Original Article

A novel machine learning-driven immunogenic cell death signature for predicting ovarian cancer prognosis

Yali Wang[1], Peng Zhao[2], Xude Sun[3], Felipe Batalini[4], Gabriel Levin[5,6], Hooman Soleymani majd[7]^, Hao Chen[8], Tingting Gao[9]

[1]Department of Obstetrics and Gynecology, Maternal and Child Health Center in Fuping County, Fuping, China; [2]Oncology Department, Xi'an Daxing Hospital, Xi'an, China; [3]Department of Anesthesia, The Second Affiliated Hospital of Air Force Medical University, Xi'an, China; [4]Department of Medical Oncology, Mayo Clinic Arizona, Phoenix, AZ, USA; [5]Division of Gynecologic Oncology, Jewish General Hospital, McGill University, Montreal, QC, Canada; [6]The Department of Gynecologic Oncology, Hadassah Medical Center, Faculty of Medicine, Hebrew University, Jerusalem, Israel; [7]Oxford University Hospitals NHS Foundation Trust, Department of Gynaecology Oncology, Churchill Hospital, Oxford, United Kingdom; [8]Department of Thoracic Surgery, Tangdu Hospital of Air Force Military Medical University, Xi'an, China; [9]Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Air Force Medical University, Xi'an, China

Contributions: (I) Conception and design: Y Wang, T Gao; (II) Administrative support: Y Wang, T Gao; (III) Provision of study materials or patients: P Zhao, H Chen; (IV) Collection and assembly of data: X Sun, H Chen; (V) Data analysis and interpretation: T Gao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Tingting Gao, MD. Department of Obstetrics and Gynecology, The Second Affiliated Hospital of Air Force Medical University, 569 Xinsi Road, Xi'an 710000, China. Email: ttgao1981@126.com.

Background: Ovarian cancer (OC) is one of the most lethal malignancies in women, primarily due to the absence of reliable predictive biomarkers and effective therapies. The complex role of immunogenic cell death (ICD) in OC remains poorly understood, despite its critical implications for enhancing immune responses against tumors. We are committed to developing and validating a novel ICD-related gene signature and producing certain guiding value for the clinical treatment of OC.

Methods: We employed single-sample gene set enrichment analysis (ssGSEA) and weighted gene coexpression network analysis (WGCNA) on The Cancer Genome Atlas (TCGA)-ovarian carcinoma dataset to identify ICD-associated genes. A combination of 10 different machine learning approaches was used to construct an ICD-related signature (ICDRS), which was then validated across multiple datasets. The model's predictive power was integrated into a clinical nomogram to predict patient outcomes. Ultimately, we assessed the reaction of various risk subgroups to screen pharmaceuticals designed to address specific risk factors in the context of personalized medicine.

Results: We identified 72 prognostic genes related to ICD. An unanimous ICDRS was developed using a 101-combination machine learning computational structure, demonstrating outstanding predictive accuracy for prognosis and clinical use. Patients categorized as low ICDRS varied from those of high ICDRS in terms of biological processes, mutation profiles, and immune cell penetration in the tumor microenvironment. In addition, potential medications that target specific subgroups at risk were identified.

Conclusions: The ICDRS presents a significant advancement for prognostication of patients with OC, facilitating refined predictions and the exploration of personalized treatment pathways. Prospective clinical trials are necessary to validate its clinical utility and expand the application of this model to other cancer types.

Keywords: Machine learning; ovarian cancer (OC); tumor microenvironment

Submitted Jan 14, 2025. Accepted for publication Feb 18, 2025. Published online Feb 26, 2025.
doi: 10.21037/tcr-2025-118
View this article at: https://dx.doi.org/10.21037/tcr-2025-118

^ ORCID: 0000-0002-7603-9986.

Introduction

Among gynecological malignancies, ovarian cancer (OC) is the leading cause of cancer-related death in women (1). Standard therapies, which include surgical removal and platinum-based chemotherapy, result in a 5-year survival rate of less than 50% for patients diagnosed with OC due to frequent recurrence and chemoresistance (2). Furthermore, despite the use of immune checkpoint inhibitors (ICIs) in the treatment of various cancers, these drugs have not been approved for OC treatment due to the lack of efficacy (3). Differences in the tumor immune microenvironment across patients and different cancers constitute one of the key factors limiting the widespread success of immunotherapies and thus represents a compelling rationale for the identification of predictable predictive biomarkers for improving personalized treatment options (4-6).

The Nomenclature Committee for Cell Death (NCCA) has updated guidelines to include twelve types of programmed cell death, among which immunogenic cell death (ICD) is recognized as a form that both initiates and modulates adaptive immunity (7). Cancer cells undergoing ICD exhibit elevated levels of specific proteins (e.g., heat shock proteins), actively secrete metabolites and cytokines (e.g., adenosine triphosphate and interleukin-6), and passively release soluble macromolecules (e.g., ribonucleic acid) (8). These cells also show increased expression of cytoskeleton-related proteins such as F-actin, which impacts the tumor microenvironment. These molecules, known as damage-associated molecular patterns (DAMPs), transform cancer cells from an immunologically dormant state to an immunostimulatory one by acting as ligands for pattern recognition receptors (PRRs) (9).

Subsequently, the interaction of these substances with PRRs enhances the recruitment and activation of antigen-presenting cells, subsequently triggering specific immune responses against tumors (10). ICD is predominantly triggered by conventional chemotherapeutic agents, radiotherapy, targeted therapies, and specific intracellular parasites (11). Research indicates that combining chemotherapy with immunotherapy enhances the immunogenic properties of tumor cells, leading to improved antitumor immune responses and significantly better therapeutic outcomes (12). However, the prognostic relevance of ICD markers in OC remains underexplored, with the associated mechanisms still being poorly understood.

Advancements in bioinformatics have led to the discovery of numerous gene signatures with potential prognostic value (13). Despite this progress, the clinical integration of these multigene signatures faces challenges such as improper application of machine learning methods, inadequate validation to represent unique and diverse patient cohorts, which ultimately leads to a lack of implementation. Addressing these issues may involve developing novel markers with improved predictive accuracy or devising a comprehensive evaluation tool that consolidates various prognostic indicators. In this study, we used artificial intelligence methods to identify and validate an ICD-related gene signature. This signature can assess prognosis, predict responses to immunotherapy and chemotherapy, and has the potential to inform clinical decision-making for patients with OC. We present this article in accordance with the TRIPOD reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-2025-118/rc).

Highlight box

Key findings

- We developed and validated an immunogenic cell death (ICD)-related signature (ICDRS) for ovarian cancer (OC), incorporating advanced machine learning techniques.
- Our signature can provide a new perspective for tailoring therapeutic strategies for patients with OC.

What is known and what is new?

- Many researches have reported the prognostic relevance of ICD markers in OC.
- Our study used a 101-combination machine learning computational structure to identify and validate an ICD-related gene signature.

What is the implication, and what should change now?

- This signature is not only helpful in providing a refined prognosis but also has potential to improve outcomes by tailoring therapeutic strategies. Prospective validation with clinical trials is needed to confirm clinical utility of the ICDRS across various clinical scenarios. Additionally, expanding the application of the ICDRS to other types of cancer could provide insights into its broader applicability and impact. It is crucial to integrate the ICDRS into standard clinical practice, in order to ensure that it contributes to more precise and effective cancer treatments.

Methods

Data acquisition and processing

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study sourced OC data from two primary databases: The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/)and

the Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). The details of the dataset characteristics, selection criteria, and processing methodologies are as follows:

(I) GEO dataset (GSE73614): this dataset has rich prognostic information, comprising107 patient records of OC. It includes messenger RNA (mRNA) expression and oncological outcomes, such as survival time and treatment response. The samples were obtained from tissues that constitute the primary tumors, while extraction and sequencing of RNA were conducted in line with uniform laboratory procedures to ensure quality data.

(II) TCGA dataset: transcriptome data with associated clinical information were extracted from TCGA for a cohort of patients with OC which includes the patient demographics, tumor stage, histological type, and survival outcomes. Somatic mutation information was also downloaded in the Mutation Annotation Format (MAF), which supplies information regarding mutation type, location, and associated gene impact.

(III) Data processing: data processing included normalization and cleaning, gene expression data handling, quality control, and analytical preparation. (I) For normalization and cleaning, we employed the "sva" package in R (The R Foundation for Statistical Computing) for batch effect mitigation to standardize data across different sequencing platforms and batches. This step was crucial for removing nonbiological variability and ensuring comparability across sample. (II) In gene expression data handling, the raw gene expression values were processed with several steps of pretreatment and included the normalization via the method robust multiarray average (RMA). Subsequently, the signals of the probes mapping to the same gene were averaged according to recent gene annotation databases. (III) For quality control, rigorous quality control measures were applied, including the exclusion of any samples with incomplete data or those that did not meet specific quality thresholds such as RNA integrity numbers (RINs). (IV) Finally, for analytical preparation, prior to analysis, data were reviewed for outliers, and the completeness of the metadata was verified. The prepared dataset was then split into training and validation subsets to facilitate the development and testing of the machine learning models, ensuring a robust evaluation of the ICD-related signature (ICDRS).

Single-sample gene set enrichment analysis (ssGSEA)

ssGSEA is commonly employed to determine the enrichment score of particular gene sets within individual samples, indicating whether these gene sets are predominantly upregulated or downregulated. For each TCGA-OC sample, the ICD score was computed using the ssGSEA function provided by the R package "GSVA".

Weighted gene coexpression network analysis (WGCNA)

WGCNA is used as a systematic biological approach to investigating the pattern of gene relationships across different samples, which aids in identifying highly interconnected gene clusters. The "WGCNA" R package was used to conduct such an analysis in this study. Analyses were initiated via the calculation of an optimal soft thresholding parameter, β, and a scale-free network was subsequently established. The weighted adjacency matrix was then transformed into a topological overlap matrix (TOM), and dissimilarity calculations was performed using dissimilarity TOM (dissTOM). Following this, a dynamic tree cut algorithm was used for gene clustering and module identification, with the module most closely associated with the ICD score being selected for further examination.

Consensus clustering

Following WGCNA, predictive genes were identified through univariate Cox regression analysis. Subsequently, k-means algorithms facilitated consensus clustering to delineate distinctive gene expression patterns linked to risk. The "ConsensusClusterPlus" module within R was employed to execute 1000 iterations (14) confirming the stability of the classifications. A composite consensus score matrix was generated, and the optimal cluster count was ascertained using the cumulative distribution function (CDF) curve (15).

Establishment of prognostic model using comprehensive machine learning approaches

The "WGCNA" R package was used to analyze the TCGA bulk RNA-sequencing (RNA-seq) data and identify a module associated with ICD. Univariate Cox regression

analysis conducted on this module pinpointed prospective predictive ICD genes within the TCGA-OC dataset. To ensure balanced clinical attributes, the dataset was partitioned into a 70% training subset and a 30% internal validation subset, with the GSE73614 dataset being used for external validation. Integration of 10 machine learning methods, including least absolute shrinkage and selection operator (LASSO), ridge, stepwise Cox, CoxBoost, random survival forest (RSF), elastic net (Enet), partial least squares regression Cox (plsRcox), SuperPC, gradient boosting machine (GBM), and survival support vector machine (SVM), was conducted. An extensive evaluation of 101 algorithm combinations was performed with a 10-fold cross-validation approach within the TCGA-OC training data for variable selection and model formulation. Each model's performance was assessed across the TCGA internal validation cohort, the combined dataset, and the GSE73614 dataset with the concordance index (C-index). Models with robust performance and clinical relevance were selected, leading to the development of the ICDRS to predict overall survival in patients with OC.

Survival analysis and construction of a predictive nomogram

The median ICDRS hazard score stratified the TCGA training, internal validation, total datasets, and GSE73614 into high- and low-risk categories. Survival rates across these categories were contrasted using Kaplan-Meier curve analysis, and compared using the log rank test. The predictive capacity of the ICDRS for overall survival (OS) in OC cases was gauged using receiver operating characteristic (ROC) curve analysis. Further evaluations included comparisons of the ICDRS with diverse clinical parameters via area under the curve (AUC) values. Moreover, the correlation between the ICDRS and clinical factors such as age, tumor stage, and radiation therapy was determined. To confirm the ICDRS's independent prognostic significance for the survival of patients with OC, both univariate and multivariate Cox regression analyses were implemented on the TCGA-OC dataset. A nomogram integrating the ICDRS with clinical traits was devised to project survival probabilities for patients with OC, with its performance being validated through ROC curves, the C-index, and calibration plots to ascertain its discrimination ability, accuracy and consistency.

Enrichment analysis

Biological functions and pathways pertinent to the high-risk and low-risk groups were elucidated through Gene Ontology (GO) categorization and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses via the "clusterProfiler" package in R. Additional analysis was carried out using the "ReactomePA" package in R. Moreover, gene set enrichment analysis (GSEA) was conducted to identify potential regulatory pathways of the ICDRS, with genes ranked according to log2 fold change (log2FC) following differential analysis via the "limma" package in R.

Analysis of immune features and immune checkpoint expression

The relationship between the ICDRS and the immune landscape within the ovarian tumor microenvironment was characterized through the quantification of 22 distinct immune cell subsets via the CIBERSORT method. The validity of CIBERSORT results was corroborated using the ESTIMATE algorithm. The link between risk score and immune checkpoint expression was depicted using the "ggstatsplot" R package.

Predicting potential medications for chemotherapy

The responsiveness to drugs was compared between patients with varying ICDRS classifications, particularly between those with high and low scores. This involved computing individual drug sensitivity scores for each patient via the "oncoPredict" R package.

Tumor somatic mutation

The study also examined the differences in single-nucleotide variations (SNVs), and waterfall plots were used to visually represent the top 20 mutated genes. Following this, the association between tumor mutational burden (TMB) and risk assessment was analyzed, with subsequent survival analysis being carried out to determine the variance in survival probabilities between individuals with high and low TMB levels.

Statistical analysis

This study utilized R software (version 4.2.2) for conducting statistical analyses. The Kruskal-Wallis test was applied for

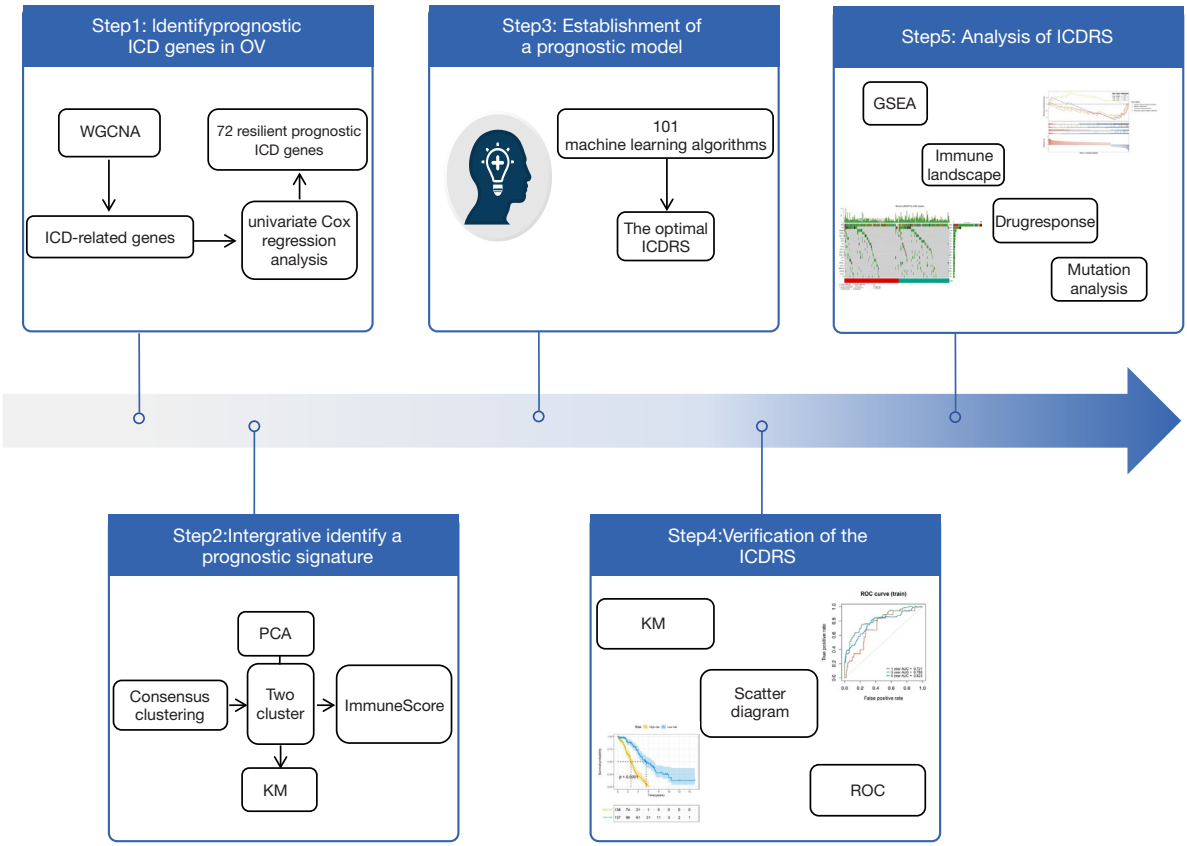

Figure 1 Overview of research methodology. GSEA, gene set enrichment analysis; ICD, immunogenic cell death; ICDRS, immunogenic cell death (ICD)-related signature; KM, Kaplan-Meier; OC, ovarian cancer; PCA, principal component analysis; ROC, receiver operating characteristic; WGCNA, weighted gene coexpression network analysis.

comparisons across multiple groups. In the univariate and multivariate Cox regression analyses of genes, a hazard ratio (HR) greater than 1 indicated a risk factor for prognosis, while an HR less than 1 signified a protective factor. Statistical significance was determined by a P value of less than 0.05.

Results

Identification of prognostic ICD genes in OC

The ssGSEA was used to assess biological pathway changes and process activity within individual TCGA-OC samples, resulting in a phenotype score (*Figure 1*). This score served as the basis for WGCNA to identify the modules strongly linked with the ICD ratings. To achieve a scale-free network, an optimal soft threshold power was established (R^2=0.890) (Figure S1A). A threshold of 150 genes per module and a MEDissThres of 0.25 were used to distinguish eight unique modules (*Figure 2A*). Notably, the MEblue module exhibited a robust correlation with ICD scores (correlation coefficient =0.86) during the analysis of bulk RNA-seq data (*Figure 2B*). A scatter plot depicting gene significance (GS) versus module membership (MM) for the blue module revealed a highly significant correlation (correlation coefficient =0.93; P=0.001), indicating a potential functional role for genes in the blue module in ICD (Figure S1B). Univariate Cox regression analysis then identified 72 robust prognostic ICD genes from 5244 genes within the blue module. Clustering of TCGA-OC samples based on these 72 genes indicated an optimal configuration of k=2 clusters (*Figure 2C-2E*). The discriminatory ability of the 72 ICD genes was demonstrated through principal component analysis (PCA), and these could effectively distinguish OC samples (*Figure 2F*). Notably, Cluster 1 showed better survival outcomes (*Figure 2G*). Moreover, Cluster 2 exhibited a significantly higher immune score than Cluster 1

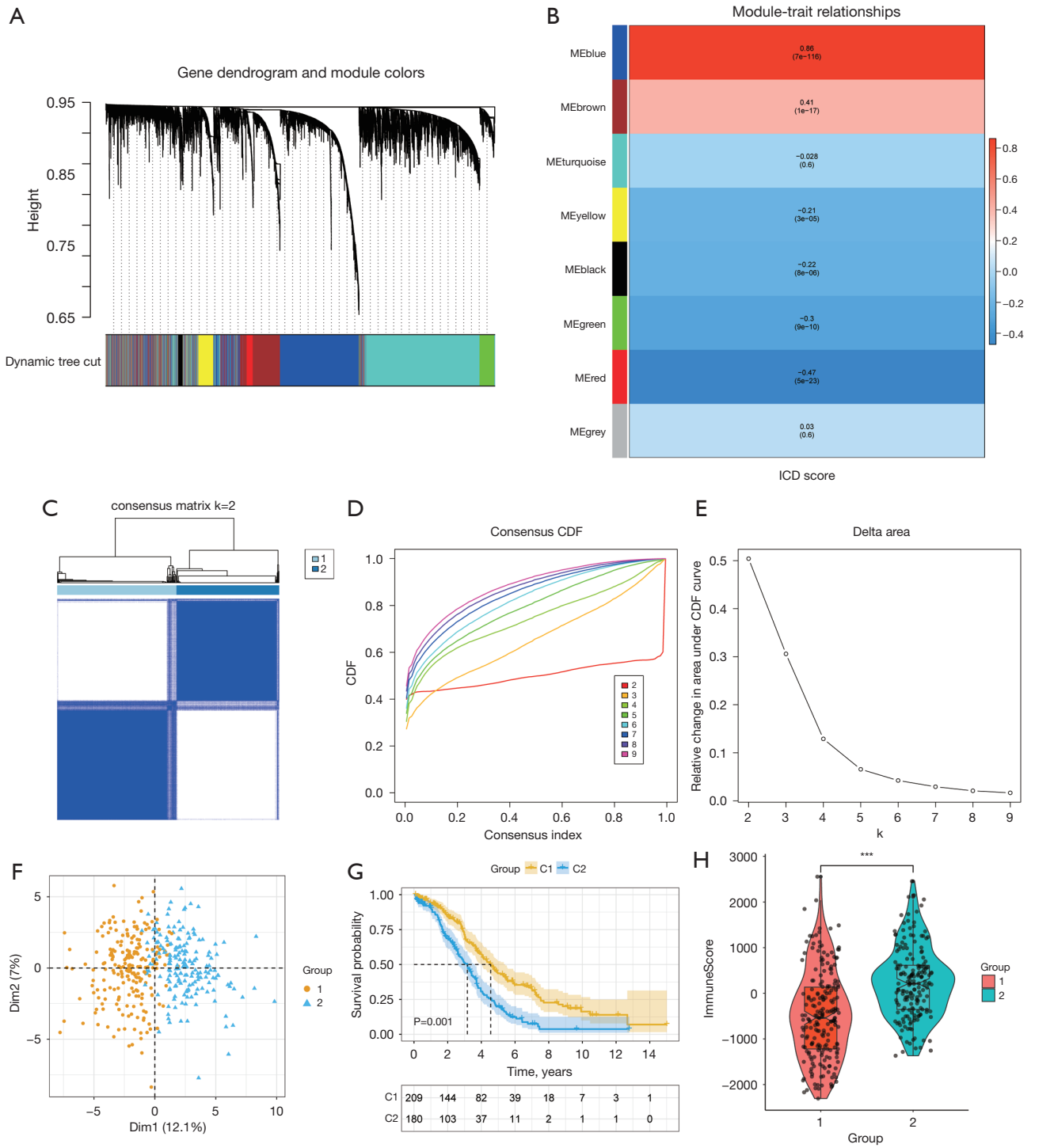

Figure 2 Gene identification and clustering analysis. (A,B) WGCNA identified modules related to ICD in the TCGA-OC dataset. (C-E) Consensus clustering depicted the classification of OC samples based on 72 prognostic genes related to ICD. (F) PCA differentiated between two main clusters of gene expression. (G) Kaplan-Meier curves compared the survival rates of the two clusters, indicating statistical differences. (H) Boxplots illustrating the differences in immune scores between the clusters and the variability in the immune microenvironment. *** means P<0.001. CDF, cumulative distribution function; ICD, immunogenic cell death; OC, ovarian cancer; WGCNA, weighted gene coexpression network analysis; PCA, principal component analysis; TCGA, The Cancer Genome Atlas.

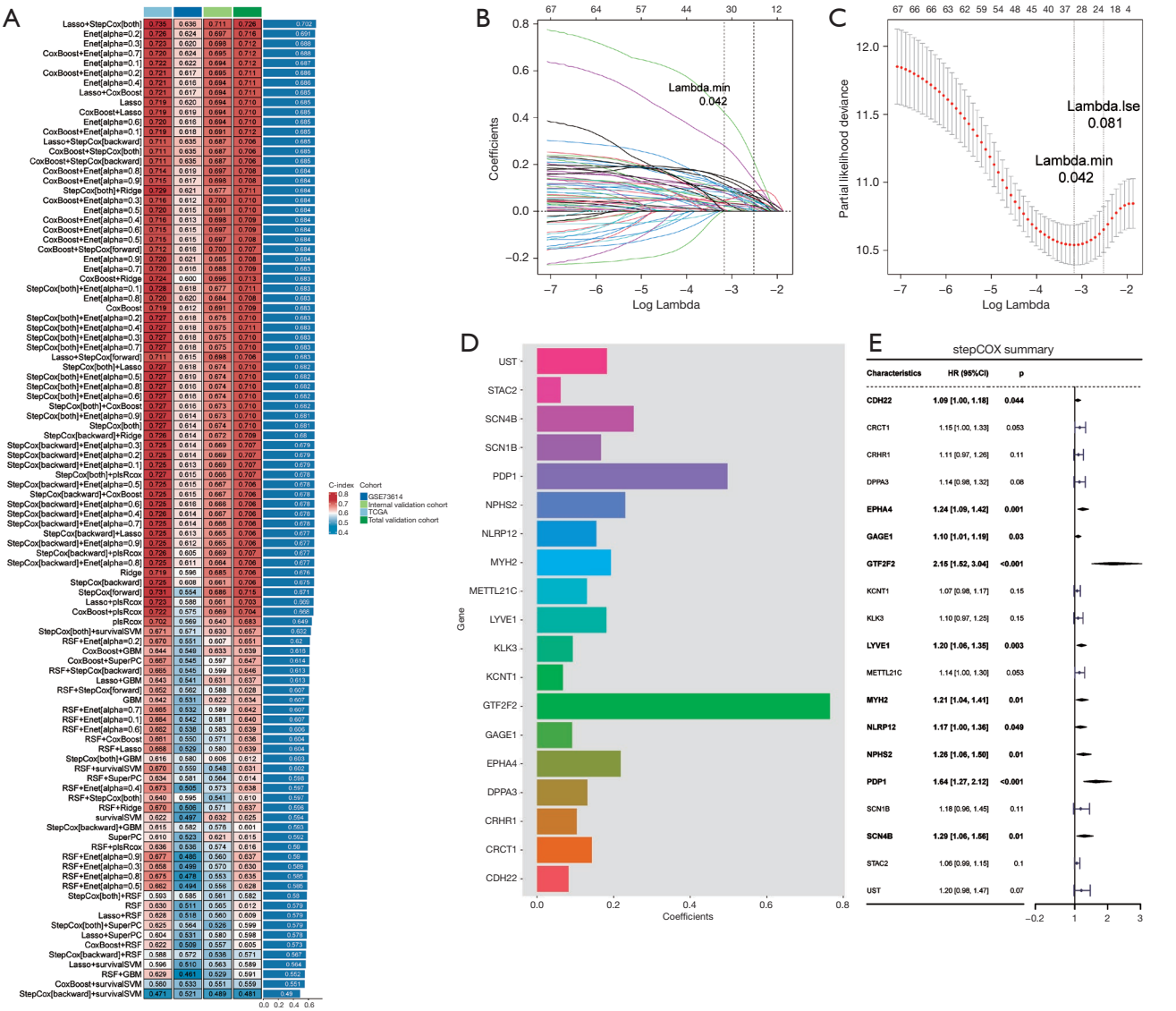

Figure 3 Development of the prognostic model. (A) C-index across all training and validation sets demonstrating the predictive accuracy of the model. The X-axis displays C-index in the range 0–1. (B,C) LASSO analysis for optimizing feature selection and model complexity. (D,E) Final selection of coefficients for ICDRS genes after stepwise Cox regression, with the most influential predictors being identified. LASSO, least absolute shrinkage and selection operator; RSF, random survival forest; GBM, gradient boosting machine; TCGA, The Cancer Genome Atlas; HR, hazard ratio; ICDRS, immunogenic cell death (ICD)-related signature; CI, confidence interval.

(*Figure 2H*) did, supporting the importance of prognostic ICD genes in predicting outcomes and revealing immune variability during OC progression.

Establishment of a prognostic model using integrated machine learning

To develop an ICDRS, we evaluated the 72 prognostic genes. Within the training set, we constructed 101 predictive models, with the C-indices being calculated for both subsets, as depicted in *Figure 3A*. The integration of LASSO and stepwise Cox achieved the highest average C-index (0.702) across all validation sets, suggesting superior efficacy. The optimal λ value, 0.042, was determined via LASSO analysis through 10-fold cross-validation to minimize the partial likelihood deviation

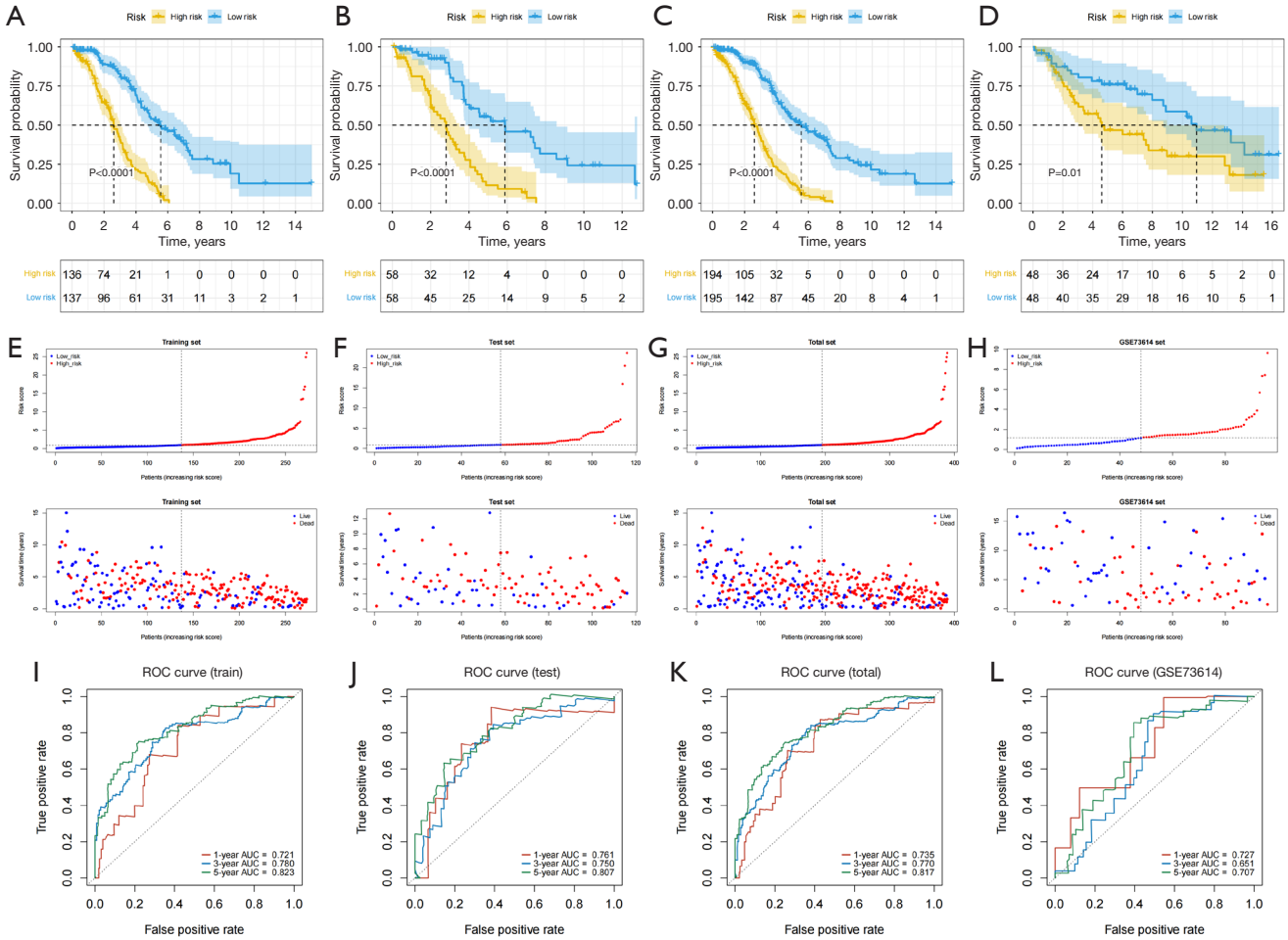

Figure 4 Validation of the ICDRS. (A-D) Kaplan-Meier survival curves for the training, testing, total, and GSE73614 cohorts, showing differences in survival based on ICDRS categorization. (E-H) Scatter plots illustrating the distribution of survival times and risk scores across the cohorts. (I-L) ROC curves for 1-, 3-, and 5-year survival predictions indicating the prognostic power of the ICDRS. ICDRS, immunogenic cell death (ICD)-related signature; ROC, receiver operating characteristic; AUC, area under the curve.

(*Figure 3B*,*3C*). Subsequently, genes with nonzero coefficients in the LASSO analysis were further refined via stepwise Cox proportional hazards regression, culminating in a final set of 19 genes (*Figure 3D*,*3E*).

Verification of the ICDRS

Risk scores were computed for each patient by integrating gene expression levels with coefficients from the Cox regression, with patients being classified into high- and low-risk groups based on the median score. The Kaplan-Meier survival method then demonstrated a significant decrease in OS for high-risk patients within the training set (TCGA-OC; P<0.0001; *Figure 4A*) and in the validation sets, including the test set (P<0.0001; *Figure 4B*), the combined set (P<0.0001; *Figure 4C*), and the GSE73614 dataset (P=0.01; *Figure 4D*). A consistent increase in risk score correlated with higher mortality rates across all datasets (*Figure 4E-4H*). ROC curve analysis verified the ICDRS's discriminative capability, with respective AUCs of 0.721, 0.780, and 0.823 for 1-, 3-, and 5-year survival predictions in the training set; 0.761, 0.750, and 0.807 in the test set; 0.735, 0.770, and 0.817 in the combined set, and 0.725, 0.651, and 0.707 in the GSE73614 dataset (*Figure 4I-4L*).

Construction and evaluation of a nomogram

Cox regression analyses of the survival data (*Figure 5A*,*5B*)

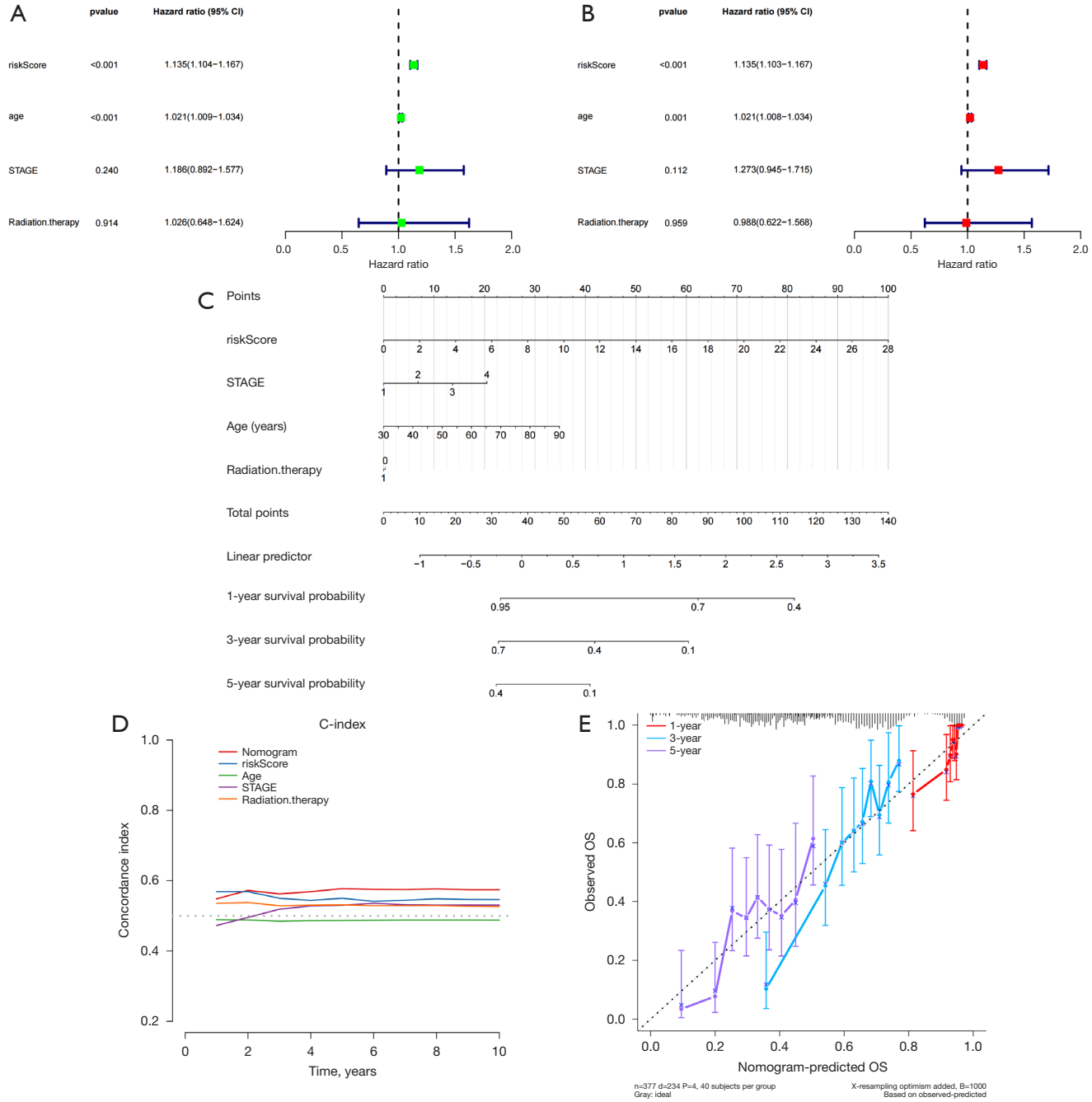

Figure 5 Nomogram integration and evaluation. (A,B) Univariate (A) and multivariate (B) cox regression analyses provided statistical foundation for the nomogram that incorporated ICDRS and clinical characteristics. (C) The nomogram integrates multiple prognostic indicators, offering a personalized survival probability estimate. (D) The concordance index (C-index) was used to evaluate the predictive accuracy of the nomogram. (E) Calibration plots were used to assess the agreement between predicted and observed outcomes to ensure model reliability. ICDRS, immunogenic cell death (ICD)-related signature; OS, overall survival.

identified the ICDRS as a significant predictor of OS in the univariate analysis (hazard ratio >1; P<0.001). It remained an independent prognostic marker for OS (HR 1.135; 95% confidence interval: 1.103–1.167; P<0.001) in the multivariate analysis, confirming its prognostic significance. A nomogram incorporating the ICDRS and clinical characteristics was constructed (*Figure 5C*) to enhance clinical utility. The nomogram's C-index confirmed its superior predictive accuracy, outperforming other clinical factors in predicting OS over 1 to 10 years (*Figure 5D*). Calibration plots showed strong concordance between the nomogram's predictions and actual survival outcomes (*Figure 5E*), suggesting that the ICDRS-based nomogram offers a reliable and precise tool for personalized prognosis estimation in patients with OC.

Enrichment analysis

Enrichment analysis was performed to identify the molecular and functional characteristics associated with high and low ICDRS. There were marked differences between the two groups, especially in extracellular matrix (ECM) organization, collagen-containing ECM, and receptor ligand activity, as per GO assessment (*Figure 6A*). KEGG analysis (*Figure 6B*) highlighted enrichment in pathways such as cytokine-cytokine receptor interaction and transforming growth factor (TGF)-β signaling, in cases with a high ICDRS score. Reactome analysis (*Figure 6C*) revealed increased involvement in G protein-coupled receptor (GPCR) ligand binding, peptide ligand-binding receptors, and ECM organization in the high-risk group. GSEA of the GO gene set (*Figure 6D*) showed that the group with a low ICDRS score was enriched in cytokine-cytokine receptor interaction, olfactory transduction, and neuroactive ligand-receptor interaction, while the group with a high ICDRS score displayed enrichment in the chemokine signaling pathway. These results suggest a potential role of ICDRS in ovarian tumorigenesis.

Immune landscape

To investigate disparities in immune cell infiltration, we employed the CIBERSORT method to estimate the infiltration frequencies in each specimen (*Figure 7A*). This analysis revealed a predominant presence of M1 macrophages and follicular helper T cells in the group with a low ICDRS score. The immunological landscape of OC samples was further assessed using the ESTIMATE algorithm, which calculated immune scores, stromal scores and ESTIMATE scores for each ICDRS subclass. An increase in risk score corresponded to a reduction in these indices (*Figure 7B-7D*). Previous research has demonstrated that enhanced expression of immune checkpoint proteins is linked to a better responsiveness to ICI treatments (16). Therefore, we analyzed immune checkpoint expression patterns across different ICDRS risk subsets, identifying correlations between ICDRS subclasses and immune checkpoint molecules. A lower risk score was positively associated with elevated expressions of *LAG3* and *CD274* (*Figure 7E,7F*).

Drug response and genomic variation landscape for ICDRS groups

In pursuit of novel therapeutic approaches for patients with OC, we identified three drugs from the Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/) that were particularly effective for those in the low-ICDRS score category. These patients exhibited a significant inverse correlation with the calculated half-maximal inhibitory concentration (IC_{50}) values of these drugs, suggesting that lower IC_{50} values were prevalent in this subgroup (*Figure 8A-8C*). To delineate the mutational landscape in OC, we generated a waterfall chart that depicted the top 20 genes with the highest mutation frequency in OC (*Figure 8D*). In both the high- and low-ICDRS score cohorts, *TP53* emerged as the most frequently mutated gene, followed by *TTN*. The analysis of TMB scores revealed a clear positive correlation with increasing risk scores (*Figure 8E*), highlighting the strong association between ICDRS and TMB. Additionally, an investigation into the prognostic implications of TMB levels revealed that patients with high TMB had a better prognosis (P=0.01) (*Figure 8F*).

Discussion

The treatment of patients with OC remains a significant challenge. The combination of limited effective therapies with a lack of reliable predictive biomarkers compromises the effectiveness of therapeutic interventions. A more in-depth investigation into the molecular mechanisms driving OC's progression and the identification of novel biomarkers or signatures that can assist in prognostic assessment are critical to enable personalized care for patients. Herein, we demonstrate the development of a sophisticated

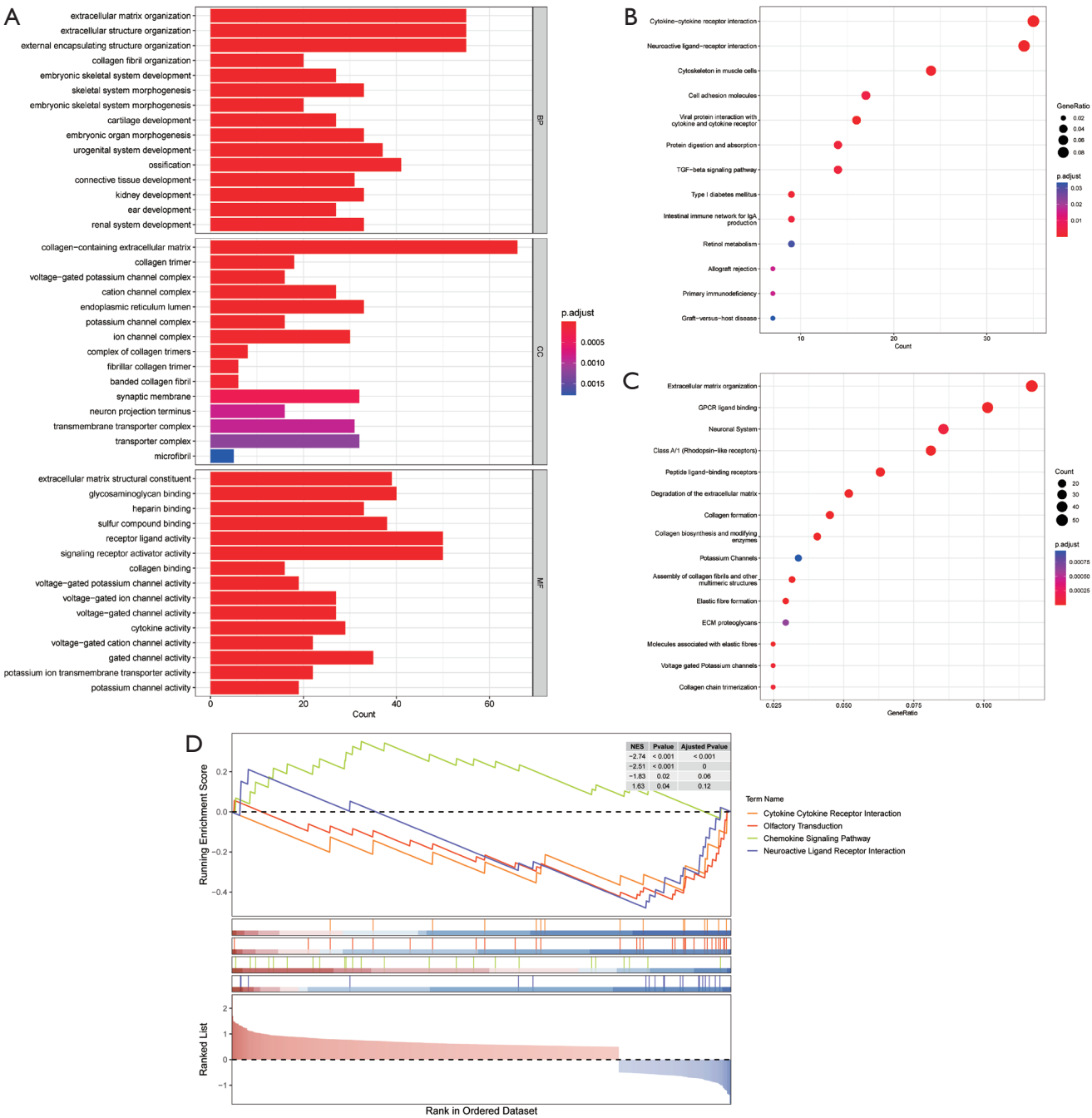

Figure 6 Pathway and functional enrichment analysis. (A-D) Analysis of GO, KEGG, Reactome, and GSEA revealing key biological processes and pathways differentiating between the high- and low-ICDRS score groups. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; GSEA, gene set enrichment analysis; ICDRS, immunogenic cell death (ICD)-related signature; BP, biological process; CC, cellular component; MF, molecular function; TGF, transforming growth factor; GPCR, G protein-coupled receptor; ECM, extracellular matrix; NES, normalization of enrichment scores.

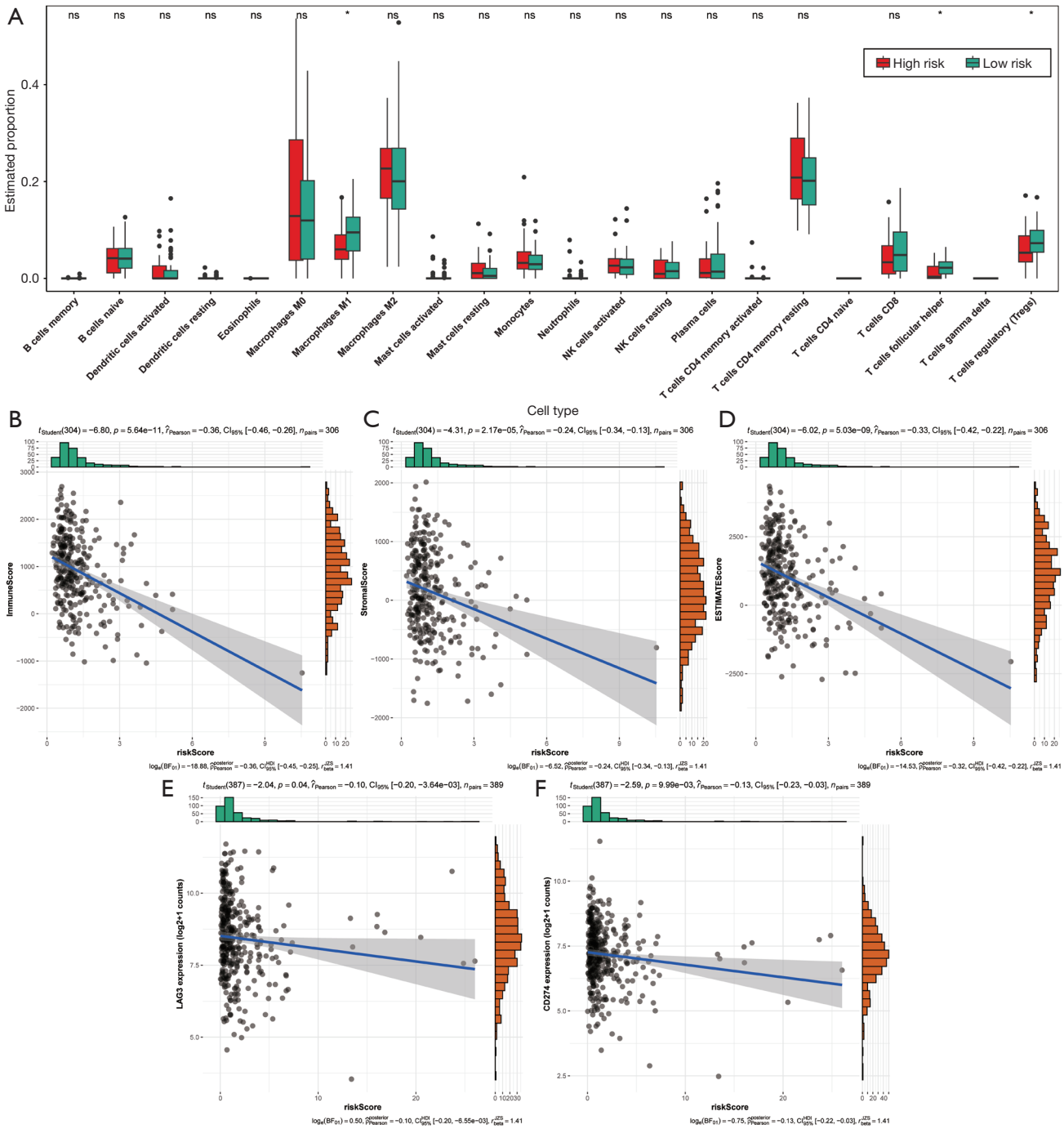

Figure 7 Immunological profiling of the OC samples. (A) Visualization of the differential immune cell infiltration patterns between the ICDRS groups. "ns" means no statistical significance, while "*" indicates P<0.05. (B-D) Immune, stromal, and ESTIMATE scores plotted to quantify the tumor microenvironment's complexity. The X-axis represents the risk score, while the y axis represents Immune, stromal, and ESTIMATE scores respectively. (E,F) Expression levels of immune checkpoints LAG3 and CD274 correlated with ICDRS risk categories, suggesting potential immunotherapy targets. On the outer x and y axes are the frequency distribution histograms of observations and risk scores, respectively. ICDRS, immunogenic cell death (ICD)-related signature; NK, natural killing; CD4, cluster of differentiation 4; CD8, cluster of differentiation 8; BF, Bayes Factor; CI, confidence interval; HDI, Highest Density Interval; JZS, Jeffreys-Zellner-Siow.

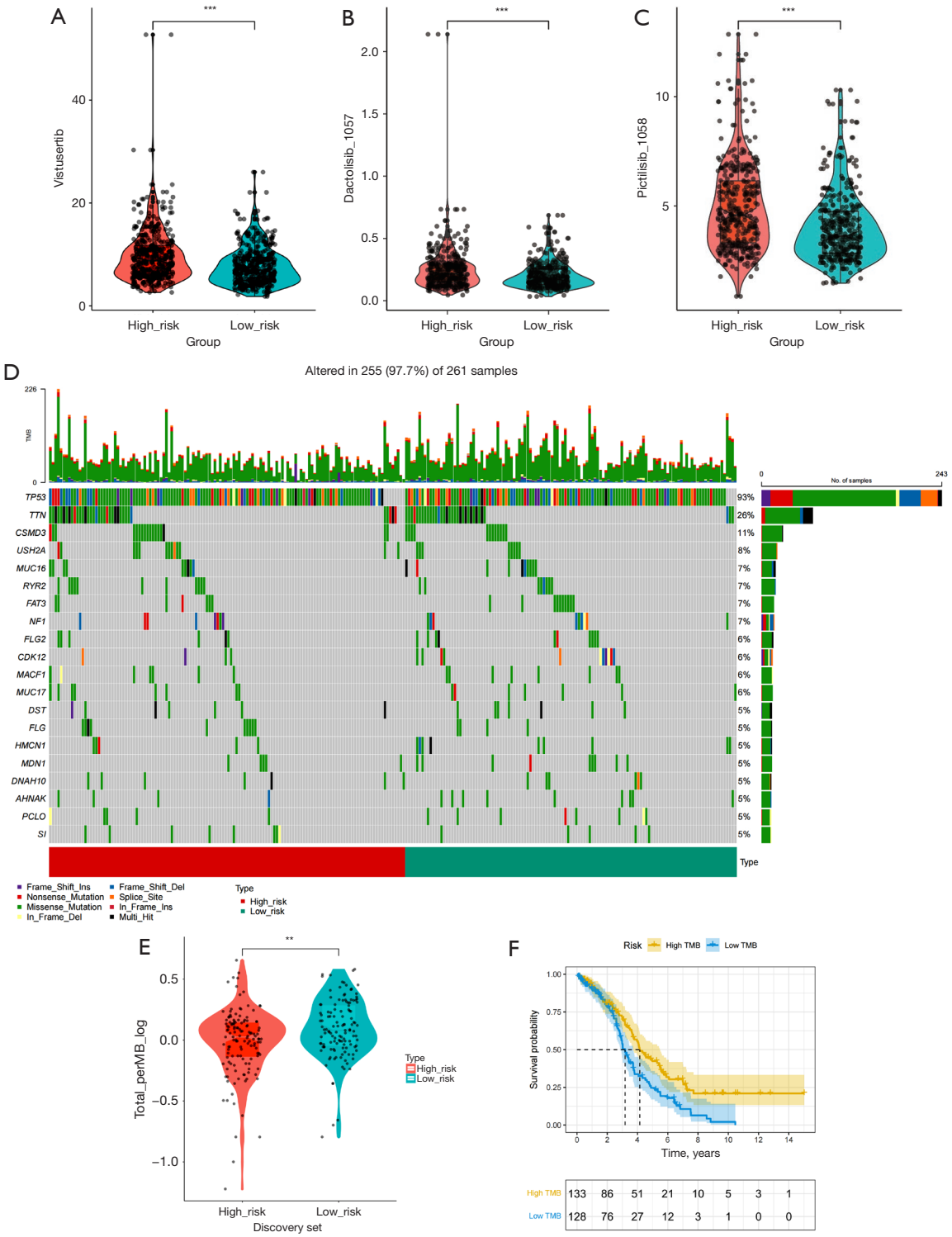

Figure 8 Drug sensitivity and genomic variation. (A-C) Drug sensitivity profiles showing the IC_{50} values for vistusertib, dactolisib, and pictilisib indicating differential responses across ICDRS groups. (D) Waterfall plot displaying the top 20 mutated genes in OC, with the genetic variations correlating with ICDRS risk. (E,F) TMB scores and their association with survival outcomes across the high and low TMB groups as demonstrated through Kaplan-Meier curves. "**" means P<0.01, while "***" indicates P<0.001. IC50, half maximal inhibitory concentration; ICDRS, immunogenic cell death (ICD)-related signature; OC, ovarian cancer; TMB, tumor mutation burden.

computational model to establish a robust predictive biomarker for OC, integrating 10 different machine learning techniques and all possible 101 permutations (Table S1). The comprehensive evaluation identified ICDRS as an exceptionally accurate predictor, demonstrating significant clinical relevance in the context of OC.

Moreover, we examined the relevant molecular mechanisms and pathways associated with the signature across multiple cohorts, finding a substantial correlation between ICDRS and both the progression and outcome of OC. GO analysis suggested involvement of the ECM in the adverse progression of OC, aligning with recent research findings (17). KEGG pathway analysis highlighted a significant link between the TGF-β signaling pathway and risk scores, with earlier study indicating that TGF-β1 may play a role in paclitaxel resistance in OC (18). Additionally, Reactome analysis indicated that genes related to GPCR may promote cell proliferation, invasive motility, and chemoresistance in OC (19). The accumulated evidence supports the hypothesis that ICDRS could influence both the development and progression of ovarian tumors.

Furthermore, these results raise novel perspectives for investigating the imbalance within the immunological microenvironment in the context of OC. By employing the CIBERSORT method, we discovered a higher abundance of M1 macrophages and follicular helper T cells in the group with low ICDRS scores. The predominance of M1 macrophages, known for their tumor-constraining effect, correlates significantly with improved survival outcome, suggesting that a robust immune response can effectively counteract tumor progression (20). Additionally, our study emphasized that epithelial-mesenchymal stimulation in OC cells can facilitate interactions between cancer cells and T cells via the LGALS3/LAG3 pathway, potentially leading to T-cell dysfunction within infiltrating tumors, thereby reducing antitumor immunity. This discovery of an inverse relationship between LAG3 expression and risk scores, in which lower risk scores were associated with better survival outcomes, underscores the complexity of the immune evasion strategies employed by tumors and the potential of targeting such pathways for therapeutic intervention (21).

Additionally, our investigation into potential therapeutic interventions identified compounds that could inhibit the progression of OC in patients categorized with a severe ICDRS classification. This innovative approach leverages the integration of precision medicine facilitated by the recognition of unique patient response profiles. Using the GDSC dataset in conjunction with sophisticated algorithms, we designed targeted treatments tailored to patients with a low ICDRS score. Our analysis identified three chemotherapeutic agents—vistusertib, dactolisib, and pictilisib—as particularly efficacious for individuals with low ICDRS scores. These Food and Drug Administration Approved (FDA)-approved drugs, currently under investigation in clinical trials, support the refinement of therapeutic strategies based on individual patient risk assessments provided by the ICDRS.

Limitations of this study include the potential for overfitting due to the extensive use of complex machine learning algorithms and concerns regarding the generalizability of our findings due to the homogeneity of the datasets used. Future studies should address these shortcomings and enhance the robustness of the models with more advanced validation techniques. Additionally, a greater population diversity is necessary to comprehensively assess the utility and applicability of ICDRS across diverse clinical settings and contribute to its routine clinical use.

In conclusion, we propose ICDRS as a promising tool to stratify patients with OC to study preventive strategies and personalized treatment. An improved prognostic model that allows for better treatment selection has potential to improve patient outcomes while decreasing costs by avoiding ineffective treatments.

Conclusions

Employing a sophisticated array of machine learning techniques, we have successfully established the ICDRS, which can significantly enhance the accuracy of prognostic assessments in OC. This signature is not only helpful in providing a refined prognosis prediction but can also aid in tailoring therapeutic strategies, hence potentially improving patient outcomes. Further studies will be needed to confirm the validity of the ICDRS in large clinical trials across various clinical scenarios. Additionally, expanding the application of the ICDRS to other types of cancer could provide insights into its broader applicability and impact. Ultimately, it would be crucial to integrate the ICDRS into standard clinical practice, in order to ensure that it contributes to more precise and effective cancer treatments.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-2025-118/rc

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-2025-118/prf

Funding: None.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-2025-118/coif). H.S.m. serves as an unpaid Associate Editor-in-Chief of *Translational Cancer Research* from January 2025 to December 2026. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74:12-49.
2. Menon U, Gentry-Maharaj A, Burnell M, et al. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. Lancet 2021;397:2182-93.
3. Monk BJ, Enomoto T, Kast WM, et al. Integration of immunotherapy into treatment of cervical cancer: Recent data and ongoing trials. Cancer Treat Rev 2022;106:102385.
4. Egen JG, Ouyang W, Wu LC. Human Anti-tumor Immunity: Insights from Immunotherapy Clinical Trials. Immunity 2020;52:36-54.
5. Perrone C, Angioli R, Luvero D, et al. Targeting BRAF pathway in low-grade serous ovarian cancer. J Gynecol Oncol 2024;35:e104.
6. Tonti N, Golia D'Augè T, Cuccu I, et al. The Role of Tumor Biomarkers in Tailoring the Approach to Advanced Ovarian Cancer. Int J Mol Sci 2024;25:11239.
7. Galluzzi L, Vitale I, Aaronson SA, et al. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. Cell Death Differ 2018;25:486-541.
8. Jhunjhunwala S, Hammer C, Delamarre L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. Nat Rev Cancer 2021;21:298-312.
9. Meier P, Legrand AJ, Adam D, et al. Immunogenic cell death in cancer: targeting necroptosis to induce antitumour immunity. Nat Rev Cancer 2024;24:299-315.
10. Zhang L, Montesdeoca N, Karges J, et al. Immunogenic Cell Death Inducing Metal Complexes for Cancer Therapy. Angew Chem Int Ed Engl 2023;62:e202300662.
11. Liu Z, Xu X, Liu K, et al. Immunogenic Cell Death in Hematological Malignancy Therapy. Adv Sci (Weinh) 2023;10:e2207475.
12. Kroemer G, Galassi C, Zitvogel L, et al. Immunogenic cell stress and death. Nat Immunol 2022;23:487-500.
13. Yang J, Zhang Y, Cheng S, et al. Anoikis-related signature predicts prognosis and characterizes immune landscape of ovarian cancer. Cancer Cell Int 2024;24:53.
14. Sabah A, Tiun S, Sani NS, et al. Enhancing web search result clustering model based on multiview multirepresentation consensus cluster ensemble (mmcc) approach. PLoS One 2021;16:e0245264.
15. Şenbabaoğlu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. Sci Rep 2014;4:6207.
16. Pardoll DM. The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 2012;12:252-64.
17. Cong S, Fu Y, Zhao X, et al. KIF26B and CREB3L1 Derived from Immunoscore Could Inhibit the Progression of Ovarian Cancer. J Immunol Res 2024;2024:4817924.
18. Guan W, Yuan J, Li X, et al. Cyclin dependent kinase 14 as a paclitaxel-resistant marker regulated by the TGF-β signaling pathway in human ovarian cancer. J Cancer 2023;14:2538-51.

19. Ha JH, Radhakrishnan R, Nadhan R, et al. Deciphering a GPCR-lncrna-miRNA nexus: Identification of an aberrant therapeutic target in ovarian cancer. Cancer Lett 2024;591:216891.
20. Cao K, Zhang G, Yang M, et al. Attenuation of Sialylation Augments Antitumor Immunity and Improves Response to Immunotherapy in Ovarian Cancer. Cancer Res 2023;83:2171-86.
21. Yakubovich E, Cook DP, Rodriguez GM, et al. Mesenchymal ovarian cancer cells promote CD8(+) T cell exhaustion through the LGALS3-LAG3 axis. NPJ Syst Biol Appl 2023;9:61.

Cite this article as: Wang Y, Zhao P, Sun X, Batalini F, Levin G, Soleymani majd H, Chen H, Gao T. A novel machine learning-driven immunogenic cell death signature for predicting ovarian cancer prognosis. Transl Cancer Res 2025;14(2):1359-1374. doi: 10.21037/tcr-2025-118