

Development of prognosis and diagnosis model for osteosarcoma based on oxidative stress-related genes

Dongli Li

Tumor Hospital of Guangxi Medical University

Tianjun Ma

First Affiliated Hospital of GuangXi Medical University

Hong Wu

Tumor Hospital of Guangxi Medical University

Wen Sun

Tumor Hospital of Guangxi Medical University

Mei He

Tumor Hospital of Guangxi Medical University

Qingfeng Song

Tumor Hospital of Guangxi Medical University

Ying Cui (✉ ycuigx@139.com)

Tumor Hospital of Guangxi Medical University

Research Article

Keywords: Oxidative Stress, prognosis, diagnosis, immunity, osteosarcoma

Posted Date: July 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1861817/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

The most frequently malignant bone tumor in children and teenagers is osteosarcoma. However, it is unclear how oxidative stress (OS) affects osteosarcoma. The purpose of this investigation was to determine how OS genes affected osteosarcoma development and prognosis. The Therapeutically Applicable Research to Generate Effective Treatments (TARGET), GenotypeTissue Expression (GTEx), and Gene Expression Omnibus (GEO) databases were used to obtain sample gene expression profiles and clinical data for this investigation. Based on distinct OS gene expression patterns, nonnegative matrix factorization (NMF) clustering was performed to identify molecular groupings. A multivariate Cox regression analysis was used to build the prognostic model. More thorough estimate of immune cells in malignant tumor tissues and prediction of treatment response. Functional analyses include Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, and Gene Set Variation Analysis (GSVA). An in-vitro investigation of OD, a Chinese medication, has been substantial. Using network pharmacology to research the possible therapeutic targets and molecular processes of *Oldenlandia diffusa* (OD) in the treatment of osteosarcoma, basing on high - and low-risk prognostic risk groups. The diagnostic models were built using multilayer perceptron (MLP) neural networks. To predict the survival of patients with osteosarcoma, a prognostic model and a nomogram were created. OS-related genes were shown to be closely linked to the immune state and predictive survival of osteosarcoma patients, according to the GSVA and Gene functional enrichment analyses. 7 active ingredients and 233 target genes were identified from OD, and 25 of the target genes intersected with osteosarcoma-related genes were thought to be prospective therapeutic targets. Network pharmacological analysis showed that the active ingredients for the treatment of osteosarcoma was closely connected with regulation of immune cells. The diagnostic models were constructed based on a 9-gene signature showing excellent recognition ability of osteosarcoma patients. The interplay between OS-related genes and immunology in the development of osteosarcoma might generate new insights into molecular processes and targeted therapeutics for patients with the disease.

Introduction

The lesion location of patients with osteosarcoma that is more prone to lung metastases and worse prognosis predominantly initiates in the metaphysis of the long bones [1]. Only 30% of individuals with lung metastases who have osteosarcoma survive for five years, according to follow-up statistics. However, it is worth mentioning that skip metastases (SM) mainly lead to an even worse prognosis of osteosarcoma [2, 3]. Several anticarcinogens that have been applied clinically didn't see great improvement in osteosarcoma patients with one or more SM [4, 5]. It was significant to emphasize the relationship between genome instability and the molecular mechanism of osteosarcoma [6].

Furthermore, oxidative stress is a physiological condition in which the cell's antioxidant defense systems are overwhelmed by high amounts of reactive oxygen species (ROS). The link between oxidative stress and cancer pathogenesis was obvious [7]. ROS production has been proposed as a crucial aspect in normal cell health [8], whereas excess ROS can induce mutagenicity, cytotoxicity, and gene expression

alterations [9]. Therefore, the oxidant/antioxidant balance plays an important role in the initiation and development of cancer [10]. The damages caused by ROS can be repaired through endogenous and exogenous antioxidants, which can enhance the immune defense and lower the risk of disease and cancer [11, 12]. The impact of oxidative stress on cancer response to therapy was worthy of further investigation. Currently, alternative antioxidant treatments for liver cancer increase the capacity of mitochondrial oxygen intake [13]. Some drugs approved for breast cancer treatment also directly or indirectly have an effect on ROS homeostasis [14]. Consequently, oxidative stress-based treatments have a significant impact on improving the therapeutic efficacy of cancer chemotherapy. Not only has cancer therapy been investigated, but numerous studies have also been conducted to investigate the fundamental mechanism of tumor start and progression based on oxidative stress [15–17].

With the advancement of tumor tissue sequencing tools, there is a pressing need to build prognostic models based on OS-related genes, displaying how to apply prognosis prediction in clinical situations. There is no such prediction model for the most common malignant bone tumor, osteosarcoma. We created a predictive model for osteosarcoma patients based on oxidative stress genes linked to immunological state and investigated its prognostic implications and diagnostic efficiency.

Materials And Methods

Data collection and oxidative stress-related genes

TARGET-OS (<https://portal.gdc.cancer.gov/projects/TARGET-OS>) was used to obtain the HTSeq-Counts data of osteosarcoma patients, which were then downloaded using the UCSC Xena browser at <https://xenabrowser.net>. on January 3, 2022, with 88 samples obtained. Additionally, a total of 803 healthy Muscle-Skeletal samples were gathered from the Genome Tissue Expression database (GTEx; <https://gtexportal.org/home/datasets>) as a control group, in order to enhance the number of healthy samples. Finding genes that were differently expressed between original data of healthy Muscle-Skeletal and osteosarcoma mRNA expression required the use of the DESeq2 R program under $|\log_2 \text{ fold change}| \geq 1$ and a standard false discovery rate (FDR) < 0.05 . Meanwhile, genes were eliminated using an average count limit of 1. The Gene Cards database (<https://www.genecards.org>) was conducted for 1402 OS-related genes with a relevance score ≥ 6 on December 15, 2021. In addition, The HTSeq-FPKM data and features of osteosarcoma patients were obtained from the TARGET database and set as the training cohort for the prognostic risk model. The GSE21257 and GSE16091 datasets were utilized as a validation cohort for the prognostic risk model, with 87 osteosarcoma patients' gene profiles and clinical information. Furthermore, the training cohort of the artificial neural network model was created by combining the TARGET and GTEx data sets. Gene expression levels were calculated using the FPKM values. The artificial neural network model was tested for effectiveness verification on GSE87437 and GSE157322, GSE33382 and GSE142162, GSE14827 and GSE21257 datasets. To eliminate batch effects described by earlier investigations, profiles from various platforms were standardized using the R package "sva". These datasets are summarized in Table 1. No ethical oversight was required, as these data were downloaded from a public repository.

Screening of prognosis-associated genes and exploratory molecular subgroup classification

Univariate Cox regression analysis was initially utilized to detect prognosis-related oxidative stress genes in osteosarcoma patients. The differentially expressed OS genes (DEOSGs) with $P < 0.05$ were identified to be related to prognosis. To simplify the final signature and prevent over-fitting, pre-screened prognosis-related OSGs were gathered to run the least absolute shrinkage and selection operator (LASSO) analysis using the glmnet R package. Using the nmf R package, additional subtypes were identified using nonnegative matrix factorization (NMF) clustering. The stromal score, immunological score, ESTIMA score, and tumor purity of distinct subgroups were calculated using the ESTIMATE method. In addition, the tumor immune estimation resource (TIMER) method was used to determine the abundance of immune cells in osteosarcoma tissues, such as B cells, CD4⁺ T cells, CD8⁺ T cells, neutrophil cells, macrophage cells, and dendritic cells.

Model construction and verification based on OSGs

The prognostic model was built using the TARGET dataset, which was performed as the training cohort. Multivariate Cox regression with stepwise regression was used to create the prognostic risk model. Each osteosarcoma patient received a prognosis risk score based on the prognostic model, with the medium operating as the dividing line between high- and low-risk groups. High-risk patients were those with risk scores that were greater than the medium. To evaluate the model's reliability, samples from the GEO database were utilized as a verification cohort. ROC curves were created using the "survivalROC" and "timeROC" R packages to assess the efficacy of our risk model in predicting the overall survival rate of patients with osteosarcoma. To evaluate outcomes between two risk groupings, the Kaplan–Meier technique and log-rank test were being used using the R Bioconductor survival package. In TCGA SARC tumor samples, TIMER was utilized to examine the mRNA expression data for the chosen OS-associated genes that were discovered to develop a predictive model and its link with tumor infiltration of six immune cell types (B cells, CD4⁺T cells, CD8⁺ T cells, neutrophils, macrophages, and dendritic cells).

Nomogram construction and validation based on the prognostic model

The prognostic model and clinical data of patients with osteosarcoma, including sex, age, tumor metastasis, and tumor location, were combined using the "rms" R package in order to create a more accurate prognosis utilizing the research's results. Furthermore, the GEO datasets were utilized to validate the prognosis prediction efficacy of the built nomogram. The concordance index (C-index) and calibration curve were used to evaluate the nomogram's predictive accuracy and discriminative capabilities.

Enrichment analysis and mechanism exploration

The "ClusterProfiler" R package was used to investigate the underlying mechanism of OSGs on the development of osteosarcoma employing Gene ontology (GO) enrichment and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. To begin, the R package limma was used to identify differentially expressed genes (DEGs) in the high-risk and low-risk groups. The underlying function was

then investigated using Gene Ontology (GO) and KEGG analyses. In addition, the DEGs were used to screen hub genes by creating a protein-protein interaction (PPI) network in the Search Tool for the Retrieval of Interacting Genes (STRING) database. In Cytoscape version 3.7.2, the Molecular Complex Detection (MCODE) plugin was used to filter modules of the PPI network (degree cutoff=2, node score cutoff=0.2, k-score=2, and max.depth=100). The top cluster with the greatest clustering score was chosen, and the ggplot2 and "clusterProfiler" R packages were used to analyze GO and KEGG pathways. Besides, to further investigate the KEGG terms, the Gene Set Variation Analysis (GSVA) tool was used.

Prediction of response to targeted therapy

The IC50 value represents a substance's ability to block particular biological or metabolic activities. The "ggplot2" and "pRRophetic" R packages were used to plot the half-maximal inhibitory concentrations (IC50) of targeted treatment medicines in the high- and low-risk groups of osteosarcoma patients. Box plots were used to display the sensitivity of the various patients.

Identification of substances and mechanisms of Oldenlandia diffusa in the Treatment of osteosarcoma by Integrating Network Pharmacology and molecular docking

Oldenlandia diffusa (OD), a well-known traditional Chinese medicine, is used to both prevent and cure a wide range of illnesses, particularly cancer. Based on the criteria of drug-likeness (DL) ≥ 0.18 and oral bioavailability (OB) $\geq 30\%$, the active constituents of OD were selected in the TCMSP database (<https://old.tcmsp-e.com/tcmsp.php>). Then, using SEA (<https://sea.bkslab.org>) databases, protein targets interacting with those prospective active substances were predicted. The differentially expressed genes (DEGs) in the high-risk and low-risk groups screening p value < 0.05 . Following their intersection with OD therapeutic targets, these targets were gathered as candidate genes. Through the use of Cytoscape software (version 3.7.2), a network was created to represent the intricate connection between OD and osteosarcoma-potential target genes. utilizing "clusterProfiler" and "ggplot2" R packages, the investigation of GO biological functions enrichment was then shown which the p-value was < 0.05 . Finally, using the survival R package, all candidate genes were subjected to a univariate Cox regression analysis. Genes having a P < 0.05 were included in the molecular docking simulation. The PubChem website (<https://pubchem.ncbi.nlm.nih.gov/>) was used to get the two-dimensional structure of the active component. The PDB database (<https://www.rcsb.org/>) was used to identify the 3D structure of the candidate genes, and the "PyMOL" software was used to remove the water molecules and small molecule ligands.

Development of the neural network models

The MLP kind of ANN model was constructed using the SPSS 20.0 Neural Network module. In this study, the input layer neurons were made up of nine OS-related genes (G6PD, ADRB1, MYC, TBXA2R, PPARG, MAP3K5, EPHX2, ACADVL, and TOR1A). As output layer neurons, six variables (normal or osteosarcoma samples, osteosarcoma or Ewing sarcoma samples, metastatic or non-metastatic samples) were applied

independently. At a 7:3 ratio, all participants were randomly assigned to a training and a test group. The test set is used to verify the outcomes and the training set is used to train the network.

Statistical analysis

For statistical analyses, R software (version 4.1.0) and SSPS software (SPSS Inc, Chicago, IL, USA) were used. The survival-related genes were discovered using univariate Cox regression analysis, and multivariate COX regression analysis was utilized to screen the prognostic prediction genes that were used to build the prognostic model. A time-dependent receiver operating characteristic (ROC) analysis was used to examine the prognostic model's prediction effectiveness. Receiver operating characteristic curves (ROC) were drawn and the area under the curve (AUROC) was determined to analyze the role of nine diagnostic indicators. The Student t-test was employed for statistical analysis when there were only two groups, and a 1-way analysis of variance (ANOVA) was flexibly utilized when there were three or more groups. A statistically significant difference was defined as a P-value of less than 0.05. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$.

Results

Identification of a molecular subgroup

The methodology depicted in Supplementary figure 1 was used to execute bioinformatics analysis of publically available datasets. With the “survival” R package, 604 DEOSGs were submitted to univariate Cox regression analysis, and 79 DEOSGs were selected as osteosarcoma prognostic-associated genes with a $P < 0.05$ (Supplementary Table 1). Following that, the LASSO algorithm was used to decrease the range of particular DEOSGs (Figure 1a-b), and 17 DEOSGs were picked. Ultimately, to generate the prognostic risk model, the multifactor Cox proportional hazard regression model employing stepwise regression was utilized to conduct multivariate analysis and identify nine candidate genes (Table 2). The NMF consensus clustering was used to make an exploratory classification of molecular groupings based on nine candidate genes (Supplementary figure 2), with two main categories ruled out satisfactorily (Figure 1c). A heatmap was utilized to portray the expression levels of nine candidate genes per patient in cluster 1 and cluster 2, as shown in Figure 1d. To assess each osteosarcoma patient's immunological status and compare the two clusters, the ESTIMATE method was employed. Figure 2a shows that the osteosarcoma tissues in cluster 2 had a higher ESTIMATE Score, Immune Score, and lower Tumor Purity than those in cluster 1. In addition, the abundance of immune cells in osteosarcoma tissues was calculated using the TIMER algorithm, resulting in cluster 2 had a significantly greater abundance of macrophages (Figure 2b). Furthermore, we analyzed the overall survival of the two clusters and discovered that Cluster 2 had significantly greater survival (Figure 2c).

Establishment of prognostic model based on the TARGET dataset

The risk score was calculated using the equation below, which was based on a multivariable Cox proportional hazard model that included nine candidate genes: Risk score = $\sum \text{exp gene}_i \times \beta_i$, in which

expgene is the expression level of each selected 9 OS-associated genes (G6PD, ADRB1, MYC, TBXA2R, PPARG, MAP3K5, EPHX2, ACADVL, and TOR1A) and β represents the gene's coefficient value. Patients with osteosarcoma in the TARGET cohorts were classified as high- or low-risk using the prognostic risk model's median risk score. A heatmap was used to display the gene expression levels of nine candidate genes per patient (Figure 3a). Furthermore, immune cell examination indicated that the high-risk group had a significantly decreased abundance of immune cells such as dendritic cells, CD8+ T cells, neutrophil cells, and macrophage cells (Figure 3b), as well as lower stromal score, immune score, ESTIMATE score, and tumor purity as compared to a low-risk group (Figure 3c). Furthermore, the Kaplan-Meier approach was used to compare two groups, revealing that patients in the low-risk group had even greater hope of surviving than patients in the high-risk group (Figure 3d). Utilizing time-dependent ROC analysis, the built prognostic model's validity was further evaluated, which showed AUCs of 0.94, 0.91, and 0.92 for 1, 3, and 5 years, significantly (Figure 3e). Investigating the relationships between nine-OS signature and tumor immunity showed that a total of eight candidate genes (G6PD, ADRB1, MYC, TBXA2R, PPARG, MAP3K5, EPHX2, and TOR1A) were involved in immune infiltration in SARC (Supplementary figure 3).

Independence testing of the prognostic model

To evaluate the independence of the built prognostic model, the clinical parameters of the patients, such as sex, age, metastatic status, and lesion location, were taken into consideration. The univariate and multivariate COX regression analysis demonstrated that risk score and metastatic status were separately linked with osteosarcoma patients' prognosis, as depicted in Figure 4a-b. At the same time, there was no link between sex, location, or age and prognosis. Furthermore, subgroup analysis by age (Figure 4c), sex (Figure 4d), and metastasis status (Figure 4e) revealed that patients with higher risk score had a considerably shorter overall survival rate, illustrating the constructed prognostic model's independent prediction efficiency. Additionally, identifying if the risk score difference between each grouping was statistically significant. There was no statistical difference in gender (Figure 4f) and age (Figure 4g) between the groupings categorized. In comparison, there were significant variations in risk scores between the metastasis and non-metastasis groupings (Figure 4h). These findings suggested that the created prognostic model might be used independently to predict the prognosis of osteosarcoma patients.

Validation of the prognostic model

Using the GEO dataset, performance of the built prognostic model is verified. The built prognostic model to compute the risk scores that were used to separate the GEO cohort into two subgroups. A heatmap depicted the expression levels of the nine candidate genes used to build the model in the GEO cohort (Figure 5a). In additionally, compared to the low-risk group, high-risk groups exhibited a significant abundance of immune cells in osteosarcoma tissues, including dendritic cells, B cells, CD8+ T cells, neutrophil cells, and macrophage cells (Figure 5b), but also the immunological status of osteosarcoma tissues in the GEO cohort, with substantial differences in stromal score, immune score, ESTIMATE score, and tumor purity identified between the high- and low-risk groups (Figure 5c). Patients in the low-risk

group had significantly superior survival rates than those in the high-risk group, according to survival analyses (Figure 5d). Furthermore, ROC analysis revealed that the AUC for 1 year, 3 years, and 5 years was 0.78, 0.73, and 0.72, respectively, according to ROC analysis (Figure 5e). These findings corroborated the created prognostic model's prognostic prediction accuracy.

Construction and validation of the nomogram

To forecast the survival rate of osteosarcoma patients at 1, 3, and 5 years, a nomogram integrating the risk score with the related clinical characteristics, such as sex, age, metastatic status, and tumor location, was created. Patients might receive a total score to predict the 1-, 3-, and 5-year overall survival as depicted in Figure 6a, with a score assigned to each item depending on the actual circumstance. The predicted accuracy of the built nomogram was validated using the TARGET and GEO cohorts. The blue, green, and red lines in Figure 6b indicate the observed survival rate. The optimum survival rate is shown by the black line, which shows a good match between the observed and optimized values in the training cohort (C-index, 0.907, [95% CI, 0.845 to 0.969], p-value, 4.855e-38). In the verification cohort, the observed survival rate and optimized survival rate also matched well (C-index, 0.753, [95% CI, 0.672 to 0.834], p-value, 7.610e-10) (Figure 6c).

Filtering of differentially expressed oxidative stress-related genes and functional enrichment analysis

The levels of gene expression in the high-risk and low-risk groups were compared first. In the training cohort, 345 DEGs were found in the two groups, as can be seen in Figure 7a, with 79 genes upregulated and 266 genes downregulated. The PPI analysis (Supplementary figure 3) further screened the first module (Figure 7b). According to GO function enrichment analysis, hub genes were more prevalent in the first module's immune-related BP, for example, positive regulation of cytokine production, neutrophil-mediated immunity, and macrophage activation (Supplementary figure 4). In addition, GO analysis showed that, in terms of biological process (BP), 345 DEGs were mainly enriched in positive regulation of leukocyte activation, T cell activation, and leukocyte cell-cell adhesion. In terms of cellular component (CC), external side of plasma membrane, membrane microdomain, and primary lysosome were enriched. Regarding molecular function (MF), 345 DEGs were mainly related to peptide binding, chemokine activity, and immunoglobulin receptor binding (Figure 7c-d). Enrichment analysis of the KEGG pathway showed that the 345 DEGs were linked to several immune-related signaling pathways, such as Cytokine-cytokine receptor interaction, Neutrophil extracellular trap formation, Complement and coagulation cascades, Chemokine signaling pathway (Figure 7e-f). The GSEA enrichment analysis demonstrated that immune-related KEGGs which helped to detect the differences in expression, such as Natural killer cell mediated cytotoxicity, antigen Processing and Presentation, Leukocyte migration were highly expressed in a low-risk group, compared with a high-risk group (Figure 8a). Harmoniously, these gene function enrichment studies revealed that immunity disorders might be closely related to osteosarcoma patients' underlying mechanism. Besides, the expression of nine candidate genes in the training cohort was shown in Figure 8b, respectively.

Response to Targeted Therapy

Based on the pRRophetic algorithm, the IC50 of six chemotherapeutic drugs (AKT.inhibitor.VIII, Axitinib, Bortezomib, MG.132, Pazopanib and Thapsigargin) were estimated in OS patients. Bortezomib, MG.132, and Pazopanib exhibited higher IC50 in patients with high-risk scores (Figure 9c-e). On the contrary, the IC50 of AKT.inhibitor.VIII, Axitinib and Thapsigargin were higher in low-risk samples (Figure 9a-b, f). These findings revealed that the two groups with varying risks had dramatically varied responses to various targeted medications.

Result of Network Pharmacology and Molecular Docking

37 OD components in all were looked up and taken from the TCMSP. Seven active chemicals were chosen from the OD substances for database creation based on the screening parameters of OB $\geq$ 30% and DL $\geq$ 0.18 (Supplementary Table 2). Meanwhile, 377 target genes were obtained from the SEA database for 7 active ingredients of OD (Supplementary Table 3). After verification, the duplicates were eliminated, leaving 233 target genes. OD target genes and osteosarcoma target genes were intersected using Venny 2.1 software to obtain 25 potential targets. The drug-components-targets-disease network was created through the "Cytoscape" software in Figure 10a, which included 34 nodes (7 compounds, 25 targets) and 76 edges. As shown in Figure 10b, in terms of biological processes, targets were mostly associated with immunity, such as, regulation of T cell activation, regulation of leukocyte cell-cell adhesion, positive regulation of lymphocyte activation, etc. Through univariate Cox regression analysis, 11 genes (CA3, CBS, EPHA3, EPHA6, G6PD, MITF, CDC25B, PTPRC, ST3GAL1, TERT, AXL) were collected for Molecular docking. Meanwhile, 3D structure of EPHA6 and ST3GAL1 couldn't be found in PubChem website. 7 active components bind 9 candidate genes to varying degrees (Supplementary Table 4). Vina scores that are lower suggest a more potent and consistent interaction between the chemical and receptor. We found that the vina scores of 20 molecular docking results were less than -6.5, besides Molecular docking results of PTPRC and 2-methoxy-3-methyl-9,10-anthraquinone, PTPRC and 2,3-dimethoxy-6-methyanthraquinone. The docking results for the four compounds with the lowest binding energies are shown in Figure 10c-f.

Construction and validation of neural network models

The diagnostic models were built using multilayer perceptron (MLP) neural networks based on the TARGET and GEO datasets. The model's architecture, which includes input layers neurons and output layers neurons, was shown in Supplementary Figure 5. Figure 11 shows the recognition ability of the MLP model. For the MLP model of constructing tumor and the normal samples based on the TARGET datasets, 100% of training and testing samples were correctly diagnosed. Besides, the AUROCs of MLP models was 1 (Figure 11a). For the MLP model of constructing tumor and the normal samples based on the GEO datasets, 97% of training samples and 100% of testing samples were correctly diagnosed. Besides, the AUROCs of MLP models was 0.976 (Figure 11b). For the MLP model of constructing osteosarcoma and the Ewing sarcoma samples, 100% of training samples, 92.2% of testing samples were correctly diagnosed. Besides, the AUROCs of MLP models was 0.985 (Figure 11c). For the MLP model of constructing metastases and the non-metastases samples, 58.9% of training samples, and

58.3% of testing samples were correctly diagnosed. Besides, the AUROCs of MLP models was 0.661 (Figure 11d).

Discussion

The youthful and older people were likely to be found relatively high risks of osteosarcoma that was mesenchymal origin [18]. Patients with metastases or recurrent illness [19] had a 5-year survival rate of just around 20% [20]. The treatment of osteosarcoma patients stressed possibly removing and eradicating the tumors [21]. Extensive surgical resection, neoadjuvant chemotherapy, and adjuvant chemotherapy are the current standard of treatment for patients with osteosarcoma [22]. Several anticarcinogens applied clinically could not improve the prognosis of those metastatic and recurrent osteosarcoma patients [5]. Consequently, the development of innovative treatment approaches is urgently required, overcome conventional treatment regimens' barriers and improve the overall survival of patients with osteosarcoma. Since the 1950s, immunotherapy has been an effective therapeutic method for treating cancer [23]. Although challenges remain on the path forward, the overall survival of patients with osteosarcoma have benefited from the development of immunotherapy in recent years. Furthermore, bioinformatics analysis of osteosarcoma may help with the development of osteosarcoma immunotherapies.

An unbalance between the generation of oxidants and antioxidant defenses is described to as "oxidative stress," and it may harm biological systems. Searching for the effect of oxidative stress in multiplication, invasion, and metastasis tumor can causes it hopeful of investigating the mechanism of osteosarcoma and improving the prognosis. Inducing imbalance of oxidative stress in tumor cells may can treat tumors more effectively [24]. However, oxidative stress favored increasing the sensitivity to treatment, which may limit antioxidant therapies' curative effect in clinical trials [25]. Moreover, some chemotherapeutic agents emphasized the feature of oxidative stress and aimed at inducing ROS. For example, the chemotherapeutic drugs 5-FU and oxaliplatin got through increasing ROS formation to induce the cytotoxicity [26]. ROS manipulation inspired effective cancer therapies that has altered redox balance [27]. Such as increasing the production of ROS levels to toxic levels and exhausting the antioxidant system capacity caused programmed cell death, which revealed a connection between oxidative stress and the diagnosis, treatment, and prognosis of many malignancies [28–30]. Unfortunately, no research has looked at how oxidative stress and clinical traits in individuals with osteosarcoma relate to one another. Consequently, based on oxidative stress-related genes (OSGs), this research develops a thorough predictive model and investigates the underlying molecular mechanism. Initially, based on the expression matrix of prognosis-related oxidative stress-related genes, we separated the osteosarcoma patients into two clusters, overall survival and immunity status showed a significant difference between the two clusters. These analysis results indicated that the immunity of osteosarcoma patients might be closely related to the different prognoses between the two clusters. In order to further forecast the prognosis of osteosarcoma patients, we created an OSGs signature utilizing the TARGET dataset. This signature has a strong ability to predict prognosis, and the GEO cohort clearly verified it. Nine genes make up the oxidative stress signature: TBXA2R, ADRB1, ACADVL, EPHX2, PPARG, TOR1A, MYC, G6PD, and MAP3K5.

In numerous primary malignancies, greater metastasis was linked to a single nucleotide variation in the TBXA2R gene, according to a phenome-wide association analysis [31]. Alterations in ADRB1 genes may work to drive tumor formation and development. The overall survival of solid tumor patients in the GDC Pan-Cancer set was reduced by increasing expressions of ADRB1 and ADRB2 genes [32]. ACADVL is a 20-exon gene that is involved in fatty acid oxidation and is postulated perhaps the result of under-expression is associated with many cancers, which may produce the same effect as over-expression of FASN [33]. Downregulation of EPHX2 in various types of cancer could be related to tumor progression [34]. Ferroptosis has two roles in the biology of tumors and is a promising target for cancer treatment [35]. One idea was that several downstream target genes of PPARG were responsible for both apoptosis and ferroptosis, respectively [36]. TOR1A was discovered as one of the GFP (Genomic Fabric Paradigm) master regulators capable of completely characterizing the transcriptome complexity of a heterogeneous prostate tumor and identifying the most significant genes in each cancer nodule [37]. Targeting MYC is an appealing option for cancer treatment since it is one of the most commonly unregulated driver genes in human cancer and has a causal role in cancer development [38]. Endothelial cells exposed to oxidants showed elevated ROS when G6PD activity was lacking [39], but G6PD-overexpressing cells continued to be resistant to redox imbalance [40]. In living creatures, MAP3K5 has pleiotropic activities via different processes that show increased ROS and reduced GSH reserves. MAP3K5 is very susceptible to oxidative stress, significantly contributes to apoptosis, and plays these roles in a variety of ways [41].

Through the use of subgroup analysis, univariate cox regression analysis, and multivariate cox regression analysis on the training cohort, the prognostic model's independence was clearly proven. Additionally, it was shown that the presence of metastases independently impacts the prognosis of individuals with osteosarcoma. The outlook for those with osteosarcoma can be better predicted by building a nomogram incorporating clinical variables of the disease, such as sex, age, metastatic state, and lesion location. The risk score predicated on the gene profile associated with oxidative stress had the highest weighted score under this nomogram that was developed. The prognostic prediction efficiency of the nomogram was demonstrably shown in perspective of 1-y, 3-y, and 5-y overall survival in both training and testing datasets, which further confirmed the prognostic model's prediction effectiveness.

Individuals' diagnoses or illness predictions may be performed using an artificial neural network (ANN) that fits a nonlinear association between input and output variables [42–46]. MLP is one of the most common methods used for network training. Thus, we have utilized the MLP network to create models for separating osteosarcoma patients from healthy Muscle-Skeletal samples from healthy Muscle-Skeletal samples, osteoblast and Ewing sarcoma samples, and osteosarcoma metastasis from non-metastasis samples of osteosarcoma patients. In the neural network modeling, there have been nine oxidative stress signatures (G6PD, ADRB1, MYC, TBXA2R, PPARG, MAP3K5, EPHX2, ACADVL and TOR1A) employed. The osteosarcoma diagnostic models had outstanding diagnostic capabilities, with accuracy levels that were higher than 80%. The diagnostic models for osteosarcoma metastasis and non-metastasis, unfortunately, have a limited sensitivity, which might result in some missed diagnosis. Finally, we constructed diagnostic models for osteosarcoma employing these gene biomarkers and MLP neural networks after evaluating the potential of the G6PD, ADRB1, MYC, TBXA2R, PPARG, MAP3K5, EPHX2, and TOR1A gene

biomarkers for osteosarcoma diagnosis. Different oxidative stress-related gene expression matrices of osteosarcoma patients showed different immunity status, more active immunity status indicated a better prognosis. Tumor microenvironment (TME) is intricate and constantly changing. Innate and adaptive immune cells are included in the TME in addition to stromal cells, fibroblasts, and endothelial cells. [47]. Immune cells significantly influence how tumors behave and respond to treatment [48]. The creation of efficient immunotherapies for the treatment of cancer patients was aided by a better knowledge of immunoregulatory mechanisms [49]. Comparing the low-risk group to the high-risk group in this research, immune cell infiltration was more noticeable in the low-risk group. OSGs may affect the immunological milieu of the tumor microenvironment, although it is yet unknown if this has any effect on the onset and progression of osteosarcoma patients. In order to identify the underlying mechanism, function analysis was then carried out. First, the function of DEGs that were filtered out between the high-risk and low-risk groups was shown by GO enrichment analysis to be substantially correlated with immunity. A number of KEGGs that associated with immunity were also more prominent in the low-risk group than in the high-risk group, according to the GSVA heatmap. Recent research has shown that enhancing oxidative stress in tumor cells may effectively trigger immunogenic cell death (ICD) and improve anticancer immunity [50]. The rationally designed tumor specific enhanced oxidative stress polymer compound (TSEOP) could effectively induce ICD in tumor cells with a well-tailored structure that increased oxidative stress in the tumor cells selectively and cooperatively without harming normal cell [51]. The relationship between oxidative stress and both apoptosis and autophagy are complex and reciprocal. Redox molecules are a double-edged sword in the regulation of cell viability [52]. In this study, some clues were found that oxidative stress affected the osteosarcoma's immune status and development. This study could help to further clarify the specific mechanism and provide a new insight into potential targeted therapies for osteosarcoma.

We investigated the fundamental mechanism in the development and therapy of osteosarcoma by bioinformatics analysis based on the link between oxidative stress and immunity. Nevertheless, some limitations of this study should not be ignored. First, the clinical characteristics of osteosarcoma patients were limited. To confirm these findings, experiment validation should be carried out next. The third reason is that this research was created as a retrospective analysis. Thus, further research in this area should be conducted to investigate the relationship between oxidative stress and osteosarcoma in great depth.

Conclusions

As a conclusion, this research gave new information on the investigation of molecular processes and targeted therapeutics for osteosarcoma and highlighted the significance of the interplay between oxidative stress and immunity on the development of osteosarcoma.

Abbreviations

OS oxidative stress

TARGET	The Therapeutically Applicable Research to Generate Effective Treatments
GTE _x	GenotypeTissue Expression
GEO	Gene Expression Omnibus
NMF	Nonnegative Matrix Factorization
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
GSVA	Gene Set Variation Analysis
OD	Oldenlandia diffusa
MLP	Multilayer Perceptron
SM	Skip Metastases
ROS	Reactive Oxygen Species

Declarations

Author Contribution Statement

DL, TM, and QS conceived and designed the study and drafted the manuscript. DL, HW revised the manuscript for important intellectual content. DL, WS, and MH collected, analyzed, and interpreted the experimental data. DL and YC grammar edited the manuscript. All authors read and approved the final manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public or commercial.

Data Availability Statement

All original data included in the manuscript are available upon reasonable request.

Acknowledgments

Not applicable

Conflict of Interest Statement

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Brown HK, Tellez-Gabriel M, Heymann D. Cancer stem cells in osteosarcoma. *Cancer Lett.* 2017;386:189–95.
2. Sajadi KR, Heck RK, Neel MD, Rao BN, Daw N, Rodriguez-Galindo C, et al. The incidence and prognosis of osteosarcoma skip metastases. *Clin Orthop Relat Res.* 2004(426):92–6.
3. Kaste SC, Pratt CB, Cain AM, Jones-Wallace DJ, Rao BN. Metastases detected at the time of diagnosis of primary pediatric extremity osteosarcoma at diagnosis: imaging features. *Cancer.* 1999;86(8):1602–8.
4. Saifuddin A, Michelagnoli M, Pressney I. Skip metastases in high-grade intramedullary appendicular osteosarcoma: an indicator of more aggressive disease? *Skeletal Radiol.* 2021;50(12):2415–22.
5. Lussier DM, Johnson JL, Hingorani P, Blattman JN. Combination immunotherapy with alpha-CTLA-4 and alpha-PD-L1 antibody blockade prevents immune escape and leads to complete control of metastatic osteosarcoma. *J Immunother Cancer.* 2015;3:21.
6. Ock CY, Hwang JE, Keam B, Kim SB, Shim JJ, Jang HJ, et al. Genomic landscape associated with potential response to anti-CTLA-4 treatment in cancers. *Nat Commun.* 2017;8(1):1050.
7. Klaunig JE. Oxidative Stress and Cancer. *Curr Pharm Des.* 2018;24(40):4771–8.
8. Alpay M, Backman LR, Cheng X, Dukel M, Kim WJ, Ai L, et al. Oxidative stress shapes breast cancer phenotype through chronic activation of ATM-dependent signaling. *Breast Cancer Res Treat.* 2015;151(1):75–87.
9. Harris CC. Interindividual variation among humans in carcinogen metabolism, DNA adduct formation and DNA repair. *Carcinogenesis.* 1989;10(9):1563–6.
10. Gupta RK, Patel AK, Kumari R, Chugh S, Shrivastav C, Mehra S, et al. Interactions between oxidative stress, lipid profile and antioxidants in breast cancer: a case control study. *Asian Pac J Cancer Prev.* 2012;13(12):6295–8.
11. Valko M, Rhodes CJ, Moncol J, Izakovic M, Mazur M. Free radicals, metals and antioxidants in oxidative stress-induced cancer. *Chem Biol Interact.* 2006;160(1):1–40.
12. Gorrini C, Harris IS, Mak TW. Modulation of oxidative stress as an anticancer strategy. *Nat Rev Drug Discov.* 2013;12(12):931–47.
13. Wang Z, Li Z, Ye Y, Xie L, Li W. Oxidative Stress and Liver Cancer: Etiology and Therapeutic Targets. *Oxid Med Cell Longev.* 2016;2016:7891574.
14. Hecht F, Pessoa CF, Gentile LB, Rosenthal D, Carvalho DP, Fortunato RS. The role of oxidative stress on breast cancer development and therapy. *Tumour Biol.* 2016;37(4):4281–91.
15. Chiang SK, Chen SE, Chang LC. The Role of HO-1 and Its Crosstalk with Oxidative Stress in Cancer Cell Survival. *Cells.* 2021;10(9).
16. Yang Y, Long X, Li K, Li G, Yu X, Wen P, et al. Development and validation of an oxidative stress-associated prognostic risk model for melanoma. *PeerJ.* 2021;9:e11258.

17. Nourazarian AR, Kangari P, Salmaninejad A. Roles of oxidative stress in the development and progression of breast cancer. *Asian Pac J Cancer Prev*. 2014;15(12):4745–51.
18. Bielack SS, Hecker-Nolting S, Blattmann C, Kager L. Advances in the management of osteosarcoma. *F1000Res*. 2016;5:2767.
19. Kager L, Zoubek A, Potschger U, Kastner U, Flege S, Kempf-Bielack B, et al. Primary metastatic osteosarcoma: presentation and outcome of patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols. *J Clin Oncol*. 2003;21(10):2011–8.
20. Gorlick R, Janeway K, Lessnick S, Randall RL, Marina N, Committee COGBT. Children's Oncology Group's 2013 blueprint for research: bone tumors. *Pediatr Blood Cancer*. 2013;60(6):1009–15.
21. Pan KL, Chan WH, Chia YY. Initial symptoms and delayed diagnosis of osteosarcoma around the knee joint. *J Orthop Surg (Hong Kong)*. 2010;18(1):55–7.
22. Thanindratarn P, Dean DC, Nelson SD, Hornicek FJ, Duan Z. Advances in immune checkpoint inhibitors for bone sarcoma therapy. *J Bone Oncol*. 2019;15:100221.
23. Hellstrom I, Sjogren HO. In vitro demonstration of humoral and cell-bound immunity against common specific transplantation antigen(s) of adenovirus 12-induced mouse and hamster tumors. *J Exp Med*. 1967;125(6):1105–18.
24. Gill JG, Piskounova E, Morrison SJ. Cancer, Oxidative Stress, and Metastasis. *Cold Spring Harb Symp Quant Biol*. 2016;81:163–75.
25. Appierto V, Tiberio P, Villani MG, Cavadini E, Formelli F. PLAB induction in fenretinide-induced apoptosis of ovarian cancer cells occurs via a ROS-dependent mechanism involving ER stress and JNK activation. *Carcinogenesis*. 2009;30(5):824–31.
26. Afzal S, Jensen SA, Sorensen JB, Henriksen T, Weimann A, Poulsen HE. Oxidative damage to guanine nucleosides following combination chemotherapy with 5-fluorouracil and oxaliplatin. *Cancer Chemother Pharmacol*. 2012;69(2):301–7.
27. Moloney JN, Cotter TG. ROS signalling in the biology of cancer. *Semin Cell Dev Biol*. 2018;80:50–64.
28. Morotti M, Zois CE, El-Ansari R, Craze ML, Rakha EA, Fan SJ, et al. Increased expression of glutamine transporter SNAT2/SLC38A2 promotes glutamine dependence and oxidative stress resistance, and is associated with worse prognosis in triple-negative breast cancer. *Br J Cancer*. 2021;124(2):494–505.
29. Kundaktepe BP, Sozer V, Durmus S, Kocael PC, Kundaktepe FO, Papila C, et al. The evaluation of oxidative stress parameters in breast and colon cancer. *Medicine (Baltimore)*. 2021;100(11):e25104.
30. Wu Z, Wang L, Wen Z, Yao J. Integrated analysis identifies oxidative stress genes associated with progression and prognosis in gastric cancer. *Sci Rep*. 2021;11(1):3292.
31. Pulley JM, Jerome RN, Ogletree ML, Bernard GR, Lavieri RR, Zaleski NM, et al. Motivation for Launching a Cancer Metastasis Inhibition (CMI) Program. *Target Oncol*. 2018;13(1):61–8.
32. Lehrer S, Rheinstein PH. The ADRB1 (Adrenoceptor Beta 1) and ADRB2 genes significantly co-express with commonly mutated genes in prostate cancer. *Discov Med*. 2020;30(161):163–71.

33. Soon PS, Libe R, Benn DE, Gill A, Shaw J, Sywak MS, et al. Loss of heterozygosity of 17p13, with possible involvement of ACADVL and ALOX15B, in the pathogenesis of adrenocortical tumors. *Ann Surg.* 2008;247(1):157–64.
34. Zhan K, Bai Y, Liao S, Chen H, Kuang L, Luo Q, et al. Identification and validation of EPHX2 as a prognostic biomarker in hepatocellular carcinoma. *Mol Med Rep.* 2021;24(3).
35. Chen X, Kang R, Kroemer G, Tang D. Broadening horizons: the role of ferroptosis in cancer. *Nat Rev Clin Oncol.* 2021;18(5):280–96.
36. Han L, Bai L, Qu C, Dai E, Liu J, Kang R, et al. PPARG-mediated ferroptosis in dendritic cells limits antitumor immunity. *Biochem Biophys Res Commun.* 2021;576:33–9.
37. Iacobas S, Iacobas DA. A Personalized Genomics Approach of the Prostate Cancer. *Cells.* 2021;10(7).
38. Duffy MJ, O'Grady S, Tang M, Crown J. MYC as a target for cancer treatment. *Cancer Treat Rev.* 2021;94:102154.
39. Leopold JA, Cap A, Scribner AW, Stanton RC, Loscalzo J. Glucose-6-phosphate dehydrogenase deficiency promotes endothelial oxidant stress and decreases endothelial nitric oxide bioavailability. *FASEB J.* 2001;15(10):1771–3.
40. Leopold JA, Zhang YY, Scribner AW, Stanton RC, Loscalzo J. Glucose-6-phosphate dehydrogenase overexpression decreases endothelial cell oxidant stress and increases bioavailable nitric oxide. *Arterioscler Thromb Vasc Biol.* 2003;23(3):411–7.
41. Sakauchi C, Wakatsuki H, Ichijo H, Hattori K. Pleiotropic properties of ASK1. *Biochim Biophys Acta Gen Subj.* 2017;1861(1 Pt A):3030–8.
42. Huang S, Xu Y, Yue L, Wei S, Liu L, Gan X, et al. Evaluating the risk of hypertension using an artificial neural network method in rural residents over the age of 35 years in a Chinese area. *Hypertens Res.* 2010;33(7):722–6.
43. Tang ZH, Liu J, Zeng F, Li Z, Yu X, Zhou L. Comparison of prediction model for cardiovascular autonomic dysfunction using artificial neural network and logistic regression analysis. *PLoS One.* 2013;8(8):e70571.
44. Harrison RF, Kennedy RL. Artificial neural network models for prediction of acute coronary syndromes using clinical data from the time of presentation. *Ann Emerg Med.* 2005;46(5):431–9.
45. Colak MC, Colak C, Kocaturk H, Sagioglu S, Barutcu I. Predicting coronary artery disease using different artificial neural network models. *Anadolu Kardiyol Derg.* 2008;8(4):249–54.
46. Hirose H, Takayama T, Hozawa S, Hibi T, Saito I. Prediction of metabolic syndrome using artificial neural network system based on clinical data including insulin resistance index and serum adiponectin. *Comput Biol Med.* 2011;41(11):1051–6.
47. Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. *Cancer Res.* 2019;79(18):4557–66.
48. Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer Lett.* 2017;387:61–8.

49. Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. *Nat Immunol.* 2013;14(10):1014–22.

50. Krysko DV, Garg AD, Kaczmarek A, Krysko O, Agostinis P, Vandenabeele P. Immunogenic cell death and DAMPs in cancer therapy. *Nat Rev Cancer.* 2012;12(12):860–75.

51. Ma S, Song W, Xu Y, Si X, Lv S, Zhang Y, et al. Rationally Designed Polymer Conjugate for Tumor-Specific Amplification of Oxidative Stress and Boosting Antitumor Immunity. *Nano Lett.* 2020;20(4):2514–21.

52. Ortona E, Maselli A, Delunardo F, Colasanti T, Giovannetti A, Pierdominici M. Relationship between redox status and cell fate in immunity and autoimmunity. *Antioxid Redox Signal.* 2014;21(1):103–22.

Tables

Table 1 Summary of data sets used in this research.

Dataset	Platform	Sample size(n)	Status/version
TARGET- OS	Illumina	88	July 23, 2019
GTEx-muscle	Illumina	803/396	June 05, 2017/ April 27, 2016
GSE21257	GPL10295	53	Feb 11, 2011
GSE16091	GPL96	34	Nov 04, 2009
GSE87437	GPL570	21	Feb 09, 2017
GSE157322	GPL15207	20	Sep 03, 2020
GSE33382	GPL10295	87	Mar 28, 2012
GSE142162	GPL16311	79	Apr 29, 2021
GSE14827	GPL570	27	Feb 01, 2010

Table 2 Genes included for construction prognostic-related gene signature

Gene symbol	coefficient	Full name	Category
TBXA2R	0.44883	Thromboxane A2 Receptor	Protein Coding
ADRB1	0.80081	Adrenoceptor Beta 1	Protein Coding
ACADVL	1.54804	Acyl-CoA Dehydrogenase Very Long Chain	Protein Coding
		Epoxide Hydrolase 2	Protein Coding
EPHX2	1.17574	Peroxisome Proliferator Activated Receptor Gamma	Protein Coding
PPARG	-0.71903	Torsin Family 1 Member A	Protein Coding
		MYC Proto-Oncogene	Protein Coding
TOR1A	-1.73223	Glucose-6-Phosphate Dehydrogenase	Protein Coding
MYC	0.59661		Protein Coding
G6PD	-0.52272		Protein Coding
			Protein Coding
MAP3K5	-1.53225	Mitogen-Activated Protein Kinase Kinase Kinase 5	Protein Coding

Figures

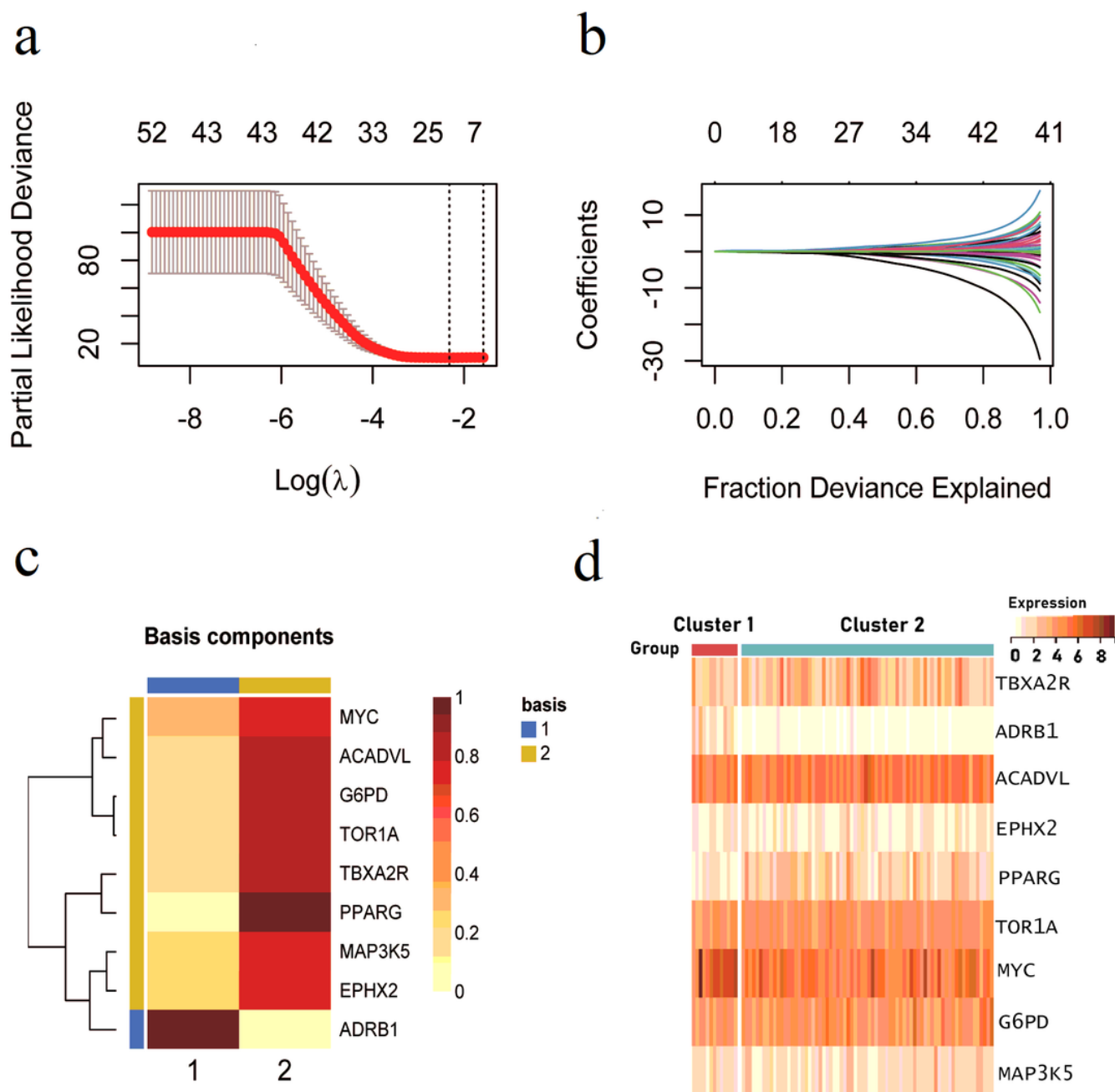


Figure 1

Screening of molecular subgroups through nonnegative matrix factorization (NMF) cluster. a-b Seventeen candidate genes screened out by LASSO analysis with minimal lambda. c Two subgroups were identified the optimal value for consensus clustering. d Heatmap visualizing the expression of candidate oxidative stress-related genes in the 2 subgroups.

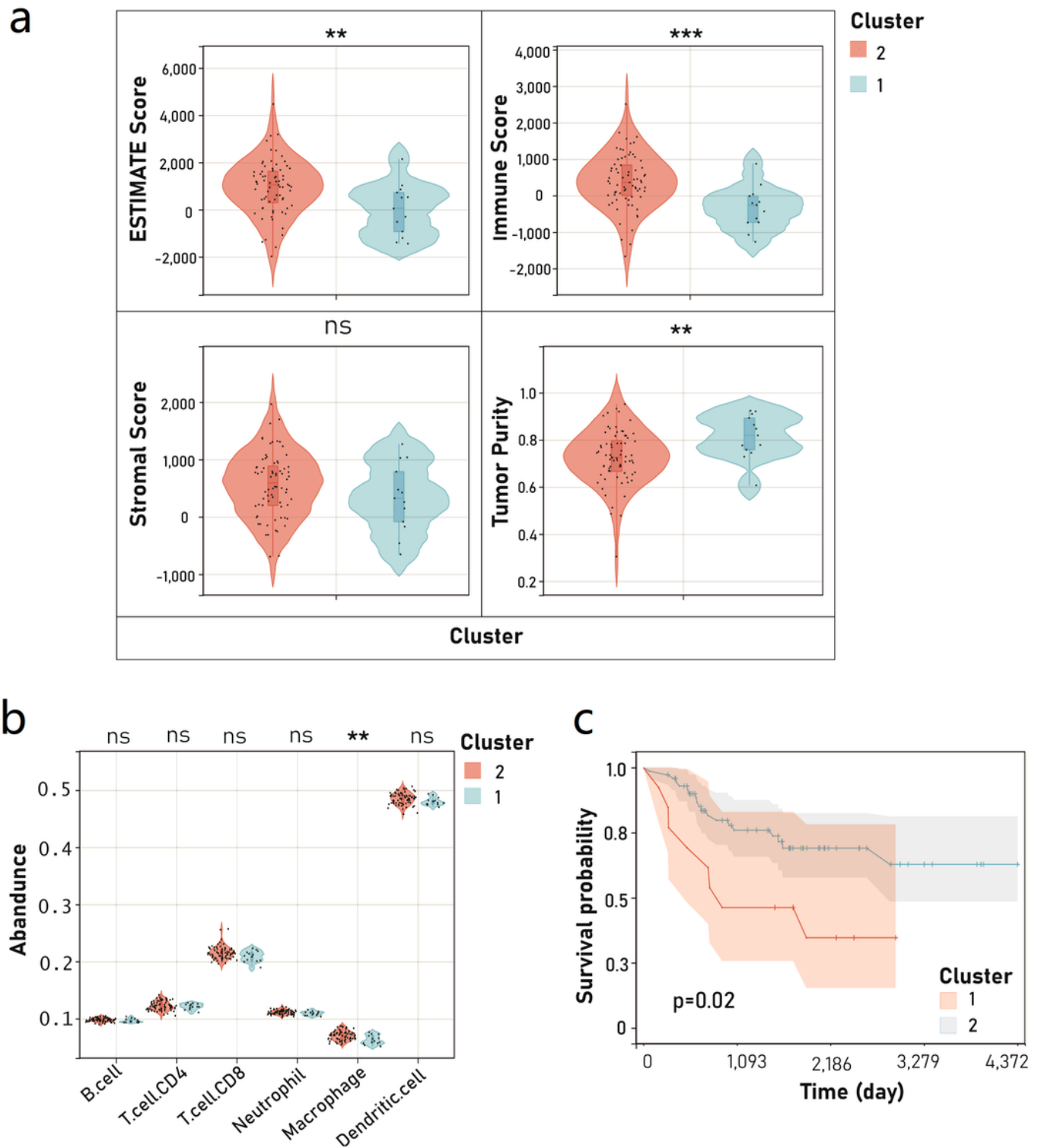


Figure 2

Comparisons between the 2 subgroups based on NMF cluster. Calculating terms of ESTIMATE score, Immune score, Stromal score, Tumor purity (a), and abundance of immune infiltrating cells (b) in tumor tissues; and survival curve of patients (c).

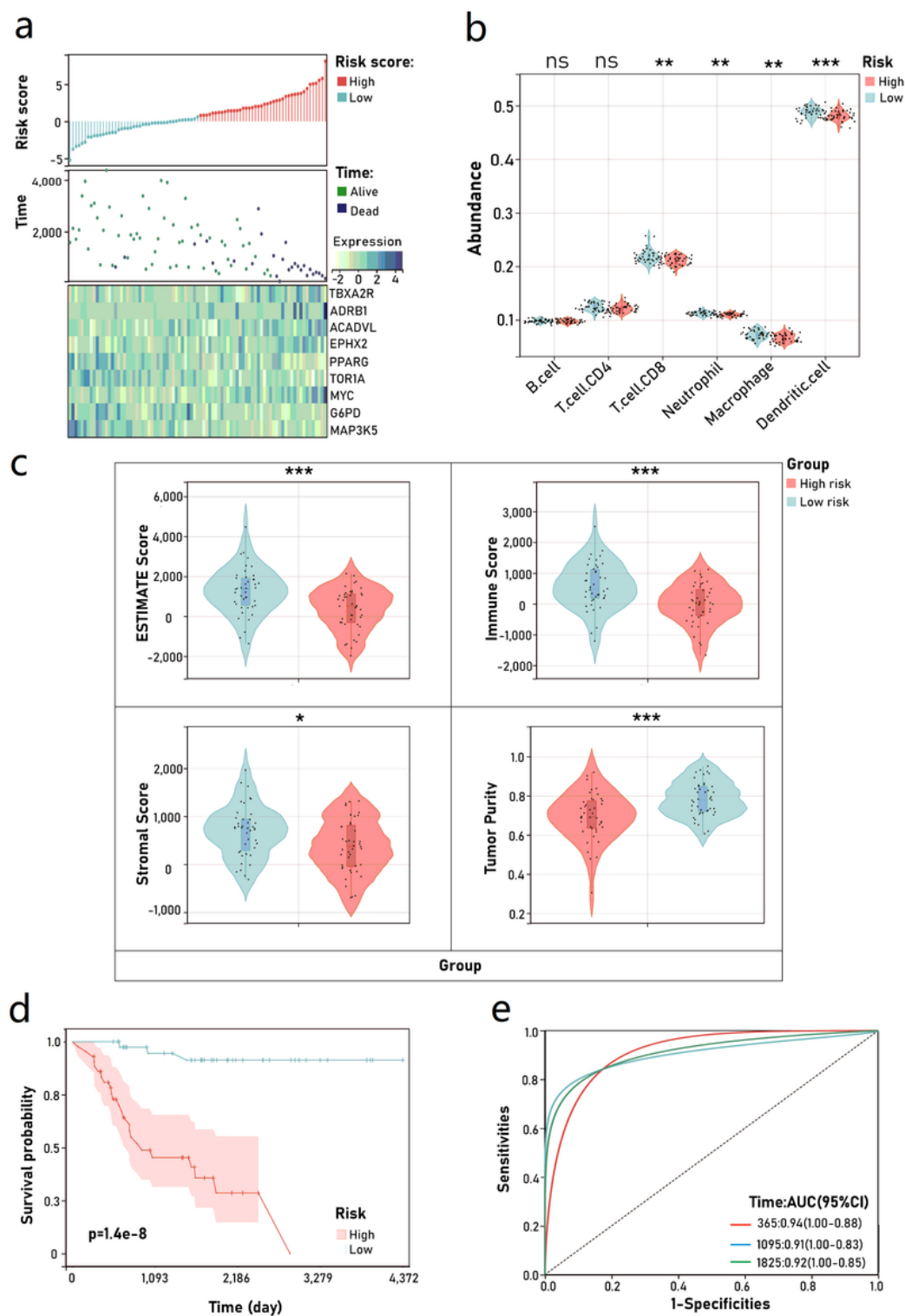


Figure 3

Construction of the prognosis risk prediction model in the training cohort. a Distribution of risk score and survival status of each osteosarcoma patient in the high risk and low risk and heatmap visualizing the expression matrix of the 9 candidate genes in the 2 groups. b-d Comparisons between the 2 groups in terms of abundance of immune filtering cells and ESTIMATE score, Immune score, Stromal score, Tumor

purity in tumor tissues and survival curve of the patients. e Time- dependent ROC curve of the constructed model.

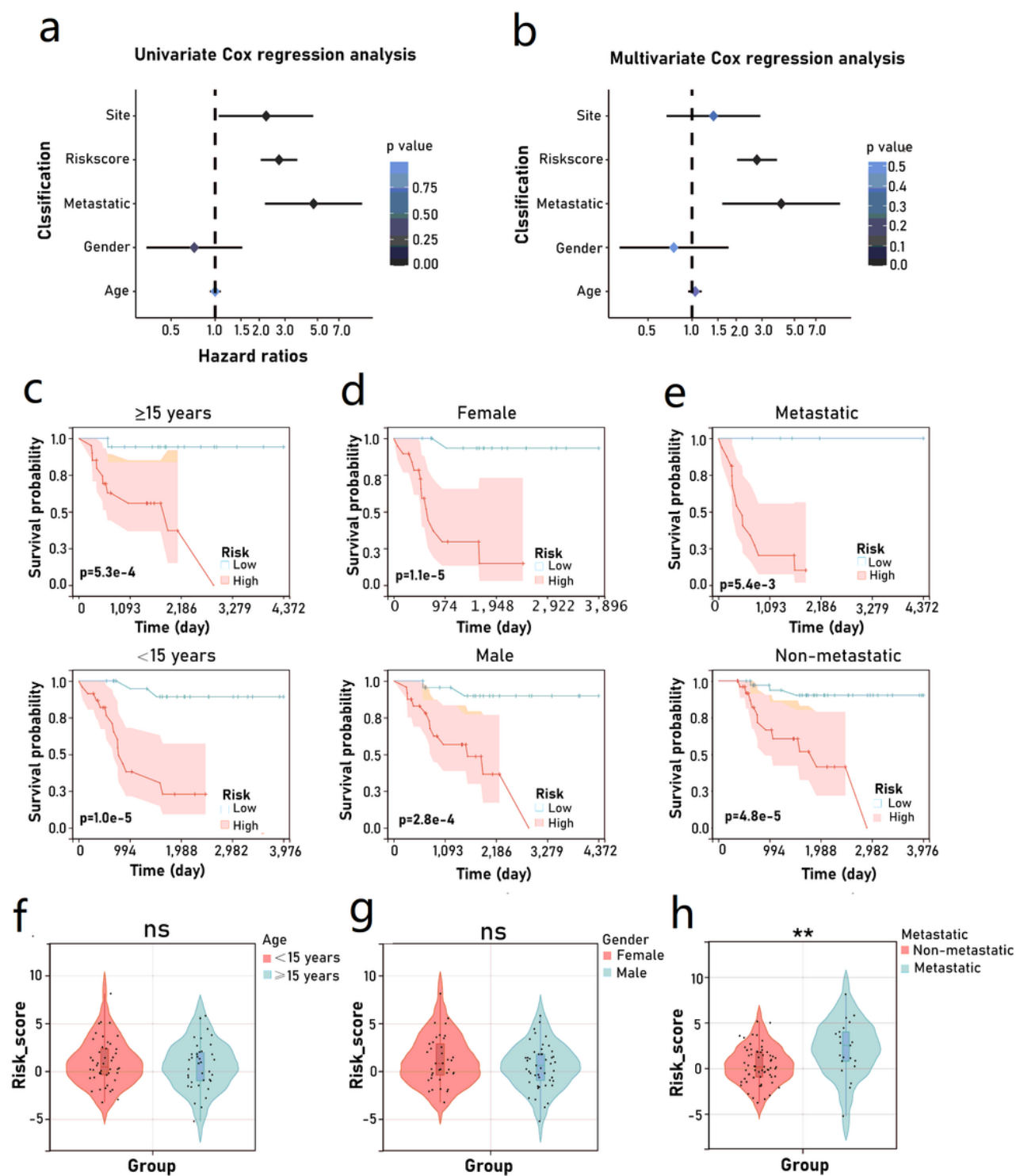


Figure 4

Independence detection of the constructed risk prediction model. a-b Independent prognostic-related factors screened out by COX regression analysis. c-e Subgroups analysis of survival curve in accordance

with different age, sex and metastasis status. f-h Subgroups analysis of risk score in accordance with different age, sex and metastasis status.

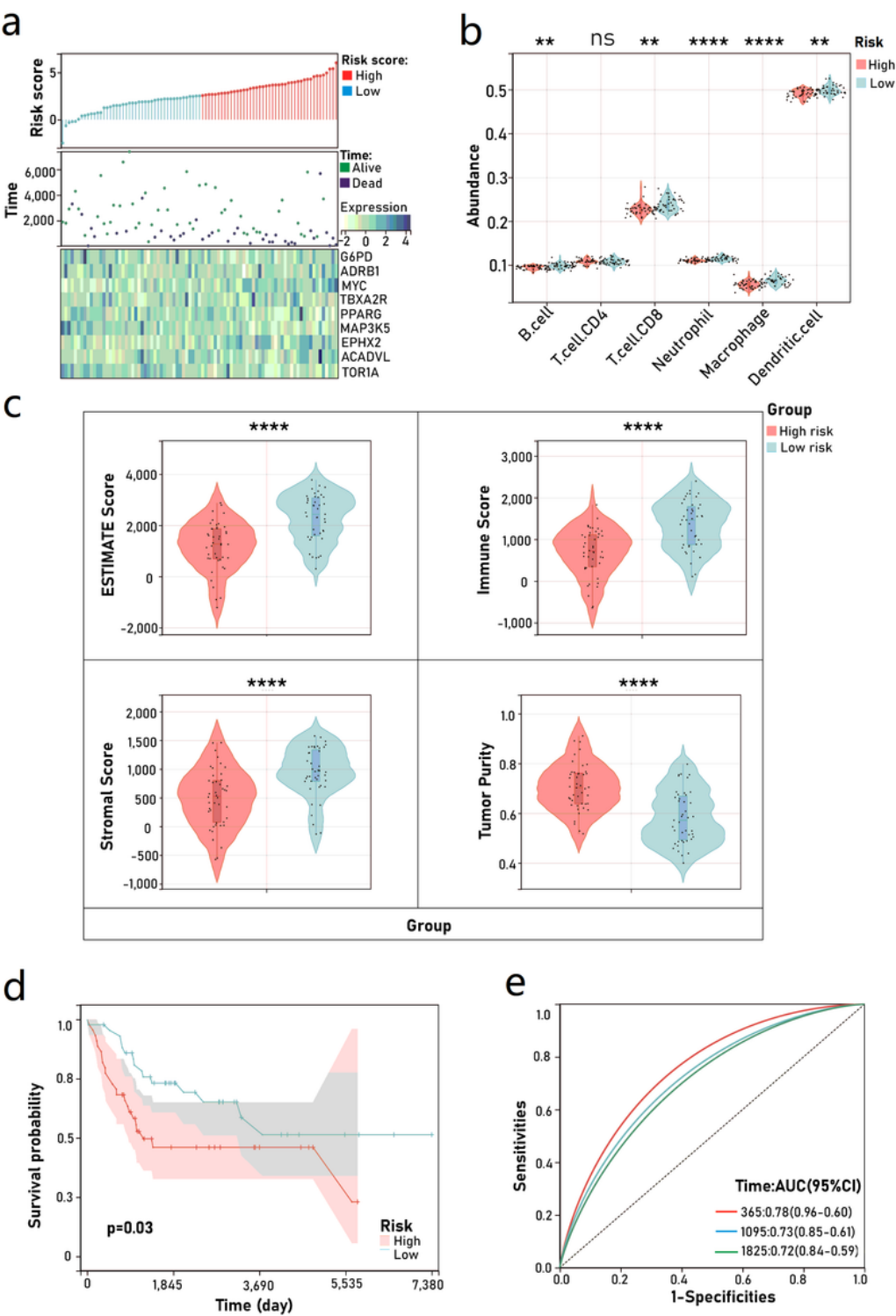


Figure 5

Validation of the constructed risk prediction model in the verification cohort. a Distribution of survival status and risk score of osteosarcoma patients in the verification cohort and heatmap visualizing the

expression matrix of 9 candidate genes in the verification cohort. b-d Comparisons between the high risk and low risk groups in terms of abundance of immune infiltrating cells and stromal score, immune score, ESTIMATE score, tumor purity in tumor and survival curve in the verification cohort. e ROC curve of the risk prediction model in the verification cohort.

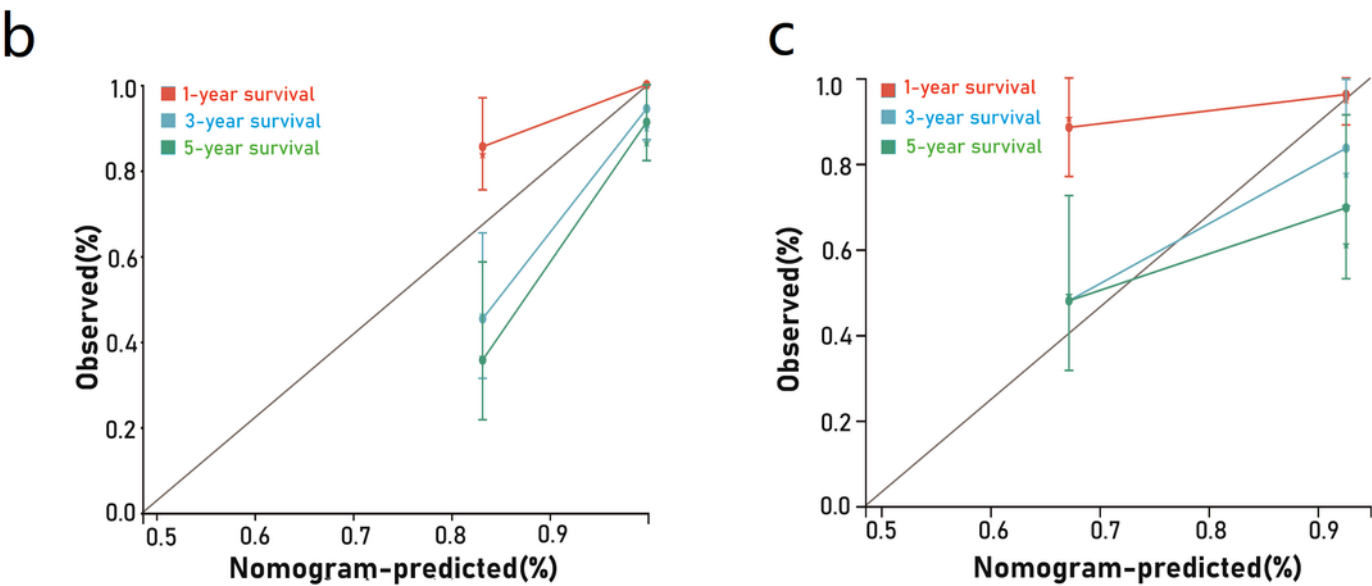
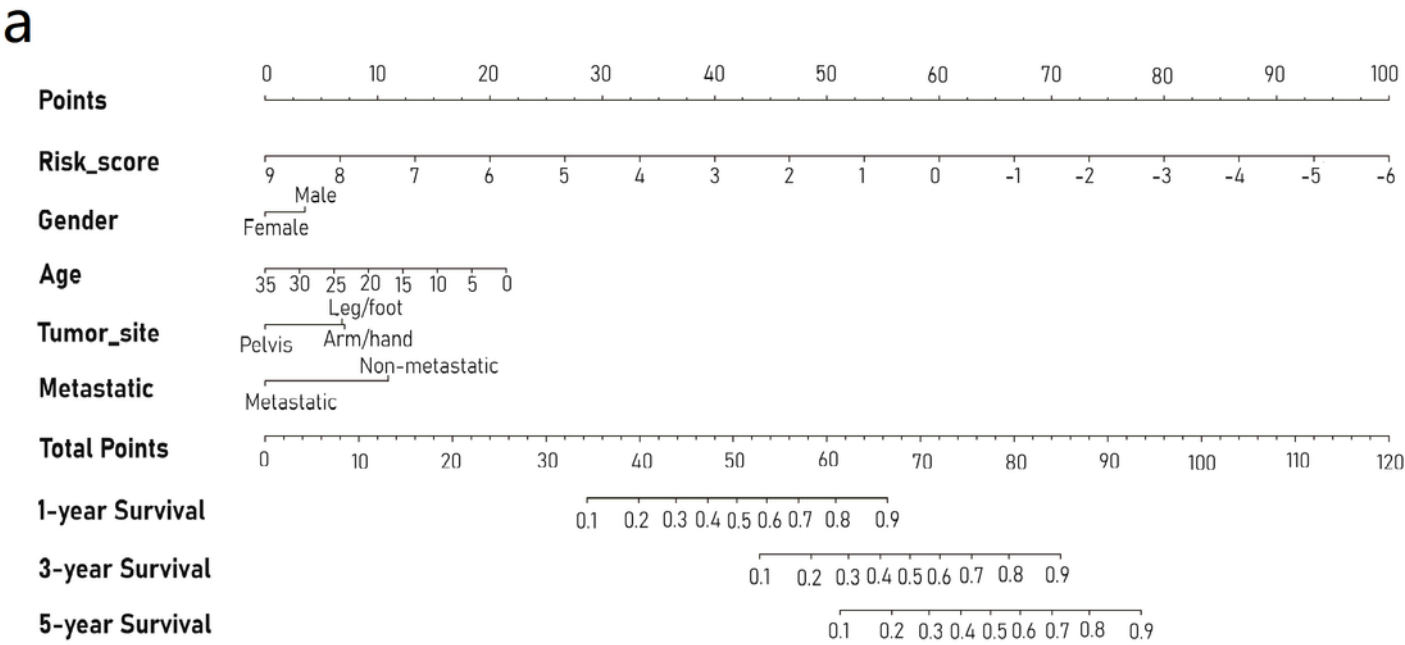


Figure 6

Construction and calibration of nomogram. a Nomogram integrating risk score and clinical characteristics. b calibration of the nomogram at 1-y, 3-y and 5-y survival in the training cohort. c Calibration of the nomogram at y-y, 3-y and 5-y survival in the verification cohort.

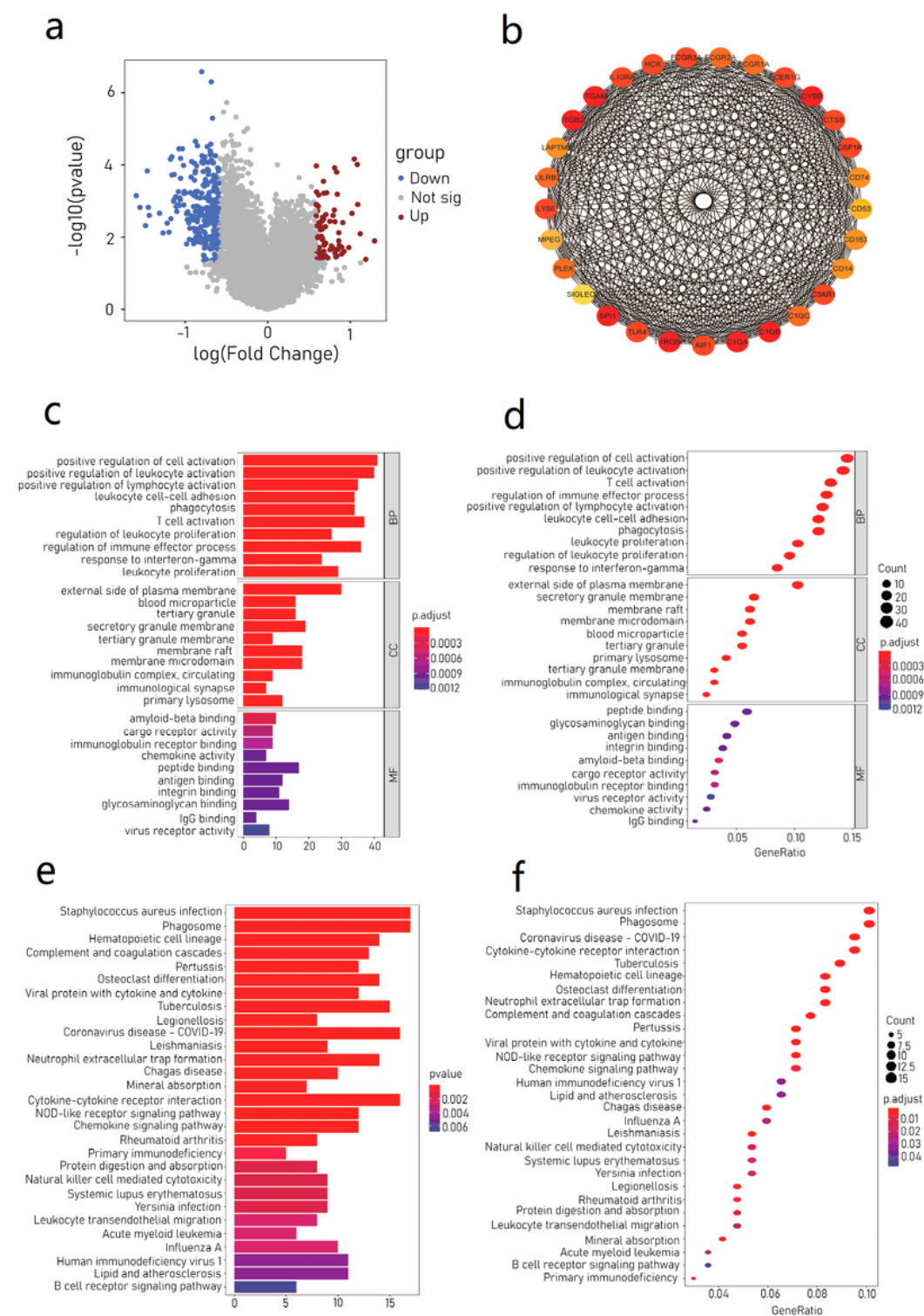


Figure 7

Differentially expressed genes analysis and gene function enrichment analyses. a Volcano plot showing the DEGs between the high risk and low risk groups. b The first module of PPI analysis of DEGs. (c) Bar plot and (d) bubble plot of gene ontology (GO) enrichment pathway. (e) Bar plot and (f) bubble plot of KEGG enrichment pathway.

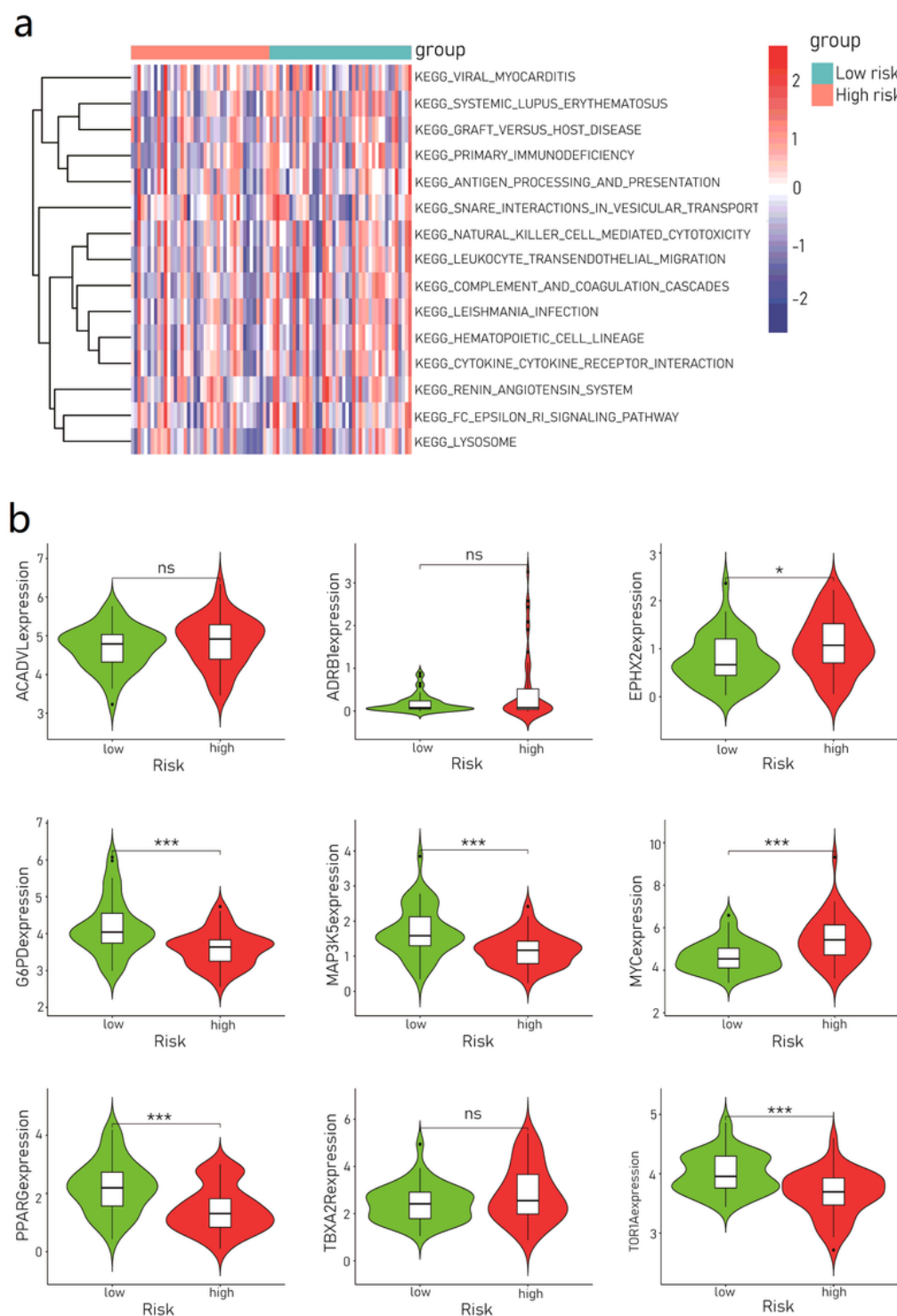


Figure 8

Enrichment pathways of GSVA. a Heatmap illustrating the result of GSVA. b Different expression of nine candidate genes in the high- and low-risk groups from the training cohort.

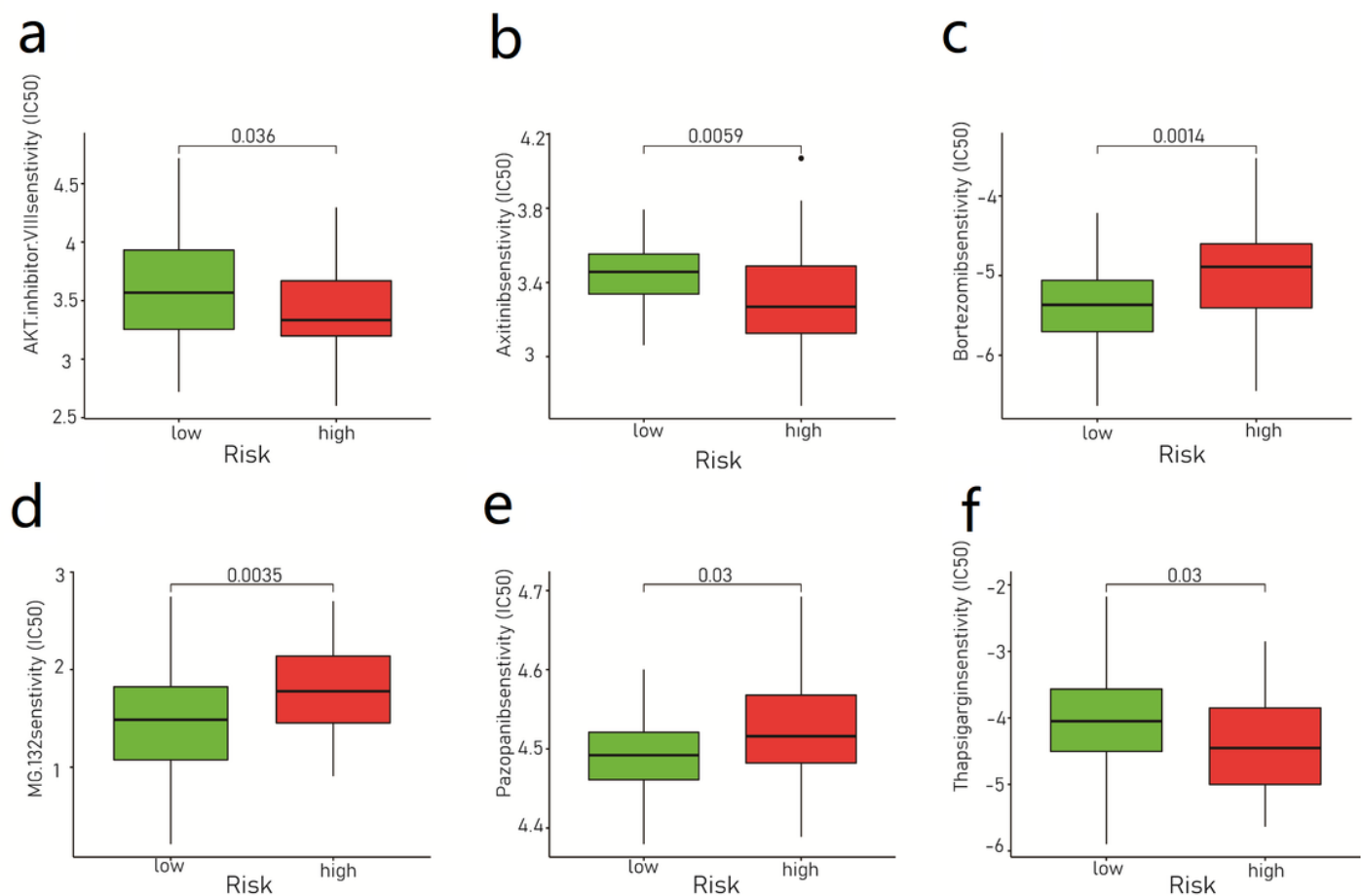


Figure 9

Prediction of response to targeted drugs. (a) AKT.inhibitor.VIII. (b) Axitinib. (c) Bortezomib. (d) MG.132. (e) Pazopanib. (f) Thapsigargin.

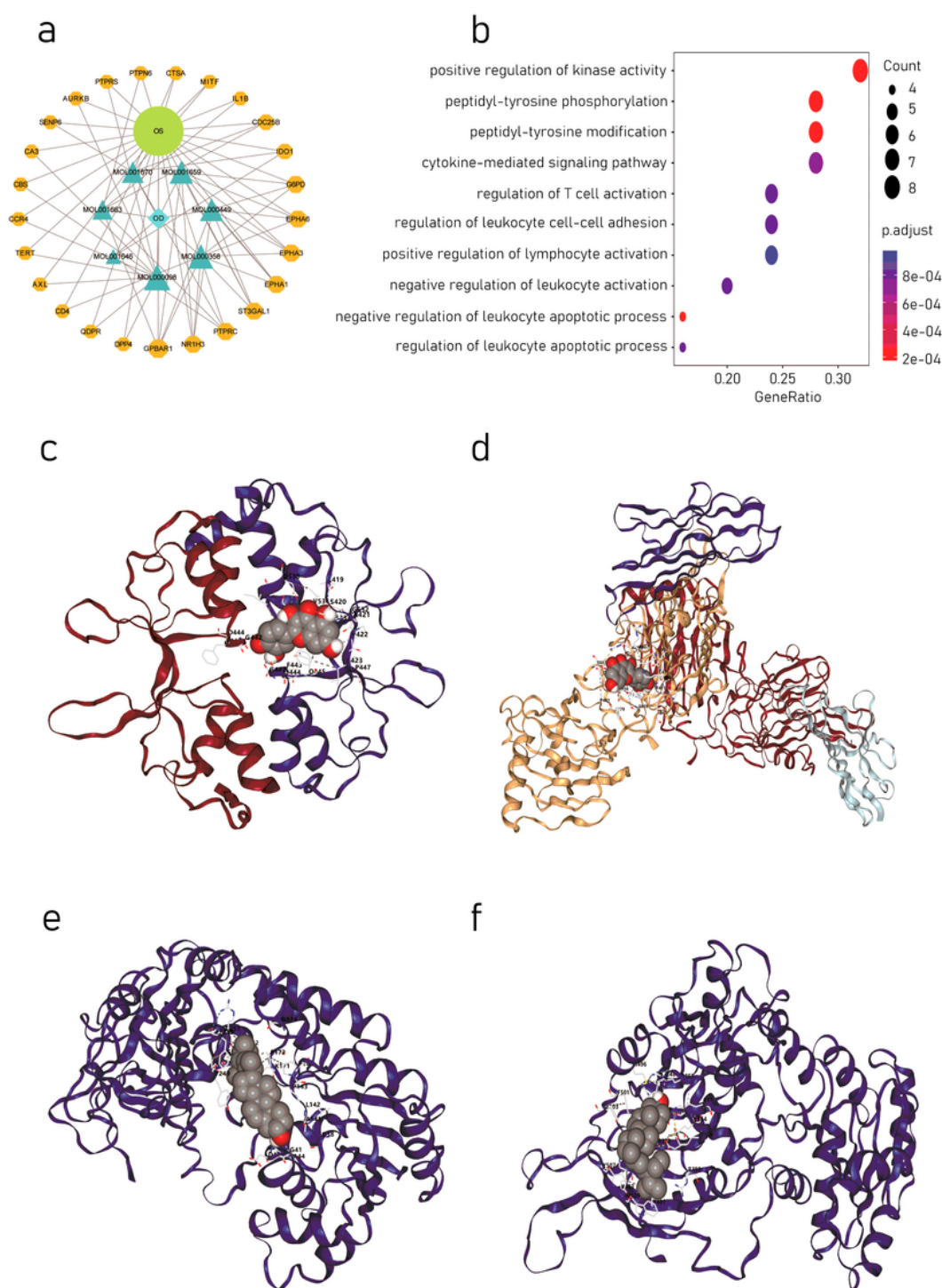


Figure 10

Further analyse of prognostic model basing on Network Pharmacology and molecular docking. a The drug-components-targets-disease network (The green ellipse node is osteosarcoma, navy blue diamond is active Ingredients of OD, yellow octagon is Targets of OD, baby blue diamond is OD). b GO enrichment analysis (the top 10 results of BP). c Molecular docking results of CBS and quercetin (binding energy -9.7). d Molecular docking results of AXL and quercetin (binding energy -9.1). e Molecular docking results

of G6PD and Poriferasterol (binding energy -8.6). f Molecular docking results of G6PD and Stigmasterol (binding energy -8.1).

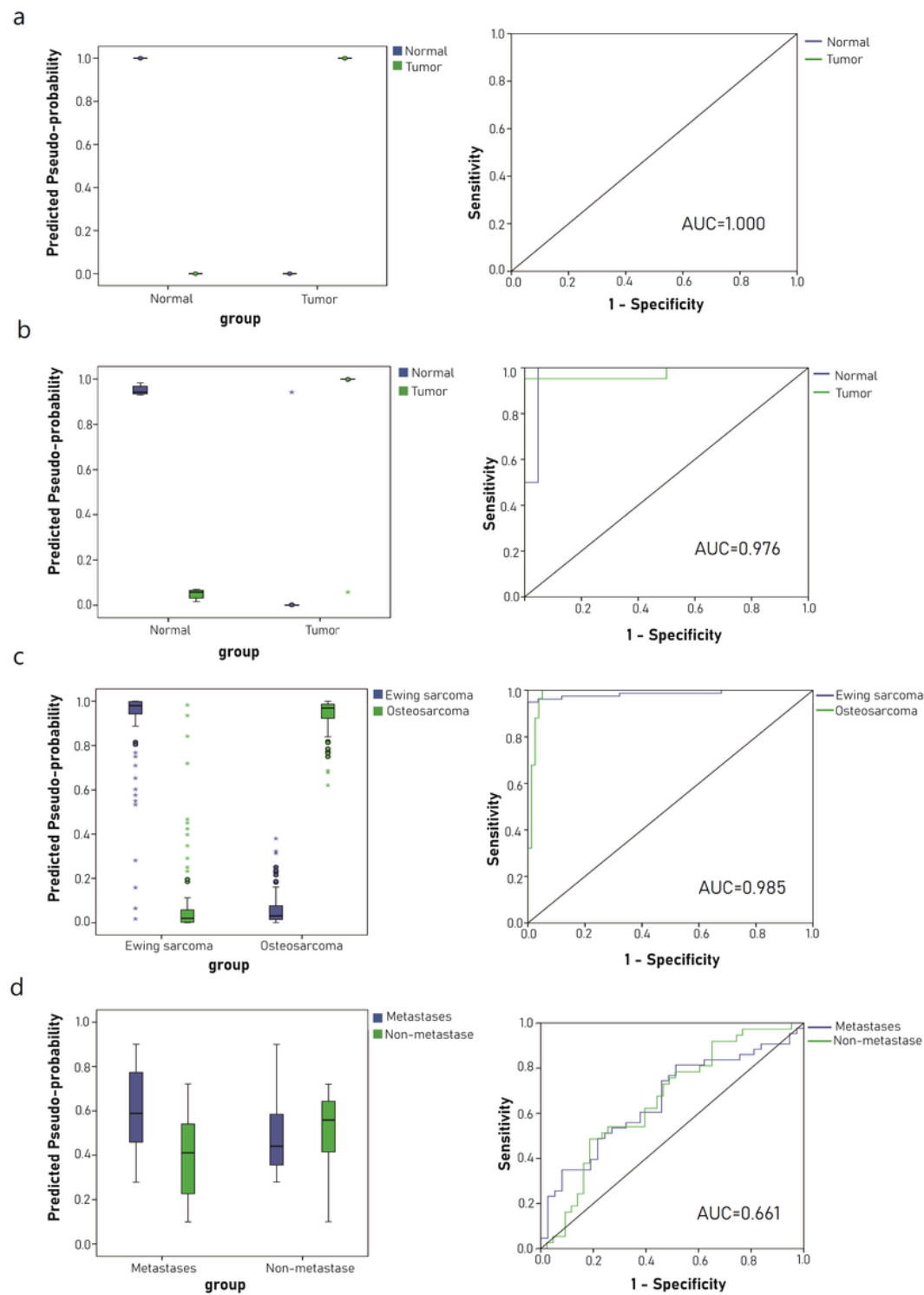


Figure 11

Rediction probability histograms and ROC curves for neural network models. a-b Normal vs Tumor based on TARGET and GEO datasets. c Ewing sarcoma vs Osteosarcoma. d Metastases vs non-metastase

samples.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile.doc](#)