Research Article

Development and Validation of an Immunotherapy-Related Prognostic Signature Based on Lymph Node Ratio for Gastric Cancer

Jianxin Li [ID], **Ting Han, Xin Wang, Yinchun Wang, Xuan Chen, Wangsheng Chen, and Qingqiang Yang** [ID]

Department of General Surgery (Gastrointestinal Surgery), The Affiliated Hospital of Southwest Medical University, Luzhou 646000, Sichuan, China

Correspondence should be addressed to Qingqiang Yang; yangqingqiang121@163.com

Background. The long-term prognosis of gastric cancer (GC) remains poor due to postoperative recurrence and metastasis. The increasing evidence show that the lymph node ratio (LNR) serves as an independent prognostic factor in patients with GC. In this study, we aimed to develop a prognostic signature for GC based on LNR. *Methods.* Survival analysis was conducted by comparing low- and high-LNR groups according to the optimal cutoff value of LNR, which was identified by receiver operating characteristic (ROC) curve analysis. Then, we identified the differentially expressed genes (DEGs) related to LNR in the training cohort of GC. Univariate Cox regression and least absolute shrinkage and selection operator (LASSO) regression were performed to construct the risk score signature. We then evaluated the risk score signature from the viewpoints of survival, clinic-pathological characteristics, tumor microenvironment (TME), tumor mutation burden (TMB), and immunotherapeutic and chemotherapeutic efficacy. *Results.* High LNR was significantly correlated with poorer prognosis and was an independent predictor of recurrence in patients with GC. Then, an eleven-gene signature that could predict the prognosis of GC patients was developed based on LNR-related DEGs in the training cohort, and the results were further confirmed in external independent cohort. In addition, the high-risk group showed aggressive clinicopathological characteristics, specific TME status, low TMB, and low immunotherapeutic sensitivity. *Conclusions.* The present study constructed an eleven-gene prognostic signature based on LNR to predict the prognosis of patients with GC and facilitate the development of individualized treatment strategy.

1. Introduction

Gastric cancer (GC) serves as one of the most commonly diagnosed solid malignant tumors and ranks fourth in cancer-related death rates around the world [1]. Surgery is the main treatment in GC, and the mortality of GC has indeed decreased with the development of medical treatment and surgical techniques over the past few decades. However, recurrence and metastasis remain the main causes of GC death, and the long-term prognosis for these patients is still unsatisfactory due to lack of effective therapeutic strategy [2]. Besides, the majority of GC patients are diagnosed at advanced tumor stage because of lack of specific early symptoms and biomarkers, which worsen the prognosis of GC [3]. Thus, exploring novel biomarkers

for the identification of higher-risk patients and directing the application of adjuvant therapy regimens is important.

Lymph node metastasis is the most common pathway of metastasis for GC, which results in poor prognosis. The 5-year survival rate for patients diagnosed with early stage is about 95%, while the 5-year survival rate sharply declines to not more than 45% for patients with advanced stage [4, 5]. Tumor invasion (T stage) and nodal status (N stage) are currently the most important prognostic factors in surgically treated GC, and the TNM staging system established by the American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC) has been widely used to determine the stage of GC and its prognosis [6]. N stage is determined by the number of regional lymph
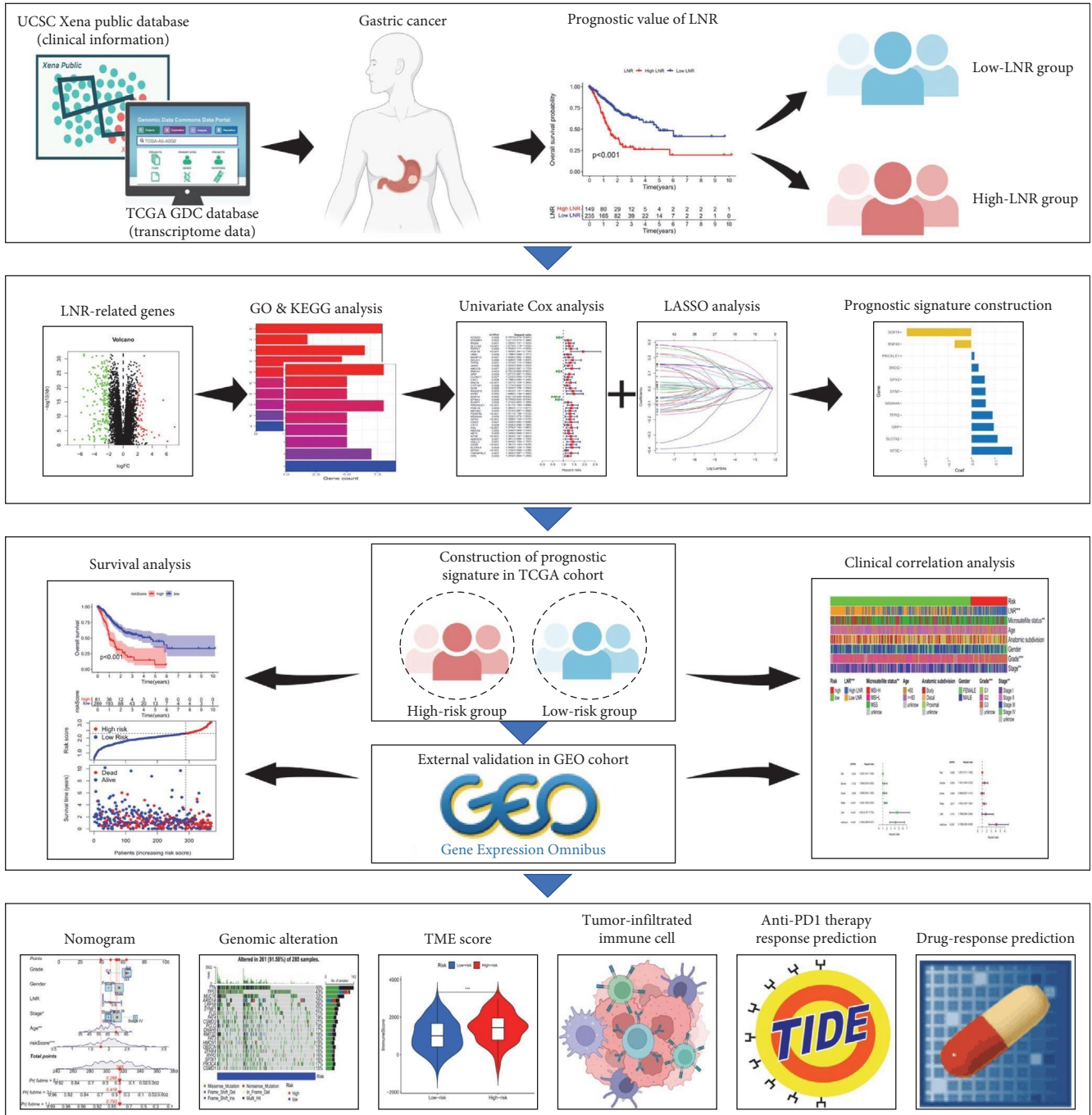
FIGURE 1: Flowchart of this study.

nodes with metastases and a minimum of 15 lymph nodes to be removed during gastrectomy is required for the accurate diagnosis of the tumor stage in GC [6]. However, certain factors have led to an insufficient number of lymph nodes being dissected in clinical practice, thereby hampering the clinical application of the AJCC staging system in GC [7]. Thus, the lymph node ratio (LNR) staging system, which is defined as the ratio of metastatic lymph nodes to the total lymph nodes examined, was developed and has been proposed as a valuable prognostic factor that is superior to the AJCC/UICC system [7, 8]. Emerging evidence exist that LNR

serves as an independent prognostic indicator of survival in patients with GC [9, 10]. In general, these studies indicated that the lymph node status is one of the key prognostic factors in GC, which deserves further exploration.

To date, no high-throughput studies have explored the potential prognostic roles of lymph node status-related signatures in GC. In the present study, we aimed to develop a novel prognostic signature for predicting the overall survival of GC patients based on LNR. To the best of our knowledge, this is the first study to construct a prognostic model correlated with LNR in GC. The workflow for this study is shown in Figure 1.

2. Materials and Methods

2.1. Data Collection.
Complete clinic-pathological parameters of 384 GC patients were obtained from the UCSC Xena (https://xenabrowser.net/datapages/) database [11, 12]. RNA-seq profiles of 375 GC tissues and 32 normal tissues were downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) [13]. Similarly, the mutation information of 431 GC patients was also downloaded from the database and combined into a mutation annotation format (MAF) file for further analysis. In addition, the microarray expression data and corresponding clinic-pathological information of 433 GC patients were downloaded from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) project for external validation cohort [14].

2.2. Identification of DEGs Related to LNR in GC.
LNR was defined as the number of metastatic lymph nodes divided by the total number of retrieved lymph nodes. The optimal cutoff value of LNR was calculated by ROC analysis, and patients with GC were divided into low- and high-LNR groups according to the cutoff value. Then, LNR-related genes which are differentially expressed between low- and high-LNR groups were identified using the "Limma" package in R based on the Wilcoxon rank-sum test. The screening criteria were defined as | log2 (fold change) | >0.585 and adjusted P-value < 0.05. Besides, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis on these LNR-related DEGs to get an insight into their biological function based on Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/) project [15].

2.3. Construction and External Validation of the Prognostic Signature.
Univariate Cox regression analysis was applied to assess the prognostic value of LNR-related genes, and genes with $P < 0.01$ were identified as prognostic genes and were incorporated into least absolute shrinkage and selection operator (LASSO) regression. LASSO regression is an algorithm usually used for fitting selecting variables in high-dimensional generalized linear models [16]. We then constructed a risk score model on the basis of the expression levels and the risk coefficients (β) of prognostic genes identified by LASSO regression. The formula of the risk score for each patient is listed as follows: risk score = $\beta_{(Gene1)} \times Exp_{(Gene1)} + \beta_{(Gene2)} \times Exp_{(Gene2)} + \ldots + \beta_{(Gene\gamma)} \times Exp_{(Gene\gamma)}$. GC patients were classified into low- and high-risk groups according to the optimal cutoff value of the risk score. Kaplan–Meier survival analysis, univariate Cox regression analysis, and multivariate Cox regression analysis were applied to assess the predictive capability of the risk score signature. In addition, independent cohort from the GEO database was used for external validation of the predictive performance of our constructed signature.

2.4. Clinical Relevance Analysis.
The correlations between the risk score and the clinic-pathological parameters of the training cohort were analyzed to determine the association of the LNR-based prognostic signature with other characteristics.

2.5. Construction of a Predictive Nomogram.
Nomograms are useful and accessible tools used for predicting survival and have been developed for predicting the prognosis of multiple tumor types [17]. In the present study, we constructed a nomogram to predict 1-, 3-, and 5-year survival of GC patients based on the clinical parameters and risk score signature. The calibration curve was conducted to assess the prediction probabilities of the nomogram. Subsequently, the ROC curve and the decision curve (DCA) were plotted to compare the performance of single and combined models [18].

2.6. Tumor Mutation Burden in Different Subgroups.
The "Maftools" package was applied to illustrate the genetic mutation differences between the low- and high-risk groups by waterfall plot [19]. Moreover, tumor mutation burden (TMB) of each patient was derived, and its correlation with risk score and prognosis of GC patients was evaluated.

2.7. Immune Microenvironment Analysis.
ESTIMATE algorithm assesses the level of infiltrating stromal and immune cells in the tumor microenvironment (TME) by calculating stromal and immune scores on the basis of specific gene expression signatures [20]. In this study, we performed Kaplan–Meier survival analysis to evaluate the correlation between the scores and the prognosis of GC patients. The correlation between the risk score and ESTIMATE score also be analyzed based on the Wilcoxon test. Besides, we explored the correlation between the tumor-infiltrated immune cell (TIIC) and risk score based on currently acknowledged algorithms, including XCELL, TIMER, QUANTISEQ, MCPcounter, EPIC, CIBERSORT, and CIBERSORT-ABS. Finally, the relationship between the risk score and immune checkpoint biomarkers (ICB) was assessed.

2.8. Tumor Immune Dysfunction and Exclusion Analysis.
Tumor immune dysfunction and exclusion (TIDE) is a computational algorithm to model two primary mechanisms of tumor immune escape—T-cell dysfunction and T-cell exclusion—which has been proved to be an effective method to predict immune checkpoint inhibitor response in cancer treatment [21]. Therefore, we applied TIDE platform (http://tide.dfci.harvard.edu) to calculate the TIDE scores of each GC patient and assessed their correlation with the LNR-based risk score signature.

2.9. Drug Sensitivity Prediction.
Drug sensitivities for GC patients were evaluated through the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org) project, a public resource for information on drug sensitivity in cancer cells and molecular markers of drug response [22]. The half-maximal inhibitory concentration (IC_{50}) values calculated by using the "pRRophetic" package in R were defined as the sensitivity indicators of the chemotherapy drugs, and the differences of IC_{50} values of the chemotherapy drugs between the low- and high-risk groups were estimated [23].

2.10. Statistical Analysis.
All statistical analyses were completed by the R software (version 4.1.1). Kaplan–Meier survival analysis and log-rank test were used to compare the

prognosis of different subgroups. The optimal cutoff for survival analysis was determined by the R package "survminer". Univariate and multivariate analyses were applied to identify independent prognostic indicators. The Wilcoxon rank-sum test was applied for the analysis of categorical variables. The χ^2 test was used to determine the differences in clinic-pathological features in different subgroups. A P-value < 0.05 was considered statistically significant.

3. Results

3.1. Clinical Significance of LNR in GC Patients. First, we used the value of the area under curve (AUC) to assess the predictive performance of LNR in all GC patients. The AUC value of LNR was 0.658 (95% CI (0.604–0.701)), and LNR of 0.318 at the maximum end was taken as the optimal cut-off value to distinguish low- and high-LNR groups (sensitivity = 60%, specificity = 66.34%) (Figure 2(a)). The time-ROC curve showed that the AUC value of 1-, 3-, and 5-years overall survival (OS) were 0.658, 0.694, and 0.661, respectively (Figure 2(b)). We also compared the AUC value of LNR with other clinical parameters and found that LNR had the most considerable AUC value (Figure 2(c)). Kaplan–Meier survival analysis showed that high LNR was significantly correlated with shorter OS and disease-free survival (DFS) than low LNR (Figures 2(d) and 2(e)). We further performed univariate and multivariate Cox regression to evaluate the prognostic effects of LNR and other common clinic-pathological features, including age, gender, grade, and stage. The results showed that LNR could serve as an independent prognostic factor to predict the OS of patients with GC (Figures 2(f) and 2(g)).

3.2. Identification of DEGs Related to LNR in GC. A total of 202 DEGs were identified between the low- and high-LNR groups with the criteria of | log2 (fold change) | >0.585 and adjusted P-value < 0.05. Among them, 184 DEGs were upregulated in the high-LNR groups, whereas 18 DEGs were downregulated (*Supplementary 1*). The expression pattern of these DEGs was depicted in the heatmap (Figure 3(a)), and the results also revealed that T stage, N stage, M stage, pathologic tumor stage, age, grade, vital status, and microsatellite status were different in a significant manner between the low- and high-LNR groups. As shown in Figure 3(b) and *Supplementary 2*, GO annotation analysis revealed these LNR-related genes were mainly enriched in immune response (GO:0006955), immunoglobulin production (GO:0002377), and inflammatory response (GO:0006954). KEGG pathway enrichment analysis found these genes were mainly enriched in the Wnt signaling pathway (hsa:04310) and calcium signaling pathway (hsa04020) (Figure 3(c) and *Supplementary 3*).

3.3. Construction and Verification of the LNR-Based Prognostic Signature. Forty-three DEGs significantly correlated with the prognosis of GC patients were identified by univariate Cox regression analysis (Figure 4(a)). Then, these prognosis-related genes were conducted into LASSO regression analysis to eliminate overfitting, and eleven feature genes (SOX14, RNF43, PRICKLE1, SNCG, GPX3, SYN1, MS4A4A, TFPI2,

GRP, SLC7A2, NT5E) were obtained to construct a prognostic signature (Figures 4(b) and 4(c)). The coefficients of each gene in the prognostic signature were displayed in Figure 4(d) and the risk score of each GC patient in the TCGA cohort was calculated. Ultimately, GC patients were classified into low- ($n = 289$) and high-risk ($n = 81$) groups according to the optimal cut-off value of risk score (2.308). The Kaplan–Meier survival analysis found that the OS of the high-risk group was significantly worse than that of the low-risk group (Figure 5(a)). The risk plot and scatter plot illustrated that the survival rate and survival time decreased with the increase in the risk score. (Figure 5(b)). In addition, univariate and multivariate Cox analyses showed that the risk score serves as a prognostic factor and was independent of age, gender, grade, tumor stage, and LNR (Figures 5(c) and 5(d)).

Moreover, we performed external validation to evaluate whether the LNR-based risk signature has clinical application value. We calculated the risk score of each patient in the external validation cohort according to the calculation formula derived from TCGA. Then, GC patients were divided into low- and high-risk groups according to the optimal cut-off value of risk score and the prognosis between the patients in low- and high-risk groups showed to be a significant difference (Figure 5(e)–5(h)). These results confirmed the clinical application value of the LNR-based signature.

3.4. Correlation of the LNR-Based Prognostic Signature with Clinical Features. The correlation between the risk score and the clinical parameters of the GC patients was then analyzed. As shown in the band diagram (Figure 6(a)) and scatter plots (Figure 6(b)), there were positive correlations between the risk score and tumor grade, LNR, microsatellite instability (MSI) status, and tumor stage. In general, these results indicate that high-risk score was significantly correlated with aggressive clinic-pathological characteristics.

3.5. Construction of a Nomogram to Predict Prognosis. Taking all the clinic-pathological characteristics and risk scores into consideration, we constructed a comprehensive nomogram to predict the prognosis of GC (Figure 7(a)). The nomogram showed that the risk score contributed significantly to the prognosis. The corresponding calibration plots in 1-, 3-, and 5-year proved that the performance of the nomogram was best in predicting the OS of GC patients (Figure 7(b)). Additionally, the prognostic capacity of the nomogram and other single features were demonstrated by the AUC value of the ROC curve. The results revealed that the nomogram showed the largest AUC (0.755) compared with other single features (Figure 7(c)). Finally, the DCA analysis demonstrated that the nomogram showed the best net benefit for predicting the prognosis of GC (Figure 7(d)).

Taken together, these results suggest that our constructed nomogram confers excellent predictive potential for prognosis of GC patients.

3.6. Somatic Mutations in Different Subgroups Based on LNR-Based Signature. Somatic mutation serves as one of the significant elements in carcinogenesis and progression. Thus, we assessed the correlation between the LNR-based signature
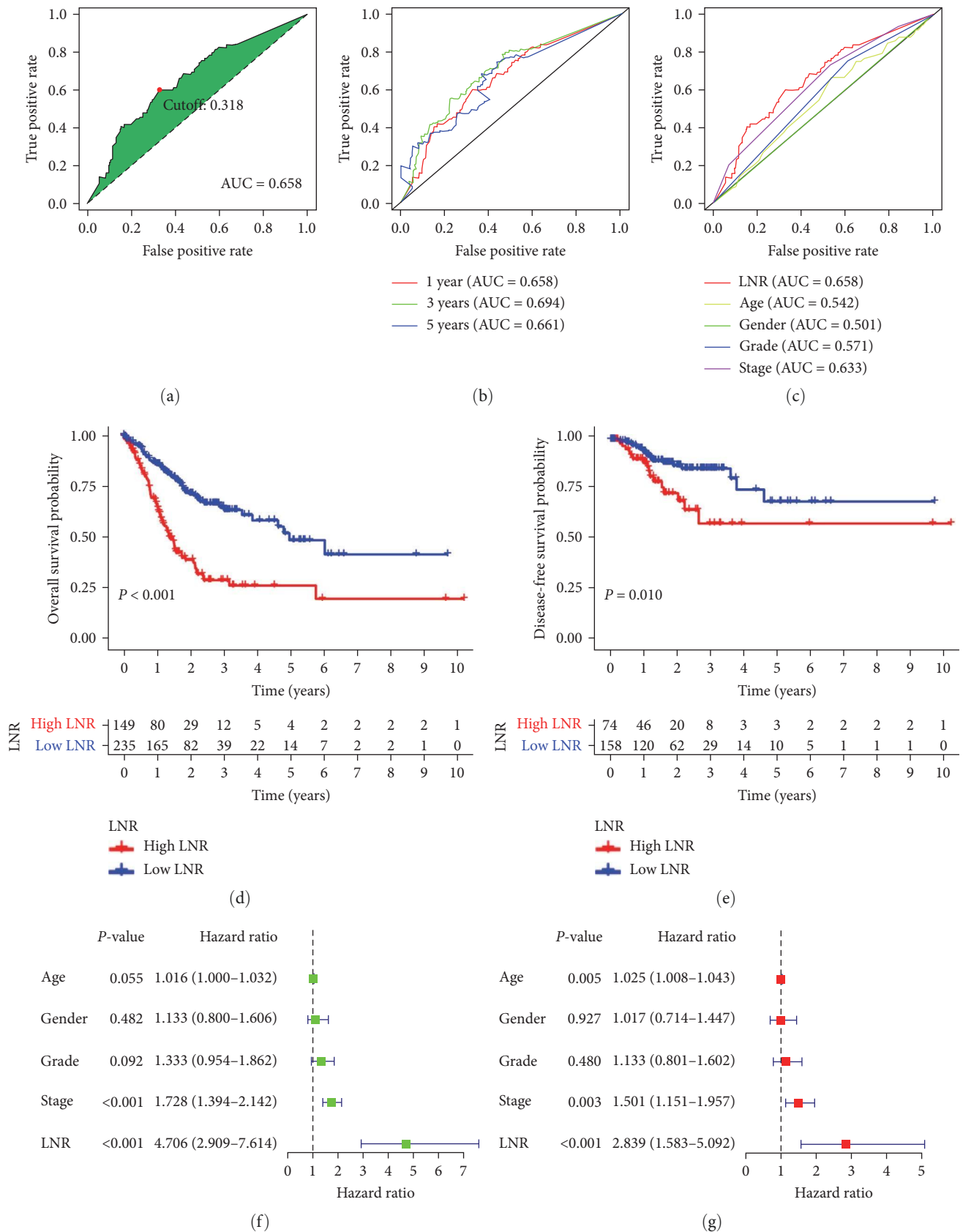
FIGURE 2: Clinical significance of LNR in GC. (a) The ROC curve for the LNR and the threshold (0.318) was taken as the cut-off value to distinguish the high- and low-risk of samples. (b) The AUC values of the 1-, 3-, and 5-year ROC curves of the LNR. (c) ROC curve analyses in predicting the overall survival of patients for LNR, age, gender, grade, and stage. (d) The overall survival time of the high-LNR group was significantly shorter than that of the low-LNR group; (e) The disease-free survival time of the high-LNR group was significantly shorter than that of the low-LNR group. (f) Univariate Cox regression of LNR and clinical parameters in GC. (g) Multivariate Cox regression of LNR and clinical parameters in GC.

Vital status***
- Deceased
- Living
- Unknown

Stage***
- Stage I
- Stage II
- Stage III
- Stage IV
- Unknown

T***
- T1
- T2
- T3
- T4
- Unknown

N***
- N0
- N1
- N2
- N3

M***
- M0
- M1

Grade***
- G1
- G2
- G3
- Unknown

Gender
- Female
- Male

Anatomic subdivision
- Body
- Distal
- Proximal
- Unknown

Age*
- <60
- ≥60
- Unknown

Microsatellite status**
- MSI-H
- MSI-L
- MSS
- Unknown

Cluster
- High LNR
- Low LNR

(a)

FIGURE 3: Continued.

and mutation profile in GC patients. As shown in the box plot, the level of TMB was significantly higher in the low-risk group than that in the high-risk group (Figure 8(a)), and the Pearson correlation analysis corroborated that the risk score was negatively correlated with TMB ($R = -0.32$, $P = 2.7$e-10) (Figure 8(b)). Survival analysis revealed that GC patients with high TMB showed better prognosis than those with low TMB (Figure 8(c)). Moreover, we further evaluated the effect of risk score and TMB on OS of GC patients considering the synergistic effect of them and found that the predictive

ability of risk score was independent of TMB (Figure 8(d)). Genes that had the most frequent mutation in low- and high-risk groups were shown in Figures 8(e) and 8(f).

In sum, the above results demonstrate that the risk score is a predictive biomarker for prognosis of GC patients that is independent of TMB and can effectively predict the TMB status.

3.7. The Relationship between the Prognostic Signature and Immune Microenvironment. We applied ESTIMATE algorithm to calculate the TME scores (immune score, stromal score, and
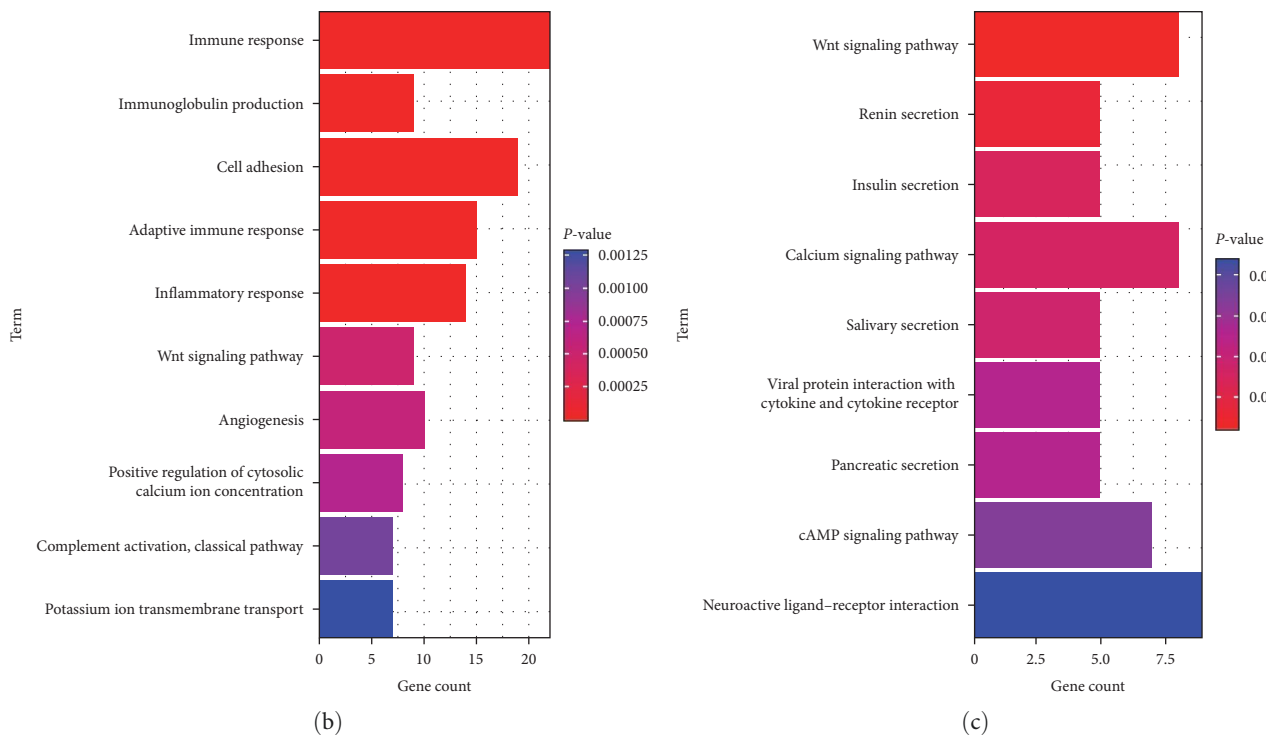
(b)

(c)

FIGURE 3: Identification of LNR-related gene. (a) Clinical heatmap of clinical parameters and DEGs between high- and low-LNR groups. (b) GO terms of LNR-related genes significantly enriched in GC. (c) KEGG pathways of LNR-related genes significantly enriched in GC.
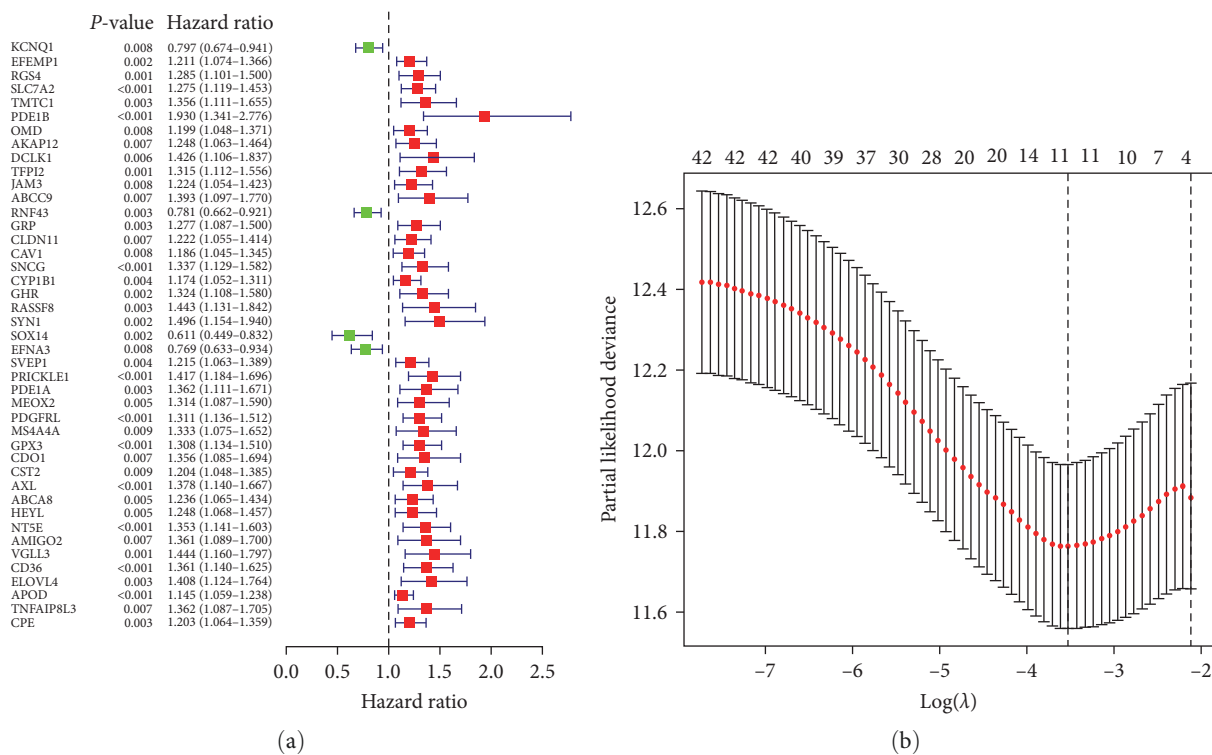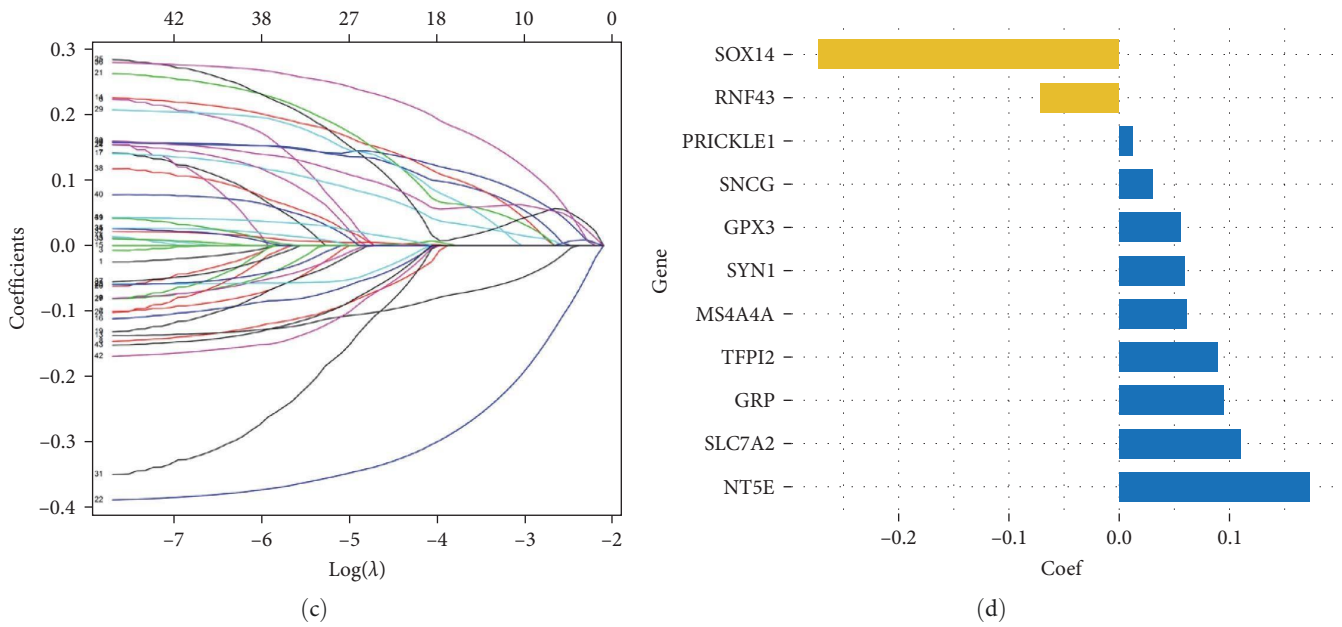

(a)

(b)

FIGURE 4: Continued.

(c)

(d)

FIGURE 4: Construction of the risk signature using LNR-related genes. (a) Univariate Cox regression analysis identified 43 LNR-related genes correlated with overall survival. (b, c) Eleven LNR-related genes were identified for the construction of a prognostic signature by a LASSO regression analysis. (d) The regression coefficient value of each gene involved in the risk signature.

(a)

(b)
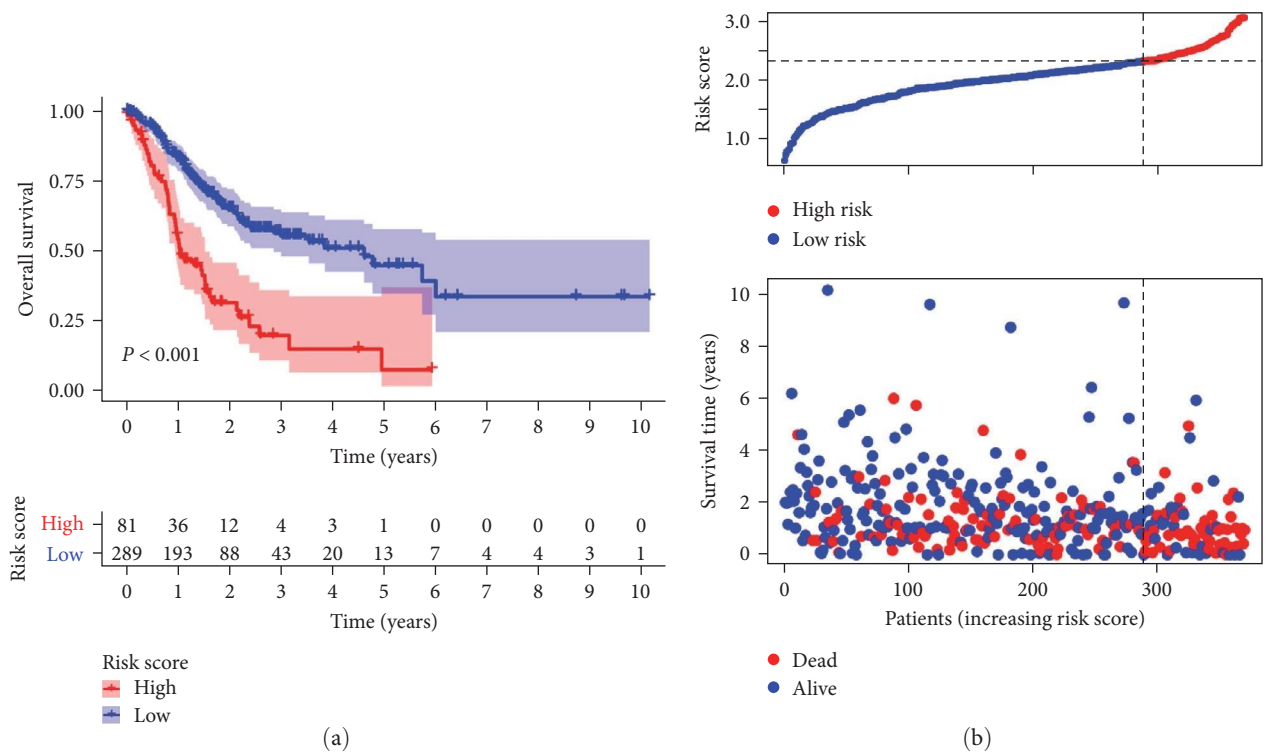
FIGURE 5: Continued.

ESTIMATE score) of each GC patient in the TCGA project. Survival analysis found that higher stromal score and ESTIMATE score were significantly correlated with a much poorer OS rate among GC patients, while immune score showed no significant correlation with OS (Figure 9(a)–9(c)). Then, we explored the correlation between the patient's TME scores and risk scores. The results revealed that immune score, stromal score and ESTIMATE score of the high-risk group were obviously higher than those of the low-risk group (Figure 9(d)–9(f)). In terms of TIIC, Spearman correlation analysis indicated that

(c)

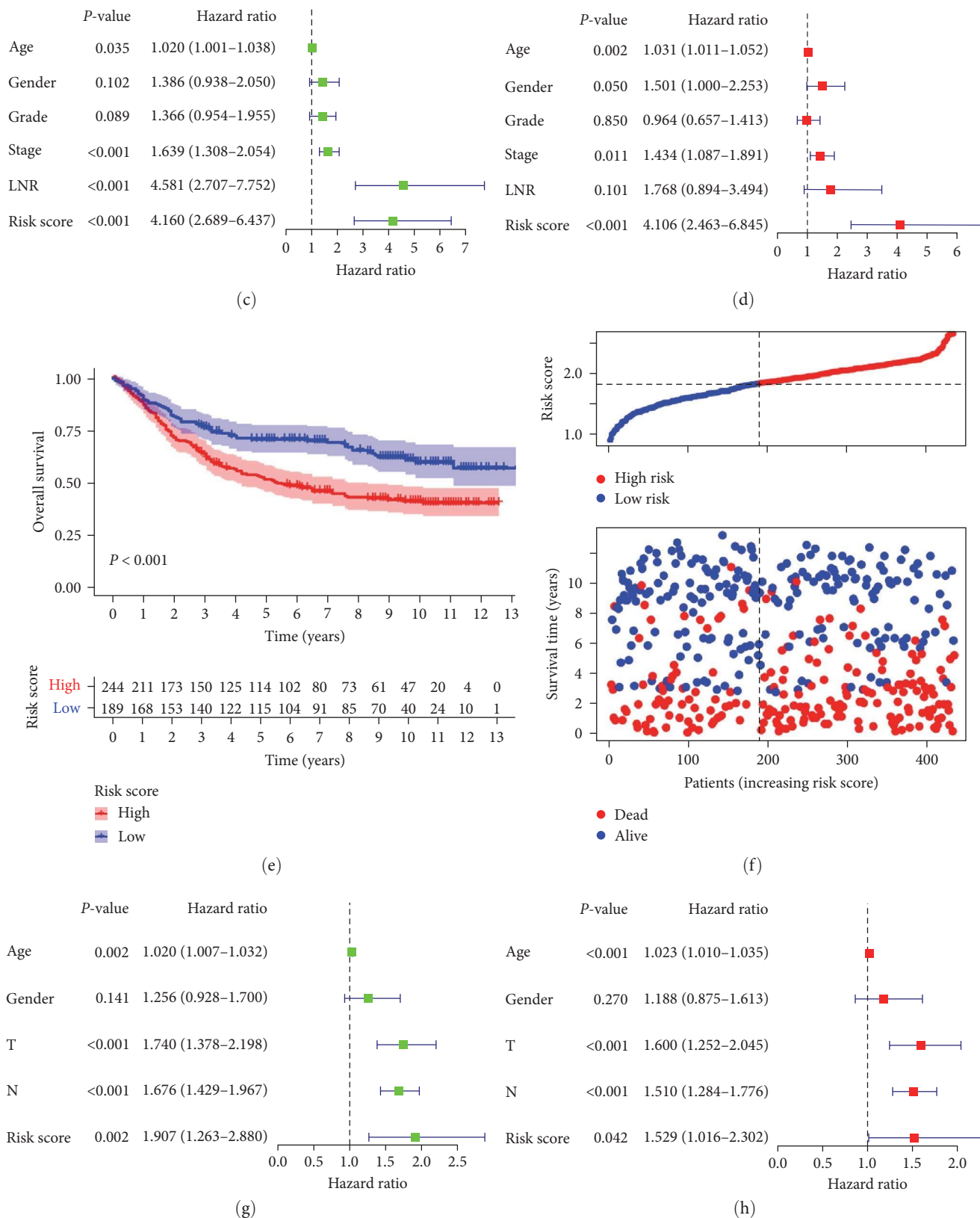(d)

(e)

(f)

(g)

(h)

Figure 5: Evaluation and external validation of the prognostic signature in predicting overall survival of GC. (a) The Kaplan–Meier survival analysis of the signature for predicting the overall survival of GC patients in TCGA cohort. (b) The distribution of the risk score and the survival status of GC patients in TCGA cohort. (c) Univariate Cox regression of risk score and clinical parameters of GC in TCGA cohort. (d) Multivariate Cox regression of risk score and clinical parameters of GC in TCGA cohort. (e) The Kaplan–Meier survival analysis of the signature for predicting the overall survival of GC patients in GEO cohort. (f) The distribution of the risk score and the survival status of GC patients in GEO cohort. (g) Univariate Cox regression of risk score and clinical parameters of GC in GEO cohort. (h) Multivariate Cox regression of risk score and clinical parameters of GC in GEO cohort.
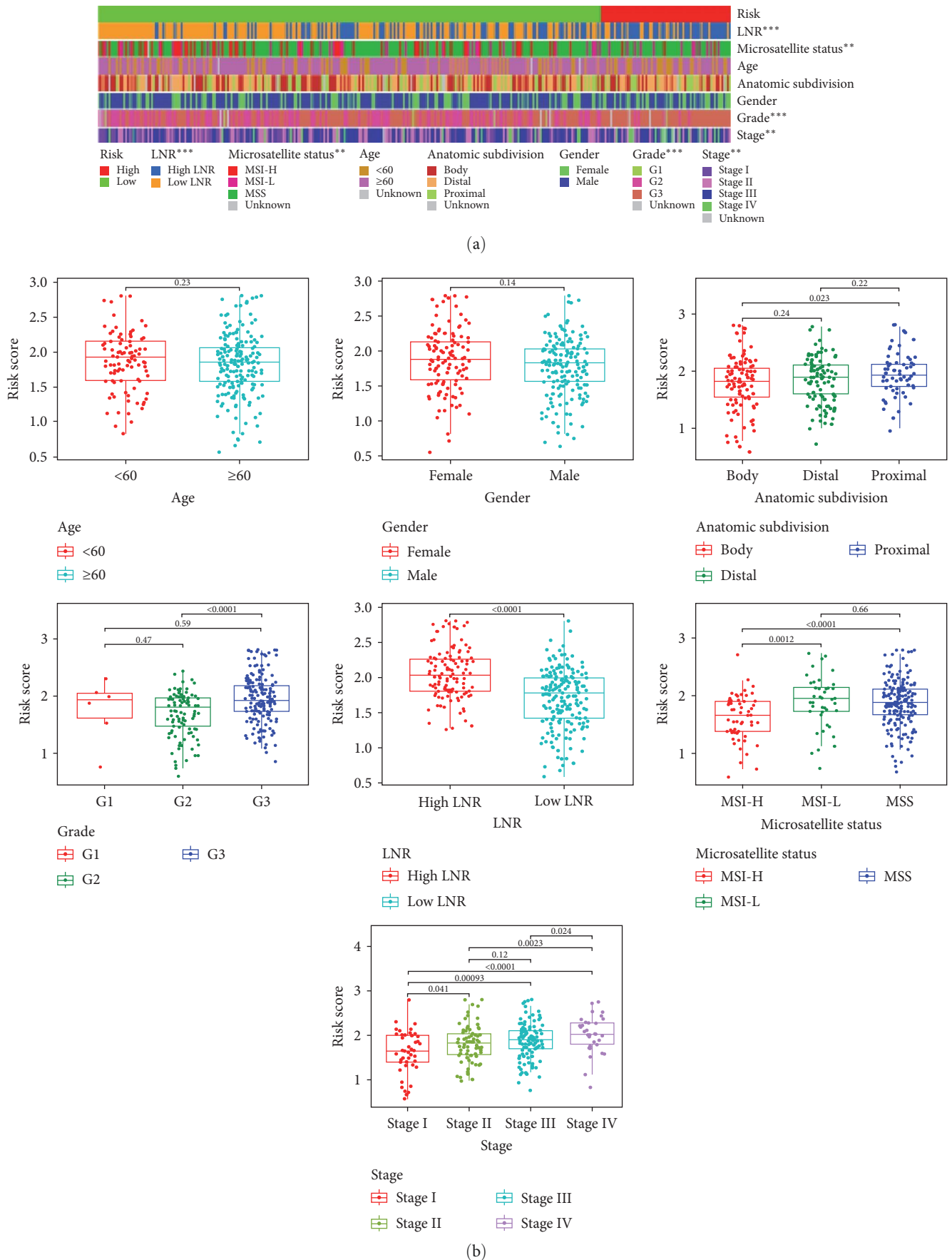
FIGURE 6: The correlation between the risk score and clinical parameters. (a) A strip chart showed the differences of clinical parameters between low- and high-risk groups. (b) The scatter diagram showed the correlations of risk score with clinical parameters in GC. **$P < 0.01$; ***$P < 0.001$.
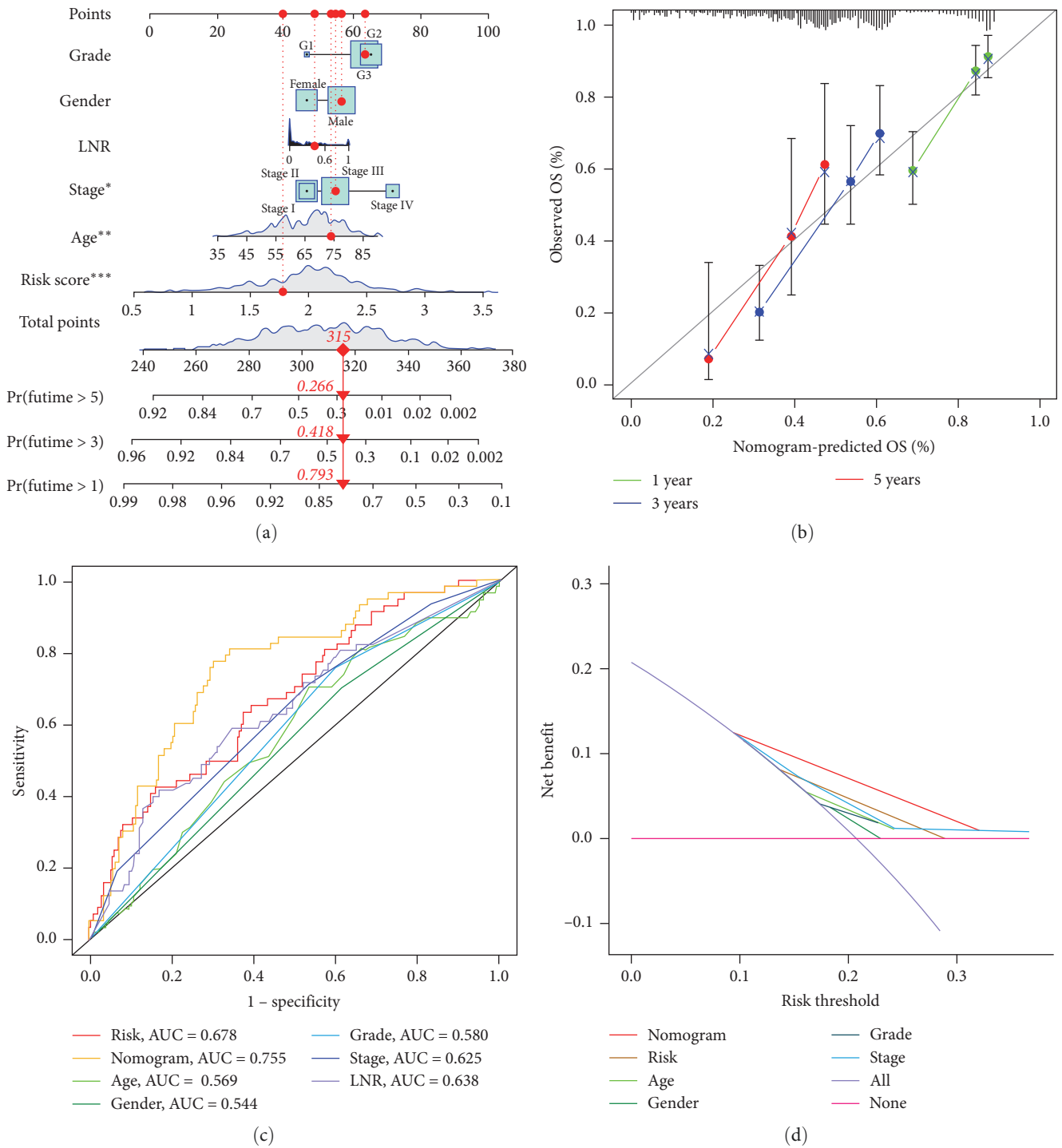
(a)

(b)

(c)

(d)

FIGURE 7: Nomogram for predicting overall survival of GC patients and decision curve analysis. (a) Nomogram to predict 1-, 3-, and 5-year overall survival probability by integrating the LNR-based risk score and clinicopathologic parameters. (b) Plot depicting the calibration of the nomogram in terms of the agreement between predicted and observed outcomes. Nomogram performance is shown by the plot relative to the dotted line, which represents perfect predictive ability. (c) A comparison of ROC curves showed the superiority of the nomogram and the risk signature. (d) Decision curve analysis of the nomogram. None: assume an event will occur in no patients; all: assume an event will occur in all patients. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

the high-risk group was more positively correlated with TIICs (Figure 9(g)). As expected, we also found the high-risk group was more positively correlated with ICBs, which usually are expressed on the surface of TIIC and suppressed its function (Figure 9(h)).
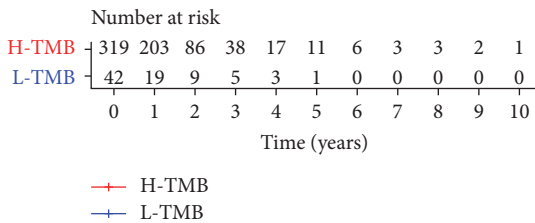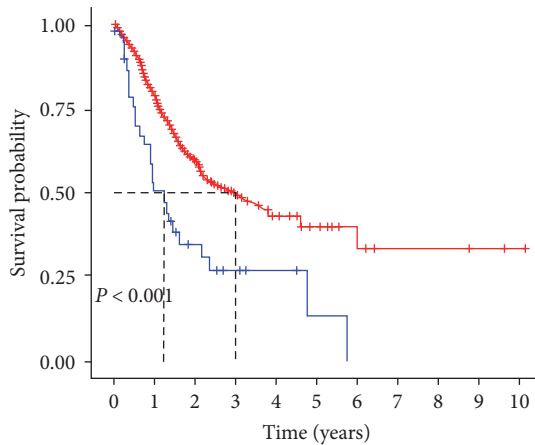
3.8. Prediction of Immunotherapy and Chemotherapy Responsiveness. We further investigated the correlation of the LNR-based signature with immunotherapy and chemotherapy responsiveness in GC. Patients in the high-risk group were featured by higher TIDE scores ($P < 0.001$, Figure 10(a)),
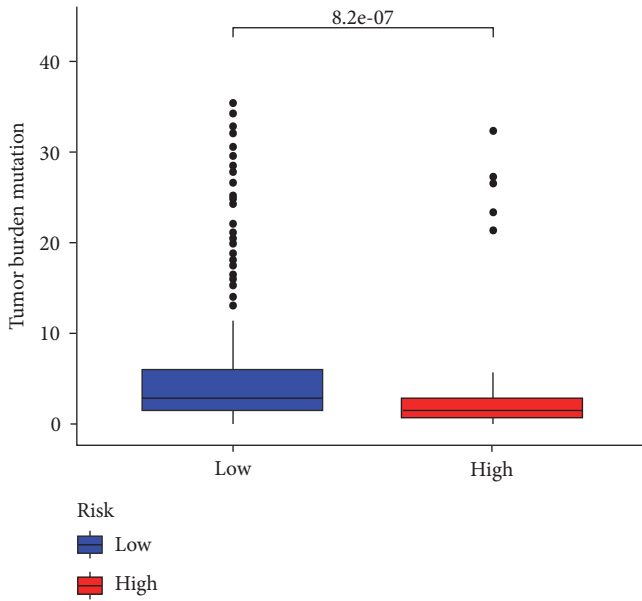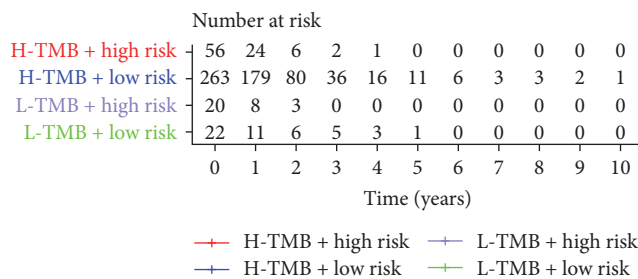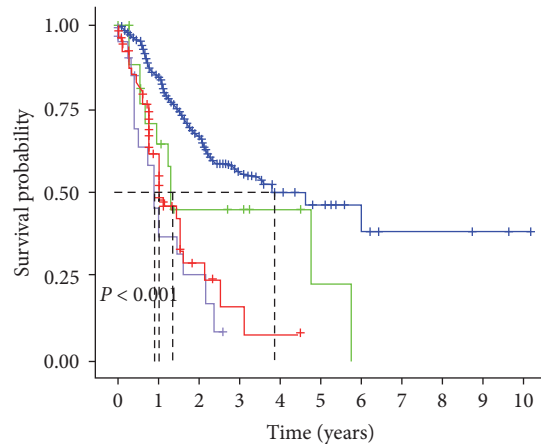
(a)
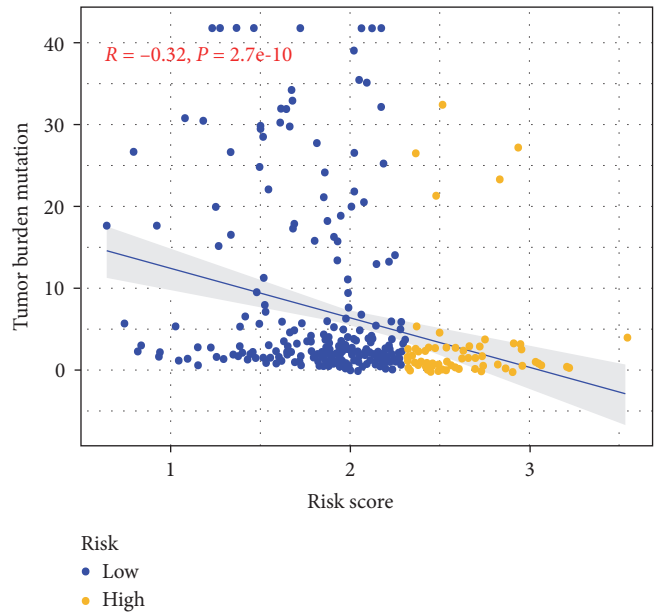
(b)

(c)

(d)

FIGURE 8: Continued.

higher dysfunction score ($P < 0.001$, Figure 10(b)), and higher exclusion score ($P < 0.001$, Figure 10(c)) compared with those in the low-risk group. Besides, the low-risk group was predicted to hold a higher proportion of immunotherapeutic responders compared with the high-risk counterpart (Figure 10(d)). Furthermore, the patients in the high-risk group were featured by lower proportion of MSI-H

(Figure 10(e)) and high-TMB (Figure 10(f)) compared with those in the low-risk group. These findings indicate that GC patients with high-risk scores exhibited less responsiveness to the immunotherapy, which may be responsible for the poor prognosis. The drug sensitivity analyses found that the IC$_{50}$ of bleomycin, JNK inhibitor VIII, lapatinib, paclitaxel, and sunitinib were remarkably decreased in patients with
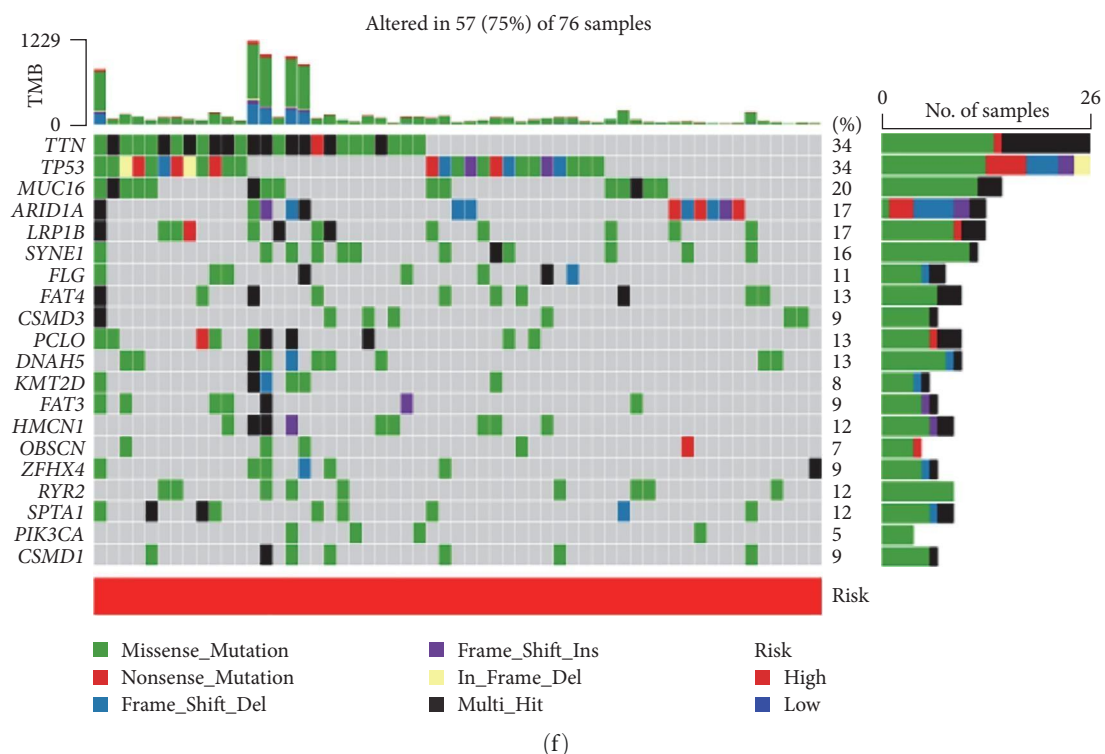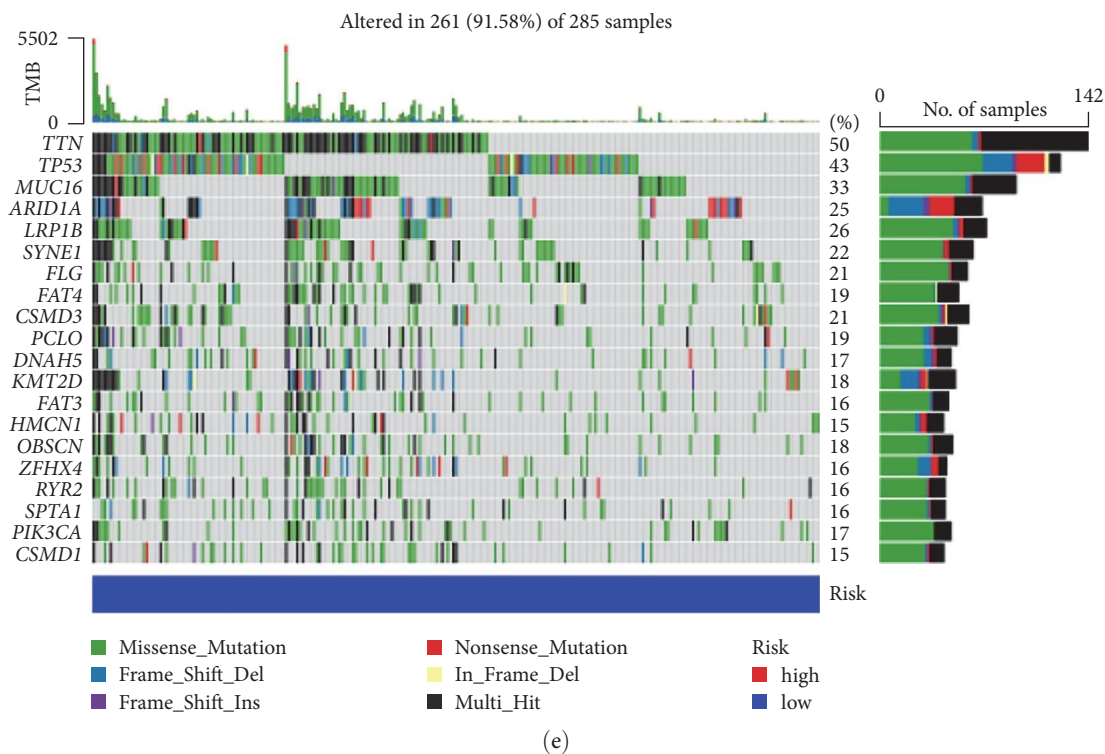
(e)

(f)

FIGURE 8: The correlation between the risk score and TMB. (a) The difference of TMB between low- and high-risk groups. (b) Pearson correlation analysis between TMB and risk score. (c) The association of TMB and GC patients' overall survival. (d) The KM curve showing overall survival after combining the risk score with TMB. (e) Mutation profile of GC patients in low-risk group. (f) Mutation profile of GC patients in high-risk group.

high-risk scores, while the IC_{50} of methotrexate in high-risk group was significantly elevated (Figure 10(g)). These results suggest that the LNR-based signature has the potential ability to predict the sensitivity of chemotherapy drugs.

4. Discussion

GC ranks as the fourth leading cause of cancer-related death rate around the world, which seriously threatens the health
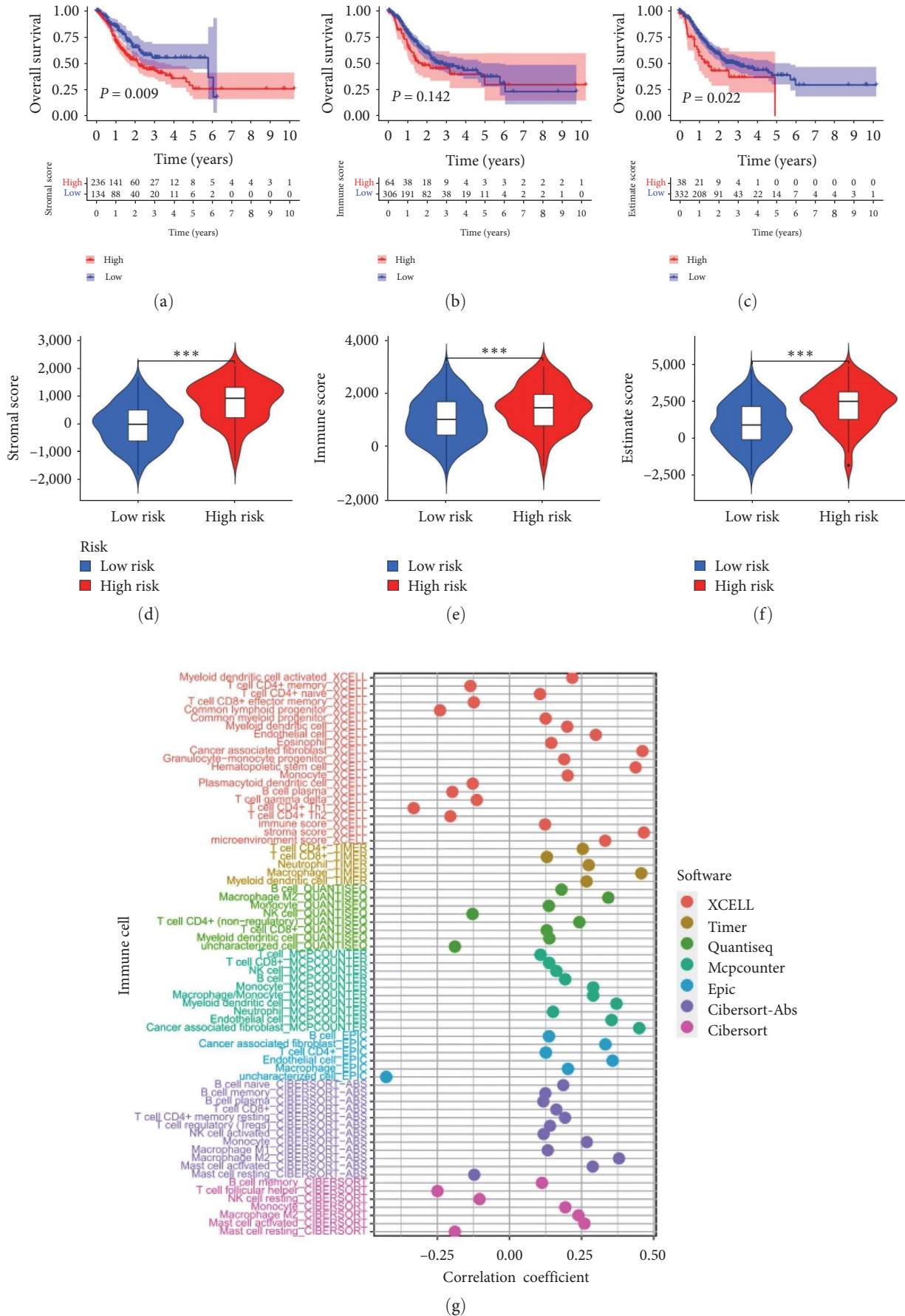
(a)

(b)

(c)

(d)

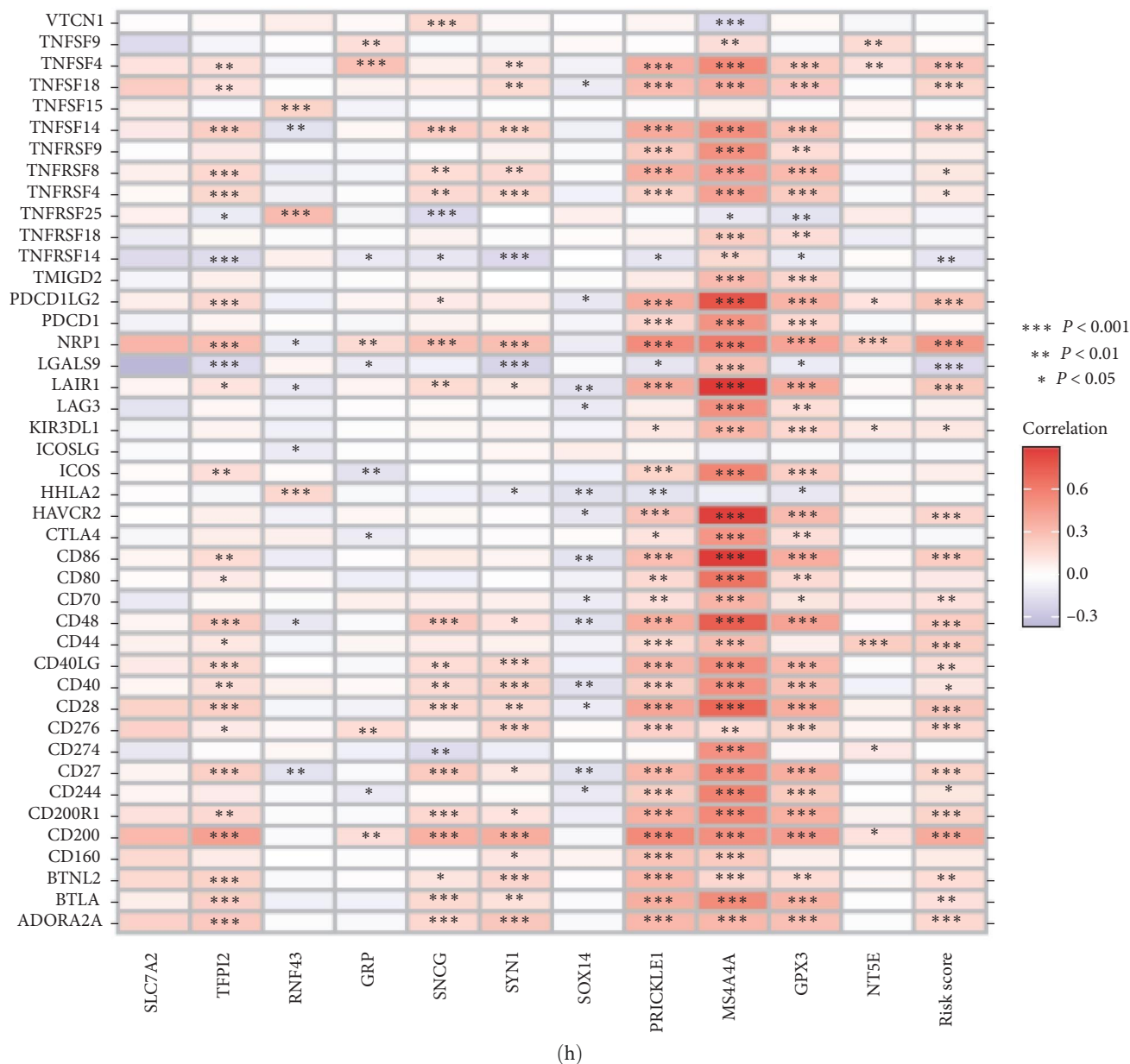(e)

(f)

(g)

FIGURE 9: Continued.

(h)

FIGURE 9: The relationship between the signature and tumor immune microenvironment. The KM survival regarding the prognostic value of (a) stromal score, (b) immune score, and (c) ESTIMATE score. The vioplot of the difference of the (d) stromal score, (e) immune score, (f) ESTIMATE score between high-and low-risk groups, (g) the relationship between risk score and TIICs in multiple databases, and (h) the correlation between the LNR-based signature and various ICBs. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.

of the public and causes a heavy burden to the social economy. Because GC is a heterogeneous disease, patients with the same clinical stage showed a different prognosis. Therefore, it is usually difficult to use the traditional TNM staging system to precisely forecast the clinical outcomes of GC [24]. Lymph node metastasis acts as the most common metastasis pathway for GC, and LNR has been proven to be an independent prognostic risk factor for GC. The present study aimed to develop a prediction model by analyzing the gene sequencing data from a large sample based on LNR.

We first identified 202 LNR-related genes from the TCGA project and eleven of which were filtered out to construct a prognostic signature for GC through univariate Cox

analysis and LASSO analysis. It is reassuring that a number of studies have been devoted to investigate the specific biological functions of the eleven LNR-related genes included in the signature. For instance, SOX14, which is a member of the SOXB2 subgroup of transcription factors implicated in neural development, has been found to be overexpressed in cervical cancer cells and could promote cell proliferation and invasion by activating the Wnt/β-catenin pathway [25]. RNF43 was reported to be significantly downregulated in gastric cancer and could suppress cell proliferation via inducing cell apoptosis [26]. PRICKLE1 is a cell polarity protein that is significantly overexpressed in GC cell lines from metastatic lesions compared with those from the primary tumor,
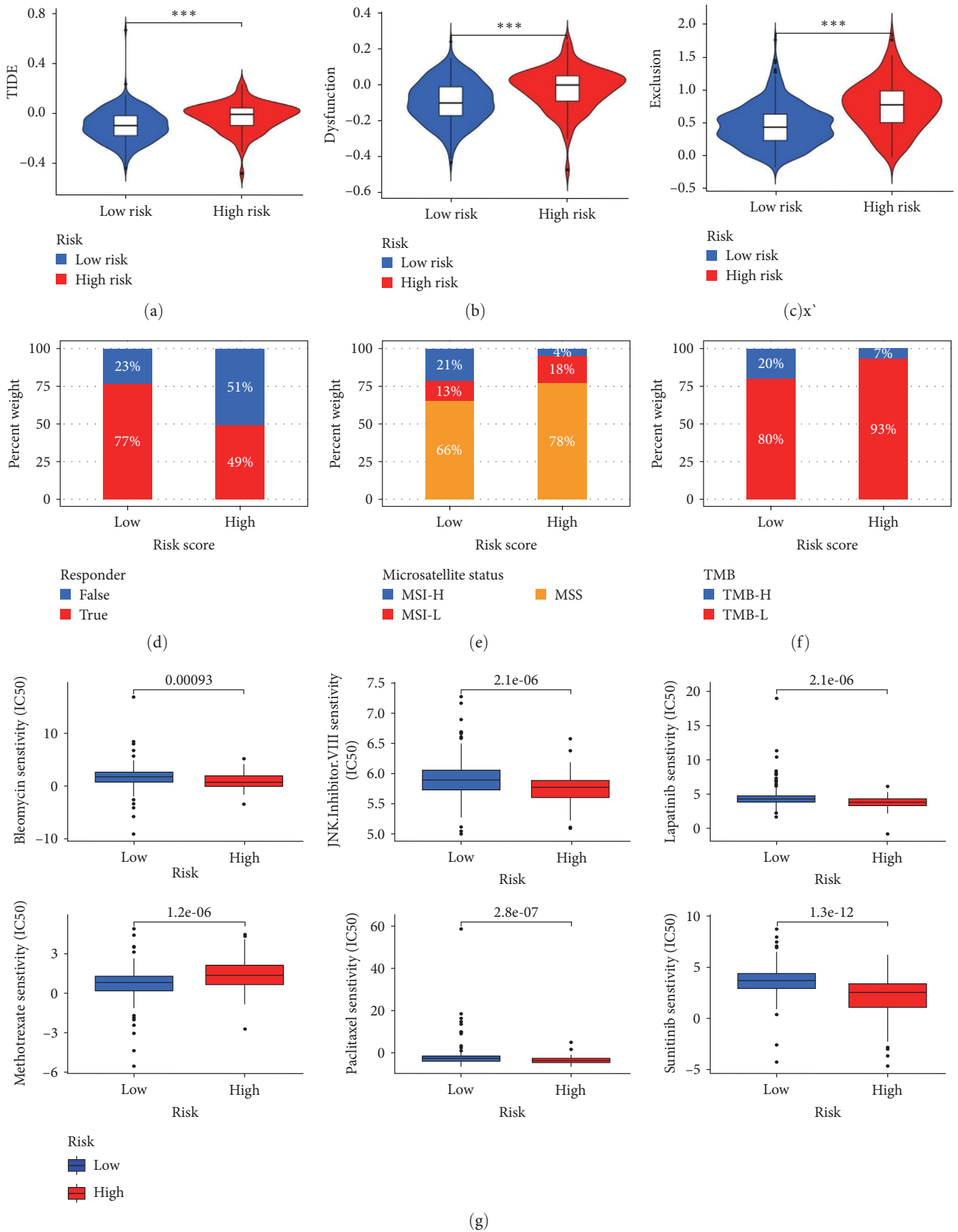
FIGURE 10: Association of LNR-base signature with immunotherapy and chemotherapy responsiveness in GC. The vioplot of the difference of the (a) TIDE score, (b) dysfunction score, (c) exclusion score between high- and low-risk groups. The barplot of the percentages of (d) immunotherapy responder predicted by the TIDE algorithm, (e) microsatellite status, (f) TMB between high-and low-risk groups in GC patients, and (g) correlations between risk scores and chemotherapeutic drugs. ***$P < 0.001$.

with the silencing of PRICKLE1 significantly inhibiting tumor metastasis through blocking mTOR signaling pathway *in vitro* [27]. SNCG, also known as BCSG1, is a member of the synuclein neuronal protein family. Yanagawa et al. [28] reported that abnormal expression and demethylation of SNCG in GC was significantly correlated with lymph node metastasis and advanced tumor stage. GPX3 hypermethylation in GC was reported to be significantly correlated with lymph node metastases and tumor relapse [29]. MS4A4A is a member of the membrane-spanning, four domain family, which has been identified as a novel cell surface marker for M2-like macrophages and plasma cells [30]. TFPI2 is a Kunitz-type serine proteinase inhibitor that has been identified as a tumor-suppressor gene, and aberrantly promoter hypermethylation of TFPI2 in GC contributed to tumor progression [31]. Overexpression of GRP and its receptors have been demonstrated in various cancer types including GC, and GRP has also been shown to act as a potent mitogen for cancer cells of diverse origins both *in vitro* and in animal models of carcinogenesis [32]. NT5E has been observed to be highly expressed in GC tissues, and overexpression of NT5E was obviously correlated with advanced clinical stage, lymph node metastasis and distant metastasis in GC patients, with the silencing of NT5E significantly suppressing cell proliferation, migration, and invasion of GC cells *in vitro* [33]. SLC7A2 is essential for transportation of arginine, lysine, and ornithine, and genetic polymorphisms in the SLC7A2 gene are significantly correlated with colorectal cancer development and progression [34]. However, the role of SYN1 in GC is rarely reported, and its role deserves further exploration. Our study found that all of these eleven genes were significantly correlated with the lymph node metastasis status of GC. Considering that traditional methods of predicting prognosis with single biomarkers or clinical data lack high sensitivity and specificity, our study comprehensively considered the expression patterns and prognostic values of these eleven LNR-related genes to generate a prognostic signature in GC. Excitingly, external validation cohort further confirmed the universal applicability of the LNR-based prognostic signature in clinical practice.

We also explored the correlation of the LNR-based risk score signature with the activity of immunity in TME of GC. As a result, high-risk score was significantly correlated with high stromal and immune scores, which means GC patients in the high-risk group possess higher levels of infiltrating stromal and immune cells in TME compared with those in low-risk group. A previous study has indicated that the infiltrating stromal cells and immune cells, together with various cytokines they secreted in TME, have the capability to modulate tumorigenesis and progression [35]. Thus, we further assessed the correlation between the abundances of TIICs and the risk score signature and found that most immune cells showed a positive correlation with the risk score. Furthermore, most of the ICBs also showed a significant positive correlation with the risk score, which indicated that the activity of immunity was more active in high-risk GC patients.

In the last few years, immunotherapy has achieved great progress as anticancer therapy for several solid tumors by reactivating the host immune response to tumors [36]. The success of immunotherapy has changed the treatment landscape of several cancers including GC. For example, pembrolizumab, a humanized IgG4 monoclonal antibody against programmed cell death-1 (PD-1), combined with trastuzumab as the first-line chemotherapy for patients with HER-2 positive advanced GC has received approval by the Food and Drug Administration (FDA) [37]. Another anti-PD-1 antibody, nivolumab, also showed satisfactory efficacy in the treatment of advanced GC [38]. However, due to the great heterogeneity in individuals, only a limited number of GC patients achieved the clinical benefit of immunotherapy, and the development of novel marker is urgently needed. Currently, the most commonly used clinical biomarkers for GC immunotherapy include MSI [39], TMB [40], and ICB [37, 38]. Besides, emerging evidence proposed that the TIDE score is a promising biomarker that could predict immunotherapy responsiveness based on pretreatment tumor profiles [21]. In the present study, we explored the correlation of the LNR-based risk signature with tumor immunotherapy responsiveness in GC based on MSI, TMB, ICB, and TIDE. As a result, we found that the proportion of immunotherapy responders was significantly higher in the low-risk group compared to the high-risk group. In sum, these results indicate that the LNR-based risk signature has the potential to help oncologists select GC patients who are more likely to benefit from immunotherapy, but whether patients in the low-risk group are more suitable for corresponding immunotherapy than the high-risk group remains for further experimental verification, the evaluation of which will be time-consuming.

5. Conclusion

To summarize, the present study successfully demonstrated that a novel signature constructed by genes that significantly correlated with the status of lymph node metastasis could predict prognosis for patients with GC and might help in distinguishing those who could benefit from antitumor immunotherapy.

Data Availability

The datasets generated and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/), The Cancer Genome Atlas (TCGA, https://www.cancer.gov/tcga), and UCSC Xena (https://xenabrowser.net/datapages/) projects.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

JL, TH, and QY conceived and designed the study. JL, TH, YW, XW, XC, and WC performed the bioinformatics

analysis and interpretation of the data. JL and TH drafted the manuscript. QY agreed to be responsible for all aspects of the work to ensure that issues of accuracy or completeness of the study were properly investigated and addressed. All authors read and approved the final manuscript. JL and TH authors have contributed equally to this work.

Supplementary Materials

Supplementary 1. Differentially expressed genes between low- and high-LNR groups.

Supplementary 2. GO enrichment analysis of LNR-related genes.

Supplementary 3. KEGG enrichment analysis of LNR-related genes.

References

[1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 Countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[2] A. P. Thrift and H. B. El-Serag, "Burden of gastric cancer," *Clinical Gastroenterology and Hepatology*, vol. 18, no. 3, pp. 534–542, 2020.

[3] K. Lyons, L. C. Le, Y. T.-H. Pham et al., "Gastric cancer: epidemiology, biology, and prevention: a mini review," *European Journal of Cancer Prevention*, vol. 28, no. 5, pp. 397–412, 2019.

[4] Z.-D. Chen, P.-F. Zhang, H.-Q. Xi, B. Wei, L. Chen, and Y. Tang, "Recent advances in the diagnosis, staging, treatment, and prognosis of advanced gastric cancer: a literature review," *Frontiers in Medicine*, vol. 8, Article ID 744839, 2021.

[5] H. Katai, T. Ishikawa, K. Akazawa et al., "Five-year survival analysis of surgically resected gastric cancer cases in Japan: a retrospective analysis of more than 100,000 patients from the nationwide registry of the Japanese Gastric Cancer Association (2001–2007)," *Gastric Cancer*, vol. 21, pp. 144–154, 2018.

[6] S. B. Edge and C. C. Compton, "The American Joint Committee on cancer: the 7th edition of the *AJCC Cancer Staging Manual* and the future of TNM," *Annals of Surgical Oncology*, vol. 17, pp. 1471–1474, 2010.

[7] Y.-C. Lee, P.-J. Yang, Y. Zhong, T. E. Clancy, M. T. Lin, and J. Wang, "Lymph node ratio-based staging system outperforms the Seventh AJCC system for gastric cancer: validation analysis with National Taiwan University Hospital cancer registry," *American Journal of Clinical Oncology*, vol. 40, no. 1, pp. 35–41, 2017.

[8] Y. Hou, X. Wang, and J. Chen, "Prognostic significance of metastatic lymph node ratio: the lymph node ratio could be a prognostic indicator for patients with gastric cancer," *World Journal of Surgical Oncology*, vol. 16, Article ID 198, 2018.

[9] J. Zhu, Z. Xue, S. Zhang et al., "Integrated analysis of the prognostic role of the lymph node ratio in node-positive gastric cancer: a meta-analysis," *International Journal of Surgery*, vol. 57, pp. 76–83, 2018.

[10] A. Bilici, F. Selcukbiricik, M. Seker et al., "Prognostic significance of metastatic lymph node ratio in patients with pN3 gastric cancer who underwent curative gastrectomy," *Oncology Research and Treatment*, vol. 42, no. 4, pp. 204–211, 2019.

[11] J. Liu, T. Lichtenberg, K. A. Hoadley et al., "An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics," *Cell*, vol. 173, no. 2, pp. 400–416. E11, 2018.

[12] J. Zhu, J. Z. Sanborn, S. Benz et al., "The UCSC cancer genomics browser," *Nature Methods*, vol. 6, pp. 239–240, 2009.

[13] J. N. Weinstein, E. A. Collisson, G. B. Mills et al., "The Cancer Genome Atlas Pan-Cancer analysis project," *Nature Genetics*, vol. 45, pp. 1113–1120, 2013.

[14] S.-J. Yoon, J. Park, Y. Shin et al., "Deconvolution of diffuse gastric cancer and the suppression of CD34 on the BALB/c nude mice model," *BMC Cancer*, vol. 20, Article ID 314, 2020.

[15] B. T. Sherman, M. Hao, J. Qiu et al., "DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update)," *Nucleic Acids Research*, vol. 50, no. W1, pp. W216–W221, 2022.

[16] Q. Wang, X. Li, Y. Wang et al., "Development and validation of a three-gene prognostic signature based on tumor microenvironment for gastric cancer," *Frontiers in Genetics*, vol. 12, Article ID 801240, 2022.

[17] H. Gittleman, A. E. Sloan, and J. S. Barnholtz-Sloan, "An independently validated survival nomogram for lower-grade glioma," *Neuro-Oncology*, vol. 22, no. 5, pp. 665–674, 2020.

[18] A. J. Vickers and E. B. Elkin, "Decision curve analysis: a novel method for evaluating prediction models," *Medical Decision Making*, vol. 26, no. 6, pp. 565–574, 2006.

[19] A. Mayakonda, D.-C. Lin, Y. Assenov, C. Plass, and H. P. Koeffler, "Maftools: efficient and comprehensive analysis of somatic variants in cancer," *Genome Research*, vol. 28, pp. 1747–1756, 2018.

[20] S. Yu, Y. Wang, K. Peng, M. Lyu, F. Liu, and T. Liu, "Establishment of a prognostic signature of stromal/immune-related genes for gastric adenocarcinoma based on ESTIMATE algorithm," *Frontiers in Cell and Developmental Biology*, vol. 9, Article ID 752023, 2021.

[21] P. Jiang, S. Gu, D. Pan et al., "Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response," *Nature Medicine*, vol. 24, pp. 1550–1558, 2018.

[22] W. Yang, J. Soares, P. Greninger et al., "Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Research*, vol. 41, no. D1, pp. D955–D961, 2013.

[23] P. Geeleher, N. Cox, and R. S. Huang, "pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels," *PLoS ONE*, vol. 9, no. 9, Article ID e107468, 2014.

[24] T. Sano, D. G. Coit, H. H. Kim et al., "Proposal of a new stage grouping of gastric cancer for TNM classification: international Gastric Cancer Association staging project," *Gastric Cancer*, vol. 20, pp. 217–225, 2017.

[25] F. Li, T. Wang, and S. Tang, "SOX14 promotes proliferation and invasion of cervical cancer cells through Wnt/β-catenin pathway," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 2, pp. 1698–1704, 2015.

[26] L. Niu, H.-Z. Qin, H.-Q. Xi, B. Wei, S.-Y. Xia, and L. Chen, "RNF43 inhibits cancer cell proliferation and could be a potential prognostic factor for human gastric carcinoma," *Cellular Physiology and Biochemistry*, vol. 36, no. 5, pp. 1835–1846, 2015.

[27] R. Zhou, Z. Tang, H. Li, X. Wang, and Y. Sun, "PRICKLE1 promotes gastric cancer metastasis by activating mTOR signaling," *American Journal of Translational Research*, vol. 13, no. 5, pp. 4266–4280, 2021.

[28] N. Yanagawa, G. Tamura, T. Honda, M. Endoh, S. Nishizuka, and T. Motoyama, "Demethylation of the *synuclein* γ gene

CpG Island in primary gastric cancers and gastric cancer cell lines," *Clinical Cancer Research*, vol. 10, no. 7, pp. 2447–2451, 2004.

[29] C. Zhou, R. Pan, B. Li et al., "*GPX3* hypermethylation in gastric cancer and its prognostic value in patients aged over 60," *Future Oncology*, vol. 15, no. 11, pp. 1279–1289, 2019.

[30] R. Sanyal, M. J. Polyak, J. Zuccolo et al., "MS4A4A: a novel cell surface marker for M2 macrophages and plasma cells," *Immunology & Cell Biology*, vol. 95, no. 7, pp. 611–619, 2017.

[31] H. Takada, N. Wakabayashi, O. Dohi et al., "Tissue factor pathway inhibitor 2 (*TFPI2*) is frequently silenced by aberrant promoter hypermethylation in gastric cancer," *Cancer Genetics*, vol. 197, no. 1, pp. 16–24, 2010.

[32] O. Patel, A. Shulkes, and G. S. Baldwin, "Gastrin-releasing peptide and cancer," *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1766, no. 1, pp. 23–41, 2006.

[33] S. Hu, F. Meng, X. Yin, C. Cao, and G. Zhang, "NT_5E is associated with unfavorable prognosis and regulates cell proliferation and motility in gastric cancer," *Bioscience Reports*, vol. 39, no. 5, Article ID BSR20190101, 2019.

[34] P. Sun, X. Zhu, M. J. Shrubsole et al., "Genetic variation in *SLC7A2* interacts with calcium and magnesium intakes in modulating the risk of colorectal polyps," *The Journal of Nutritional Biochemistry*, vol. 47, pp. 35–40, 2017.

[35] Q. Luo, F. He, and J. Cao, "A stromal and immune cell infiltration-based score model predicts prognosis and chemotherapy effect in colorectal cancer," *International Immunopharmacology*, vol. 99, Article ID 107940, 2021.

[36] E. Picard, C. P. Verschoor, G. W. Ma, and G. Pawelec, "Relationships between immune landscapes, genetic subtypes and responses to immunotherapy in colorectal cancer," *Frontiers in Immunology*, vol. 11, Article ID 369, 2020.

[37] Y. Y. Janjigian, A. Kawazoe, P. Yañez et al., "The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer," *Nature*, vol. 600, pp. 727–730, 2021.

[38] Y. Y. Janjigian, K. Shitara, M. Moehler et al., "First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial," *The Lancet*, vol. 398, no. 10294, pp. 27–40, 2021.

[39] D. T. Le, J. N. Durham, K. N. Smith et al., "Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade," *Science*, vol. 357, no. 6349, pp. 409–413, 2017.

[40] F. Wang, X. L. Wei, F. H. Wang et al., "Safety, efficacy and tumor mutational burden as a biomarker of overall survival benefit in chemo-refractory gastric cancer treated with toripalimab, a PD-1 antibody in phase Ib/II clinical trial NCT02915432," *Annals of Oncology*, vol. 30, no. 9, pp. 1479–1486, 2019.