10.18637/jss.v033.i01

15. 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(13):3001–8. doi: 10.1093/bioinformatics/bti422

16. Yu G, Wang LG, Han Y, He QY. Clusterprofiler: An R Package for Comparing Biological Themes Among Gene Clusters. *OMICS* (2012) 16(5):284–7. doi: 10.1089/omi.2011.0118

17. Vickers AJ, Elkin EB. Decision Curve Analysis: A Novel Method for Evaluating Prediction Models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361

18. Tin Kam H. The Random Subspace Method for Constructing Decision Forests. *IEEE Trans Pattern Anal Mach Intell* (1998) 20(8):832–44. doi: 10.1109/34.709601

19. Zeng D, Li M, Zhou R, Zhang J, Sun H, Shi M, et al. Tumor Microenvironment Characterization in Gastric Cancer Identifies Prognostic and Immunotherapeutically Relevant Gene Signatures. *Cancer Immunol Res* (2019) 7(5):737–50. doi: 10.1158/2326-6066.CIR-18-0436

20. Bruce JP, Yip K, Bratman SV, Ito E, Liu FF. Nasopharyngeal Cancer: Molecular Landscape. *J Clin Oncol* (2015) 33(29):3346–55. doi: 10.1200/JCO.2015.60.7846

21. Farias TP, Dias FL, Lima RA, Kligerman J, de Sa GM, Barbosa MM, et al. Prognostic Factors and Outcome for Nasopharyngeal Carcinoma. *Arch Otolaryngol Head Neck Surg* (2003) 129(7):794–9. doi: 10.1001/archotol.129.7.794

22. Luan J, Wang J, Su Q, Chen X, Jiang G, Xu X, et al. Meta-Analysis of the Differentially Expressed microRNA Profiles in Nasopharyngeal Carcinoma. *Oncotarget* (2016) 7(9):10513–21. doi: 10.18632/oncotarget.7013

23. Bartel DP. MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell* (2004) 116(2):281–97. doi: 10.1016/S0092-8674(04)00045-5

24. Xue WQ, Qin HD, Ruan HL, Shugart YY, Jia WH. Quantitative Association of Tobacco Smoking With the Risk of Nasopharyngeal Carcinoma: A Comprehensive Meta-Analysis of Studies Conducted Between 1979 and 2011. *Am J Epidemiol* (2013) 178(3):325–38. doi: 10.1093/aje/kws479

25. Wu L, Wang J, Zhu D, Zhang S, Zhou X, Zhu W, et al. Circulating Epstein-Barr Virus microRNA Profile Reveals Novel Biomarker for Nasopharyngeal Carcinoma Diagnosis. *Cancer Biomark* (2020) 27(3):365–75. doi: 10.3233/CBM-190160

26. Bissey PA, Law JH, Bruce JP, Shi W, Renoult A, Chua MLK, et al. Dysregulation of the MiR-449b Target TGFBI Alters the TGFbeta Pathway to Induce Cisplatin Resistance in Nasopharyngeal Carcinoma. *Oncogenesis* (2018) 7(5):40. doi: 10.1038/s41389-018-0050-x

27. Wang S, Zhang R, Claret FX, Yang H. Involvement of microRNA-24 and DNA Methylation in Resistance of Nasopharyngeal Carcinoma to Ionizing Radiation. *Mol Cancer Ther* (2014) 13(12):3163–74. doi: 10.1158/1535-7163.MCT-14-0317

28. Bian LH, Duan JL, Zhou C, Shen GW, Wang XY, Yang Y, et al. MicroRNA19b Inhibitors can Attenuate the STAT3 Signaling Pathway in NPC C6661 Cells. *Mol Med Rep* (2020) 22(1):51–6. doi: 10.3892/mmr.2020.11112

29. Sabarimurugan S, Kumarasamy C, Baxi S, Devi A, Jayaraj R. Systematic Review and Meta-Analysis of Prognostic microRNA Biomarkers for Survival Outcome in Nasopharyngeal Carcinoma. *PloS One* (2019) 14(2):e0209760. doi: 10.1371/journal.pone.0209760

30. Zhang H, Zou X, Wu L, Zhang S, Wang T, Liu P, et al. Identification of a 7-microRNA Signature in Plasma as Promising Biomarker for Nasopharyngeal Carcinoma Detection. *Cancer Med* (2020) 9(3):1230–41. doi: 10.1002/cam4.2676

31. Liu N, Chen NY, Cui RX, Li WF, Li Y, Wei RR, et al. Prognostic Value of a microRNA Signature in Nasopharyngeal Carcinoma: A microRNA Expression Analysis. *Lancet Oncol* (2012) 13(6):633–41. doi: 10.1016/S1470-2045(12)70102-X

32. Wang T, Wu J, Wu Y, Cheng Y, Deng Y, Liao J, et al. A Novel microRNA-Based Signature Predicts Prognosis Among Nasopharyngeal Cancer Patients. *Exp Biol Med (Maywood)* (2021) 246(1):72–83. doi: 10.1177/1535370220958680

33. Liu N, Tang LL, Sun Y, Cui RX, Wang HY, Huang BJ, et al. MiR-29c Suppresses Invasion and Metastasis by Targeting TIAM1 in Nasopharyngeal Carcinoma. *Cancer Lett* (2013) 329(2):181–8. doi: 10.1016/j.canlet.2012.10.032

34. Zhang JX, Qian D, Wang FW, Liao DZ, Wei JH, Tong ZT, et al. MicroRNA-29c Enhances the Sensitivities of Human Nasopharyngeal Carcinoma to Cisplatin-Based Chemotherapy and Radiotherapy. *Cancer Lett* (2013) 329 (1):91–8. doi: 10.1016/j.canlet.2012.10.033

35. Zhao Y, Wang P, Wu Q. miR-1278 Sensitizes Nasopharyngeal Carcinoma Cells to Cisplatin and Suppresses Autophagy via Targeting ATG2B. *Mol Cell Probes* (2020) 53:101597. doi: 10.1016/j.mcp.2020.101597

36. Cui Z, Pu T, Zhang Y, Wang J, Zhao Y. Long non-Coding RNA LINC00346 Contributes to Cisplatin Resistance in Nasopharyngeal Carcinoma by Repressing miR-342-5p. *Open Biol* (2020) 10(5):190286. doi: 10.1098/rsob.190286