A prognostic model for STAD based on hypoxia- and immune-related genes

Keywords
prognostic factor, immunotherapy, hypoxia, stomach adenocarcinoma, immune microenvironment

Abstract
Introduction
Featured with vast heterogeneity, gastric cancer (GC) is one of the leading causes of cancer-related deaths. A specific prognostic model is necessary for the improvement of clinical treatment strategies. Hypoxia is a common feature in the tumor microenvironment that promotes tumor progression. However, the current evaluation of hypoxic tumor immune microenvironment in GC is still inadequate.

Material and methods
With sequence data and single nucleotide variants data obtained from The Cancer Genome Atlas-STAD dataset as well as hypoxia- and immune-related genes acquired from MsigDB and ImmPort, a hypoxia-immune-based gene signature of gastric adenocarcinoma (STAD) was built by Cox regression analysis. Riskscore could be used as an independent prognostic factor.

Results
Receiver operating characteristic curve and survival curve showed the accuracy of the model. Pearson correlation analysis showed that DUSP1, one of the hypoxia- and immune-related feature genes, was positively correlated with immune cell scores and immune-related function scores. In addition, low-risk group peers were found to be in higher immune infiltration status and had higher immunophenoscore as demonstrated by single-sample GSEA, indicating better response to immune checkpoint inhibitors (ICI) treatment among the low-risk group. q-PCR results showed that DUSP1, IGFBP1, CGB5, GPC3 and EGF were significantly highly expressed in STAD cells, while FAM3D and FGF8 were significantly down-regulated.

Conclusions
Overall, our study not only paves the way for future studies focusing on hypoxia and immune microenvironment but also improves STAD patient’s prognosis and their response to immunotherapy.
A prognostic model for STAD based on hypoxia- and immune-related genes

Running title: The construction of a STAD prognostic model

Abstract
Featured with vast heterogeneity, gastric cancer (GC) is one of the leading causes of cancer-related deaths. A specific prognostic model is necessary for the improvement of clinical treatment strategies. Hypoxia is a common feature in the tumor microenvironment that promotes tumor progression. However, the current evaluation of hypoxic tumor immune microenvironment in GC is still inadequate. With sequence data and single nucleotide variants data obtained from The Cancer Genome Atlas-STAD dataset as well as hypoxia- and immune-related genes acquired from MsigDB and ImmPort, a hypoxia-immune-based gene signature of gastric adenocarcinoma (STAD) was built by Cox regression analysis. Riskscore could be used as an independent prognostic factor. Receiver operating characteristic curve and survival curve showed the accuracy of the model. Pearson correlation analysis showed that DUSP1, one of the hypoxia- and immune-related feature genes, was positively correlated with immune cell scores and immune-related function scores. In addition, low-risk group peers were found to be in higher immune infiltration status and had higher immunophenoscore as demonstrated by single-sample GSEA, indicating better response to immune checkpoint inhibitors (ICI) treatment among the low-risk group. q-PCR results showed that DUSP1, IGFBP1, CGB5, GPC3 and EGF were significantly highly expressed in STAD cells, while FAM3D and FGF8 were significantly down-regulated. Overall, our study not only paves the way for future studies focusing on hypoxia and immune microenvironment but also improves STAD patient’s prognosis and their response to immunotherapy.

Keywords: stomach adenocarcinoma; hypoxia; immune microenvironment; prognostic factor; immunotherapy
1. Introduction

Gastric cancer (GC) is a highly invasive and heterogeneous malignancy that ranked 5th in cancer morbidity with over 1 million patients per year according to the Global Cancer Statistics 2020[1]. As the predominant histological subtype, stomach adenocarcinoma (STAD) accounts for 90%-95% of GC. Most patients display no symptoms before the disease develops to middle or even advanced stage, and only a handful of them are diagnosed at an early stage. Despite breakthroughs in surgery, adjuvant chemotherapy and targeted therapy that dragged down GC mortality rate, the overall survival and quality of survival of advanced GC patients remain poor due to limited therapy and medication[2]. Hence, novel biomarkers and therapeutic targets are in need to improve the prognosis of GC patients.

Tumor immune microenvironment (TIME) is the most important factor in the study of immunotherapy response, and findings on immune-related components and relevant genes are helpful enough to inspire new insights into the research and development of immune checkpoint inhibitors (ICIs). Hypoxia is a common feature in the tumor microenvironment (TME), which can affect survival, proliferation, invasion, metastasis, drug resistance and angiogenesis of tumor cells by regulating cellular and biological activities[3]. It has been confirmed by numerous studies that there is a link between hypoxia and tumor immunology. Qiu et al.[4] showed that HIF-1 (hypoxia-inducible factor-1, a key mediator of tumor metabolism, is involved in tumor-related inflammatory signaling) can promote ENTPD2 (ectonucleoside triphosphate diphosphohydrolase-2, a family of nucleoside triphosphate-diphosphohydrogenase enzymes, can slow cancer growth and improve the efficiency and efficacy of ICIs) expression, which enhances the progression of hepatocellular carcinoma and suppresses myeloid-derived cell differentiation. Meanwhile, the inhibition of ENTPD2 improves the ICI efficiency. Another study of GC conducted by Yun et al.[5] found that HIF-1 can inhibit M1 polarization and function of macrophages by inhibiting miR-30c (a member of the microRNA family, a
tumor suppressor in human cancers) expression to decrease mTOR (mechanistic target of rapamycin, an atypical protein kinase of the PI3K-related kinase family, has a key role in various biological processes such as cell proliferation, survival, autophagy, metabolism and immunity) activity as well as glycolysis in macrophages in GC patients. Therefore, the exploration of hypoxia- and immune-related prognostic markers in STAD would benefit the prognosis prediction and subsequent treatment of STAD patients.

We aimed to develop a novel model based on hypoxia- and immune-related genes to predict prognosis as well as immune infiltration of STAD patients. With relevant STAD data obtained in the public dataset (The Cancer Genome Atlas (TCGA)-STAD), hypoxia- and immune-related genes were analyzed, feature genes screened, prognostic models built, and the link between riskscore and clinicopathological characteristics as well as prognosis validated. In addition, the relationship between riskscore and immune infiltration as well as mutation status were investigated to guide treatment STAD treatment and improve prognosis.

2 Materials and methods

2.1 Dataset collection

TCGA-STAD cohort including 363 samples was downloaded from TCGA, among which there were 343 STAD tissue samples and 30 normal gastric tissue samples. In addition, single-nucleotide variants (SNV) and corresponding clinical data of STAD patients were downloaded as well. Hypoxia-related genes in STAD were obtained in MsigDB (https://www.gsea-msigdb.org/gsea/msigdb/), while immune-related genes in STAD were retrieved from ImmPort (https://www.immport.org).

2.2 Identification of differentially expressed genes (DEGs) related to hypoxia and immunity in STAD

To identify hypoxia- and immune-related DEGs in STAD, edgeR package[6] was
used for differential expression analysis on STAD and normal samples in TCGA-STAD cohort to acquire STAD-related DEGs (|logFC|>1 and FDR<0.05). Subsequently, by intersecting DEGs and hypoxia-related genes/immune-related genes, the hypoxia- and immune-related DEGs in STAD were identified for subsequent analysis.

2.3 The building and validation of hypoxia-immune-based prognostic model of STAD

Referring to the survival information of STAD samples retrieved from TCGA-STAD cohort, those with survival time longer than 30 days were randomly split into training and validation sets in the ratio of 7:3. To select hypoxia- and immune-related genes that affect survival of STAD patients, univariate Cox regression analysis was performed on hypoxia- and immune-related genes in the training cohort using R package survival[7] (P<0.05). To prevent overfitting, the genes acquired from univariate Cox regression underwent LASSO regression using glmnet[8]. 10-fold cross-validation yielded an optimal penalty parameter lambda (λ) to remove genes with high correlation and reduce model complexity. Multivariate Cox regression analysis was performed using R package survival to obtain prognostic genes related to hypoxia and immunity in STAD. Based on the genes required, the prognostic prediction model was built as manifested below: (Riskscore= 0.1968*DUSP1 + 0.0545*IGFBP1 + 0.0587*CGB5 - 0.0775*FAM3D - 0.0991*FGF8 + 0.0971*GPC3 + 0.0655*EGF).

Riskscore of the training cohort were calculated based on the prognostic prediction model. The sample cohort was divided into high- and low-risk groups according to the median. Combined with the survival time and riskscore of the samples, the survival status of STAD patients was plotted using the R package ggplot2[9]. Subsequently, the survival curve of patients in the high- and low-risk groups was plotted using the R package survival to evaluate prognostic differences. Finally, the R package timeROC[10] was used to draw the receiver operating characteristic (ROC) curve of riskscore for predicting the 1-, 3-, and 5-year survival.
of STAD patients, the results of which were then validated in the validation set.

2.4 Validation of the independence of riskscore

Both univariate and multivariate Cox regression analyses were performed based on clinicopathologic characteristics, such as stage, gender, tumor (T), node (N), metastasis (M). Finally, to evaluate the clinical value of riskscore, nomograms were developed using rms[11] to predict 1-, 3-, and 5-year survival in STAD patients and calibration curves were drawn to assess the accuracy.

2.5 Gene Oncology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses of DEGs

To investigate the differences in biological functions and signaling pathways involved in the high- and low-risk groups, package edgeR was used for differential expression analysis to obtain DEGs in the high- and low-risk groups. After that, these DEGs were subject to enrichment analysis using the R package clusterProfiler[12], and the most significantly enriched GO terms and KEGG pathways were selected with a threshold of $P<0.05$.

2.6 Immune microenvironment analysis

The microenvironment has been proven to impact the prognosis of tumor patients. Single sample Gene Set Enrichment Analysis (ssGSEA) was performed using the R package GSVA to investigate TME differences in STAD patients of different riskscores, based on which the association between immune infiltration status and riskscore was analyzed. Subsequently, the correlation of hypoxia-related genes and immune cell scores as well as immune-related function scores were calculated, showing how hypoxia-related genes impact immune infiltration status.

2.7 Tumor mutation analysis and immunophenoscore (IPS)

With gene mutation frequency of high- and low-risk groups calculated based on the mutation data in TCGA-STAD cohort, waterfall plots were drawn using the R
package GenVisR[13] to present the mutational landscape of STAD patients. Chi-square test was used to determine the differences in top 5 gene mutation frequencies in both high- and low-risk groups. IPS of STAD patients were downloaded from The Tree Care Industry Association (TCIA) database for the comparison between high- and low-risk groups by using t-test.

2.8 Cell culture

Human gastric mucosal epithelial cell (CTCC-022-HUM) and human gastric adenocarcinoma epithelial cells NUGC-3 (CTCC-0488-Luc2) and AGS (CTCC-001-0038) were purchased from MEISEN CELL. All the above cell lines were cultured in RPMI-1640 medium supplemented with 10% fetal bovine serum at 37 °C in a 5% CO2 incubator.

2.9 Quantative real-time polymerase chain reaction (qRT-PCR)

Total RNA was extracted from cells using TRIzol reagent (Life Technologies, USA), and the concentration of RNA was determined using a NanoDrop 2000 system (Thermo Fisher Scientific, Inc., USA). Total RNA was reverse transcribed using PrimeScript RT Master Mix (Takara, P.R., Japan) according to kit instructions. The amount of mRNA expression was measured using miScript SYBR Green PCR Kit (Qiagen, Germany). The expression levels of DUSP1, IGFBP1, CGB5, FAM3D, FGF8, and GPC3 were determined by qRT-PCR on a Bio-Rad CFX96 real-time PCR detection system (Bio-Rad Laboratories, Hercules, USA). β-actin was used as a standardized endogenous control. Results $2^{-\Delta\Delta Ct}$ value was used to compare the relative gene expression between the control group and the experimental group. Primer sequences were listed as following: DUSP1[14] (Forward: 5'-GCCACCATCTGCTTGCATTACC-3’; Reverse: 5'-ATGATGCTTCGCTCTGCTTAC-3’), IGFBP1[15] (Forward: 5'-TTGGGACGCCATCAGTACCTA-3’; Reverse: 5'-TTGGCTAACCTCTCTACGACTCT-3’), CGB5[16] (Forward: 5’-CGCTGTGGGACTCAGGTGCTG-3’; Reverse: 5’-CGCTGTGGGACTCAGGTGCTG-3’).
5'-CGCTGTGGACTCAGGTGTGCTG-3'), FAM3D[17] (Forward:
5'-GTAAAAGCCCCTTTGAGCAGT-3'; Reverse:
5'-GGCCATCCCTCGTATTTGTGCAGA-3'), FGF8[18] (Forward:
5'-CCCCTTGCAAGGCTCACT-3'; Reverse:
5'-CCCCTTCTGTCGCTACATC-3'); GPC3[19] (Forward:
5'-GTGGCTTGGCTGCTACATC-3'; Reverse:
5'-TCCACGAGTTCTTGCTCATT-3'), EGF[20] (Forward:
5'-CAGGGAAGATGACCACCACT-3'; Reverse:
5'-CATGTTCCCACCAGTGCTGT-3'), β-actin[14] (Forward:
5'-CTCCACTGCGCGCGCTGTG-3'; Reverse:
5'-GCTGTACCTCACCAGCTTC-3').

3 Results

3.1 Identification of hypoxia- and immune-related prognostic genes in STAD patients

Hypoxic TME is associated with poor outcomes and survival. Hypoxic lesions are formed when intravascular oxygen delivery in tumor is insufficient to meet the metabolic needs of tumor cells. Genes expressed differently under this circumstance are known as hypoxia-related genes[21,22]. Hypoxia affects tumor immunity by favoring immune evasion and resistance[23]. Highly hypoxic TME also impairs the maturation and activity of dendritic cells (DCs) and natural killer (NK) cells[23]. A positive correlation is found between the hypoxic TME and metabolic changes in cells of the immune system[23]. These findings indicate that hypoxia regulation can be a target in ICI therapy. We retrieved 401 hypoxia- and immune-related genes from MsigDB and ImmPort, respectively, and overlapped them with DEGs. Eventually, 243 hypoxia- and immune-related DEGs were obtained for subsequent prognostic model construction (Figure 1).

3.2 The construction and evaluation of the STAD prognostic model
Tumor samples in TCGA-STAD cohort were randomly split into the training set and validation set at the ratio of 7:3. Referring to both expression data of 243 candidate genes in the training set and the corresponding clinical information, samples with survival time >30 days were screened. A univariate Cox regression analysis screened 29 candidate genes significantly associated with survival. Subsequently, optimal penalty parameter $\lambda$ was selected using ten-fold cross-validation. When $\lambda = -3$, the overfitting was relatively small and complexity of the model was the lowest. 13 important candidate genes were finally selected out (Figure 2A-B). After multivariate Cox regression analysis, a 7-gene hypoxia--immune-based signature was constructed for STAD prognosis (Figure 2C). After calculating the riskscore of each sample, samples in training set were divided into high- and low-risk groups by the median riskscore (Figure 3A). According to scatter plot of survival status, higher riskscores correlated with poorer survival status (Figure 3B). The expression levels of prognostic signature genes in the high- and low-risk groups were presented using heat maps (Figure 3C). In addition, better prognosis was found in low-risk groups (Figure 3D). Subsequently, 1-, 3-, and 5-year survival of STAD patients predicted by riskscore were plotted using ROC curves, with AUC values of 0.66, 0.73, and 0.64, respectively (Figure 3E). The predictive performance of riskscore was proven. Later, we conducted validation in the validation set, and the trend was consistent with the validation set, with better outcomes for patients in low-risk group (Figure 3F). The AUC of 1-, 3-, and 5-year curves were 0.71, 0.71, and 0.78, respectively (Figure 3G). In general, an effective prognostic prediction model based on immune-related and hypoxia-related genes in STAD was constructed here.

3.3 Riskscore can be an independent prognostic factor for STAD patients

Univariate and multivariate Cox regression analyses were conducted combined with riskscore and clinicopathologic characteristics such as gender and stage. Univariate Cox regression analysis identified two independent prognostic factors for STAD: N and riskscore (Figure 4A). According to multivariate Cox regression analysis, age, N, M, and riskscore could be independent prognostic factors (Figure
Overall, riskscore may be an independent prognostic factor in STAD patients. Subsequently, after plotting a nomogram that predicted 1-, 3-, and 5-year survival (Figure 4C), corresponding calibration curves were plotted and showed the high predictive accuracy of the nomogram (Figure 4D-F). Based on above findings, riskscore obtained from the prognostic model containing 7 hypoxia- and immune-related genes could serve as an independent predictor.

3.4 GO and KEGG enrichment analyses

With 1340 DEGs yielded from the differential analysis, GO and KEGG enrichment analyses were performed. According to the GO analysis, most of these genes were enriched in receptor-ligand activity, collagen-containing extracellular matrix, muscle system process, signaling receptor activator activity, receptor activator activity, and axon development (Figure 5A). KEGG analysis unveiled enrichments in Neuroactive ligand-receptor interaction, cAMP signaling pathway, Chemical carcinogenesis-receptor activation, Calcium signaling pathway and Vascular smooth muscle contraction (Figure 5B).

3.5 Differential analysis of TME

TME is closely related to the progression and prognosis of cancer patients[24]. After detecting the infiltration level of immune cells in the TME of each STAD patient using ssGSEA, immune infiltration and immunocompetence difference between two risk groups were also compared. The results identified more immune-related components as well as higher immunocompetence in the low-risk group (Figure 6A-C). Subsequently, the correlation between riskscore and immune function as well as immune cell scores was calculated. The results found that riskscore was strongly negatively correlated with APC_co_inhibition (APC: antigen presenting cell), MHC_class_I (MHC: major histocompatibility complex), T_cells_co-stimulation, APC_co_stimulation, Th2_cells (T helper 2 cells), Cytolytic_activity, Inflammation-promoting, Tfh (T follicular helper cells), Check-point, Th1_cells (T helper 1 cells), pDCs (Plasmacytoid dendritic cells), and...
T-cell co-inhibition. This indicated that patients with lower risk score are more likely to benefit from immunotherapy (Table 1).

According to previous studies, DUSP1 inhibits ERK kinase activity, thereby preventing the overactivation of HIF-1, a key player in regulating oxygen delivery in hypoxia tumor cells and metabolic adaptation to hypoxia[25]. Hypoxia affects tumor immune status by facilitating immune evasion and resistance[23]. Therefore, here we calculated the correlation between DUSP1 and a variety of immune-related cells and functions. DUSP1 was positively correlated not only with immune-related cell scores such as Mast_cell and Neutrophils but also with immune-related function scores such as CCR and type_II_INF_REPONSE. These observations added weight to the finding that, via repressing hypoxia, DUSP1 may inhibit immune evasion and resistance of tumor cells to suppress cancer progression (Figure 7A-D).

3.6 Tumor mutation analysis and IPS differences in high- and low-risk groups

The top 30 most frequently mutated genes, according to the SNV mutation data from TCGA-STAD cohort, were selected to draw waterfall plots in two risk groups in the validation set (Figure 8A-B). According to the results, the high-risk group was identified with a higher mutation frequency of the top 30 genes, with the top 5 genes being TP53, TTN, MUC16, LRP1B, and SYNE1. As for the low-risk group, five genes that topped the mutation frequency rank were TTN, TP53, ARID1A, MUC16, and SYNE1. Subsequently, we performed a Chi-square test on the above-mentioned genes to investigate whether their mutant frequency varied in two groups. The results presented that the mutation frequencies of TTN and ARID1A were substantially elevated in the high-risk group (Table 2).

IPS is a measurement of how oncology patients respond to ICI therapy, and higher scores indicate a better response to immunotherapy. After downloading IPS of STAD patients from TCIA, we plotted differences between 4 IPSs across high- and low-risk groups. According to the results, IPS, IPS-CTLA4 blocker score, IPS-PD1/PDL1/PDL2 blocker score, and IPS-CTLA4/PD1/PDL1/PDL2 blocker score were substantially increased in the low-risk group (Figure 8C-F). Thus, patients in
the low-risk group were more likely to benefit from ICI treatment.

3.7 Validation of expression levels of prognostic signature genes in STAD

To verify whether prognostic signature genes were significantly different between STAD and normal gastric cells, we performed q-PCR analysis. As expected, the expression levels of DUSP1, IGFBP1, CGB5, GPC3 and EGF in STAD cells were significantly higher than those in normal gastric cells, while the expression levels of FAM3D and FGF8 were significantly lower (Figure 9).

4. Discussion

TME is a complex network crucial to the development of cancer. Among elements within such a complex network, hypoxia is considered to be the most relevant factor with tumor response[26]. Prevalent in most solid tumors, hypoxia bears strong relationship with drug resistance and poor prognosis as it promotes local and systemic tumor progression by activating angiogenesis[27]. Besides, the hypoxic TME can suppress the efficacy of the immune response[28]. It has been demonstrated that hypoxic TME can inactivate immune effector cells and promote the activity of immunosuppressive cells, with such an immunosuppression state further enhanced by immune evasion and tumor cell adaptability to hypoxia[29-32]. Therefore, according to the characteristics of tumor hypoxia and immunity, this study retrieved relevant data from public databases and by adopting bioinformatics analysis, built a STAD prognostic model related to hypoxia and immunity.

Hypoxia- and immune-related genes in STAD were analyzed and screened by bioinformatics methods in this study, which yielded a total of 7 prognostic feature genes (DUSP1, IGFBP1, CGB5, FAM3D, FGF8, GPC3 and EGF). Immune-related genes included CGB5, FAM3D, FGF8 and EGF. CGB5 (chorionic gonadotropin-β 5) is a protein-coding gene that is mainly involved in invasive melanoma and ectopic pregnancy. Qin et al.[33] found that CGB5 is a critical gene in STAD patients’ prognosis prediction and is associated with a variety of immune cells. As a member of
FAM3, FAM3D (Family with sequence similarity 3 member D), is a member of the cytokine-like family and plays an important role in cell proliferation, and it has been proven to be associated with intestinal inflammation and notably highly expressed in gastrointestinal inflammation[34,35]. FGF (fibroblast growth factor) family genes are now included in anti-cancer therapy as targets to combat chemoresistance in many different malignancies. In a study conducted by Jomrich et al.[36], FGF8 is highly expressed in adenocarcinoma of the esophagogastric junction and shorter overall survival accompanies with higher FGF8 expression. EGF and its receptors are crucial to tumors development. Li et al.[37] discovered that, in basal-like breast cancer, the inhibition of the EGF signaling pathway enhances PD-L1 (Programmed Cell Death-Ligand 1, a type I transmembrane protein of 40kDa, promotes proliferation of T cells with antigen specificity) stability and the therapeutic effect of PD-1 (Programmed Death 1, an important immunosuppressive molecule) blockade, thus facilitating tumor-infiltrating cytotoxic T-cell immune response. It can be seen that these biomarkers can be used to assess the extent of clinical disease progression. In future molecular or clinical studies, these markers may also have the potential to be used as therapeutic targets to improve the prognosis of patients with various tumors, including gastric adenocarcinoma. In the clinical treatment setting, prognostic biomarkers can also help to decide whether or how to aggressively pursue therapeutic interventions.

Two hypoxia-related genes used to build the prognostic model were DUSP1 and IGFBP1. DUSP (Dual-specificity protein phosphatase1), a subfamily of the protein tyrosine phosphatase (PTP) superfamily, is a negative regulator of HIF-1α[38]. It has been demonstrated that DUSP1 plays an important role in regulating cell proliferation, tumorigenesis and drug resistance. CASC9-EZH2-DUSP1 regulatory axis is confirmed to regulate p-ERK (ERK, extracellular signal-regulated kinase, belongs to the mitogen-activated protein kinase family and is responsible for fundamental cellular processes, including cell proliferation and differentiation) expression, thus promoting gefitinib resistance in non-small cell lung cancer[39]. IGFBP (Insulin-like growth factor-binding protein) is a series of cystine-rich proteins that can bind to IGFs
in serum and regulate cell proliferation[40]. After analyzing IGFBP family genes in GC, Liu et al.[41] found that IGFBP1 is highly expressed in GC, and its higher expression level is accompanied by a shorter survival time. Xu et al.[15] found that IGFBP1 may regulate renal cell clear cell carcinoma progression and immune infiltration by mediating the biological function of monocyte. Thus, the hypoxia signature genes screened in this study are closely related to tumor immune infiltration as well as prognosis.

In recent years, immunotherapy has become more popular, since it can effectively treat cancer and uplift cancer patients by affecting the interaction between the human immune system and cancer[42]. Although the application of immunotherapy is a great success in the treatment of various tumors, its efficacy is not ideal due to the limited response to immunotherapy. Hence, biomarkers indicating the response condition to immunotherapy becomes the focus of current studies. There is a close link between tumor transformation and somatic mutations. Therefore, it is important to explore TMB and tumor immune infiltration for the development of immunotherapy regimens and the prognosis of patients. Currently, the research focus of mounting studies is placed on the relationship between TMB and tumor immune infiltration. Ren et al.[43] analyzed data from GC patients in public databases using the TIDE algorithm and found patients with low stromal cell infiltration scores had higher TMB, microsatellite instability, and sensitivity to ICIs. Wang et al.[44] found that, when using the PD-1 inhibitor toripalimab to treat advanced GC patients, the response rate is significantly higher in patients with high TMB [14.6 versus 4.0 months, HR=0.48 (96% CI 0.24 - 0.96), \( P=0.038 \)], while PD-L1 expression was not associated with patient survival. In this study, we combined TMB score with tumor immune infiltration and found that patients with low riskscore had higher immune infiltration, immune activity and TMB. It indicated that immunotherapy may yield better results among GC patients with low riskscores, which was consistent with previous studies. However, hypoxia triggers a series of events, such as promoting tumor growth, enhancing tumor immune escape and stimulating tumor angiogenesis[45]. It has been found that hypoxia enhances HLA-G expression in
tumors\[46,47\]. HLA-G is a marker of tumor immune escape. Murdaca et al.\[48\] found that HLA-G expression corresponded to a lower poor survival rate in patients with stage III gastric adenocarcinoma, which may be related to immune escape mechanisms of cancer cells\[49\]. This may be one of the reasons for the failure of immunotherapy and chemoresistance in some cancer patients\[50\]. Exploiting tumor hypoxia may be a potential cancer treatment strategy.

In summary, with data retrieved from public databases, this study analyzed hypoxia- and immune-related genes in STAD patients and built a 7-gene prognostic model with stable predictive performance. The prognostic risk score of patients can not only act as an independent prognostic factor but also as a reference for patients receiving immunotherapy. At the same time, the feature genes of the prognostic model in this study are closely related to the occurrence and development of tumors and can be used as potential targets for STAD treatment, which can be used as an entry point to conduct a more in-depth study of the pathogenesis of STAD. Unfortunately, there are still some limitations in this study. For instance, retrospective data in public databases used in this study that may affect the stability of the model due to inherent selection bias. Clinical data will subsequently be introduced through collaboration with other institutes to fully investigate the potential value of the 7-gene signature in clinical practice. Besides, multi-prospective studies and in vivo and in vitro experiments are needed to deeply explore the link between risk score, TMB and immune infiltration to fully prove the accuracy of prognostic models.

**Declaration**

**Data availability**

The datasets generated during and analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

**Funding**

This study was funded by The Science and Technology Project of The Health Planning Committee of Sichuan Province, China, grant number: 21PJ147.
Conflict of interest
All authors confirm no conflicts of interest in this work.

Ethics approval and consent to participate
Not applicable.

Reference
49 Liu, L., Wang, L., Zhao, L., He, C. & Wang, G. The Role of HLA-G in Tumor Escape: Manipulating the Phenotype and Function of Immune Cells. Front Oncol 10, 597468,

Figure legends

**Figure 1** Upset plot of the screening of hypoxia- and immune-related prognostic genes in STAD patients

**Figure 2** Identification of hypoxia- and immune-related prognostic genes
(A) Survival-related genes in STAD when log $\lambda$ approaches 0 in the LASSO Cox regression model; (B) Selection interval of the penalty parameter $\lambda$ for the minimum goodness-of-fit; (C) Forest plot of multivariate Cox regression analysis of feature genes in the best prognostic model.

**Figure 3** Validation of the STAD hypoxia- and immune-related prognostic model
(A) Distribution of the high- and low-risk groups samples with the median riskscore as the cut-off point; (B) Scatterplot of riskscore and survival status in the training set; (C) Survival analysis of the high- and low-risk groups in the training set; (D) Riskscore was utilized to predict ROC curves of 1-, 3-, and 5-year survival of patients in the training set; (E) Survival analysis of the high- and low-risk groups in the validation set; (F) Riskscore for predicting the ROC curves of 1-, 3-, and 5-year survival of patients in the validation set.

**Figure 4** Riskscore can be used as an independent prognostic factor in STAD patients
(A) Univariate Cox regression analysis of riskscore and clinicopathological characteristics; (B) Multivariate Cox regression analysis of riskscore and clinicopathological characteristics; (C) Nomogram predicting 1-, 3-, and 5-year survival in STAD patients; (D-F) Calibration curves of 1-, 3-, and 5-year survival.

**Figure 5** GO and KEGG enrichment analyses of high- and low-risk groups
(A-B) Bubble map of GO and KEGG enrichment analyses.

**Figure 6** ssGSEA of high- and low-risk groups
(A) Heat map depicts the expression of 29 immune cell gene sets in high- and low-risk groups; (B) Violin plot of immune cell fraction differences in high- and low-risk groups; (C) Violin plot of immunocompetence-related gene differences in high- and low-risk groups.

**Figure 7 Correlation analysis of DUSP1 with immune-related cell and function scores**

(A) type_I_II_INF_REPONSE; (B) Mast_cell; (C) Neutrophils; (D) CCR.

**Figure 8 Tumor mutation analysis and IPS difference in high- and low-risk groups**

(A) Tumor mutation burden (TMB) waterfall plot of top 30 mutated genes in high-risk group; (B) TMB waterfall plot of top 30 mutated genes in low-risk group; (C) IPS difference between high- and low-risk groups; (D) IPS-CTLA4 blocker scores; (E) IPS-PD1/PDL1/PDL2 blocker scores, and (F) IPS-CTLA4/PD1/PDL1/PDL2 blocker scores in both groups.

**Figure 9 The mRNA expression levels of DUSP1, IGFBP1, CGB5, GPC3, EGF, FAM3D and FGF8 in different cell lines were measured by qRT-PCR (*P < 0.05).**
### Table 1 Correlation analysis of risk score with immune cell scores and immune function score

<table>
<thead>
<tr>
<th>Name</th>
<th>Correlation</th>
<th>P.val</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC_co_inhibition</td>
<td>1.042e-04</td>
<td>-0.26</td>
</tr>
<tr>
<td>MHC_class_I</td>
<td>4.628e-04</td>
<td>-0.24</td>
</tr>
<tr>
<td>T_cell_co-stimulation</td>
<td>2.582e-03</td>
<td>-0.2</td>
</tr>
<tr>
<td>APC_co_stimulation</td>
<td>4.68e-03</td>
<td>-0.19</td>
</tr>
<tr>
<td>Th2_cells</td>
<td>6.188e-03</td>
<td>-0.19</td>
</tr>
<tr>
<td>Cytolytic_activity</td>
<td>6.793e-03</td>
<td>-0.18</td>
</tr>
<tr>
<td>Inflammation-promoting</td>
<td>2.102e-02</td>
<td>-0.16</td>
</tr>
<tr>
<td>Tfh</td>
<td>1.736e-02</td>
<td>-0.16</td>
</tr>
<tr>
<td>Check-point</td>
<td>2.757e-02</td>
<td>-0.15</td>
</tr>
<tr>
<td>Th1_cells</td>
<td>3.078e-02</td>
<td>-0.15</td>
</tr>
<tr>
<td>pDCs</td>
<td>3.544e-02</td>
<td>-0.14</td>
</tr>
<tr>
<td>T_cell_co-inhibition</td>
<td>4.71e-02</td>
<td>-0.14</td>
</tr>
</tbody>
</table>

### Table 2 Chi-square test results of the top 5 genes ranked by mutation frequencies in the high- and low-risk groups

<table>
<thead>
<tr>
<th>ID</th>
<th>X-SQUARED</th>
<th>P-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTN</td>
<td>6.32213</td>
<td>0.01192</td>
</tr>
<tr>
<td>TP53</td>
<td>0.16617</td>
<td>0.68354</td>
</tr>
<tr>
<td>ARID1A</td>
<td>7.78352</td>
<td>0.00527</td>
</tr>
<tr>
<td>MUC16</td>
<td>1.662</td>
<td>0.19733</td>
</tr>
<tr>
<td>SYNE1</td>
<td>2.67463</td>
<td>0.10196</td>
</tr>
<tr>
<td>LRP1B</td>
<td>0.51928</td>
<td>0.47115</td>
</tr>
</tbody>
</table>
A

B

C

<table>
<thead>
<tr>
<th>Gene</th>
<th>Hazard Ratio (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUSP1</td>
<td>1.22 (1.00 - 1.48)</td>
<td>0.051</td>
</tr>
<tr>
<td>IGFBP1</td>
<td>1.06 (1.00 - 1.12)</td>
<td>0.071</td>
</tr>
<tr>
<td>CGB5</td>
<td>1.06 (1.01 - 1.12)</td>
<td>0.028 *</td>
</tr>
<tr>
<td>FAM3D</td>
<td>0.93 (0.86 - 0.99)</td>
<td>0.033 *</td>
</tr>
<tr>
<td>FGF8</td>
<td>0.91 (0.82 - 1.00)</td>
<td>0.043 *</td>
</tr>
<tr>
<td>GPC3</td>
<td>1.10 (1.01 - 1.20)</td>
<td>0.033 *</td>
</tr>
<tr>
<td>EGF</td>
<td>1.07 (0.99 - 1.15)</td>
<td>0.087</td>
</tr>
</tbody>
</table>

# Events: 92; Global p-value (Log-Rank): 1.0454e-05
AIC: 84.05; Concordance Index: 0.67
A

B

Neuronal-and-receptor interaction
Calcium signaling pathway
AMPK signaling pathway
Chemokine signaling pathway
neprilysin-mediated natriuretic peptide activation
vasodilator-stimulated phosphoprotein signaling pathway
Adenosine signaling in cardiac myocytes
Complement and coagulation cascades
Regulation of actin cytoskeleton
Protein digestion and absorption
cGMP-PKG signaling pathway
Pancreatic secretion
Diabetes mellitus
Dilated cardiomyopathy
Cancer signaling pathways and inhibition
Hyperpigmentation or hyperpigmentation-related disorders
Antiphospholipid antibody syndrome
Cancerous neoplasm of unspecified behavior
Erb1 signaling
Salivary secretion
Sphingomyelinosis, genetic disorders
Adenosine A2A receptor signaling pathway
Signaling molecule and receptor activity
Drug metabolism – xenobiotic metabolism
Metabolism of xenobiotics by cytochrome P450
Calcium ion transport
Gastrointestinal acid secretion
Regulation of lipoprotein metabolism
Eicosanoid metabolism
Nicotinic acid metabolism
Porphyrin metabolism
Fat digestion and absorption
Peroxisome and glucuronate metabolism
Interconversions of fatty acids