Original Research

A novel microRNA-based signature predicts prognosis among nasopharyngeal cancer patients

Tianyu Wang[1,*], Jian Wu[1,*], Yun Wu[2,*], Yin Cheng[1], Yue Deng[1], Jianchun Liao[1], Huanhai Liu[1] and Hu Peng[1]

[1]Department of Otolaryngology-Head and Neck Surgery, Changzheng Hospital, Second Military Medical University, Shanghai 200003, China; [2]Department of Otorhinolaryngology-Head and Neck Surgery, Jiangsu Taizhou People's Hospital, Taizhou 225300, China; *Tianyu Wang, Jian Wu and Yun Wu are contributed equally to this paper.

Corresponding authors: Huanhai Liu. Email: liuhuanhaiok@smmu.edu.cn; Hu Peng. Email: penghudy@smmu.edu.cn

*These authors contributed equally to this paper

Impact statement

Nasopharyngeal cancer is one of the most common malignant tumors in the head and neck. Identification of promising miRNA biomarkers might benefit a lot to the detection of nasopharyngeal carcinoma. A three-miRNA signature (has-miR-142-3p, has-miR-29c, and has-miR-30e) was obviously associated with the overall survival of nasopharyngeal carcinoma patients. The model has better clinical independence and has better clinical prediction effect when combined with clinical characteristics. Our results revealed that a three-miRNA signature was a potential novel prognostic biomarker for nasopharyngeal carcinoma.

Abstract

Nasopharyngeal cancer is one of the most common malignant tumors in the head and neck. Identification of promising miRNA biomarkers might benefit a lot to the detection of nasopharyngeal carcinoma. miRNA expression profile and clinical information were obtained from two microarray profiling data sets from the Gene Expression Omnibus (GEO) database. miRNA signature model was constructed via univariate Cox survival analysis, multivariate Cox survival analysis, and least absolute shrinkage and selection operator Cox regression analysis. Kaplan–Meier curve, area under the curve (AUC), decision curve analysis, Box plot, and nomogram were used to evaluate the prognosis of the model to patients. 67 up-regulated and 93 down-regulated miRNAs were identified from GEO microarray data sets ($P < 0.05$). A three-miRNA signature (has-miR-142-3p, has-miR-29c, and has-miR-30e) was obviously associated with the overall survival of nasopharyngeal carcinoma patients ($P < 0.001$). The AUCs for the signature were 0.74, 0.7 for the training set and external validation set. The AUC of disease free survival and distant metastasis-free survival were also high. The model has better clinical independence and has better clinical prediction effect when combined with clinical characteristics ($P < 0.0001$). Compared with the published models, our model had a higher AUC. Our results revealed that a three-miRNA signature was a potential novel prognostic biomarker for nasopharyngeal carcinoma.

Keywords: Nasopharyngeal cancer, miRNA signature, Gene Expression Omnibus, area under the curve

Experimental Biology and Medicine **2021; 246: 72–83. DOI: 10.1177/1535370220958680**

Introduction

The incidence of nasopharyngeal cancer is closely related to genetic factors, Epstein–Barr virus (EBV) infection and environmental factors. Early diagnosis is the most effective means to save patients' lives and improve their quality of life.[1] Unfortunately, the onset of nasopharyngeal carcinoma is hidden and has a strong tendency to metastasize. According to statistics, about 75% of patients have reached the advanced stage at the time of treatment, becoming the main reason for the failure of nasopharyngeal cancer treatment.[2] Therefore, screening the tumor markers of nasopharyngeal carcinoma, striving for early detection, and selecting the best treatment scheme have important clinical significance for the diagnosis and treatment of nasopharyngeal carcinoma.

The study of tumor markers is a hot topic along with the whole course of tumor research. At present, various tumor markers have been found and have been widely used in clinical practice. Alpha-fetoprotein is used to diagnose and monitor liver cancer, prostatespecific antigen is used for the diagnosis of prostate cancer, and humanpapillomavirus is used to screen for cervical cancer. EBV antibody and EBV-

DNA are commonly used in clinical screening of nasopharyngeal carcinoma, and the positive rate is correlated with the clinical stage, TNM Classification of Malignant Tumors (TNM) stage, and survival rate of nasopharyngeal carcinoma,[3] but the sensitivity and specificity are far from the clinical requirements. The ideal tumor marker should have the characteristics of easy detection, stable expression, early change, high sensitivity and specificity, and be able to reflect the progress and progression of the tumor.[4] At present, tumor markers that fully meet the standard have not been found.

Studies have found that more than 50% of miRNAs are located in tumor-related genomic regions, and chromosomal abnormalities directly lead to changes in the copy number of miRNA genes, resulting in misregulation of miRNA expression in a variety of tumors and playing the role of oncogenes or tumor suppressor genes.[5]miR-205 can be easily differentiated between squamous cell lung cancer and non-small cell lung cancer, with a sensitivity of 96% and specificity of 90%.[6] High expression of miR-205 may increase the sensitivity of breast cancer cells to the molecular boot-to-drug tyrosine kinase inhibitors gefitinib and lapatinib.[7] It was found that mir-216b was downregulated in nasopharyngeal carcinoma tissues and cells, and inhibited the proliferation of nasopharyngeal carcinoma cells *in vitro* by inhibiting AKT and ERK pathways.[8]miR-200a promotes epithelium-mesenchymal transition of nasopharyngeal carcinoma cells by regulating boot genes ZEB2 and p-catenin, and differentiate into stem-like cells, so as to obtain the ability of infiltration and metastasis.[9,10]EBV-miR-BART8-3p expression was significantly higher in human nasopharyngeal carcinoma and is a potential therapeutic target for nasopharyngeal carcinoma.[11] In summary, miRNA has a promising application prospect in the study of tumor molecular markers. However, integrating multiple genes may be a more reliable predictor of tumor prognosis.

In this study, bioinformatics technology was used to analyze nasopharyngeal carcinoma miRNA expression profiles in Gene Expression Omnibus (GEO) database, and univariate and multivariate COX survival analysis was used to identify miRNA signature. We hope that this miRNA prognosis model can be used as a potential diagnostic indicator for patients with nasopharyngeal carcinoma.

Materials and methods

Data collection

The miRNA expression profiles and clinical data from GSE32960 and GSE70970 were obtained from the GEO (www.ncbi.nlm.nih.gov/geo) database.[12] The GSE32960 data set contains 312 non-distant metastasis nasopharyngeal cancer and 18 non-cancerous nasopharyngitis biopsy samples. The median follow-up time was 62.1 months (IQR 47.7–71.5). All these samples were collected from the Sun Yat-sen University Cancer Center (Guangzhou, China) between 16 January 2003 and 25 February 2006. The clinical staging was classified according to the criteria of the United States Joint Committee on Cancer Staging (Seventh Version). The GSE70970 data set contains a total of 246 nasopharyngeal cancer patients from Princess Margaret Cancer Center (Toronto, Canada). Samples with a survival time of less than one month were removed and ComBat was used to delete the batch effect. The data set information is shown in Table 1. The work flow chart is shown in Figure 1.

Table 1. Clinical information of the two data sets.

Characteristic	GSE32960 ($n = 312$)	GSE70970 ($n = 246$)
Age (years)		
≤50	201	116
>50	111	130
Survival status		
No	238	176
Yes	74	70
Disease free status		
No	217	159
Yes	95	87
Metastasis status		
No	246	211
Yes	66	35
Gender		
Female	79	71
Male	233	175
pathologic_T		
T1	66	74
T2	89	50
T3	71	52
T4	86	67
pathologic_N		
N0	44	49
N1	148	83
N2	72	90
N3	48	23
Tumor Stage		
Stage I	12	–
Stage II	86	–
Stage III	91	–
Stage IV	123	–
Cocurrent chemotherapy		
No	44	120
Yes	268	126
Radiotherapy boosting		
No	163	–
Yes	149	–

miRNA differential expression analysis

The tumor samples and non-tumor samples from the GSE32960 data set were subjected to differential analysis through the R software limma package, and miRNAs with $|\log 1.5| > 1$ and false discovery rate < 0.05 were defined as differentially expressed miRNAs.

miRNA prognosis model

miRNAs that were significant in the univariate Cox survival analysis and remained significant after inclusion of clinical covariates were used as candidate gene sets. Eighty percent of the GSE32960 data sets were randomly selected as the training set. Based on the "glmnet"[13] software package in R, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed for 200 times, and miRNA with the highest occurrence frequency

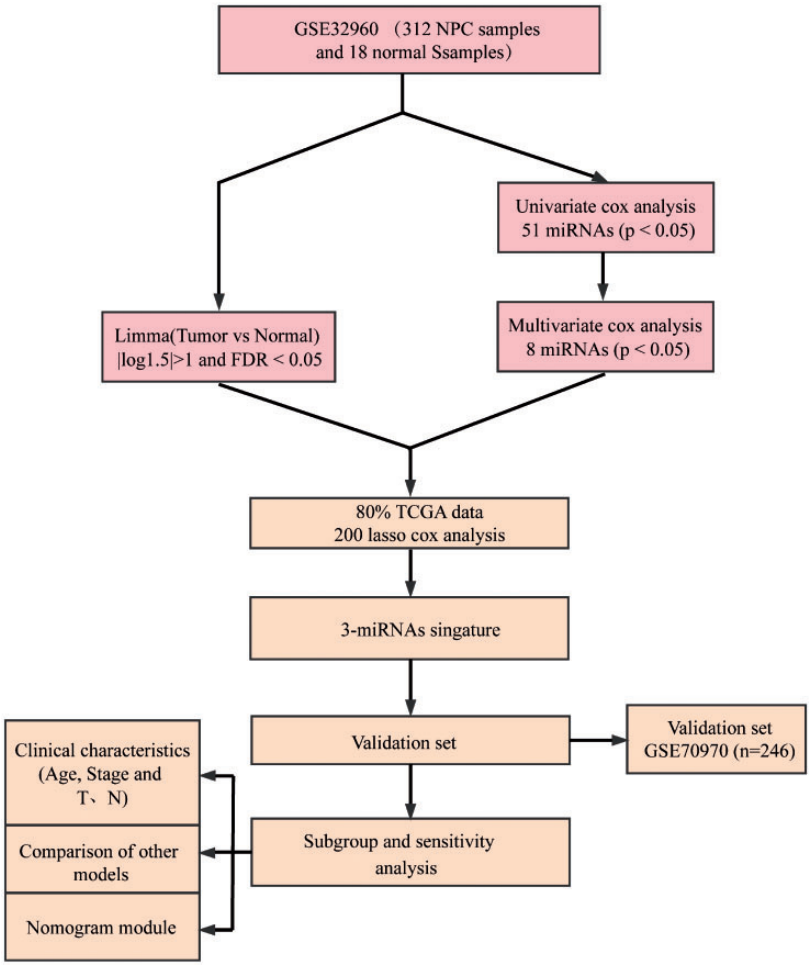

Figure 1. Work flow chart. (A color version of this figure is available in the online journal.)

(higher than 150) was selected as nasopharyngeal cancer samples to establish the best prognosis model. The optimal value of the parameter was determined through 10 cross-validation, and finally the relative expression of each prognostic miRNA and its correlation coefficients were determined by multivariate cox analysis, and the prognostic miRNA risk score for each sample was calculated. The risk score for miRNA signatures was calculated as follows:

$$\sum_{i=1}^{n} (\mathrm{coef}_i \times \mathrm{Expr}_i)$$

where Expr_i is the expression value of the corresponding gene sample, coef_i is the cox regression coefficient of multivariate factors.

Evaluate the prognosis of patients with miRNA model

According to the risk score associated with the miRNA model, patients from different data sets were divided into low-risk and high-risk groups. The R package "timeROC"[14] analysis was then used to calculate one-, three- and five-year area under the curve (AUC) of disease-free survival (DFS), overall survival (OS), and distant metastasis-free survival (DMFS). Kaplan–Meier (KM) survival analysis was used to assess the survival differences between different clinicopathological characteristics, between the high/low-risk groups and between the Stage I/II *vs.* Stage III/IV groups in the data sets mentioned above. The "survival" package in R was used to perform a two-sided log-rank test and univariate and multivariate Cox regression analyses.

Correlation between prognostic model and clinicopathological features

Boxplot was used to show the relationship between risk score and the corresponding clinicopathological characteristics, including age, Stage stages, T, N, etc., and the statistical significance was analyzed by chi-square test.

Nomogram

First, univariate and multivariate Cox regression analyses were performed to identify the appropriate items to construct the nomogram. Value, hazard ratio (HR), and 95% confidence interval (CI) of each variable, was determined

using the "forestplot" software package in R.[15] We found that the miRNA signature model, N Stage, T Stage, and Gender were the only four independent prognostic factors that could be used to predict survival. Thus, by using the "RMS" package in R,[16] four independent prognostic factors were used to construct the nomogram. The AUC was used to analyze and compare the predictive models of clinical outcomes. In addition, decision curve analysis (DCA) was used to measure the suitability of the nomograms we established for clinical application.

Comparison with published models

By referring to the literature, we finally selected three prognostic risk models: 4-miRNA signature,[17] 4-miRNA signature,[18] and 4-miRNA signature[19] for comparison with our 3-miRNA model, and evaluated them by KM curve, receiver operating characteristic (ROC) curve, restricted mean survival (RMS), and decision curve analysis (DCA) curve.

Results

Correlation between nasopharyngeal carcinoma samples and clinical features

First, according to the OS time and clinical characteristics of the sample, lymph node metastasis N, invasion degree T, age, TMN Stage, radiotherapy, and chemotherapy were, respectively, analyzed for KM prognosis. The results showed that lymph node metastasis N, invasion degree T, TNM Stage, age, and Gender have a significant impact on prognosis (Figure S1).

Identification of miRNAs significantly correlated with prognosis

Univariate Cox proportional risk regression model was used for the expression and survival data of 872 miRNAs in GSE32960 samples. The R package was used to survival coxph function, and $P < 0.05$ was selected as the threshold value. Finally, 51 miRNAs with significant prognosis were obtained (Figure 2(a)). Furthermore, the significant clinical features of lymph node metastasis (N Stage), invasion level (T Stage), TNM Stage, Age, and Gender were used as covariables for multivariate survival analysis of 51 miRNA. The significance threshold was selected as 0.05, and eight miRNAs were finally obtained (Figure 2(b)). Then, based on the expression levels of these 872 miRNAs, the differences between tumor samples and non-tumor samples were calculated by limma package. Finally, significant differences were found in 150 miRNAs, among which 67 were up-regulated and 93 were down-regulated (Figure 2(c)), among the 8 miRNAs obtained by multivariate survival analysis, 6 miRNAs showed significant differences in down-regulation, and 1 miRNA showed significant differences in up-regulation. The heat map also showed differential expression of miRNA (Figure 2(d)). Finally, the seven differentially expressed miRNAs associated with prognosis were used for subsequent studies.

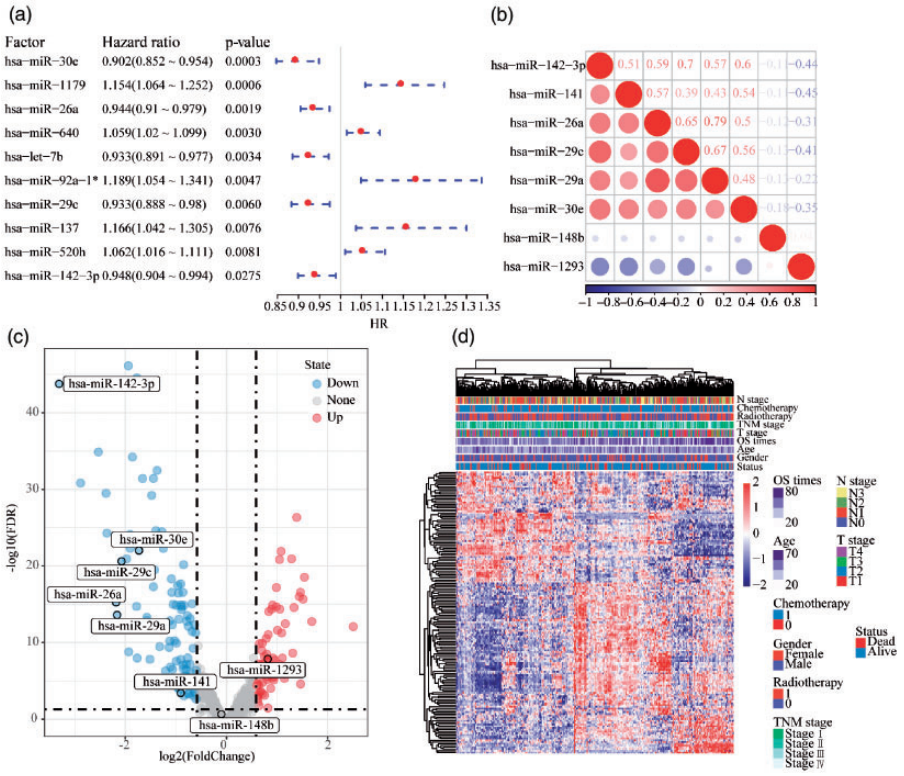

Figure 2. miRNAs with significantly different prognosis. (a) Univariate analysis of significant miRNA forest map. (b) Multivariate analysis of significant miRNA-related heatmaps. Red represents positive correlation and blue represents negative correlation. (c) Volcano map of differential expression miRNA. Red is up-regulated miRNAs, blue is down-regulated miRNAs. (d) Heatmap of differential expression miRNA. (A color version of this figure is available in the online journal.)

3-miRNA signature analysis

In order to reduce the number of miRNAs, LASSO regression analysis was used for dimensionality reduction. Eighty percent of the samples were randomly selected from GSE32960 for LASSO analysis, and 10× cross-validation was used to conduct LASSO analysis for 200 times, and the frequency of each probe in the 200 times was counted (Figure 3(a)). Finally, three miRNAs with frequency greater than 150 were selected (Table 2), and KM curve analysis of the three miRNAs showed that these three genes could significantly distinguish the GSE32960 samples from the two groups of high and low risk (Figure 3(b) to (d)). A formula was obtained to calculate the risk score for every patient from the expression values of the three hub miRNAs, weighted by the regression coefficient. $\text{RiskScore}_3 = -0.124 \times \exp^{\text{hsa-miR-142-3p}} - 0.219 \times \exp^{\text{hsa-miR-29c}} - 0.336 \times \exp^{\text{hsa-miR-30e}}$.

With this risk score formula, nasopharyngeal cancer patients were divided into low-risk or high-risk groups. Furthermore, the survival difference between the two groups of patients was plotted. Notably, compared with patients with low-risk scores of the three miRNAs, patients with high-risk scores had a shorter OS. Also, high expression of hsa-mir-142-3p, hsa-mir-29c, and hsa-mir-30e were associated with low risk, which were protective factors (Figure 4(a)). Furthermore, to compare the sensitivity and specificity of prediction, ROC analysis was performed and the AUC of one year is 0.74 (Figure 4(b)). Finally, we conducted zscore for RiskScore, dividing the samples greater than zero into the high-risk group and the samples less than zero into the low-risk group. KM prognosis curve showed significant difference in prognosis (Figure 4(c)). Similarly, the five years AUCs of the risk score in RFS and DMFS were 0.68 and 0.7, respectively (Figure 5 (a) and (b)). KM curve of RFS and DMFS also had significant difference between high-risk group and low-risk group (Figure 5(c) and (d)).

Robustness of 3-miRNA signature

To evaluate the predictive value of miRNA signature, the same model and the same coefficient as the training set were used in an external validation set (GSE70970). With this risk score formula, nasopharyngeal cancer patients were divided into low-risk or high-risk groups. Furthermore, the survival difference between the two groups of patients was plotted. Notably, compared with

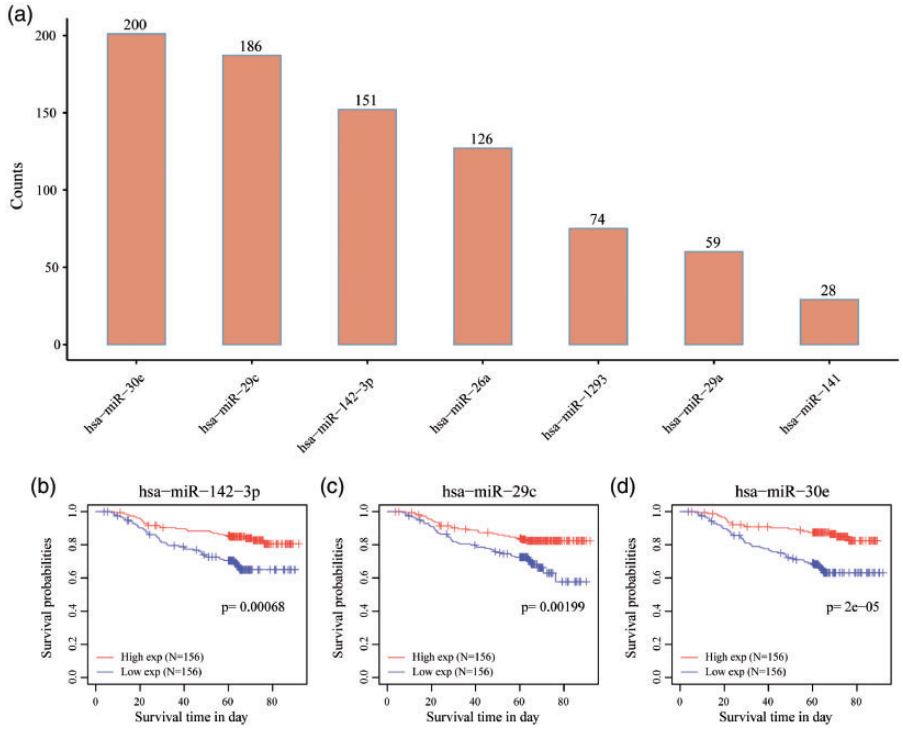

Figure 3. Screening target miRNAs. (a) The frequency distribution of miRNA selected by 200 LASSO features, the horizontal axis represents genes, and the vertical axis represents the frequency of occurrence. (b) to (d) KM prognosis curve of has-miR-142-3p, has-miR-29c, and has-miR-30e. (A color version of this figure is available in the online journal.)

Table 2. 3-miRNA signature.

Symbol	coef	HR	*P* value	Low 95%CI	High 95%CI
hsa-miR-142-3p	−0.124	0.883	0.482	0.624	1.249
hsa-miR-29c	−0.219	0.803	0.225	0.564	1.144
hsa-miR-30e	−0.336	0.715	0.022	0.536	0.953

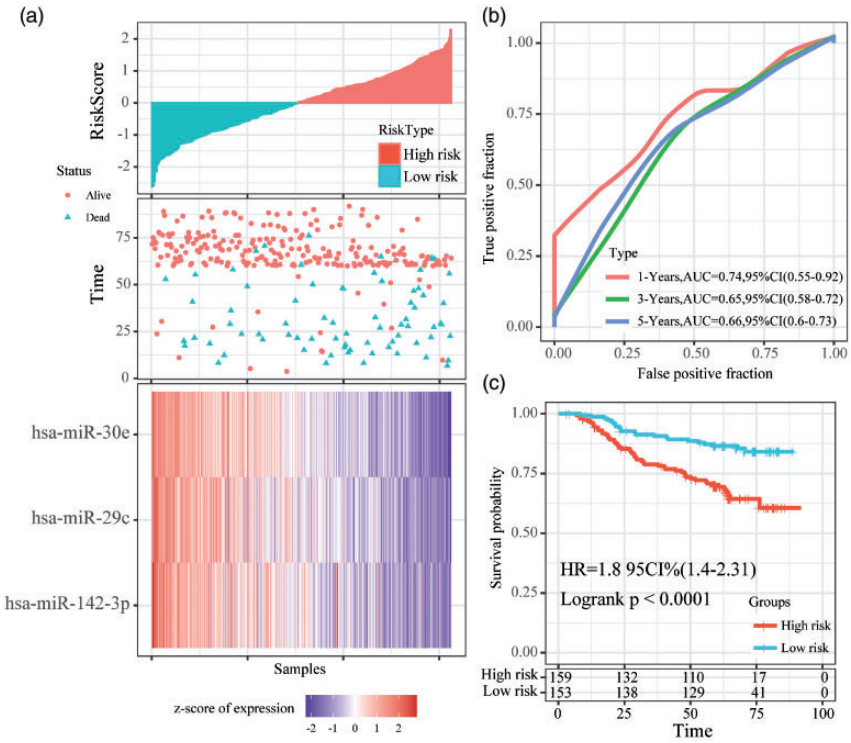

Figure 4. Prognosis curve of 3-miRNA signature in training set. (a) Risk score, survival time and survival status, and 3-miRNA signature in the training set. (b) ROC curve and AUC of 3-miRNA signature in training set (GSE32960 data set). Abscissa means false positive fraction, ordinate means true positive fraction. (c) KM survival curve of 3-miRNA signature in training set (GSE32960 data set). Abscissa means time, ordinate means survival probability. (A color version of this figure is available in the online journal.)

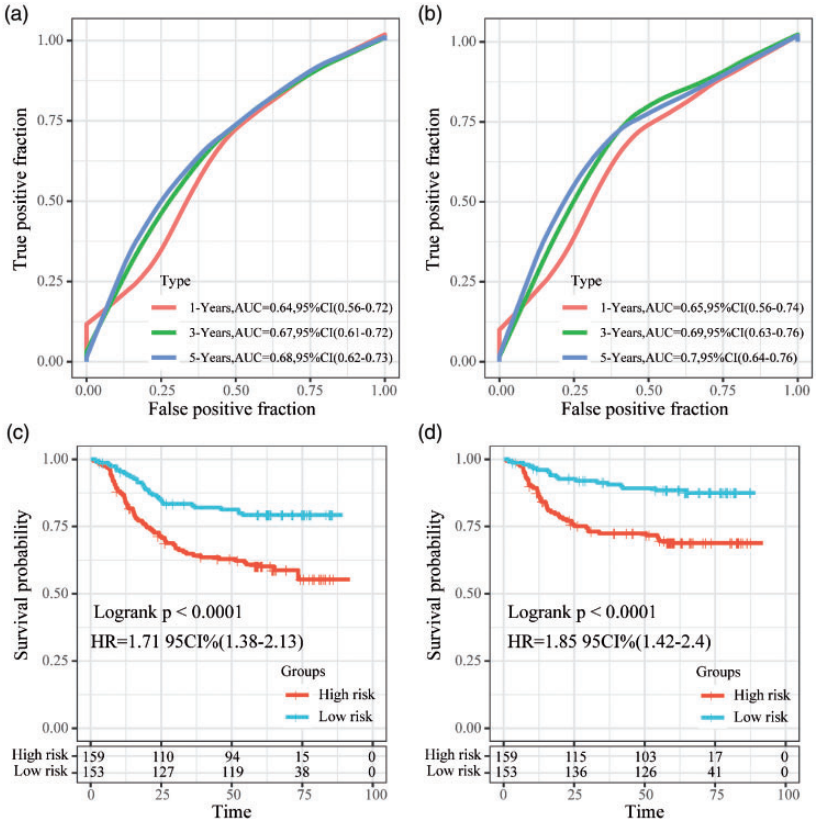

Figure 5. Prognostic capability of 3-miRNA signature. (a) DFS ROC curve and AUC of 3-miRNA signature in training set (GSE32960 data set). Abscissa means false positive fraction; ordinate means true positive fraction. (b) DMFS-ROC curve and AUC of 3-miRNA signature in training set (GSE32960 data set). Abscissa means false positive fraction; ordinate means true positive fraction. (c) KM-DFS curve distribution of 3-miRNA signature in training set (GSE32960 data set). Abscissa means time, ordinate means survival probability. (d) KM- DMFS curve distribution of 3-miRNA signature in training set (GSE32960 data set). Abscissa means time, ordinate means survival probability. (A color version of this figure is available in the online journal.)

patients with low-risk scores of the three miRNAs, patients with high-risk scores had a shorter OS. Similarly, high expression of hsa-miR-142-3p, hsa-miR-29c, and hsa-miR-30e were associated with low risk, which were protective factors (Figure 6(a)). ROC analysis was performed and the AUC of one years is 0.70 (Figure 6(b)). KM OS prognosis curve showed significant difference in prognosis (Figure 6 (c)). Those data implied that 3-miRNA signature had a good robustness.

Analysis of clinicopathological features in the training set

Clinicopathological data were collected from the GSE32960 data set, including age, Gender, T Stage, N Stage, Stage stage, and the differences between high- and low-risk groups were analyzed. Chi-square test results show that T staging of patients with nasopharyngeal cancer significantly related, reveal more low-risk patients T1 Stage, although we did not find significant differences in the Stage stage, but it can be observed from the Stage I to Stage IV high-risk patients increased in turn (Figure 7). Patients with different clinical features (age, male, Stage III + IV, T1 + T2, T3 + T4, N0 + N1, N2 + N3, chemotherapy and radiotherapy) can be significantly distinguished between high- and low-risk group using RiskScore, which further shows that our model in different clinical signs also still has good prediction ability (Figure 8).

Independence of 3-miRNA signature model

In order to identify the independence of the 3-miRNA signature model in clinical application, univariate and multivariate COX regression was used to analyze the relevant HR (95%CI of HR, *P* value) in the GSE32960 data set. In the TCGA data set, univariate COX regression analysis found that Age, T Stage, N Stage, Gender, Stage, and RiskScore were all significantly correlated with survival (Figure 9(a)), while the corresponding multivariate COX regression analysis found that RiskScore, Gender, T Stage, and N Stage were significantly correlated with survival. The results showed that 3-miRNA signature, Gender, T staging, and N staging were independent prognostic factors that could be used to predict the survival rate of nasopharyngeal cancer patients (Figure 9(b)).

A nomogram shows the results of the risk model visually and effectively. Based on the results of univariate and multivariate cox analysis, we constructed a nomogram model combining the four independent prognostic factors Gender, T Stage, N Stage, and RiskScore, and provided a quantitative method to predict the OS prognosis of nasopharyngeal cancer patients at three and five years (Figure 10(a)). In addition, the calibration diagram shows that the nomogram has similar performance to the ideal model (Figure 10(b)). Finally, the ROC analysis results also showed that our nomogram had a high potential for clinical application (Figure 10(c)). The results of DCA also

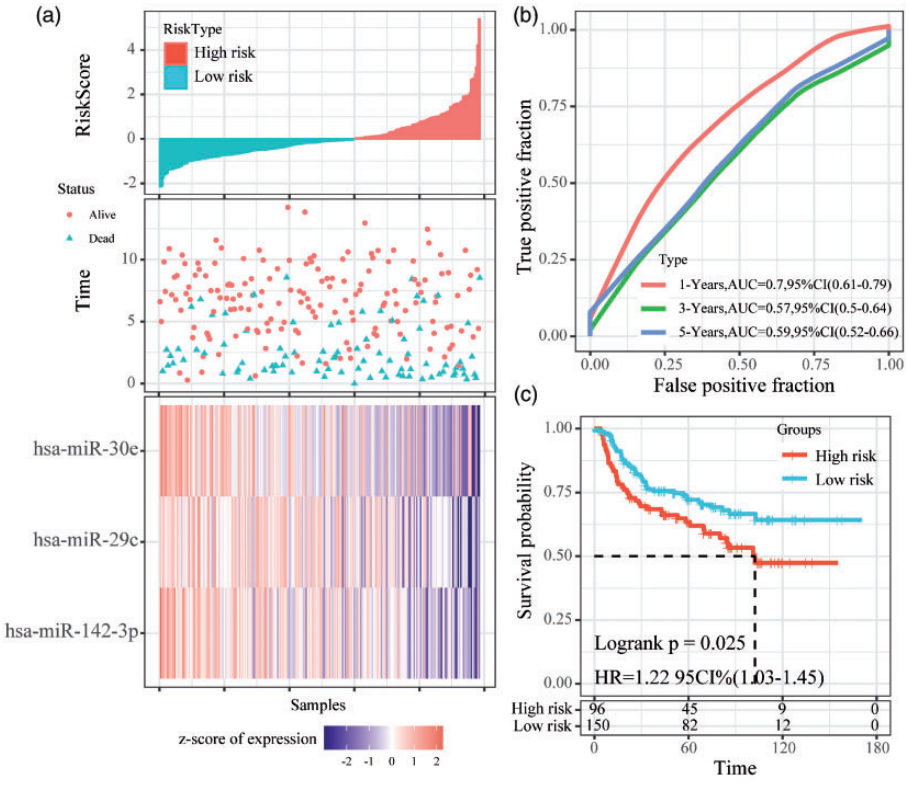

Figure 6. Validation of 3-miRNA signature prognosis. (a) Risk score, survival time and survival status, and 3-miRNA signature in the test set (GSE70970 data set). (b) ROC curve and AUC of 3-miRNA signature in test set (GSE70970 data set). Abscissa means false positive fraction; ordinate means true positive fraction. (c) KM survival curve of 3-miRNA signature in test set (GSE70970 data set). Abscissa means time, ordinate means survival probability. (A color version of this figure is available in the online journal.)

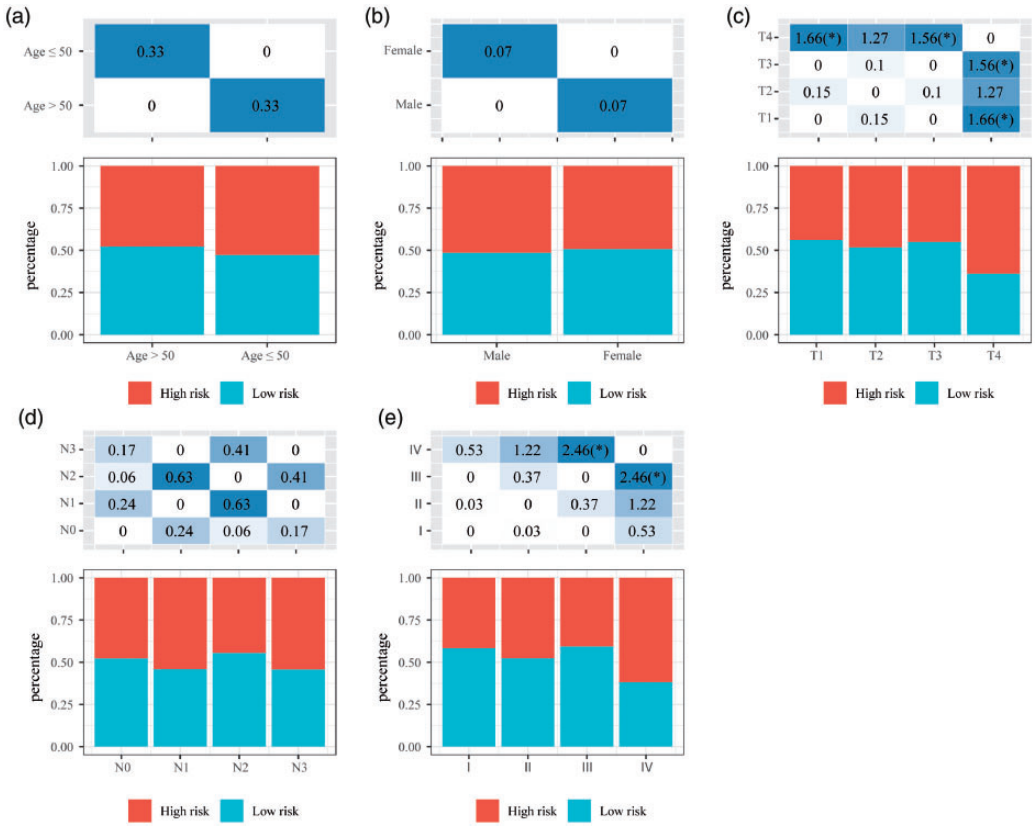

Figure 7. Analysis of clinicopathological features in the training set. (a) Sample distribution of different age groups (age > 50 and age ≤ 50) in high- and low-risk groups. (b) Sample distribution of different genders (female and male) in high- and low-risk groups. (c) Sample distribution of T Stage (T1–T4) in high- and low-risk groups. (d) Sample distribution of N Stage (N0–N3) in high- and low-risk groups. (e) Sample distribution of Stage stage (Stages I–IV) in high- and low-risk groups. (A color version of this figure is available in the online journal.)

showed that our nomogram had a high potential for clinical application (Figure 10(d)).

Advantages of 3-miRNA signature

By referring to the literature, we finally selected three prognostic risk models: 4-miRNA signature,[17] 4-miRNA signature,[18] and 4-miRNA signature[19] for comparison with our 3-miRNA model. In order to make the models comparable, we calculated the risk score of each nasopharyngeal cancer in the GSE32960 data using the same method based on the corresponding miRNAs in the three models. ROC of each model was evaluated, and the samples were divided into high-risk and low-risk groups according to the median risk score, and the differences in OS prognosis between the two groups were calculated. Only the KM curve results of Liu *et al.*'s model showed no significant difference, and the results of Zhao *et al.*'s model and Liu *et al.*'s model were found to be worse than our 3-miRNAs model, and Zhang *et al.*'s model was similar to our results through AUC comparison (Figure 11(a) to (c)). The restricted mean survival curves of these models were further compared. It can be seen that our model has the highest C-index among the four models, which is more advantageous in long-term survival prediction (Figure 11(d)). At the same time, we compared the 3-miRNA signature and the prediction effect of the three models through the DCA curve, and the results showed that the performance of our model was significantly better than that of Zhao and Liu *et al.*, and slightly better than that of Zhang *et al.* (Figure 11(e)).

Discussion

Considering that the abnormal expression of miRNA can affect the molecular functions and biological processes of a variety of tumors, many attempts have been made to use miRNA as a biomarker to accurately predict the diagnosis and prognosis of cancers.[20–23] Up to now, several studies have also focused on the discovery of miRNAs as biomarkers for nasopharyngeal cancer diagnosis. Huo Zhang *et al.* identified a 7-miRNA signature in plasma for nasopharyngeal cancer detection.[24] Zhang *et al.* successfully identified a four-miRNA signature using an integrated bioinformatics analysis for predicting the prognosis of patients with nasopharyngeal cancer.[19] Wen *et al.* identified two miRNA signatures (8-miRNA and 16-miRNA signatures) with high diagnostic accuracy for nasopharyngeal cancer from whole blood of patients.[25] The lack of consistency between these results and the limited overlap with our results may be due to different initial screening methods, different subject sizes, or sample processing methods. Secondly, due to the excessive miRNA in the published signatures, the clinical application is limited. Subsequently, we constructed a three-miRNA signature to predict the prognosis of nasopharyngeal cancer patients. The risk score of the three miRNAs revealed a better

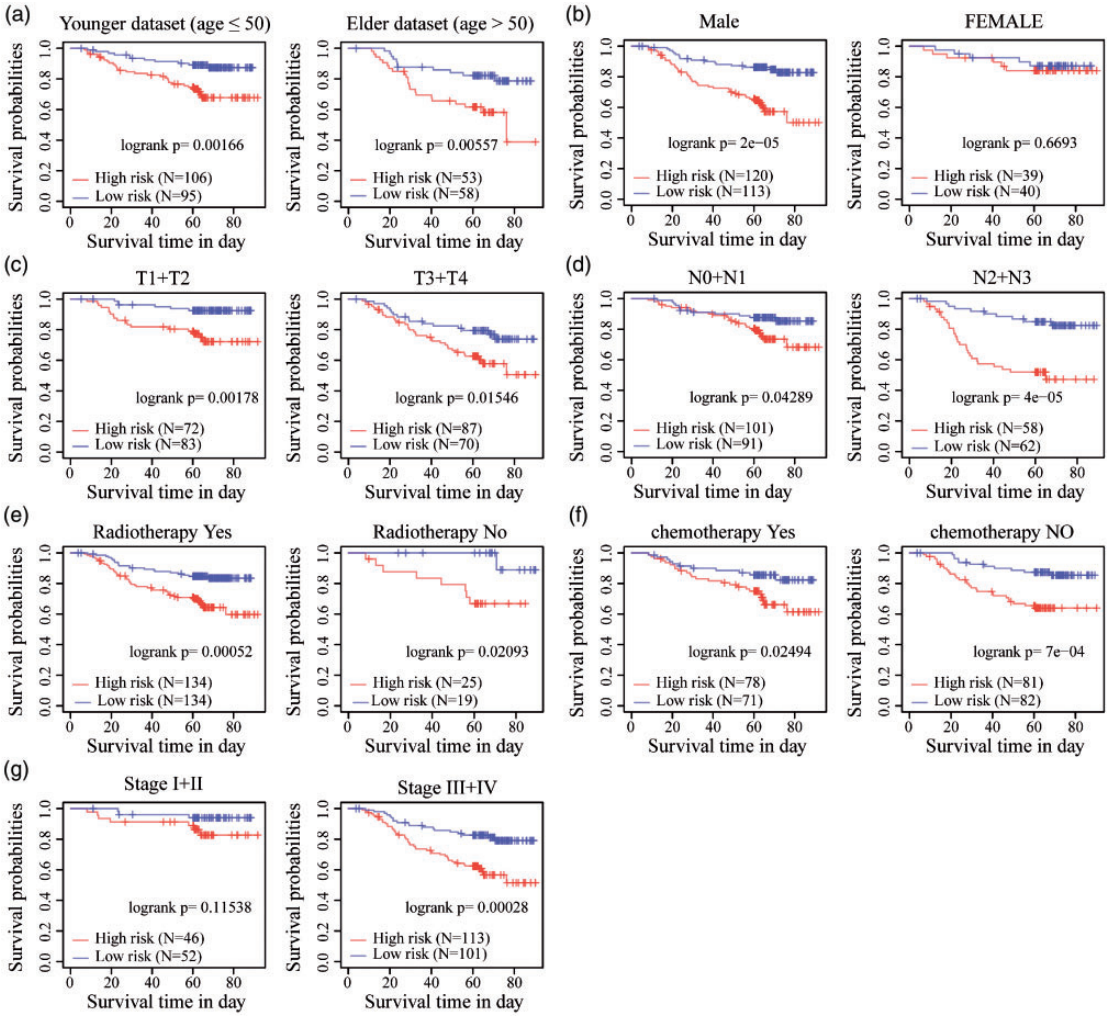

Figure 8. OS prognosis KM curve of clinical characteristics. (a) KM curve of age ≤ 50 y and age > 50 y. (b) KM curve of female and male. (c) KM curve of T1/T2 stage and T3/T4 stage. (d) KM curve of N0/N1 stage and N2/N3 stage. (e) KM curve of radiotherapy. (f) KM curve of chemotherapy. (g) KM curve of early stage (Stage I/II) and late stage (Stage III/IV). (A color version of this figure is available in the online journal.)

prediction of survival than did TNM Stage, T Stage, N Stage, and sex alone with regard to DMFS, OS, RFS, and DFS.

miRNA plays pivotal role in biological processes including cell proliferation, metastasis, differentiation, development, and apoptosis.[26,27] The ability to bind complementary sequences in 3′-UTR of various target mRNAs leading to direct mRNA degradation or translational repression.[28] It is speculated that the miRNAs identified in the manuscript may promote some important signaling pathways, such as nuclear factor-Kappa B activation and PI3K/AKT signal pathway to promote the progression of nasopharyngeal carcinoma, but the mechanism is unknown. Although potential biomarkers of miRNA have been identified for nasopharyngeal cancer diagnosis, the function of miRNA in nasopharyngeal cancer carcinogenesis and progression are still in its infancy. miR-142-3p is over-expressed in nasopharyngeal cancer tissues and cell lines, and miR-142-3p promotes cell proliferation in nasopharyngeal carcinoma by down-regulating SOCS6 expression.[29] But a study found that miR-142-3p as a key suppressive regulator in nasopharyngeal cancer metastasis.[30]miR-29c suppresses invasion and metastasis by targeting TIAM1 in nasopharyngeal carcinoma.[31] A study illustrated ectopic restoration of miR-29c substantially enhanced the sensitivity of nasopharyngeal cancer cells to IR and cisplatin treatment by promoting apoptosis.[32]miR-30e-5p was lowly expressed in nasopharyngeal cancer and inhibits proliferation and metastasis of nasopharyngeal carcinoma cells by targeting USP22.[33] Those evidences are still limited, but it could still indicate the feasibility of our signature, and our findings may provide for the functional studies of the three miRNAs found in nasopharyngeal carcinoma.

This study has several limitations. First of all, this study is based on GEO public data set, which is retrospective. Therefore, the performance of the 3-miRNA signature needs to be verified in future clinical studies. Second, the current diagnostic model only treats miRNA expression as a single set of data. Therefore, the combination of more molecular omics data, such as mRNA expression, CpG methylation, and genomic information, may help improve model accuracy. Third, more in-depth functional studies should be performed.

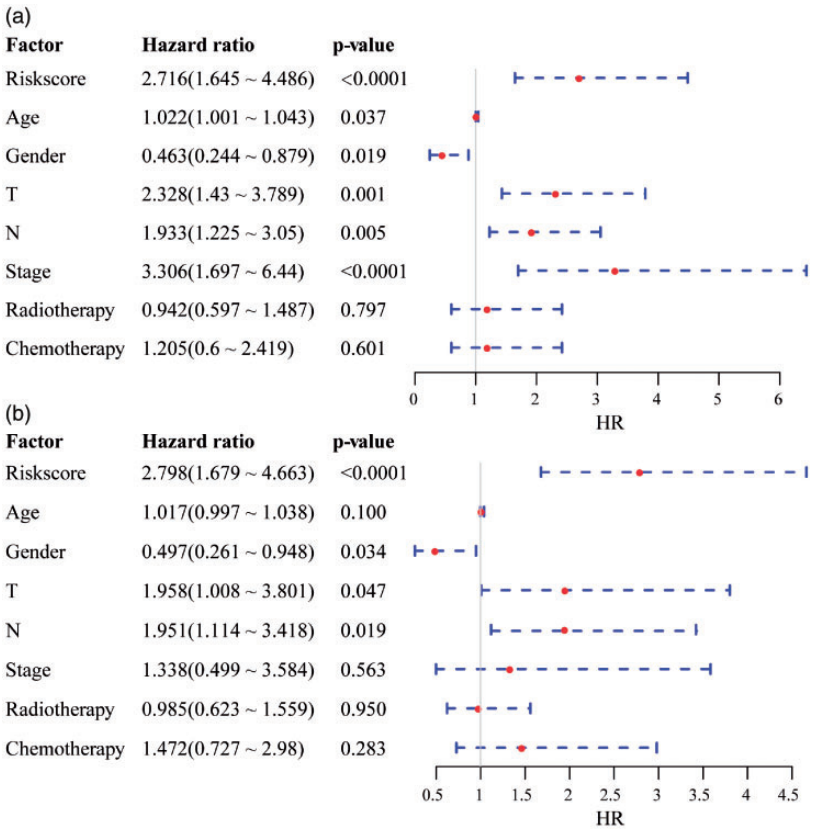

Figure 9. Independence of 3-miRNA signature. (a) Forest maps for univariate survival analysis. (b) Forest maps of multivariate survival analysis. (A color version of this figure is available in the online journal.)

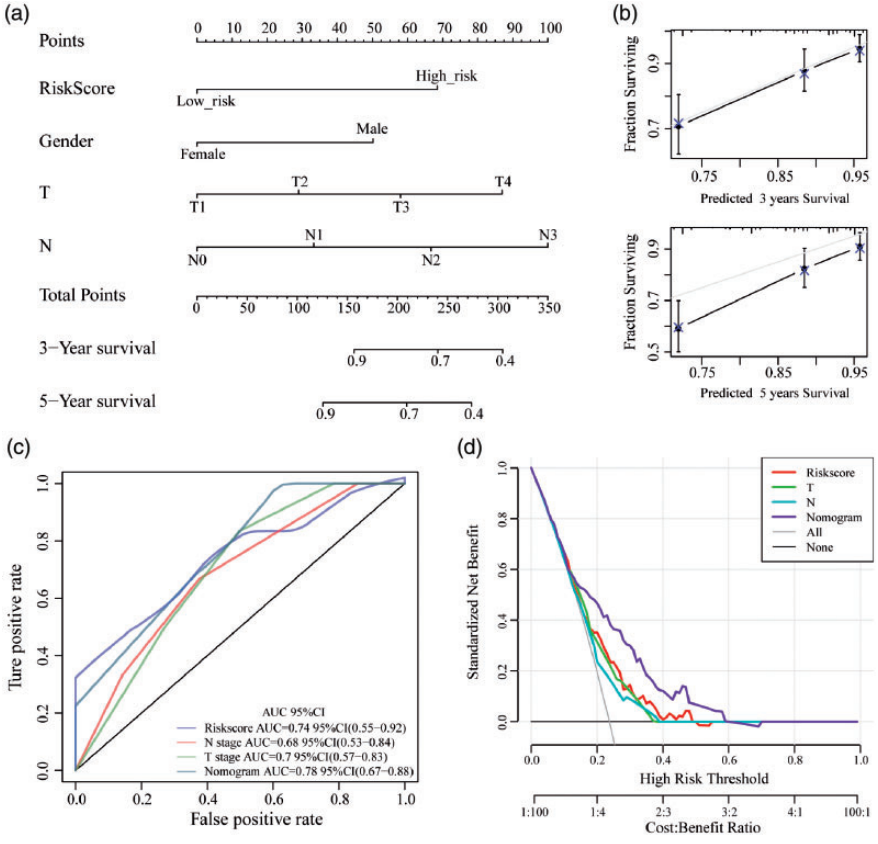

Figure 10. Independence of 3-miRNA signature model. (a) Nomogram of T + N + Gender + RiskScore. (b) The calibration plots for predicting patient three-year and five-year OS. Nomogram-predicted probability of survival is plotted on the x-axis; actual survival is plotted on the y-axis. (c) ROC curves of T, N, RiskScore, and Nomogram models. (d) DCA curve. (A color version of this figure is available in the online journal.)

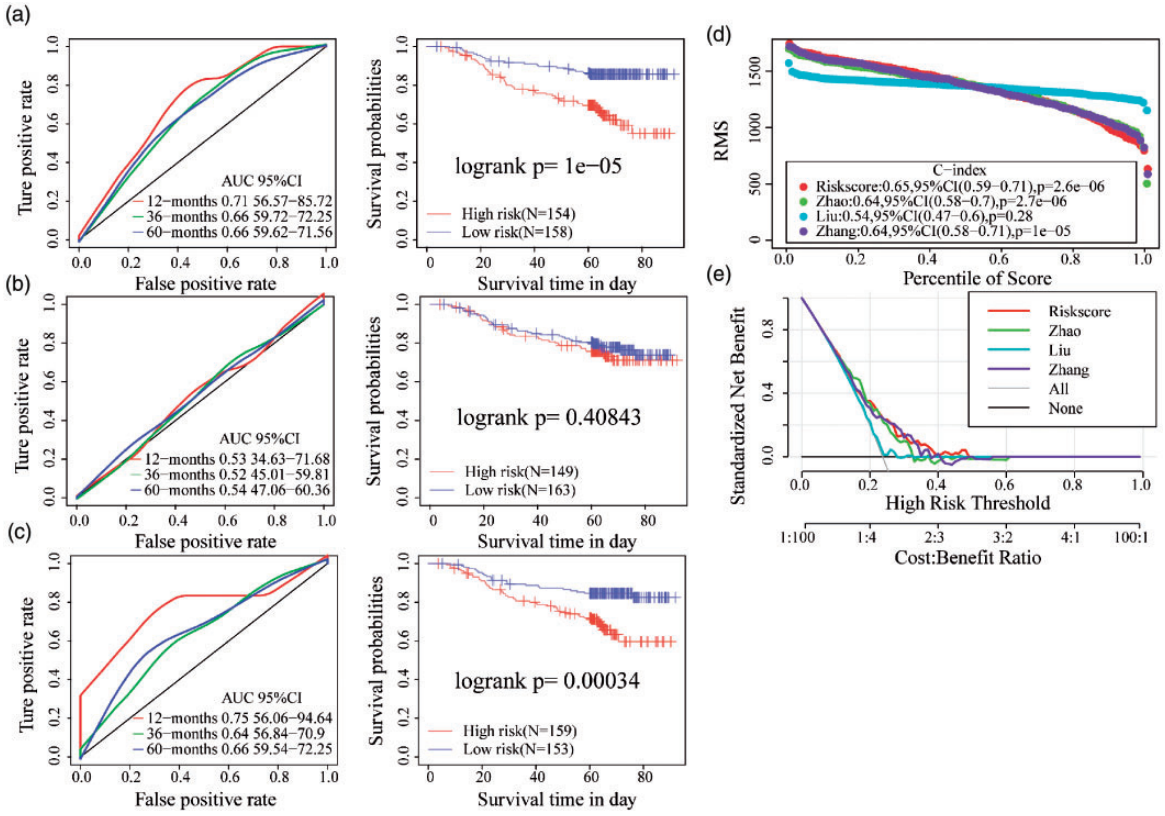

Figure 11. Superiority of the model. (a) AUC curve and prognosis KM curve of Zhao *et al*. model in training set. (b) AUC curve and prognosis KM curve of Liu *et al*. model in training set. (c) AUC curve and prognosis KM curve of Zhang *et al*. model in training set. (d) RMS curve of the four model. (e) DCA curve of the four models. (A color version of this figure is available in the online journal.)

Conclusions

In summary, we identified 3-miRNA signatures for nasopharyngeal cancer detection. Although virtual clinical application still has a long way to go, considering its convenience and low impact on health, miRNA panel could be combined with some traditional strategies in the near future to help disease screening and improve clinical outcome for nasopharyngeal cancer patients.

Authors' contributions: All authors participated in the design, interpretation of the studies and analysis of the data, and review of the manuscript. TYW and JW conducted the analysis and interpretation of data. TYW, YD, and JW performed statistical analysis. TYW drafted the manuscript. YW and YC acquisited of data. YW contributed to revision of manuscript for important intellectual content. HP and JCL performed conception and design of the research. HHL and HP obtained funding.

DECLARATION OF CONFLICTING INTERESTS

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by Science Project of Shanghai Municipal Commission of Health and Family Planning (201540173), Shanghai sailing program (16YF1403400), National Natural Science Foundation of China (81702643, 81541038, 81670905 and 81870702).

ORCID iD

Hu Peng https://orcid.org/0000-0002-5799-9703

SUPPLEMENTAL MATERIAL

Supplemental material for this article is available online.

REFERENCES

1. Lo KW, To KF, Huang DP. Focus on nasopharyngeal carcinoma. *Cancer Cell* 2004;**5**:423–8
2. Wei WI, Sham JS. Nasopharyngeal carcinoma. *Lancet* 2005;**365**:2041–54
3. Lin JC, Wang WY, Chen KY, Wei YH, Liang WM, Jan JS, Jiang RS. Quantification of plasma Epstein–Barr virus DNA in patients with advanced nasopharyngeal carcinoma. *N Engl J Med* 2004;**350**:2461–70
4. Dalton WS, Friend SH. Cancer biomarkers – an invitation to the table. *Science* 2006;**312**:1165–8
5. Esquela-Kerscher A, Slack FJ. Oncomirs – microRNAs with a role in cancer. *Nat Rev Cancer* 2006;**6**:259–69
6. Lebanony D, Benjamin H, Gilad S, Ezagouri M, Dov A, Ashkenazi K, Gefen N, Izraeli S, Rechavi G, Pass H, Nonaka D, Li J, Spector Y, Rosenfeld N, Chajut A, Cohen D, Aharonov R, Mansukhani M. Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma. *J Clin Oncol* 2009;**27**:2030–7

7. Negrini M, Nicoloso MS, Calin GA. MicroRNAs and cancer–new paradigms in molecular oncology. *Curr Opin Cell Biol* 2009;**21**:470–9
8. Deng M, Tang H, Zhou Y, Zhou M, Xiong W, Zheng Y, Ye Q, Zeng X, Liao Q, Guo X, Li X, Ma J, Li G. miR-216b suppresses tumor growth and invasion by targeting KRAS in nasopharyngeal carcinoma. *J Cell Sci* 2011;**124**:2997–3005
9. Xia H, Ng SS, Jiang S, Cheung WK, Sze J, Bian XW, Kung HF, Lin MC. miR-200a-mediated downregulation of ZEB2 and CTNNB1 differentially inhibits nasopharyngeal carcinoma cell growth, migration and invasion. *Biochem Biophys Res Commun* 2010;**391**:535–41
10. Xia H, Cheung WK, Sze J, Lu G, Jiang S, Yao H, Bian XW, Poon WS, Kung HF, Lin MC. miR-200a regulates epithelial-mesenchymal to stem-like transition via ZEB2 and beta-catenin signaling. *J Biol Chem* 2010;**285**:36995–7004
11. Lin C, Zong J, Lin W, Wang M, Xu Y, Zhou R, Lin S, Guo Q, Chen H, Ye Y, Zhang B, Pan J. EBV-miR-BART8-3p induces epithelial-mesenchymal transition and promotes metastasis of nasopharyngeal carcinoma cells through activating NF-κB and Erk1/2 pathways. *J Exp Clin Cancer Res* 2018;**37**:283
12. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets – update. *Nucleic Acids Res* 2013;**41**:D991–5
13. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics* 2019;**11**:123
14. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013;**32**:5381–97
15. Shim SR, Kim SJ. Intervention meta-analysis: application and practice using R software. *Epidemiol Health* 2019;**41**:e2019008
16. Eng KH, Schiller E, Morrell K. On representing the prognostic value of continuous gene expression biomarkers with the restricted mean survival curve. *Oncotarget* 2015;**6**:36308–18
17. Zhao L, Fong AHW, Liu N, Cho WCS. Molecular subtyping of nasopharyngeal carcinoma (NPC) and a microRNA-based prognostic model for distant metastasis. *J Biomed Sci* 2018;**25**:16
18. Liu N, Cui RX, Sun Y, Guo R, Mao YP, Tang LL, Jiang W, Liu X, Cheng YK, He QM, Cho WC, Liu LZ, Li L, Ma J. A four-miRNA signature identified from genome-wide serum miRNA profiling predicts survival in patients with nasopharyngeal carcinoma. *Int J Cancer* 2014;**134**:1359–68
19. Zhang S, Yue W, Xie Y, Liu L, Li S, Dang W, Xin S, Yang L, Zhai X, Cao P, Lu J. The fourmicroRNA signature identified by bioinformatics analysis predicts the prognosis of nasopharyngeal carcinoma patients. *Oncol Rep* 2019;**42**:1767–80
20. Yang Z, Yin H, Shi L, Qian X. A novel microRNA signature for pathological grading in lung adenocarcinoma based on TCGA and GEO data. *Int J Mol Med* 2020;**45**:1397–408
21. Zhao X, Cui L. A robust six-miRNA prognostic signature for head and neck squamous cell carcinoma. *J Cell Physiol* 2020; DOI: 10.1002/jcp.29723
22. Zhao E, Bai X. Nomogram based on microRNA signature contributes to improve survival prediction of clear cell renal cell carcinoma. *BioMed Res Int* 2020;**2020**:7434737
23. Wu C, Tong L, Wu C, Chen D, Chen J, Li Q, Jia F, Huang Z. Two miRNA prognostic signatures of head and neck squamous cell carcinoma: a bioinformatic analysis based on the TCGA dataset. *Cancer Med* 2020;**9**:2631–42
24. Zhang H, Zou X, Wu L, Zhang S, Wang T, Liu P, Zhu W, Zhu J. Identification of a 7-microRNA signature in plasma as promising biomarker for nasopharyngeal carcinoma detection. *Cancer Med* 2020;**9**:1230–41
25. Wen W, Mai SJ, Lin HX, Zhang MY, Huang JL, Hua X, Lin C, Long ZQ, Lu ZJ, Sun XQ, Liu SL, Yang Q, Zhu Q, Wang HY, Guo L. Identification of two microRNA signatures in whole blood as novel biomarkers for diagnosis of nasopharyngeal carcinoma. *J Transl Med* 2019;**17**:186
26. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004;**5**:522–31
27. Bushati N, Cohen SM. microRNA functions. *Annu Rev Cell Dev Biol* 2007;**23**:175–205
28. Shin VY, Chu KM. MiRNA as potential biomarkers and therapeutic targets for gastric cancer. *World J Gastroenterol* 2014;**20**:10432–9
29. Qi X, Li J, Zhou C, Lv C, Tian M. MiR-142-3p suppresses SOCS6 expression and promotes cell proliferation in nasopharyngeal carcinoma. *Cell Physiol Biochem* 2015;**36**:1743–52
30. Li Y, He Q, Wen X, Hong X, Yang X, Tang X, Zhang P, Lei Y, Sun Y, Zhang J, Wang Y, Ma J, Liu N. EZH2-DNMT1-mediated epigenetic silencing of miR-142-3p promotes metastasis through targeting ZEB2 in nasopharyngeal carcinoma. *Cell Death Differ* 2019;**26**:1089–106
31. Liu N, Tang LL, Sun Y, Cui RX, Wang HY, Huang BJ, He QM, Jiang W, Ma J. MiR-29c suppresses invasion and metastasis by targeting TIAM1 in nasopharyngeal carcinoma. *Cancer Lett* 2013;**329**:181–8
32. Zhang JX, Qian D, Wang FW, Liao DZ, Wei JH, Tong ZT, Fu J, Huang XX, Liao YJ, Deng HX, Zeng YX, Xie D, Mai SJ. MicroRNA-29c enhances the sensitivities of human nasopharyngeal carcinoma to cisplatin-based chemotherapy and radiotherapy. *Cancer Lett* 2013;**329**:91–8
33. Ma YX, Zhang H, Li XH, Liu YH. MiR-30e-5p inhibits proliferation and metastasis of nasopharyngeal carcinoma cells by target-ing USP22. *Eur Rev Med Pharmacol Sci* 2018;**22**:6342–9

(Received June 30, 2020, Accepted August 17, 2020)