Development of a prediction model for radiotherapy response among patients with head and neck squamous cell carcinoma based on the tumor immune microenvironment and hypoxia signature

Guang-Li Zhu | Kai-Bin Yang | Cheng Xu | Rui-Jia Feng | Wen-Fei Li | Jun Ma

Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangdong Key Laboratory of Nasopharyngeal Carcinoma Diagnosis and Therapy, Guangzhou, P. R. China

Correspondence
Jun Ma, Department of Radiation Oncology, Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, 651 Dongfeng Road East, Guangzhou 510060, P. R. China.
Email: majun2@mail.sysu.edu.cn

Funding information
This work was supported by grants from the Key-Area Research and Development Program of Guangdong Province (2019B020230002), the Natural Science Foundation of Guangdong Province (2017A030312003), the Medical Science and Technology Research Fund of Guangdong Province (A2019418), and the Overseas Expertise Introduction Project for Discipline Innovation (111 Project, B14035).

Abstract

Introduction: The immune system and hypoxia are major factors influencing radiosensitivity in patients with different cancer types. This study aimed at developing a model to predict radiotherapy response in patients with head and neck squamous cell carcinoma (HNSCC) based on the tumor immune microenvironment and hypoxia signature.

Materials and Methods: We first evaluated the hypoxia status and tumor immune microenvironment in the Cancer Genome Atlas (TCGA) cohort by using transcriptomic data. Differentially expressed genes (DEGs) were identified between the “high immunity and low hypoxia” and “low immunity and high hypoxia” groups and those DEGs significantly associated with disease-specific survival in the univariate Cox regression analysis were selected as the prognostic DEGs. We selected the immune hypoxia–related genes (IHRGs) by intersecting prognostic DEGs with immune and hypoxia gene sets. We used the IHRGs to train a multivariate Cox regression model in the TCGA cohort, based on which we calculated the IHRG prognostic index (IHRGPI) for each patient and validated its efficacy in predicting radiotherapy response in the Gene Expression Omnibus cohorts. Furthermore, we explored potential mechanisms and effective combinational treatment strategies for different IHRGPI groups.

Results: Five IHRGs were used to construct the IHRGPI, which was used to dichotomize the cohorts. The patients with lower IHRGPI showed a better radiotherapy response across different cohorts and endpoints, including overall survival, progression-free survival, and recurrence-free survival ($p<0.05$). Patients with higher IHRGPI showed greater hypoxia and lesser immune cell infiltration. A lower IHRGPI indicated a better immunotherapy response, while a higher IHRGPI indicated a better chemotherapy response.
1 | INTRODUCTION

Head and neck squamous cell carcinoma (HNSCC) is a heterogeneous group of tumors that originate from the mucosal epithelium of head and neck, including the lip, oral cavity, larynx, and pharynx (excluding the nasopharynx). It is the most common type of head and neck cancer and the sixth leading cancer worldwide. In 2018, its incidence and mortality were reported to be 890,000 and 450,000, respectively. However, hitherto, no effective screening method is available for HNSCC, and more than 50% of the cases were found to be locoregionally advanced (stage III–IV) at diagnosis. Surgery, radiotherapy, and chemotherapy are the major treatment modalities for non-metastatic HNSCC. Surgery and radiotherapy are the mainstays in most cases. For high-risk patients who have undergone resection, postoperative adjuvant radiotherapy alone or chemoradiotherapy is recommended. Considering individual differences in response to radiotherapy, the optimal criteria for treatment decisions remained unestablished. However, individual differences in radiosensitivity greatly limit the therapeutic effect of radiation. Therefore, current treatment plans might be inadequate for the management of radioresistant HNSCC. Furthermore, since radiotherapy has deleterious effects on adjacent normal organs and tissues, which might impair phonation or swallowing, it might not be the optimal choice for radioresistant HNSCC patients. Recently, radiotherapy techniques have improved greatly. The advent of intensity-modulated radiotherapy (IMRT) has resulted in a better dose-volume distribution and greatly reduced toxicities. However, compared with two-dimensional radiotherapy or three-dimensional conventional radiotherapy, IMRT did not achieve better recurrence-free survival (RFS), disease-free survival (DFS), and overall survival (OS). Under such circumstances, increasing the strength of combinatory treatment modalities or switching to other treatment modalities might be necessary. Thus, considering radiosensitivity during treatment decisions is necessary for choosing the optimal curative and safe approach.

In the traditional theory regarding the biological effects of irradiation, the factors affecting radiotherapy response are known as the “5 Rs”: Repair of radiation-induced DNA damage, Redistribution/Reassortment of cell cycles, Regeneration/Repopulation of survived cells, Reoxygenation and cellular intrinsic Radiosensitivity. Oxygen is a radiosensitizer because it plays a key role in the indirect radiation effect, in which radiation induces free radical generation to attack the target molecules. Compared with euoxic cells, hypoxic cells are 2–3 times more resistant to radiation. With the rapid development of immunotherapy, the 6th R, Reactivation of the immune system, was proposed in 2019. Radiation attacks DNA and induces DNA damage repair, during which the tumor cells become more visible to the immune system, and the tumor microenvironment is altered. Innate and adaptive immunity is initiated to cause a specific immune response against the tumor and immunogenic cell death. Besides the local effect, the irradiated tumor becomes an in situ vaccine, inducing a systemic immune response. Non-irradiated metastases also shrink under the attack of the immune system, which is known as the abscopal effect. The findings associated with the 6th R are changing our understanding of radiosensitivity, which was previously restricted to tumors themselves.

The development of a model to predict radiotherapy response among HNSCC patients could guide individualized treatment decisions and improve prognosis. With the rapid development of sequencing techniques, many gene signatures were developed to predict radiosensitivity. However, most of them were purely statistically driven and not based on radiobiological theories; therefore, the underlying mechanism is hard to explain. Furthermore, most did not consider the impact of the tumor immune microenvironment, because they were developed in cell lines. Hence, in this study, we tried to use transcriptomic data to develop and validate an immune hypoxia-related gene prognostic index (IHRGPI) based on the tumor immune microenvironment and hypoxia signature to predict radiotherapy response in HNSCC patients, investigate the underlying mechanisms, and explore potential effective combinatory treatment strategies.

2 | MATERIALS AND METHODS

2.1 | Data acquisition and preparation

Figure 1 presents the workflow of this study. The Cancer Genome Atlas (TCGA) HNSCC cohort consists of 501
HNSCC patients, providing 546 HNSCC samples including 500 primary tumor samples, two metastatic tumor samples, and 44 para-cancer samples. The cohort’s RNA-Seq data were downloaded from the TCGA database (https://portal.gdc.cancer.gov) by using the R package “TCGAbiolinks” on June 1, 2021. The corresponding clinicopathological information was downloaded from the cBioPortal database (https://www.cbioportal.org/). The microarray gene expression data and clinicopathological information of two HNSCC cohorts were downloaded from Gene Expression Omnibus (GEO) database (accession numbers: GEO cohort-1 and GEO cohort-2, GSE117973 and GSE39366, respectively). Disease-specific survival (DSS) and RFS were the only available endpoint for GEO cohort-1 and cohort-2 respectively. Therefore, DSS, which was available for both the TCGA cohort and the GEO cohort-1, was chosen as the primary clinical endpoint. The other endpoints, including RFS, PFS, and OS data from GEO cohort-2, GEO cohort-1 and TCGA cohort, respectively, were used to investigate the model’s generalizability for the prediction of the radiotherapy effect on the other endpoints. Primary tumor samples from patients who received radiotherapy and whose complete follow-up data were available were included in the analysis. The study strictly followed the principles of the Declaration of Helsinki and data access policies of the TCGA and GEO database. The requirements of patients’ informed consent and institutional review board approval were waived given the datasets were de-identified and publicly available.

2.2 Identification of hypoxia and tumor microenvironment status

To investigate the patients’ hypoxia status, the “WINTER” hypoxia gene set containing 95 hypoxia-related genes upregulated in head and neck tumor samples was downloaded from the molecular signature database (MSigDb version 7.4, https://www.gsea-msigdb.org/gsea/msigdb/cards/WINTER_HYPOXIA_UP.html). The included TCGA samples were divided
into two groups per the Euclidian distance of their expression patterns of the hypoxia gene sets, which were visualized by heatmap and t-distributed Stochastic Neighbor Embedding (t-SNE) with the R packages “heatmap” and “Rtsne.” The group with the obviously upregulated hypoxia gene set was annotated as the “high hypoxia” group, while the other as the “low hypoxia” group. To validate the hypoxia status, the differences in the expression levels of the genes involved in the HIF1 signal transduction pathway were compared between the two groups by using the limma algorithm. Genes with a false discovery rate (FDR) adjusted P-value < 0.05 and an absolute log2 (fold change) > 1 were considered as differentially expressed genes (DEGs). The genes involved in the HIF1 signal transduction pathway were downloaded from the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (https://www.kegg.jp; ID:04066). Finally, the difference in DSS between the two groups was evaluated using Kaplan–Meier curves with the log-rank test.

To infer the patients’ tumor microenvironment status, the Estimation of STromal and Immune cells in Ma lignant Tumor tissues using Expression data (ESTIMATE) was applied for the TCGA cohort. Using ESTIMATE, the immune and stromal scores were calculated to estimate the degree of immune and stromal cell infiltration in tumor tissues based on the gene expression profile. The patients were divided according to the cutoff of the ESTIMATE immune score determined by maximally selected log-rank statistics with the R package “survminer.” The group with the higher immune score was annotated as the “high immunity” group, while the other as the “low immunity” group. To explore differences in the tumor microenvironment between the two groups, the proportions of different types of tumor-infiltrating cells from the transcriptomic data of each tumor sample were estimated using the Microenvironment Cell Populations-counter (MCP-counter) algorithm, which allows quantification of the absolute abundance of eight immune and two stromal cell populations and between-sample comparison. Finally, the difference in DSS between the two groups was evaluated based on Kaplan–Meier curves with the log-rank test.

2.3 Identification of IHR subtypes and their association with survival

Based on the findings obtained from the aforementioned steps, the TCGA cohort was further divided into three groups by using a two-dimensional index combining the tumor microenvironment and hypoxia status: including “high immunity and low hypoxia”, “low immunity and high hypoxia”, and “mixed” groups. Finally, the difference in DSS between the “high immunity and low hypoxia” and “low immunity and high hypoxia” groups was evaluated using Kaplan–Meier curves with the log-rank test.

2.4 IHRGPI development and validation

To build a prognostic model for radiotherapy response prediction in the TCGA cohort, first, the “limma” algorithm was used to identify DEGs between the “high immunity and low hypoxia” and “low immunity and high hypoxia” groups (absolute log2 [fold change] > 1; FDR-adjusted P-value < 0.05). Next, univariate Cox regression analysis was performed to select prognostic DEGs that were significantly associated with DSS (P < 0.05). The hallmark hypoxia gene set was retrieved from MSigDb, and an immune gene set was compiled by combining all immune-related gene sets retrieved from InnateDB (http://www.inнатedб.com). Prognostic DEGs that were present in both the immune and hallmark hypoxia gene sets were selected as the IHRGs for model training. To improve comparability affected by platform differences, the expression levels of the genes in the TCGA and GEO datasets were individually transformed to z-scores.

A multivariate Cox regression model was trained with the above selected IHRGs in the TCGA cohort. IHRGPI was the sum of the products of the expression value of each gene for each patient and their coefficients in the trained multivariate Cox model. To investigate whether the IHRGPI score is an independent prognostic factor of radiotherapy response, univariate and multivariate Cox regression analyses incorporating common clinical variables and IHRGPI were performed for the TCGA cohort. According to IHRGPI with 0 as the cut-off value, patients in each cohort were dichotomized into radiosensitive (IHRGPI<0) and radioresistant (IHRGPI≥0) groups. GEO cohort-1 served as an external validation source. To reduce the bias introduced by confounding factors, 1:1 propensity score matching was performed for T and N stages in the TCGA cohort and GEO cohort-1 respectively. The prognostic power of the IHRGPI group in predicting radiotherapy response was evaluated using Kaplan–Meier curves with the log-rank test and univariate Cox regression model in both TCGA and GEO cohorts. Furthermore, to explore the generalizability of IHRGPI to other endpoints, the prognostic significance was explored with OS, PFS, and RFS as an
Comprehensive analysis of hypoxic, tumor-microenvironmental, and molecular characteristics in the two IHRGPI groups

To illustrate the mechanism underlying the differences in radiotherapy response in the two IHRGPI groups, first, differential expression analysis was performed with the limma algorithm\(^18\) for both groups in the TCGA cohort to identify radiosensitivity-associated DEGs (absolute log2[fold change] >1; FDR-adjusted P-value <0.05). The radioresistant group was set as the control group. Furthermore, functional enrichment analysis was performed with these DEGs by using the R package “clusterProfiler”.\(^22\) Radiosensitivity-associated DEGs were first mapped to Gene Ontology (GO, \url{http://geneontology.org/}) terms and KEGG (\url{https://www.genome.jp/kegg/}) pathways. Next, the correlation between the genes adopted in the trained model was explored using GO terms. Finally, gene set enrichment analysis (GSEA) was performed to identify the enriched gene sets in both groups. The gene sets utilized in the GSEA were retrieved from MSigDb, including GO (C5) and Hallmark (H) gene sets. To analyze the difference in the tumor immune microenvironment between the two groups, CIBERSORT\(^23\) (\url{https://cibersort.stanford.edu/}) was used to estimate and compare the relative abundance of different types of immune or stromal cells from the gene expression profile. Regarding genetic mutational analyses, the somatic mutational data of the TCGA cohort were downloaded from TCGA database by using the R package “TCGAbiolinks”.\(^10\) Somatic mutations in the IHRGPI groups were analyzed and visualized, and the tumor mutational burden (TMB) was calculated using the R package “maftools”.\(^24\) TMB was defined as the total number of non-synonymous somatic mutations per megabase (Mb) of the tumor genome. Finally, the ChIP-X Enrichment Analysis 3 (ChEA3) tool\(^25\) was used to perform transcription factor enrichment analysis to identify transcription factors significantly associated with the selected IHRGs.

Exploration of treatment strategies for two IHRGPI groups

To explore the potential effective combinational strategies of radiotherapy and immunotherapy, the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm\(^26\) (\url{http://tide.dfci.harvard.edu}) was used to predict the patients’ response to immune checkpoint inhibitors. Tumors with higher levels of cytotoxic T cells (CTLs) achieved immune escape by inhibiting CTL functions, while tumors with lower levels of CTLs achieved immune escape by inhibiting CTL infiltration. The TIDE score takes both conditions into consideration to evaluate the potential of tumor immune escape based on gene expression profiles. A higher TIDE score indicates a higher immune evasion phenotype and worse response to immune-checkpoint inhibitors. Next, the expression levels of several well-studied targets for immune checkpoint inhibitors were compared, including \(\text{CTLA4, PD1, PDL1, B7H3, B7H4, B7H5 (VISTA), BTLA, LAG3, TIM3, TIGIT, PVRIG, A2aR, CD73, and NKG2A}^{27,28}\), between the two groups.

To explore the potential effective combination strategies of radiotherapy and chemotherapy, the R package “pRRophetic” was used to predict individual sensitivity to three commonly used chemotherapeutic drugs for HNSCC: cisplatin, gemcitabine, and docetaxel. The “pRRophetic” algorithm\(^29\) was developed based on the public pharmacogenomics database Genomics of Drug Sensitivity in Cancer (\url{https://www.cancerrxgene.org}). It built a ridge regression model from gene expression profile and drug sensitivity data.

Statistical analysis

Statistical analysis was performed using the R software (v.4.0.1, \url{https://www.r-project.org/about.html}). The independent Student’s \(t\)-test was used to compare differences in continuous variables following normal distribution between two groups, while the Wilcoxon test was performed for continuous variables that did not follow normal distribution. The chi-square test or Fisher’s exact test was used to test differences in categorical variables between two groups. Multiple testing was adjusted by the FDR method.

RESULTS

Hypoxia status and immunophenotype of the TCGA cohort

A total of 274 primary tumor samples from eligible patients in the TCGA cohort were included in the analysis. To infer the hypoxia status of the TCGA cohort, we first clustered the samples into two groups according to the Euclidian distance of their expression patterns in the “WINTER” hypoxia gene sets. Subsequently, 148 and 126 patients were clustered into cluster 1 and 2, respectively (Figure 2A). The expression level of the “WINTER” hypoxia gene sets was significantly upregulated in cluster 2 than in cluster 1. Since these gene sets consisted of
FIGURE 2  Identification of hypoxia and tumor microenvironment status and their association with response to radiotherapy in the TCGA cohort. (A) Heatmap showing the expression pattern of hypoxia gene sets in two clusters. (B) t-SNE plots of all patients separated into two clusters labeled by color. (C) Expression changes of genes involved in the HIF1 KEGG pathway (high hypoxia group vs. low hypoxia group). (D) Heatmap showing quantification of eight immune and two stromal populations estimated by MCP-counter. (E) Kaplan–Meier curves of disease-specific survival for patients in two hypoxia groups. (F) Kaplan–Meier curves of disease-specific survival for patients in two immunity groups. (G) Kaplan–Meier curves of disease-specific survival for patients in three hypoxia-immunity groups.
upregulated genes associated with hypoxia in HNSCC, cluster 1 and 2 were annotated as the “low hypoxia” and “high hypoxia” groups, respectively. We also generated a t-SNE plot colored by different clusters to visualize the relevance of the clusters, which separated them into two discrete groups (Figure 2B). In addition, we further analyzed the expression changes of the 26 genes involved in the HIF1 signaling pathway, of which 14 genes were differentially expressed between the two groups (Figure 2C). Six genes (85.7%) associated with increased oxygen delivery and five genes (71.4%) associated with reduced oxygen consumption were upregulated in the “high hypoxia” group in response to hypoxia. The differences in the expression of genes associated with the HIF1 pathway further validates the hypoxia status. The “low hypoxia” group showed a significantly better DSS than the “high hypoxia” group (log-rank test, \( P = 0.023 \)) (Figure 2E). The differences in survival suggest that patients with lower hypoxia status responded better to radiotherapy.

The tumor microenvironment status of the patients from the TCGA cohort was inferred from the immune scores calculated by using the ESTIMATE algorithm, which estimated the degree of immune cell infiltration in tumors. According to the cutoff value of the immune score determined by maximally selected log-rank statistics, 119 and 155 patients from the TCGA cohort were divided into “high immunity” and “low immunity” groups, respectively (Figure S1). As estimated by MCP-counter, the “high immunity” group showed a significantly higher abundance of eight immune cell populations, namely, T cells, CD8 T cells, cytotoxic lymphocytes, B lineage, NK cells, monocytic lineage cells, and myeloid dendritic cells, while both groups shared similar distribution patterns for the abundance of neutrophils, endothelial cells and fibroblasts (Figure 2D). The Kaplan–Meier curve showed that the DSS of the “high immunity” group was better than that of the “low immunity” group (\( P = 0.013 \)); thus, the higher immune status was associated with better response to radiotherapy (Figure 2F).

### 3.2 IHR subtypes of the TCGA cohort

By combining the tumor immune microenvironment and hypoxia status of the patients from the TCGA cohort into a two-dimensional index, we further divided the patients into “high immunity and low hypoxia,” “low immunity and high hypoxia,” and “mixed” groups, which included 87, 94, and 93 patients, respectively. The Kaplan–Meier curve (Figure 2G) showed a statistically significant survival difference only between “high immunity and low hypoxia” and “low immunity and high hypoxia” groups (\( P = 0.003 \)). These survival differences indicate that a combination of hypoxia and immune status can better stratify patients in terms of their response to radiotherapy.

### 3.3 Construction and validation of the IHRGPI

A total of 274 eligible patients from the TCGA cohort were included in the training set to train an IHRGPI model to predict the response to radiotherapy. The clinicopathological characteristics of the patients in the TCGA cohort are presented in Table S1. Locoregionally advanced HNSCC patients predominated in the TCGA cohorts, with the stage III-IV HNSCC patients accounting for 76.6%. The GEO cohorts were used as the validation set. Tables S2 and S3 demonstrate the clinicopathological characteristics of the 52 and 100 patients in GEO cohort-1 and GEO cohort-2, respectively. The groups showed no significant differences in age and sex. The mean age (SD) was 59.18 (10.64), 60.02 (7.92) and 57.05 (11.29) years in the TCGA cohort, GEO cohort-1, and GEO cohort-2, respectively, demonstrating similar distributions (\( p = 0.15 \)), while the proportion of female patients was 21.5%, 21.2%, and 28.0%, respectively, in three cohorts. Patients with locoregionally advanced HNSCC also predominated in GEO cohort-1 and cohort-2, accounting for 75% and 90% of the cohorts, respectively.

A total of 2665 IHR DEGs were identified between the “high immunity and low hypoxia” and “low immunity and high hypoxia” groups from the TCGA cohort, of which 311 prognostic DEGs were significant independent prognostic factors of DSS in the univariate Cox regression. By intersecting the 324 prognostic DEGs with the hallmark hypoxia gene set from MSigDb and the immune-related gene set from InnateDB, five IHRGs (BCL2, SERPINE1, CAV1, CXCR4, and F3) were finally included in the multivariate Cox regression model (Figure 3A). We constructed a model for each patient, which was calculated by the formula

\[
\text{IHRGPI} = \text{BCL2} \times (0.373) + \text{SERPINE1} \times (0.225) + \text{CAV1} \times (0.077) + \text{CXCR4} \times (0.026) + \text{F3} \times (0.113).
\]

Univariate Cox regression showed that HRGPI and HPV status were independent prognostic factors of DSS, while only IHRGPI was an independent prognostic factor in multivariate Cox regression (Figure 3B,C).

With 0 as the cutoff value for the IHRGPI, we dichotomized the patients from the TCGA cohort into the radiosensitive and radioresistant groups, which included 123 and 151 patients, respectively. After 1:1 matching, 105 patients remained in each group (Table S4). Kaplan–Meier survival analysis showed a significantly better DSS in the radiosensitive group (unmatched cohort: \( p = 0.005 \); matched cohort: \( p = 0.032 \)) (Figure 3D). Univariate Cox regression showed that the hazard ratio (HR) of the
FIGURE 3 Construction and validation of the IHRGPI model. (A) Venn diagram showing the intersection between the immune-hypoxia-related prognostic differentially expressed genes, hypoxia gene sets, and immune gene sets. (B) Forest plot demonstrating the univariate Cox regression of IHRGPI and common clinicopathological characteristics. (C) Forest plot demonstrating the multivariate Cox regression of IHRGPI and the common clinicopathological characteristics. (D) Kaplan–Meier curves for disease-specific survival in the TCGA cohort. (E) Kaplan–Meier curves for disease-specific survival in GEO cohort-1. (F) Kaplan–Meier curves for progression-free survival in GEO cohort-1. (G) Kaplan–Meier curves for recurrence-free survival in GEO cohort-2. (H) Kaplan–Meier curves for overall survival in TCGA cohort. IHRGPI, immune-hypoxia-related gene prognostic index; DSS, disease-specific survival; PFS, progression-free survival; RFS, recurrence-free survival; OS, overall survival.
radioresistant group over the radiosensitive group was 2.18 (95% CI: 1.25–3.80; \( p = 0.007 \)) and 2.00 (95% CI: 1.05–3.79; \( p = 0.036 \)) in the unmatched and matched TCGA cohorts, respectively. Next, we applied the IHRGPI calculation to the GEO cohort-1 with DSS as an endpoint for validation. Twenty-eight and twenty-four patients were divided into the radiosensitive and radioresistant groups, respectively, with 0 as the cutoff value of the IHRGPI. After 1:1 matching, 20 patients remained in each group (Table S5). Kaplan–Meier survival analysis demonstrated a better DSS in the radiosensitive group in the matched cohort (\( p = 0.019 \)) (Figure 3E), with a corresponding HR of 5.28 (95% CI: 1.12–24.97; \( p = 0.036 \)). Finally, the survival of the radiosensitive group remained significantly superior in the Kaplan–Meier analysis when the endpoint adopted was OS (\( p = 0.007 \)) in the TCGA cohort, PFS (\( p = 0.048 \)) in GEO cohort-1, and RFS (\( p = 0.02 \)) in GEO cohort-2 (Figure 3F,H).

### 3.4 Hypoxic, tumor-microenvironmental, and molecular signatures in the two IHRGPI groups

A total of 2675 radiosensitivity-associated DEGs, including 1490 upregulated genes and 1185 downregulated genes, were identified between the radiosensitive and radioresistant groups (Figure S2). In the GO enrichment analysis (Supplementary Table), the radiosensitivity-associated DEGs were mainly enriched in immune-related activities, including complement activation, humoral immune response, immunoglobulin-mediated immune response, B-cell-mediated immunity (Figure 4A,B). In the KEGG enrichment analysis, the DEGs were mainly enriched in cytokine–cytokine receptor interaction, retinol metabolism, and hemopoietic cell lineage (Figure S3). The association between the expression changes of each IHRG and different GO terms is demonstrated in Figure S4. Next, we performed GSEA with GEO and hallmark gene sets from MSigDb. The gene sets enriched in the radiosensitive group were mainly involved in immune-related activities, including antigen-binding, B cell and T cell-associated immune response, immunoglobulin-mediated immune response, and complement activation (Figure 4C). In contrast, the hypoxia, angiogenesis, epithelial-mesenchymal transition, and TNFα signaling genes via \( NfκB \) were mainly enriched in the radioresistant group (Figure 4D).

In evaluations of the intergroup differences in immune characteristics, the ESTIMATE immune score was significantly higher in the radiosensitive group (Figure S5). In analyses using Cibersort to assess the immune landscape in both groups, memory B cells, naïve B cells, plasma cells, CD8 T cells, follicular helper T cells, and regulatory T cells were more abundant in the radiosensitive group, while M0 macrophages and activated mast cells were more abundant in the radioresistant group (Figure 5A,B).

We also analyzed the mutational signatures and identified the top 10 mutated genes of both groups (Figure 5C,D). Missense mutations predominated in both groups (Figure S6). \( TP53, TTN, SYNE1, \) and \( FAT1 \) were present among the top 10 mutated genes of both groups. The mutation rates of \( CDKB2A, DNAH5, \) and \( NOTCH1 \) were higher in the radioresistant group, while those of \( MUC16, PIK3CA, CSMD3, \) and \( FLG \) were higher in the radiosensitive group. The TMB did not differ significantly between the two groups (\( p = 0.33 \)) (Figure S7). Finally, the top 10 transcriptional factors associated with the five selected IHRGs using the ChEA3 tool are listed in Table S6, including \( FOXD1, ZNF469, NR3C1, PPARG, SMAD3, MEOX1, BCL6B, HIF1A, \) \( ELK3, \) and \( FOSL1, \) of which \( HIF1A \) was elevated significantly in the radioresistant group in the TCGA cohort (\( p < 0.01 \)), GEO cohort-1 (\( p = 0.01 \)) and GEO cohort-2 (\( p < 0.01 \)).

### 3.5 Treatment strategies

To explore potentially effective combinations of radiotherapy and other treatment strategies, we calculated the TIDE score for each patient. The TIDE score was significantly lower in the radiosensitive group (\( p = 0.019 \)), indicating that a better response to radiotherapy correlates with a better response to immunotherapy (Figure 6A). Moreover, the proportion of predicted responders to immune checkpoint inhibitors was higher in the radiosensitive group (37.4%) than in the radioresistant group (27.2%). However, there was no statistically significant intergroup difference in the TMB. As for the differences in the expression levels of the targets of well-studied immune checkpoint inhibitors, 

In the expression levels of the targets of well-studied immune checkpoint inhibitors was higher in the radioresistant group (37.4%) than in the radioresistant group (27.2%). However, there was no statistically significant intergroup difference in the TMB. As for the differences in the expression levels of the targets of well-studied immune checkpoint inhibitors, 

The mutation rates of \( CDKB2A, DNAH5, \) and \( NOTCH1 \) were higher in the radioresistant group, while those of \( MUC16, PIK3CA, CSMD3, \) and \( FLG \) were higher in the radiosensitive group. The TMB did not differ significantly between the two groups (\( p = 0.33 \)) (Figure S7). Finally, the top 10 transcriptional factors associated with the five selected IHRGs using the ChEA3 tool are listed in Table S6, including \( FOXD1, ZNF469, NR3C1, PPARG, SMAD3, MEOX1, BCL6B, HIF1A, \) \( ELK3, \) and \( FOSL1, \) of which \( HIF1A \) was elevated significantly in the radioresistant group in the TCGA cohort (\( p < 0.01 \)), GEO cohort-1 (\( p = 0.01 \)) and GEO cohort-2 (\( p < 0.01 \)).

### 4 DISCUSSION

Radiotherapy is one of the three major treatment modalities for HNSCC, and establishing models to predict
patients’ radiotherapy response can facilitate individualized treatment decisions. Although several predictive models have been reported,\(^6\)\(^–\)\(^9\) most of them were purely statistically driven without a radiobiological foundation, so they were difficult to explain. Intrinsic radiosensitivity is not the only determinant factor of the response to

| ID          | Description                                                                 |
|-------------|-----------------------------------------------------------------------------|
| GO:0006958  | complement activation, classical pathway                                     |
| GO:0002455  | humoral immune response mediated by circulating immunoglobulin               |
| GO:0006956  | complement activation                                                        |
| GO:0016064  | immunoglobulin mediated immune response                                      |
| GO:0006959  | humoral immune response                                                      |
| GO:0019724  | B cell mediated immunity                                                     |
| GO:0002449  | lymphocyte mediated immunity                                                 |
| GO:0002377  | immunoglobulin production                                                    |

FIGURE 4 Functional enrichment analysis. (A) GOCircle plot showing the top eight GO terms based on the adjusted \(p\)-values. (B) GOcluster plot showing the clustering of GO terms based on Euclidean distances and average linkage. (C) Gene sets enriched in the radiosensitive group by GSEA. (D) Gene sets enriched in the radioresistant group by GSEA. GSEA, gene set enrichment analysis.
FIGURE 5  The immune and mutational landscape. (A) Boxplot comparing the proportions of different types of immune cells in tumor tissues estimated by Cibersort between the two IHRGPI groups (ns: not significant, *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001). (B) 100% stacked bar plot comparing the proportion of different type of immune cells in each patient between the two IHRGPI groups. (C) Top 10 mutated genes and their mutational types in the radiosensitive group. (D) Top 10 mutated genes and their mutational types in the radioresistant group. IHRGPI, immune-hypoxia-related gene prognostic index
radiotherapy; the tumor immune microenvironment and hypoxia also play important roles in this process. Reoxygenation and reactivation of the immune system are among the well-known traditional “5 Rs” and the newly proposed “6th R”, respectively. Thus, we developed the IHRGPI to predict HNSCC patients’ response to radiotherapy based on the tumor immune microenvironment and hypoxia signature by using the transcriptomic data in the TCGA database. The IHRGPI was proven to be a valid predictor of the radiosensitivity of HNSCC, showing better survival in the radiosensitive group after treatment with radiotherapy in both TCGA and GEO cohorts.

Five IHRGs were included in the IHRGPI. Among them, BCL2 is an anti-apoptotic protein. However, previous studies on the effect of BCL2 on have yielded contradictory findings, precluding its validation as a reliable biomarker of prognosis or radiosensitivity. Theoretically, increased expression of BCL2, an apoptosis suppressor gene, will block apoptosis and lead to resistance to chemoradiotherapy. However, higher BCL2 expression was instead associated with a better response to radiotherapy in this study. These contradictory findings can be attributed to multiple mechanisms. Because apoptosis is regulated by a complex network, pro-apoptotic signals should be taken into consideration while interpreting these findings. In our study, the ratio of BCL2 to BAX, a well-known pro-apoptotic gene, was not significantly different between the two IHRGPI groups in the TCGA cohort (p = 0.99), indicating a balance between pro-apoptotic and anti-apoptotic signals. On the other hand, BCL2 overexpression was reported to block DNA double-strand break repair via non-homologous end-joining and enhance radiosensitivity.
SERPINE1 also plays an important role in the HIF1 pathway by increasing oxygen delivery in response to hypoxia, and was also included in previously proposed gene signatures to predict locoregional tumor control for HNSCC patients treated with postoperative radiotherapy or the response to neo-adjuvant chemoradiotherapy before surgery for locally advanced esophageal cancer. CAV1 encodes the main component of the lipid domains on plasma membranes known as caveolae. Although it promoted radioresistance in cancer cell lines, loss of its stromal expression was also reported to mediate the resistance of prostate cancer to radiotherapy. In this study, upregulated CAV1 was associated with a better response to radiotherapy. Since bulk RNA sequencing of HNSCC tissue provides the average gene expression of both tumor cells and stromal cells, our result may mainly reflect the influence of stromal CAV1 expression on the radiosensitivity of HNSCC, and methodologies with higher resolution may better elucidate our findings. CXCR4 is a specific chemokine receptor for stromal cell-derived factor-1 that is expressed on cancer cells and many other cells in the tumor microenvironment, and has many important functions including promoting tumor progression and metastasis and facilitating angiogenesis. Although post-irradiation CXCR4-mediated angiogenesis is associated with radioresistance and recurrence after the radiotherapy through rescuing the damaged tumor vasculature caused by radiation, upregulated CXCR4 was associated with better response to radiotherapy in this study. CXCR4-mediated angiogenesis may improve the hypoxia status of the tumor in the pre-treatment sample and thereby sensitize the cancer cells to radiation. F3 encodes coagulation factor III, a cell-surface glycoprotein that can initiate blood coagulation cascade in the extrinsic coagulation pathway, and has also been reported to participate in tumor angiogenesis, but its association with radioresistance remains to be elucidated.

For all five genes, HIF1A was one of the top 10 transcriptional factors in the transcription factor enrichment analysis by ChEA3. HIF1A is a master regulator of the cellular and systematic response to hypoxia. HIF1A expression was elevated significantly in the radioresistant group in the three cohorts. GSEA also showed that hypoxia and angiogenesis were mainly enriched in the radioresistant group.

The complex interactions between and synergistic effects of radiotherapy and anticancer immunity have been extensively explored and elucidated, and reactivation of the immune system was lately proposed as the “6th R”. In this study, patients in “high immunity” group showed better response to radiotherapy than those in “low immunity” group. Thus, our results further confirmed the role of immunity in radiotherapy. Many immune-related activities, including complement activation and humoral immune response, were also significantly enriched between the two IHRGPI groups in both GO and GSEA functional enrichment analysis, and CD8+ T cells, the most powerful effector cells in the anticancer immune response, were more abundant in the tumor microenvironment of HNSCC patients in both the “high immunity” and radioresistant groups in comparison with the “low immunity” and radioresistant groups, respectively. Taken together, these findings indicated that anticancer immunity did play an important role in HNSCC patients’ response to radiotherapy, and an immune-active tumor microenvironment favored a better response to radiotherapy among HNSCC patients.

Advances in radiotherapy techniques, such as the advent of IMRT, have facilitated significantly better dose distribution and constraint. However, they did not significantly improve prognosis in most studies. Therefore, we sought to identify an optimal combinational treatment strategy for patients with different radiosensitivity. Our results suggest that radioresistant patients might benefit more from chemotherapy, while radiosensitive patients might benefit more from immunotherapy. Therefore, for patients who receive radiotherapy and have lower IHRGPI, immunotherapy could be recommended, and chemotherapy might be omitted or de-intensified, while chemotherapy could not be omitted in those with high IHRGPI. Notably, although the TIDE scores suggested that the radiosensitive group generally responded better to the immunotherapy, the expression levels of B7H3 and CD73 were higher in the radioresistant group, indicating that the radioresistant group might benefit more from B7H3 and CD73 inhibitors.

This study had several limitations. First, we did not evaluate the effectiveness of radiotherapy in the two IHRGPI groups among all patients, and we only evaluated the prognostic significance of the IHRGPI groups among all patients receiving radiotherapy in each cohort. For non-metastatic HNSCC, surgery remains the mainstay of treatment strategies, while radiotherapy is usually adjuvant therapy for high-risk postoperative patients. Therefore, patients receiving radiotherapy usually bear more factors indicating a poor prognosis. Because the indications for adjuvant radiotherapy are complicated, and information regarding the rationale for treatment decisions is insufficient, evaluating the effectiveness of radiotherapy in two IHRGPI groups among all patients respectively will introduce more confounding factors that could not be adjusted. Second, we did not explore the effectiveness of the combination of radiotherapy and targeted therapy, since that reliable biomarkers for the response to cetuximab, an epithelial growth factor receptor inhibitor commonly used in treating HNSCC, are not yet been available.
In conclusion, we developed and validated the IHRGPI in HNSCC patients, which was based on the tumor microenvironment and hypoxia signature and can facilitate prediction of the response to radiotherapy. We also explored the potential mechanisms underlying the differences between two IHRGPI groups. Moreover, our model provides individualized suggestions for combinational treatment strategies involving different IHRGPI groups.

AUTHOR CONTRIBUTIONS
GLZ, KBY, and CX have made contributions to the design of the work, the acquisition and analysis of data, drafting and revision of the work. RJF and WFL have made contributions to the design and revision of the work. JM have made contributions to the design and revision of the work, supervision of the work and funding acquisition. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS
The results published here are in part based on data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

CONFLICT OF INTEREST
The authors declare that they have no competing interests.

DATA AVAILABILITY STATEMENT
The data analyzed in this study were obtained from the Cancer Genome Atlas (TCGA) at TCGA-HNSC and Gene Expression Omnibus (GEO) at GSE117973 and GSE39366.

ETHICS STATEMENT
The requirements of patients’ informed consent and institutional review board approval were waived given the datasets were de-identified and publicly available.

ORCID
Guang-Li Zhu https://orcid.org/0000-0002-5895-1349
Cheng Xu https://orcid.org/0000-0002-4503-6029
Wen-Fei Li https://orcid.org/0000-0001-7605-3609
Jun Ma https://orcid.org/0000-0002-1137-9349

REFERENCES
1. Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. Int J Cancer. 2019;144(8):1941-1953.
2. Gujral DM, Nutting CM. Patterns of failure, treatment outcomes and late toxicities of head and neck cancer in the current era of IMRT. Oral Oncol. 2018;86:225-233. https://www.sciencedirect.com/science/article/pii/S1368837518303324
3. Steel GG, McMillan TJ, Peacock JH. The 5Rs of radiobiology. Int J Radiat Biol. 1989;56(6):1045-1048.
4. Rockwell S, Dobrucki LT, Kim EY, Morrison ST, van Vu T. Hypoxia and radiation therapy: past history, ongoing research, and future promise. Curr Mol Med. 2009;9(4):442-458.
5. Boustanli J, Grapin M, Laurent P-A, Apestoh L, Mirjolet C. The 6th R of radiobiology: reactivation of anti-tumor immune response. Cancer. 2019;11(6):860. https://www.mdpi.com/2072-6694/11/6/860
6. Eschrich SA, Pramana J, Zhang H, et al. A gene expression model of intrinsic tumor radiosensitivity: prediction of response and prognosis after chemoradiation. Int J Radiat Oncol Biol Phys. 2009;75(2):489-496. https://www.sciencedirect.com/science/article/pii/S0360301609009195
7. Kim SI, Kang JW, Noh JK, et al. Gene signature for prediction of radiosensitivity in human papillomavirus-negative head and neck squamous cell carcinoma. Radiat Oncol J. 2020;38(2):99-108. https://pubmed.ncbi.nlm.nih.gov/33012153/
8. Foy J-P, Bazire L, Ortiz-Cuaran S, et al. A 13-gene expression-based radioresistance score highlights the heterogeneity in the response to radiation therapy across HPV-negative HNSCC molecular subtypes. BMC Med. 2017;15(1):165.
9. Liu J, Han M, Yue Z, et al. Prediction of radiosensitivity in head and neck squamous cell carcinoma based on multiple omics data. Front Genet. 2020;11:960.
10. Colaprico A, Silva TC, Olsen C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. Nucleic Acids Res. 2016;44(8):ec71.
11. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401-404.
12. Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal. 2013;6(269):pl1.
13. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30(1):207-210. https://academic.oup.com/nar/article/30/1/207/1332640
14. Walter V, Yin X, Wilkerson MD, et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. PLOS One. 2013;8(2):e56823.
15. Winter SC, Buffa FM, Silva P, et al. Relation of a hypoxia meta-gene derived from head and neck cancer to prognosis of multiple cancers. Cancer Res. 2007;67(7):3441-3449. https://pubmed.ncbi.nlm.nih.gov/17409455/15
16. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550.
17. Liberezon A, Subramanian A, Pincheck B, Thorvaldsdottir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739-1740. https://academic.oup.com/bioinformatics/article/27/12/1739/257711
18. Ritchie ME, Hipson B, Di W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47. https://academic.oup.com/nar/article/43/7/e47/2414268
19. Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28(1):27-30.
20. Yoshihara K, Shahmoradgoli M, Martinez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.

21. Becht E, Giraldo NA, Lacroix L, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol*. 2016;17(1):218. https://genomembiology.biomedcentral.com/articles/10.1186/s13059-016-1070-5

22. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics J Integr Biol*. 2012;16(5):284-287.

23. Newman AM, Steen CB, Liu CL, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat Biotechnol*. 2019;37(7):773-782. https://www.nature.com/articles/s41587-019-0114-2

24. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maptools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756.

25. Keenan AB, Torre D, Lachmann A, et al. ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res*. 2019;47(W1):W212-W224. https://academic.oup.com/nar/article/47/W1/W212/549769

26. Jiang P, Gu S, Pan D, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med*. 2018;24(10):1550-1558.

27. Marin-Acevedo JA, Kimbrough EO, Lou Y. Next generation of immune checkpoint inhibitors and beyond. *J Hematol Oncol*. 2021;14(1):45. https://jhoonline.biomedcentral.com/articles/10.1186/s13045-021-01056-8/tables/1

28. Qin S, Xu L, Yi M, Yu S, Wu K, Luo S. Novel immune checkpoint targets: moving beyond PD-1 and CTLA-4. *Mol Cancer*. 2019;18(1):155. https://molecular-cancer.biomedcentral.com/articles/10.1186/s12943-019-1091-2

29. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLOS One*. 2014;9(9):e107468.

30. Wilson GD, Saunders M, Dische S, Richman P, Daley F, Bentzen SM. Bcl-2 expression in head and neck cancer: an enigmatic prognostic marker. *Int J Rad Oncol Biol Phys*. 2001;49(2):435-441.

31. Oing C, Tennstedt P, Simon R, et al. BCL2-overexpressing prostate cancer cells rely on PARP1-dependent end-joining and are sensitive to combined PARP inhibitor and radiation therapy. *Cancer Lett*. 2018;423:60-70. https://www.sciencedirect.com/science/article/pii/S0304383518301964

32. Hwang K-T, Kim K, Chang JH, et al. BCL2 regulation according to molecular subtype of breast cancer by analysis of the cancer genome atlas database. *Cancer Research and Treatment*. 2018;50(3):658-669.

33. Wang Q, Gao F, May WS, Zhang Y, Flagg T, Deng X. Bcl2 negatively regulates DNA double-strand-break repair through a nonhomologous end-joining pathway. *Mol Cell*. 2008;29(4):488-498.

34. Pavón MA, Arroyo-Solera I, Céspedes MV, Casanova I, León X, Mangues R. uPA/uPAR and SERPIN1E1 in head and neck cancer: role in tumor resistance, metastasis, prognosis and therapy. *Oncotarget*. 2016;7(35):57351-57366. https://www.oncotarget.com/articlenotices/1034/text/#bib79

35. Schmidt S, Linge A, Zwanenburg A, et al. Development and validation of a gene signature for patients with head and neck carcinomas treated by postoperative radio(chemo)therapy. *Clin Cancer Res*. 2018;24(6):1364-1374.

36. Wang J, Yu P, Luo J, Sun Z, Yu J, Wang J. Transcriptomic and microRNA expression profiles identify biomarkers for predicting neo-chemoradiotherapy response in esophageal squamous cell carcinomas (ESCC). *Front Pharmacol*. 2021;12:626972.

37. Codenotti S, Marampon F, Triggiani L, et al. Caveolin-1 promotes radioresistance in rhabdomyosarcoma through increased oxidative stress protection and DNA repair. *Cancer Lett*. 2021;505:1-12.

38. Panic A, Kettele J, Reis H, et al. Progression-related loss of stromal caveolin 1 levels fosters the growth of human PC3 xenografts and mediates radiation resistance. *Sci Rep*. 2017;7:41138.

39. Santagata S, Ieranò C, Trotta AM, et al. CXCR4 and CXCR7 signaling pathways: a focus on the cross-talk between cancer cells and tumor microenvironment. *Front Oncol*. 2021;11:59138.

40. Martin BJ. Inhibiting vasculogenesis after radiation: a new paradigm to improve local control by radiotherapy. *Semin Radiat Oncol*. 2013;23(4):281-287.

41. Brown JM. Vasculogenesis: a crucial player in the resistance of solid tumours to radiotherapy. *Br J Radiol*. 2014;87(1035):20130686.

42. Belting M, Dorrell MI, Sandgren S, et al. Regulation of angiogenesis by tissue factor cytoplasmic domain signaling. *Nat Med*. 2004;10(5):502-509. https://www.nature.com/articles/nm1037

43. Hisada Y, Mackman N. Tissue factor and cancer: regulation, tumor growth, and metastasis. *Semin Thromb Hemost*. 2019;45(4):385-395.

44. Semenza GL. Oxygen sensing, homeostasis, and disease. *N Engl J Med*. 2011;365(6):537-547.

45. Herrera FG, Bourhis J, Coukos G. Radiotherapy combination opportunities leveraging immunity for the next oncology paradigm. *CA Cancer J Clin*. 2017;67(1):65-85.

46. Raskov H, Orhan A, Christensen JP, Gögenur I. Cytotoxic CD8+ T cells in cancer and cancer immunotherapy. *Br J Cancer*. 2021;124(2):359-367.

**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:** Zhu G-L, Yang K-B, Xu C, Feng R-J, Li W-F, Ma J. Development of a prediction model for radiotherapy response among patients with head and neck squamous cell carcinoma based on the tumor immune microenvironment and hypoxia signature. *Cancer Med*. 2022;11:4673-4687. doi: 10.1002/cam4.4791