Research Article

A Novel 7-Methylguanosine (m7G)-Related Gene Signature for Overall Survival Prediction in Patient with Clear Cell Renal Cell Carcinoma

Yongxin Fu, Jiawu Wang, Zhiya Hu, Yang Gou, Yisen Li, and Qing Jiang ⓘ

Department of Urology, The Second Affiliated Hospital of Chongqing Medical University, Chongqing, China

Correspondence should be addressed to Qing Jiang; 300899@hospital.cqmu.edu.cn

Clear cell renal cell carcinoma (ccRCC) is the most common pathology type of renal cancer that has an abysmal prognosis. Although a crucial role for 7-methylguanosine modification in cancer cell development has been reported, its role in ccRCC remains uncertain. This study was conducted to determine the efficacy of predictive biomarkers based on m7G-related genes in ccRCC. Firstly, we extracted clinical data and gene expression profiles of ccRCC patients from publicly accessible databases. It identified that 22 of the m7G-related 34 genes were related to overall survival, and 5 of the 22 genes were significantly expressed differently in tumor tissues. Based on Lasso regression analysis, five optimal genes (CYFIP2, EIF4A1, NUDT1, NUDT10, and NUDT4) were chosen to build a new predictive risk model in the TCGA cohort. Validation was carried out with the E-MTAB-1980 cohort. Then, a prognostic nomogram was erected, including the m7G-related gene risk score, age, histological grade, and stage status. Further studies and analysis showed that immune cell infiltration might be associated with the m7G-related risk genes. In addition, the relationship between gene expression and drug response was evaluated by the Pearson correlation test. Therefore, the risk signature with five selected m7G-related genes may be a promising prognostic biomarker and contribute to standardized prognostic assessment for ccRCC.

1. Introduction

Ninety percent of all renal malignancies are renal cell carcinomas (RCC), a prevalent malignancy in the urinary system [1]. ccRCC, the most common subtype of RCC, accounts for approximately 70% of RCC and is also one of the most aggressive subtypes [2]. The early symptoms of renal clear cell carcinoma are not obvious, with only 6–10% of patients presenting with typical symptoms such as backache, an abdominal mass, or hematuria [3]. More than about 30% of ccRCC patients already have metastasis or local progression at the first diagnosis [4], so it is necessary to diagnose ccRCC at an early stage. Although surgery is the best therapy method for ccRCC, nearly 30% of patients develop local recurrence and metastasis after surgically removing local ccRCC [5]. Despite the advent of targeted drugs and immune checkpoint inhibitors, patients with metastatic renal cancer still have a low overall survival rate. Hence, it is crucial to discover effective and novel biomarkers to provide novel molecular targets for adjuvant therapies and improve patient outcomes.

In the last decade, the role of RNAs in cellular processes has attracted more and more attention. Over 150 types of RNA modifications have been identified to date [6, 7], including N7-methylguanosine, N1-methyladenosine,2′-O-methylation, N6-methyladenosine, and 5-methylcytosine [8–10]. In general, M7G, an RNA modification with a positive charge [11], is present in the eukaryotic mRNA, tRNA, rRNA, and microRNAs [12], which is required for nearly all phages that participate in the expression of mRNAs, such as transcription elongation [13], pre-mRNA splicing [14], polyadenylation [15], nuclear export [16], and translation [17]. Although some literature reported that m7G modification is linked to the development of various cancers,

including lung cancer [18] and colon cancer [19], the relationship between m7G RNA methylation regulators and ccRcc remains widely unclear.

We first downloaded mRNA expression and clinical data from public databases. Then, selected five optimal genes and constructed a predictive model. The model's predictive value has been confirmed via various survival analyses. Moreover, we also validated the accuracy of the model in the E-MTAB-1980 cohorts. Furthermore, we set up a nomogram to enhance individualized prognosis assessment. Finally, we elucidate the underlying mechanism of ccRCC through functional analysis and explore the relationship between risk genes and chemosensitivity.

2. Materials and Methods

2.1. Data Acquisition. We acquired the RNA sequence data of 539 tumors and 72 normal tissues of ccRCC and clinical information ($n = 537$) in the TCGA database. In addition, 101 patients in the E-MTAB-1980 cohort with clinico-pathological information and RNA sequence data were also downloaded from the ArrayExpress database. Reference [20]. Our analysis excluded patients who had no follow-up days during the survival period. Thirty-four m7G-related genes were obtained from earlier studies [21], and gene sets: gomf-m7G-5-pppn-diphosphatase-activity, gomf-RNA-7-methylguanosine-cap-binding, and gomf-RNA-cap-binding (https://www.gsea-msigdb.org/gsea/index.jsp). Basic clinical information is summarized in Table 1. A flow chart of our overall study is displayed in Figure 1.

2.2. Prognostic Model Construction and Validation. Differentially expressed genes (DEGs) in ccRCC normal samples and tumor samples in the TCGA cohort were processed using the "Limma" R package, and the filter conditions were set (fdrFilter <0.05, logFCfilter >1). Then, a univariate Cox analysis of survival outcomes was performed to screen the genes with prognostic values. The "Venn" R package was used to select m7G-related DEGs with predictive values through DEGs and prognostically valuable genes. Furthermore, we used the "pheatmap" package to display heatmaps and the "Survival" package to show forest maps to represent differences between groups. In the TCGA cohort, the "Survival" and the "glmnet" packages were applied for prognostic risk characteristics. The risk score of each patient was estimated as follows: Risk score $= \sum$ (gene expression $\times$ corresponding coefficient. Then, we calculated the risk score of each patient according to the above formula and divided all patients into high-risk and low-risk groups with the median score. "Survminer" and "Survival" of the R language were introduced to evaluate overall survival (OS) based on the Kaplan–Meier (K–M) method. The "timeRoc" of the R package was applied to assess the accuracy of the related gene model. The "pheatmap" of the R package was used to describe the risk scores and corresponding survival times. The R packages "Rtsne" and "GGplot2" were used to investigate whether there was a difference in distribution between the two groups of at-risk

TABLE 1: Basic clinical characteristics of patients in the TCGA cohort and the E-MTAB-1980 Cohort.

	TCGA cohort (537)	E-MTAB 1980 cohort (101)
Age (years)		
≤65	352	57
>65	185	44
Gender		
Female	191	24
Male	346	77
Grade		
G1	14	13
G2	230	59
G3	207	22
G4	78	5
Gx	5	
Unknown	3	2
Stages		NA
I	269	
II	57	
III	125	
IV	83	
Unknown	3	
T Stage		
T1	275	68
T2	69	11
T3	182	21
T4	11	1
N Stage		
No	240	94
Not no	17	7
Nx	280	
M Stage		
M0	426	89
M1	79	12
Mx	30	
Unknown	2	

patients. We further processed the same analyses to validate the model's predictive performance in the E-MTAB-1980 cohorts.

2.3. The Construction of a Predictive Nomogram Using Risk Scores and Clinical Factors. We performed the analysis of univariate and multivariate Cox regression further to examine the relationship between OS and clinical factors. Then, establishing a nomogram including a risk score and the independent prognostic clinical characteristics, we can assess the probability of survival at 1, 3, and 5 years for ccRCC patients. In order to measure this nomogram's predictive precision, calibrate curve analysis was used.

2.4. Functional and Immune Infiltration Enrichment Analyses. Using the package ClusterProfiler, we used GO and KEGG to analyze risk-related DEGs. The ESTIMATE algorithm assessed the association between risk score and stromal score or immune score. "GSVA" and "GSEABASE" were utilized to assess immune functions and cells among risk-related DEGs.
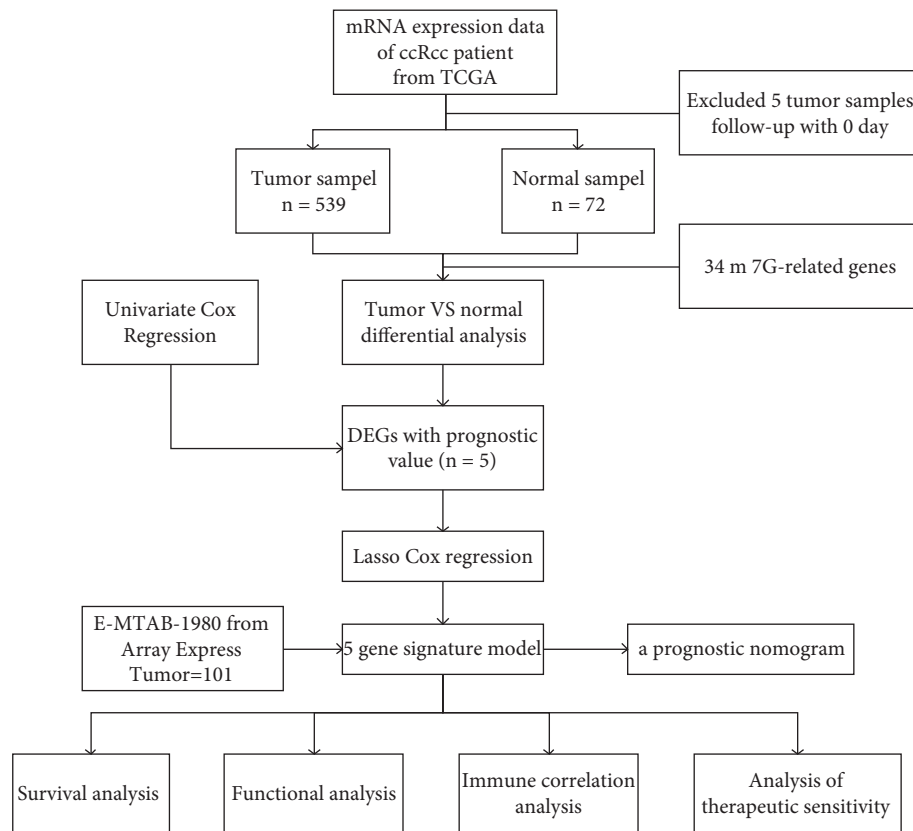
Figure 1: The overall workflow of this study.

2.5. Drug Response Analysis. We obtained the NCI-60 data from the CellMiner platform and used the Pearson correlation test to assess the relation of m7G-related gene expressions with drug response.

2.6. Statistics Analysis. Statistical analyses were conducted using the R software (version 4.1.0) and the PERL programming language (version 5.32.1). Statistical significance is set at $P < 0.05$, and P values follow a double-tailed distribution. To calculate the adjusted P value, Benjamini-Hochberg was invoked.

3. Results

3.1. Differential Expression of MRGs. Most of the m7G-related expressed genes (5/34, 14.7%) were differentially expressed between nontumorous tissues and tumorous tissues in the TCGA cohort (adjusted $P < 0.05$, logFCfilter >1). We screened out 22 genes associated with OS ($P < 0.01$) via univariate Cox regression analysis. Thus, five m7G-related genes were identified, as indicated by the Venn diagram (Figure 2(a)). Afterward, the heatmap demonstrated the expression levels in normal and tumor samples (Figure 2(b)). Furthermore, we calculated the P values, hazard ratios, and 95% confidence intervals (CI) of each gene displayed in forest plots (Figure 2(c)). Finally, we visualized five prognostic m7G-related DEG interactions with correlation networks (Figure 2(d)).

3.2. Generation of the Predictive Risk Model and Survival Analyses in the TCGA Database. By lasso regression analysis, we discovered that five m7G-related genes fit the model (Figure 3(a) and 3(b)). Among them, EIF4A1, NUDT1, and NUDT10 are high-risk genes, while CYFIP2 and NUDT4 are assigned to low-risk genes. In order to establish the prognosis model based on the expression levels of the selected genes, we computed risk scores using the following formula: risk score = (−0.532106574444019*CYFIP2) + (0.257284742647963*EIF4A1) + (0.210867606644546* NUDT1) + (0.0273145419183592*NUDT10) + (−0.254461172952311*NUDT 4). Five m7G-related genes model was identified in the TCGA cohort. We categorized all patients based on the median risk scores into two groups: high-risk ($n = 262$) and low-risk ($n = 263$). (Figure 4(a)). The results of PCA and t-SNE analysis of MRGs indicated that different risk groups of patients were entirely distributed differently (Figures 4(b) and 4(c)). Furthermore, we noticed that the high-risk patients died more frequently than the low-risk patients when we took survival outcomes into consideration (Figure 4(d)). Therefore, the predictive capacity for the clinical prognosis of the model was evaluated using the K-M survival and time-dependent ROC analyses, respectively. The results are as follows: survival rates for patients assigned to the high-risk group were significantly worse than those in the low-risk group according to the K-M survival curve ($P = 4.097e - 14$) (Figure 4(e)). During 1 year, 0.748 AUC was recorded, followed by 0.694 at 2 years, and 0.704 at 3 years (Figure 4(f))
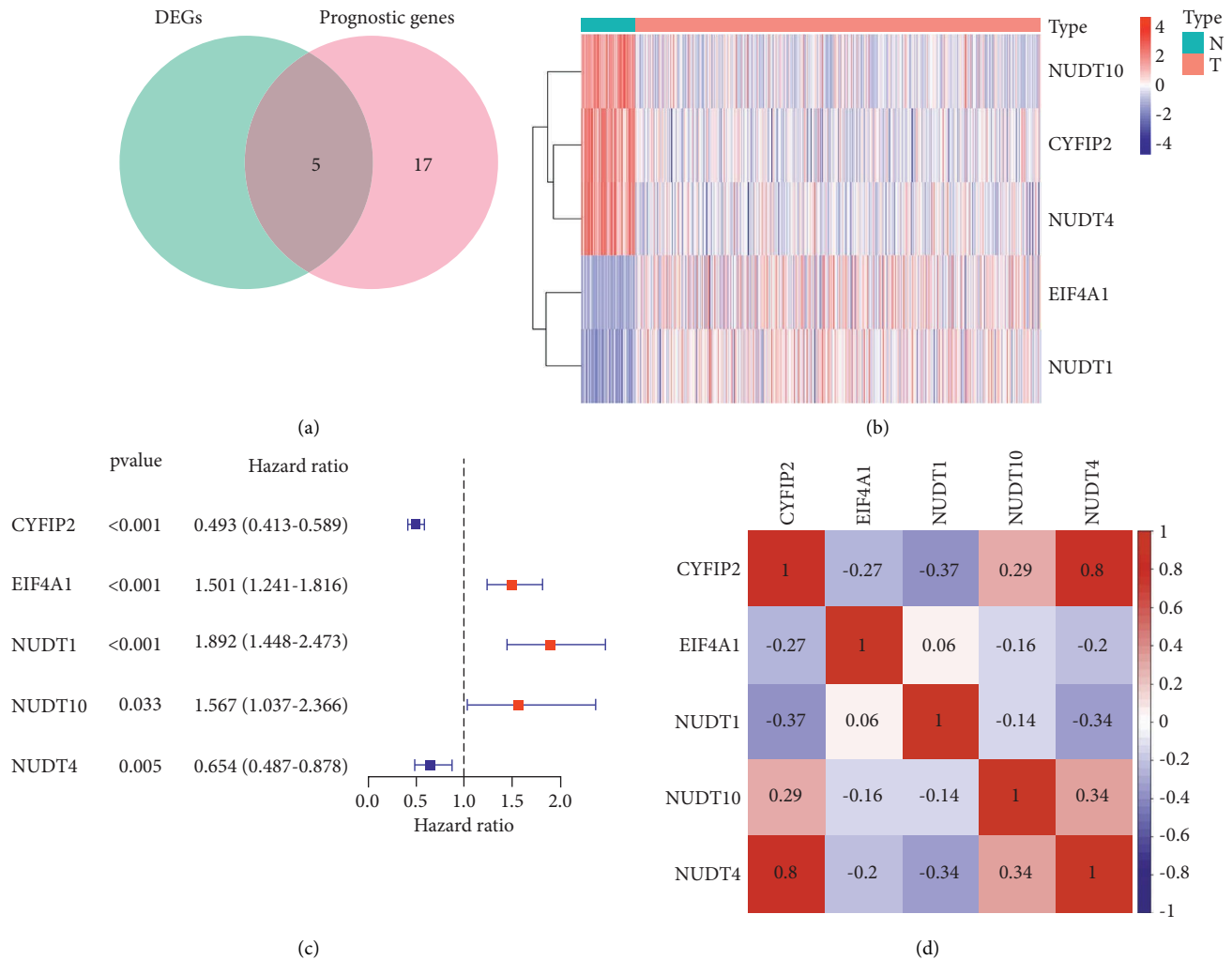
(a)

(b)

(c)

(d)

FIGURE 2: Intersection of differentially expressed m7G-related genes and survival-associated m7G-related genes. (a) Venn diagram showing the five m7G-related genes associated with survival between differentially expressed genes and prognostic genes. The green circle represents differentially expressed genes, and the pink circle represents prognostic genes. (b) Heatmap showing the five m7G-related associated with survival. The red rectangle represents tumor tissue, and the blue rectangle represents normal tissue. (c) The forest figure of the 5 key m7G-related genes in univariate Cox regression. The blue line represents a 95% confidence interval. The position of the square represents the hazard ratio. (d) The correlation network of 5 m7G-related genes, in which different colors represent the correlation coefficients. The red represents a positive relationship, and the blue represents a negative relationship.

and reached 0.748 at risk, 0.766 at stage, 0.645 at grade, 0.641 at age, and 0.506 at gender in the ROC analysis (Figure 4(g)).

3.3. Survival Analyses of the E-MTAB-1980 Cohorts for Verification. We also calculated the risk score for the E-MTAB-1980 cohorts for validation ($n = 101$). In light of the median risk score, patients were divided into two sets: those at high-risk ($n = 50$) and those at low-risk ($n = 51$). (Figure 5(a)). MRGs were reduced in dimension by PCA and T-SNE analysis, and then we found that different risk groups were distributed in two different directions (Figures 5(b) and 5(c)). Additionally, the death rates in the high-risk group were higher than those in the low-risk group. (Figure 5(d)). In addition, we performed analyses of K-M survival and time-dependent ROC curves, which suggested that both were significant in assessing the prognostic value of the risk

score model. Survival curves showed that patients in the high-risk group had significantly worse OS than patients in the low-risk group ($P = 1.53e - 03$) (Figure 5(e)). The area under the curve (AUC) was 0.617 at 1 year, at 2 years it reached 0.622, and at 3 years it reached 0.659 (Figure 5(f)) and reached 0.617 at risk, 0.757 at stage, 0.681 at grade, 0.609 at age, and 0.522 at gender in the ROC analysis (Figure 5(g)).

3.4. The Independent Prognostic Evaluation of Five Genes. We depicted to use a heatmap to illustrate the expression levels of the five selected m7G genes and clinical features in the two low-high groups, which suggested the significant differences in clinical features between the low and high groups (Figure 6(a)). Toward a better understanding of the predictive value of risk scores and other clinical features in ccRCC patients, the TCGA cohort was subjected to both
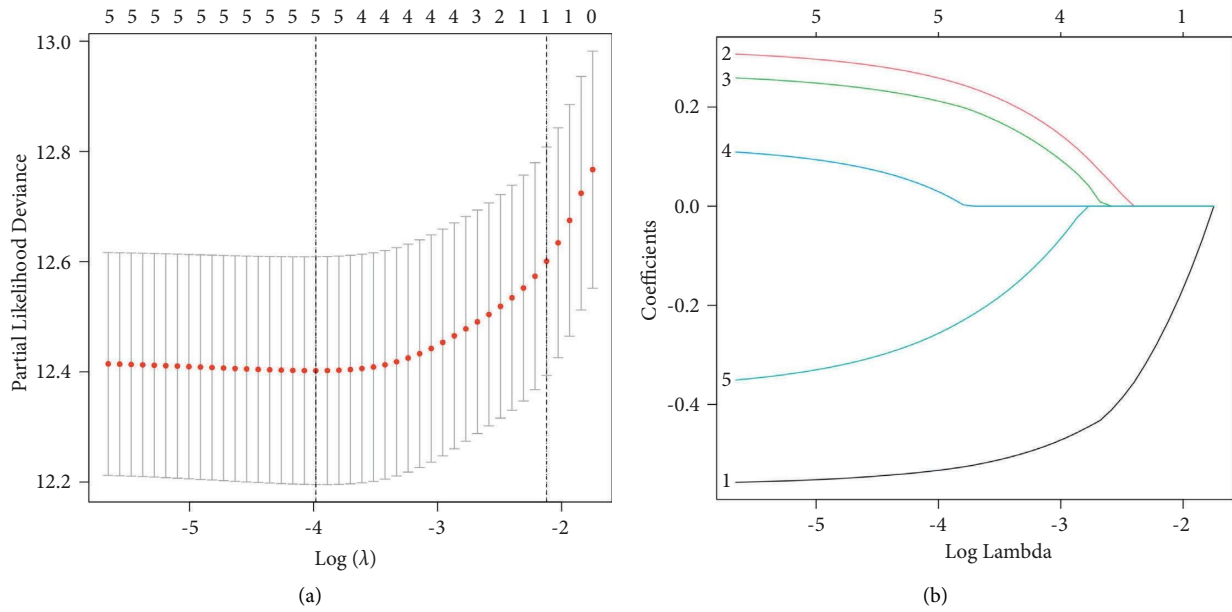
(a)

(b)

Figure 3: Identification of five optimal MRGs. (a) Partial likelihood deviance was plotted against log (λ). The vertical dotted lines indicate the λ value with minimum error. The largest λ value is where the deviation is within one standard error (SE) of the minimum. (b) Least absolute shrinkage and selection operator (Lasso) coefficient profiles of MRGs in ccRCC.
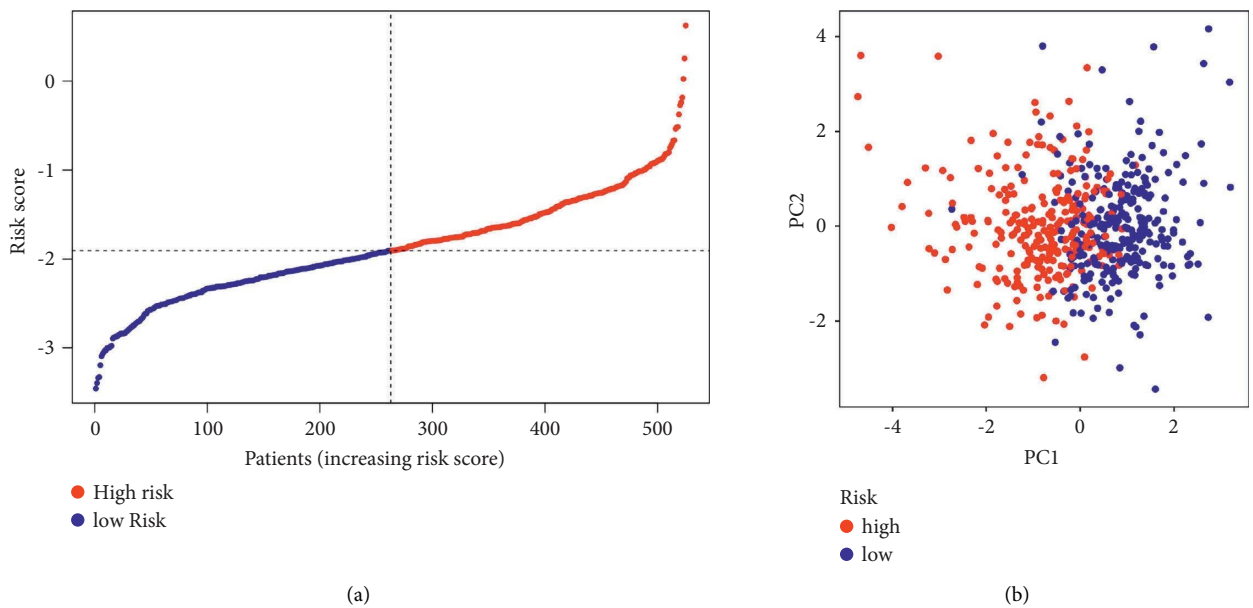

(a)

(b)
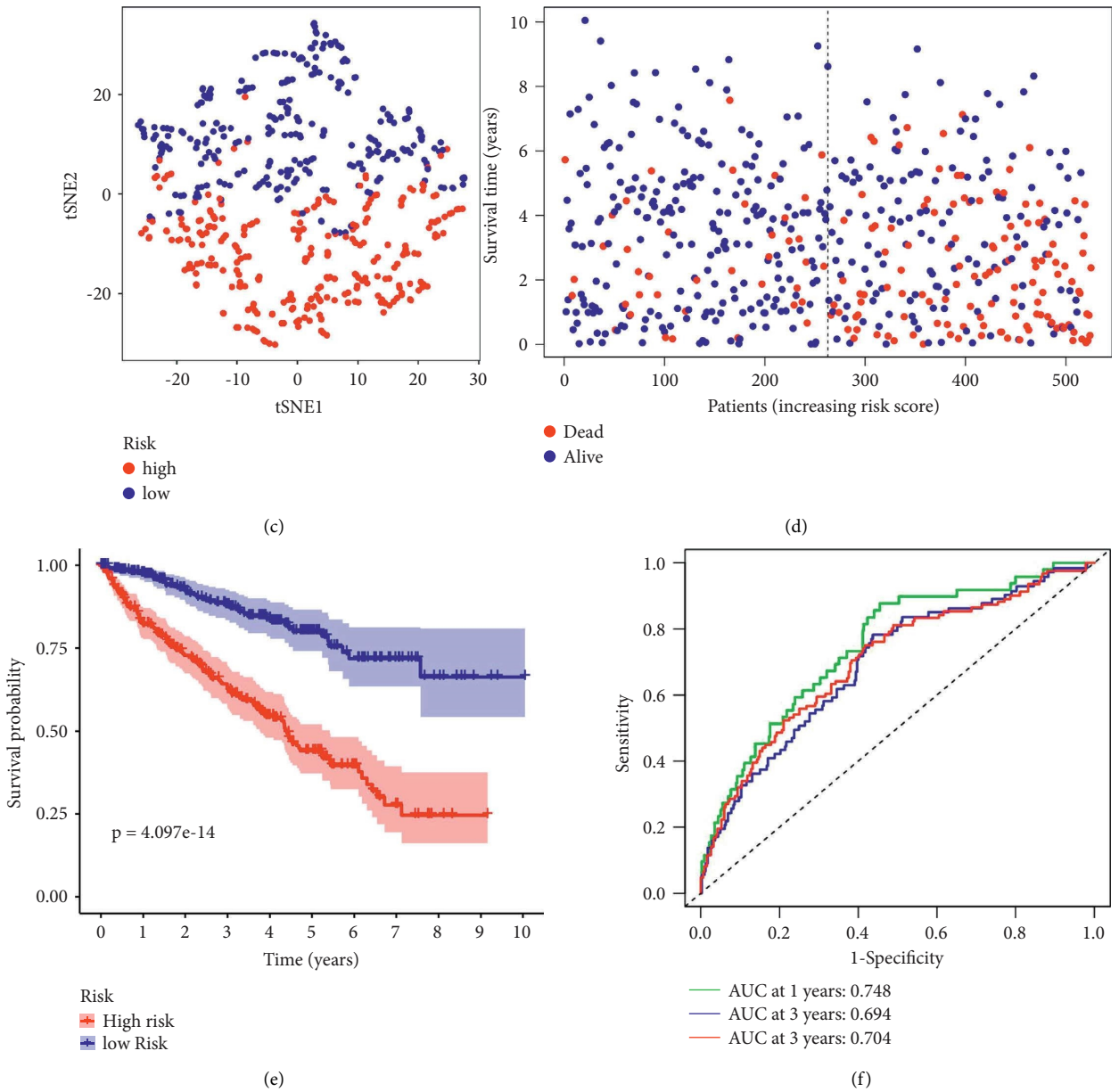
Figure 4: Continued.

(c)

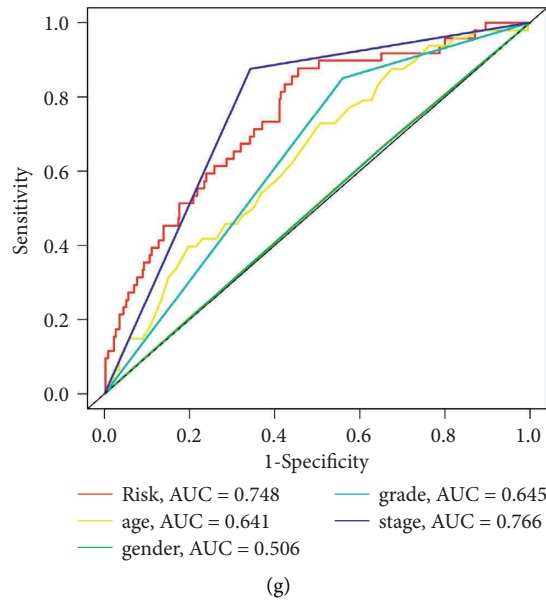

(d)

(e)

(f)

FIGURE 4: Continued.

(g)

Figure 4: Survival analyses of the 5-gene model in the TCGA cohort. (a) The distribution of the risk scores in the TCGA cohort. (b) Principal component analysis plots display the established gene signature expression distribution in different risk groups. (c) *t*-Distributed stochastic neighbor embedding plots reveal the patients' distribution in different risk groups. (d) The distributions of the risk scores and corresponding survival times of all patients in the TCGA cohort. (e) OS-based K-M survival curves for the patients in the high and low-risk groups in the TCGA cohort. (f) AUC of time-dependent ROC curves verified the prognostic performance of the risk score in the TCGA cohort. (g) AUC of clinical ROC curves verified the prognostic performance of the risk score in the TCGA cohort.
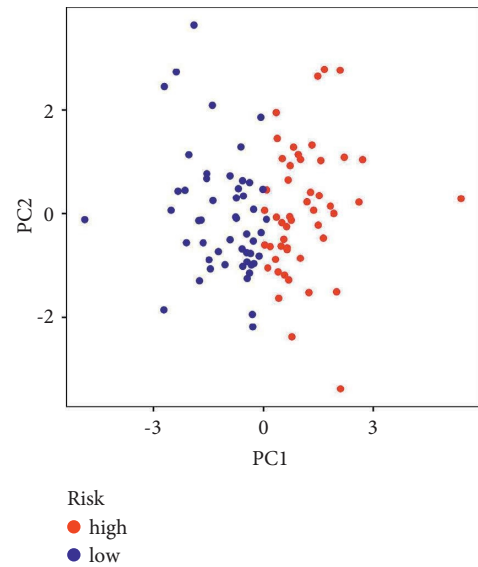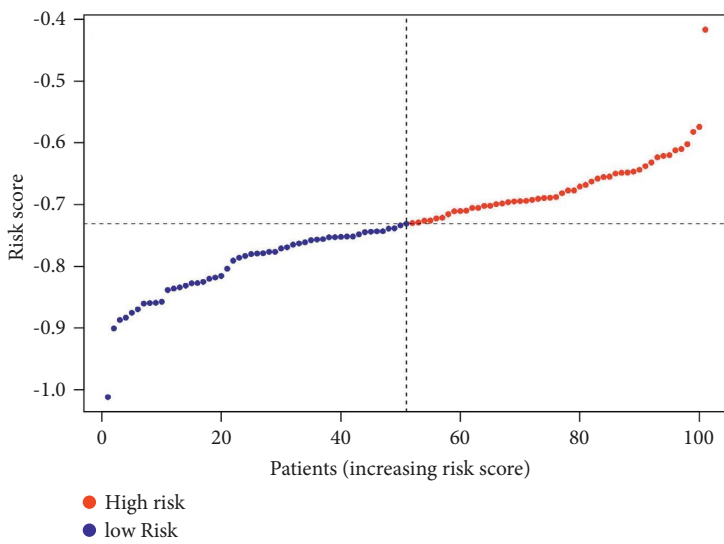

(a)

(b)

Figure 5: Continued.

(c)

(d)

p = 8.529e-03

Risk
High risk
low Risk

(e)

AUC at 1 years: 0.811
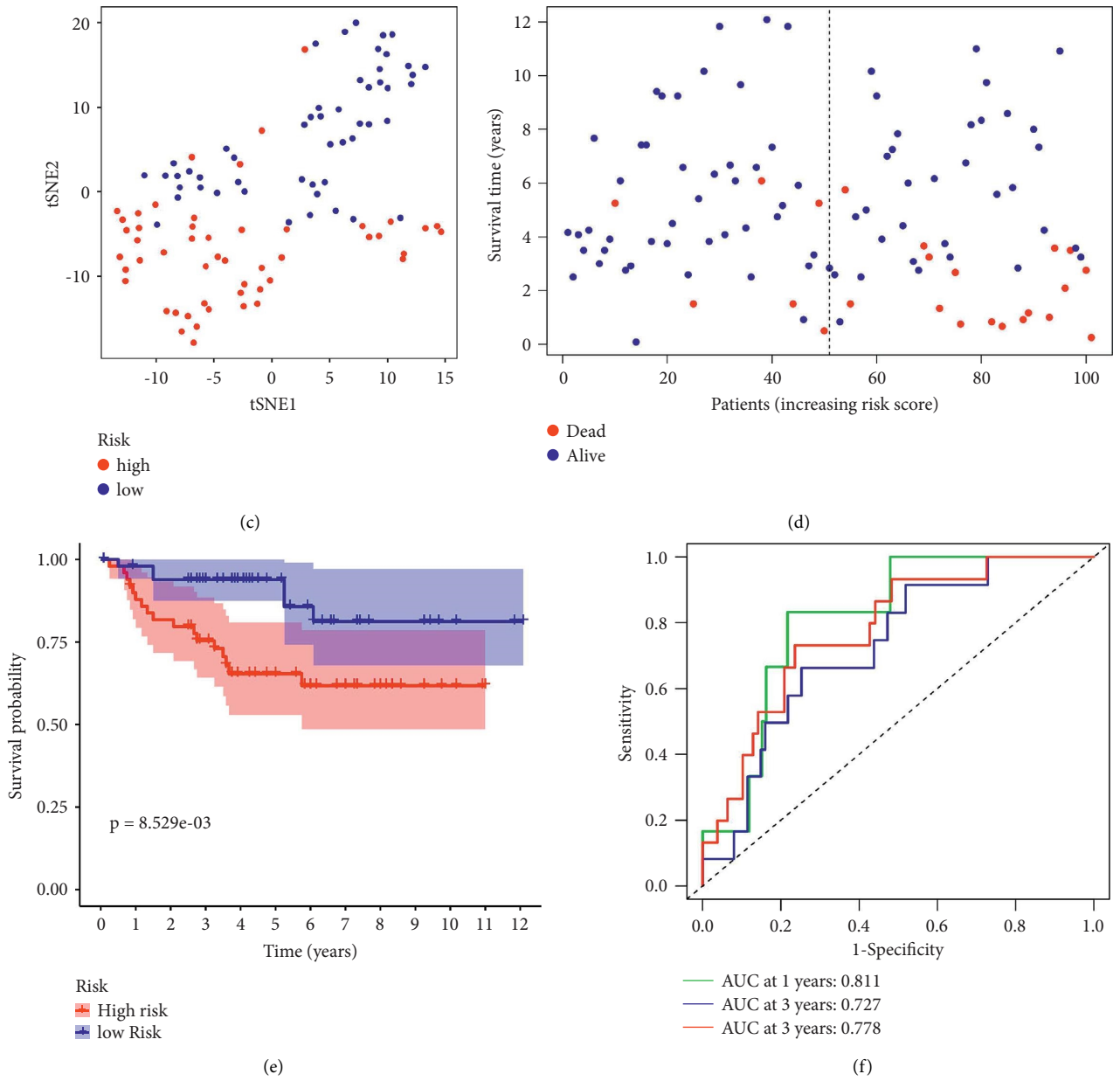AUC at 3 years: 0.727
AUC at 3 years: 0.778
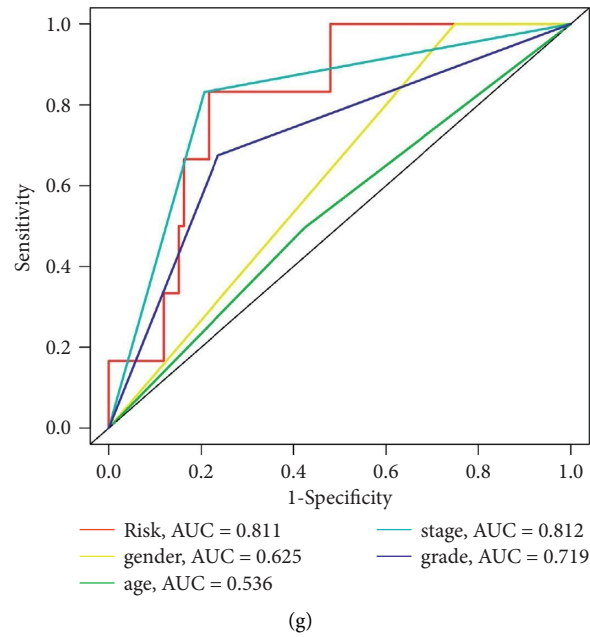
(f)

FIGURE 5: Continued.

(g)

Figure 5: Validation of the 5-gene model in the E-MTAB-1980 cohort. (a) The distribution of the risk scores in the E-MTAB-1980 cohort. (b) Principal component analysis plots display the established gene signature expression distribution in different risk groups. (c) *t*-Distributed stochastic neighbor embedding plots reveal the patients' distribution in different risk groups. (d) The distributions of the risk scores and corresponding survival times of all patients in the E-MTAB-1980 cohort. (e) OS-based K-M survival curves for the patients in the high- and low-risk groups in the E-MTAB-1980 Cohort. (f) AUC of time-dependent ROC curves verified the prognostic performance of the risk score in the E-MTAB-1980 cohort. (g) AUC of clinical ROC curves verified the prognostic performance of the risk score in the E-MTAB-1980 cohort.
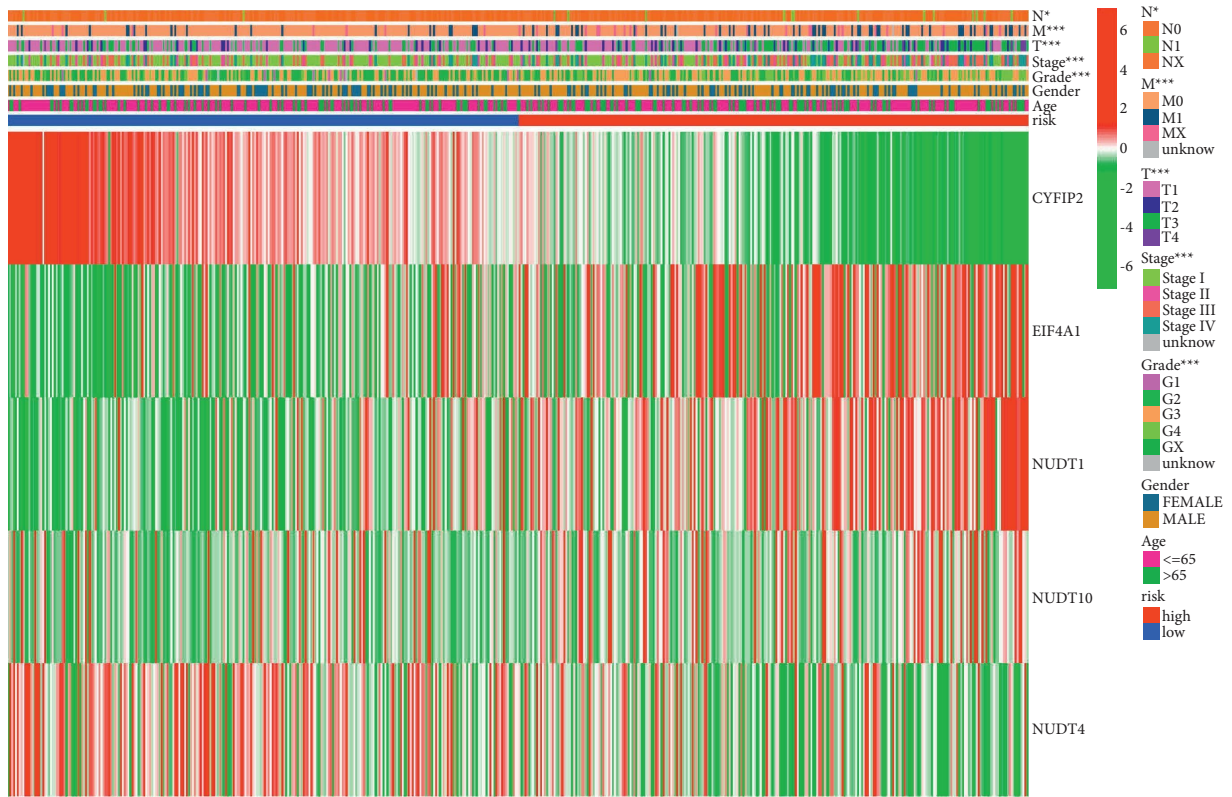
univariate and multivariate Cox regression analyses (Figures 6(b) and 6(c)). We omitted the N-stage in our further analyses because most patients lacked clinical information on the N-stage. A prediction model was developed using four independent prognostic factors: age, risk score, grade, and stage. A nomogram predicting the OS probability over the course of 1, 3, and 5 years was estimated using the five m7G-related genes (Figure 6(d)) and assessed its predictive power (Figures 6(e) and 6(g)). The calibration curves proved that the nomogram had a positive effect on prognosis.

3.5. Functional Analyses in TCGA Cohort. To address the potential biological differences, we conducted GO and KEGG enrichment analyses on the DEGs between the different risk groups. The GO enrichment analysis of risk-related DEGs revealed significant correlations between humoral immune response, immunoglobulin complex, and antigen binding (Figure 7(a), *P* adjust <0.05). The KEGG pathway analysis demonstrated that DEGs had significantly enriched the following pathways: cytokine-cytokine receptor interaction, viral protein interaction with cytokine, and cytokine receptor and mineral absorption (Figure 7(b), *P* adjust <0.05).
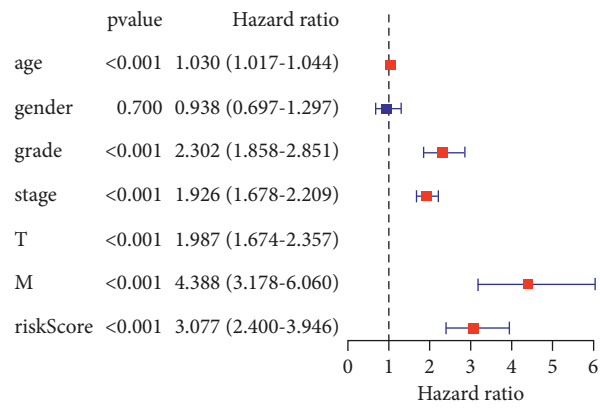
3.6. Immune Infiltration ssGSEA in TCGA Cohort. The TCGA cohort was evaluated by ESTIMATE, which manifested a positive correlation between risk scores, immune

scores ($P < 0.001$; Figure 8(a)), and stromal scores ($P = 0.03$; Figure 8(b)). On six of the 16 immune cells, the high-risk group demonstrated significantly higher infiltration of activated dendritic cells (aDCs), CD8+ T cells, helper T cells 2 (Th2 cells), helper T cells 1 (Th1 cells), follicular helper T cells (Tfh), and tumor-infiltrating lymphocyte (TIL). In contrast, iDCs and mast cells showed the opposite pattern ($P < 0.05$, Figure 8(c)). In the immunopathways analysis, seven pathways were positively associated with risk scores, such as CCR, check-point, cytolytic-activity, inflammation-promoting, parainflammation, T cell co-stimulation, and type I IFN response. In comparison, the type II IFN response had a negative effect ($P < 0.05$, Figure 8(d)). The signature significantly correlates with immune infiltration, based on our findings.
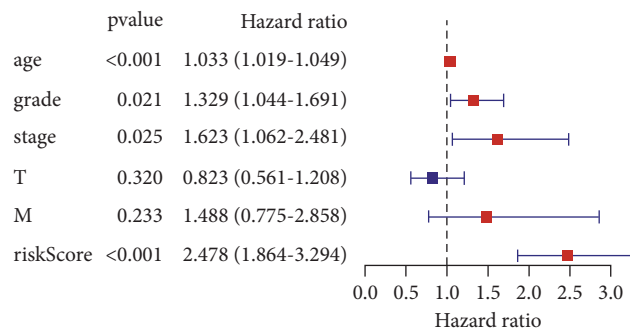
3.7. Expressions of MRGs Correlated with Therapeutic Response. We incorporated the data on cancer cell expression and the effectiveness of FDA-approved drugs in our study in order to investigate the clinical applications of the m7G-related gene signature. CYFIP2-overexpressed cancer cells were demonstrated to be more sensitive to the following drugs: nelarabine, melphalan, idarubicin, and so on. The elevated NUDT4 expression levels in cancer cells increased the sensitivity to selumetinib, but drug resistance to everolimus was correlated with them. Overexpression of NUDT1 in cancer cells makes them more susceptible to carboplatin and triethylenemelamine, and cancer cells expressing higher levels of EIF4A1 were more sensitive to cladribine (Figure 9).

(a)

	pvalue	Hazard ratio
age	<0.001	1.030 (1.017-1.044)
gender	0.700	0.938 (0.697-1.297)
grade	<0.001	2.302 (1.858-2.851)
stage	<0.001	1.926 (1.678-2.209)
T	<0.001	1.987 (1.674-2.357)
M	<0.001	4.388 (3.178-6.060)
riskScore	<0.001	3.077 (2.400-3.946)

(b)

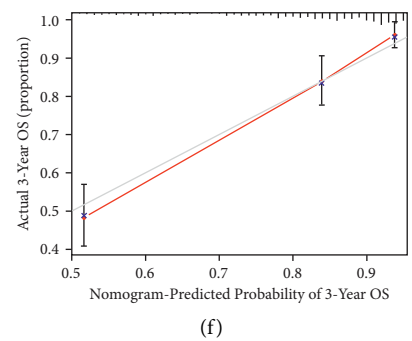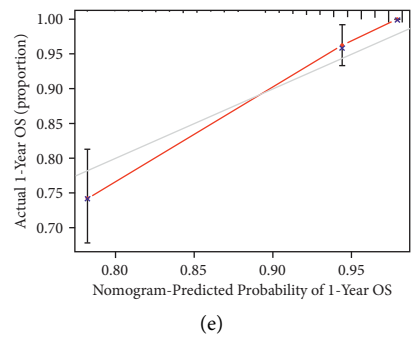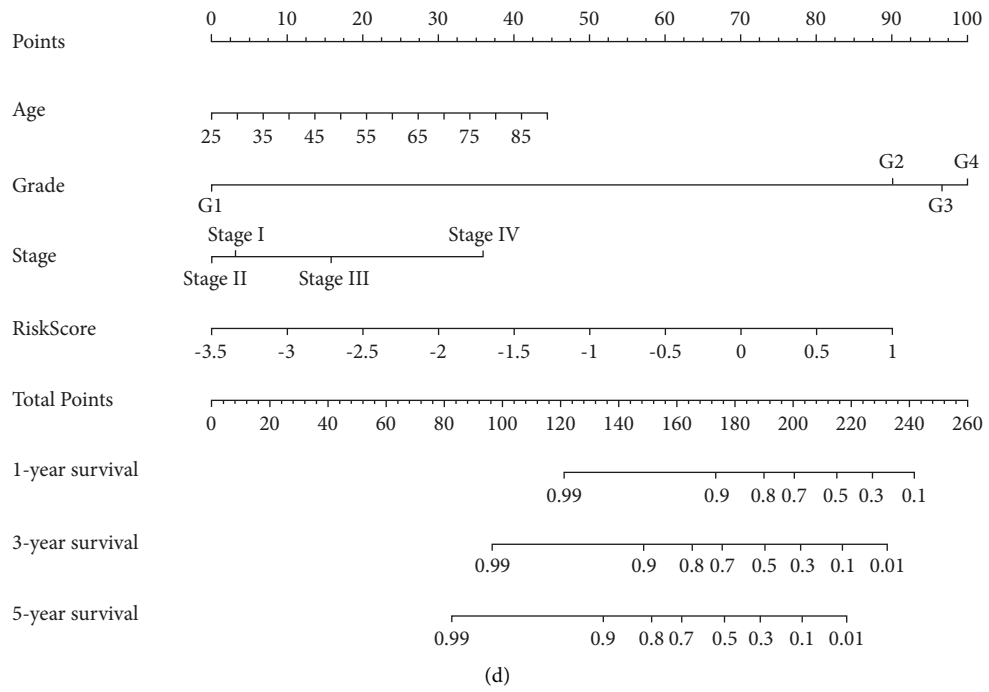|	pvalue	Hazard ratio
age	<0.001	1.033 (1.019-1.049)
grade	0.021	1.329 (1.044-1.691)
stage	0.025	1.623 (1.062-2.481)
T	0.320	0.823 (0.561-1.208)
M	0.233	1.488 (0.775-2.858)
riskScore	<0.001	2.478 (1.864-3.294)

(c)

FIGURE 6: Continued.

(d)
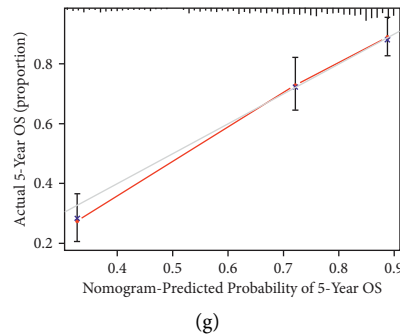

(e)

(f)

FIGURE 6: Continued.

(g)

FIGURE 6: Construction of a new prognostic nomogram. (a) Heatmap and the clinicopathologic characters of the low- and high-risk groups. $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. (b) Univariate Cox regression analysis. (c)Multivariate Cox regression analysis. (d) Prognostic nomogram for predicting the survival of patients with ccRCC. The blue line represents a 95% confidence interval. The position of the square represents the hazard ratio. (e-g) Calibration curves of the nomogram for predicting survival at 1, 3, and 5 years. The nomogram prediction accuracy is higher if the actual curve is closer to the ideal curve. The red line represents the actual curve and the black line represents the ideal line.

4. Discussion

Clear cell renal cell carcinoma is a fatal adult renal cancer [22], with a higher risk of recurrence and metastasis [23]. m7G and m6A are both common types of internal RNA modifications. Interestingly, numerous studies have revealed that m6A RNA methylation regulators are connected with the development and progression of cancer in humans [24], such as bladder cancer [25], hepatocellular carcinoma [26], and breast cancer [27]. However, there are few similar studies on m7G. Therefore, this study explored the impact of m7G RNA methylation regulators on ccRCC and constructed an m7G-related signature and nomogram for prognostic prediction of ccRCC.

First, we filtered out the differentially expressed m7G-related genes with prognostic values. Then we constructed the signature of five m7G-related genes, including CYFIP2, EIF4A1, NUDT1, NUDT10, and NUDT4. The expression of CYFIP2 and NUDT4 was downregulated, whereas EIF4A1, NUDT1, and NUDT10 expressions were upregulated in the ccRCC high-risk group. The cytoplasmic FMR1-interacting protein (CYFIP) family was initially recognized as a protein associated with brain diseases [28]. Recent research demonstrated that Cyfip 2 might play a critical role and have potential functions in cancers, such as gastric cancer [29] and colorectal adenocarcinoma [30]. Previous studies indicated CYFIP2 was downregulated in ccRCC patients due to its DNA methylation, and it was involved in the metabolic reprogramming, the EMT pathway, and immune infiltration processes in ccRCC by its DNA methylation [31]. The eukaryotic initiation factor 4A (eIF4A) family plays a vital role in many cancers [32–34]. In human cancers like gastric cancers and breast cancers, ZBP1 expression is increased, which also correlates with poor prognoses [35, 36]. EIF4A1 is a necessary component of EIF4F, which was considered a direct connection between essential steps in cancer development and translation initiation [37]. Moreover, accumulating studies have linked EIF4A1 to malignant phenotypes of tumor cells and tumor-specific survival [38, 39]. NUDT1, NUDT4, and NUDT10 belong to the nudix

hydroxylase (NUDT) family. The expression of NUDT1 and NUDT10 is increased and the expression of NUDT4 is decreased in ccRCC [40]. Up to now, the association between the NUDT family and tumorigenesis has been unclear, and studies about their role in ccRCC are rare [41–43]. Previous studies indicated that NUDT1 might provide diagnostic and prognostic value for KIRC [44]. A study revealed that NUDT1 is activated by HIF2αtranscription, thereby inhibiting oxidative stress and promoting the progression of ccRCC [45]. The roles of NUDT4 and NUDT10 remain ambiguous in ccRCC tumor progression and metastasis.

An analysis using Cox regression displayed that risk score, age, stage, and grade could act as the independent predictive factors for patients with ccRCC. These independent predictive factors were used to build a nomogram to improve the ability to predict the prognosis for ccRCC. The calibration curves demonstrated that the nomogram was fairly accurate in predicting the actual OS.

After performing GO, KEGG, and immune infiltration analyses on differential risk DEG genes, many biological pathways and functions were identified as contributing to immunity. It might be reasonable to conclude that m7G modification is closely related to tumor immunity. Lines of evidence demonstrate that the progression of ccRCC is associated with the presence of different types of immune cells and immune functions [46]. It has been proven that RNA methylation contributes to tumor immunity [47]. But the relationship between m7G modification and tumor immunity is still developing, and a deeper exploration of the mechanism is needed.

To facilitate the clinical transformation of the five genes model, we further identified FDA-approved sensitive drugs that express high levels of CYFIP2, EIF4A1, NUDT1, NUDT4, and NUDT10 in multiple cancer cell types. A high expression of CYFIP2 is related to cancer cells' sensitivity to nelarabine, Melphalan, and so on. Nelarabine is a process for obtaining the deoxyguanosine analog 9-β-Darabinofuranosylguanine(Ara-G) [48, 49]. Nelarabine has preferential cytotoxicity to T lymphoblastic cells through an
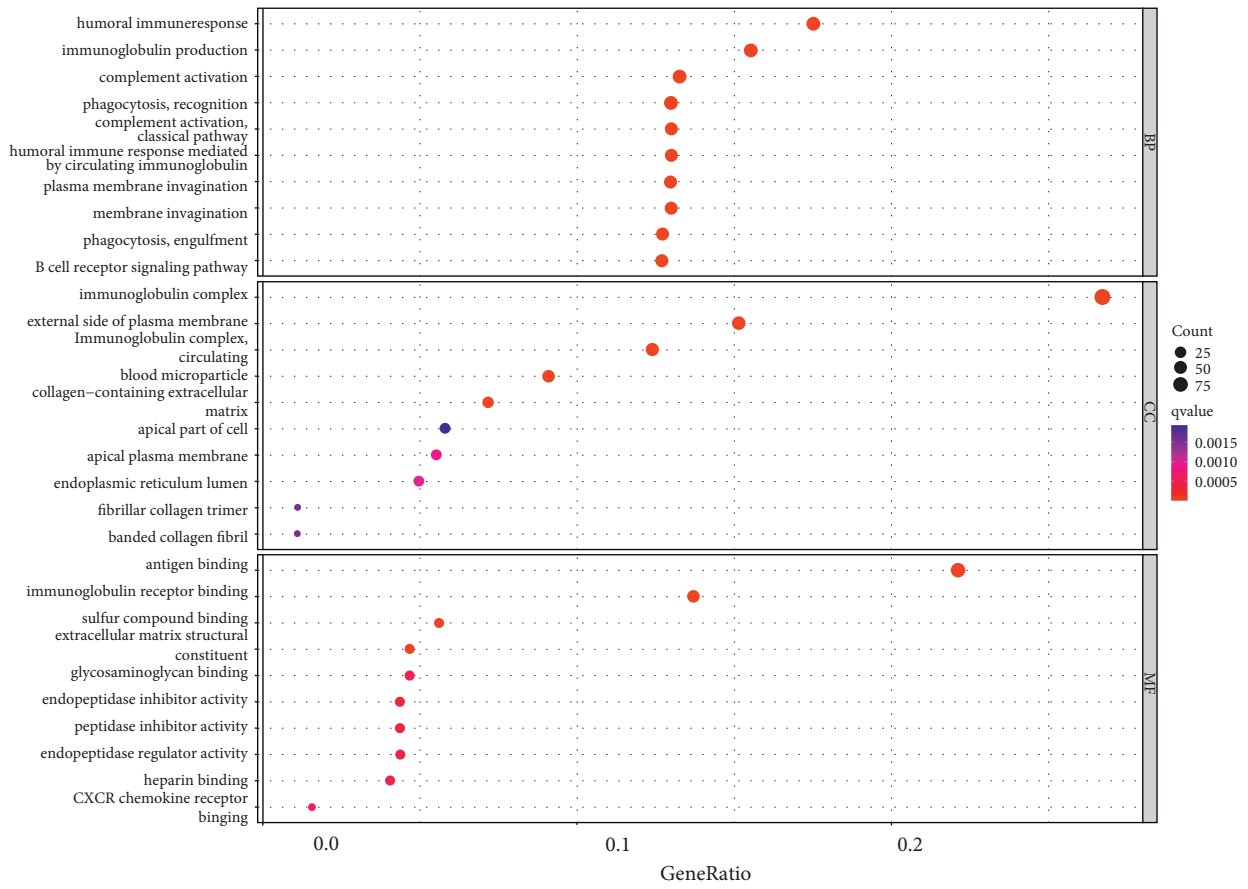
(a)
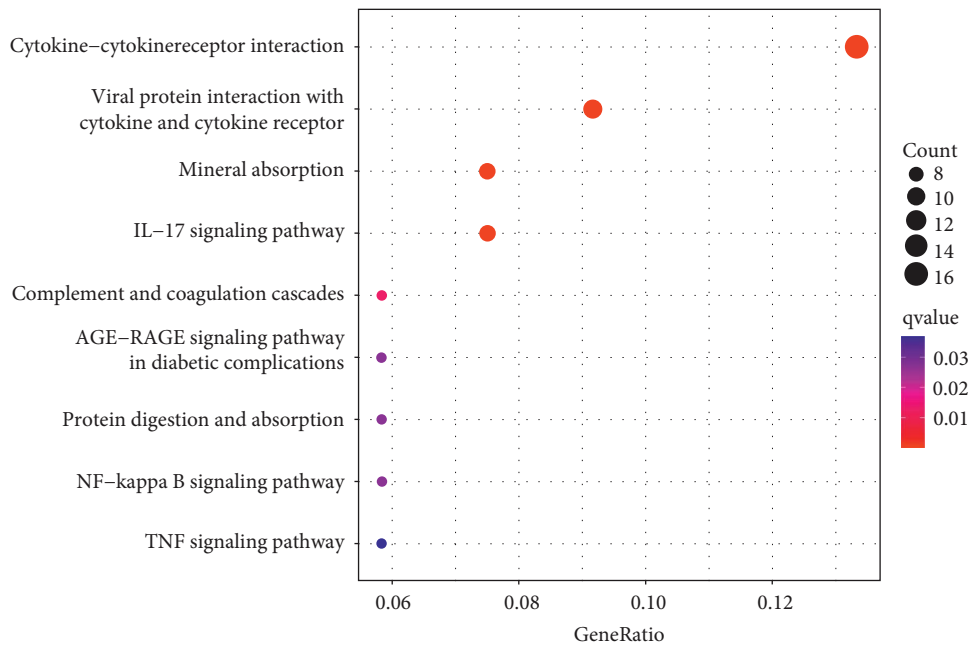
(b)

FIGURE 7: Results of GO and KEGG analysis of the TCGA cohort. (a) GO enrichment analysis revealed the biological processes and molecular functions involved in the differentially expressed genes. The GO enrichment analysis is divided into three parts, including cellular component (CC), molecule function (MF), and biological processes (BP). (b) KEGG analysis shows the signaling pathways involved in the differentially expressed genes.
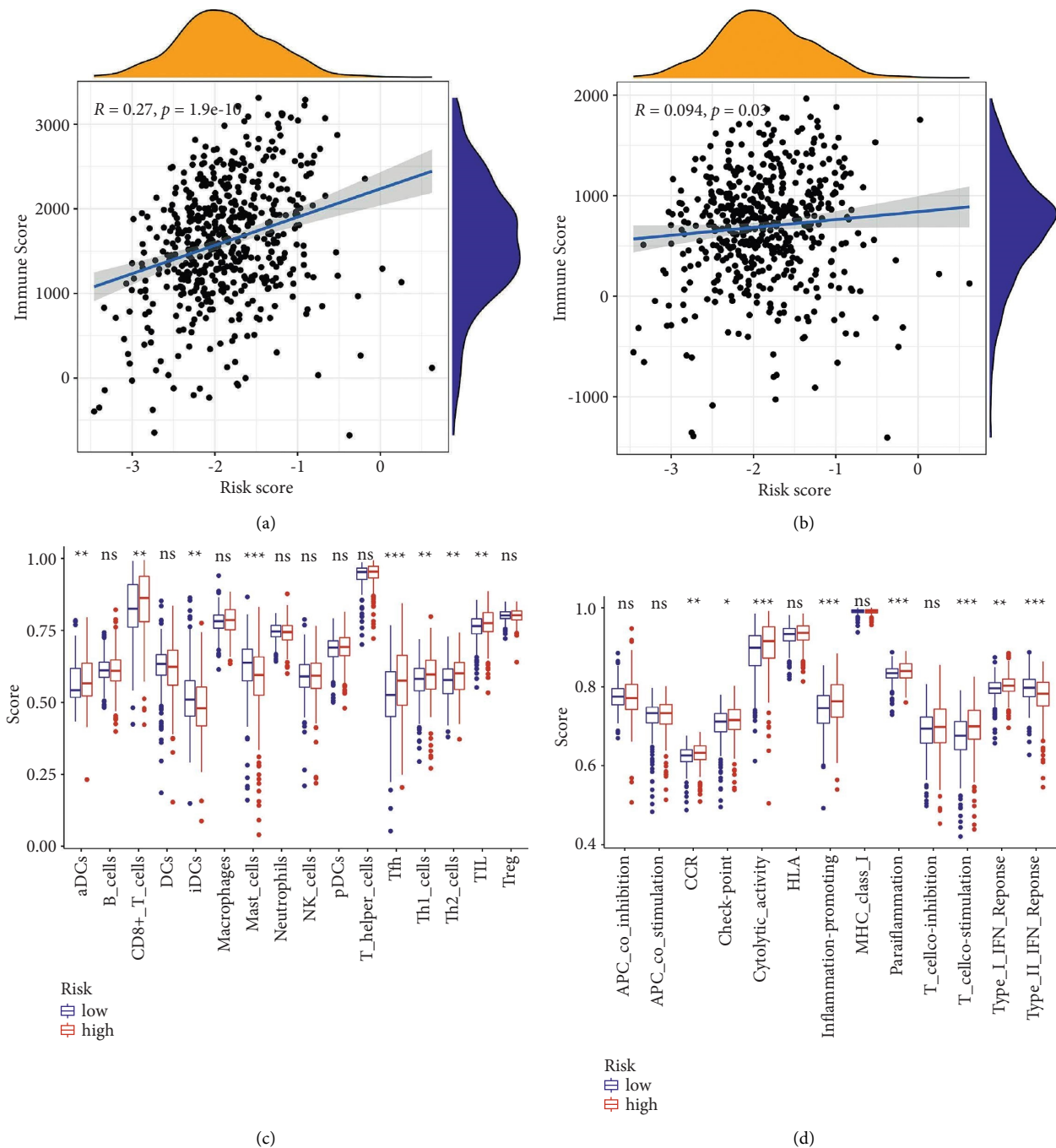
(a)

(b)

(c)

(d)

FIGURE 8: Results of ssGSEA immune infiltration in the TCGA cohort. The association between risk score and (a) immune score or (b) stromal score. (c) 16 immune cell ssGSEA scores between different risk groups. (d) 13 immune-related functional ssGSEA scores between different risk groups. Adjusted P values are: ns, not significant; $^*P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$.

accumulation of Ara-GTP, thereby inhibiting ribonucleotide reductase and subsequent DNA synthesis [50, 51]. According to the score comparisons of immune cells and immune function, we observed that high-risk patients possessed higher scores in helper T cells 2 (Th2 cells), CD8+ T cells, follicular helper T cells (Tfh), helper T cells 1 (Th1 cells), and T cell co-stimulation, suggesting a potential therapeutic agent in ccRCC with nelarabine. However, no literature has been found on the treatment of kidney cancer

by nelarabine. Melphalan plus fludarabine is a feasible and effective RIC regimen for allogeneic SCT in metastatic RCC [52]. Increasing NUDT4 expression has been related to the resistance of tumor cells to everolimus. Nevertheless, it has been associated with increased sensitivity to selumetinib. Everolimus, a mammalian target of rapamycin (mTOR) inhibitor, is the standard second or third line therapy in patients with ccRCC [53]. SELUMETINIB inhibits MEK1/2, a mitogen-activated protein kinase, which can abrogate
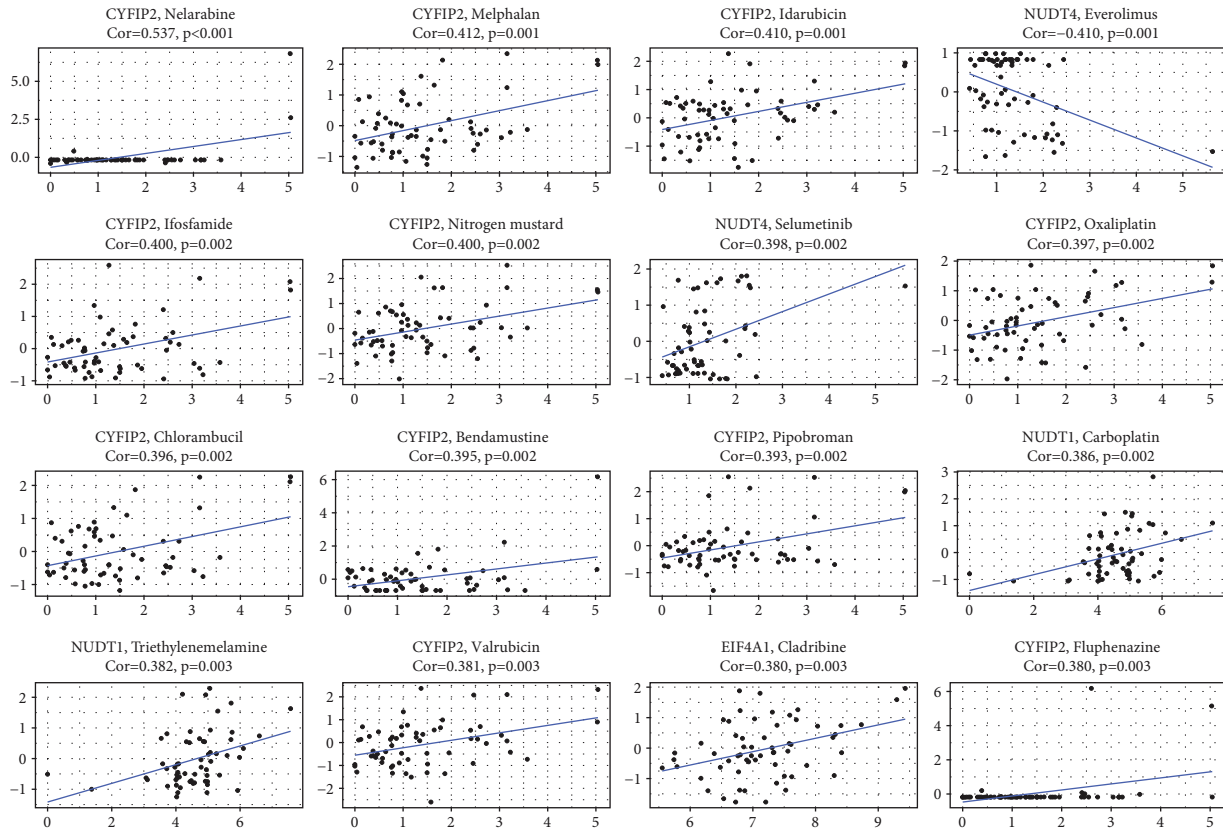
FIGURE 9: Correlation of m7G-related genes expressions with drug response. The relationship between drug sensitivity and CYFIP2, EIF4A1, NUDT1, NUDT1, and NUDT10 expression.

resistance, leading to improved antitumor efficacy in renal cell carcinoma [54]. NUDT1 expression was positively correlated with carboplatin and triethylenemelamine sensitivity in cancer cells. The increased expression of EIF4A1 was connected with higher sensitivity to cladribine in cancer cells. Carboplatin, triethylenemelamine, and cladribine are some of the most widely used cancer chemotherapeutics, but ccRCC is not sensitive to chemotherapy [55].

Currently, a large amount of literature on gene markers predicting the prognosis of ccRCC patients can be found. Some studies have constructed models without internal or external validation [56, 57] or only constructed prediction models without nomogram diagrams [57, 58]. Wenwei Chen and colleagues included age, sex, grade, and stage in univariate regression analysis but did not include TNM clinical indicators [59]. Zhao et al. constructed a nomograph diagram, including age, grade, stage, and risk score. Risk-score is classified as low and high rather than specific numerical values [56]. Compared with other gene predictive models of ccRCC patients, our research makes up for these deficiencies and has specific innovations and advantages. First of all, it is undeniable that our study is the first comprehensive study of m7G-related mRNA in ccRCC. In addition, the nomogram diagrams contain risk scores and clinical indicators that can further effectively predict the outcome of patients with ccRCC. Furthermore, Zhao et al. only evaluated the accuracy of the nomogram diagram, but not

the accuracy of the gene-related model [56]. However, we evaluated the accuracy of the model using ROC. The closer the AUC is to 1.0, the higher the authenticity of the predictive model. Wang et al. also estimated the model's accuracy using ROC, and the risk AUC was 0.704 [60]. Our study demonstrates that the risk AUC is 0.748 in the TCGA cohort and the risk AUC is 0.811 in the E-MTAB-1980 cohort. It shows that our model is more accurate and effective. This study still has some flaws in its process and methodology. Firstly, the ccRCC cohort is relatively small, and the samples lack complete clinical information. Secondly, how those m7G-related genes interact with each other and which pathway is involved in ccRCC needs further study. Finally, the molecular mechanism of m7G-related genes in ccRCC remains to be conducted both in internal and external experiments.

5. Conclusion

Collectively, we constructed five m7G-related gene predictive models, which performed well in predicting survival prognosis, immune microenvironment, and sensitivity to drugs in ccRCC. Then, a prognostic nomogram for survival prediction based on five m7G-related genes could improve survival estimates for ccRCC patients. However, the potential mechanism of m7G-related genes in ccRCC remains unclear and needs further exploration.

Data Availability

All datasets are publicly available, and we can find all datasets here: https://portal.gdc.cancer.gov/and https://www.ebi.ac.uk/arrayexpress.

Conflicts of Interest

In this work, none of the authors has conflicts of interest.

Authors' Contributions

Qing Jiang designed and supervised the study. Yongxin Fu, Jiawu Wang, and Zhiya Hu organized the data. Yongxin Fu and Jiawu Wang performed the analysis. Yongxin Fu checked the statistical method. Yongxin Fu, Jiawu Wang, Zhiya Hu, and Yang Gou prepared the figures. Yongxin Fu, Jiawu Wang, Zhiya Hu, Yang Gou, and Yisen Li co-wrote this manuscript. All authors agree on the submitted version.

Acknowledgments

References

[1] B. Ljungberg, S. C. Campbell, H. Y. Cho et al., "The epidemiology of renal cell carcinoma," *European Urology*, vol. 60, no. 4, pp. 615–621, 2011.

[2] H. Moch, A. L. Cubilla, P. A. Humphrey, V. E. Reuter, and T. M. Ulbright, "The 2016 WHO classification of tumours of the urinary system and male genital organs-Part A: renal, penile, and testicular tumours," *European Urology*, vol. 70, no. 1, pp. 93–105, 2016.

[3] B. Ljungberg, K. Bensalah, S. Canfield et al., "EAU guidelines on renal cell carcinoma: 2014 update," *European Urology*, vol. 67, no. 5, pp. 913–924, 2015.

[4] A. Sánchez-Gastaldo, E. Kempf, A. González Del Alba, and I. Duran, "Systemic treatment of renal cell cancer: a comprehensive review," *Cancer Treatment Reviews*, vol. 60, pp. 77–89, 2017.

[5] S. Gulati and U. Vaishampayan, "Current state of systemic therapies for advanced renal cell carcinoma," *Current Oncology Reports*, vol. 22, no. 3, p. 26, 2020.

[6] S. R. Jaffrey, "An expanding universe of mRNA modifications," *Nature Structural & Molecular Biology*, vol. 21, no. 11, pp. 945-946, 2014.

[7] S. Zaccara, R. J. Ries, and S. R. Jaffrey, "Reading, writing and erasing mRNA methylation," *Nature Reviews Molecular Cell Biology*, vol. 20, no. 10, pp. 608–624, 2019.

[8] I. A. Roundtree, M. E. Evans, T. Pan, and C. He, "Dynamic RNA modifications in gene expression regulation," *Cell*, vol. 169, no. 7, pp. 1187–1200, 2017.

[9] K. Du, L. Zhang, T. Lee, and T. Sun, "m(6 A RNA methylation controls neural development and is involved in human diseases," *Molecular Neurobiology*, vol. 56, no. 3, pp. 1596–1606, 2019.

[10] M. T. Bohnsack and K. E. Sloan, "Modifications in small nuclear RNAs and their roles in spliceosome assembly and function," *Biological Chemistry*, vol. 399, no. 11, pp. 1265–1276, 2018.

[11] Y. Furuichi, "Discovery of m(7)G-cap in eukaryotic mRNAs," *Proceedings of the Japan Academy, Series B*, vol. 91, no. 8, pp. 394–409, 2015.

[12] C. Enroth, L. D. Poulsen, S. Iversen, F. Kirpekar, A. Albrechtsen, and J. Vinther, "Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing," *Nucleic Acids Research*, vol. 47, no. 20, 2019.

[13] Y. Pei and S. Shuman, "Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5," *Journal of Biological Chemistry*, vol. 277, no. 22, pp. 19639–19648, 2002.

[14] M. M. Konarska, R. A. Padgett, and P. A. Sharp, "Recognition of cap structure in splicing in vitro of mRNA precursors," *Cell*, vol. 38, no. 3, pp. 731–736, 1984.

[15] D. R. Drummond, J. Armstrong, and A. Colman, "The effect of capping and polyadenylation on the stability, movement and translation of synthetic messenger RNAs in Xenopus oocytes," *Nucleic Acids Research*, vol. 13, no. 20, pp. 7375–7394, 1985.

[16] J. D. Lewis and E. Izaurflde, "The role of the cap structure in RNA processing and nuclear export," *European Journal of Biochemistry*, vol. 247, no. 2, pp. 461–469, 1997.

[17] S. Muthukrishnan, G. W. Both, Y. Furuichi, and A. J. Shatkin, "5'-Terminal 7-methylguanosine in eukaryotic mRNA is required for translation," *Nature*, vol. 255, no. 5503, pp. 33–37, 1975.

[18] J. Ma, H. Han, Y. Huang et al., "METTL1/WDR4-mediated m(7)G tRNA modifications and m(7)G codon usage promote mRNA translation and lung cancer progression," *Molecular Therapy*, vol. 29, no. 12, pp. 3422–3435, 2021.

[19] Y. Liu, Y. Zhang, Q. Chi, Z. Wang, and B. Sun, "RETRACTED: methyltransferase-like 1 (METTL1) served as a tumor suppressor in colon cancer by activating 7-methyguanosine (m7G) regulated let-7e miRNA/HMGA2 axis," *Life Sciences*, vol. 249, Article ID 117480, 2020.

[20] Y. Sato, T. Yoshizato, Y. Shiraishi et al., "Integrated molecular analysis of clear-cell renal cell carcinoma," *Nature Genetics*, vol. 45, no. 8, pp. 860–867, 2013.

[21] C. Tomikawa, "7-Methylguanosine modifications in transfer RNA (tRNA)," *International Journal of Molecular Sciences*, vol. 19, no. 12, p. 4080, 2018.

[22] G. Scelo and T. L. Larose, "Epidemiology and risk factors for kidney cancer," *Journal of Clinical Oncology*, vol. 36, no. 36, pp. 3574–3581, 2018.

[23] M. B. Amin, M. B. Amin, P. Tamboli et al., "Prognostic impact of histologic subtyping of adult renal epithelial neoplasms: an experience of 405 cases," *The American Journal of Surgical Pathology*, vol. 26, no. 3, pp. 281–291, 2002.

[24] S. Wang, C. Sun, J. Li et al., "Roles of RNA methylation by means of N(6)-methyladenosine (m(6)A) in human cancers," *Cancer Letters*, vol. 408, pp. 112–120, 2017.

[25] M. Chen, Z. Y. Nie, X. H. Wen, Y. H. Gao, H. Cao, and S. F. Zhang, "m6A RNA methylation regulators can contribute to malignant progression and impact the prognosis of bladder cancer," *Bioscience Reports*, vol. 39, no. 12, Article ID BSR20192892, 2019.

[26] N. Qu, X. Bo, B. Li et al., "Role of N6-methyladenosine (m(6) A) methylation regulators in hepatocellular carcinoma," *Frontiers in Oncology*, vol. 11, Article ID 755206, 2021.

[27] F. Liu, X. Yu, and G. He, "m6A-Mediated tumor invasion and methylation modification in breast cancer microenvironment," *Journal of oncology*, vol. 2021, Article ID 9987376, 17 pages, 2021.

[28] S. Abekhoukh and B. Bardoni, "CYFIP family proteins between autism and intellectual disability: links with Fragile X syndrome," *Frontiers in Cellular Neuroscience*, vol. 8, 2014.

[29] S. Jiao, N. Li, S. Cai, H. Guo, and Y. Wen, "Inhibition of CYFIP2 promotes gastric cancer cell proliferation and chemoresistance to 5-fluorouracil through activation of the Akt signaling pathway," *Oncology Letters*, vol. 13, no. 4, pp. 2133–2140, 2017.

[30] R. S. Jackson, Y. J. Cho, S. Stein, and P. Liang, "CYFIP2, a direct p53 target, is leptomycin-B sensitive," *Cell Cycle*, vol. 6, no. 1, pp. 95–103, 2007.

[31] J. Tong, X. Meng, Q. Lv et al., "The downregulation of prognosis- and immune infiltration-related gene CYFIP2 serves as a novel target in ccRCC," *International Journal of General Medicine*, vol. 14, pp. 6587–6599, 2021.

[32] A. Wolf and M. Hatzfeld, "A role of plakophilins in the regulation of translation," *Cell Cycle*, vol. 9, no. 15, pp. 3045–3050, 2010.

[33] A. Fukao, Y. Mishima, N. Takizawa et al., "MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans," *Molecular Cell*, vol. 56, no. 1, pp. 79–89, 2014.

[34] S. Sridharan, M. Robeson, D. Bastihalli-Tukaramrao et al., "Targeting of the eukaryotic translation initiation factor 4A against breast cancer stemness," *Frontiers in Oncology*, vol. 9, 2019.

[35] C. Gao, X. Guo, A. Xue, Y. Ruan, H. Wang, and X. Gao, "High intratumoral expression of eIF4A1 promotes epithelial-to-mesenchymal transition and predicts unfavorable prognosis in gastric cancer," *Acta Biochimica et Biophysica Sinica*, vol. 52, no. 3, pp. 310–319, 2020.

[36] A. Modelska, E. Turro, R. Russell et al., "The malignant phenotype in breast cancer is driven by eIF4A1-mediated changes in the translational landscape," *Cell Death & Disease*, vol. 6, no. 1, Article ID e1603, 2015.

[37] J. Pelletier, J. Graff, D. Ruggero, and N. Sonenberg, "Targeting the eIF4F translation initiation complex: a critical nexus for cancer development," *Cancer Research*, vol. 75, no. 2, pp. 250–263, 2015.

[38] S. Nagel, L. Venturini, C. Meyer et al., "Multiple mechanisms induce ectopic expression of LYL1 in subsets of T-ALL cell lines," *Leukemia Research*, vol. 34, no. 4, pp. 521–528, 2010.

[39] S. Liang, Y. Zhou, Y. Chen, G. Ke, H. Wen, and X. Wu, "Decreased expression of EIF4A1 after preoperative brachytherapy predicts better tumor-specific survival in cervical cancer," *International Journal of Gynecological Cancer*, vol. 24, no. 5, pp. 908–915, 2014.

[40] Y. Wang, F. Wan, K. Chang, X. Lu, B. Dai, and D. Ye, "NUDT expression is predictive of prognosis in patients with clear cell renal cell carcinoma," *Oncology Letters*, vol. 14, no. 5, pp. 6121–6128, 2017.

[41] A. R. Dallosso, S. Dolwani, N. Jones et al., "Inherited predisposition to colorectal adenomas caused by multiple rare alleles of MUTYH but not OGG1, NUDT1, NTH1 or NEIL 1, 2 or 3," *Gut*, vol. 57, no. 9, pp. 1252–1255, 2008.

[42] W. C. Cho, A. S. Chow, and J. S. Au, "MiR-145 inhibits cell proliferation of human lung adenocarcinoma by targeting EGFR and NUDT1," *RNA Biology*, vol. 8, no. 1, pp. 125–131, 2011.

[43] P. Garre, V. Briceño, R. M. Xicola et al., "Analysis of the oxidative damage repair genes NUDT1, OGG1, and MUTYH in patients from mismatch repair proficient HNPCC families (MSS-HNPCC)," *Clinical Cancer Research*, vol. 17, no. 7, pp. 1701–1712, 2011.

[44] J. Xie, L. Cui, S. Pan, D. Liu, F. Liu, and Z. Liu, "Metabolic understanding of the genetic dysregulation in the tumor microenvironment of kidney renal clear cell carcinoma," *Disease Markers*, vol. 2022, Article ID 6085072, 17 pages, 2022.

[45] J. Shi, Z. Xiong, K. Wang et al., "HIF2α promotes tumour growth in clear cell renal cell carcinoma by increasing the expression of NUDT1 to reduce oxidative stress," *Clinical and Translational Medicine*, vol. 11, no. 11, 2021.

[46] D. Bai, H. Feng, J. Yang, A. Yin, A. Qian, and H. Sugiyama, "Landscape of immune cell infiltration in clear cell renal cell carcinoma to aid immunotherapy," *Cancer Science*, vol. 112, no. 6, pp. 2126–2139, 2021.

[47] Q. Lan, P. Y. Liu, J. Haase, J. L. Bell, S. Hüttelmaier, and T. Liu, "The critical role of RNA m(6)A methylation in cancer," *Cancer Research*, vol. 79, no. 7, pp. 1285–1292, 2019.

[48] D. J. DeAngelo, D. Yu, J. L. Johnson et al., "Nelarabine induces complete remissions in adults with relapsed or refractory T-lineage acute lymphoblastic leukemia or lymphoblastic lymphoma: cancer and Leukemia Group B study 19801," *Blood*, vol. 109, no. 12, pp. 5136–5142, 2007.

[49] J. Kurtzberg, "The long and winding road of the clinical development of Nelarabine," *Leukemia and Lymphoma*, vol. 48, no. 1, pp. 1-2, 2007.

[50] C. O. Rodriguez, C. M. Stellrecht, and V. Gandhi, "Mechanisms for T-cell selective cytotoxicity of arabinosylguanine," *Blood*, vol. 102, no. 5, pp. 1842–1848, 2003.

[51] R. A. Larson, "Three new drugs for acute lymphoblastic leukemia: nelarabine, clofarabine, and forodesine," *Seminars in Oncology*, vol. 34, no. 6 Suppl 5, pp. S13–S20, 2007.

[52] N. T. Ueno, Y. C. Cheng, G. Rondón et al., "Rapid induction of complete donor chimerism by the use of a reduced-intensity conditioning regimen composed of fludarabine and melphalan in allogeneic stem cell transplantation for metastatic solid tumors," *Blood*, vol. 102, no. 10, pp. 3829–3836, 2003.

[53] M. Meskawi, R. Valdivieso, P. Dell'Oglio, V. Trudeau, A. Larcher, and P. I. Karakiewicz, "The role of Everolimus in renal cell carcinoma," *Journal of kidney cancer and VHL*, vol. 2, no. 4, pp. 187–194, 2015.

[54] C. M. Diaz-Montero, F. J. Mao, J. Barnard et al., "MEK inhibition abrogates sunitinib resistance in a renal cell carcinoma patient-derived xenograft model," *British Journal of Cancer*, vol. 115, no. 8, pp. 920–928, 2016.

[55] P. Makhov, S. Joshi, P. Ghatalia, A. Kutikov, R. G. Uzzo, and V. M. Kolenko, "Resistance to systemic therapies in clear cell renal cell carcinoma: mechanisms and management strategies," *Molecular Cancer Therapeutics*, vol. 17, no. 7, pp. 1355–1364, 2018.

[56] G. J. Zhao, Z. Wu, L. Ge et al., "Ferroptosis-related gene-based prognostic model and immune infiltration in clear cell renal cell carcinoma," *Frontiers in Genetics*, vol. 12, Article ID 650416, 2021.

[57] S. Ma, M. Zhao, J. Fan et al., "Analysis of ferroptosis-related gene expression and prognostic factors of renal clear cell carcinoma based on TCGA database," *International Journal of General Medicine*, vol. 14, pp. 5969–5980, 2021.

[58] Y. Hong, M. Lin, D. Ou, Z. Huang, and P. Shen, "A novel ferroptosis-related 12-gene signature predicts clinical

prognosis and reveals immune relevancy in clear cell renal cell carcinoma," *BMC Cancer*, vol. 21, no. 1, p. 831, 2021.

[59] W. Chen, W. Lin, L. Wu, A. Xu, C. Liu, and P. Huang, "A novel prognostic predictor of immune microenvironment and therapeutic response in kidney renal clear cell carcinoma based on necroptosis-related gene signature," *International Journal of Medical Sciences*, vol. 19, no. 2, pp. 377–392, 2022.

[60] J. Wang, C. Zhang, W. He, and X. Gou, "Effect of m(6)A RNA methylation regulators on malignant progression and prognosis in renal clear cell carcinoma," *Frontiers in Oncology*, vol. 10, p. 3, 2020.