

ORIGINAL RESEARCH

Construction and validation of m⁶A RNA methylation regulators associated prognostic model for gastrointestinal cancer

Yandong Miao¹  | Bin Su²  | Xiaolong Tang¹  | Jiangtao Wang¹  |
Wuxia Quan³  | Yonggang Chen⁴  | Denghai Mi^{1,5} 

¹The First Clinical Medical College of Lanzhou University, Lanzhou, China

²Department of Oncology, The 920th Hospital of the Chinese People's Liberation Army Joint Logistic Support Force, Kunming, China

³Qingyang People's Hospital, Qingyang, China

⁴Lanzhou University Second Hospital, Lanzhou, China

⁵Gansu Academy of Traditional Chinese Medicine, Lanzhou, China

Correspondence

Denghai Mi, The First Clinical Medical College of Lanzhou University, No. 1 Donggang West Road, Chengguan District, Lanzhou 730000, China.
Gansu Academy of Traditional Chinese Medicine, Gansu Provincial Hospital of Traditional Chinese Medicine, No. 418 Guazhou Road, Qilihe District, Lanzhou, China.
Email: mi.dh@outlook.com

Yonggang Chen, Lanzhou University Second Hospital, No.82, Cuiyingmen, Chengguan District, Lanzhou 730030, China.
Email: chenyg15@lzu.edu.cn

Abstract

N⁶-methyladenosine (m⁶A) RNA methylation is correlated with carcinogenesis and dynamically possessed through the m⁶A RNA methylation regulators. This paper aimed to explore 13 m⁶A RNA methylation regulators' role in gastrointestinal cancer (GIC) and determine the risk model and prognosis value of m⁶A RNA methylation regulators in GIC. We used several bioinformatics methods to identify the differential expression of m⁶A RNA methylation regulators in GIC, constructed a prognostic model, and carried out functional enrichment analysis. Eleven of 13 m⁶A RNA methylation regulators were differentially expressed in different clinicopathological characteristics of GIC, and m⁶A RNA methylation regulators were nearly associated with GIC. We constructed a risk model based on five m⁶A RNA methylation regulators (METTL3, FTO, YTHDF1, ZC3H13, and WTAP); the risk score is an independent prognosis biomarker. Moreover, the five m⁶A RNA methylation regulators can also forecast the 1-, 3- and 5-year overall survival through a nomogram. Furthermore, four hallmarks of oxidative phosphorylation, glycolysis, fatty acid metabolism, and cholesterol homeostasis gene sets were significantly enriched in GIC. m⁶A RNA methylation regulators were related to the malignant clinicopathological characteristics of GIC and may be used for prognostic stratification and development of therapeutic strategies.

KEYWORDS

bioinformatic analysis, epigenetics, gastrointestinal carcinoma, m⁶A methylation, RNA modification

1 | INTRODUCTION

It is generally accepted that epigenetics mainly involves DNA methylation, nucleosome remodelling and histone modification, and non-coding RNAs are essential to the genesis of cancer [1]. In the past decade, RNA sequencing (RNA-seq) has become an essential tool for transcriptional genome-wide analysis of differential gene expression [2]. Many RNA modifications have been recognized in various RNAs, such as mRNAs, tRNAs, micro-RNAs, small nuclear RNAs, rRNAs, and long non-coding RNA [3–6]. It is reported that these RNA modifications have several modes, including 5-methylcytosine, N1-methyladenosine, N₆,2'-O-

dimethyladenosine (m⁶A), N7-methyladenosine, and 2'-O-methylation [7, 8]. The m⁶A modification is the first determined and most widespread pattern of mRNA methylation in eukaryotes [9–12].

RNA modification is mediated through 'writers'-methyltransferases, 'readers'-binding proteins, and 'erasers'-demethylases [13]. m⁶A methylation regulators made up of 'writers', such as methyltransferase like 3 (METTL3), KIAA1429, RNA binding motif protein 15 (RBM15), methyltransferase like 4 (METTL14), zinc finger CCCH domain-containing protein 13 (ZC3H13), and WT1-related protein (WTAP). 'Readers' such as YTH N⁶-methyladenosine RNA binding protein 1 (YTHDF1), heterogeneous nuclear ribonucleoprotein C (HNRNPC), YTH

Yandong Miao, Bin Su, and Xiaolong Tang have equal contributions to this study.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. *IET Systems Biology* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

domain-containing 1 (YTHDC1), YTH domain-containing 2 (YTHDC2), and YTH N6-methyladenosine RNA binding protein 2 (YTHDF2). ‘Erasers’ such as α -ketoglutarate-dependent dioxygenase alkB homologue 5 (ALKBH5) and fat mass- and obesity-associated protein (FTO) [7, 13]. m⁶A RNA methylation displayed an invertible and crucial biological process by regulation of ‘readers’, ‘writers’, and ‘erasers’. The authentication of m⁶A RNA methylation regulators has remarkably enhanced the character of m⁶A modification in the gene expression regulation. The crucial function of m⁶A RNA methylation regulators is also observed in various cancers, including ovarian cancer, acute myeloid leukemia, renal clear cell carcinoma, glioma, and glioblastoma [14–18].

Emerging evidence demonstrated that m⁶A methylation was closely associated with the occurrence, development, metastasis, and angiogenesis of gastrointestinal cancer (GIC) [19]. METTL3 promotes tumour angiogenesis in gastric cancer and facilitates CRC growth by preventing SOX2 mRNA degradation [20, 21]. Oncogene c-Myc promotes YTHDF1 expression in CRC [22]. Besides, YTHDF1 inhibited the Wnt/ β -catenin pathway activity of CRC cells [23]. Nevertheless, literature was deficient in comprehensively analysing the expression of m⁶A RNA methylation regulators in GIC with various clinicopathological features and their role and prognosis value in the malignant development of GIC.

In the present research, we comprehensively analysed 13 extensively reported m⁶A RNA regulators [16, 17, 24, 25] using mRNA sequencing data of the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) datasets in GIC. The data of TCGA is the training cohort, and the data of GEO is the validation cohort. We obtained the expression data of each m⁶A modification regulator concerning various clinicopathological features. Then, a specific prognostic model was constructed and filtered out by five key m⁶A RNA methylation regulators, and a risk score was used to classify and predict the GIC patients' prognosis. We also discovered a conspicuously increased expression of FTO and ZC3H13 in GIC with malignant clinicopathological characteristics and compared with the previous research. Our work provided a novel risk signature as an independent prognostic biomarker for GIC patients. The flow chart of our study is manifested in Figure 1.

2 | MATERIALS AND METHODS

2.1 | Data sources

Both the RNA-sequencing (RNA-seq) and clinicopathological data of gastrointestinal tumours, including stomach adenocarcinoma (STAD), colon adenocarcinoma (COAD), rectosigmoid junction cancer, and rectum adenocarcinoma (READ), were downloaded from the TCGA database (Data release 23.0 - April 7, 2020, <https://tcga-data.nci.nih.gov/tcga/>). mRNA expression was selected from the matrix file obtained from the RNA-seq data. GSE39582 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39582>), and

GSE87211, (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87211>), obtained from GEO (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) for the validation studies [26, 27]. Gene chips' original data were normalized using the Robust Multi-Array Average (RMA) algorithm provided by R-package ‘limma’ [28]. We used the Perl script and R-package ‘sva’ to merge the different microarray data from GSE39582 and GSE87211 [29] and reduce the heterogeneity. The m⁶A RNA methylation regulators are obtained from previously published literature [13, 15, 17]. 13 m⁶A RNA methylation regulators were obtained. Data annotation and extracting gene expression of m⁶A RNA methylation regulators from the TCGA and GEO data were done through the R software (version 3.6.2). The data are obtained from TCGA and GEO databases, strictly following the publication guidelines approved by TCGA and GEO. Therefore, there is no requirement for ethics committee approval.

2.2 | Identification m⁶A RNA methylation regulators' expression in GIC

Perl script was used for data extraction, and integration and the Wilcoxon test were utilised to screen the expression of m⁶A RNA methylation regulators in GIC with disparate clinicopathological features. $p < 0.05$ are set as the cut-offs. $p < 0.05$ annotated ‘*’, $p < 0.01$ annotated ‘**’ and $p < 0.001$ annotated ‘***’. Bidirectional hierarchical clustering analysis and a drawing heat map were performed by R-package ‘pheatmap’. R-package ‘vioplot’ was used to draw the violin plot.

2.3 | Accordance clustering of m⁶A RNA methylation regulators to determine two clusters in GIC and perform survival analysis

To explore the function of m⁶A RNA methylation regulators in GIC, we divided GIC patients into two groups through the R-package ‘ConsensusClusterPlus’ (50 iterations, Pearson correlation resample rate is 0.8, <http://www.bioconductor.org/>) based on the m⁶A RNA methylation regulators' expression. Principal component analysis analysis is implemented through the R-package ‘limma’ using R software to research the gene expression modes in different GIC clusters. The survival analysis was performed by the Kaplan-Meier method and log-rank test ($p < 0.05$).

2.4 | Exploring the prognostic value and constructing a prognosis model of m⁶A RNA methylation regulators in GIC

We performed a univariate Cox analysis to evaluate the prognostic value of m⁶A RNA methylation regulators in GIC. The least absolute shrinkage and selection operator (Lasso) Cox regression analysis was utilised to construct the optimal model of m⁶A RNA methylation regulators. Each patient's risk score was computed by the following formula: Risk

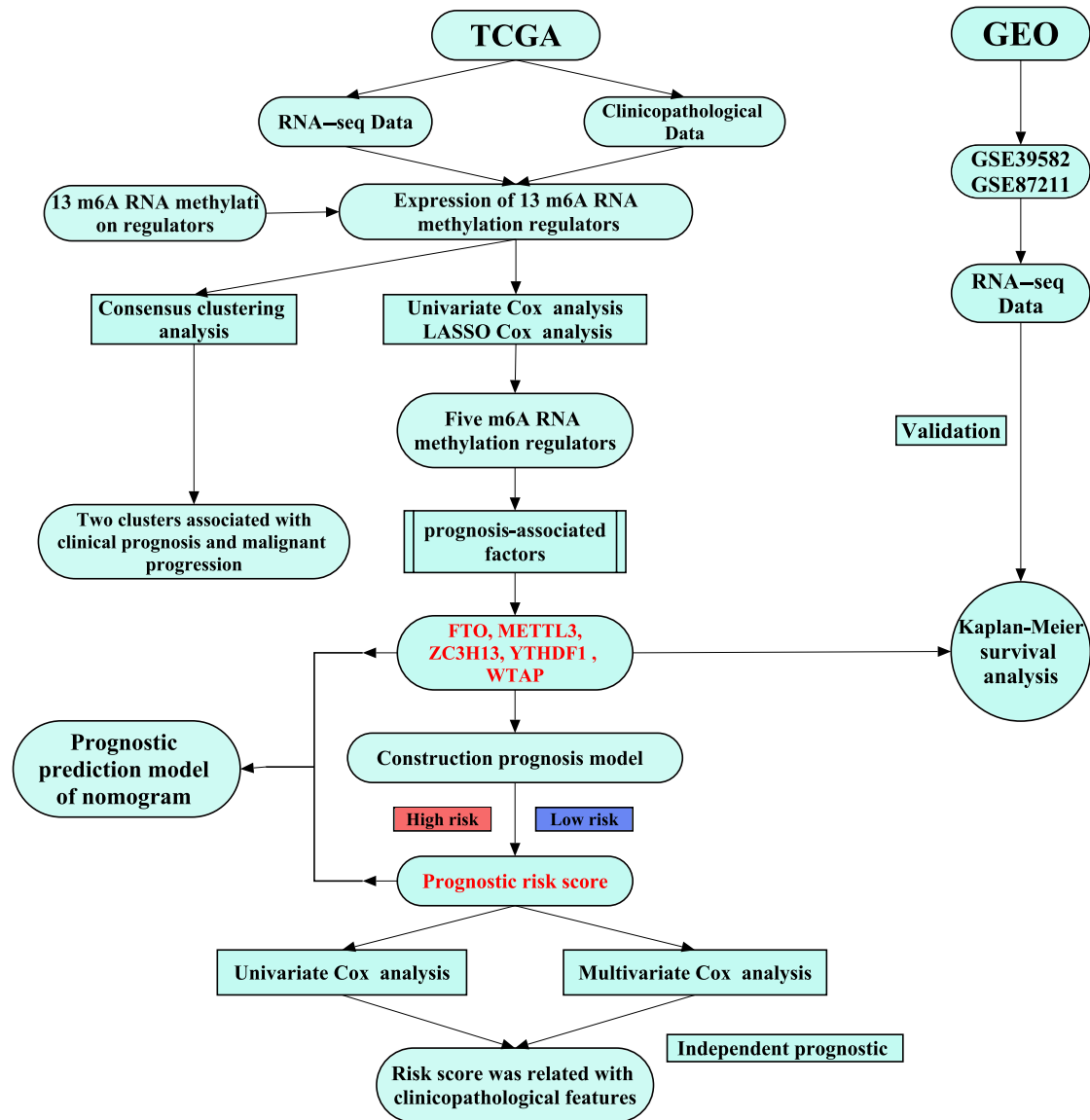


FIGURE 1 A flow chart of the study design and analysis

score = $\sum_{j=1}^n Coefj * Xj$, $Coefj$ demonstrating the coefficient and Xj indicating the relative expression levels of every m⁶A RNA methylation regulator standardized by z-score [30]. The median risk score is chosen as a cut-off value of the bisection GIC cohort.

2.5 | Clinical correlation analysis of the risk score

The R-package ‘survival ROC’ was utilised to draw the Receiver Operating Characteristic (ROC) curve, which was used to test the sensitivity and specificity of the risk score to predict survival. The prognostic accuracy is evaluated by the area under the curve (AUC) value. Depending on the different patients' risk scores, we drew the survival state diagram, risk curve, and heat map. Univariate and multivariate Cox

proportional regression analyses are used to perform independent prognostic analysis of m⁶A RNA methylation regulators. R-package ‘beeswarm’ is used to conduct clinical correlation analysis. The prognostic m⁶A RNA methylation regulators have verified in the GSE87211 and GSE39582 cohorts.

2.6 | Drawing the nomogram

The expression of FTO, METTL3, WTAP, YTHDF1, and ZC3H13 was used to construct a nomogram via R-package ‘Formula’, ‘foreign’, ‘ggplot2’, ‘Hmisc’, ‘lattice’, and ‘rms’. Furthermore, a calibration curve was utilised to evaluate the consistency between actual and predicted survival. Besides, we evaluated the performance of the model in predicting prognosis through the consistency index (c-index). The C-index value of 0.5 and 1.0 represents a random possibility and an

exceptional capacity to predict survival with the model, respectively.

2.7 | Correlation and functional enrichment analysis of m⁶A RNA methylation regulators in GIC

The correlation analysis of m⁶A RNA methylation regulators is performed by the R-package ‘corrplot’. Functional annotation and enrichment analysis of five m⁶A RNA methylation regulators are based on R-package ‘clusterProfiler,’ ‘ggplot2,’ ‘enrichplot,’ and ‘org.Hs.for example.db’, which classified Gene Ontology (GO) to the Biological Processes (BP), false discovery rate (FDR) <0.05 is used as the cut-off. The Gene Set Enrichment Analysis (GSEA) is used to execute the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways and hallmark of GIC. Permutation number is set to 1,000, and FDR <0.25 was recognized as statistically significant [31].

3 | RESULTS

3.1 | Patients' features

TCGA GIC cohort consisted of 1067 patients. 821 patients were included in our analysis and patients with incomplete clinical information and those with survival time of less than 90 days, were excluded. The GSE39582 cohort included 585 patients, 19 colon mucosa, and 566 stage I-IV COAD patients [26]. The GSE87211 cohort consisted of 363 patients, 160 rectal mucosa, and 203 rectal tumour patients [27]. Nine hundred forty-eight patients were obtained by merging these two cohorts. Using the same exclusion criteria as the TCGA GIC cohort, 701 CRC patients were included in our validation analysis. The clinical features of the patients are listed in Table S1.

3.2 | The expression of m⁶A RNA methylation regulators was correlative with clinicopathological traits in GIC

Because m⁶A RNA methylation regulators play a significant role in the cancer occurrence and progression, we investigated the expression level of 13 m⁶A RNA methylation regulators in various pathological features of GIC, containing tumour status (normal and tumour), pathological stage (early stage- stage I and II; advanced stage-stage III and IV), and regional lymph node metastasis (N₀ and N₁₋₃). A total of 1099 RNA-seq was obtained; 83 (7.5%) samples were normal and 1016 samples (92.5%) were tumours. 11 m⁶A RNA methylation regulators are abnormally expressed in GIC tissues patients (Figure 2a). Compared with the normal tissue, expressions of METTL3, YTHDF1, YTHDF2, KIAA1429, ZC3H13, WTAP, FTO, HNRNPC, RBM15, and YTHDC1 were up-regulated, while ALKBH5 expression was down-regulated in GIC tissue samples (Figure 2b).

Then, we investigated the relationship between the expression of 13 m⁶A RNA methylation regulators and pathological stage, regional lymph node metastasis in GIC, respectively. Six m⁶A RNA methylation regulators were differently expressed between the advanced and early stages of GIC (Figure 2c). The expressions of ZC3H13, FTO, and RBM15 are up-regulated in the advanced stage, while the expression of METTL3, WTAP, HNRNPC are up-regulated in the early stage (Figure 2d). Similarly, five m⁶A RNA methylation regulators are differentially expressed between the N₀ and N₁₋₃ subgroup of GIC (Figure 2e). ZC3H13 and FTO were highly expressed in the N₁₋₃ subgroup, while METTL3, KIAA1429, and HNRNPC were highly expressed in the N₀ subgroup (Figure 2f).

3.3 | Accordance clustering of m⁶A RNA methylation regulators identified two gastrointestinal cancer clusters

Depending on the expression correspondence of m⁶A RNA methylation regulators, with clustering stability-enhancing from $k = 2$ to 10 in the TCGA dataset, $k = 4$ seemed to be the right choice (Figure 3a,b). However, we noticed that the inter-group correlation is significantly higher in the $k = 4$ group than in the $k = 2$ group (Figure S1, S2). So, we divided the expression of m⁶A RNA methylation regulators into two groups according to $k = 2$ and contrasted the clinicopathological characteristics of the two subgroups by $k = 2$, named cluster 1 and cluster 2, then used Principal component analysis to contrast the transcriptional profile between cluster 1 and cluster 2 subgroups. There are apparent differences between them (Figure 3c). Cluster 2 subgroup contains more lymph node metastasis patients than cluster 1 (Figure 3d, Table S2, $p < 0.001$). Furthermore, we observed a shorter overall survival (OS) in the cluster 2 subgroup than the cluster 1 subgroup ($p < 0.05$, Figure 3e).

3.4 | Prognostic value of m⁶A RNA methylation regulators and construction of the prognosis model

A univariate Cox analysis on the expression of the TCGA dataset was implemented to investigate the prognostic value of m⁶A RNA methylation regulators in GICs. The results demonstrated that nine out of 13 m⁶A RNA methylation-related genes are conspicuously related to OS ($p < 0.05$). FTO and ZC3H13 are high-risk prognosis-related genes (Hazard ratio [HR] > 1, $p < 0.05$), while METTL14, METTL3, WTAP, YTHDC1, YTHDF2, YTHDF1, and HNRNPC are protective prognosis-related genes (HR < 1, $p < 0.05$). The nine prognosis-related m⁶A RNA methylation genes showed in Figure 4a, Table S3.

We utilised the LASSO Cox analysis method to better forecast GICs' clinical outcomes to analyse the nine prognosis-associated m⁶A RNA methylation in the TCGA dataset. Five prognosis-related m⁶A RNA methylation regulators were filtered out through the minimum criteria, and the LASSO algorithm acquired the coefficient (Figure 4b-d). The coefficient of each

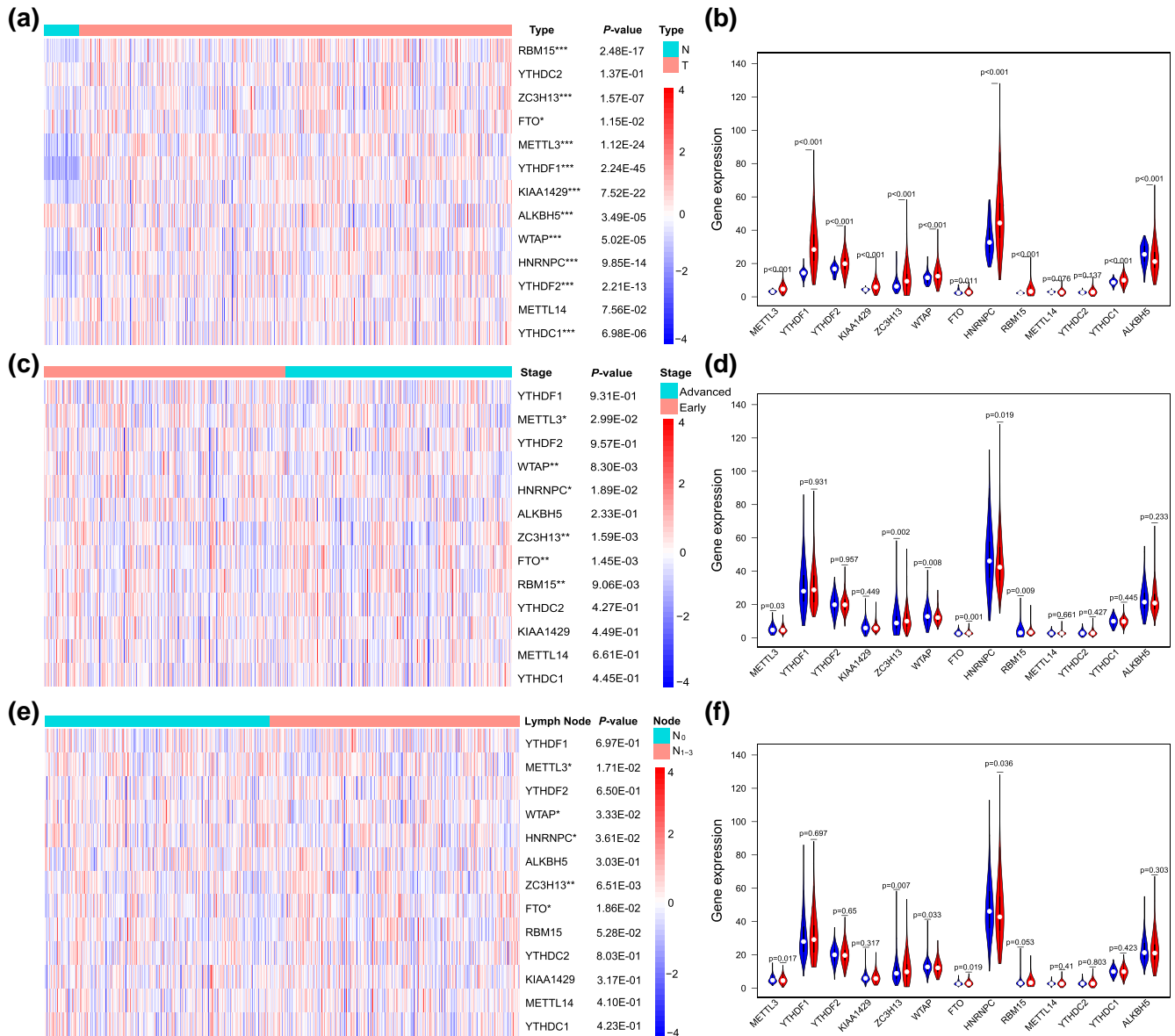


FIGURE 2 Expression of m⁶A RNA methylation regulators in gastrointestinal carcinoma tissues with diverse clinicopathological traits (a, c, e) The heatmaps of 13 m⁶A RNA methylation regulators in various pathological features (a) for normal and tumour status, (c) for the advanced and early stage, and (e) for N₀ and N₁₋₃. (b, d, f) The expression level of 13 m⁶A RNA methylation regulators in diverse pathological features (b) for normal and tumour status, (d) for early and advanced stage, and (f) for N₀ and N₁₋₃. **p* < 0.05, ***p* < 0.01, and ****p* < 0.001. The red fusiformis represents the tumour tissue (b), advanced stage (d), and N₁₋₃ (f), while the blue fusiformis represents the normal tissue (b), early-stage (d), and N₀ (f), respectively

gene in GIC is utilised to compute the risk score of training (TCGA) and validation dataset (GEO). The prognosis model is constructed according to five m⁶A RNA methylation genes, FTO, METTL3, ZC3H13, YTHDF1, and WTAP. The full name, function, and coefficient of these genes are shown in Table S4.

Based on the expression level of the m⁶A RNA methylation regulator and the risk coefficient of every gene, the risk score of each patient is calculated. Risk score = 0.030 × expression of WTAP + (−0.013) × expression of YTHDF1 + 0.012 × expression of ZC3H13 + (−0.103) × expression of METTL3 + 0.148 × expression of FTO. According to the median risk score, the risk score was used to predict the prognosis and divide patients into high-risk and

low-risk subgroups. The OS of the low-risk group is higher than the high-risk group in both the TCGA (Figure 4e) and GEO datasets (Figure 4f, *p* < 0.05).

3.5 | The risk score had a significantly association with the clinicopathological features of GIC

A heatmap was used to show the expression of the five m⁶A RNA methylation related-genes in high- and low-risk patients from the TCGA GICs dataset (Figure 5a). Distributions of the risk score in GIC patients and the relationship between risk

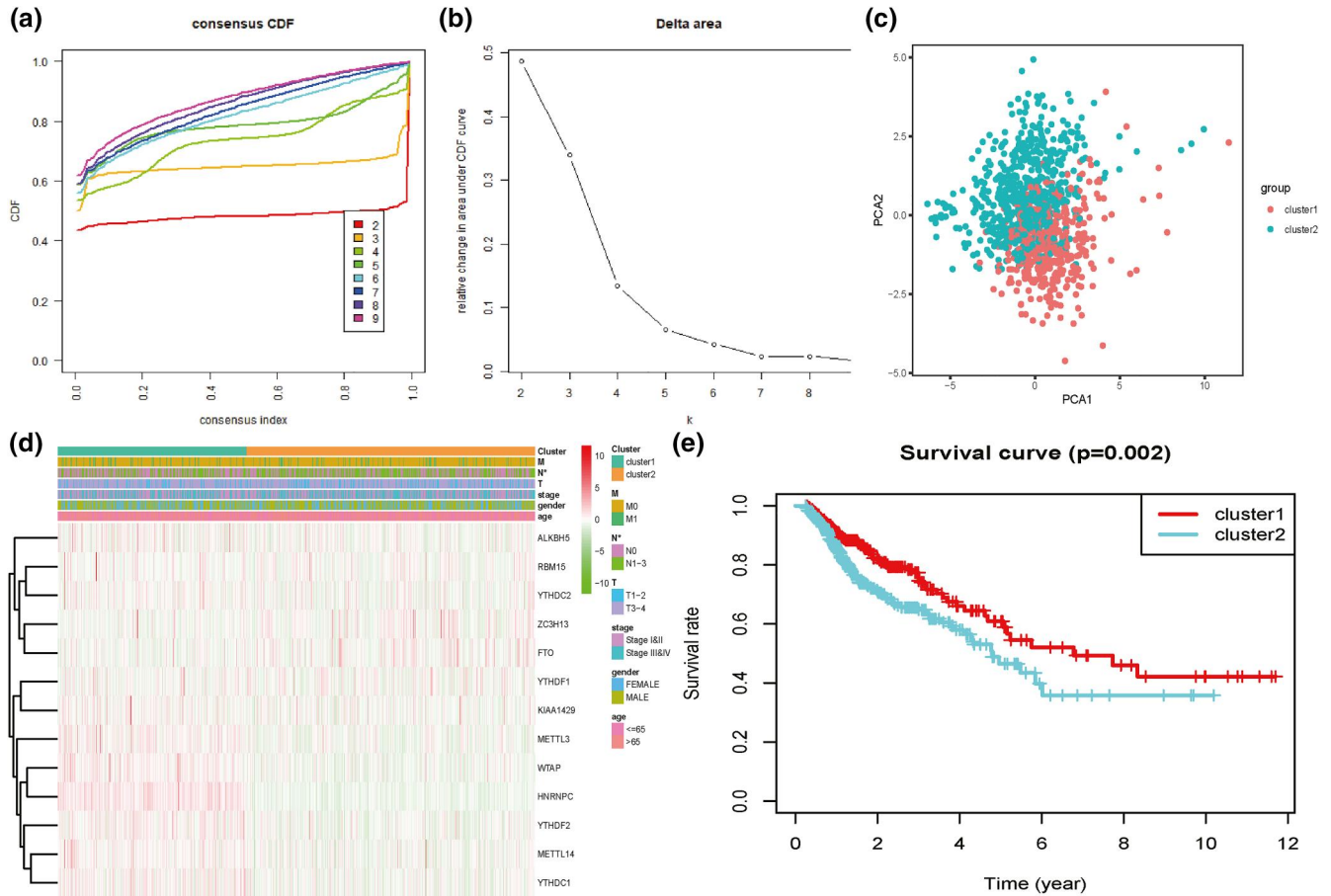


FIGURE 3 Different clinicopathological traits and overall survival (OS) of gastrointestinal cancer (GIC) in the cluster 1/2 subgroups. (a) Consensus clustering cumulative distribution function, CDF for $k = 2-10$. (b) Relative alteration in the area under the CDF curve for $k = 2-10$. (c) Principal component analysis (PCA) of the total RNA expression profile in the the Cancer Genome Atlas (TCGA) dataset. GIC in cluster 1 is marked with red, and cluster 2 is marked with blue. (d) The clinicopathological characteristics of the two clusters (cluster half) are defined through the consensus expression of m^6A RNA methylation regulators. (e) Kaplan–Meier OS curves of TCGA gastrointestinal carcinoma patients in cluster1 and 2. * $p < 0.05$. CDF, cumulative distribution function; PCA, principal component analysis

score and survival time are visualised in Figure 5b. As the risk score enhanced, the survival time of the patients decreased, and death numbers increased. High-risk genes (FTO and ZC3H13) were more probably expressed in the high-risk group, while protective genes (YTHDF1, METTL3, and WTAP) tended to express in the low-risk group (Figure 5a). FTO expression was higher in the advanced stage group and N_{1-3} (Figure 5c, 5d, $p < 0.001$). Similarly, ZC3H13 showed a high expression in the advanced stage group, T_{3-4} , and N_{1-3} (Figure 5e–g, $p < 0.05$). On the contrary, WTAP was high expressed in the early stage, and N_0 (Figure 5j, 5k, $p < 0.05$) METTL3 expression was also higher in the N_0 subgroup (Figure 5l, $p < 0.001$). The risk score was higher in the advanced stage group and N_{1-3} (Figure 5h, 5i, $p < 0.05$). 1-year, 3-year, and 5-year AUC value of ROC for risk score was 0.668, 0.658, and 0.687, respectively; prognostic accuracy of N-stage and pathological stage was higher than other clinical characters (Figure 5m–o). These results indicate that the risk scores and FTO and ZC3H13 are closely related to the malignant clinicopathological features and can accurately predict the GICs patients' outcomes.

Furthermore, univariate and multivariate Cox analyses were performed on the TCGA dataset to identify whether the risk

score was an independent prognostic factor. Then we validated in the GEO dataset. The older patients, N_{1-3} , M_1 , and high risk-scores were associated with the poor OS both in the univariate and multivariate independent prognostic Cox analysis ($p < 0.01$, Figure 5p, 5q, Table 1). Similar results are discovered in the GEO dataset (Figure 5r, Table 2). Overall, these were found to validate the risk score, based on five m^6A RNA methylation regulators, and could independently predict prognosis and closely associated with t GIC patients' malignancy.

3.6 | The prognostic prediction models

Nomogram is an effective tool that has been used to quantitatively identify the individuals' risk in the clinical setting through combining multiple risk elements [32]. We designed a nomogram to predict 1-, 3- and 5-year OS probability by combining five m^6A RNA methylation regulators' signatures (Figure 6a). We computed every patient's total point by summing up the number of points for five m^6A RNA methylation regulators, which might aid to appropriate practitioners to make clinical decisions for GIC patients. The calibration curve demonstrated that the actual

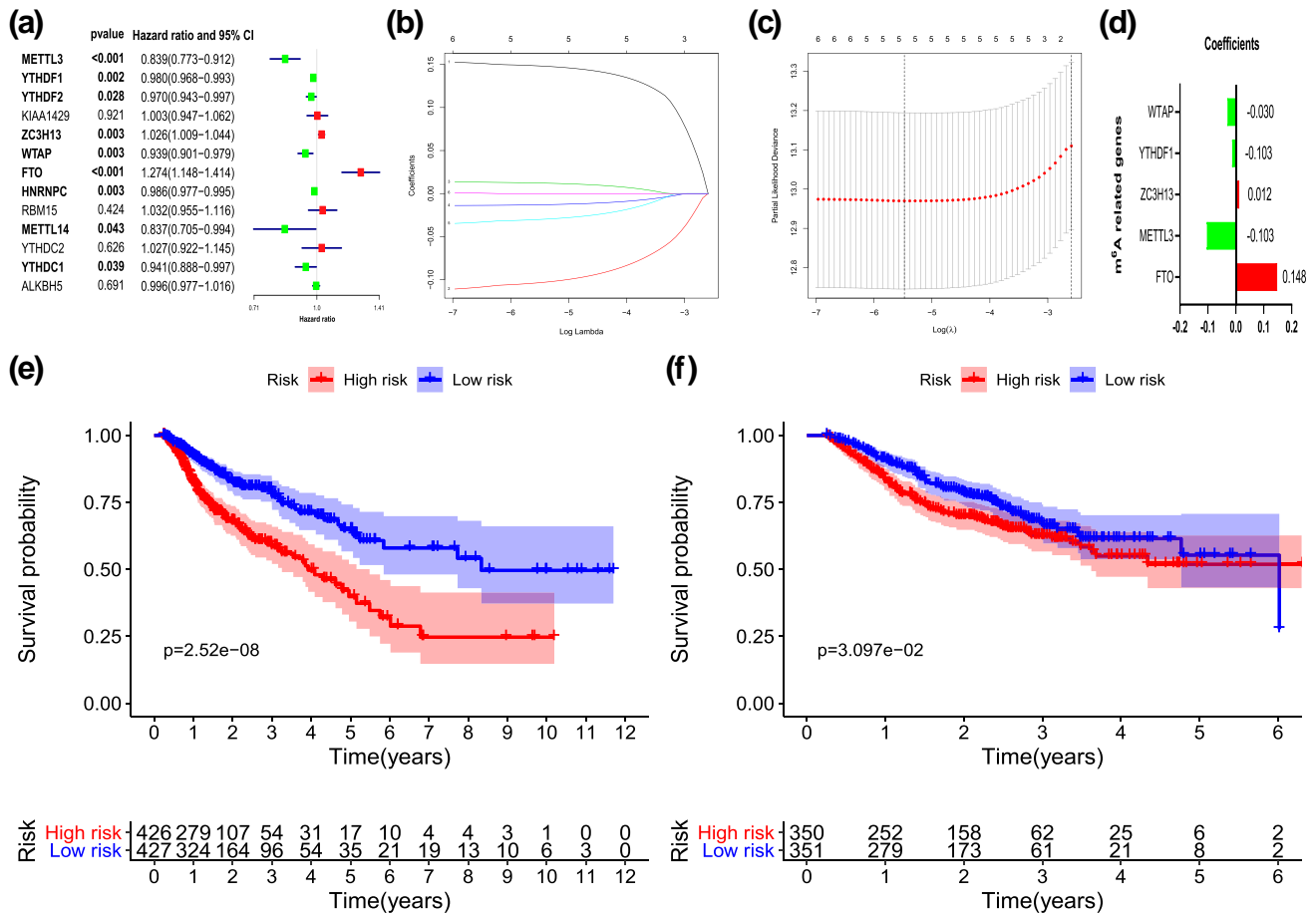


FIGURE 4 Risk model with five m⁶A RNA methylation regulators and survival analysis. (a)-(d) The screening process of the five m⁶A RNA methylation regulators. (a) Hazard ratio (HR) 95% confidence intervals, CI of 13 m⁶A RNA methylation regulators, are calculated by univariate Cox regression analysis. Nine prognosis-related m⁶A RNA methylation genes are manifested in bold black. HR > 1 is marked as a red box, and HR < 1 is marked as a green box. (b), (c) The process of calculating the coefficient by the least absolute shrinkage and selection operator (LASSO) Cox regression. (d) The coefficient of five m⁶A RNA methylation regulators (e), (f) Kaplan–Meier overall survival (OS) curves for patients in the the Cancer Genome Atlas (TCGA) (e) and Gene Expression Omnibus (GEO). (f) datasets are assigned to high- and low-risk groups according to the median risk score value. HR, Hazard ratios; CI, confidence intervals; TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; LASSO, least absolute shrinkage, and selection operator; OS, overall survival

survival rates in the TCGA data set matched well with the predicted 1-, 3- and 5- year survival rates; the C-index is 0.728 (Figure 6b, 6d, 6f). The nomogram is also verified in the GEO cohorts, the C-index is 0.734, and calibration curves of 1-, 3- and 5- years are shown in Figure 6c, 6e, 6g.

3.7 | Correlation and function enrichment analysis of m⁶A RNA methylation regulators

We also analysed these regulators' relationship to better appreciate the 13 m⁶A RNA methylation regulators (Figure 7a). The expressions of YTHDC1, YTHDF1, and YTHDF2 are significantly associated with each other. HNRNPC had a significant correlation with YTHDC1, YTHDC2, and YTHDF2 in GIC. WTAP expression was also prominently correlated with the 'writers' of KIAA1429, RBM15, METTL14, and METTL3 in GIC. Moreover, the expressions of FTO and ZC3H13 are positively related to each other; FTO was not correlated with WTAP, YTHDF1, YTHDF2, and negatively

correlated with METTL3 and HNRNPC; ZC3H13 was not correlated with WTAP, METTL3, and negatively correlated with HNRNPC. These findings were consistent with that of the FTO and ZC3H1 which were risky genes, while the WTAP, YTHDF1, YTHDF2, METTL3, and HNRNPC were the protective genes.

We further used functional GO enrichment analysis to understand the previous five m⁶A RNA methylation regulators' pathways. A total of 80 GO terms of BP are identified to be significant (FDR < 0.05). These genes are categorised into several essential processes, mainly containing RNA modification, mRNA methylation, RNA methylation, and mRNA metabolic process regulation. Top-20 GO terms of BP are shown in Figure 7b. GSEA analysis was performed that altered the genes associated with several common KEGG pathways, including focal adhesion (NES = 2.17, $p = 0.002$), Extracellular matrix receptor interaction (NES = 2.14, $p = 0.002$), regulation of actin cytoskeleton (NES = 2.13, $p = 0.000$), and dilated cardiomyopathy (NES = 2.02, $p = 0.004$; Figure 7c). Moreover, GSEA revealed that the malignant hallmarks of tumours,

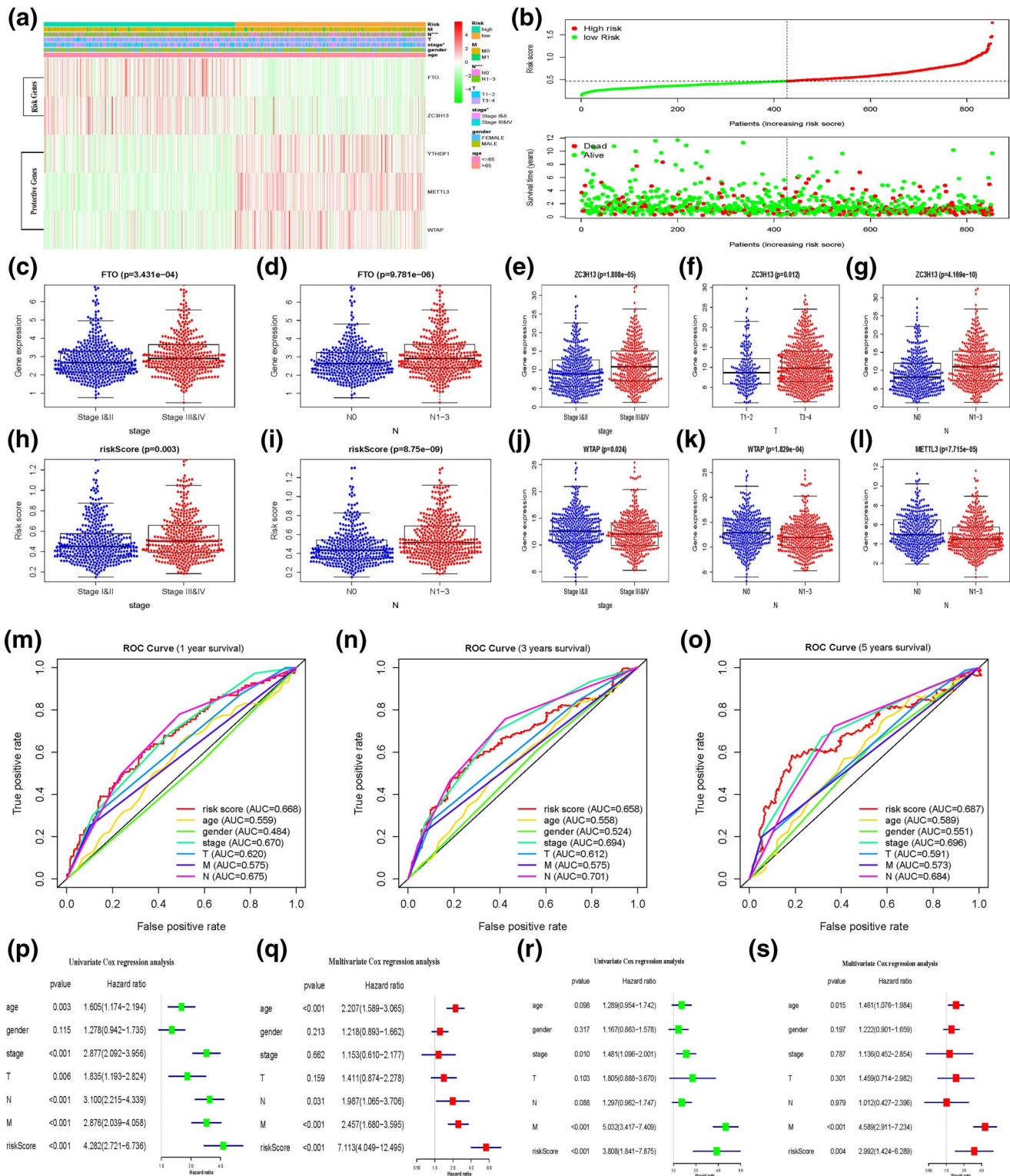


FIGURE 5 Association between the risk score and clinicopathological characteristics in gastrointestinal carcinoma. (a) The heatmap manifested the expression levels of the five m^6A RNA methylation regulators in high and low-risk groups. The distribution of clinicopathological characteristics is compared between the high- and low-risk groups, $*p < 0.05$, $***p < 0.001$. (b) The distribution of the risk score and survival time of the patient, as well as the status of the gastrointestinal cancer (GIC). The black dotted line was the optimum cutoff dividing patients into high-risk and low-risk groups. (c–l) The association between the clinicopathological characteristics and expression of m^6A RNA methylation regulators. The expression level of FTO in different Stages (c) and N subgroups (d). The ZC3H13 expression in distinct Stages (e), T (f), and N subgroup (g). The risk score in different stages (h) and N subgroups (i). The expression level of WTAP in diverse Stages (j) and N subgroups (k). (l) The METTL3 expression in different N subgroups. (m–o) Receiver Operating Characteristic (ROC) analysis of the specificity and sensitivity of the overall survival (OS) for the combination of the risk score and clinical features in the the Cancer Genome Atlas (TCGA) gastrointestinal carcinoma. (p) Univariate and (q) Multivariate independent prognostic Cox regression analysis in the TCGA database. (r) Univariate and (s) Multivariate independent prognostic Cox regression analysis in the Gene Expression Omnibus (GEO) database. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus

TABLE 1 The prognostic value of different clinical characters in the training group

Variables	Univariate prognostic analysis			Multivariate prognostic analysis		
	HR	95% CI	<i>p</i> -value	HR	95% CI	<i>p</i> -value
Age (≤ 65 years VS > 65 years)	1.605	1.174–2.194	2.99E–4	2.207	1.589–3.065	2.35E–06
Gender (female vs. male)	1.280	0.942–1.735	0.115	1.218	0.893–1.662	0.213
Stage (I/II vs. III/IV)	2.877	2.092–3.956	7.87E–11	1.153	0.610–2.177	0.662
T (1–2 vs. 3–4)	1.835	1.193–2.824	0.006	1.411	0.873–2.278	0.159
N (0 vs. 1–3)	3.100	2.215–4.339	4.27E–11	1.987	1.065–3.706	0.031
M (0 vs. 1)	2.876	2.039–4.058	1.79E–09	2.457	1.680–3.595	3.64E–06
Risk score	4.282	2.721–6.736	3.18E–10	7.113	4.049–12.495	8.80E–12

Abbreviation: HR, hazard ratio.

TABLE 2 The prognostic value of different clinical characters in the validation group

Variables	Univariate prognostic analysis			Multivariate prognostic analysis		
	HR	95% CI	<i>p</i> -value	HR	95% CI	<i>p</i> -value
Age (≤ 65 years VS > 65 years)	1.289	0.954–1.742	0.098	1.461	1.076–1.984	0.015
Gender (female vs. male)	1.167	0.863–1.578	0.317	1.223	0.901–1.659	0.197
Stage (I/II vs. III/IV)	1.481	1.096–2.001	0.011	1.136	0.452–2.854	0.787
T (1–2 vs. 3–4)	1.805	0.888–3.670	0.103	1.459	0.714–2.982	0.301
N (0 vs. 1–3)	1.297	0.962–1.747	0.088	1.012	0.427–2.396	0.979
M (0 vs. 1)	5.031	3.417–7.409	2.77E–16	4.589	2.911–7.234	5.38E–11
Risk score	3.808	1.841–7.875	3.10 E–04	2.992	1.424–6.289	0.004

Abbreviation: HR, hazard ratio.

containing oxidative phosphorylation (NES = 1.89, $p = 0.019$), glycolysis (NES = 1.68, $p = 0.013$), fatty acid metabolism (NES = 1.68, $p = 0.038$), and cholesterol homeostasis (NES = 1.62, $p = 0.039$) were significantly related to the five m⁶A RNA methylation regulators (Figure 7d).

4 | DISCUSSION

Conventional epigenetics concentrated on histone modifications, nucleosome remodelling, DNA methylation, and non-coding RNAs has diverse functions in early cancer detection, prediction, progression, prognosis, and response to treatment [33, 34]. In this study, we illustrated that the expression of m⁶A RNA methylation regulators, another field of epigenetics, is also closely related to the prognosis and malignancy of GIC. Firstly, 10 of 13 m⁶A RNA methylation regulators are substantially up-regulated in the GIC patients. We discovered that ZC3H13 and FTO were both positively associated with the group of advanced stage and N₁₋₃. Then, we identified that the cluster 1/2 subgroups influenced the prognosis and clinicopathological features of GIC. We also designed a prognostic risk model according to five m⁶A RNA methylation regulators and classified the GIC patients into high- and low-risk subgroups. We further identified risk scores and high-risk genes (FTO and ZC3H13) closely associated with the malignant clinicopathological feature and can accurately predict the GICs

patients' outcomes. The five m⁶A RNA methylation regulators can predict the probability of 1-, 3- and 5-year OS through a nomogram and are approximately associated with BP, hub signalling pathways, and GIC's malignant hallmarks.

This paragraph thoroughly analysed the expression of 13 m⁶A RNA methylation regulators in GIC with different clinicopathological traits. Previous research has reported that the eraser FTO was influential in promoting gastric cancer and FTO polymorphisms associated with colorectal adenomas in African-Americans [35, 36]; ALKBH5 was down-regulated in colorectal cancer FTO which positively correlated with poor prognosis in colorectal cancer patients [39]. ALKBH5 altered cell adhesion of bladder cancer cells by regulating the ITGA6 expression [37]. FTO and ALKBH5 showed demethylase activity towards N (6)-methyladenosine in RNA [9]; these results indicate that ALKBH5 and FTO may adjust demethylation of disparate methylation targets in GIC, and it deserves to be explored in future research studies. The expression of ZC3H13 and RBM15 was significantly increased in the advanced stage, indicating potential functions of ZC3H13 and RBM15 in GIC malignancy. METTL3 promotes the translation of many subgroups of oncogenic mRNAs and acts as an oncogene, sustains SOX2 expression of CRC cells by an m⁶A-IGF2BP2-dependent mechanism, and facilitates tumour progression [21, 38]. The previous report has found that carbonic anhydrase IV targeting the WTAP-WT1-TBL1 axis to inhibit Wnt signalling pathway [39] and ZC3H13 may suppress invasion

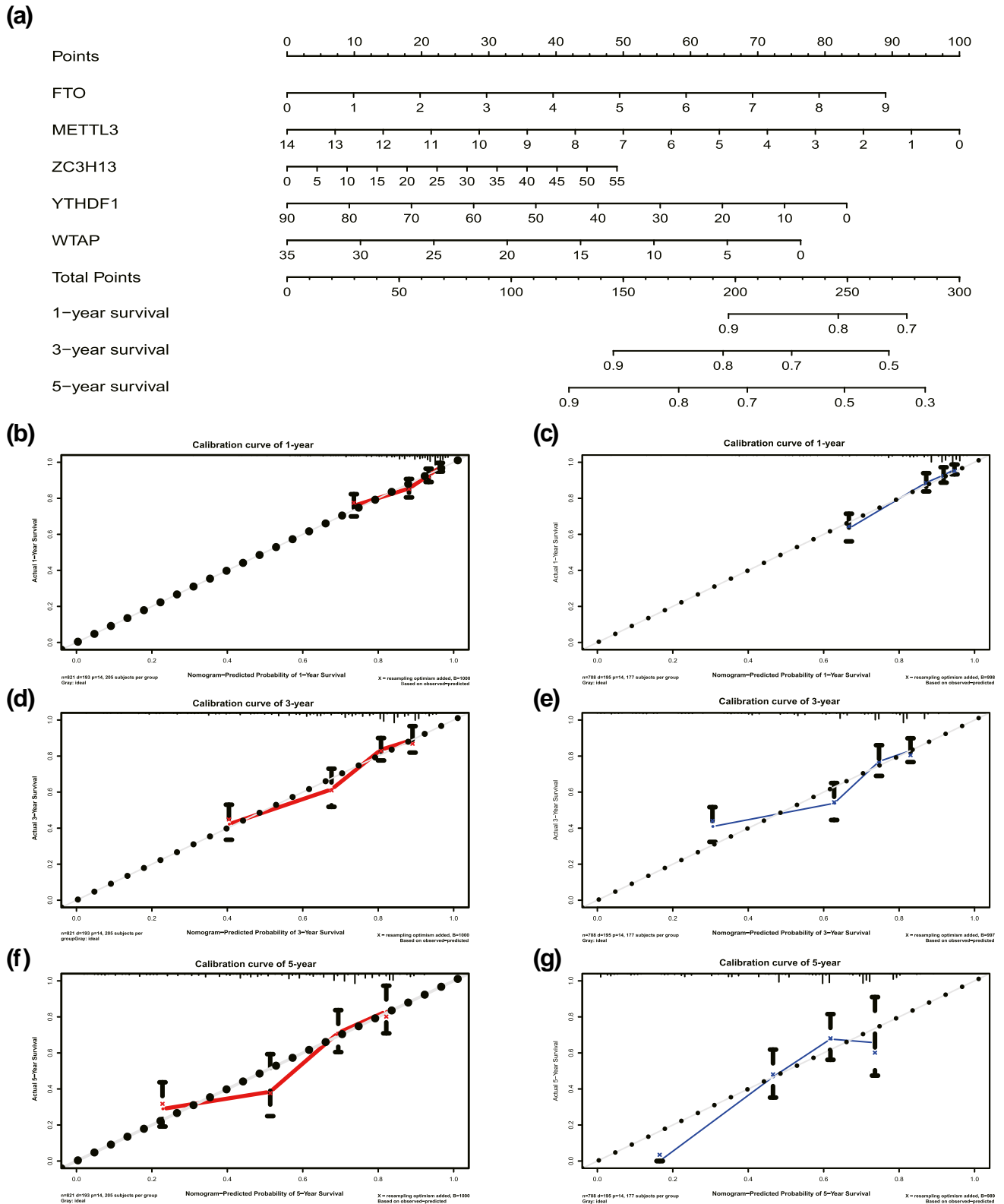


FIGURE 6 The nomogram to anticipate prognostic possibilities in gastrointestinal carcinoma. (a) The nomogram for predicting 1-, 3- and 5-year overall survival (OS) of gastrointestinal carcinoma by expression of five m⁶A RNA methylation regulators, (b–g) Calibration curves assess consistency between actual and predicted OS for 1, 3, and 5 years. (b, d, f) The Cancer Genome Atlas (TCGA) dataset, (c, e, g) Gene Expression Omnibus (GEO) dataset. TCGA, The Cancer Genome Atlas; GEO, Gene Expression Omnibus; OS, overall survival

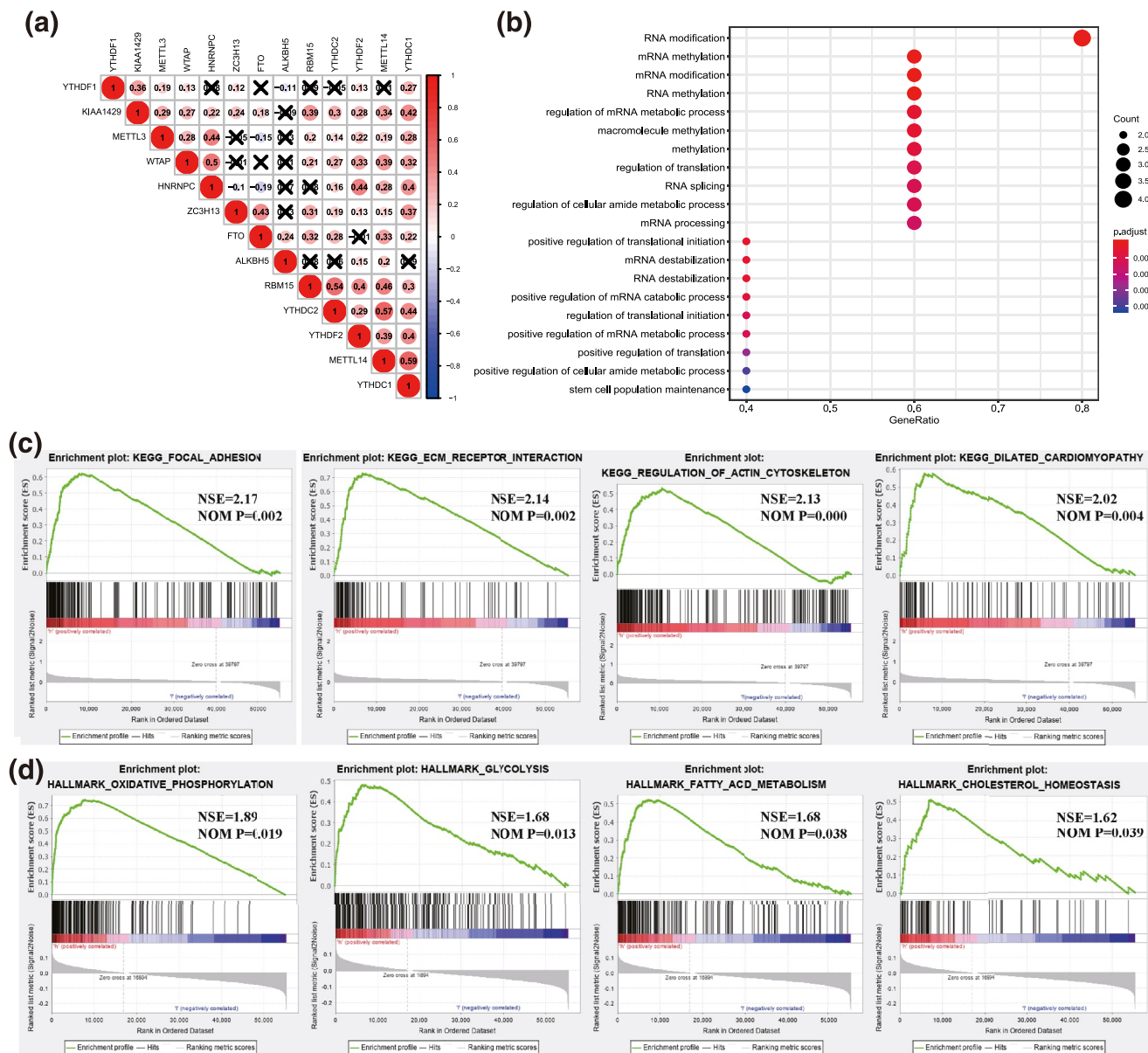


FIGURE 7 Correlation and functional enrichment analysis of m⁶A RNA methylation regulators in gastrointestinal carcinoma. (a) Spearman correlation analysis of the 13 m⁶A modification regulators. Colour dots represent the Pearson correlation coefficient. (b) Functional annotation of the genes using the Gene Ontology (GO) terms of Biological Processes (BP). (c) Gene Set Enrichment Analysis (GSEA) analysis of the genes in high-risk groups for the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway. (d) GSEA analysis of the genes in high-risk groups for hallmarks of malignant tumours. GO, Gene Ontology; GSEA, Gene Set Enrichment Analysis

and proliferation of CRC inactivating regulator of the Ras-ERK signalling pathway [40]. RBM15 mutations also may be crucial in the pathogenesis of borderline/malignant phyllodes tumours [41]. These findings are also advantageous to investigate novel therapeutic methods of GIC by distinguishing the expression of each m⁶A methylation regulator.

In the present research, the prognostic model had a moderate prognostic value in GIC. The patients in the high-risk group had an obviously shorter survival time than patients in the low-risk group. Expression of FTO and ZC3H13 increased in GIC with malignant clinicopathological features, such as advanced stage, sizeable primary tumour, and lymph

node metastasis. The risk score was also obviously related to the malignant clinicopathological features in GIC. These results were consistent with the previous studies [37]. 1-year, 3-year, and 5-year AUC values of ROC for the risk score are consistent with a previous study that the AUC of the ROC curve is 0.68 for the ability of YKL-40 to predict GIC [42]. Both univariate and multivariate COX regression analyses confirmed that malignant clinicopathological traits were related to poor prognosis in GIC patients from TCGA and GEO databases.

Subsequently, we established a nomogram to forecast the individual patient's clinical outcomes. The nomogram was a

trustworthy tool to quantitatively measure individual risk by merging and illustrating risk elements utilised in cancer prognosis, including CRC [43]. In addition to traditional clinicopathological traits (e.g., TNM (T: Tumor; N: Node; M: Metastasis) stage and histological subtype), gene biomarkers can also be combined with a predictive nomogram model to predict clinical outcomes [44, 45]. In the present research, we indicated that a nomogram containing five m⁶A RNA methylation regulators could well predict 1-, 3- and 5-year GIC patients' survival possibilities.

Finally, some BP and pathways related to the occurrence and progression of GIC were confirmed. It was known that RNA modification participated in the regulation Phosphatidylinositol 3-kinase/Protein kinase B/mammalian target of rapamycin (PI3K/AKT/mTOR) and ErbB Pathways in GIC [46], mRNA modification, mRNA methylation, RNA methylation, and regulation of mRNA metabolic process to promote the proliferation of cancer [47, 48]. Focal adhesion has been reported to involve the proliferation, differentiation, and apoptosis of cells related to colorectal carcinogenesis [49]. Previous research showed an essential requirement for oxidative phosphorylation in tumour progression [50]. More and more evidence suggested that glycolysis, fatty acid metabolism, and cholesterol homeostasis played an essential role in cancer development and cancer cell proliferation [51–53]. Our study has consistently shown that the five m⁶A RNA methylation regulators are also associated with these processes, signalling pathways, and tumours—hallmarks manifesting that they are associated with the malignancy of GIC.

This research discovered that the prognostic model obtained according to the five m⁶A RNA methylation regulators had conspicuous prognostic value in GIC patients. Some improvements to the model's programing aspect can be brought through additional hierarchy levels to represent arranged activities in more detail. However, there are also several deficiencies to this study. First, the present study is purely analysing big data by the bioinformatics method, and further experimental and clinical research might be necessary about these m⁶A RNA methylation regulators in the clinical application. Second, our study patients are mainly foreigners, which may lead to a potential risk of selection bias. Some of the antecedently overlooked m⁶A RNA methylation regulators might have the opportunity to become additional biomarkers for GIC. Finally, this research enhanced our cognition about the sophisticated reciprocities of GIC and might detect novel therapeutic targets.

5 | CONCLUSION

Our research comprehensively clarified the expression, molecular function, and prognostic value of five m⁶A RNA methylation regulators in GIC. The expression of the m⁶A-RNA methylation regulatory factor was significantly associated with the clinicopathological traits of GIC, and it is also significantly associated with the biological process of promoting the malignant progress of GIC and the increase of gene expression level in the signal pathway. Our research also

highlights the essential role of the risk score as an independent prognostic marker for GIC. These findings provide an extensive viewpoint for further research on the role of m⁶A RNA methylation regulators in the GIC's pathogenesis and as potential markers for GIC diagnosis and treatment.

ACKNOWLEDGMENTS

We thank the public databases for data availability. We are grateful to the School of Foreign Languages professors at Lanzhou University for their help in the linguistic embellishment of this manuscript. The authors would like to express their gratitude to those who contributed to this study. The authors received no funding for this work.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT


The authors certify that all the original data in this research could be obtained from a public database. Both the RNA-sequencing (RNA-seq) and clinicopathology of the gastrointestinal tumor, including STAD, COAD, rectosigmoid junction cancer, and READ, were obtained from the TCGA database (Data release 22.0 - 16 January 2020, <https://tcga-data.nci.nih.gov/tcga/>). The expression of mRNA extracted from the matrix file was obtained from the RNA-seq data. The data access GSE39582 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39582>) and GSE87211, (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE87211>), were obtained from GEO (GEO, <https://www.ncbi.nlm.nih.gov/geo/>). Other data used to support the findings of this study are included within the supplementary information files.

ORCID

Yandong Miao  <https://orcid.org/0000-0002-1429-8915>

Bin Su  <https://orcid.org/0000-0003-3626-166X>

Xiaolong Tang  <https://orcid.org/0000-0001-9229-6424>

Jiangtao Wang  <https://orcid.org/0000-0002-1222-164X>

Wuxia Quan  <https://orcid.org/0000-0002-1898-850X>

Yonggang Chen  <https://orcid.org/0000-0002-9017-9416>

Denghai Mi  <https://orcid.org/0000-0002-8643-4496>

REFERENCES

1. Dawson, M.A., Kouzarides, T.: Cancer epigenetics: from mechanism to therapy. *Cell*. 150(1), 12–27 (2012)
2. Stark, R., Grzelak, M., Hadfield, J.: RNA sequencing: the teenage years. *Nat. Rev. Genet.* 20(11), 631–656 (2019)
3. Kirchner, S., Ignatova, Z.: Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat. Rev. Genet.* 16(2), 98–112 (2015)
4. Dragomir, M.P., Knutsen, E., Calin, G.A.: SnapShot: unconventional miRNA functions. *Cell*. 174(4), 1038 (2018)
5. Smith, E.R., et al.: The little elongation complex regulates small nuclear RNA transcription. *Mol. Cell*. 44(6), 954–965 (2011)
6. Zhang, Q., Shalaby, N.A., Buszczak, M.: Changes in rRNA transcription influence proliferation and cell fate within a stem cell lineage. *Science*. 343(6168), 298–301 (2014)
7. Zaccara, S., Ries, R.J., Jaffrey, S.R.: Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* 20(10), 608–624 (2019)

8. Daffis, S., et al.: 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature*. 468(7322), 452–456 (2010)
9. Shen, L., et al.: Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu. Rev. Biochem.* 83, 585–614 (2014)
10. Mauer, J., Jaffrey, S.R.: FTO, m(6) Am, and the hypothesis of reversible epitranscriptomic mRNA modifications. *FEBS Lett.* 592(12), 2012–2022 (2018)
11. Vu, L.P., Cheng, Y., Kharas, M.G.: The biology of m(6)A RNA methylation in normal and malignant hematopoiesis. *Cancer. Discov.* 9(1), 25–33 (2019)
12. Chen, X.Y., Zhang, J., Zhu, J.S.: The role of m(6)A RNA methylation in human cancer. *Mol. Cancer.* 18(1), 103 (2019)
13. Yang, Y., et al.: Dynamic transcriptomic m(6)A decoration: writers, erasers, readers and functions in RNA metabolism. *Cell. Res.* 28(6), 616–624 (2018)
14. Liu, T., et al.: The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. *Nucleic Acids Res* (2020)
15. Kwok, C.T., et al.: Genetic alterations of m(6)A regulators predict poorer survival in acute myeloid leukemia. *J. Hematol. Oncol.* 10(1), 39 (2017)
16. Wang, J., et al.: Effect of m(6)A RNA methylation regulators on malignant progression and prognosis in renal clear cell carcinoma. *Front. Oncol.* 10, 3 (2020)
17. Chai, R.C., et al.: m(6)A RNA methylation regulators contribute to malignant progression and have clinical prognostic impact in gliomas. *Aging.* 11(4), 1204–1225 (2019)
18. Mallm, J.P., et al.: Glioblastoma initiating cells are sensitive to histone demethylase inhibition due to epigenetic deregulation. *Int. J. Cancer.* 146(5), 1281–1292 (2020)
19. Hu, B.B., et al.: N(6)-methyladenosine (m(6)A) RNA modification in gastrointestinal tract cancers: roles, mechanisms, and applications. *Mol. Cancer.* 18(1), 178 (2019)
20. Wang, Q., et al.: METTL3-mediated m(6)A modification of HDGF mRNA promotes gastric cancer progression and has prognostic significance. *Gut* (2019)
21. Li, T., et al.: METTL3 facilitates tumor progression via an m(6)A-IGF2BP2-dependent mechanism in colorectal carcinoma. *Mol. Cancer.* 18(1), 112 (2019)
22. Nishizawa, Y., et al.: Oncogene c-Myc promotes epitranscriptome m(6)A reader YTHDF1 expression in colorectal cancer. *Oncotarget.* 9(7), 7476–7486 (2018)
23. Bai, Y., et al.: YTHDF1 regulates tumorigenicity and cancer stem cell-like activity in human colorectal carcinoma. *Front. Oncol.* 9, 332 (2019)
24. Xu, L.C., Pan, J.X., Pan, H.D.: Construction and validation of an m6A RNA methylation regulators-based prognostic signature for esophageal cancer. *Cancer. Manag. Res.* 12, 5385–5394 (2020 Jul 6)
25. Wu, X., et al.: Prognostic value of an m6A RNA methylation regulator-based signature in patients with hepatocellular carcinoma. *BioMed Res. Int.* 2020, 2053902 (2020)
26. Marisa, L., et al.: Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS. Med.* 10(5), e1001453 (2013)
27. Hu, Y., et al.: Colorectal cancer susceptibility loci as predictive markers of rectal cancer prognosis after surgery. *Genes. Chromosom. Cancer.* 57(3), 140–149 (2018)
28. Ritchie, M.E., et al.: Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic. Acids. Res.* 43(7), e47 (2015)
29. Zhang, S., et al.: DNA methylation exploration for ARDS: a multi-omics and multi-microarray interrelated analysis. *J. Transl. Med.* 17(1), 345 (2019)
30. Liu, Y., et al.: Prognostic implications of autophagy-associated gene signatures in non-small cell lung cancer. *Aging.* 11 (2019)
31. Miao, Y., et al.: Prognostic implications of metabolism-associated gene signatures in colorectal cancer. *PeerJ.* 8, e9847 (2020)
32. Won, Y.W., et al.: A nomogram to predict brain metastasis as the first relapse in curatively resected non-small cell lung cancer patients. *Lung. cancer.* 88(2), 201–207 (2015)
33. Wong, C.C., et al.: Epigenomic biomarkers for prognostication and diagnosis of gastrointestinal cancers. *Semin. Cancer. Biol.* 55, 90–105 (2019)
34. Selaru, F.M., et al.: Epigenetic events in gastrointestinal cancer. *Am. J. Gastroenterol.* 104(8), 1910–1912 (2009)
35. Xu, D., et al.: FTO expression is associated with the occurrence of gastric cancer and prognosis. *Oncol. Rep.* 38(4), 2285–2292 (2017)
36. Nock, N.L., et al.: FTO polymorphisms are associated with adult body mass index (BMI) and colorectal adenomas in African-Americans. *Carcinogenesis.* 32(5), 748–756 (2011)
37. Liu, X., et al.: Expression patterns and prognostic value of m(6)A-related genes in colorectal cancer. *Am. J. Transl. Res.* 11(7), 3972–91 (2019)
38. Choe, J., et al.: mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature.* 561(7724), 556–560 (2018)
39. Zhang, J., et al.: Carbonic anhydrase IV inhibits colon cancer development by inhibiting the Wnt signalling pathway through targeting the WTAP-WT1-TBL1 axis. *Gut.* 65(9), 1482–1493 (2016)
40. Zhu, D., et al.: ZC3H13 suppresses colorectal cancer proliferation and invasion via inactivating Ras-ERK signaling. *J. Cell. Physiol.* 234(6), 8899–8907 (2019)
41. Garcia-Dios, D.A., et al.: MED12, TERT promoter and RBM15 mutations in primary and recurrent phyllodes tumours. *Br. J. Cancer.* 118(2), 277–284 (2018)
42. Allin, K.H., et al.: Cancer risk by combined levels of YKL-40 and C-reactive protein in the general population. *Br. J. Cancer.* 106(1), 199–205 (2012)
43. Renfro, L.A., et al.: Clinical Calculator for Early Mortality in Metastatic Colorectal Cancer: an analysis of patients From 28 clinical trials in the Aide et Recherche en Cancerologie Digestive Database. *J. Clin. Oncol.* 35(17), 1929–1937 (2017)
44. Reichling, C., et al.: Artificial intelligence-guided tissue analysis combined with immune infiltrate assessment predicts stage III colon cancer outcomes in PETACC08 study. *Gut* (2019)
45. Wang, J., et al.: Construction prognosis model based on autophagy-related gene signatures in hepatocellular carcinoma. *Biomarkers. Med.* 14(13), 1229–1242 (2020)
46. Zhao, Y., et al.: m1A regulated genes modulate PI3K/AKT/mTOR and ErbB pathways in gastrointestinal cancer. *Trans. oncol.* 12(10), 1323–1333 (2019)
47. Liu, L., et al.: m(6)A mRNA methylation regulates CTNBN1 to promote the proliferation of hepatoblastoma. *Mol. Cancer.* 18(1), 188 (2019)
48. Yu, S., et al.: N(6)-Methyladenosine: a novel RNA imprint in human cancer. *Front. Oncol.* 9, 1407 (2019)
49. Murata, T., et al.: Localization of FAK is related with colorectal carcinogenesis. *Int. J. Oncol.* 32(4), 791–6 (2008)
50. Maiuri, M.C., Kroemer, G.: Essential role for oxidative phosphorylation in cancer progression. *Cell. Metabol.* 21(1), 11–12 (2015)
51. Altman, B.J., Stine, Z.E., Dang, C.V.: From Krebs to clinic: glutamine metabolism to cancer therapy. *Nat. Rev. Cancer.* 16(10), 619–634 (2016)
52. Currie, E., et al.: Cellular fatty acid metabolism and cancer. *Cell. Metabol.* 18(2), 153–161 (2013)
53. Kuzu, O.F., Noory, M.A., Robertson, G.P.: The role of cholesterol in cancer. *Cancer. Res.* 76(8), 2063–2070 (2016)

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Miao, Y., et al.: Construction and validation of m⁶A RNA methylation regulators associated prognostic model for gastrointestinal cancer. *IET Syst. Biol.* 16(2), 59–71 (2022). <https://doi.org/10.1049/syb2.12040>