A Prognostic Risk Model of a Novel Oxidative Stress-Related Signature Predicts Clinical Prognosis and Demonstrates Immune Relevancy in Lung Adenocarcinoma

Xing Huang, Zhichao Lu, Min He, Yipeng Feng, Shaorong Yu, Bo Shen, Jianwei Lu, Pingping Wu, Banzhou Pan, Hanlin Ding, Chen Chen, and Yidan Sun

1Department of Pathology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & Nanjing Medical University Affiliated Cancer Hospital, China
2Research Center of Clinical Medicine, Affiliated Hospital of Nantong University, Nantong 226001, China
3Department of Thoracic Surgery, Nanjing Medical University Affiliated Cancer Hospital & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, 21009 Nanjing, China
4Jiangsu Key Laboratory of Molecular and Translational Cancer Research, Cancer Institute of Jiangsu Province, Nanjing, China
5The Fourth Clinical College of Nanjing Medical University, Nanjing, China
6Department of Oncology, Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research & The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing 210000, China
7The Comprehensive Cancer Centre of Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, China
8Department of Oncology, First Teaching Hospital of Tianjin University of Traditional Chinese Medicine, Tianjin, China

Correspondence should be addressed to Chen Chen; chenchen881021@outlook.com and Yidan Sun; huiranzhongyi@126.com

Received 18 June 2022; Revised 10 October 2022; Accepted 13 October 2022; Published 18 November 2022

Lung adenocarcinoma (LUAD) is among the most prevalent malignant lung cancers with a poor prognosis due to high invasiveness and lethality despite multiple treatments. Since the lung is an important organ associated with oxidative stress, and it has been confirmed that oxidative stress represents a potential cancer-specific depletion, it is of important significance to investigate and evaluate the clinical value of oxidative stress mechanisms regulating tumor cell apoptosis. Furthermore, there are few studies on the impact of the microenvironment on reaction to immune-checkpoint inhibitors (ICIs) in patients with LUAD. Based on the TCGA-LUAD dataset, which is stratified into a training set as well as a validation set in a ratio of 2:1, this investigation constructs and validates a prognostic predictive power of a gene signature model of oxidative stress-related prognostic signatures. To ascertain the differences between the high-risk score group and the low-risk score group in tumor-infiltrating lymphocytes and patients’ response to ICI therapy. This oxidative stress-related prognostic gene signature is composed of MAP3K19 and NTSR1 and is an independent prognosis-related factor in the LUAD group. The outcome of patients having a low risk score is better, and the difference was statistically significant, and individuals with a low risk score had a larger number of infiltrating immune cell distribution in the tumor microenvironment, which was closely related to clinical outcome. Our study suggests that the synergistic effect of oxidative stress-related prognostic gene markers-MAP3K19 and NTSR1 has clinical significance in the prognosis identification and immunotherapy of LUAD patients. Thus, the results may help to better intersect the oxidative stress-related mechanisms in clinical value in LUAD but requires prospective validation.
1. Introduction

Lung cancer is the major important cause of tumor-related mortalities worldwide [1, 2]. LUAD is the most significant subtype. Although many clinical studies have confirmed that multitargeted drugs and immunotherapy can prolong the overall survival (OS) and improve the objective response rate (ORR) of LUAD patients [3–5], the rapid progression of the disease due to multi-drug resistance is currently very difficult. Few targeted treatment options [6–9]. With the in-depth research in many aspects [10, 11], it is beneficial to discover small molecule inhibitors. Furthermore, significant literature confirmed the hypothesis that the tumor microenvironment (TME) enhances tumor growth through paracrine signaling [12]. So, more investigation should be taken to improve the outcome among LUAD patients.

Redox homeostasis is crucial in not only the survival of normal cells but also cancerous cells. Oxidative stress (OS) is predominantly triggered by an imbalance between cellular antioxidant mechanisms and metabolically generated oxidative free radical species [13, 14]. This imbalance eventually causes the excessive buildup of reactive oxygen species (ROS) within body cells, leading to irreversible or reversible injury to the body [15]. Nevertheless, numerous tumors have increased levels of ROS and exhibited signs of chronic oxidative stress as a result of oncogenic injury, hypoxia, metabolic malfunctions, and proteotoxic stress [13]. Increased ROS is hypothesized to enhance the progression of tumors at subthal levels by inducing the mutations and changing cell signaling [16]. Nonetheless, to block excessive oxidative injury, tumors generally upregulate the antioxidant pathways [17]. Consequently, numerous cancerous cells are hypersensitive to perturbation of ROS levels. Excessive oxidative stress is known to aggravate the cytotoxic impacts of chemotherapy, and efforts are being made to enhance ROS generation in these environments [18, 19].

In this study, we attempted to obtain oxidative stress-related expression profiling data for LUAD from The Cancer Genome Atlas (TCGA) database and aggregated clinical information and transcriptomes from 445 patients in TCGA-LUAD with complete clinical informatics data and express the data, dividing it into a training set (2/3 of the total, n = 296) as well as a test set (1/3 of the total, n = 149). We then performed univariate Cox proportional hazards regression in both the training and the validation sets, respectively, to identify genes with prognostic values utilizing the expression data of 147 oxidative stress-related genes. Cox p values < 0.05 indicated a coexpression network of 35 oxidative stress DEGs associated with LUAD overall survival, with clinical information and transcriptomic expression data. Based on the LASSO algorithm, the prognosis-related gene signatures composed of 3 and 8 oxidative stress-related genes were screened in the training set and the test set, respectively. Therefore, we obtained a LUAD prognostic risk model composed of MAP3K19 and NTSR1 by intersection difference analysis. We further validated the subgroup prognostic risk model by Kaplan-Meier (KM) analysis as well as receiver-operating characteristic curve (ROC) analysis. Additionally, the link between the risk model and the TME was ascertained by both the ESTIMATE R software package and the CIBERSORT tool. Lastly, differentially infiltrating immune cells were discovered in the two risk groups, and MAP3K19 and NTSR1 were found to be statistically significant with immune-related genes. In conclusion, our study suggests that oxidative stress-related risk models may provide a viable prognostic tool and important function in the modulation of immune cell distribution in the LUAD tumor microenvironment. Our hypothesis is that this oxidative stress-related prognostic signature in LUAD. The aim of our investigation was to create and validate an oxidative stress-related LUAD prognostic model and to explore its predictive effect on poor prognosis in LUAD. Assess and validate prognostic power and its independent prognostic value. Our goal was to guide the clinical application of this oxidative stress-related prognostic signature in LUAD.

2. Materials and Methods

Figure 1 presents the flow chart for the bioinformatics analysis. R language software (Version 4.0.3) [20] accomplished all the statistical analyses, and p < 0.05 denoted a statistically significant difference without a designated setting.

2.1. Data Acquiring and Cleaning. Transcriptomic data (reads), as well as relevant clinical data encompassing survival status, age, sex, grade, and stage of LUAD patients, were retrieved from TCGA (https://portal.gdc.cancer.gov/) database. In total, 551 samples were incorporated in this investigation (54 normal lung samples and 497 LUAD samples). The initial expression data were normalized with the aid of the trim mean of M values (TMM) algorithm in the "limma" package [21], and genes with an average expression not exceeding 1 were excluded. The "limma" package was additionally utilized for MTG differential expression analysis. The criteria for identifying differentially expressed genes in our study were |logFC| > 1 and adj. p < 0.05. The read counts were transformed to TPM values, and a log2(x + 1) conversion was conducted for further analyses since the TPM values were identical to the microarray values.

We use the k-fold cross-validation method to divide the sample set into k mutually exclusive subsets of similar size, each of which keeps the data distribution as consistent as possible. Then, each time the union of k – 1 subsets is used as the training set, and the remaining subset is used as the test set, so that k sets of training/testing sets can be obtained. Perform k training and testing times, and return the average of the k test results.

2.2. Identification of Differentially Expressed Oxidative Stress-Related Genes with Prognostic Value in LUAD Tissue and Normal Lung Tissue. As per the search term "oxidative stress," from the OMIM database, NCBI gene function module, and the GeneCard database, a total of 9469 human genes which are linked to oxidative stress were collected. Based on this, we intersected these oxidative stress-related genes with...
the differentially expressed genes in the TCGA-LUAD database to set and screen 147 oxidative stress-related genes with significant expression differences for subsequent analysis. We used a string-generated protein interaction (PPI) network database (version 11.0) of 147 oxidative stress-related genes differentially expressed in the TCGA-LUAD dataset to construct a molecular interaction network for analyzing closely interacting differential genes. Then, the PPI was exported, and the Cytoscape [22] software was used for further analysis, the network properties of each node were calculated, and the MCODE [23] and Cytohubba [24] were used to mine the hub nodes based on the degree of the nodes. The high level of linkage may have an extremely important function in the modulation of the whole biological process, which deserves further study.

2.3. Creation and Validation of a Prognostic Oxidative Stress-Related Signature for LUAD. We randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into a training set (2/3 of the total, n = 296) and a test set (1/3 of the total, n = 149). We then performed univariate Cox proportional hazards regression in the training set and validation set, respectively, to discover genes with prognostic value, utilizing the expression data of 147 DE-oxidative stress-related genes. A Cox p value < 0.05 connoted a substantial correlation with overall survival (OS). The degree of correlation and prognostic value were selected using the "Venn" R package for cross-analysis and genes affecting prognostic value. To avert overfitting, all genes with p values < 0.05 were subjected to a least absolute shrinkage and selection operator (LASSO) analysis using the package glmnet. LASSO regression is usually a regularization method for high-dimensional predictor selection. The hazard system used LASSO Cox proportional hazards to build a score-identifying gene signature to anticipate OS models for LUAD patients. Prediction scores are weighted sums using developed genes, with coefficient regularization by LASSO. After being filtered by the LASSO model, the selected genes are constructed by the multivariate Cox proportional hazards model to construct an immune-related risk model: risk score = level of gene a\(^{\ast}\) coefficient a + level of gene b\(^{\ast}\) coefficient b + level of gene c\(^{\ast}\) coefficient c + ⋯ + level of gene n\(^{\ast}\) coefficient n. The risk score in the model represents the prognosis of LUAD patients; the smaller the risk score, the better the prognosis. Patients were categorized into two risk groups utilizing the median risk score that served as a cutoff value. The "pheatmap" R package used scatter diagrams to show how the risk scores and survival durations of all patients were distributed. The "stats" R package used principal component analysis (PCA) to measure the gene expression of an established signature. Kaplan-Meier (K-M) survival analysis and time-dependent ROC analysis based on OS were carried out utilizing the "survival" package, the "survminer" package [25], and the "timeROC" package [26] in R to measure the prognostic accuracy of the gene signature in the derivation set and validate it in the validation set. Kaplan-Meier survival curves were used to derive the predictive power, whereas a log-rank p value < 0.05 denoted a statistical significance (utilizing package survival and survminer). To ascertain predictive power of the immune signature’s, time-dependent ROC curves (package survivalROC) were employed. Subsequently, logistic regression was used for correlation analysis of dichotomous clinical subgroup variables.
2.4. Pathway and Function Enrichment Analysis. To investigate the biological value of these differentially expressed IRGs, we utilized DAVID Bioinformatics Resources 6.8 to conduct a pathway and function enrichment analysis. The visualization procedure was carried out via the package “ggplot2” [25].

2.5. Gene Set Enrichment Analysis (GSEA) of Oxidative Stress Prognostic Model in LUAD. GSEA is a computational approach utilized to ascertain if a predefined set of genes exhibits statistical differences between two biological states, which is typically employed to obtain the expression in a dataset sample. Changes in the pathway as well as activity of biological processes were analyzed [23]. To investigate the variations in biological processes between the two groups of samples, utilizing the gene expression profile dataset, we retrieved the reference gene set “c2.cp.kegg.v7.4.entrez.gmt” from the MSigDB database, using the R package “clusterProfiler.” The GSEA method is included in “enrichment analysis and visualization of datasets. p value < 0.05 was considered statistically significant.

Gene set variation analysis (GSVA) [27] is a nonparametric unsupervised analysis method, which is mainly used to evaluate the microarray by transforming the expression matrix of genes between various samples into the expression matrix of gene sets between samples. Transcriptome gene set enrichment findings assess if various metabolic pathways are enriched across various samples. To investigate the biological process variation occurring in the two groups of samples, we utilized the R package GSVA to carry out the GSVA procedure based on the gene expression profile dataset and downloaded the reference “c2.cp.kegg.v7.4.entrez” gene set from the MSigDB database to compute the enrichment score of every sample in each pathway in the dataset, combined with the R package limma to screen significantly different pathways, the GSVA enrichment results were visualized based on the heatmap utilizing the R package pheatmap, and p value < 0.05 denoted a statistically significant significance.

2.6. Establishment and Verification of Nomogram. Nomogram has been extensively utilized for anticipating cancer-related prognosis. This approach allows individualized approximates of the likelihood of recurrence, mortality, or drug adherence. Using the prognostic model, this investigation developed the nomogram in the TCGA-LUAD cohort by incorporating the above-stated clinical parameters to anticipate the OS probably over 1, 3, and 5 years.

Table 1: Clinical characteristics of patients in TCGA-LUAD internal training set and validation set.

| Characteristics | Train (N = 296) | Test (N = 149) | Total (N = 445) | p value | FDR |
|-----------------|----------------|---------------|----------------|---------|-----|
| Age             |                |               |                |         |     |
| Mean ± SD       | 64.86 ± 10.00  | 65.05 ± 10.21 | 64.93 ± 10.06  |         |     |
| Median [min-max]| 66.00 [33.00, 88.00] | 66.00 [41.00, 86.00] | 66.00 [33.00, 88.00] |         |     |
| Gender          |                |               |                |         |     |
| Female          | 162 (36.40%)   | 81 (18.20%)   | 243 (54.61%)   |         |     |
| Male            | 134 (30.11%)   | 68 (15.28%)   | 202 (45.39%)   |         |     |
| Stage           |                |               |                |         |     |
| Stage I         | 157 (35.28%)   | 87 (19.55%)   | 244 (54.83%)   |         |     |
| Stage II        | 68 (15.28%)    | 36 (8.09%)    | 104 (23.37%)   |         |     |
| Stage III       | 54 (12.13%)    | 21 (4.72%)    | 75 (16.85%)    |         |     |
| Stage IV        | 17 (3.82%)     | 5 (1.12%)     | 22 (4.94%)     |         |     |
| T               |                |               |                |         |     |
| T1              | 108 (24.27%)   | 49 (11.01%)   | 157 (35.28%)   |         |     |
| T2              | 152 (34.16%)   | 80 (17.98%)   | 232 (52.13%)   |         |     |
| T3              | 24 (5.39%)     | 14 (3.15%)    | 38 (8.54%)     |         |     |
| T4              | 12 (2.70%)     | 6 (1.35%)     | 18 (4.04%)     |         |     |
| M               |                |               |                |         |     |
| M0              | 198 (44.49%)   | 100 (22.47%)  | 298 (66.97%)   |         |     |
| M1              | 17 (3.82%)     | 5 (1.12%)     | 22 (4.94%)     |         |     |
| MX              | 81 (18.20%)    | 44 (9.89%)    | 125 (28.09%)   |         |     |
| N               |                |               |                |         |     |
| N0              | 186 (41.80%)   | 105 (23.60%)  | 291 (65.39%)   |         |     |
| N1              | 54 (12.13%)    | 27 (6.07%)    | 81 (18.20%)    |         |     |
| N2              | 49 (11.01%)    | 17 (3.82%)    | 66 (14.83%)    |         |     |
| N3              | 2 (0.45%)      | 0 (0.0e+0)    | 2 (0.45%)      |         |     |
| NX              | 5 (1.12%)      | 0 (0.0e+0)    | 5 (1.12%)      |         |     |
| Features     | p-value     | Hazard ratio (95%CI)                  |
|--------------|-------------|--------------------------------------|
| HAO1         | 1.1e-289    | 21.69 (18.37-25.60)                  |
| HTR1B        | 1.2e-273    | 0.01 (0.01-0.02)                     |
| APOB         | 1.2e-268    | 13.46 (11.64-15.57)                  |
| SLCA65       | 4.6e-211    | 11.41 (9.78-13.30)                   |
| SCA2NA22     | 1.3e-209    | 2.06 (1.97-2.16)                     |
| MUC2         | 5.3e-194    | 7.34 (6.43-8.37)                     |
| PRODH2       | 9.4e-166    | 2.0e-13 (2.4e-14-1.6e-12)            |
| KCNA4        | 2.7e-164    | 4.2e-4 (4.2e-4-7.4e-4)               |
| FOXL2        | 1.4e-121    | 138.42 (91.67-209.01)                |
| NPSR1        | 1.6e-94     | 92.01 (59.87-141.41)                 |
| MRGPRX3      | 4.4e-69     | 56298651.63 (7687524.49-412296335.40) |
| TNR          | 3.8e-34     | 0.07 (0.04-0.10)                     |
| FGFR5        | 1.2e-33     | 0.37 (0.32-0.44)                     |
| MAP3K19      | 8.1e-29     | 0.11 (0.07-0.16)                     |
| NRIH4        | 4.0e-28     | 3.07 (2.51-3.75)                     |
| GABBR3       | 1.9e-24     | 2.33 (1.96-2.76)                     |
| DNAH8        | 3.5e-24     | 3.70 (2.81-4.88)                     |
| CA6          | 3.5e-23     | 6.6e-4 (4.8e-6-8.9e-52)              |
| F11          | 3.6e-22     | 22.50 (12.16-41.62)                  |
| CHRM1        | 1.1e-20     | 2.33 (1.96-2.76)                     |
| DDX25        | 7.6e-19     | 8.66 (5.37-13.96)                    |
| NOS2P2       | 2.0e-17     | 1.8e-14 (1.2e-17-2.6e-11)            |
| TTC29        | 2.0e-16     | 3.45 (2.57-4.63)                     |
| HYDIN        | 7.9e-15     | 51.34 (19.01-138.64)                 |
| GAABRA5      | 1.8e-14     | 18.52 (8.78-39.05)                   |
| GIP          | 1.0e-13     | 0.46 (0.38-0.57)                     |
| C10orf90     | 3.9e-13     | 0.04 (0.02-0.10)                     |
| TNFRSF11B    | 5.7e-12     | 0.41 (0.32-0.53)                     |
| ANGPTL3      | 3.7e-11     | 5.47 (3.31-9.06)                     |
| CPN1         | 3.8e-11     | 0.06 (0.03-0.14)                     |
| KCNQ9        | 4.5e-9      | 0.43 (0.32-0.57)                     |
| NEUROD1      | 5.3e-8      | 0.20 (0.11-0.35)                     |
| UGT3A1       | 4.5e-8      | 0.32 (0.21-0.48)                     |
| TRIM22       | 4.8e-8      | 0.27 (0.17-0.43)                     |
| PTFRQ        | 7.1e-5      | 0.34 (0.20-0.57)                     |
| UPB1         | 8.7e-5      | 0.10 (0.03-0.32)                     |
| CLCA1        | 9.9e-4      | 4.53 (3.84-11.13)                    |
| SLC7A3       | 3.1e-3      | 0.28 (0.12-0.65)                     |

TCGA-LUAD Train data set

(a)

Figure 2: Continued.
Figure 2: Univariate Cox analysis based on TCGA-LUAD internal training set and validation set and screening of differentially oxidative stress genes associated with LUAD prognosis. (a) Forest plot displays the findings of univariate Cox analysis of differential oxidative stress genes associated with LUAD prognosis in the training set; (b) forest plot demonstrates the findings of univariate Cox analysis of differential oxidative stress genes associated with LUAD prognosis in the training set; (c) Venn shows 35 prognosis-related differential oxidative stress genes at the intersection of the two sets.
Figure 3: Continued.
Furthermore, calibration curves evaluated the fitness between actual survival statuses with visualized survival status of the developed nomogram via bootstrap methods (1,000 replicates). The values of prognosis evaluation between risk signature, stage, and the nomogram were compared via ROC curves at 1 year, 3 years, and 5 years, correspondingly.

2.7. Computation of Immune Infiltration Score Was Premised on the Gene Groups Previously Determined. We downloaded the unified as well as the standardized pan-cancer dataset from the UCSC (https://xenabrowser.net/) database: TCGA TARGET GTEx (PANCAN, N = 19131, G = 60499). Additionally, we extracted gene expression data in every sample and further screened the sample sources: Primary Blood-Derived Cancer-Peripheral Blood (TCGA-LAML), Primary Tumor, Metastatic TCGA-SKCM, Primary Blood-Derived Cancer-Bone Marrow, Primary Solid Tumor, and Recurrent Blood-Derived Cancer-Bone Marrow samples, further log2 (x + 0.001) transformation was carried out on every expression value; in addition, we also extracted the gene expression profile of every tumor, respectively, and mapped the expression profile. On the GeneSymbol, the R software package ESTIMATE (version 1.0.13, https://bioinformatics.mdanderson.org/public-software/estimate/, doi:10.1038/ncomms3612) [28] was further used to calculate the gene expression in each tumor, stromal, immune, and ESTIMATE scores for each patient.

2.8. GO, KEGG, and Immune Infiltration Enrichment Analyses for Risk-Related DEGs. In accordance with the risk grouping, normalized gene expression matrixes of the derivation set as well as the validation set generated above were employed with the “limma” R package to detect risk-related DEGs with the cut-off criteria of |logFC| ≥ 1 and adj. p value < 0.05, correspondingly. Risk-related DEGs were analyzed with both GO and KEGG utilizing the “clusterProfiler” R package. Next, single-sample GSEA (ssGSEA) for immune infiltration was adopted with the “GSVA” R package to ascertain the infiltration score of immune cells and the immune-related roles.

2.9. Calculation of Immune Infiltration Score. CIBERSORT played a crucial function in calculating an absolute immune infiltrate score for the primary tumor samples. The default CIBERSORT parameters were instrumental in generating the curated CIBERSORT signature matrix. It roughly yielded the expected relative abundances. We performed a pan-cancer analysis of the CIBERSORT score, including the correlation of immune score in pan-cancer, analyzed the correlation of MAP3K19 and NTSR1 with the distribution of various immune cells for LUAD patients, and visualized them with heatmaps and scatter plots. The tumor purity,
as well as stromal scores, were determined utilizing the Esti-
mation of Stromal and Immune Scores (ESTIMATE) based on
RNA-seq data and global proteomic data.

2.10. Quantitative Reverse Transcription-Polymerase Chain
Reaction (qRT-PCR). Total RNA extracted from lung adeno-
carcinoma tissue and para-cancerous tissue with Trizol
Reagent (Invitrogen, USA) were reverse transcribed with
HiScript III 1st Strand cDNA Synthesis Kit (Vazyme,
China). Next, HiScript II One Step RT-PCR Kit (Vazyme,
China) and qRT-PCR analysis were used for detecting
cDNA expression levels, and GAPDH was used as internal
reference. Primers were shown as follows: GAPDH, forward
(F): 5′-AATGGGCAGCCGTTAGGAAA-3′, reverse (R):
GCCCAATACGACCAAATCAGAG-3′; NTR1, forward
(F): 5′-TCATCGCCTTTGTGGTCTGCT-3′, reverse (R):
TGGTTGCTGGACACGCTGTCG-3′; MAP3K19, forward
(F): 5′-AGGAGTTCGACCAAGATGGTG-3′, reverse (R):
GGTCGAAAACTCTTCTGTCCTG-3′.

2.11. Immunohistochemistry (IHC) Staining. After depara-
ffinization and dehydrating the tissue sections, they were sub-
jected to epitope retrieval, treated with H2O2, and blocked
against nonspecific bindings. The tissues were then incu-
bated overnight with anti-NTSR1 antibodies (1:100,
Abcam, ab217134) and anti-MAP3K19 antibodies (1:100,
Invitrogen, PA5-29285) at 4°C. Subsequently, the tissue sec-
tions were incubated with secondary antibodies (1:1000,
Proteintech, SA00001-2) for two hours at ambient tempera-
ture. The signal was detected with an enhanced DAB stain-
ing kit (Proteintech, China).

2.12. Statistical Analysis. The Spearman correlation test was
performed to investigate the link between two variables that
were nonlinearly linked. The Student’s t-test, on the other
hand, was utilized in comparing the normally distributed
data, whereas the chi-square test was carried out to contrast
pairwise and categorical features in various subgroup. Uni-
variate as well as multivariate Cox regression analyses
assessed the influences of the immune signature and numer-
ous clinic-pathological parameters on the survival of
patients. The package pheatmap was vital in plotting heat-
maps. A two-sided p value of <0.05 was statistically signi-
ficant. Kruskal-Wallis test was employed for one independent var-
iable with two or more levels and an ordinal dependent var-
able. K-M analysis measured the proportion of individuals
living for a particular period, whereas the log-rank test eval-
uated the significance of differences. In these investigations,
statistical analysis was performed by R software (4.0.0). A
two-tailed p value of <0.05 was statistically significant unless
otherwise stated.

3. Result

3.1. The Characteristics of Patients. RNA-sequencing profiles of
a total of 445 LUAD sample as well as clinic-pathological
data were obtained from the UCSC Xena TCGA-LUAD
dataset. Then, randomization was conducted to divide them

Table 2: List of GO and KEGG enrichment analysis results of differential oxidative stress hub genes with close interaction.

| Ontology | ID         | Description                                                                 | p value     | p.adjust | Q value |
|----------|------------|------------------------------------------------------------------------------|-------------|----------|---------|
| BP       | GO:0008217 | Regulation of blood pressure                                                 | 4.92e-05    | 0.008    | 0.004   |
|          | GO:0090276 | Regulation of peptide hormone secretion                                     | 7.32e-05    | 0.008    | 0.004   |
|          | GO:0030072 | Peptide hormone secretion                                                    | 1.26e-04    | 0.008    | 0.004   |
|          | GO:0007187 | G protein-coupled receptor signaling pathway, coupled to cyclic nucleotide second messenger | 1.36e-04    | 0.008    | 0.004   |
|          | GO:0046883 | Regulation of hormone secretion                                              | 1.52e-04    | 0.008    | 0.004   |
|          | GO:0030658 | Transport vesicle membrane                                                   | 0.003       | 0.039    | 0.024   |
|          | GO:0005788 | Endoplasmic reticulum lumen                                                  | 0.006       | 0.039    | 0.024   |
| CC       | GO:0031045 | Dense core granule                                                           | 0.008       | 0.039    | 0.024   |
|          | GO:0042629 | Mast cell granule                                                           | 0.009       | 0.039    | 0.024   |
|          | GO:0030133 | Transport vesicle                                                            | 0.010       | 0.039    | 0.024   |
|          | GO:0048018 | Receptor ligand activity                                                     | 6.48e-04    | 0.014    | 0.008   |
|          | GO:005179  | Hormone activity                                                             | 9.68e-04    | 0.014    | 0.008   |
| MF       | GO:0005184 | Neuropeptide hormone activity                                                | 0.011       | 0.067    | 0.036   |
|          | GO:0071855 | Neuropeptide receptor binding                                                | 0.013       | 0.067    | 0.036   |
|          | GO:0042056 | Chemoattractant activity                                                     | 0.015       | 0.067    | 0.036   |
|          | hsa04923   | Regulation of lipolysis in adipocytes                                         | 4.83e-04    | 0.005    | 0.001   |
|          | hsa04080   | Neuroactive ligand-receptor interaction                                       | 7.00e-04    | 0.005    | 0.001   |
| KEGG     | hsa04024   | cAMP signaling pathway                                                       | 0.007       | 0.032    | 0.009   |
|          | hsa04950   | Maturity onset diabetes of the young                                         | 0.016       | 0.056    | 0.017   |
|          | hsa04913   | Ovarian steroidogenesis                                                      | 0.031       | 0.064    | 0.019   |
Figure 4: Continued.
into two groups, namely, the training set (331 patients) and the test set (166 patients) (Table 1).

3.2. Prognosis-Related Differential Oxidative Stress Gene Signatures Are Independent Prognostic Factors for LUAD Patients. To determine the prognostic value of differentially oxidative stress genes, we did a univariate Cox regression analysis. There were 100 and 56 xenooxidative stress genes significantly linked to the overall survival (OS) in the training and validation sets, respectively (Figures 2(a) and 2(b)). A total of 35 differentially oxidative stress genes associated with prognosis were obtained by intersecting them (Figure 2(c)).

3.3. Construction of PPI Network of DEGs Related to Oxidative Stress. We first performed a differential analysis between 59 normal samples and 535 LUAD samples in the TCGA-LUAD dataset using the limma algorithm and obtained a total of 1607 differentially expressed genes (|logFC| > 2, adj. p value < 0.05). Based on this differential result, 9469 previously identified oxidative stress genes were intersected, and 147 differentially expressed oxidative stress genes were obtained (Figure 3(a)). The STRING database constructed a protein-protein interaction network (PPI) to reflect the intermolecular interactions, and the largest confidence interaction score was established at 0.4, which was analyzed and visualized by the Network Analyzer tool of Cytoscape (v3.7.2) (Figure 3(b)). PPI network modules were screened for closely linked genes using the MCODE plugin and visualized (Figure 3(c)). At the same time, the Cyto-Hubba plugin was used to screen the Top20 closely linked genes (Figure 3(d)). The intersection of the two approaches was shown by the Venn diagram, obtaining 18 closely related differentially oxidative stress-related genes (Figure 3(e)), including HIST2H2AB, HIST1H2BC, SCG3, HIST1H2A1, CHGA, HIST1H3J, HIST1H1B, HIST1H2BM, HIST1H1A, SCG2, HIST1H4F, HIST1H2AE, GHRHR, KLK3, NEUROD1, NPY, HIST1H2BB, HIST1H2BH, and CGA.

3.4. Functional Enrichment Analysis. We analyzed these 18 closely related differentially oxidative stress-related genes. Gene Ontology (GO) enrichment analysis employing the above genes suggests that these hub genes exist in the membrane of transport vesicles, endoplasmic reticulum lumen, mast cell granules, and transport vesicles, which can affect blood pressure, hormone secretion, and G protein. Coupling cyclic, coupling receptor signaling pathway, nucleotide second messenger, and other functions play a regulatory role and have certain applications for receptor-ligand activity, neuropeptide hormone activity, neuropeptide receptor binding, and chemotactic activity value. The enrichment of BP set was mostly concentrated in oxidative phosphorylation, mitochondrial translation elongation, mitochondrial...
Figure 5: Continued.
translation termination, translation termination, and purine ribonucleoside triphosphate metabolism. CC enrichment was mostly concentrated in the inner mitochondrial membrane, mitochondrial protein complexes, mitochondrial matrix, organelle ribosomes, and mitochondrial ribosomes. The enrichment of MF was mostly concentrated in the structural components of ribosomes, proton transmembrane, transporter activity, electron transfer activity, cytochrome-c oxidase activity, and heme-copper terminal oxidase activity. This suggests that the Hub gene may be related to the transmembrane transport of cells in terms of molecular structure.

Combined with all the above enrichment results, we speculate that key genes related to oxidative stress may play a role in the interaction process and related mechanisms between cells. The KEGG pathway enrichment results indicated that differential oxidative stress hub genes were enriched in modulation of lipolysis in adipocytes, neuroactive ligand-receptor interaction, CAMP signaling pathway, maturity onset diabetes of the young, ovarian steroidogenesis, and other pathways. This result suggests that the differentially expressed oxidative stress-related genes can affect the occurrence and progression of LUAD via the above potential biological functions and molecular pathways (Figures 3(f) and 3(g) and Table 2).

3.5. Establishment and Validation of Prognostic Gene Signature Associated with Oxidative Stress in TCGA-LUAD.

To avert overfitting, we additionally performed a LASSO-Cox analysis. In order to avoid the influence of confounding factors, we first performed LASSO-Cox analysis on 147 differentially oxidative stress genes in the training set and established 8 gene signatures, including C10orf90, CIDEC, MUC2, FGF5, KRT6A, DLEC1, MAP3K19, and NOS2P2. Patients were categorized into high- and low-risk groups as per the median risk score with each group containing 167 participants. To tune parameter selection through the LASSO model, we utilized cross-validation (Figure 4(a)). The coefficient profiles of LASSO for the eight prognostic differentially oxidative stress-related genes were shown (Figure 4(b)). At the same time, we show eight prognostic differential oxidative stress gene signature risk analysis graphs (Figure 4(d)). The upper graph shows the distribution as well as the median of risk scores in the training set and the middle graph shows the distribution of individuals in various risk groups, and the lower graph shows heatmap of differential expression of 3 genes in the two risk groups; (e) the prognostic performance of risk scores in the TCGA-LUAD training set over 1, 3, and 5 years is validated utilizing AUC of time-dependent ROC curves; (f) based on the training set, multivariate Cox proportional hazards regression forest plot for a 3-gene prognostic model.
Figure 6: Continued.
Furthermore, LASSO-Cox analysis was conducted on the above 35 prognosis-related differential oxidative stress genes in the training set and validation set, correspondingly. A 3-gene prognostic model was predicted among the genes in the training set, including NTSR1, MAP3K19, and NOS2P2 as predictors of patient prognosis (Figures 5(a) and 5(b)). Not only K-M survival but also time-dependent ROC analysis findings were significant. The high-risk group had a greater mortality risk in contrast with the low-risk group ($p = 0.02$, HR = $1.60 (1.09, 2.35)$) (Figure 5(c)). We show 3 prognostic differential oxidative stress gene signature risk analysis graphs, the upper graph shows the distribution and median of risk scores in the TCGA-LUAD training set, and the heatmap reflects the expression differences of model genes (Figure 5(d)). The AUC reached 0.56 (0.66-0.46) at 1 year, 0.58 (0.68-0.48) at 3 years, and 0.65 (0.80-0.49) at 5 years (Figure 5(e)). In the multivariate Cox regression analysis, NTSR1, MAP3K19, and NOS2P2 were incorporated. Multivariate Cox proportional hazards regression based on the training set suggested that only NTSR1 and MAP3K19 were jointly used as predictors of poor prognosis at high risk of LUAD, with significant statistical significance (Figure 5(f)).

An 8-gene prognostic model was predicted among the genes in the validation set, including SLC6A5, SNORA71D, PIK3C2G, KRT6B, IGF2BP3, NTSR1, KLK6, and MAP3K19 as predictors of patient prognosis (Figures 6(a) and 6(b)). Both K-M survival and time-dependent ROC analysis findings were significant. The high-risk group had a greater mortality risk in comparison with the low-risk group ($p = 1.2e-6$, HR = $4.12 (2.22, 7.63)$) (Figure 6(c)). We show risk analysis plots for eight prognostic differential oxidative stress gene signatures, the upper panel shows the distribution and median of risk scores in the TCGA-LUAD training set, the middle panel shows the distribution of individuals in each risk group, and the lower panel shows heatmap of differential expression of 8 genes in both risk groups; (e) AUC of time-dependent ROC curves validates the prognostic performance of risk scores in the TCGA-LUAD training set at 1, 3, and 5 years.

In the prediction model constructed based on LASSO-Cox analysis of 147 differentially oxidative stress-related genes and 35 prognostic differentially oxidative stress-related genes in the validation set, both MAP3K19 and NOS2P2 were found to be prognostic predictors, although NOS2P2 in the further multivariate Cox regression analysis results were suggested to have no statistically significant effect on prognosis. More interestingly, LASSO-Cox analysis based on 35 prognostic differentially oxidative stress-related genes found that MAP3K19 and NTSR1 could serve as prognostic predictors in both training and validation sets. Based on the above analysis, we believe that MAP3K19 and NTSR1 may have potential application value in predicting the prognosis of LUAD, and more abundant analysis is urgently needed to evaluate its application value and important clinical significance in LUAD.
Survival probability

\[ \text{HR} = 3.53, 95\% \text{CI} (2.40, 5.19) \]
\[ p = 1.1 \times 10^{-11} \]

Risk score

- L
- H

- Sensitivities
- 1-Specificities

Time:AUC(95\%CI)

- 365:0.75(0.84-0.66)
- 1095:0.73(0.81-0.65)
- 1825:0.76(0.86-0.66)

Figure 7: Continued.
Points

Age ($p = 0.27$)

Gender ($p = 0.43$)

Stage ($p = 0.10$)

T ($p = 0.14$)

M ($p = 0.25$)

N ($p < 0.01$)

NTSR1 ($p = 0.04$)

MAP3K19 ($p < 0.001$)

Total points

Linear Predictor

Probability of 1 year

Probability of 3 year

Probability of 5 year

Figure 7: Continued.
3.6. MAP3K19 and NTSR1 Prognostic Independence of Clinical Characteristics. To further ascertain the prognostic value and predictive performance of the MAP3K19 and NTSR1 gene signatures, we first conducted a Cox regression analysis based on 445 patient samples with complete clinical information in TCGA-LUAD. We included clinical factors including age at diagnosis, gender, tumor pathological grade, lymph node status, distant metastasis, TNM stage, and gene signatures MAP3K19 and NTSR1 (high- and low-expression groups divided by median value) in the training set-related prognostic variables to predict LUAD overall survival; (d) calibration plot for internal validation of nomograms for overall survival prognostic status for LUAD at 1, 3, and 5 years of training set; (e) nomograms effectively integrate and demonstrate the ability of prognostic variables in TCGA-LUAD to predict overall survival in LUAD; (f) calibration plot for internal validation of nomograms for overall survival prognostic status for LUAD at 1, 3, and 5 years of the TCGA-LUAD dataset.

The above results suggest that two gene marker prognostic factors—MAP3K19 and NTSR1—have important clinical significance for the prognosis of LUAD. Thus, we performed independent analyses for MAP3K19 and NTSR1 in TCGA-LUAD. Unpaired differential analysis between normal samples and LUAD samples indicated that NTSR1 was significantly overexpressed in LUAD (Figure 8(a)), while MAP3K19 was significantly underexpressed in LUAD (Figure 8(d)). The results of differential expression analysis of paired samples were consistent (Figures 8(b) and 8(e)). Subsequent ROC curve analysis suggested that NTSR1 (AUC: 0.601 (0.532–0.671), Figure 8(c)) and MAP3K19 (AUC: 0.710 (0.633–0.787), Figure 8(f)) had a good diagnostic performance for the differential diagnosis of LUAD samples.
Figure 8: Differential expression analysis and diagnostic efficacy verification of oxidative stress-related 2-gene marker prognostic factors MAP3K19 and NTSR1 in TCGA-LUAD. (a) Unpaired differential analysis of NTSR1 in LUAD between normal and LUAD samples indicated significantly high expression; (b) paired differential analysis of NTSR1 in LUAD and normal samples suggested significantly high expression; (c) ROC curve validated MAP3K19. It has good diagnostic performance for the differential diagnosis of LUAD samples; (d) between normal and LUAD samples, the unpaired differential analysis of MAP3K19 in LUAD suggests significantly lower expression in LUAD and significantly lower expression in LUAD; (e) MAP3K19 in LUAD paired difference analysis with normal samples indicated significantly low expression; (f) ROC curve verified that MAP3K19 has good diagnostic performance for the differential diagnosis of LUAD.
Survival probability

TCGA-LUAD

Overall survival
HR = 1.67 (1.25–2.23)

Pathologic stage: Stage II & Stage III & Stage IV

Overall survival
HR = 1.61 (1.11–2.33)

NTSR1

Low
High

Primary therapy outcome: PR & CR

Overall survival
HR = 1.80 (1.17–2.75)

NTSR1

Low
High

N stage: N1 & N2 & N3

Overall survival
HR = 1.63 (1.08–2.46)

NTSR1

Low
High

T stage: T2 & T3 & T4

Overall survival
HR = 1.47 (1.06–2.04)

NTSR1

Low
High

M stage: M1

Overall survival
HR = 3.41 (1.06–10.95)

NTSR1

Low
High

Figure 9: Continued.
3.7. Differential Analysis and Enrichment Analysis Based on Oxidative Stress-Related 2 Gene Signatures. All of the above analyses confirmed the superior predictive performance of MAP3K19 and NTSR1 for poor prognosis in LUAD, and the following analysis focused on exploring how the oxidative stress-related 2 gene signature might lead to poor prognosis in LUAD.

In order to additionally explore the molecular mechanism engaged in the identification of high-risk populations with poor prognosis of LUAD in the oxidative stress prognostic model composed of MAP3K19 and NTSR1, we separately analyzed the significant differences between two risk groups in the prognostic model constructed by the LASSO algorithm in the training set and the validation set. We performed differential analysis on the two risk groups defined by the median risk of the training set (Figure 11(a)) and validation set (Figure 11(b)) prediction model, respectively, showing the expression of significantly different genes as a heatmap. The 92 prognostic risk genes of oxidative stress coexisting in both sets were obtained by the Venn diagram (Figure 11(c)). These genes were then subjected to enrichment analysis. GSVA analysis shows enriched entries with $|\logFC| > 0.5$ and adj-$p < 0.05$ as a heatmap. We then analyzed the reference gene set retrieved from the MSigDB database in “c2.cp.v7.2.symbols.gmt,” “h.all.v7.2.symbols.gmt” gene set, and immune cell infiltration-related gene set. Enrichment scores for each gene functional pathway were obtained utilizing the GSVA package in R. The heatmap visualizes the differential gene expression between the two risk groups in the training set as well as the validation set. It was found that the gene set enrichment results of “h.all.v7.2.symbols.gmt” (Figure 11(d)) and “c2.cp.v7.2.symbols.gmt” (Figure 11(e)) based on the MSigDB database were concentrated in UNFOLDED PROTEIN RESPONSE, GLYCOLYSIS, MTORC1 SIGNALING, MYC TARGETS V1, MYC TARGETS V2, and other functions and ways. Immune-related “immune.gmt” (Figure 11(f)) gene set enrichment results focused on Eosinophil, Natural.killer.cell, Immature..B.cell, Activated.B.cell, Mast.cell, and Type.1.-T.helper .cell.

Next, the GO (Table 3) and KEGG (Table 4) enrichment analysis results are shown by bubble plots (Figures 12(a) and 12(b)) and bar graphs (Figures 12(c) and 12(d)). BP was significantly enriched in oxidative phosphorylation, mitochondrial translational elongation, mitochondrial translational termination, translational termination, purine ribonucleoside triphosphate metabolic process, and other
**Figure 10**: qRT-PCR and IHC analysis of model genes. (a–c) NTSR1 mRNA and protein expression were significantly increased in LUAD samples; (d–f) MAP3K19 expression in LUAD samples was significantly lower than that in paratumor normal tissue.
Figure 11: Continued.
Figure 11: Continued.
Figure 11: Continued.
items. CC was significantly enriched in mitochondrial inner membrane, protein complex, matrix, ribosome, organelar ribosome, and other entries. MF was significantly enriched in the structural components of the ribosome, proton transmembrane transporter activity, cytochrome-c oxidase activity, electron transfer activity, heme-copper terminal oxidase activity, and other items. KEGG was significantly enriched in oxidative phosphorylation, thermogenesis, ribosome, Parkinson’s disease, cardiac muscle contraction, and other pathways. A pathview diagram with differentially expressed genes colored for the arachidonic acid metabolism pathway is displayed in (Figures 12(e)–12(i)), including hsa00190, hsa03010, hsa04260, hsa04714, and hsa05012.

### Table 3: GO enrichment analysis of coexpressed genes based on the LUAD oxidative stress-related prognostic risk model.

| Ontology | ID      | Description                               | BgRatio | p value   | p.adjust | Q value |
|----------|---------|-------------------------------------------|---------|-----------|-----------|---------|
| BP       | GO:0006119 | Oxidative phosphorylation                  | 145/18670 | 8.76e-10  | 1.28e-06  | 1.17e-06 |
|          | GO:0070125 | Mitochondrial translational elongation     | 88/18670 | 5.21e-09  | 2.78e-06  | 2.53e-06 |
|          | GO:0006415 | Translational termination                  | 104/18670 | 1.97e-08  | 7.19e-06  | 6.56e-06 |
|          | GO:0009205 | Purine ribonucleoside triphosphate metabolic process | 335/18670 | 2.83e-08  | 7.42e-06  | 6.76e-06 |
|          | GO:0005743 | Mitochondrial inner membrane               | 473/19717 | 3.79e-18  | 7.50e-16  | 6.10e-16 |
|          | GO:0098798 | Mitochondrial protein complex              | 262/19717 | 1.10e-10  | 1.09e-08  | 8.85e-09 |
| CC       | GO:0005759 | Mitochondrial matrix                       | 469/19717 | 1.45e-08  | 9.60e-07  | 7.81e-07 |
|          | GO:0000313 | Organellar ribosome                        | 87/19717  | 1.03e-07  | 4.09e-06  | 3.33e-06 |
|          | GO:0005761 | Mitochondrial ribosome                     | 87/19717  | 1.03e-07  | 4.09e-06  | 3.33e-06 |
|          | GO:0003735 | Structural constituent of ribosome         | 202/17697 | 3.60e-06  | 9.07e-04  | 7.84e-04 |
|          | GO:0015078 | Proton transmembrane transporter activity   | 133/17697 | 3.07e-05  | 0.004     | 0.003   |
| MF       | GO:0009055 | Electron transfer activity                 | 114/17697 | 1.66e-04  | 0.011     | 0.009   |
|          | GO:0004129 | Cytochrome-c oxidase activity              | 28/17697  | 2.69e-04  | 0.011     | 0.009   |
|          | GO:0015002 | Heme-copper terminal oxidase activity      | 28/17697  | 2.69e-04  | 0.011     | 0.009   |
Table 4: KEGG enrichment analysis of coexpressed genes based on the LUAD oxidative stress-related prognostic risk model.

| Ontology | ID     | Description                       | BgRatio | p value | p.adjust | Q value |
|----------|--------|-----------------------------------|---------|---------|----------|---------|
| KEGG     | hsa00190 | Oxidative phosphorylation         | 133/8076| 4.47e-08| 3.35e-06| 2.82e-06|
|          | hsa04714 | Thermogenesis                     | 231/8076| 4.98e-07| 1.87e-05| 1.57e-05|
|          | hsa03010 | Ribosome                          | 158/8076| 2.71e-05| 6.77e-04| 5.70e-04|
|          | hsa05012 | Parkinson disease                 | 249/8076| 6.90e-05| 0.001    | 0.001   |
|          | hsa04260 | Cardiac muscle contraction        | 87/8076 | 1.25e-04| 0.002    | 0.001   |

3.8. Gene Set Enrichment Analysis of Oxidative Stress Gene Model in LUAD. In order to additionally explore the molecular processes engaged in the identification of high-risk populations with poor prognosis of LUAD in the oxidative stress prognostic model composed of MAP3K19 and NTSR1, GSEA is a computational method utilized to ascertain if a predefined set of genes shows statistical differences between two biological states, which is typically utilized to estimate expression in a dataset sample. Changes in the pathway as well as biological process activity were analyzed [23]. To investigate the variations in biological mechanisms between the two groups of samples, employing the gene expression profile dataset, we obtained the reference gene set “c2.cp.kegg.v7.4.entrez.gmt” from the MSigDB database, utilizing the R package “clusterProfiler.” The GSEA method is included in enrichment analysis and visualization of datasets. p value < 0.05 was considered statistically significant (Figure 13 and Table 5).

3.9. Pan-Cancer Analysis of Immune Cells and Immune Infiltration Based on MAP3K19 and NTSR1. The infiltration abundance of immune cells was analyzed by CIBERSORT, and the correlation heatmap in pan-cancer showed that MAP3K19 (Figure 14(a)) and NTSR1 (Figure 14(b)) were significantly correlated with immune infiltration in more types of tumors. Based on TCGA LUAD transcription profile and CIBERSORT, we derived the proportions of 22 tumor-infiltrating immune cells. In our study, the use of RNA-seq data as well as global proteomic data inferred MAP3K19 and NTSR1 for pan-cancer tumor purity, immune score, and stromal score. From this, significantly correlated immune infiltration scores were identified, and we analyzed the scores for stromal cell levels, tumor purity, and immune cell infiltration levels in cancer tissues calculated with ESTIMATE expression which was substantially positively linked to immune score, ESTIMATE score, and stromal score, respectively. The results suggest that NTSR1 (Figures 14(c)–14(e)) and MAP3K19 (Figures 14(f)–14(h)) were closely related to stromal cell level, tumor purity, and immune cell infiltration level in the lung tumor microenvironment and may affect the prognosis of LUAD by changing the tumor immune microenvironment.

4. Discussion

Globally, lung cancer is the major cause of cancer-related mortalities, whereas LUAD is the most prevalent histological subtype of the disease. About two-thirds of LUAD have activated oncogenes. Most oncogene mutations often activate downstream signaling pathways through oxidative stress pathways and states, ultimately leading to lethal malignancies, including LUAD [24, 29, 30]. Molecularly targeted therapy significantly improves survival in patients with treatment-targeted lesions compared to conventional chemotherapy [31, 32]. However, the clinical efficacy of targeted drugs has been improved due to the lack of appropriate small molecules to bind major tumor-causing gene mutations.

For the survival of both normal cells and cancerous cells, redox hemostasis is fundamental. Nonetheless, numerous malignancies have increased levels of reactive oxygen species (ROS) and exhibit signs of chronic oxidative stress as a result of oncogenic injury, metabolic malformations, hypoxia, and proteotoxic stress [33]. The increased ROS at the sublethal level is implicated in the enhancement of tumor development via triggering mutations and altering cell signaling [34]. Because traditional cytotoxic agents additionally influence normal tissues, targeted approaches that induce catastrophic oxidative stress selectively in malignant cells would avail a better therapeutic window [35].

The LUAD cohort of TCGA availed both the expression and clinical data in this study. Among oxidative stress-related genes, we did differential expression analysis as well as univariate Cox analysis to screen 32 prognostic DEGs.
Figure 12: Continued.
Figure 12: Continued.
from 147 differentially expressed oxidative stress-related genes and utilized Lasso-penalized Cox regression analysis constructed 2 gene markers associated with prognosis. At the same time, by constructing a PPI network, we analyzed the distribution of 18 differentially oxidative stress hub genes in transport vesicles, endoplasmic reticulum lumen, etc., affecting blood pressure, hormone secretion, nucleotide second messengers, and other functions to regulate oxidative stress effect. To construct and validate the oxidative stress gene signature affecting the diagnosis and prognosis of LUAD, we randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into the training set (2/3 of the total, n = 296) and the test set (1/3 of the total, n = 149); the grouping is normally distributed. We first performed a univariate Cox regression analysis of gene expression premised on the expression data of 147 oxidative stress-related genes in the training and test set, correspondingly. Thirty-five genes with prognostic values coexisting in both sets were identified. Then, we performed LASSO-Cox proportional hazards regression based on 35 prognostic-related differential oxidative stress genes in the training and validation set, correspondingly, and found that MAP3K19 and NTSR1 in both sets showed a better prediction of poor prognosis in LUAD ability. Therefore, the prognostic independence of 2-gene signature prognostic factors, MAP3K19 and NTSR1, was further analyzed in combination with clinical features. A nomogram was constructed to effectively integrate prognostic variables and validated with calibration. Next, the difference analysis of MAP3K19 and NTSR1 in TCGA-LUAD and the correlation analysis of clinical variables indicated that both the high expression of NTSR1 and the low expression of MAP3K19 had a better diagnostic performance for the diagnosis of LUAD and the identification of poor prognosis. At the same time, MAP3K19 and NTSR1 were found to be significantly correlated with clinical variables such as LUAD pathological stage and TNM grade. Prognosis-related differential oxidative stress gene signatures are independent prognostic factors in patients with LUAD.

The importance of tumor immune activity on tumorigenesis and development, as well as individual variation at the gene level, has attracted more and more researchers to focus on the significance of differential genes that may be useful in distinguishing pRCC patients with heterogeneous responses and predicting prognosis potentially meaningful. The ESTIMATE analysis of pan-cancer species analysis of MAP3K19 and NTSR1 in TCGA database systematically recorded the abundance of 22 tumor-infiltrating immune compartments in LUAD samples through the CIBERSORT algorithm and integrated it with the MAP3K19 and NTSR1
Figure 13: Continued.
Figure 13: Continued.
Figure 13: Gene set enrichment analysis of oxidative stress prognostic gene signature in LUAD. (a) JAK_STAT_SIGNALING_PATHWAY (ES = 0.5561, NP = 0.0000); (b) SYSTEMIC_LUPUSERYTHEMATOSUS (ES = 0.6502, NP = 0.0040); (c) LEISHMANIA_INFECTION (ES = 0.5961, NP = 0.0041); (d) NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY (NP = 0.0064); (e) AUTOIMMUNE_THYROID_DISEASE (ES = 0.7506, NP = 0.0021); (f) VIRAL_MYOCARDITIS (ES = 0.6040, NP = 0.0000); (g) CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION (ES = 0.4820, NP = 0.0142); (h) T_CELL_RECEPTOR_SIGNALING_PATHWAY (ES = 0.5250, NP = 0.0000); (i) ALLOGRAFT_REJECTION (ES = 0.7815, NP = 0.0021).
molecular profiles to analyze the degree of immune infiltration in LUAD. Subsequently, further immune score, stromal score, and ESTIMATE score in the pan-cancer TME revealed the potential roles of MAP3K19 and NTSR1 in regulating stromal/immune scores and gene expression in tumors. This suggests that the oxidative stress prognostic model composed of MAP3K19 and NTSR1 may be involved in the molecular mechanism in the identification of high-risk populations with poor prognosis of LUAD.

Studies have shown that MAP3K19 level is elevated in COPD and bronchoalveolar lavage macrophages from patients with IPF. At the level of target gene transcription or protein synthesis, molecular studies have confirmed that MAP3K19 inhibitors are linked to pirfenidone or nintedanib. At the same time, MAP3K19 significantly attenuated bleomycin-induced pulmonary fibrosis [36] and is a central mediator of cigarette smoke-induced pulmonary inflammation and lower airway destruction [37]. In studies on lung cancer, targeting MAP3K19 has been reported to prevent human lung myofibroblast activation in vitro and in a humanized SCID model of idiopathic pulmonary fibrosis [38]. On the other hand, it can also phosphorylate MAP2K, thereby activating ERK as well as JNK kinases and increasing fibrosis viability [39]. The mechanism of action of NTSR1 in various tumors has also been reported many times [40]. Effects, underlying mechanisms, and clinical roles of NTSR1 on gastric adenocarcinoma cell proliferation and invasion. Interfering with NTSR1 expression exhibits anti-invasive effects through the Jun/miR-494/SOCS6 axis in glioblastoma cells [41]. The mechanism of action of NTSR1 in various tumors has also been reported many times. Effects, underlying mechanisms, and clinical roles of NTSR1 on gastric adenocarcinoma cell proliferation and invasion. Interfering with NTSR1 expression causes anti-invasive effects through the Jun/miR-494/SOCS6 axis in glioblastoma cells. NTSR1 and Wnt/β-catenin enhance tumor growth in glioblastoma [42]. Furthermore, NTSR1 methylation is linked to the lateral and noninvasive progression of colorectal tumors, whereas lowered levels of methylation might enhance the malignant potential via activation of NTSR1 [43].

In conclusion, this study constructed and validated a 2-gene oxidative stress-related LUAD prognostic model, MAP3K19 and NTSR1, which were significantly correlated with clinical variables (including LUAD pathological stage and TNM grade) and significantly affected the infiltration of immune cells in the tumor microenvironment (TME) of LUAD. And they are all involved in the process of oxidative stress and the energy metabolism network of ROS. The results of enrichment analysis based on the biological functions of GO, KEGG, and GSEA and pathway signaling patterns suggest that the molecular expression of MAP3K19 and NTSR1 and other immune cells help in the process of oxidative stress, even though conclusive evidence is still unavailable. Studies have shown that during apoptosis, immune cells are attracted and aggregated by a set of signals that enhance programmed cell death [44, 45]. In terms of bioinformatics, various studies have revealed a possible link between tumor and immune infiltration [46, 47]. Additionally, to palliative targeted therapy, monotherapy with new immunotherapies, such as immune checkpoint inhibitors (ICIs), has also demonstrated quite successful outcomes in some individuals with advanced LUAD [48]. In this study, through risk group-based immune annotation analysis, we found that macrophages, Tregs, and other types of immune cells and costimulation of immune-related roles were significantly enriched in both cohorts, suggesting that there may be potential regulatory mechanisms.

This study found some limitations. In a bioinformatics study, the weakness of the absence of experimental as well as clinical validation remains, and the utilization of alternative cutoff criteria, statistical methodologies, or analytical tools may provide varied results. Furthermore, building a prognostic model by focusing on a single marker may result in the deletion of numerous other potential prognostic genes. In conclusion, we created a novel 2-gene signature associated with oxidative stress that was shown to be an independent prognostic predictor of OS in LUAD. Through functional annotation analysis, the gene signature was associated with tumor immunity; nonetheless, its underlying process is not clear and needs to be explored further.

**Table 5: Results of gene set enrichment analysis of gene signatures in prognostic models of oxidative stress.**

| Term                                               | ES   | NES  | p value | FDR  |
|----------------------------------------------------|------|------|---------|------|
| VIRAL_MYOCARDITIS                                  | 0.604| 2.0023| 0.0117  | 0.04 |
| AUTOIMMUNE_THYROID_DISEASE                         | 0.7506| 2.0069| 0.0021  | 0.0149|
| ASTHMA                                             | 0.7919| 2.0269| 0.002   | 0.0168|
| NATURAL_KILLER_CELL MEDIATED_CYTOTOXICITY         | 0.5455| 1.9577| 0.0064  | 0.0191|
| JAK_STAT_SIGNALING_PATHWAY                         | 0.5561| 2.0488| 0.0259  | 0.023 |
| LEISHMANIA_INFECTION                               | 0.5961| 1.8795| 0.004   | 0.0266|
| ALLOGRAFT_REJECTION                                | 0.7815| 1.9236| 0.0021  | 0.0269|
| T_CELL_RECEPTOR_SIGNALING_PATHWAY                  | 0.525 | 1.8848| 0.0288  | 0.0149|
| SYSTEMIC_LUPUSERYTHEMATOSUS                       | 0.6502| 1.897 | 0.004   | 0.0293|
| CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION            | 0.482 | 1.8142| 0.0142  | 0.046 |
| TOLLLIKE_RECEPTOR_SIGNALING_PATHWAY                | 0.4956| 1.8005| 0.0151  | 0.0487|
Figure 14: Continued.
Figure 14: Continued.
The TCGA-LUAD cohort availed both the clinical data and the expression data in this study. Among oxidative stress-related genes, we did differential expression analysis and univariate Cox analysis in order to screen 32 prognostic DEGs from 147 differentially expressed oxidative stress-related genes and utilized Lasso-penalized Cox regression analysis constructed 2 gene markers associated with prognosis. At the same time, by constructing a PPI network, we analyzed the distribution of 18 differentially oxidative stress hub genes in transport vesicles, endoplasmic reticulum lumen, etc., affecting blood pressure, hormone secretion, nucleotide second messengers, and other functions to regulate oxidative stress effect. To construct and validate the oxidative stress gene signature affecting the diagnosis and prognosis of LUAD, we randomly divided 445 patients with complete clinical informatics data in TCGA-LUAD into the training set (2/3 of the total, \( n = 296 \)) and the test set (1/3 of the total, \( n = 149 \)); the grouping is normally distributed. We first performed a univariate Cox regression analysis of gene expression premised on the expression data of 147 oxidative stress-related genes in the training and test sets, respectively. Thirty-five genes with prognostic values coexisting in both sets were identified. Then, we performed LASSO-Cox proportional hazards regression based on 35 prognostic-related differential oxidative stress genes in the training and test sets, respectively. Thirty-five genes with prognostic values coexisting in both sets were identified. Then, we performed LASSO-Cox proportional hazards regression based on 35 prognostic-related differential oxidative stress genes in the training and test sets, respectively, and found that MAP3K19 and NTSR1 in the two sets showed a better prediction of poor prognosis in LUAD ability. Therefore, the prognostic independence of 2-gene signature prognostic factors, MAP3K19 and NTSR1, was further analyzed in F

![Figure 14: Immune cell infiltration analysis. (a, b) CIBERSORT analysis of MAP3K19 (a) and NTSR1 (b) heatmaps related to the infiltration abundance of immune cells in pan-cancer; (c–h) ESTIMATE analyzed the scores of tumor purity, stromal cell level, and immune cell infiltration level in tumor tissue and showed NTSR1 (c–e) and MAP3K19 (f–h) expression with immune score, ESTIMATE score, and stromal score, respectively, with a scatterplot positively correlated.](image-url)
**Figure 15: Continued.**
Figure 15: Continued.
Neutrophils

TCGA-LUAD (N = 500): $p = 0.03, r = 0.10$

Macrophages_M0

TCGA-LUAD (N = 500): $p = 1.2 \times 10^{-3}, r = 0.14$

**Figure 15: Continued.**
combination with clinical features. A nomogram was constructed to effectively integrate prognostic variables and validated with calibration. Next, the difference analysis of MAP3K19 and NTSR1 in TCGA-LUAD and the correlation analysis of clinical variables indicated that both the high expression of NTSR1 and the low expression of MAP3K19 had a better diagnostic performance for the diagnosis of LUAD and the identification of poor prognosis. At the same time, MAP3K19 and NTSR1 were found to be significantly correlated with clinical variables such as LUAD pathological stage and TNM grade. Prognosis-related differential oxidative stress gene signatures are independent prognostic factors in patients with LUAD. ESTIMATE analysis results of pan-cancer species analysis of MAP3K19 and NTSR1 in TCGA database. The abundances of 22 tumor-infiltrating immune compartments in LUAD samples were systematically recorded by the CIBERSORT algorithm and integrated with MAP3K19 and NTSR1 molecular profiles to analyze the degree of immune infiltration in LUAD. Subsequently, further immune score, stromal score, and ESTIMATE score in the pan-cancer TME revealed the potential roles of MAP3K19 and NTSR1 in regulating stromal/immune scores.

**Figure 15:** Correlation analysis of immune cell infiltration. (a, b) MAP3K19 was positively correlated with T cells CD4 memory resting and B cells memory; (c, d) MAP3K19 was negatively correlated with mast cells activated and dendritic cells activated; (e–g) NTSR1 was positively correlated with neutrophils, macrophages M0, and T cells gamma delta; (h) NTSR1 was negatively linked to mast cells resting.
and gene expression in tumors. This suggests that the oxidative stress prognostic model composed of MAP3K19 and NTSR1 may be involved in the molecular mechanism in the identification of high-risk populations with poor prognosis of LUAD.

Data Availability
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest
The authors declare that they have no conflicts of interest.

Authors’ Contributions
Xing Huang and Zhichai Lu conducted statistical analyses of the data and prepared the draft manuscript. Min He, Chen Chen, and Hanlin Ding edited the manuscript. Yipeng Feng and Yidan Sun provide critical comments to the manuscript. All authors checked and proofread the final version of the manuscript.

Acknowledgments
This work was funded by the National Natural Science Foundation of China (grant nos. 81902334) and the General project of Cancer Hospital Affiliated to Nanjing Medical University (no. ZM201921).

References
[1] J. W. Neal and G. W. Sledge, “Successes, toxicities and challenges in solid tumours,” Nature Reviews Clinical Oncology, vol. 11, pp. 627–628, 2014.
[2] R. Herbst, D. Morgensztern, and C. Boshoff, “The biology and management of non-small cell lung cancer,” Nature, vol. 553, pp. 446–454, 2018.
[3] T. S. Bowen, G. Schuler, and V. Adams, “Skeletal muscle wasting in cachexia and sarcopenia: molecular pathophysiology and impact of exercise training,” Journal of Cachexia, Sarcopenia and Muscle, vol. 6, no. 3, pp. 197–207, 2015.
[4] T. Cheng, M. Xu, H. Zhang et al., “KLHDC8A expression in association with macrophage infiltration and oxidative stress predicts unfavorable prognosis for glioma,” Oxidative Medicine and Cellular Longevity, vol. 2022, Article ID 2694377, 14 pages, 2022.
[5] W. Wang, S. Ren, Z. Wang, C. Zhang, and J. Huang, “Increased expression of TTC21A in lung adenocarcinoma infers favorable prognosis and high immune infiltrating level,” International Immunopharmacology, vol. 78, article 106077, 2020.
[6] D. R. Camidge, W. Pao, and L. V. Sequist, “Acquired resistance to TKIs in solid tumours: learning from lung cancer,” Nature Reviews Clinical Oncology, vol. 11, pp. 473–481, 2014.
[7] C. Chong and P. Jänne, “The quest to overcome resistance to EGFR-targeted therapies in cancer,” Nature Medicine, vol. 19, pp. 1389–1400, 2013.
[8] J. J. Lin, G. J. Riely, and A. T. Shaw, “Targeting ALK: precision medicine takes on drug resistance,” Cancer Discovery, vol. 7, no. 2, pp. 137–155, 2017.
[9] Q. Yao, X. Ge, Z. Lu, J. Shi, J. Shen, and J. Chen, “HAUS augmin-like complex subunit 1 influences tumour microenvironment and prognostic outcomes in glioma,” Journal of Oncology, vol. 2022, Article ID 8027686, 19 pages, 2022.
[10] J. R. Lieffers, O. F. Bathe, K. Fassbender, M. Winget, and V. E. Baracos, “Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery,” British Journal of Nutrition, vol. 107, no. 6, pp. 931–936, 2012.
[11] W. Wang, Z. Lu, M. Wang et al., “The cuproptosis-related signature associated with the tumor environment and prognosis of patients with glioma,” Frontiers in Immunology, vol. 13, article 998236, 2022.
[12] Y. Hamaguchi, T. Kaido, S. Okumura et al., “Proposal for new diagnostic criteria for low skeletal muscle mass based on computed tomography imaging in Asian adults,” Nutrition, vol. 32, no. 11-12, pp. 1200–1205, 2016.
[13] S. T. McSorley, D. H. Black, P. G. Horgan, and D. C. McMillan, “The relationship between tumour stage, systemic inflammation, body composition and survival in patients with colorectal cancer,” Clinical Nutrition, vol. 37, no. 4, pp. 1279–1285, 2018.
[14] H. W. Jung, J. W. Kim, J. Y. Kim et al., “Effect of muscle mass on toxicity and survival in patients with colon cancer undergoing adjuvant chemotherapy,” Supportive Care in Cancer, vol. 23, no. 3, pp. 687–694, 2015.
[15] J. M. Kyriakis and J. Axruch, “Mammalian MAPK Signal Transduction Pathways Activated by Stress and Inflammation: A 10-Year Update,” Physiological Reviews, vol. 92, no. 2, pp. 689–737, 2012.
[16] G. Malietzis, N. Johns, H. O. Al-Hassi et al., “Low musclearity and myosteatosis is related to the host systemic inflammatory response in patients undergoing surgery for colorectal cancer,” Annals of Surgery, vol. 263, no. 2, pp. 320–325, 2016.
[17] B. K. Pedersen and M. A. Febbraio, “Muscles, exercise and obesity: skeletal muscle as a secretory organ,” Nature Reviews Endocrinology, vol. 8, no. 8, pp. 457–465, 2012.
[18] A. S. Tolmacheva and G. A. Nevinsky, “Essential Protective Role of Catalytically Active Antibodies (Abzymes) with Redox Antioxidant Functions in Animals and Humans,” Antioxidant Functions in Animals and Humans, 2012.
[19] M. Diaz and P. Casali, “Somatic immunoglobulin hypermutation,” Current Opinion in Immunology, vol. 14, no. 2, pp. 235–240, 2002.
[20] Team RC, R: a language and environment for statistical computing, R Foundation for Statistical Computing VAfhwR-po, 2005.
[21] M. C. Burger, C. Zhang, P. N. Harter et al., “CAR-engineered NK cells for the treatment of glioblastoma: turning innate effectors into precision tools for cancer immunotherapy,” Frontiers in Immunology, vol. 10, p. 2683, 2019.
[22] A. J. Cruz-Jentoft, G. Bahat, J. Bauer et al., “Sarcopenia: revised European consensus on definition and diagnosis,” Age and Ageing, vol. 48, no. 1, pp. 16–31, 2019.
[23] T. J. Osterman, M. Terry, and R. S. Miller, “Improving cancer data interoperability: the promise of the minimal common,” JCO Clinical Cancer Informatics, vol. 4, pp. 993–1001, 2020.
I. C. Jones, M. S. Espindola, R. Narayanan et al., “Low expression of RGL4 is associated with a poor prognosis and immune infiltration in lung adenocarcinoma patients,” International Immunopharmacology, vol. 83, article 106454, 2020.

A. K. M. Kasambara and P. Biecek, “survminer: drawing survival curves using ‘ggplot2’. R package version 0.4.8,” 2020, https://CRAN.R-project.org/package=survminer.

P. Blanche, J.-F. Dartigues, and H. Jacqmin-Gadda, “Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks,” Statistics in Medicine, vol. 32, no. 30, pp. 5381–5397, 2013.

G. M. DeNicola, F. A. Karreth, T. J. Humpton et al., “Y. Sun, Y. Zhang, S. Ren et al., ‘Targeting MAP3K19 prevents human lung myofibroblast activation both in vitro and in a humanized SCID model of idiopathic pulmonary fibrosis,’ Scientific Reports, vol. 9, article 19796, 2019.

V. T. Hoang, K. Nyswaner, P. Torres-Ayuso, and J. Brognard, “The protein kinase MAP3K19 phosphorylates MAP2Ks and thereby activates ERK and JNK kinases and increases viability of KRAS-mutant lung cancer cells,” Journal of Biological Chemistry, vol. 295, no. 25, pp. 8470–8479, 2020.

D. A. Levine and The Cancer Genome Atlas Research Network, “Integrated genomic characterization of endometrial carcinoma,” Nature, vol. 497, no. 7447, pp. 67–73, 2013.

Q. Ou-yan, X. He, A. Yang, B. Li, and M. Xu, “Interference with NTSR1 Expression Exerts an Anti-Invasion Effect via the Jun/miR-494/SOCS6 Axis of Glioblastoma Cells,” Cellular Physiology and Biochemistry, vol. 49, no. 6, pp. 2382–2395, 2018.

R. A. Brooks, G. F. Fleming, R. R. Lastra et al., “Current recommendations and recent progress in endometrial cancer,” CA: A Cancer Journal for Clinicians, vol. 69, no. 4, pp. 258–279, 2019.

S. Sud, J. Holmes, M. Eblan, R. Chen, and E. Jones, “Clinical characteristics associated with racial disparities in endometrial cancer outcomes: a surveillance, epidemiology and end results analysis,” Gynecologic Oncology, vol. 148, no. 2, pp. 349–356, 2018.

M. R. Elliott and K. S. Ravichandran, “The dynamics of apoptotic cell clearance,” Developmental Cell, vol. 38, no. 2, pp. 147–160, 2016.

E. A. Ross, A. Devitt, and J. R. Johnson, “Macrophages: the good, the bad, and the glutony,” Frontiers in Immunology, vol. 12, article 70816, 2021.

Y. Wang, F. Ren, Z. Song, X. Wang, and X. Ma, “Multimomics profile and prognostic gene signature of m6A regulators in uterine corpus endometrial carcinoma,” Journal of Cancer, vol. 11, no. 21, pp. 6390–6401, 2020.

S. Zuo, M. Wei, S. Wang, J. Dong, and J. Wei, “Pan-cancer analysis of immune cell infiltration identifies a prognostic Immune-Cell Characteristic Score (ICCS) in Lung Adenocarcinoma,” Frontiers in Immunology, vol. 11, article 1218, 2020.

J. Li, L. Zhu, Y. Shi, J. Liu, L. Lin, and X. Chen, “m6A demethylase FTO promotes hepatocellular carcinoma tumorigenesis via mediating PKM2 demethylation,” American Journal of Translational Research, vol. 11, no. 9, pp. 6084–6092, 2019.