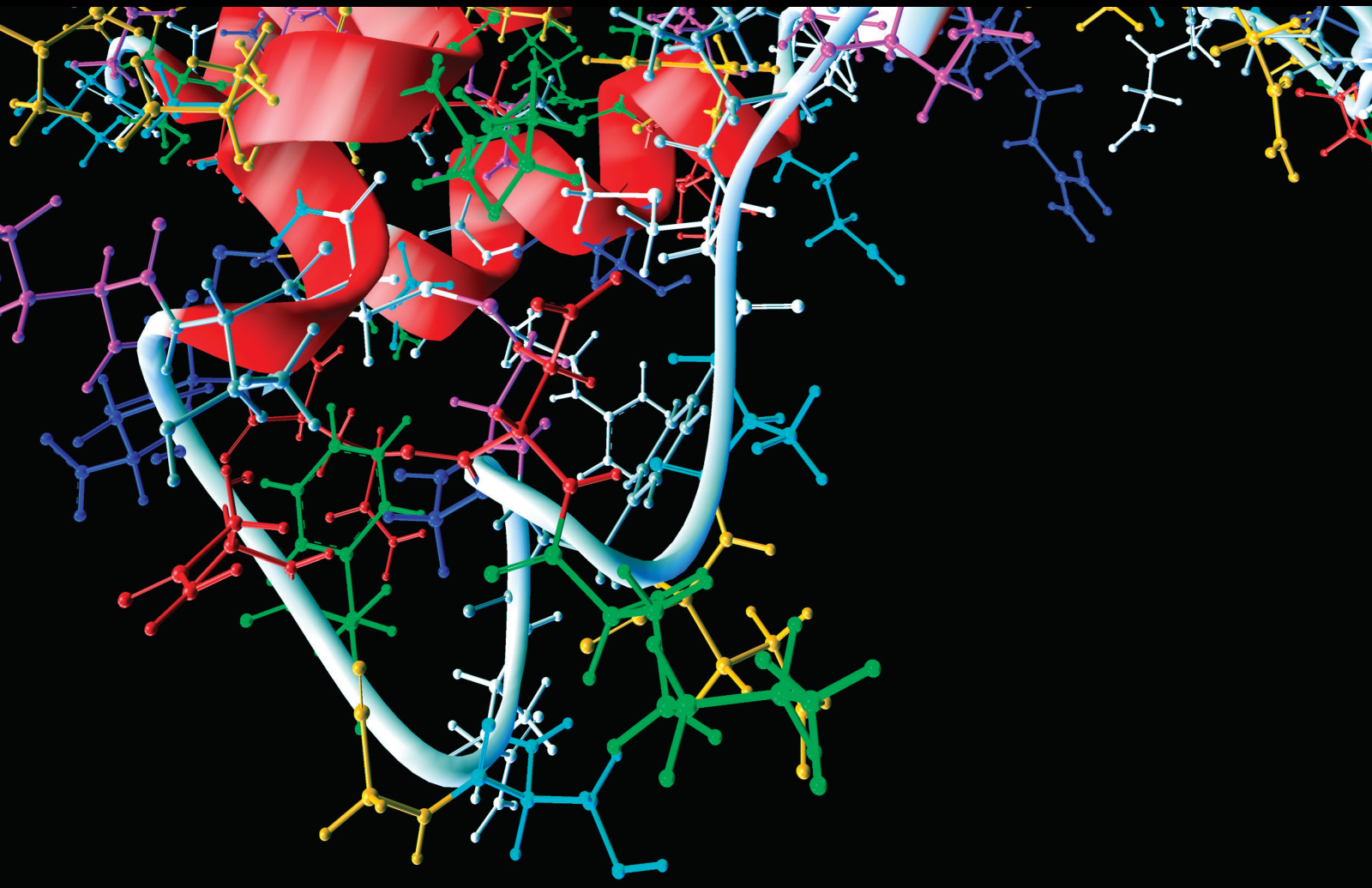
Computational and Mathematical Methods in Medicine

Machine Learning Approaches Based on Multiscale Data in Diagnosis, Treatment and Prognosis of Diseases

Lead Guest Editor: Jincheng Wang
Guest Editors: Xudong Zhang, Xiaohan Ren, Hongye Wang, Xiaoye Xu, and Neelam Yadav



Machine Learning Approaches Based on Multiscale Data in Diagnosis, Treatment and Prognosis of Diseases

Computational and Mathematical Methods in Medicine

Machine Learning Approaches Based on Multiscale Data in Diagnosis, Treatment and Prognosis of Diseases

Lead Guest Editor: Jincheng Wang
Guest Editors: Xudong Zhang, Xiaohan Ren, Hongye Wang, Xiaoye Xu, and Neelam Yadav

Contents

Contents

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 1318817, 13 pages
https://doi.org/10.1155/2023/1318817

Research Article

Circular RNA_HIPK3-Targeting miR-93-5p Regulates KLF9 Expression Level to Control Acute Kidney Injury

Zha Zhengbiao,[1] Chen Liang,[2] Zheng Zhi,[1] and Pan Youmin[1]

[1]*Department of Cardiovascular Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, China*
[2]*Department of Infection, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, China*

Correspondence should be addressed to Pan Youmin; panyoumincs@126.com

Received 5 August 2022; Revised 6 October 2022; Accepted 11 October 2022; Published 16 February 2023

Academic Editor: Jincheng Wang

Acute kidney injury (AKI) is a clinical syndrome caused by various reasons that results in the rapid decline of renal function in a short period of time. Severe AKI can lead to multiple organ dysfunction syndrome. Circular RNA HIPK3 (circHIPK3) derived from the *HIPK3* gene is involved in multiple inflammatory processes. The present research was performed to explore the function of circHIPK3 on AKI. The AKI model was established by ischemia/reperfusion (I/R) in C57BL/6 mice or hypoxia/reoxygenation (H/R) in HK-2 cells. The function and mechanism of circHIPK3 on AKI were explored via biochemical index measurement; hematoxylin and eosin (HE) staining; 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT); flow cytometry; enzyme-linked immunosorbent assay (ELISA); western blot; quantitative real-time polymerase chain reaction (RT-qPCR); detection of reactive oxygen species (ROS) and adenosine triphosphate (ATP); and luciferase reporter assays. circHIPK3 was upregulated in kidney tissues of I/R-induced mice and in H/R-treated HK-2 cells, while the microRNA- (miR-) 93-5p level was decreased in H/R-stimulated HK-2 cells. Furthermore, circHIPK3 silencing or miR-93-5p overexpression could reduce the level of proinflammatory factors and oxidative stress and recover the cell viability in H/R-stimulated HK-2 cells. Meanwhile, the luciferase assay showed that Krüppel-like transcription factor 9 (KLF9) was the downstream target of miR-93-5p. Forced expression of KLF9 blocked the function of miR-93-5p on H/R-treated HK-2 cells. Knockdown of circHIPK3 improved the renal function and reduced the apoptosis level *in vivo*. In conclusion, circHIPK3 knockdown alleviated oxidative stress and apoptosis and inhibited inflammation in AKI via miR-93-5p-mediated downregulation of the KLF9 signal pathway.

1. Introduction

Renal ischemia/reperfusion injury is an important factor leading to acute kidney injury (AKI), delayed recovery of transplanted renal function, and even acute renal failure in severe cases [1, 2]. In the process of renal ischemia, endothelial cells are damaged to induce the production of endothelin, the expression of various receptors, the oxygen-free radical production, and Ca^{2+} overload. These various factors together cause renal microcirculation disorder and hemodynamic changes. In the process of renal reperfusion, cellular structure destruction and metabolic disorders are more serious than these during ischemia, and renal tubular epithelial cells located in the medulla trigger inflammatory responses and lead to cell edema [3–5]. At present, due to the unclear pathogenesis of renal ischemia/reperfusion, there is no specific treatment to reduce its incidence and improve its cure rate. Therefore, understanding the molecular pathogenesis of renal ischemia/reperfusion injury can effectively contribute to the progression, identify the preventive treatment, and block the incidence of renal ischemia/reperfusion injury.

Plenty of researches show that circRNA is involved in the biological processes such as mRNA splicing and transcription and RNA degradation and translation and is closely related to the occurrence and development of a variety of human diseases such as cancers, nervous system disorders, and cardiovascular diseases [6, 7]. Functionally, circRNAs have competitive binding sites for miRNA, which

Table 1: Primers used in the qRT-PCR analysis.

Gene	Forward (5′-3′)	Reverse (5′-3′)
circHIPK3	TTCAACATGTCTACAATCTCGGT	ACCATTCACATAGGTCCGT
miR-93-5p	ACCATTCACATAGGTCCGTGAGCTGCCC	CTCAACTGGTGTCGTGGA
Klf9	CGAGCGGCTGCGACTACCTG	GGGCTGTGGGAAGGACTCGAC
U6	GCTTCGGCAGCACATATACT	GTGCAGGGTCCGAGGTATTC
Gapdh	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA

can regulate miRNA activity and ultimately affect the expression of downstream target genes of miRNA [8]. circRNAs participate in the regulation of organ ischemia/reperfusion injury. For instance, circVMA21 inhibits apoptosis and inflammation in sepsis-associated AKI models through the miR-9-3p/SMG1 axis [9]. circ_0114427 can regulate ATF3 by sponging miR-494 to be involved in AKI [10]. Forced expression of circYAP1 reduces inflammation and ROS production in ischemia/reperfusion-treated cells by sponging to miR-21-5p [11]. These findings suggest that circRNAs are new molecular targets for precision diagnosis and treatment of ischemia/reperfusion diseases. circHIPK3 (has_circRNA_000284) is produced from exon2 of the HIPK3 gene and the most abundant spliceosome in the circRNAs of the HIPK3 gene [12]. circHIPK3 has been demonstrated to be dysregulated in various diseases, such as cancers [13, 14], diabetes [15], and fibrosis [13–19]. More importantly, circHIPK3 can sponge diverse miRNAs to regulate the diseases progression. circHIPK3 sponges miR-382-5p to regulate DUSP1 to alleviate neuronal apoptosis and inflammatory response in spinal cord injury [20]. circHIPK3 modulates cell proliferation and migration through the miR-124/AKT3 axis in esophageal squamous cell carcinoma [14]. However, the role of circHIPK3 in renal ischemia/reperfusion injury is still unknown.

Here, we found that circHIPK3 was upregulated in AKI. Silencing of circHIPK3 reduced oxidative stress, apoptosis, and inflammation in both the mouse model and HK-2 cells. Furthermore, we identified miR-95-3p as the direct target of circHIPK3, and Krüppel-like transcription factor 9 (KLF9) would be a downstream protein of the circHIPK3/miR-95-3p axis. In summary, we found that circHIPK3 knockdown could alleviate renal AKI via regulating the miR-95-3p/KLF9 signaling pathway.

2. Materials and Methods

2.1. Animals. C57BL/6 mice (6-8 weeks, ~25 g, male) were purchased from Charles River (Beijing, China) and kept at the animal center for more than one week before the following experimental procedures. All experimental mice were fed at 22°C in standard breeding cages with a standard condition. The Laboratory Animal Ethics Committee of Tongji Hospital approved all experimental procedures using the C57BL/6 mice.

2.2. AKI Model Establishment. Mice were subjected to acute ischemia and reperfusion (I/R) treatment to induce ischemic renal AKI as in a previous report [21]. Mice were anesthetized with 50 mg/kg (body weight) pentobarbital sodium. After deep anesthesia, the back of the mouse was opened with a 1-5 cm long incision, and both renal pedicles were clipped with a micro noninvasive hemostatic clamp for 25 min and then treated with reperfusion for 6 h. Mice in the sham group received the same operation without clamping of kidney pedicles. After the operation, the backs were sutured and the mice were put back into the cage for recovery. In addition, to explore the role of circHIPK3 in vivo, adenovirus vector plasmids including AAV9-si-circHIPK3 and AAV9-si-negative control (AAV9-si-NC) were bought from GenePharma (Shanghai, China) and injected into mice by tail vein 21 days before I/R surgery. Mice were intraperitoneally injected with pentobarbital sodium and were euthanatized via excessive anesthesia with a dose of 250 mg/kg [22]. Their blood and kidney tissues were harvested and processed for biochemical, pathological, and immunoblotting analyses.

2.3. Biochemical Index Measurement. After reperfusion, about 2 mL of blood samples was gathered from each mouse via the posterior orbital venous plexus. The blood samples were then centrifuged at $4{,}000 \times g$ for 15 min, and the serum samples were collected and used for the determination of urea nitrogen (BUN), uric acid (UA), and creatinine abundance (CREA) using a Roche Cobas C111 analyzer (Roche, Switzerland). At least three biological repeats were performed for statistical comparisons of BUN, UA, and CREA between different groups.

2.4. Hematoxylin and Eosin (H&E) Staining. Kidney tissues were isolated and stained with H&E (G1120, Solarbio, Beijing, China). The tissues were fixed in 4% paraformaldehyde (P1110, Solarbio) for 1 day and embedded in paraffin (YA0012, Solarbio). The tissue was sectioned about 4 μm. The slices were stained with hematoxylin, eosin, and toluidine blue and then were used to observe the pathological changes of the tissues under a light microscope (BX43, Olympus, Tokyo, Japan).

2.5. Cell Culture. The human renal proximal tubular epithelial cell line (HK-2) was incubated at 37°C in 10% FBS (SH30413.03, Hyclone, South America) contained DMEM (SH30285.FS, Hyclone) in 5% CO_2 and 95% air. The cell hypoxia/reoxygenation (H/R) model was established as in a previous report [23]. Briefly, HK-2 cells were exposed to hypoxia 1% oxygen, 94% nitrogen, and 5% CO_2 for 6 h in DMEM. Subsequently, the culture plates were placed in a

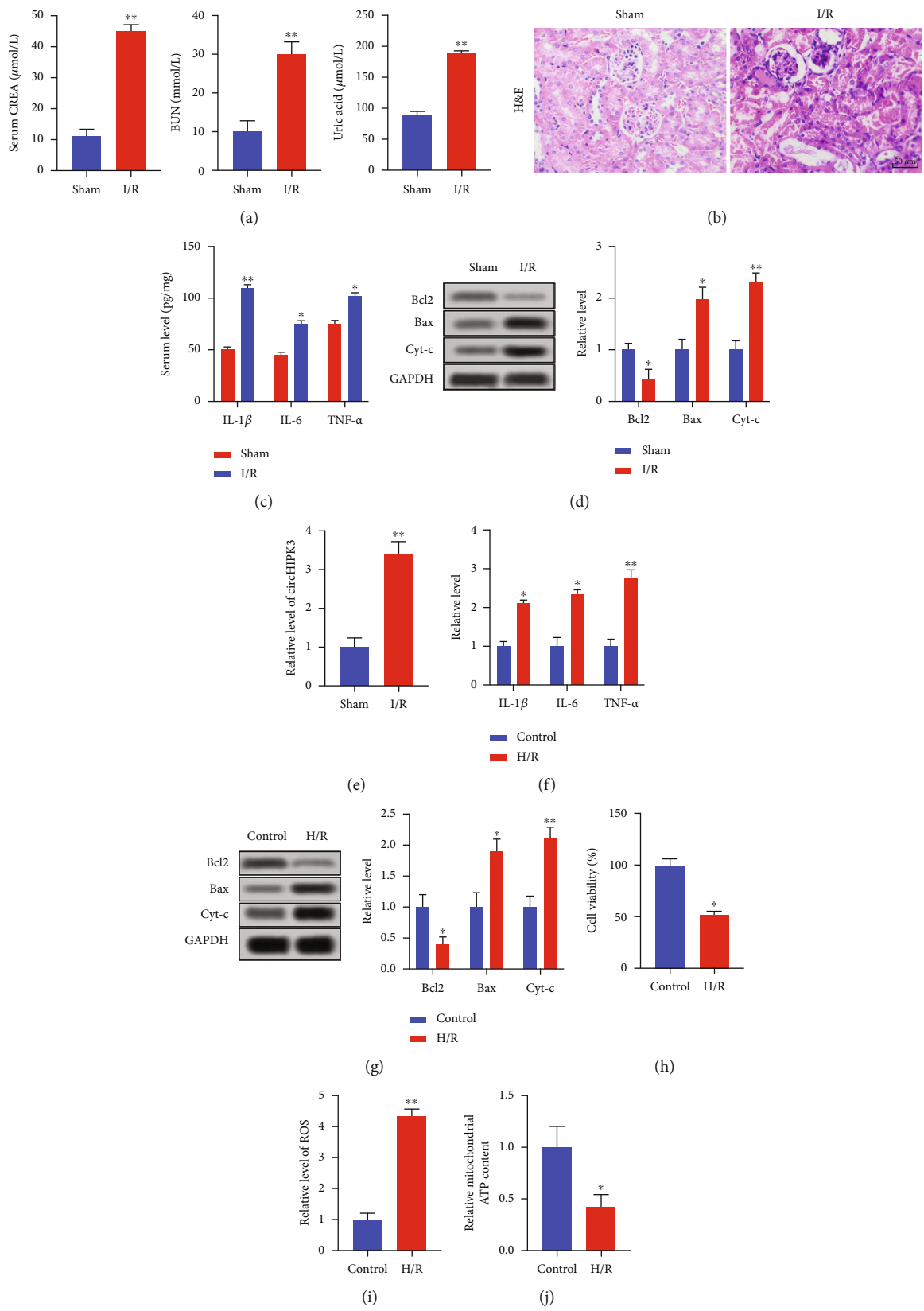
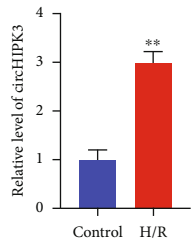


Figure 1: Continued.



(k)

Figure 1: circHIPK3 is upregulated in acute kidney injury. (a) The level of serum CREA, BUN, and uric acid in I/R and sham groups. (b) Representative H&E images are shown in I/R and sham groups. (c) The level of IL-1β, IL-6, and TNF-α was detected via ELISA. (d) The level of apoptosis-associated protein (Bcl2, Bax, and Cyt-c) was measured via western blot. (e) The level of circHIPK3 was assessed in the I/R and sham groups. (f) H/R treatment was performed in the HK-2 cell, and the level of IL-1β, IL-6, and TNF-α was detected via ELISA. (g) The level of apoptosis-associated protein was measured via western blot. (h). Cell viability was detected via MTT. (i) The production of ROS was evaluated via the Reactive Oxygen Species Assay Kit. (j) The content of ATP was measured via the ATP Assay Kit. (k) The level of circHIPK3 was assessed in the H/R and control groups. $^{*}P < 0.05$ and $^{**}P < 0.01$ vs. Control. The experiments were repeated at least three times. The results were expressed as mean ± SD.

normal cell incubator (5% CO_2 and 95% air) for reoxygenation for 6 h.

2.6. 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium Bromide (MTT) Assay. HK-2 cells with an inoculation density of 1×10^5 cells/well were sowed into 96-well plates and incubated at 37°C in 5% CO_2 for 24 h. After cells were incubated with 10 μL of MTT solution (C0009S, Beyotime, Shanghai, China) for 4 h, the culture supernatant was removed, and each well was appended with 100 μL of DMSO to dissolve the crystals. The absorbance at 570 nm was detected using the microplate reader (Tecan, Switzerland).

2.7. Flow Cytometry. HK-2 cells were plated into 24-well plates at a density of 2.5×10^5 cells/well and cultured at 37°C in 5% CO_2 overnight. Then, cells were gathered, washed with phosphate buffer saline (PBS) (C0221A, Beyotime), resuspended with 0.5 mL of bind buffer, and stained with 5 μL of propidium iodide (PI) (P1304MP, Thermo Fisher Scientific, Waltham, MA, USA) and 5 μL of Annexin V/FITC (BMS306FI-100, Thermo Fisher Scientific) at room temperature for 15 min. The cell apoptosis was determined on a FACScan flow cytometer using CellQuest software (BD Biosciences, NJ, USA).

2.8. Enzyme-Linked Immunosorbent Assay (ELISA). The concentrations of interleukin- (IL-) 1β, IL-6, and tumor necrosis factor- (TNF-) α in HK-2 cells were detected via the Human IL-1β ELISA Kit (PI305, Beyotime), Human IL-6 ELISA Kit (PI330, Beyotime), and Human TNF-α ELISA Kit (PT518, Beyotime), respectively. The concentrations of IL-1β, IL-6, and TNF-α in kidney tissues were measured via the Mouse IL-1β ELISA Kit (PI301, Beyotime), Mouse IL-6 ELISA Kit (PI326, Beyotime), and Mouse TNF-α ELISA Kit (PT512, Beyotime), respectively.

2.9. Western Blot. The protein levels were assessed via western blotting. The protein samples from HK-2 cells and kidney tissues were lysed with the RIPA lysis buffer (P0013B, Beyotime), and the protein concentration was determined via the BCA kit (P0012S, Beyotime, China). 50-60 μg of protein samples was separated via sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 12% gel and transferred to nitrocellulose membranes (HF18002S25, Merck, NJ, USA). The blot was then incubated in a blocking solution (5% defatted dry milk (P0216, Beyotime)), dissolved in a phosphate buffer saline buffer (PBS, C0221A, Beyotime) followed by washes with PBS, and incubated overnight in PBS containing antibodies (anti-Bcl2 (12789-1-AP, Proteintech, Wuhan, China), anti-Bax (50599-2-Ig, Proteintech), anti-Cytochrome c (Cyt-c) (66264-1-Ig, Proteintech), anti-KLF9 (ab227920, Abcam, Cambridge, UK), and anti-GAPDH (10494-1-AP, Proteintech)) diluted to 1 : 500. The blots were then washed with PBST (PBS containing Tween 20 (ST825, Beyotime)) and incubated for 1 h in PBS containing peroxidase-conjugated anti-mouse/rabbit IgG (PR30009, Proteintech) diluted to 1 : 8000. The blot was developed with the chemiluminescence ECL kit (PE0010, Solarbio, Beijing, China) using the Omega Lum W Chemiluminescence Imaging System (Aplegen, USA), and the optical density was measured using the Image Studio Software. The results were normalized via the loading control (GAPDH).

2.10. Quantitative Real-Time Polymerase Chain Reaction. Total RNAs from HK-2 cells and kidney tissues were collected via the TRIzol solution (15596026, Invitrogen, CA, USA). 500 ng of RNA was reverse transcribed into cDNA using the HiScript II Q RT SuperMix (R223-01, Vazyme, Nanjing, China). Quantitative real-time PCR was performed in an ABI 7500 Fast System using the qPCR SYBR Green Master Mix (4309155, Thermo Fisher, MA, USA). The PCR amplification conditions were 94°C for 10 min, 94°C for 10 s, and 60°C for 45 s for 40 cycles. *GAPDH* acted as the internal reference. The expressions of genes were calculated with the comparative threshold cycle method ($2^{-\Delta\Delta CT}$ method), in which $\Delta\Delta CT = \Delta CT_{treatment} - \Delta CT_{control}$ and $\Delta CT = Ct_{target} - Ct_{reference}$. The primers are shown in Table 1.

2.11. Reactive Oxygen Species (ROS) Production Detection. The Cellular ROS Assay Kit (ab186027) was purchased from

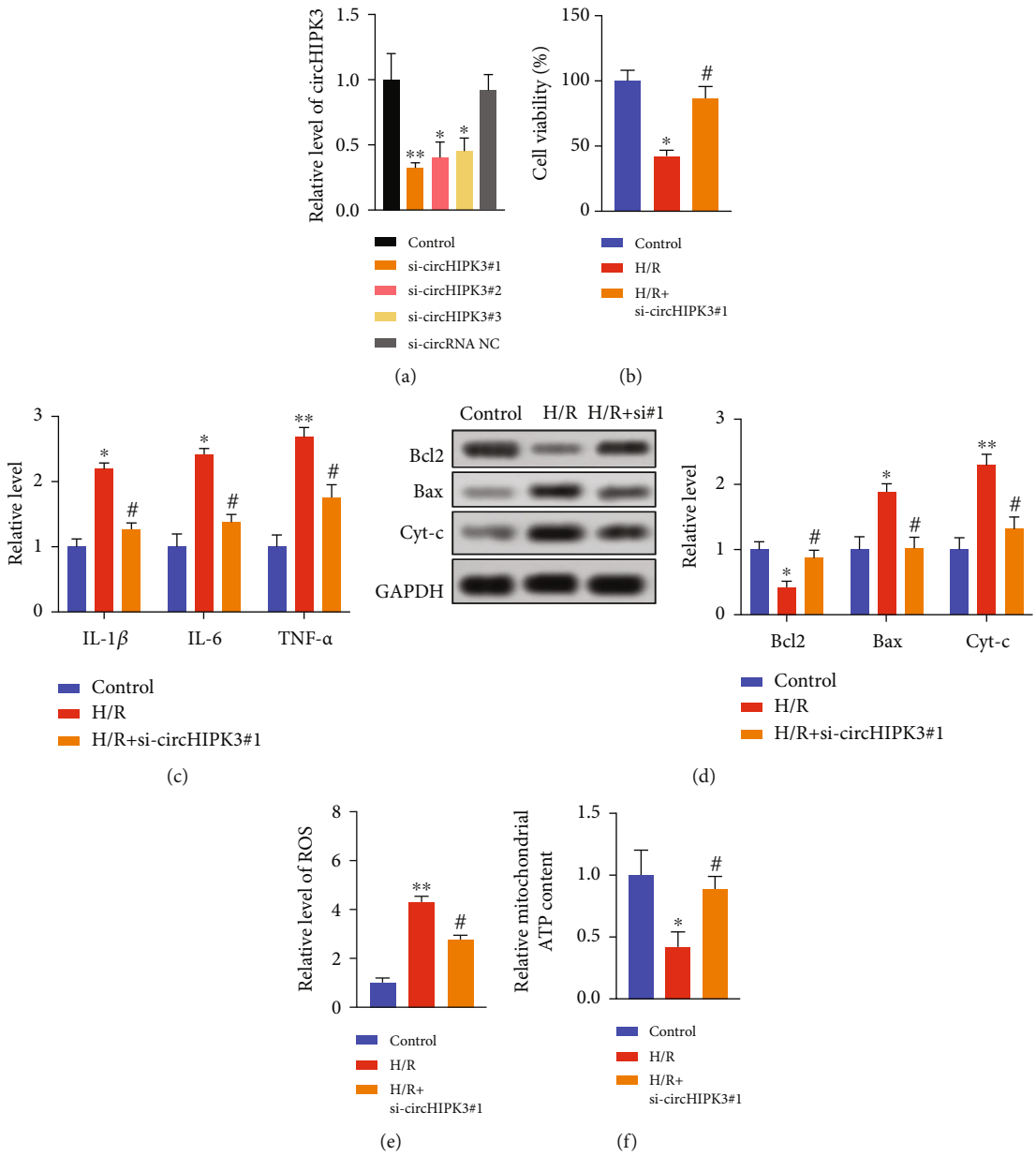


Figure 2: Knockdown of circHIPK3 alleviates the kidney injury *in vitro*. (a) The transfection effectiveness was detected via qRT-PCR. $n = 5$. (b) Cell viability was measured in different groups. (c) The level of IL-1β, IL-6, and TNF-α was detected via ELISA. (d) The level of apoptosis-associated protein (Bcl2, Bax, and Cyt-c) was measured via western blot. (e) The production of ROS was evaluated. (f) The content of ATP was measured via the ATP Assay Kit. $^{*}P < 0.05$ and $^{**}P < 0.01$ vs. Control group and $^{\#}P < 0.05$ vs. H/R group. The experiments were repeated at least three times. The results were expressed as mean ± SD.

Abcam. Cells were cultured and collected at 10^4 cells/100 μL per well. Furthermore, 100 μL/well of ROS Red Working Solution was added into the cell plate and incubated in a 37°C/5% CO_2 incubator for 1 h. Fluorescence activity was measured at Ex/Em = 520/605 nm (cut-off 590 nm) with the bottom read mode. The final results were normalized by the control group.

2.12. ATP Measurement Assay. The ATP Assay Kit (S0026, Beyotime, China) was employed to measure the cellular ATP content. Briefly, HK-2 cells were seeded in 12-well plates. After treatment, the cells were lysed with buffer and centrifuged at 12000 × g for 5 min at 4°C. The supernatant was collected. Luminance was measured via a monochromator microplate reader (Tecan). The data were normalized by the control group.

2.13. Luciferase Reporter Assay. The binding sites between circHIPK3 and miR-93-5p, as well as these between miR-93-5p and KLF9, were predicated by the starBase databases

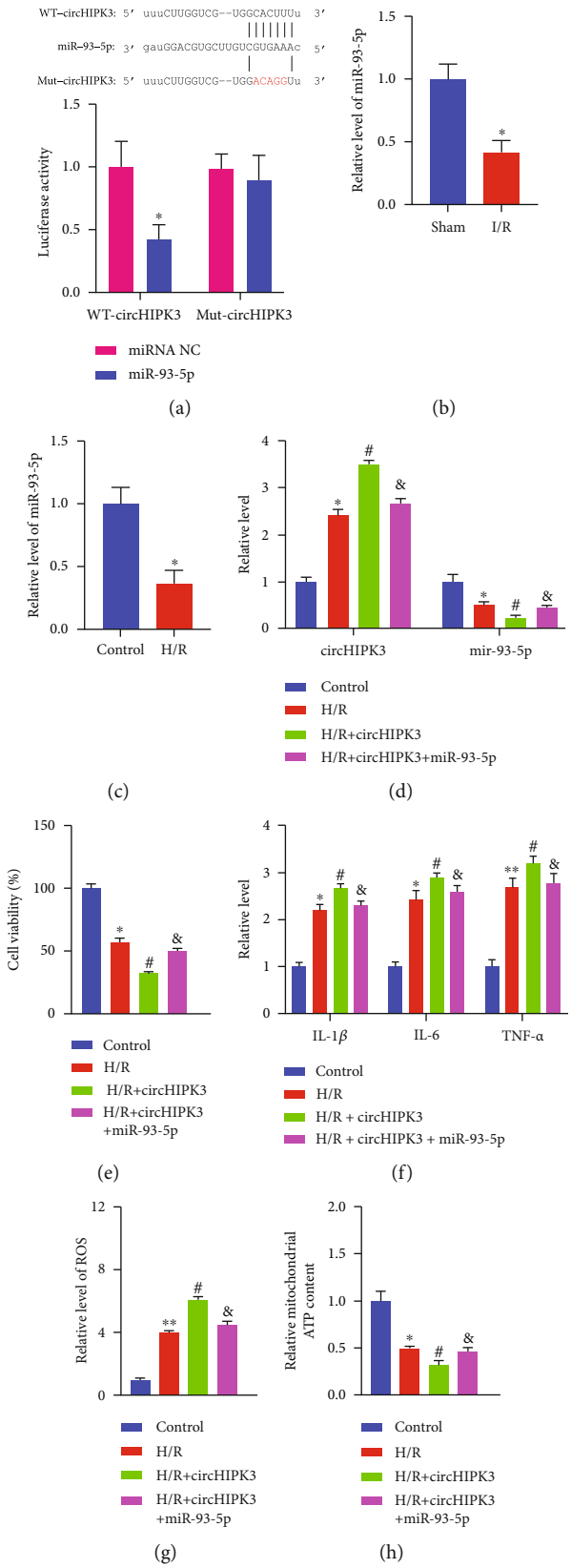
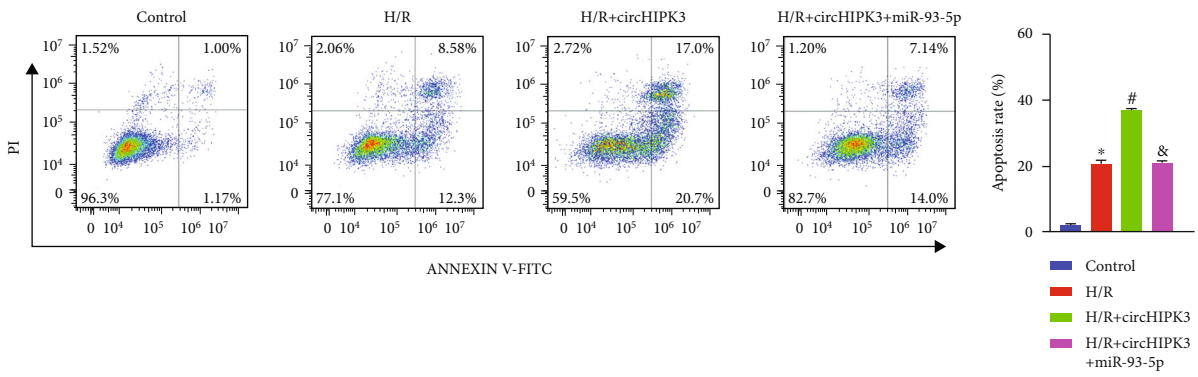


Figure 3: Continued.



(i)

Figure 3: circHIPK3 interacting with miR-93-5p regulates AKI progression. (a) Schematic diagram of bioinformatics prediction binding sites (upper). Luciferase assay report for circHIPK3 and miR-93-5p (lower). (b, c) The level of miR-93-5p was measured *in vivo* and *in vitro*. (d) The level of circHIPK3 and miR-93-5p was detected via qRT-PCR. (e) Cell viability was measured in different groups. (f) The level of IL-1β, IL-6, and TNF-α was detected via ELISA. (g) The production of ROS was evaluated. (h) The content of ATP was measured via ATP Assay Kit. (i) The apoptosis rate of the HK-2 cell after different treatments was detected via flow cytometry. $^{*}P < 0.05$ vs. Control group, $^{\#}P < 0.05$ vs. H/R group, and $^{\&}P < 0.05$ vs. H/R+circHIPK3 group. The experiments were repeated at least three times. The results were expressed as mean ± SD.

(https://starbase.sysu.edu.cn/starbase2/index.php). HEK293 cells were transfected with 20 mmol/L of miR-93-5p mimic or miR-NC (GenePharma) together with circHIPK3-WT/circHIPK3-mutation or KLF9-WT/KLF9-mutation. Luciferase activity was determined with the Dual Luciferase Reporter Assay Kit (E1910, Promega, Beijing, China) on GloMax 20/20 (Promega) at 48 h after the transfection.

2.14. Statistical Analysis. All data were analyzed as the mean ± SD. Statistical significances were measured via the paired Student's *t*-test between two groups and the one-way analysis of variance (ANOVA) for more than two groups followed by the post hoc Bonferroni test using the SPSS 26.0 software (IBM, Armonk, New York, USA). $P < 0.05$ was regarded as a significant difference.

3. Results

3.1. circHIPK3 Was Upregulated In Vivo and In Vitro. circHIPK3 plays an important role in I/R injury which occurs in a variety of organs [24–27]. To investigate the abnormal gene expression of *circHIPK3* in response to renal AKI, mice were subjected to renal I/R. Comparing with the sham group, the serum levels of CREA, BUN, and UA were significantly upregulated in the I/R group (Figure 1(a)). Then, the kidney was removed and H&E staining was performed, and tubular dilatation and necrosis were observed in the I/R group (Figure 1(b)). The serum levels of inflammatory cytokines IL-1β, IL-6, and TNF-α were upregulated in the I/R group (Figure 1(c)). The decreased expression of Bcl2 and increased expression of Bax and Cytochrome c (Cyt-c) were discovered, suggesting the inducing apoptosis in the I/R group (Figure 1(d)). The upregulation of circHIPK3 was found in the I/R group compared with the sham group (Figure 1(e)). In addition, HK-2 cells underwent hypoxia/reoxygenation (H/R) or control treatment. Compared with the control group, the levels of IL-1β, IL-6, and TNF-α were markedly upregulated in H/R-treated HK-2 cells (Figure 1(f)). The increased apoptosis level was assessed by detecting the expression of Bcl2, Bax, and Cyt-c (Figure 1(g)). The MTT assay showed the decreased cell viability in H/R-treated HK-2 cells compared with the control group (Figure 1(h)). Meanwhile, H/R induced the increased production of ROS (Figure 1(i)) and decreased level of ATP (Figure 1(j)). Consistent with *in vivo* results, the circHIPK3 level was elevated after H/R (Figure 1(k)). These results suggested that the expression of circHIPK3 was upregulated in I/R-treated mice and in H/R-treated HK-2 cells.

3.2. Silencing of circHIPK3 Could Alleviate the AKI In Vitro. To explore the function of circHIPK3, si-circHIPK3 and si-circRNA NC were transfected into HK-2 cells. All three si-circHIPK3 inhibited the expression of circHIPK3, among which si-circHIPK3#1 displayed the best interference efficiency (Figure 2(a)). Thus, si-circHIPK3#1 was used in the following experiments. Silencing of circHIPK3 improved cell viability in H/R-treated HK-2 cells (Figure 2(b)). si-circHIPK3#1 prevented the increased level of inflammatory cytokines in HK-2 cells induced by H/R treatment (Figure 2(c)). Besides, knockdown of circHIPK3 could significantly reduce the protein expression of Bax and Cyt-c with enhanced expression of the Bcl2 protein in H/R-treated HK-2 cells (Figure 2(d)). circHIPK3 downregulation reversed ROS generation (Figure 2(e)) and recovered the ATP content in H/R-treated HK-2 cells (Figure 2(f)). In summary, knockdown of circHIPK3 improved the AKI in HK-2 cells.

3.3. circHIPK3 Interacting with miR-93-5p Participated in AKI. Emerging researches have shown that circRNAs can function as miRNA sponges to regulate downstream targets [28]. Bioinformatics starBase databases predicted that there were binding sites between circHIPK3 and miR-93-5p (Figure 3(a)). To verify the targeting relationship between

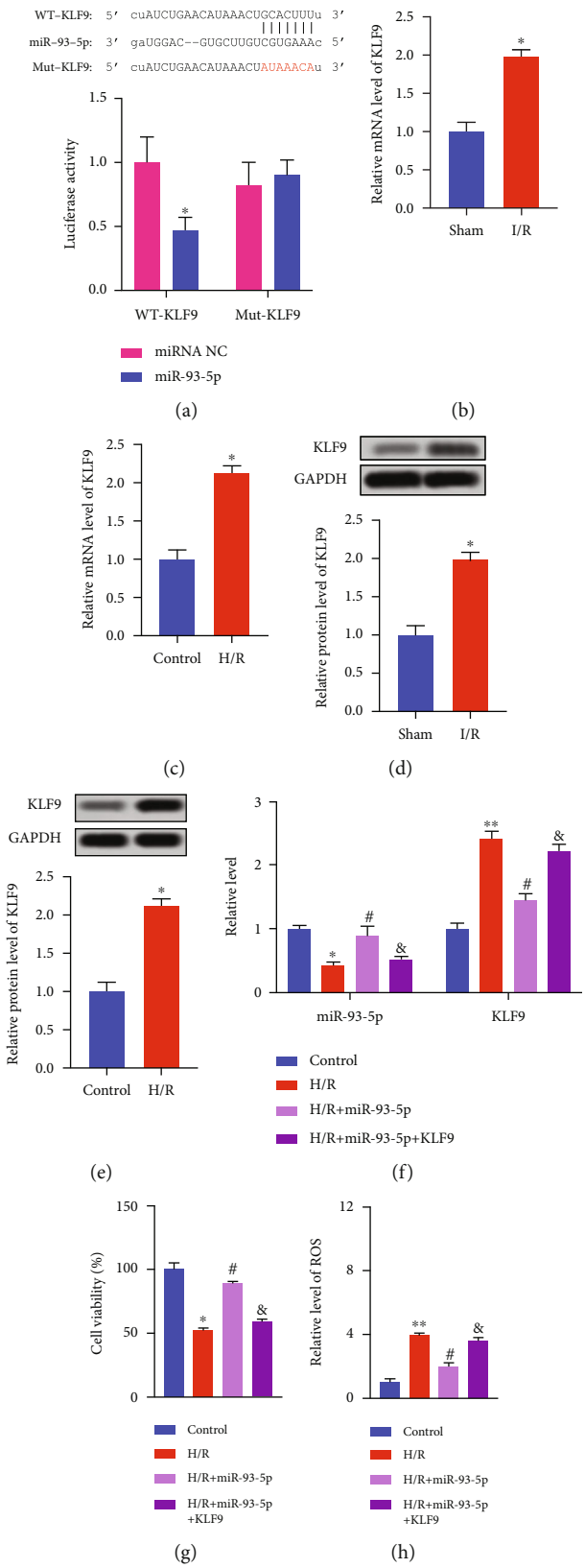
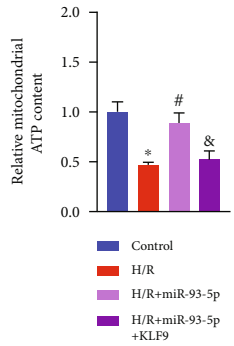
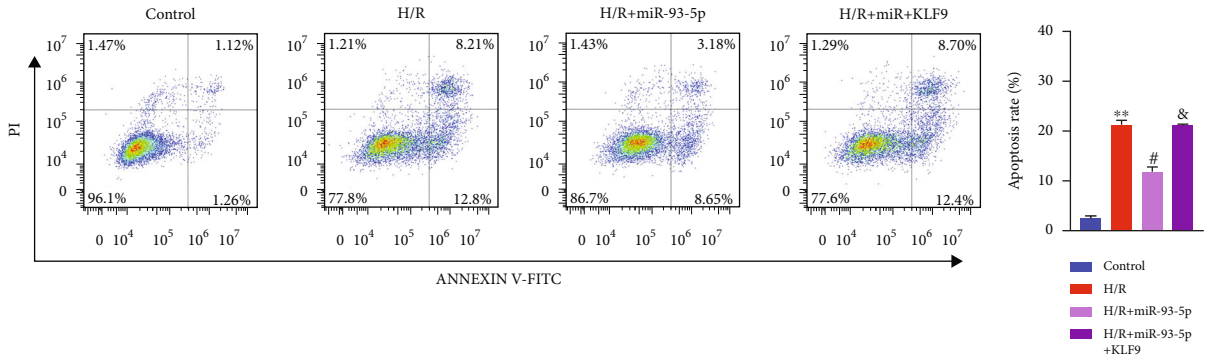


Figure 4: Continued.



(i)



(j)

Figure 4: KLF9 is a target of miR-93-5p. (a) Schematic diagram of bioinformatics prediction binding sites (upper). Luciferase assay report for KLF9 and miR-93-5p (lower). (b, c) The mRNA level of KLF9 was measured *in vivo* and *in vitro*. (d, e) The protein level of KLF9 was measured *in vivo* and *in vitro*. (f) The level of miR-93-5p and KLF9 was detected via qRT-PCR. (g) Cell viability was measured in different groups. (h) The production of ROS was evaluated. (i) The content of ATP was measured via ATP Assay Kit. (j) The apoptosis rate of HK-2 cell after different treatments was detected via flow cytometry. $^{*}P < 0.05$ and $^{**}P < 0.01$ vs. Control group, $^{\#}P < 0.05$ vs. H/R group, $^{\&}P < 0.05$ vs. H/R+KLF9 group, and $^{\&}P < 0.05$ vs. H/R+KLF9 group. The experiments were repeated at least three times. The results were expressed as mean ± SD.

circHIPK3 and miR-93-5p, mutant circHIPK3 and wild-type circHIPK3 were constructed. Then, we performed luciferase reporter assay in HEK293 cells. As shown in Figure 3(a), the decreased luciferase activity was observed in cotransfection with WT-circHIPK3 and the miR-93-5p mimic, suggesting a direct bind between circHIPK3 and miR-93-5p. Additionally, the level of miR-93-5p was downregulated both *in vivo* and *in vitro* (Figures 3(b) and 3(c)). circHIPK3 overexpression plasmid and miR-93-5p mimic were cotransfected into HK-2 cells. The circHIPK3 and miR-93-5p level was detected via qRT-PCR (Figure 3(d)). Forced expression of circHIPK3 further aggravated H/R-induced cell viability (Figure 3(e)), inflammatory cytokine level (Figure 3(f)), ROS level (Figure 3(g)), and ATP content (Figure 3(h)), and the miR-93-5p mimic could block the effect of the circHIPK3 plasmid on these indexes (Figures 3(e)–3(h)). Flow cytometry results also showed that the apoptosis rate was further induced by circHIPK3 which was blocked by the miR-93-5p mimic (Figure 3(i)). The above results showed that circHIPK3 could act as a sponge for miR-93-5p in I/R-stimulated HK-2 cells.

3.4. miR-93-5p Binding with KLF9 Regulated Kidney Injury In Vitro. Next, we forecasted that KLF9 may be a target of miR-93-5p, which was verified via the luciferase assay (Figure 4(a)). Meanwhile, the mRNA level of KLF9 was upregulated both *in vivo* and *in vitro* (Figures 4(b) and 4(c)). Western blot assay exhibited the similar results (Figures 4(d) and 4(e)). We added miR-93-5p mimics and the KLF9 vector into HK-2 cells and found that KLF9 recovered the level of miR-93-5p and KLF9 (Figure 4(f)). miR-93-5p mimics showed the similar function with si-circHIPK3 on cell viability, ROS level, ATP content, and apoptosis rate in H/R-treated HK-2 cells, while overexpression of KLF9 inhibited the favorable effect of miR-93-5p mimics on H/R-treated HK-2 cells (Figures 4(g)–4(j)). Taken together, KLF9 might be a target gene for miR-93-5p.

3.5. circHIPK3 Regulated AKI by Mediating miR-93-5p/KLF9 Axis In Vivo. To verify our results, we constructed the adenovirus vector plasmid, AAV9-si-circHIPK3/AAV9-si-negative control (AAV9-si-NC). The mice were injected with AAV9-si-NC or AAV9-si-circHIPK3 21 days before I/R surgery by tail vein. After operation, the kidneys in all groups were collected, and the levels of circHIPK3, miR-93-5p, and KLF9 were detected. circHIPK3 injection induced the upregulation of circHIPK3 and KLF9 and inhibited the expression of miR-93-5p (Figure 5(a)). The serum levels of

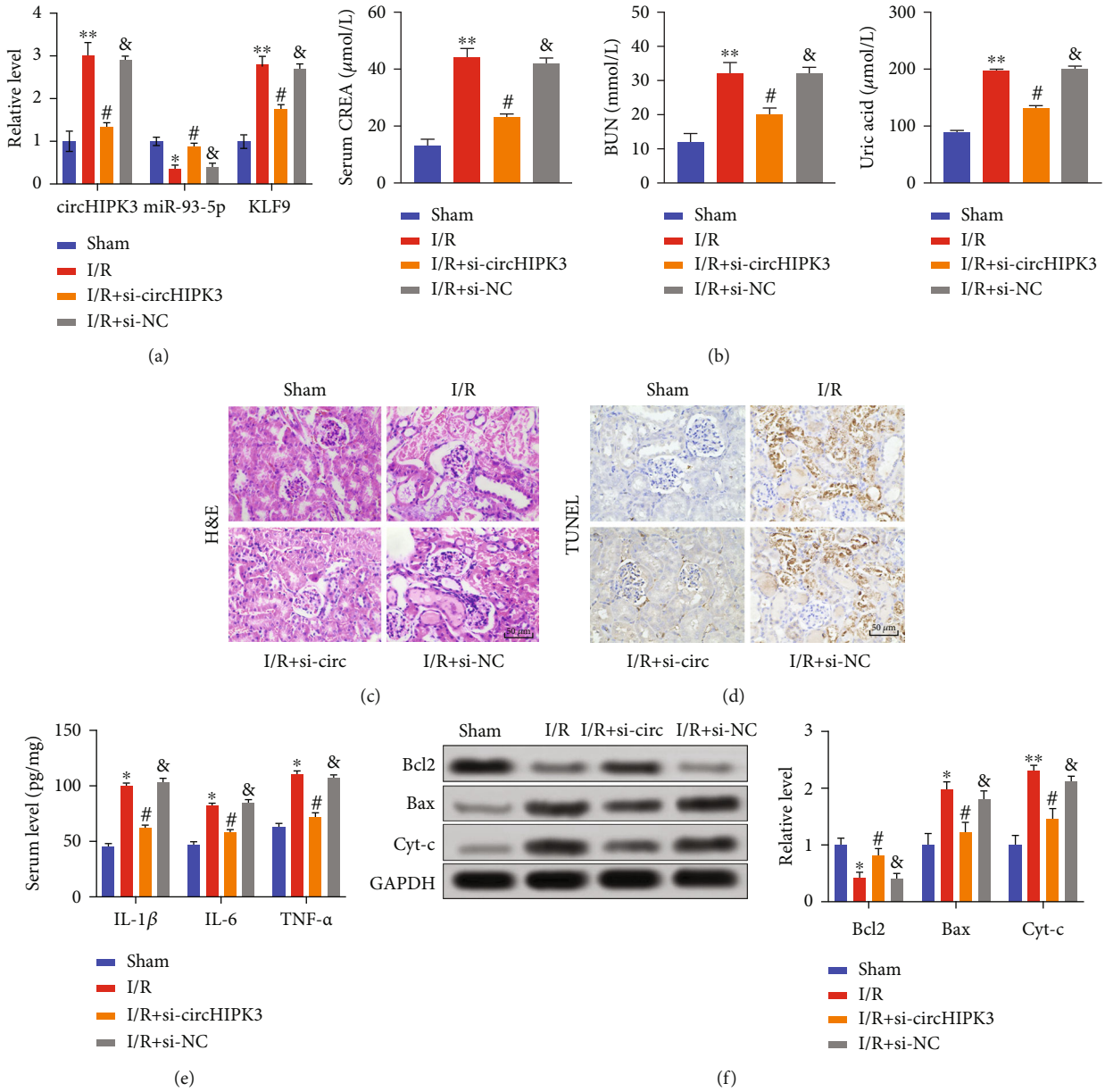


Figure 5: circHIPK3 is implicated in AKI via regulating the miR-93-5p/KLF9 axis. (a) AAV9 vector plasmid was injected into mice before operation. The level of circHIPK3, miR-93-5p, and KLF9 was detected via qRT-PCR. (b) The level of serum CREA, BUN, and uric acid in different groups. (c) Representative H&E images are shown in different groups. The section shows acute tubular necrosis characterized by loss of tubular epithelial cells. (d) Representative TUNEL staining images are shown in different groups. (e) The level of IL-1β, IL-6, and TNF-α was detected via ELISA. (f) The level of apoptosis-associated proteins (Bcl2, Bax, and Cyt-c) was measured via western blot. $^{*}P < 0.05$ and $^{**}P < 0.01$ vs. Sham group, $^{\#}P < 0.05$ vs. I/R group, and $^{\&}P < 0.05$ vs. I/R+circHIPK3 group. The experiments were repeated at least three times. The results were expressed as mean ± SD.

CREA, BUN, and UA were significantly upregulated in the I/R group, which was recovered by si-circHIPK3 infection (Figure 5(b)). AAV9-si-circHIPK3 injection mice showed improved tubular dilatation and necrosis compared with the AAV9-si-NC group after I/R (Figure 5(c)). Knockdown of circHIPK3 inhibited I/R-induced apoptosis in vivo, which was indicated by TUNEL staining and western blot assay (Figures 5(d) and 5(f)). AAV9-si-circHIPK3 injection decreased the levels of proinflammatory factors in the I/R group (Figure 5(e)). Together, silencing of circHIPK3 could alleviate AKI via the miR-93-5p/KLF9 signal pathway.

4. Discussion

When severe ischemia reperfusion injury occurs in the kidney, it will not only cause glomerular injury and renal failure but also cause chronic renal tubulointerstitial fibrosis [29, 30]. With the long-term development of fibrosis, renal injury caused by ischemia reperfusion can turn into chronic kidney disease, even the end of life. Appropriate animal models can simulate the dynamic changes in the development of the disease and may reveal the molecular regulatory mechanisms in patients [3, 4, 31]. Here, we demonstrated that circHIPK3

was involved in renal AKI by controlling apoptosis, oxidative stress, and inflammation via the miR-93-5p/KLF9 axis.

The role of circRNAs in disease has received increasing attention. Although the role of circRNAs in certain kidney diseases has been reported, such as lupus nephritis and hypertensive nephropathy [32, 33], their function in renal AKI is unclear. circHIPK3 plays an important role in the physiological and pathological processes of multiple organs, especially organ ischemia/reperfusion injury. Wu et al. [34] demonstrated that circHIPK3 was abnormally elevated in the myocardial infarction mouse heart. Interference the circHIPK3 expression could reduce infarction size and apoptosis. The author also revealed that circHIPK3 could regulate the Rac1/PI3K/AKT signal pathway by sponging miR-93-5p. Cheng et al. [35] showed that circHIPK3 could control AKT3 via sponging miR-29b in myocardial ischemia/reperfusion injury. Furthermore, Qiu et al. [25] disclosed that circHIPK3 could promote autophagy and apoptosis in H/R-induced cardiomyocytes by regulating ATF7 via sponging miR-20-5b. Our data showed that circHIPK3 was upregulated in I/R- and H/R-induced renal AKI. Silencing of circHIPK3 could inhibit apoptosis, oxidative stress, and inflammation *in vivo* and *in vitro*, and circHIPK3 knockdown also downregulated the ROS level and recovered the ATP content. These data suggested that circHIPK3 contributed to the oxidation stress, inflammation, and apoptosis during I/R-induced AKI *in vivo* and *in vitro*.

In previous research, miR-93-5p participated in AKI progression by regulating the downstream target. circFANCA controlled LPS-induced cell injury by sponging miR-93-5p. Furthermore, miR-93-5p overexpression inhibited LPS-treated HK-2 cell injury by targeting OXSR1. circFANCA triggered OXSR1 expression by binding with miR-93-5p [36]. miR-93-5p was also a downstream target of lncRNA NEAT1, and inhibition of NEAT1 blocked LPS-induced damage by inducing miR-93-5p expression in HK-2 cells [37]. Juan et al. [38] also exhibited that exosome-miR-93-5p is involved in sepsis-induced AKI via the influence of pyroptosis in renal epithelial cells. Here, the bioinformatics website predicted that miR-93-5p could interact with circHIPK3. The luciferase assay showed that circHIPK3 could bind with miR-93-5p. The miR-93-5p mimic blocked the proapoptosis and inflammation of circHIPK3 on HK-2 cells, which suggested that miR-93-5p was a downstream target of circHIPK3.

KLF9 is a transcription factor that regulates oxidative stress. It is involved in a wide range of neural development, cell differentiation, proliferation, and apoptosis [39–41]. As a target of miR-218-5p, KLF9 controlled the autophagy, apoptosis, and oxidative stress by regulating the JAK2/STAT3 signaling pathway in rheumatoid arthritis synovial fibroblasts [39]. Furthermore, KLF9 induced mitochondrial dysfunction and ROS accumulation in high glucose-induced SH-SY5Y cells via bupivacaine neurotoxicity [42]. Yan et al. [40] also demonstrated that KLF9 could sharpen the oxidative stress and ROS level in ischemia injury in cardiomyocytes. In our research, we found that KLF9 was a downstream target of miR-93-5p, which was involved in circHIPK3-regulating AKI progression.

5. Conclusion

Our research indicated that the circHIPK3 was upregulated in I/R-induced mice and H/R-treated HK-2 cells. Silencing of circHIPK3 improved cell apoptosis, oxidative stress, and inflammation in H/R-treated HK-2 cells. Furthermore, circHIPK3 initiated the reduction of miR-93-5p expression and sponged to miR-93-5p. miR-93-5p overexpression could block the injury induced by circHIPK3 in I/R-treated HK-2 cells. As the downstream target of miR-93-5p, KLF9 was involved in circHIPK3/miR-93-5p-regulating AKI progression. At present, we have exposed the function of circHIPK3 knockdown *in vivo*. To further explore and verify the downstream signal pathway, we gained or lost the function of miR-93-5p and KLF9 *in vivo*. Furthermore, we would collect the plasma of clinical renal failure patients to analyze the potential of circHIPK3 as a biomarker in AKI, which provided an underlying diagnosis and treatment target for AKI.

Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

Conflicts of Interest

The authors declare that they have no conflict of interests.

Authors' Contributions

Pan Youmin worked on the experimental design. Zha Zhengbiao, Chen Liang, and Zheng Zhi wrote this article and conducted the experiments, analyzed the data, and interpreted the data. All authors read and approved the final manuscript.

References

[1] S. Aiello, M. A. Podestà, P. Y. Rodriguez-Ordonez et al., "Transplantation-induced ischemia-reperfusion injury modulates antigen presentation by donor renal $CD11c^{+}F4/80^{+}$ macrophages through IL-1R8 regulation," *Journal of the American Society of Nephrology*, vol. 31, no. 3, pp. 517–531, 2020.

[2] J. J. Yan, J. H. Ryu, H. Piao et al., "Granulocyte colony-stimulating factor attenuates renal ischemia-reperfusion injury by inducing myeloid-derived suppressor cells," *Journal of the American Society of Nephrology*, vol. 31, no. 4, pp. 731–746, 2020.

[3] E. F. Carney, "Ferroptotic stress promotes the AKI to CKD transition," *Nature Reviews Nephrology*, vol. 17, no. 10, p. 633, 2021.

[4] E. F. Carney, "The molecular genetics of AKI," *Nature Reviews Nephrology*, vol. 17, no. 1, p. 14, 2021.

[5] X. Sun, Z. Wei, Y. Li et al., "Renal denervation restrains the inflammatory response in myocardial ischemia-reperfusion injury," *Basic Research in Cardiology*, vol. 115, no. 2, p. 15, 2020.

[6] G. J. Goodall and V. O. Wickramasinghe, "RNA in cancer," *Nature Reviews Cancer*, vol. 21, no. 1, pp. 22–36, 2021.

[7] E. Zlotorynski, "The innate function of circular RNAs," *Nature Reviews Molecular Cell Biology*, vol. 20, no. 7, p. 387, 2019.

[8] Z. Zhao, W. Sun, Z. Guo, J. Zhang, H. Yu, and B. Liu, "Mechanisms of lncRNA/microRNA interactions in angiogenesis," *Life Sciences*, vol. 254, article 116900, 2020.

[9] Y. Shi, C. F. Sun, W. H. Ge, Y. P. du, and N. B. Hu, "Circular RNA VMA21 ameliorates sepsis-associated acute kidney injury by regulating miR-9-3p/SMG1/inflammation axis and oxidative stress," *Journal of Cellular and Molecular Medicine*, vol. 24, no. 19, pp. 11397–11408, 2020.

[10] Y. Cao, X. Mi, D. Zhang, Z. Wang, Y. Zuo, and W. Tang, "Transcriptome sequencing of circular RNA reveals a novel circular RNA-has_circ_0114427 in the regulation of inflammation in acute kidney injury," *Clinical Science*, vol. 134, no. 2, pp. 139–154, 2020.

[11] T. Huang, Y. Cao, H. Wang et al., "Circular RNA YAP1 acts as the sponge of microRNA-21-5p to secure HK-2 cells from ischaemia/reperfusion-induced injury," *Journal of Cellular and Molecular Medicine*, vol. 24, no. 8, pp. 4707–4715, 2020.

[12] Q. Zheng, C. Bao, W. Guo et al., "Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs," *Nature Communications*, vol. 7, no. 1, p. 11215, 2016.

[13] L. Qi, B. Sun, B. Yang, and S. Lu, "circHIPK3 (hsa_circ_0000284) promotes proliferation, migration and invasion of breast cancer cells via miR-326," *Oncotargets and Therapy*, vol. 14, pp. 3671–3685, 2021.

[14] D. Yao, S. Lin, S. Chen, and Z. Wang, "circHIPK3 regulates cell proliferation and migration by sponging microRNA-124 and regulating serine/threonine kinase 3 expression in esophageal squamous cell carcinoma," *Bioengineered*, vol. 13, no. 4, pp. 9767–9780, 2022.

[15] K. Shan, C. Liu, B. H. Liu et al., "Circular noncoding RNA HIPK3 mediates retinal vascular dysfunction in diabetes mellitus," *Circulation*, vol. 136, no. 17, pp. 1629–1642, 2017.

[16] G. Chen, Y. Shi, M. Liu, and J. Sun, "circHIPK3 regulates cell proliferation and migration by sponging miR-124 and regulating AQP3 expression in hepatocellular carcinoma," *Cell Death & Disease*, vol. 9, no. 2, p. 175, 2018.

[17] Y. Wu, J. Luan, C. Jiao et al., "circHIPK3 exacerbates folic acid-induced renal tubulointerstitial fibrosis by sponging miR-30a," *Frontiers in Physiology*, vol. 12, article 715567, 2021.

[18] K. Zeng, X. Chen, M. Xu et al., "circHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7," *Cell Death & Disease*, vol. 9, no. 4, p. 417, 2018.

[19] J. X. Zhang, J. Lu, H. Xie et al., "circHIPK3 regulates lung fibroblast-to-myofibroblast transition by functioning as a competing endogenous RNA," *Cell Death & Disease*, vol. 10, no. 3, p. 182, 2019.

[20] X. Yin, W. Zheng, L. He, S. Mu, Y. Shen, and J. Wang, "circHIPK3 alleviates inflammatory response and neuronal apoptosis via regulating miR-382-5p/DUSP1 axis in spinal cord injury," *Transplant Immunology*, vol. 73, article 101612, 2022.

[21] S. Tang, X. Xie, M. Wang, L. Yang, and W. Wei, "Protective effects of asiaticoside on renal ischemia reperfusion injury in vivo and in vitro," *Bioengineered*, vol. 13, no. 4, pp. 10235–10243, 2022.

[22] J. W. Dutton, J. E. Artwohl, X. Huang, and J. D. Fortman, "Assessment of pain associated with the injection of sodium pentobarbital in laboratory mice (*Mus musculus*)," *Journal of the American Association for Laboratory Animal Science*, vol. 58, no. 3, pp. 373–379, 2019.

[23] X. Zhao, D. Wang, S. Wan, X. Liu, W. Wang, and L. Wang, "The suppression of Pin1-alleviated oxidative stress through the p38 MAPK pathway in ischemia- and reperfusion-induced acute kidney injury," *Oxidative Medicine and Cellular Longevity*, vol. 2021, Article ID 1313847, 15 pages, 2021.

[24] Y. Deng, J. Wang, G. Xie, X. Zeng, and H. Li, "circ-HIPK3 strengthens the effects of adrenaline in heart failure by miR-17-3p - ADCY6 axis," *International Journal of Biological Sciences*, vol. 15, no. 11, pp. 2484–2496, 2019.

[25] Z. Qiu, Y. Wang, W. Liu et al., "circHIPK3 regulates the autophagy and apoptosis of hypoxia/reoxygenation- stimulated cardiomyocytes via the miR-20b-5p/ATG7 axis," *Cell Death Discovery*, vol. 7, no. 1, p. 64, 2021.

[26] Y. Wang, R. Zhao, C. Shen et al., "Exosomal circHIPK3 released from hypoxia-induced cardiomyocytes regulates cardiac angiogenesis after myocardial infarction," *Oxidative Medicine and Cellular Longevity*, vol. 2020, Article ID 8418407, 19 pages, 2020.

[27] B. Yan, Y. Zhang, C. Liang et al., "Stem cell-derived exosomes prevent pyroptosis and repair ischemic muscle injury through a novel exosome/circHIPK3/ FOXO3a pathway," *Theranostics*, vol. 10, no. 15, pp. 6728–6742, 2020.

[28] D. W. Thomson and M. E. Dinger, "Endogenous microRNA sponges: evidence and controversy," *Nature Reviews Genetics*, vol. 17, no. 5, pp. 272–283, 2016.

[29] J. A. Kellum, P. Romagnani, G. Ashuntantang, C. Ronco, A. Zarbock, and H. J. Anders, "Acute kidney injury," *Nature Reviews Disease Primers*, vol. 7, no. 1, p. 52, 2021.

[30] H. Scholz, F. J. Boivin, K. M. Schmidt-Ott et al., "Kidney physiology and susceptibility to acute kidney injury: implications for renoprotection," *Nature Reviews Nephrology*, vol. 17, no. 5, pp. 335–349, 2021.

[31] C. Tang, J. Cai, X. M. Yin, J. M. Weinberg, M. A. Venkatachalam, and Z. Dong, "Mitochondrial quality control in kidney injury and repair," *Nature Reviews Nephrology*, vol. 17, no. 5, pp. 299–318, 2021.

[32] J. Jin, H. Sun, C. Shi et al., "Circular RNA in renal diseases," *Journal of Cellular and Molecular Medicine*, vol. 24, no. 12, pp. 6523–6533, 2020.

[33] H. Ren and Q. Wang, "Non-coding RNA and diabetic kidney disease," *DNA and Cell Biology*, vol. 40, no. 4, pp. 553–567, 2021.

[34] Y. Wu, M. Wu, J. Yang et al., "Silencing circHIPK3 sponges miR-93-5p to inhibit the activation of Rac1/PI3K/AKT pathway and improves myocardial infarction-induced cardiac dysfunction," *Frontiers in Cardiovascular Medicine*, vol. 8, article 645378, 2021.

[35] N. Cheng, C. Jin, P. Jin, D. Zhu, and Z. Hou, "High glucose protects cardiomyocytes against ischaemia/reperfusion injury by suppressing myocardiocyte apoptosis via circHIPK3/miR-29b/AKT3 signalling," *Journal of Cellular and Molecular Medicine*, vol. 25, no. 13, pp. 6137–6147, 2021.

[36] H. Li, X. Zhang, P. Wang, X. Zhou, H. Liang, and C. Li, "Knockdown of circ-FANCA alleviates LPS-induced HK2 cell injury via targeting miR-93-5p/OXSR1 axis in septic acute kidney injury," *Diabetology & Metabolic Syndrome*, vol. 13, no. 1, p. 7, 2021.

[37] J. Yang, L. Wu, S. Liu, X. Hu, Q. Wang, and L. Fang, "Long non-coding RNA NEAT1 promotes lipopolysaccharide-induced injury in human tubule epithelial cells by regulating

miR-93-5p/TXNIP axis," *Medical Microbiology and Immunology*, vol. 210, no. 2-3, pp. 121–132, 2021.

[38] C. X. Juan, Y. Mao, Q. Cao et al., "Exosome-mediated pyroptosis of miR-93-TXNIP-NLRP3 leads to functional difference between M1 and M2 macrophages in sepsis-induced acute kidney injury," *Journal of Cellular and Molecular Medicine*, vol. 25, no. 10, pp. 4786–4799, 2021.

[39] M. Chen, M. Li, N. Zhang, W. Sun, H. Wang, and W. Wei, "Mechanism of miR-218-5p in autophagy, apoptosis and oxidative stress in rheumatoid arthritis synovial fibroblasts is mediated by KLF9 and JAK/STAT3 pathways," *Journal of Investigative Medicine*, vol. 69, no. 4, pp. 824–832, 2021.

[40] Q. Yan, B. He, G. Hao et al., "KLF9 aggravates ischemic injury in cardiomyocytes through augmenting oxidative stress," *Life Sciences*, vol. 233, article 116641, 2019.

[41] S. N. Zucker, E. E. Fink, A. Bagati et al., "Nrf2 amplifies oxidative stress via induction of Klf9," *Molecular Cell*, vol. 53, no. 6, pp. 916–928, 2014.

[42] H. Li, Y. Weng, L. Lai et al., "KLF9 regulates PRDX6 expression in hyperglycemia-aggravated bupivacaine neurotoxicity," *Molecular and Cellular Biochemistry*, vol. 476, no. 5, pp. 2125–2134, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 4164232, 23 pages
https://doi.org/10.1155/2023/4164232

Research Article

Pain-Related Gene Solute Carrier Family 24 Member 3 Is a Prognostic Biomarker and Correlated with Immune Infiltrates in Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma: A Study via Integrated Bioinformatics Analyses and Experimental Verification

Shuguang Zhou,[1,2,3] **Qinqin Jin**,[1,2] **Hui Yao**,[1,2] **Jie Ying**,[1,2] **Lu Tian**,[1,2] **Xiya Jiang**,[1,2] **Yinting Yang**,[1,2] **Xiaomin Jiang**,[1,2] **Wei Gao**,[1,2] **Weiyu Zhang**,[1,2] **Yuting Zhu**,[1,2] **and Wujun Cao**[4]

[1]*Department of Gynecology, Anhui Province Maternity and Child Healthcare Hospital, Hefei, Anhui 230001, China*
[2]*Department of Gynecology, Anhui Medical University Affiliated Maternity and Child Healthcare Hospital, Hefei, Anhui 230001, China*
[3]*Department of Gynecology, Linquan Maternity and Child Healthcare Hospital, Fuyang, Anhui 236400, China*
[4]*Department of Clinical Laboratory, Anhui Province Maternity and Child Healthcare Hospital, Hefei, Anhui 230001, China*

Correspondence should be addressed to Shuguang Zhou; zhoushuguang@ahmu.edu.cn

Received 23 September 2022; Revised 13 October 2022; Accepted 24 November 2022; Published 7 February 2023

Academic Editor: Jincheng Wang

The aim of this study was to explore cervical carcinoma and screen a suitable gene as the biomarker used for prognosis evaluation as well as pain therapy. Low expression levels of solute carrier family 24 member 3 (SLC24A3) was involved in the appearance and development of numerous malignancies. Nevertheless, the prognostic value of SLC24A3 expression with cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) patients remains uncertain. During the present study, SLC24A3 expression in CESC was retrieved from TCGA, GEO, and MSigDB databases. Based on TCGA and GEO profiles, we performed survival and difference analyses about SLC24A3 both in two GEO (GSE44001 and GSE63514) and TCGA-CESC cohorts (all $p < 0.05$), indicating that SLC24A3 was low expressed in tumors and associated with higher overall survival in CESC patients. Additionally, we programmed a series of analyses, including genomic profiling, enrichment analysis, immune infiltration analysis, and therapy-related analysis to identify the mechanism of the SLC24A3 in the process of cancer in CESC. Meanwhile, qRT-PCR was used to validate that the expression of SLC24A3 mRNA in Hela and SiHa cell lines was significantly lower than in PANC-1 and HUCEC cell lines. Our finding elucidated that the SLC24A3, a sodium-calcium regulator of cells, is an indispensable factor which can significantly influence the prognosis of patients with CESC and could provide novel clinical evidence to serve as a potential biological indicator for future diagnosis and pain therapy.

1. Introduction

Cervical carcinoma is the principal cause of mortality in female, specifically in some developing countries, and they are the second most universal cancer type in gynecology [1–5]. According to statistics, over half a million women are diagnosed with cervical carcinoma and the disease results in more than 300,000 deaths worldwide each year [6, 7]. Despite the increase in the incidence of cervical adenocarcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) is still the most common type of cervical carcinoma. In relatively impoverished

countries which lack formal screening programs and effective early diagnosis, CESC incidence and mortality are extremely high [8]. Recently, despite various studies have suggested that abnormally expressed tumor markers may be related to cancer occurrence and development tremendous endeavors to explore novel biomarkers, CESC continues to be a severe health problem among females. Hence, it is a matter of urgency to search novel prognostic biomarkers for female with CESC in order to improve the prognosis and reduce the mortality rate in CESC and to develop effective treatment indications for patients.

Cancer pain, simultaneously, could be the most common and unbearable symptoms in cancer patients, which has been shown to be associated with depression, anxiety, and reduced quality of life in cancer survivors of varied diagnoses [9]. Currently, the occurrence of cancer has been confirmed to be related to oncogenes and accompanied by pain in the body [10, 11]. Pain can be a common and uncontrollable symptom for tumor patients. It seriously affects the physical and psychological functions of patients [12, 13]. Although effective pain relief can be achieved in up to 90% of patients with cancer, numerous studies have shown that pain remains inadequately controlled in many patients [14, 15]. Thus, the study of pain-related genes in cancer patients can provide a basis for the development of targeted drugs for tumor therapy, which is of great significance to decrease pain in cancer patients and enhance their quality of life. CESC pain continues to be a major clinical challenge. Despite decades of thorough study and continuous efforts, the underlined cellular and molecular mechanisms remain elusive and the clinical approaches for treating CESC and cancer-related pain are limited. Therefore, this study attempts to explore new biomarkers that not only have a significant function in the prognosis and treatment of CESC patients but also have the potential to effectively control cancer pain.

2. Materials and Methods

2.1. Data Collection and Description. The Cancer Genome Atlas (TCGA) database (http://cancergenome.nih.gov/abouttcga/policies/publicationguidelines) was applied to download gene expression quantification data and clinical information of female with CESC. Nevertheless, because the TCGA database lacks normal tissue data matched with cervical cancer, the total number of samples is only three. Hence, the Genotype-Tissue Expression (GTEx) Portal (http://xena.ucsc.edu/) was applied to obtain the expression values of the normal cervix tissue. The Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) was applied to download gene expression and clinical data of patients in GSE44001 and GSE63514 cohorts. The Molecular Signatures Database (*MSigDB*) (http://software.broadinstitute.org/gsea/msigdb/index.jsp) was used to extract pain-related genes. Since the data were downloaded from public databases, there is no need to get approval from ethics committee.

2.2. Identification of SLC24A3-Related Genes. RNA-seq data information of 306 tumor cases and 13 normal cases was obtained from the TCGA and GTEx official website. Differential expression genes (DEGs) were screened via "edgeR" R package by comparing SLC24A3 expression of the tumor and normal cases. We chose the p value < 0.05 and |log FoldChange| (|logFC|) > 1 as the cut-off criteria. We calculated the Spearman coefficients of the DEGs and key pain candidate genes, while defining DEGs with p value < 0.05 as key pain candidate genes.

2.3. Differential Methylation Analysis. 306 tumor and 13 normal cases were chosen to perform differential methylation analysis. Probes containing single-nucleotide polymorphism (SNPs), probes with over 10% missing values, and probes in chromosome X were excluded from the analysis. And the "imputeTS" R package was applied to impute the missing values of the selected central pattern generators (CPGs). Differentially methylated CpGs (DMCs) were determined via the Wilcoxon rank-sum test, and the p values were adjusted via the false discovery rate (FDR) approach. If the mean methylation difference was >0.2 with an FDR of 5%, DMCs were reported.

2.4. Inference of Infiltrating Cells in CESC. The xCell algorithm, an R package (version 1.1.0), was applied to quantify the abundance of tumor stromal cells with CESC patients, which evaluates the integrated levels of 64 stromal cell types. xCell, a gene signature-based approach, combines the advantages of gene set enrichment with deconvolution methods, which well fits with RNA-seq and microarray data. And the gene expression data were prepared according to the xCell instructions, run R package, and undertaken using the xCell signature ($N = 64$) with 1000 permutations.

2.5. Drug Sensitivity Analysis. The Genomics of Drug Sensitivity in Cancer (GDSC) [16] database (https://www.cancerrxgene.org/) was used to download transcription data for almost 1000 cancer cell lines, drug response measurements as area under the curve (AUC) for antitumor drugs in cancer cell lines, and targets/pathways of drugs. The Spearman correlation analysis was performed to calculate the association between drug sensitivity and risk score. Then, we regarded |Rs| > 0.2 and applied Benjamini and Hochberg to adjust FDR and regarded an FDR < 0.05 as significant correlation. "pRRophetic" R package [17] was used to calculate the half-maximum inhibitory concentration (IC_{50}) for comparing the drug sensitivity.

2.6. TMB and IPS Analysis. The "mafools" R package [18] was applied to calculate tumor mutation burden (TMB) from datasets of somatic mutations. The immunophenoscore (IPS) from The Cancer Immunome Atlas is a machine learning-based immune response score. The higher IPS score is associated with better immune response [19].

2.7. Cell Lines. Hela and SiHa are human cervical cancer cell lines which were used as the test groups. Among them, Hela was a cell line of cervical adenocarcinoma and SiHa was a cell line of cervical squamous cell carcinoma. We utilized PANC-1, a cell line of pancreatic cancer, as a positive control for SLC24A3 expression. Meanwhile, we utilized HUCEC, a cell line of normal cervical, as a negative control for SLC24A3 expression. All of them were obtained from the

American Type Culture Collection (ATCC, Manassas, VA, United States). We used DMEM contained with 10% FBS (Gibco, Grand Island, NY) to incubate Hela and SiHa cells. Then, we used RPMI 1640 supplementing 10% FBS (Gibco, Grand Island, NY) to incubate PANC-1 and HUCEC cells. The cells are cultured in an incubator at 37°C with 5% CO2.

2.8. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR). We used TRIzol reagent (Invitrogen, Carlsbad, CA) to collect and lyse cells. Then, M-MLV reverse transcriptase (Promega, Madison, WA) was utilized to obtain cDNA. Utilizing a LightCycler 480 II RT-PCR System (Roche, Basel, Switzerland), qRT-PCR was carried out as follows: 95°C for half a minute and then 40 cycles of 95°C for 5 seconds and 60°C for half a minute. In our study, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) gene was utilized as an internal control; sequences of primers amplified in this study were contained in Table S3.

2.9. Data Processing and Analysis. In this study, R programming language was the major tool for processing data. The DEGs between CESC samples and healthy controls were acquired via the "EdgeR" R package, whereas $|\log FC| > 2$ was screened as differentially expressed pain-related genes. The box plot and the forest plot were conducted by the "ggpubr" and "forestplot" R packages, respectively. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were done via the "enrichplot" R package. Survival curves and survival analysis were conducted by the "survminer" as well as "survival" R packages. "Clusterprofiler" R package was applied for quantification of gene set activity on the gene set enrichment analysis (GSEA). The version of R bioconductor utilized in this study was R 4.1.2. The qualitative data were analyzed by t-test via SPSS 26.0.

3. Results

3.1. Screening for Pain-Related Genes in Cervical Cancer. The flow chart of this study is presented in Figure 1. We performed difference analysis and survival analysis according to the CESC cohort from the TCGA database (combine with GTEx), including 306 tumor cases with survival data, RNA-seq data, and clinicopathological data. Then, we processed and analyzed 2 cervical squamous cell carcinoma datasets (GSE44001 and GSE63514) from the GEO platform, identifying differential expression genes and survival data between cancer and normal tissues, respectively. By integrating survival data from the TCGA and GEO cohorts, a total of 5 (DIAPH3, SLC24A3, NUP62CL, C3orf70, and CD177) DEGs ($p < 0.05$) were screened to perform the further GSEA analysis.

There were 117 pain-related genes acquired from *MSigDB* (Table S1). According to the results of GSEA analysis, there was only one gene enriched in pain-related pathway, which is SLC24A3 (Figure 2(a)). As illustrated in box plot, SLC24A3 expression in tumor groups was all considerably lower than nontumor groups in TCGA-CESC and GSE63514 cohorts (all $p < 0.05$; Figures 2(b) and 2(c)). The results demonstrated that downregulation of SLC24A3 in tumor tissues was greatly associated with better overall survival (log rank $p = 0.0019$; Figure 2(d)), disease-free survival (log rank $p = 0.049$; Figure 2(e)), and progression-free survival (log rank $p = 0.049$; Figure 2(f)) among patients with CESC.

Additionally, correlation analysis indicated that the SLC24A3-expressed levels were positively related to the SLC6A4-expressed level both in TCGA and GSE63514 cohorts (all $p < 0.05$; Figures 2(g) and 2(h)). Among them, SLC6A4 is a classical pain-related gene. In the same vein, we performed expression levels and survival analyses with SLC6A4 (Figure S1 A-E), which outlined similar results with SLC24A3. Finally, we applied the Human Protein Atlas (HPA) database to validate the SLC24A3 expression in pathological and healthy cervical tissues, finding that there was an obviously differential expression of SLC24A3 in pathological cervical tissues and healthy cervical tissues (Figure S1 F-H).

3.2. Mutation Characteristics of SLC24A3 Genome. Mutation panorama of SLC24A3 showed that somatic mutation rate was 2.08% and accounted for 5 sites in SLC24A3 (Figure 3(a) and Figure S2 A). However, SLC24A3 with copy number variations (CNV) of these amplifications had lower survival probability than those CNV deletion and without CNV mutations ($p < 0.0001$; Figure 3(b)), suggesting that genetic alteration of CNV amplification mutation might play certain functional role in CESC. Then, we used R project to perform differential methylation analysis in both CESC and normal cervix tissues, finding that hypermethylated SLC24A3 was adversely associated with SLC24A3 expression (Figure 3(c)) and hypermethylated SLC24A3 considerably contributed to a well overall survival probability in CESC patients ($p = 0.0097$; Figure 3(d)). Box plot showed the significantly correlation between methylation status of SLC24A3 and CNV deletion ($p = 0.0044$; Figure 3(e)). Moreover, to ascertain how CNV impacted the expression of SLC24A3 among CESC patients, we compared the SLC24A3-expressed levels between CNV amplification mutations, deletion mutations, and normal samples and showed that the CNV amplification mutations was significantly increased in CESC (Figure 3(f)). By the same token, we performed analyses about survival and expression of SLC24A3 derived from single-nucleotide variant (SNV) mutations, and the results were similar as CNV (Figure S2 B-C). Clinical characteristics of TCGA CESC data between the CNV and SNV groups are shown in Table S2.

Then, we performed the univariate Cox regression analysis using age, grade, stage, TNM status, body mass index (BMI), methylation status, and SLC24A3-expressed levels as inputs; our analysis proved that methylation status and the expression of SLC24A3 were significantly associated with the survival outcomes ($p < 0.05$; Figure 3(g)). Subsequently, we also performed the multivariate Cox regression analysis using stage, methylation status, and SLC24A3-expressed levels as inputs; the results exposed that the SLC24A3-expressed level was a robust and independent predictor of enhanced prognosis (hazard ratio: 0.214-0.725, $p = 1.93e - 03$; Figure 3(h)).

3.3. KEGG/GO/GSEA Biological Process Enrichment. Volcano plot indicated differential SLC24A3-related genes. Notably,

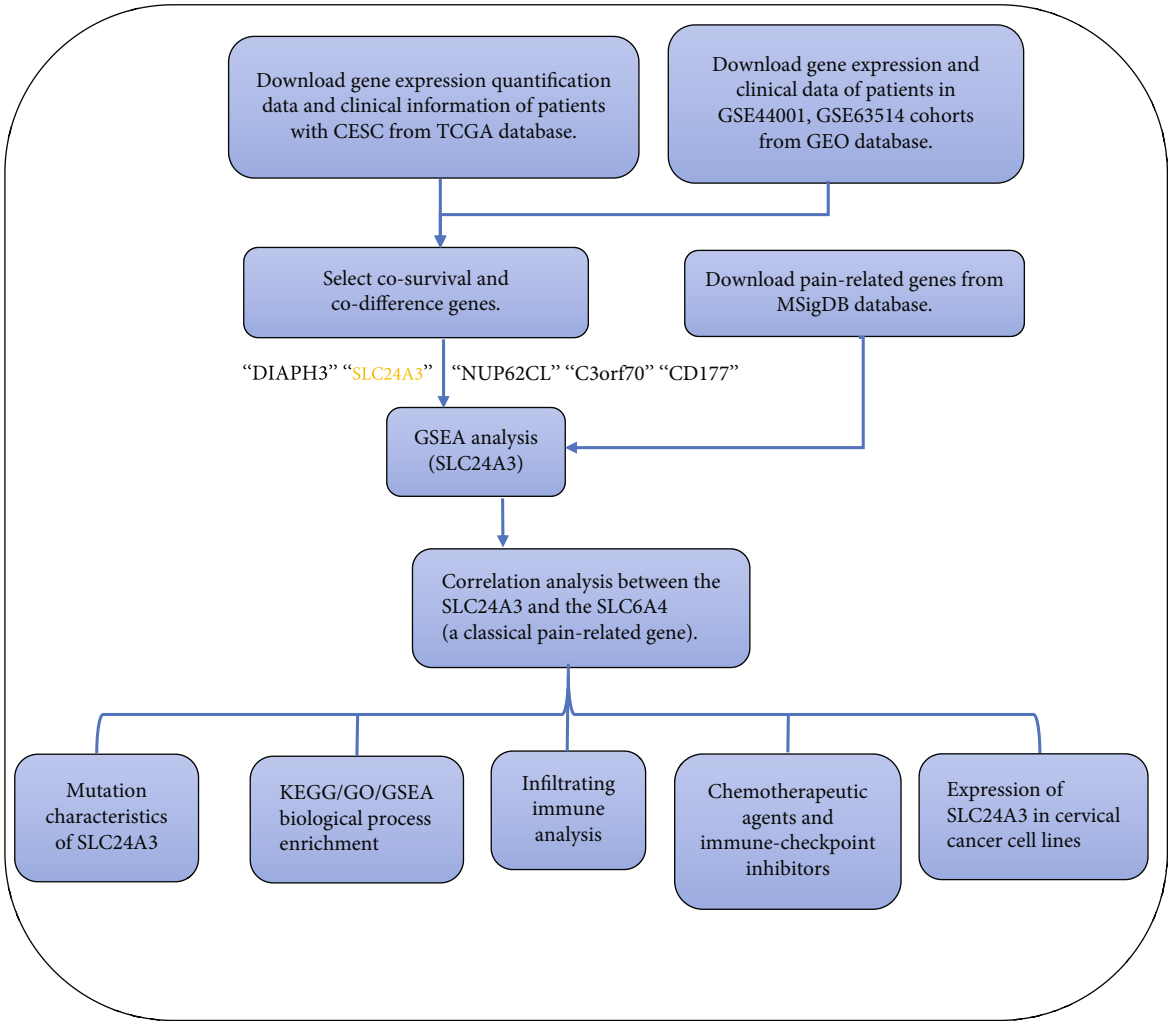
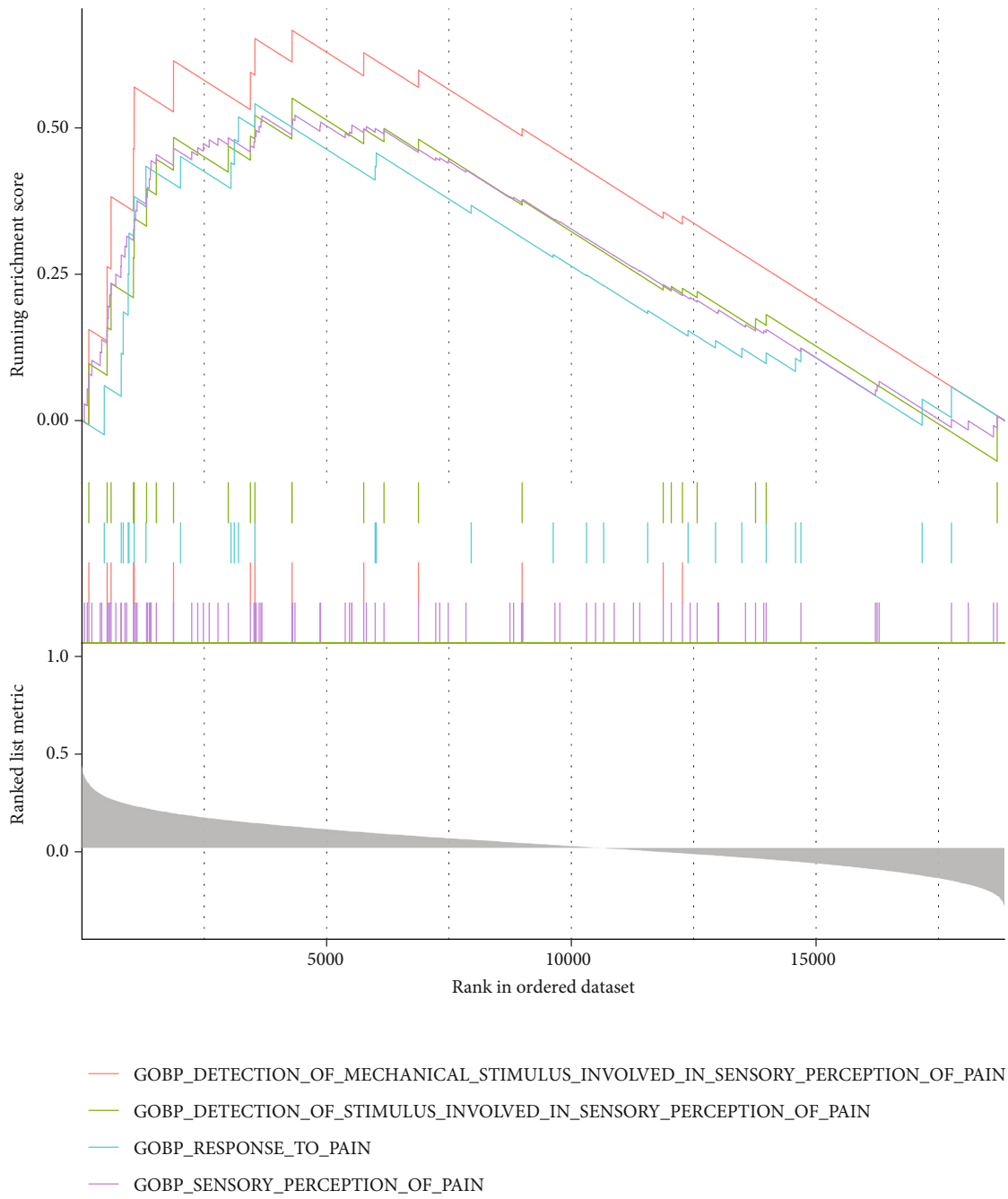


Figure 1: The flow chart of this study.

there were 1,771 DEGs overlapped among the three cohorts, which included 774 upregulated genes and 937 downregulated genes ($|\log FC| > 1$ and $p < 0.05$; Figure 4(a)). The KEGG pathway enrichment of SLC24A3 interactive genes revealed that cysteine and methionine metabolism, glutathione metabolism, glyoxylate and dicarboxylate metabolism, purine metabolism, pyrimidine metabolism, and pyruvate metabolism were the significantly enriched pathways (Figure 4(b)). In addition, GO analysis results indicated that SLC24A3-related genes were mostly enriched in the calcium signaling pathway, ECM-receptor interaction, cAMP signaling pathway, protein digestion and absorption, etc. at biological process (BP) levels; extracellular matrix organization, external encapsulating structure organization, extracellular structure organization etc. at cellular components (CC) levels; and gated channel activity, extracellular matrix structural constituent, glycosaminoglycan binding, antigen binding, etc. at molecular function (MF) levels (Figures 4(c)–4(e)). GSEA identified a set of genes associated with DNA repair, mitochondrion organization, NCRNA metabolic process, and reactome cell cycle mitotic (Figure 4(f)).

3.4. Associations of SLC24A3 with Tumor-Infiltrating Immune Cells and Immune-Related Gene Sets. In our study, the comprehensive levels of 64 immunity and stromal cell types of 274 CESC samples were calculated. Among immune infiltration analysis of 64 immune cells, the xCell algorithm results elucidated that proportion of activated B cell ($p = 1.44e-03$), activated dendritic cell ($p = 1.72e-02$), CD56 bright natural killer cell ($p = 5.62e-03$), central memory CD4 T cell ($p = 4.5e-05$), central memory CD8 T cell ($p = 1.12e-06$), effector memory CD4 T cell ($p = 2.16e-05$), effector memory CD8 T cell ($p = 1.33e-02$), gamma delta T cell ($p = 4.73e-02$), immature B cell ($p = 4.66e-03$), immature dendritic cell ($p = 5.23e-05$), macrophage ($p = 1.03e-03$), mast cell ($p = 2.92e-02$), MDSC ($p = 2.05e-03$), memory B cell ($p = 3.37e-02$), natural killer cell ($p = 2.88e-04$), natural killer T cell ($p = 2.41e-03$), neutrophil ($p = 1.33e-02$), plasmacytoid dendritic cell ($p = 5.08e-04$), regulatory T cell ($p = 1.27e-03$), T follicular helper cell ($p = 4.25e-06$), type 1 T helper cell ($p = 6.77e-05$), type 2 T helper cell ($p = 4.58e-02$), a total of 22 immune cells, was enriched in SL24A3 high-expressed group (Figure 5(a)). This



(a)

Figure 2: Continued.

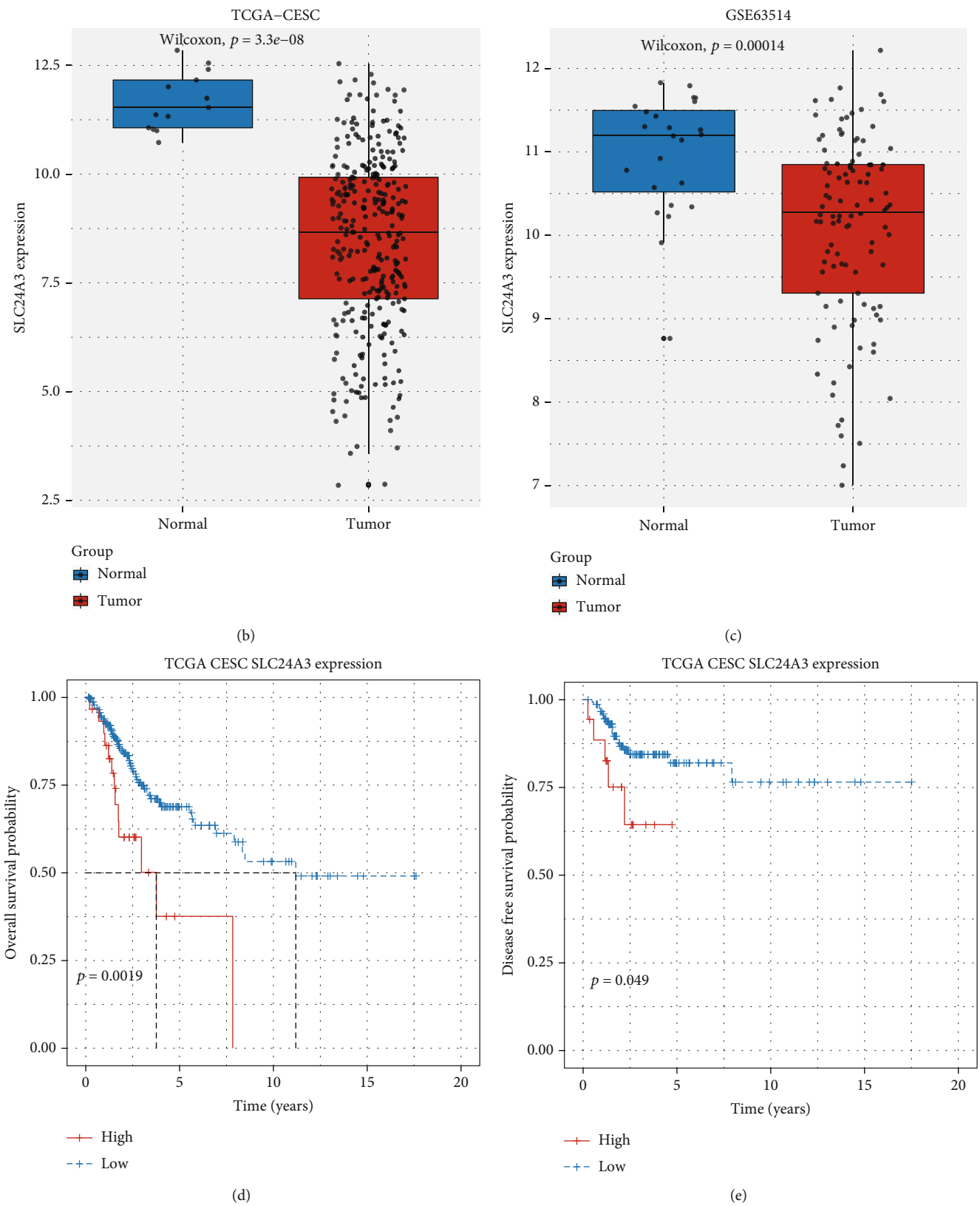


Figure 2: Continued.

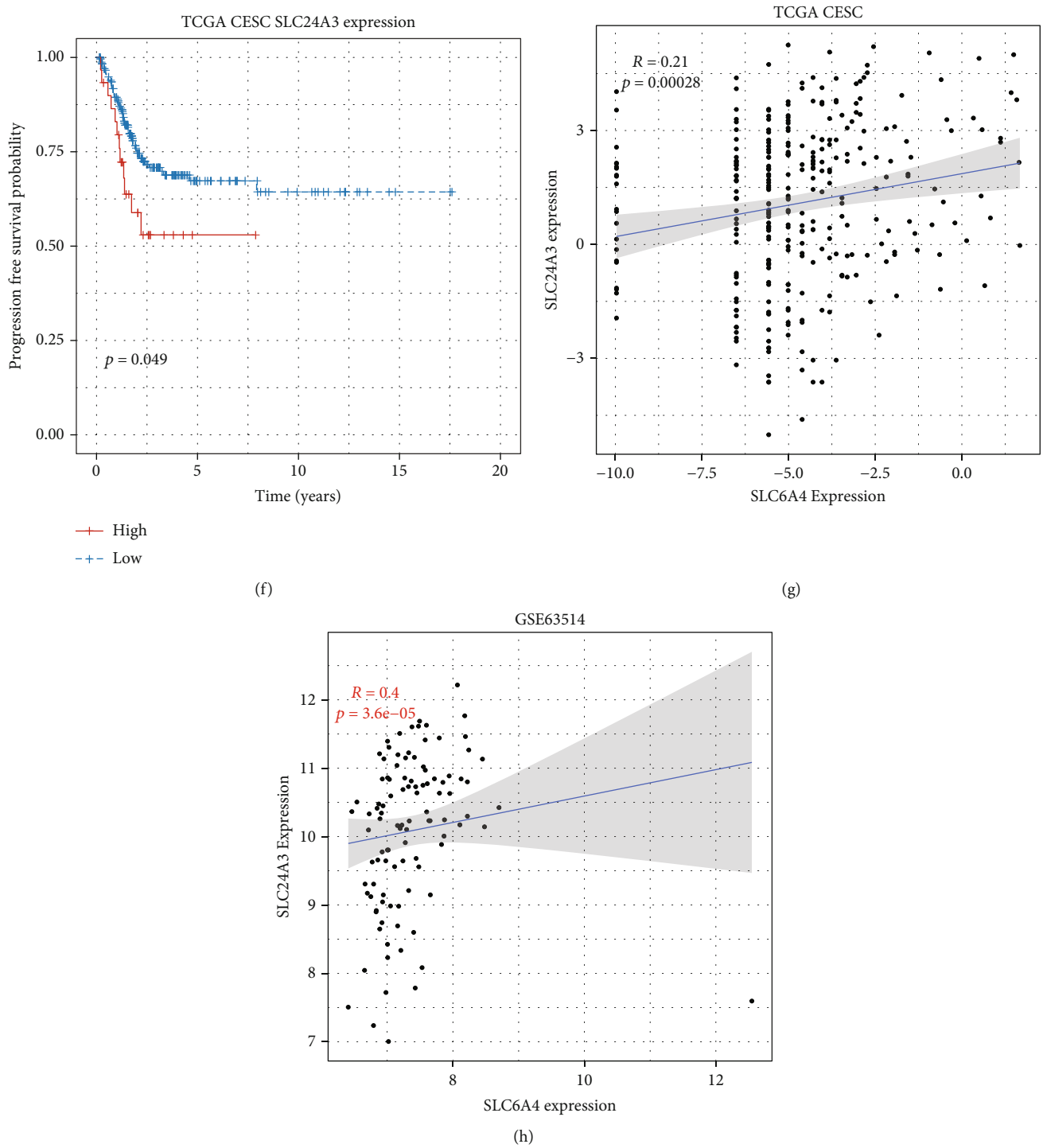


Figure 2: Pain-related gene selection: (a) GSEA for SLC24A3. (b, c) Expression levels of SLC24A3 between tumor and normal cases with CESC patients both in TCGA-CESC and GSE63514 cohorts. (d–f) Overall survival, disease-free survival, and progression survival comparisons between high and low SLC24A3 groups in TCGA database. (g, h) Correlation analysis between SLC24A3 and SLC6A4 in TCGA and GSE63514 cohorts.

evidence suggested a significant correlation between SLC24A3 and tumor-immune infiltration. Additionally, the results of correlation analysis between SLC24A3-expressed levels and antigen-presenting cells gene sets indicated that the expression levels of PSMB5, PSMB6, PSMB7, and PSMB8 were increased in SLC24A3 low-expressed group (Figures 5(b) and 5(e)). Moreover, the expression level of SLC24A3 was positively correlated to epithelial-mesenchymal transition (EMT), which implied that SLC24A3 was a contributor to EMT regulation in CESC (Figure 5(f)).

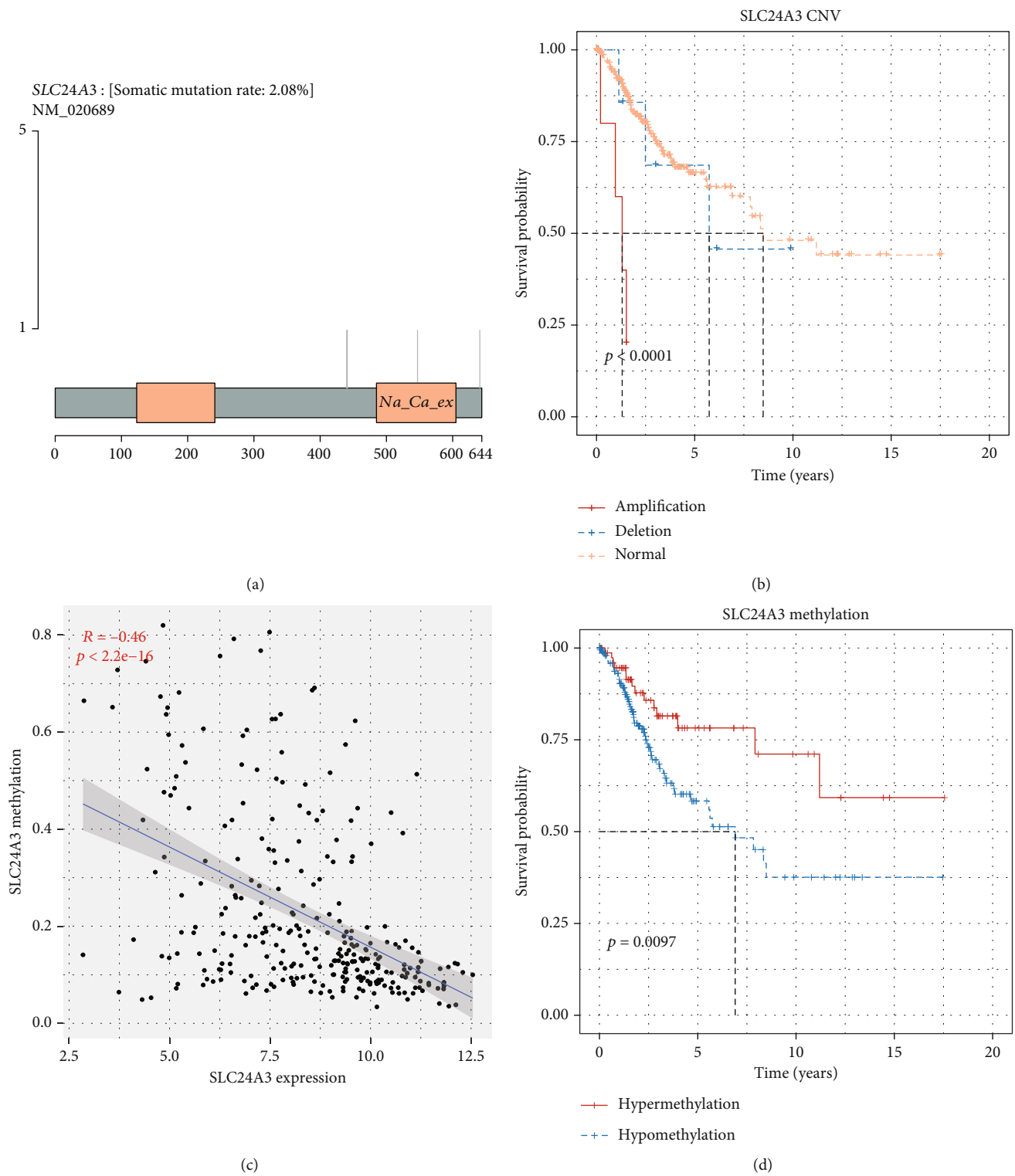
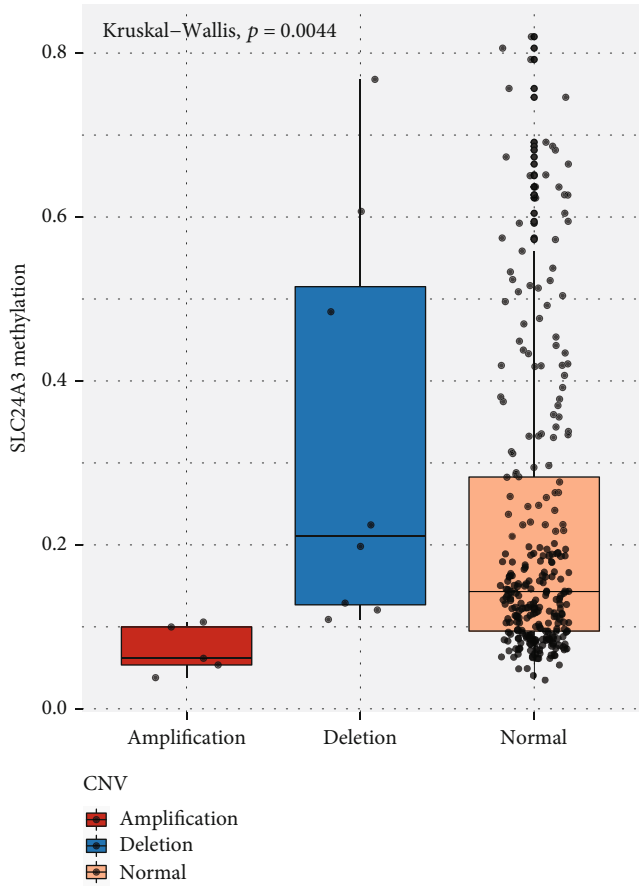
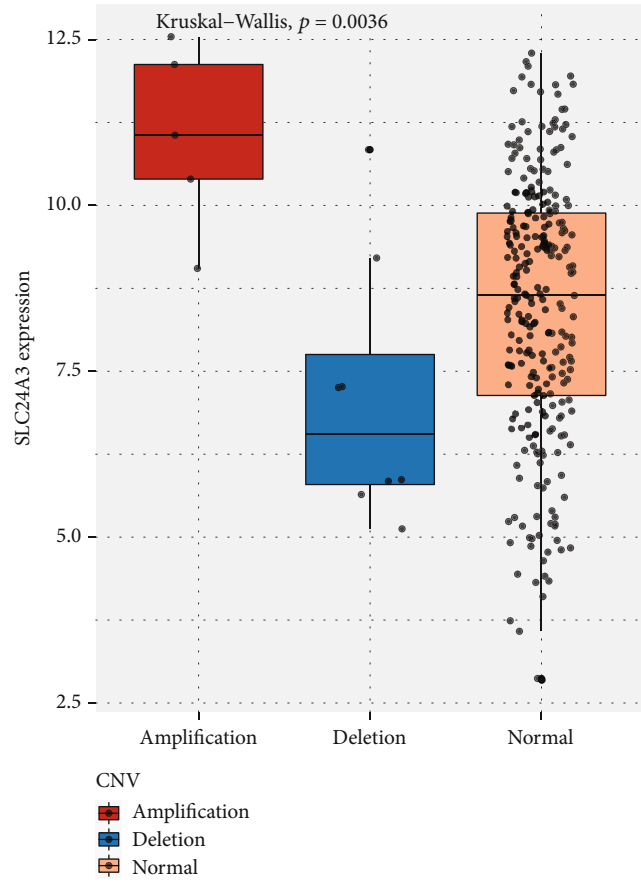


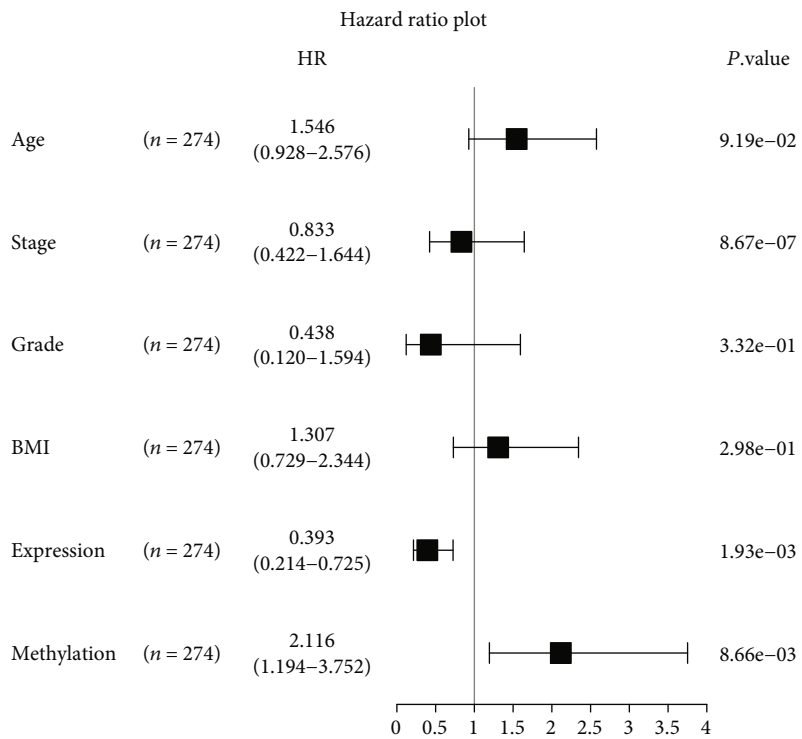
Figure 3: Continued.



(e)



(f)



(g)

Figure 3: Continued.

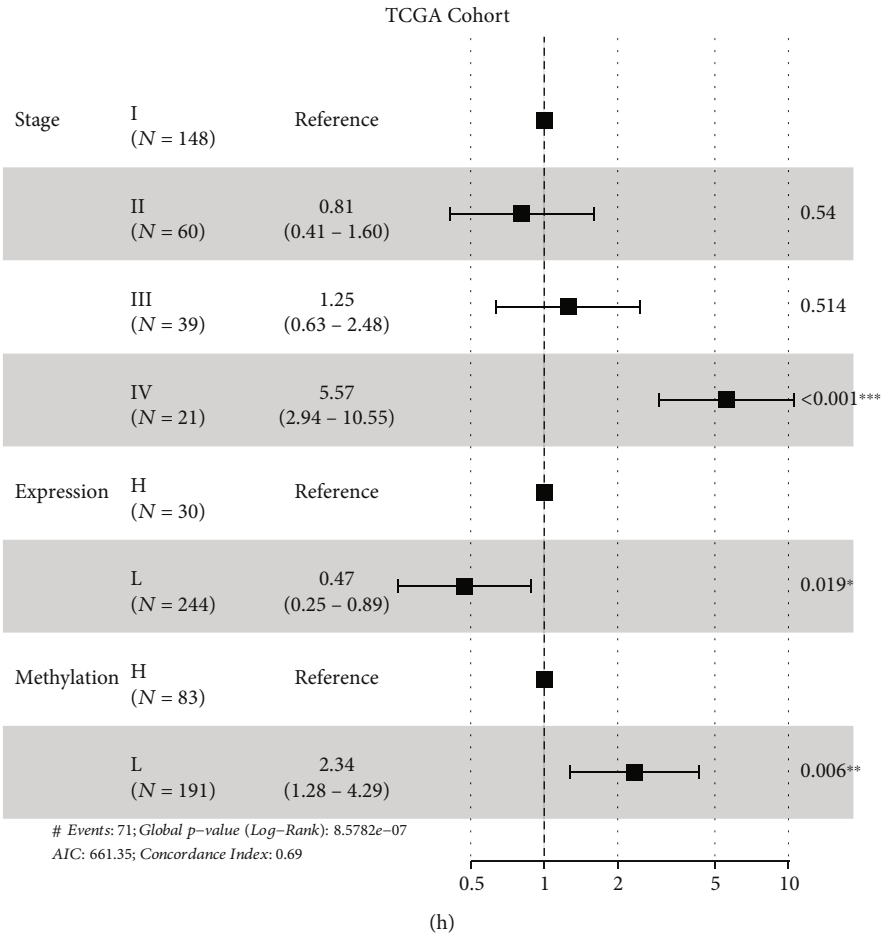
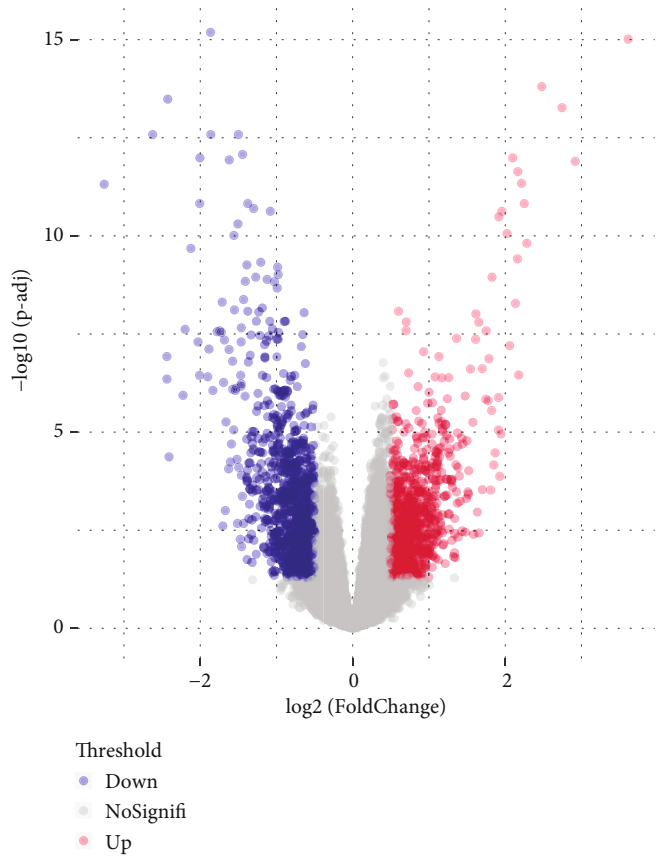


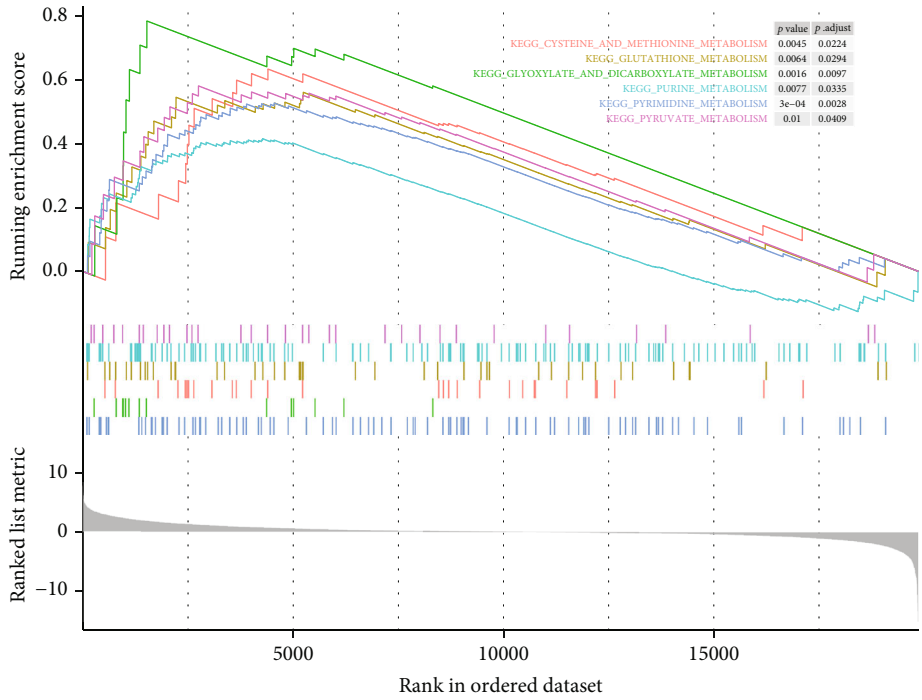
Figure 3: Landscape of genetic and expression variations of SLC24A3 in CESC. (a) The mutation frequency of SLC24A3 in CESC. (b) The Kaplan-Meier survival curve estimated in accordance with amplification, deletion, and normal SLC24A3-mutated tumor of CNV. (c) The association between copy number of variation derived from methylation and expression of SLC24A3. (d) The overall survival curve of methylation status of SLC24A3. (e) Box plot showed the relation between methylation status of SLC24A3 and CNV amplification (red), deletion (blue), and non-CNV (orange) in the TCGA-CESC cohort. (f) Box plot showed the relation between SLC24A3 expression and CNV amplification (red), deletion (blue), and non-CNV (orange) in the TCGA-CESC cohort. (g, h) The forest plots for univariate Cox regression analysis and multivariate Cox regression analysis in TCGA-CESC cohort.

3.5. The Prediction of SLC24A3 for Response to Chemotherapeutic Agents and Immune Checkpoint Inhibitors (ICIs). To fully comprehend the influence of the risk score on drug response, we estimated the relationship between the risk score and the response to drugs in tumor cell lines. So, we applied the Spearman correlation analysis to find 64 greatly associated pairs between risk score and drug sensitivity in GDSC database (Figure 6(a)). 42 pairs of which exhibited that drug sensitivity associated with the risk score, which included CCT007093, imatinib, and DMOG. 22 pairs of drug showed resistance associated with the risk score, including mitomycin C and tipifarnib. Furthermore, analyzing the signaling pathways of the genes targeted by these drugs, we identified that drugs whose sensitivity was related to the high-risk score were mainly targeting PI3K/MTOR, JNK and p38, and DNA replication signaling pathways. Conversely, the drug whose sensitivity was related to the low-risk score was mostly targeting metabolism, cell cycle, and WNT signaling pathway (Figure 6(b)).

Thereafter, we performed a study to investigate the effects of SLC24A3 with chemotherapies and immunotherapies. We utilized "pRRophetic" R package to predict therapeutic effect of 168 drugs. Among them, only 3 IC_{50} of drugs were significantly correlated with SLC24A3 expression, including mitomycin C, bleomycin, and gemcitabine (Figures 6(c)–6(e)). That meant worse expression of SLC24A3 could enhance the effect of chemotherapy drug treatment. We applied the IPS to predict ICI treatment effectiveness among CESC patients categorized as mutation and nonmutation group according to the mutation of SLC24A3 (Figure 6(g)). The nonmutation group showed superior IPS than the mutation group, which represented superior immunogenicity, thus implying a greater response to ICIs. But CESC patients with mutations of SLC24A3 had higher TMB than those without mutations (Figure 6(f)). Together, the SLC24A3 showed a considerable therapeutic value in predicting ICI treatment effect and outstanding potential drug targets.

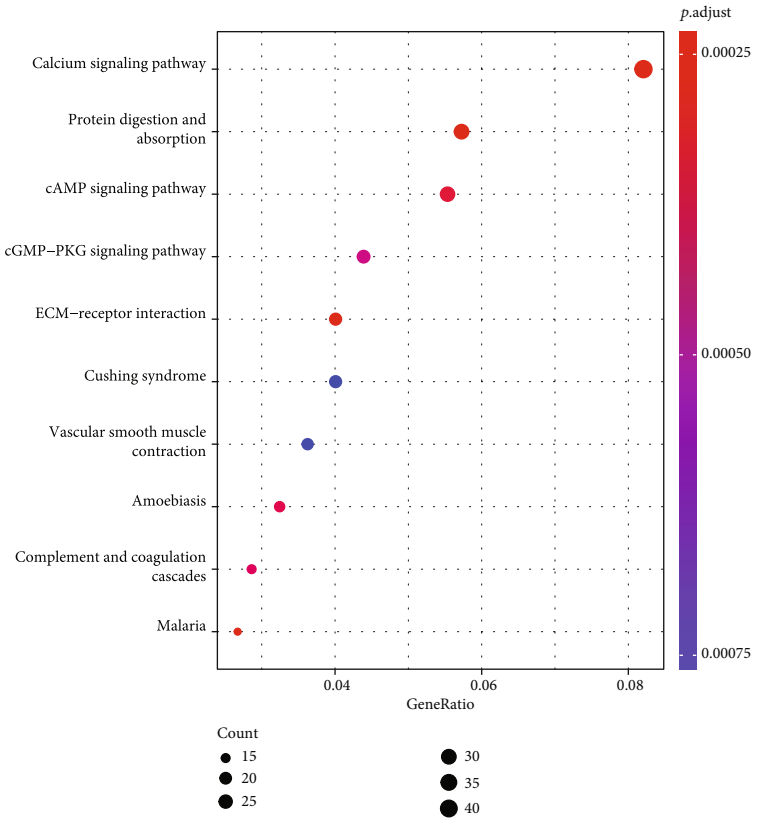


(a)

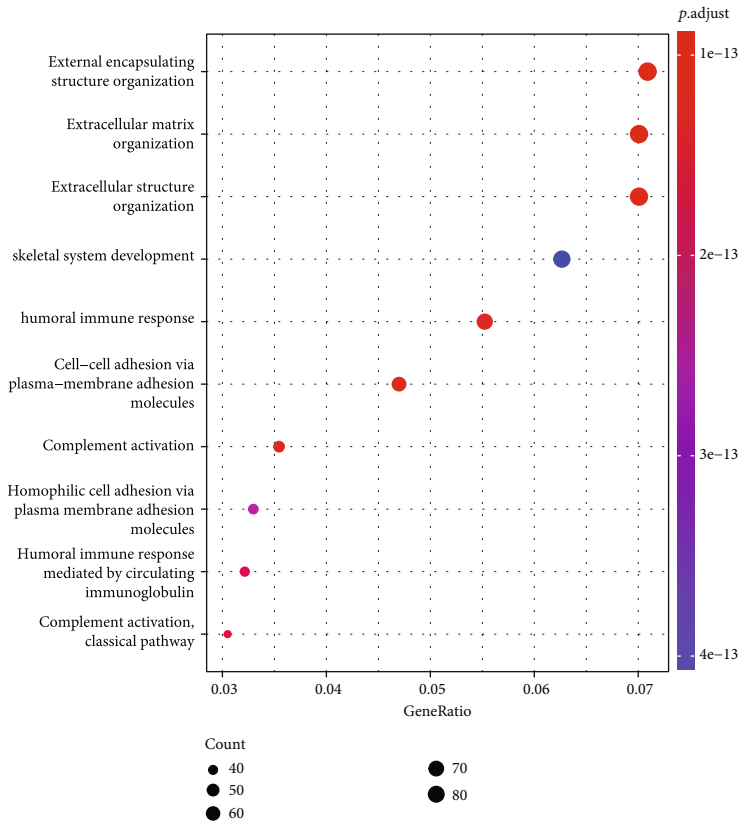


(b)

Figure 4: Continued.

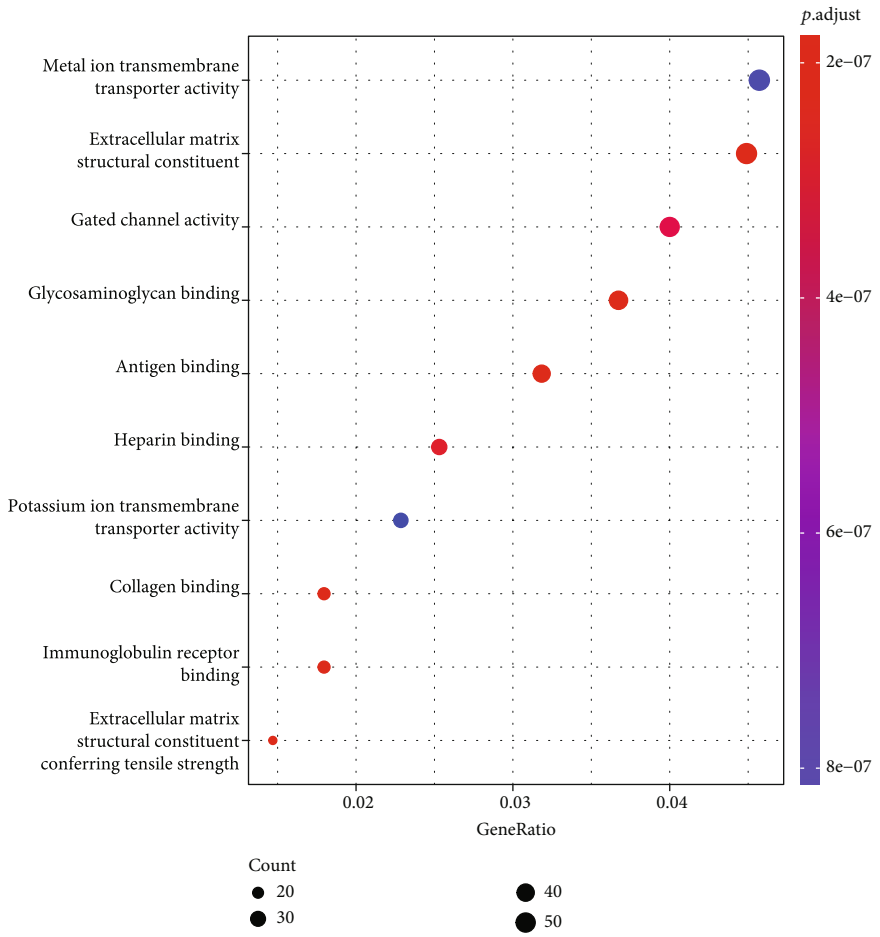


(c)

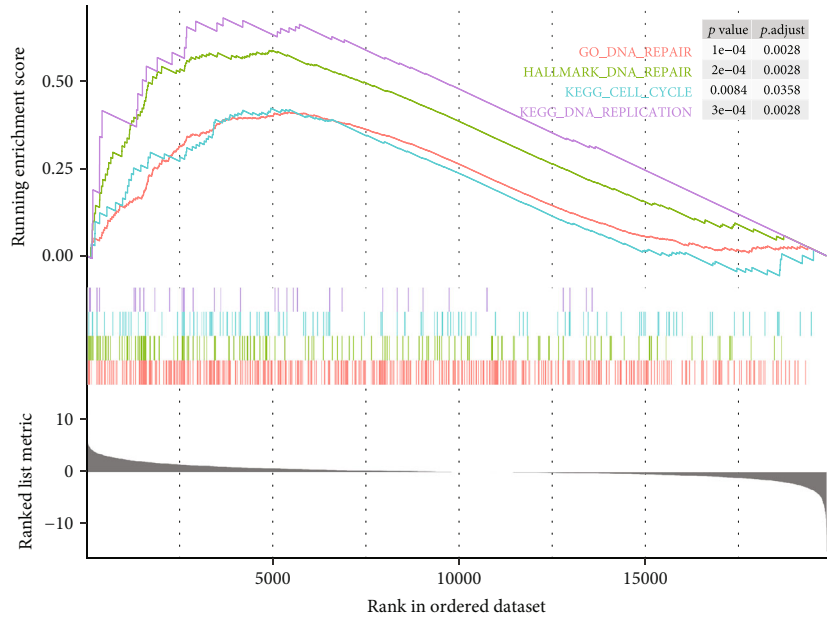


(d)

Figure 4: Continued.



(e)



(f)

Figure 4: Functional enrichment of SLC24A3-related genes. (a) Volcano plot of differential SLC24A3-related genes and $|\log FC| > 1$ with p value < 0.05 were utilized as screening criteria for DEGs. (b) KEGG signal pathway enrichment analysis. (c) Biological process enrichment analysis, (d) cell component enrichment analysis, and (e) molecular function enrichment analysis. (f) Similarly, GSEA identified a set of genes associated with SLC24A3.

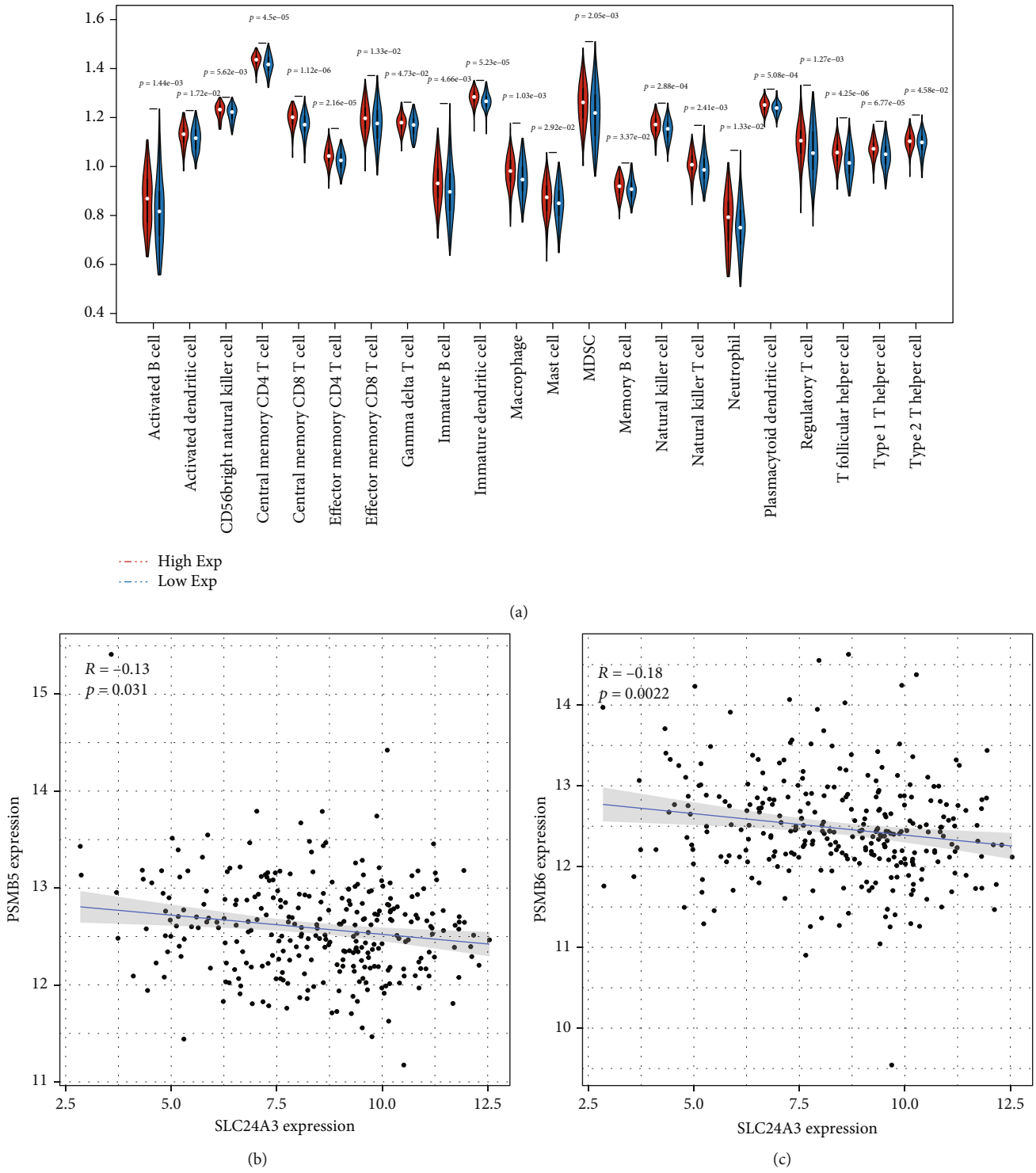


Figure 5: Continued.

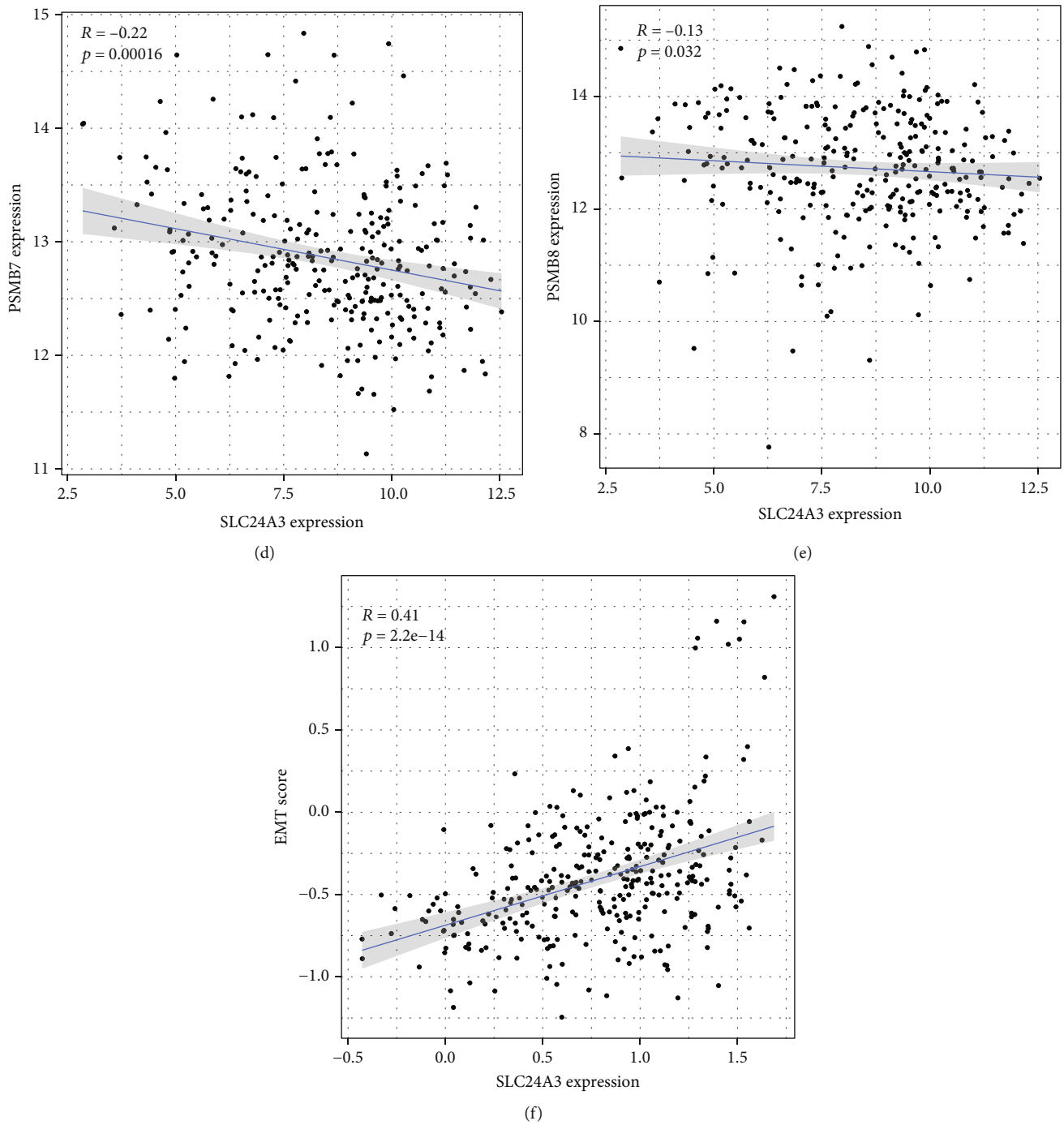
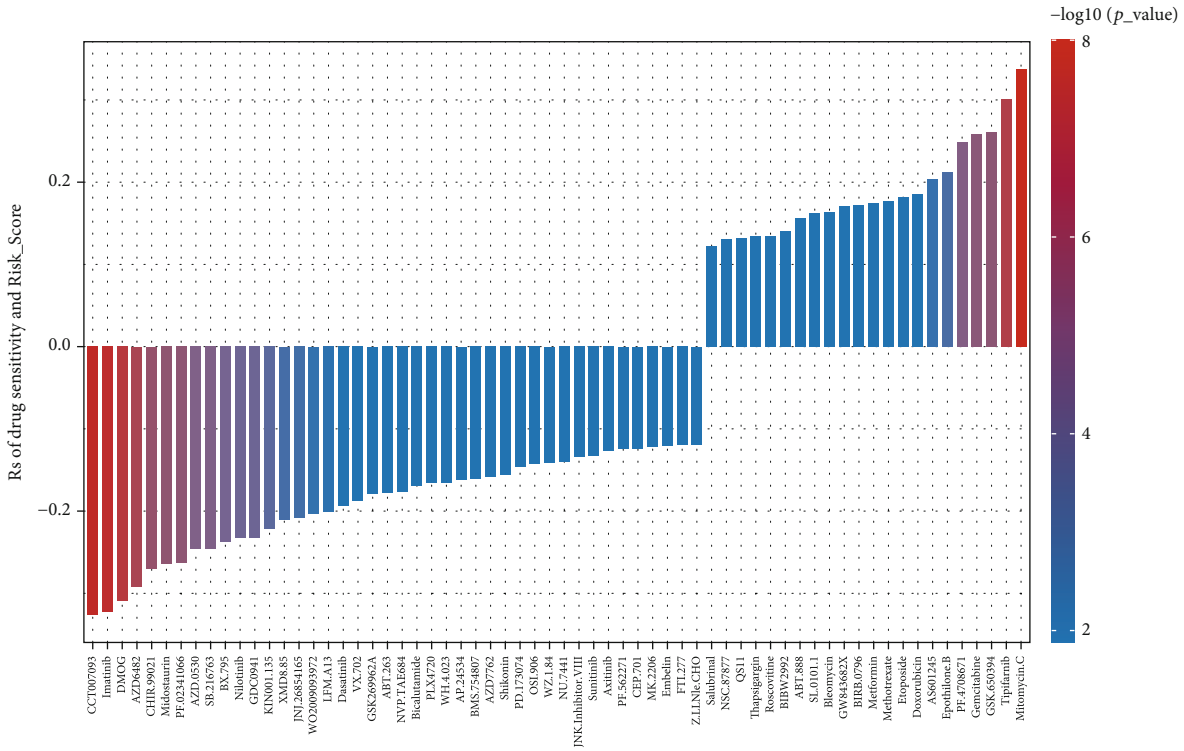


Figure 5: The correlation between SLC24A3 and immune infiltration (xCell). (a) Violin plot of association between the abundance of immune cells and SLC24A3 expression in CESC. (b–e) The association between SLC24A3 expression and antigen-presenting cell gene set. (f) The association between SLC24A3 expression and EMT score.

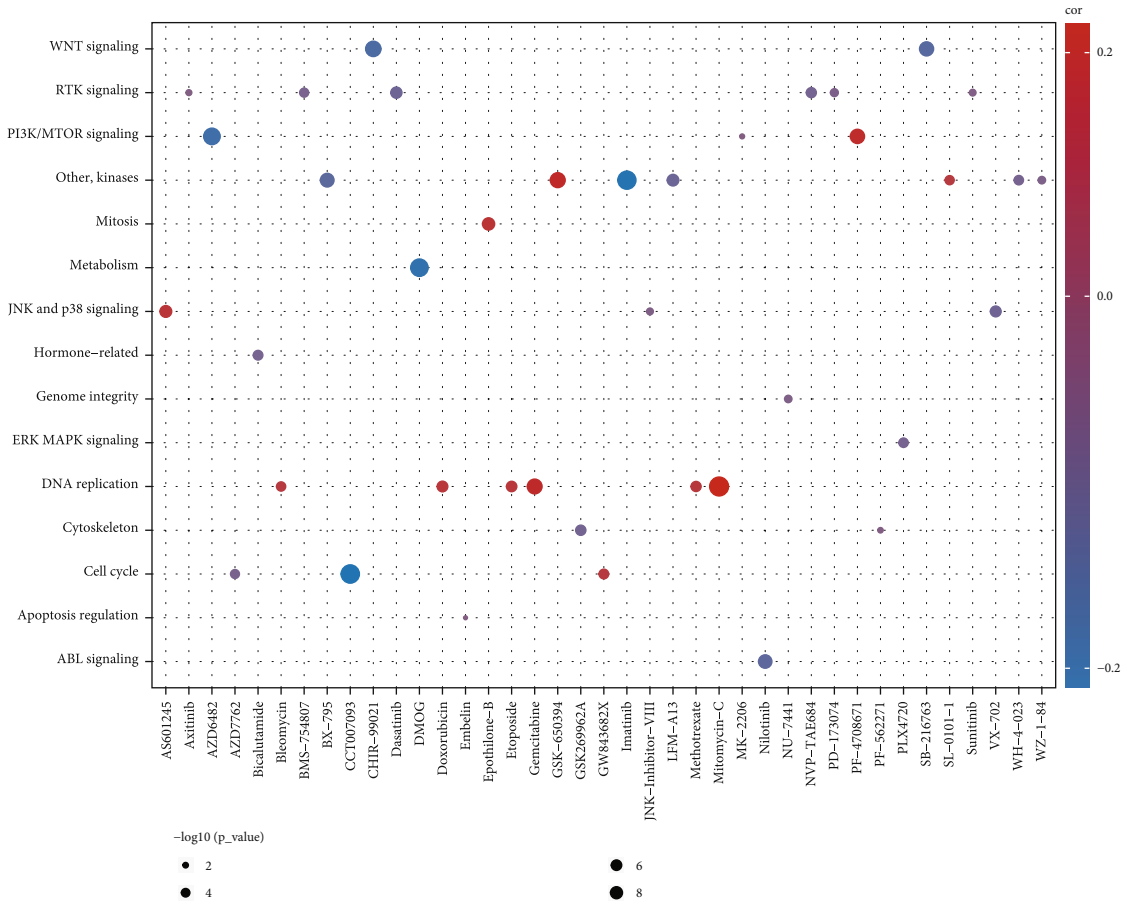
3.6. Expression of SLC24A3 in Cervical Cancer Cell Lines. qRT-PCR revealed that the levels of SLC24A3 mRNA were obviously downregulated in Hela and SiHa cell lines compared with PANC-1 and HUCEC cell lines (Figure 7, $p < 0.01$). That consisted with our bioinformatics analysis and further validated that the SLC24A3 was lower expressed among patients with CESC.

4. Discussion

Cervical carcinoma is the main reason of death among female in developing countries and areas, and there are over half a million new cases and 300,000 fatalities per year [20]. It accounts for about 10-15% of all woman cancer-related deaths globally and is the second most cause of cancer

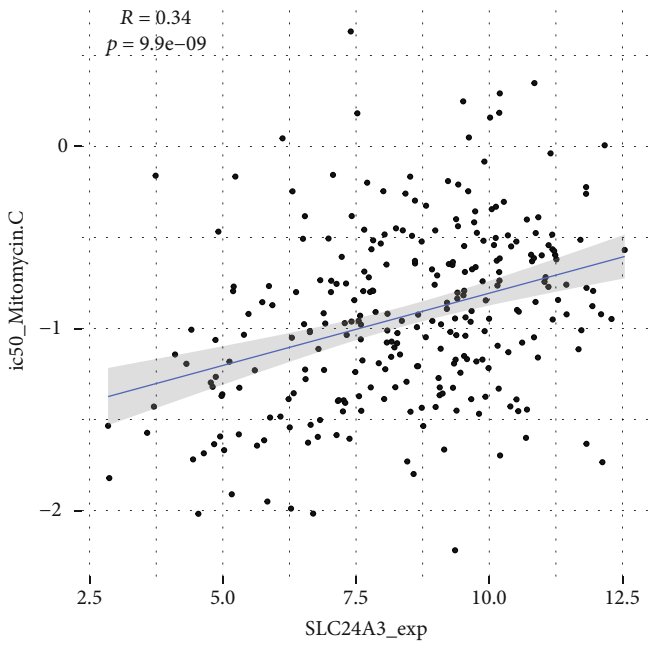


(a)

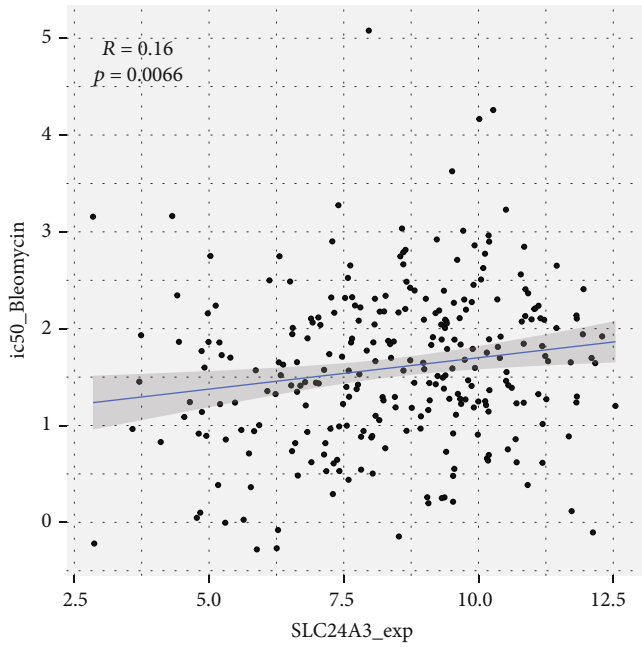


(b)

Figure 6: Continued.

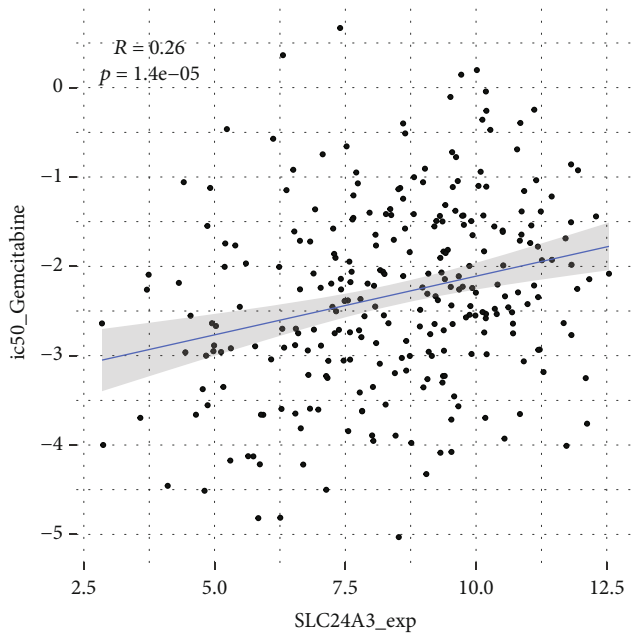


(c)

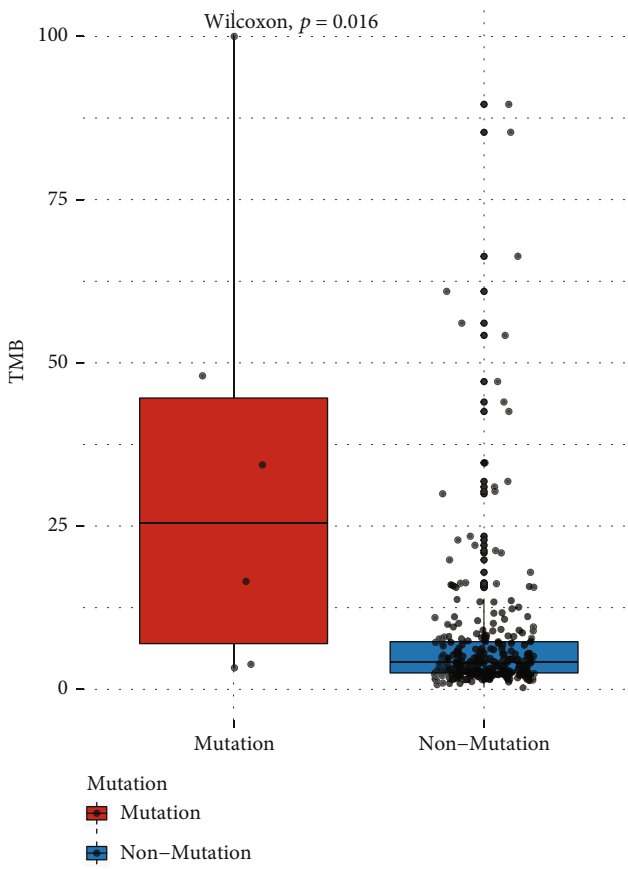


(d)

Figure 6: Continued.

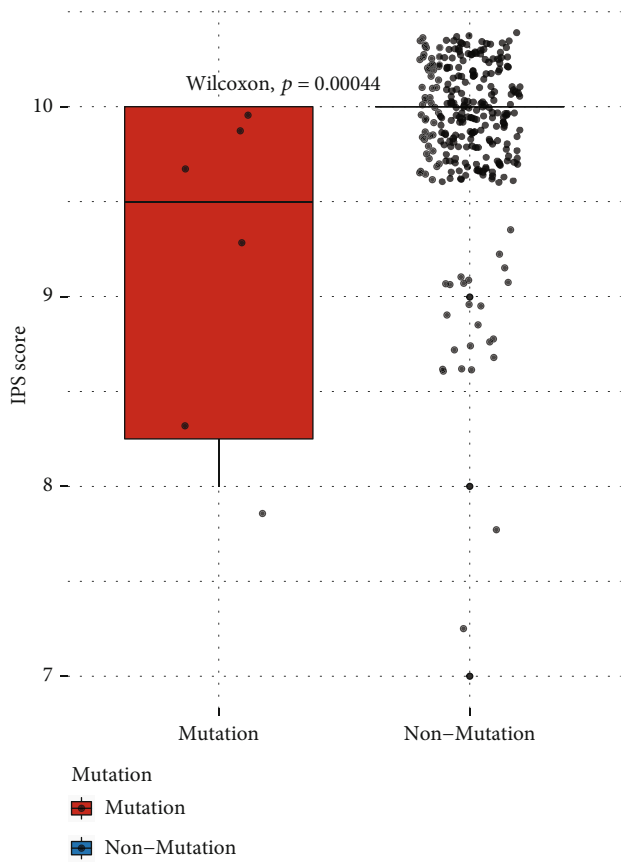


(e)



(f)

Figure 6: Continued.



(g)

Figure 6: The prediction of SLC24A3 for response to chemotherapeutic agents and immune checkpoint inhibitors. (a) The association between risk score and drug sensitivity estimated by the Spearman analysis. (b) Signaling pathways targeted by drugs that are sensitive (blue) or resistant (red) to the risk score. (c–e) The correlation between SLC24A3 expression and drug sensitivity in CESC. (f, g) Box plots showed differences in TMB (f) and IPS (g) between the SLC24A3 mutation (red) and nonmutation (blue) groups in the TCGA-CESC cohort.

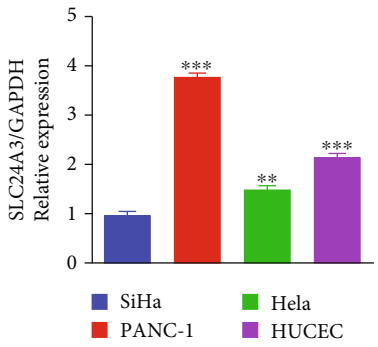


Figure 7: Expression of SLC24A3 mRNA in cervical cancer cell lines. The expression levels of SLC24A3 mRNA were significantly decreased in Hela and SiHa cell lines compared with the PANC-1 and HUCEC groups. Error bars represent ±SD. $^{*}p < 0.05$, $^{**}p < 0.01$, and $^{***}p < 0.001$. PANC-1 was utilized as a positive control. HUCEC was utilized as a negative control.

fatalities in woman, after breast cancer [21, 22]. Due to the lack of efficient early noninvasive screening, 80% of people have progressed to the invasive cancer stage of cervical cancer when were diagnosed. Meanwhile, surgery is vital for partly cancer; however, 10-80% of people suffer pain after diverse kinds of surgery. As far as we know, cancer pain refers to pain directly caused by tumor and the most turbulent influence on the quality of life of patients. Previous studies have reported that tumor invades or compresses nerve root, nerve trunk, nerve plexus, or nerve, invades brain and spinal cord, invades periosteum or bone; invades substantial organs and cavernous organs, invades or blocks vascular system, and causes local necrosis, ulceration, inflammation, etc., which can result in severe pain in all the above cases [23, 24]. Of cause, the conventional treatment of cancer pain is on basis of the WHO recommended "three-step therapy"; unfortunately, it lacks effectiveness at the terminal stage [25–27]. Although an increasing number of advances are being made in cancer treatment and diagnosis, the results of treatment are not very satisfactory as we expected. At the same time, promising insights into biomarker are now arising, and have the potential to effectively control cancer pain. Therefore, in this essay, we explore a new biomarker SLC24A3, which has functions not only in prognostic as well as treatment but also in relief cancer pain of patients with CESC.

Firstly, we downloaded CESC cohort from TCGA and GEO databases. In this study, by integrating the two GEO and TCGA-CESC cohorts, we found five difference expression genes (DIAPH3, SLC24A3, NUP62CL, C3orf70, and CD177). Then, we retrieved pain-related gene from MSigDB database. The GSEA analysis revealed that only one gene was enriched in the pain pathway. Thus, we successfully screened out the gene SLC24A3. This study manifested that SLC24A3 had a considerable difference of expression in tumor and normal tissues of CESC, and we also proved that low-expressed SLC24A3 has well survival rates in patients with CESC. Intriguingly, we found that SLC24A3 has significantly positive correlation with a classical pain gene—SLC6A4. Previous studies have indicated that the human 5-HT transporter (5-HTT) gene (SLC6A4) features some polymorphisms in its promoter region (5-HTTLPR) that impact the 5-HTT expression [28, 29]. Increasing evidence exhibited that serotonin (5-HT) is hugely correlated with pain regulation [30]. Given that SLC6A4 was highly associated with pain regulation and SLC24A3 has positively correlation with SLC6A4, we hypothesized that SLC24A3 was associated with cancer pain in CESC.

To verify the genetic alterations in SLC24A3 in CESC, we drew a panorama gene mutation showing the location and number of somatic mutations in SLC24A3. Copy number variations (CNV) are a usual form of genetic instability, which could directly alter the expression level of particular genes, which plays essential roles in susceptibility to disease [31, 32]. In addition, somatic (cancer-only) single-nucleotide variants (SNVs) as another form of mutation are the simplest type of mutation; however, their identification in DNA sequencing data is interfaced by tumor heterogeneity and sequencing, germline polymorphisms, and analysis errors [33, 34]. No matter mutations derived from CNV or SNV, all results demonstrated that the amplification mutations are positively relative with the expression level of SLC24A3. Likewise, the Kaplan-Meier survival curve revealed that SLC24A3 presenting with amplification mutations was associated with worse survival rate. Subsequently, we investigated the effect of genomic features on methylation and found that the SLC24A3 low-expressed group was highly methylated. According to the findings of univariate and multivariate Cox regression analyses, the SLC24A3-expressed level was an independent protective factor for CESC.

To further understand the functional role of the SLC24A3, GO, KEGG, and GSEA were performed, and the results indicated that SLC24A3 was greatly enriched in biological processes such as DNA repair, mitochondrion organization, NCRNA metabolic process, and reactome cell cycle mitotic.

Immune infiltration of tumors is intimately related to clinical consequences in CESC. Tumor-infiltrating immune cells (TIIC) form an ecosystem in the tumor microenvironment to modulate cancer development and have exhibited underlying prognostic value [35]. Tumor cells interact with their microenvironment and are affected by signals coming from stromal, inflammatory, immune, and endothelial cells. Certainly, tumors are usually infiltrated by diverse kinds of lymphocytes, mast cells, or macrophages [36]. Conventional wisdom is that lymphocytes could control cancer outcome, whereas the latter produce factors that promote tumor growth and maintain chronic inflammation [37]. To examine whether the expression of SLC24A3 could reflect the immune microenvironment, we performed immune infiltration of tumors analysis by using xCell algorithm. And the result implied that there were 22 immune cells enriched in the group with SLC24A3 highly expressed. In addition, correlation analysis results elucidated that SLC24A3 was a certain contributor to PSMB5, PSMB6, PSMB7, PSMB8, and EMT regulation in CESC. These changes were involved in angiogenesis in the tumor microenvironment, strengthening the invasion and metastasis of CESC cells [38, 39].

Drug sensitivity analysis was performed to validate whether risk score could predict patients' sensitivity to chemotherapy. The results of drug analysis indicated that the therapeutic effect of mitomycin C, bleomycin, and gemcitabine could be improved when the SLC24A3 was low expressed. TMB and IPS were applied to evaluate the potential response of immune checkpoint blockade (ICB) therapies. The immunotherapy results revealed that the higher the TMB and IPS, the better the immunotherapy effect. In summary, these findings may be beneficial to clinical decision-making.

The results of qRT-PCR demonstrated that the SLC24A3 mRNA was lower expressed in Hela and SiHa cell lines compared with PANC-1 (a positive control) and HUCEC (a negative control) cell lines. It was noteworthy that SLC24A3 mRNA was significantly higher expressed in PANC-1 and serves as human pancreatic cancer cell lines, compared with HUCEC cell lines. This suggested that SLC24A3 was differentially expressed in different cancer type compared with normal control tissues. As far as we know, however, no previous research has investigated whether SLC24A3 is involved in occurrence and development of pancreatic cancer. Hence, we were unable to understand how SLC24A3 exerted its function in pancreatic cancer during occurrence and development. Nonetheless, that did not influence our conclusion that the experimental results of qRT-PCR consisted with the above bioinformatics analysis obviously.

Further, in the existing studies, scientists have revealed that Nckx3 (gene SLC24A3), one of the six isoforms of Na +/K+/Ca2+ exchanger family, is able to operate either in forward or in reverse mode relying on both ion gradients and membrane potential [40–42]. Besides, transcripts of Nckx3 have been discovered in diverse other excitable body tissues with abundant smooth muscle, including the lung, intestine, aorta, and uterus [43]. Naureen et al. have reported in early research that SLC24A3 was associated with multisite chronic pain [44]. Meanwhile, SLC24A3 has been associated with some cancers. Previous studies have illustrated that SLC24A3 was significantly differentially expressed in colon cancer, ovarian cancer, and cervical cancer compared with normal control tissues [45, 46]. Gormley et al. demonstrated that SLC24A3 is a pain-related gene [47]; unfortunately, there is limited literature on the direct connection between SLC24A3 and cancer pain, and its involvement in the mechanism of initiation and progression of cancer pain is yet unclear. The existing research about SLC24A3 mainly

focuses on its own chemical mechanism, although we lack detailed clinical and experimental data to accurately estimate whether SLC24A3 has the potential to alleviate the biomarker of CESC pain. Based on the available research results, we have reason to believe that upregulating the expression of SLC24A3 will effectively alleviate the cancer pain in CESC. In other words, cancer pain is triggered precisely because SLC24A3 is low expressed in CESC, which may be due to the alteration of inflammation-related immune mechanisms. Just as we found, the deletion of Nckx3 exacerbated experimental DSS-induced mouse colitis through the p53/NF-κB pathway [48]. At the same time, based on the research we did to explore the connection between SLC24A3 and tumor-infiltrating immune cells and immune-related gene sets, we found that SLC24A3 can affect the tumor-immune microenvironment, which indicates that SLC24A3 may control the body's pain response by immunomodulation. Thus, taking all our results into consideration, the SLC24A3, as a pain-related gene, may indeed be a well biological indicator for CESC patients. This introduces a promising direction for our subsequent research, for the propose of further elucidating and understanding the mechanism of SLC24A3 in relieving cancer pain. Verifying whether SLC24A3 is appropriate as a biological indicator in early diagnosis, evaluation of prognosis and relief cancer pain of CESC patients is extremely essential.

In conclusion, firstly, we retrieved five pain-related genes from TCGA, GEO, and *MSigDB* databases. Then, SLC24A3 was obtained, enriched in the pain signaling pathway by GSEA. Subsequently, genomic profiling, enrichment analysis, immune infiltration, and chemotherapy and immunotherapy analysis were performed. All those gene bioinformatics suggested that SLC24A3 could be a possible molecular targeting mechanism for the prevention and pain treatment of CESC, which needs to be tested in further clinical trials.

Admittedly, some limitations indeed exist in this study. First, gene expression profile utilized in present study was from different platforms; this discrepancy may produce bias in the analysis process. Moreover, experiments to illustrate the underlying function and mechanism of SLC24A3 in the therapy of cervical cancer pain were not performed yet. Last but not least, we need to conduct in vitro model and experimental studies to validate the assumption we proposed according to the functions of the SLC24A3, along with the predicted drugs. Therefore, the function of SLC24A3 of CESC is encouraging enough to warrant advanced exploration.

Data Availability

Data Collection and Description. The Cancer Genome Atlas (TCGA) database (http://xena.ucsc.edu/) was applied to download gene expression quantification data and clinical information of female with CESC. Nevertheless, because the TCGA database lacks normal tissue data matched with cervical cancer, the total number of samples is only three. Hence, the Genotype-Tissue Expression (GTEx) Portal was applied to obtain the expression values of the normal cervix tissue. The Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo) was applied to download gene expression and clinical data of patients in GSE44001 and GSE63514 cohorts. The Molecular Signatures Database (MSigDB) (http://software.broadinstitute.org/gsea/msigdb/index.jsp) was used to extract pain-related genes. Since the data were downloaded from public databases, there is no need to get approval from ethics committee.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

SZ, QJ, and HY conceived, designed, and supervised the study. JY, LT, XJ, and YY drafted the manuscript and performed the data analysis and visualization. XJ, WG, WZ, and YZ collected the data. WC conducted molecular biology experiments. All authors devoted to data interpretation, manuscript preparation, editing, and review. Shuguang Zhou, Qinqin Jin, and Hui Yao contributed equally to this work.

Acknowledgments

The authors are grateful to the volunteer patients and their families who have participated in this study. This work was supported by the Applied Medicine Research Project of Hefei Health Commission (Grant No. HWKJ2019-172-14), the Research Fund Project of Anhui Medical University (Grant No. 2020xkj236), and the Natural Science Foundation of Higher Education Institutions of Anhui Province (Grant No. KJ2021A0352).

Supplementary Materials

Figure S1: correlation analysis of SLC6A4 in GSE63514 and TCGA-CESC cohorts. (A, B) Expression levels of SLC6A4 between tumor and normal cases both in TCGA-CESC and GSE63514 cohorts. (C–E) Overall survival, disease-free survival, and progression survival comparisons between high and low SLC6A4 groups in TCGA-CESC database. (F–H) The SLC24A3 expression in pathological cervical tissues and healthy cervical tissues from the Human Protein Atlas database. Figure S2: analysis of SNV mutations of SLC24A3. (A) The alteration frequency of SLC24A3. (B) The Kaplan-Meier curve showed overall survival probability of patients with (red) or without (blue) SNV mutations of SLC24A3 in TCGA-CESC cohort. (C) Box plot showed the relation between SLC24A3 expression and SNV mutations in TCGA-CESC cohorts. Supplementary Table 1: 117 pain-related genes from *MSigDB*. Supplementary Table 2: clinical characteristics of TCGA CESC data between the CNV and SNV groups. Supplementary Table 3: primer sequence of SLC24A3 used in qRT-PCR. *(Supplementary Materials)*

References

[1] D. Magee, S. Bachtold, M. Brown, and P. Farquhar-Smith, "Cancer pain: where are we now?," *Pain Management*, vol. 9, no. 1, pp. 63–79, 2019.

[2] H. J. Long, N. N. I. Laack, and B. S. Gostout, "Prevention, Diagnosis, and Treatment of Cervical Cancer," *Mayo Clinic Proceedings*, vol. 82, no. 12, pp. 1566–1574, 2007.

[3] J. A. Paice and B. Ferrell, "The management of cancer pain," *CA: a Cancer Journal for Clinicians*, vol. 61, no. 3, pp. 157–182, 2011.

[4] T. A. Berman and J. T. Schiller, "Human papillomavirus in cervical cancer and oropharyngeal cancer: one cause, two diseases," *Cancer*, vol. 123, no. 12, pp. 2219–2229, 2017.

[5] P. A. Cohen, A. Jhingran, A. Oaknin, and L. Denny, "Cervical cancer," *The Lancet*, vol. 393, no. 10167, pp. 169–182, 2019.

[6] P. Tsikouras, S. Zervoudis, B. Manav et al., "Cervical cancer: screening, diagnosis and staging," *Journal of BUON*, vol. 21, no. 2, pp. 320–325, 2016.

[7] S. Zhang, M. McNamara, and P. Batur, "Cervical cancer screening: what's new? Updates for the busy clinician," *The American Journal of Medicine*, vol. 131, no. 6, pp. 702.e1–702.e5, 2018.

[8] M. Arbyn, E. Weiderpass, L. Bruni et al., "Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis," *The Lancet Global Health*, vol. 8, no. 2, pp. e191–e203, 2020.

[9] J. B. Epstein and C. Miaskowski, "Oral pain in the cancer patient," *Journal of the National Cancer Institute. Monographs*, vol. 2019, article lgz003, no. 53, 2019.

[10] M. Zhai, S. Yang, S. Lin et al., "Distinct gene expression patterns of ion channels and cytokines in rat primary sensory neurons during development of bone cancer and cancer pain," *Frontiers in Molecular Neuroscience*, vol. 14, article 665085, 2021.

[11] C. C. Reyes-Gibby, M. R. Spitz, S. Yennurajalingam et al., "Role of inflammation gene polymorphisms on pain severity in lung cancer patients," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 18, no. 10, pp. 2636–2642, 2009.

[12] H. Breivik, N. Cherny, B. Collett et al., "Cancer-related pain: a pan-European survey of prevalence, treatment, and patient attitudes," *Annals of Oncology*, vol. 20, no. 8, pp. 1420–1433, 2009.

[13] B. M. Scarborough and C. B. Smith, "Optimal pain management for patients with cancer in the modern era," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 3, pp. 182–196, 2018.

[14] C. B. Simone 2nd, N. Vapiwala, M. K. Hampshire, and J. M. Metz, "Cancer patient attitudes toward analgesic usage and pain intervention," *The Clinical Journal of Pain*, vol. 28, no. 2, pp. 157–162, 2012.

[15] S. Deandrea, M. Montanari, L. Moja, and G. Apolone, "Prevalence of undertreatment in cancer pain. A review of published literature," *Annals of Oncology*, vol. 19, no. 12, pp. 1985–1991, 2008.

[16] W. Yang, J. Soares, P. Greninger et al., "Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Research*, vol. 41, pp. D955–D961, 2013.

[17] P. Geeleher, N. Cox, and R. S. Huang, "pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels," *PLoS One*, vol. 9, no. 9, 2014.

[18] A. Mayakonda, D. C. Lin, Y. Assenov, C. Plass, and H. P. Koeffler, "Maftools: efficient and comprehensive analysis of somatic variants in cancer," *Genome Research*, vol. 28, no. 11, pp. 1747–1756, 2018.

[19] P. Charoentong, F. Finotello, M. Angelova et al., "Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade," *Cell Reports*, vol. 18, no. 1, pp. 248–262, 2017.

[20] A. Buskwofie, G. David-West, and C. A. Clare, "A review of cervical cancer: incidence and disparities," *Journal of the National Medical Association*, vol. 112, no. 2, pp. 229–232, 2020.

[21] M. Saei Ghare Naz, N. Kariman, A. Ebadi, G. Ozgoli, V. Ghasemi, and F. F. Rashidi, "Educational interventions for cervical cancer screening behavior of women: a systematic review," *Asian Pacific Journal of Cancer Prevention*, vol. 19, no. 4, pp. 875–884, 2018.

[22] L. W. Musselwhite, C. M. Oliveira, T. Kwaramba et al., "Racial/ethnic disparities in cervical cancer screening and outcomes," *Acta Cytologica*, vol. 60, no. 6, pp. 518–526, 2016.

[23] R. A. Swarm, J. A. Paice, D. L. Anghelescu et al., "Adult cancer pain, version 3.2019, NCCN clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 17, no. 8, pp. 977–1007, 2019.

[24] M. S. Gallaway, J. S. Townsend, D. Shelby, and M. C. Puckett, "Pain among cancer survivors," *Preventing Chronic Disease*, vol. 17, article E54, 2020.

[25] C. Ripamonti and E. Bandieri, "Pain therapy," *Critical Reviews in Oncology/Hematology*, vol. 70, no. 2, pp. 145–159, 2009.

[26] G. Deng, "Integrative medicine therapies for pain management in cancer patients," *Cancer Journal*, vol. 25, no. 5, pp. 343–348, 2019.

[27] S. Kurtin and A. Fuoto, "Pain management in the cancer survivor," *Seminars in Oncology Nursing*, vol. 35, no. 3, pp. 284–290, 2019.

[28] A. Kimura, H. Yamasaki, H. Ishii, H. Yoshida, M. Shimizu, and T. Mori, "Effects of polymorphisms in the serotonin transporter promoter-linked polymorphic region on postthoracotomy pain severity," *Journal of Pain Research*, vol. 14, pp. 1389–1397, 2021.

[29] M. U. Lie, B. Winsvold, J. Gjerstad et al., "The association between selected genetic variants and individual differences in experimental pain," *Scand J Pain*, vol. 21, no. 1, pp. 163–173, 2021.

[30] E. E. Benarroch, "Descending monoaminergic pain modulation: bidirectional control and clinical relevance," *Neurology*, vol. 71, no. 3, pp. 217–221, 2008.

[31] M. Zhao, Q. Wang, Q. Wang, P. Jia, and Z. Zhao, "Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives," *BMC Bioinformatics*, vol. 14, Supplement 11, p. S1, 2013.

[32] C. Xie and M. T. Tammi, "CNV-seq, a new method to detect copy number variation using high-throughput sequencing," *BMC Bioinformatics*, vol. 10, no. 1, p. 80, 2009.

[33] N. D. Roberts, R. D. Kortschak, W. T. Parker et al., "A comparative analysis of algorithms for somatic SNV detection in cancer," *Bioinformatics*, vol. 29, no. 18, pp. 2223–2230, 2013.

[34] D. Xing, L. Tan, C. H. Chang, H. Li, and X. S. Xie, "Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, no. 8, article e2013106118, 2021.

[35] T. Li, J. Fu, Z. Zeng et al., "TIMER2.0 for analysis of tumor-infiltrating immune cells," *Nucleic Acids Research*, vol. 48, no. W1, pp. W509–W514, 2020.

[36] Y. Zhang and Z. Zhang, "The history and advances in cancer immunotherapy: understanding the characteristics of tumor-

infiltrating immune cells and their therapeutic implications," *Cellular & Molecular Immunology*, vol. 17, no. 8, pp. 807–821, 2020.

[37] C. Jochems and J. Schlom, "Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity," *Experimental Biology and Medicine*, vol. 236, no. 5, pp. 567–579, 2011.

[38] M. A. Nieto, "Epithelial-mesenchymal transitions in development and disease: old views and new perspectives," *The International Journal of Developmental Biology*, vol. 53, no. 8-10, pp. 1541–1547, 2009.

[39] D. A. Hume, "The many alternative faces of macrophage activation," *Frontiers in Immunology*, vol. 6, 2015.

[40] L. Citterio, M. Simonini, L. Zagato et al., "Genes involved in vasoconstriction and vasodilation system affect salt-sensitive hypertension," *PLoS One*, vol. 6, no. 5, article e19620, 2011.

[41] A. H. Jalloul, R. T. Szerencsei, and P. P. M. Schnetkamp, "Cation dependencies and turnover rates of the human K^+-dependent Na^+-Ca^{2+} exchangers NCKX1, NCKX2, NCKX3 and NCKX4," *Cell Calcium*, vol. 59, no. 1, pp. 1–11, 2016.

[42] H. Yang, Y. M. Yoo, E. M. Jung, K. C. Choi, and E. B. Jeung, "Uterine expression of sodium/potassium/calcium exchanger 3 and its regulation by sex-steroid hormones during the estrous cycle of rats," *Molecular Reproduction and Development*, vol. 77, no. 11, pp. 971–977, 2010.

[43] H. Yang, T. H. Kim, H. H. Lee, K. C. Choi, and E. B. Jeung, "Distinct expression of the calcium exchangers, NCKX3 and NCX1, and their regulation by steroid in the human endometrium during the menstrual cycle," *Reproductive Sciences*, vol. 18, no. 6, pp. 577–585, 2011.

[44] Z. Naureen, L. Lorusso, P. Manganotti et al., "Genetics of pain: from rare Mendelian disorders to genetic predisposition to pain," *Acta Biomed*, vol. 91, no. 13-S, article e2020010, 2020.

[45] N. Al-Henhena, S. A. Khalifa, R. P. Ying et al., "Evaluation of chemopreventive potential of Strobilanthes crispus against colon cancer formation in vitro and in vivo," *BMC Complementary and Alternative Medicine*, vol. 15, no. 1, p. 419, 2015.

[46] X. Qu, Z. Shi, J. Guo, C. Guo, J. Qiu, and K. Hua, "Identification of a novel six-gene signature with potential prognostic and therapeutic value in cervical cancer," *Cancer Medicine*, vol. 10, no. 19, pp. 6881–6896, 2021.

[47] P. Gormley, B. S. Winsvold, D. R. Nyholt, M. Kallela, D. I. Chasman, and A. Palotie, "Migraine genetics: from genome-wide association studies to translational insights," *Genome Medicine*, vol. 8, no. 1, 2016.

[48] D. N. Tran, S. M. Go, S. M. Park, E. M. Jung, and E. B. Jeung, "Loss of Nckx3 exacerbates experimental DSS-induced colitis in mice through p53/NF-κB pathway," *International Journal of Molecular Sciences*, vol. 22, no. 5, p. 2645, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 8958962, 17 pages
https://doi.org/10.1155/2023/8958962

Research Article

Identification of a Five Immune Term Signature for Prognosis and Therapy Options (Immunotherapy versus Targeted Therapy) for Patients with Hepatocellular Carcinoma

Xiaoyun Bin,[1] Zongjiang Luo,[2] Jianchu Wang,[2] and Sufang Zhou[1,3]

[1]*Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Guangxi Medical University, Nanning, Guangxi 530021, China*
[2]*Department of Hepatobiliary Surgery, Affiliated Hospital of Youjiang Medical University for Nationalities, Guangxi 533000, China*
[3]*Key Laboratory of Longevity and Aging-Related Diseases of Chinese Ministry of Education, Center for Translational Medicine, Guangxi Medical University, Nanning, Guangxi, China*

Correspondence should be addressed to Sufang Zhou; zsf200000@163.com

Received 25 August 2022; Revised 15 October 2022; Accepted 17 October 2022; Published 2 February 2023

Academic Editor: Xiaohan Ren

Background. Immune microenvironment implicated in liver cancer development. Nevertheless, previous studies have not fully investigated the immune microenvironment in liver cancer. *Methods*. The open-access data used for analysis were obtained from The Cancer Genome Atlas (TCGA-LIHC) and the International Cancer Genome Consortium databases (ICGC-JP and ICGC-FR). R program was employed to analyze all the data statistically. *Results*. First, the TCGA-LIHC, ICGC-FR, and ICGC-JP cohorts were selected for our analysis, which were merged into a combined cohort. Then, we quantified 53 immune terms in this combined cohort with large populations using the ssGSEA algorithm. Next, a prognostic approach was established based on five immune principles (CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Th2.cells, and B.cells) was established, which showed great prognostic prediction efficiency. Clinical correlation analysis demonstrated that high-risk patients could reveal higher progressive clinical features. Next, to examine the inherent biological variations in high- and low-risk patients, pathway enrichment tests were conducted. DNA repair, E2F targets, G2M checkpoints, HEDGEHOG signaling, mTORC1 signaling, and MYC target were positively correlated with the risk score. Examination of genomic instability revealed that high-risk patients may exhibit a higher tumor mutation burden score. Meanwhile, the risk score showed a strong positive correlation with the tumor stemness index. In addition, the Tumor Immune Dysfunction and Exclusion outcome indicated that high-risk patients could be higher responsive to immunotherapy, whereas low-risk patients may be higher responsive to Erlotinib. Finally, six characteristic genes DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13 were identified for risk group prediction. *Conclusions*. In summary, our study identified a signature as a useful tool to indicate prognosis and therapy options for liver cancer patients.

1. Introduction

Every year, around 840, 000 new cases of liver cancer are diagnosed and 780, 000 people lose their lives as a result of this disease. Hepatocellular carcinoma is the most common pathological subtype among all cases, accounting for 90% of primary liver cancer [1]. As a multifactorial and multicause disease, the genesis and progression of liver cancer are linked to numerous risk variables, such as genetics, lifestyle, and environmental aspects [2]. At present, surgery is still the first-line therapy option for liver cancer patients in the early stage [3]. However, considering its insidious symptoms, a substantial portion of patients have been in an advanced stage when first diagnosed, therefore, leading to loss of the appropriate timing of surgery [4]. Meanwhile, due to the characteristic of rapid progression and early metastasis, liver cancer patients tend

Table 1: Baseline information of the patients in TCGA-LIHC.

Features	Numbers	Percentage (%)
Age		
≤65	235	62.3
>65	141	37.4
Unknown	1	0.3
Gender		
Female	122	32.4
Male	255	67.6
Grade		
G1	55	14.6
G2	180	47.7
G3	124	32.9
G4	13	3.4
Unknown	5	1.3
Stage		
Stage I	175	46.4
Stage II	87	23.1
Stage III	86	22.8
Stage IV	5	1.3
Unknown	24	6.4
T-stage		
T1	185	49.1
T2	95	25.2
T3	81	21.5
T4	13	3.4
Unknown	3	0.8
N-stage		
N0	257	68.2
N1	4	1.1
Unknown	116	30.1
M-stage		
M0	272	72.1
M1	4	1.1
Unknown	101	26.8

Table 2: Baseline information of the patients in ICGC-FR.

Features	Numbers	Percentage (%)
Age		
≤65	205	55.6
>65	164	44.4
Gender		
Female	76	20.6
Male	293	79.4
T-stage		
T1	54	14.6
T2	65	17.6
T3	40	10.8
T4	1	0.3
Unknown	209	56.6
N-stage		
N0	160	43.4
Unknown	209	56.6
M-stage		
M0	159	43.1
M1	1	0.3
Unknown	209	56.6

Table 3: Baseline information of the patients in ICGC-JP.

Features	Numbers	Percentage (%)
Age		
≤65	98	37.7
>65	162	62.3
Gender		
Female	68	26.2
Male	192	73.8
Stage		
Stage I	40	15.4
Stage II	117	45.0
Stage III	80	30.8
Stage IV	23	8.8

to have a poor prognosis [2]. Therefore, it is important to find novel molecules and directions with implications for liver cancer diagnosis and therapy.

The tumor microenvironment (TME) significantly affects the biological process in cancer progression [5]. Immune cells and status are the essential components in TME. Different factors, such as cytokines, chemokines, and others, can help cancer cells shape their microenvironment to support their growth [6]. Meanwhile, the reprogramming of other cells surrounding cancer cells plays a decisive role in tumor survival and progression [7]. An example is that TME can enhance immunosuppressive M2 monocyte-derived macrophages by secreting cytokines such as IL-4, which allows the tumor to grow and progress because monocyte-derived macrophages can account for 50% of the tumor mass [8]. Also, through a CXCL13/CXCR5/NFB/p65/miR-934 positive feedback mechanism, Zhao et al. demonstrated that tumor-derived exosomal miR-934 might stimulate macrophage M2 polarizing action to increase liver metastasis of colorectal malignancy [9]. The crosstalk between immune factors and other systems can make TME more complex [10]. In TME, cancer cells showed different metabolic modes from normal cells, like the "Warburg" effect, which has a broader significance in regulating tumor immunity [10]. This energetic interaction involving tumor and immune cells results in metabolic conflict in the tumor ecosystem, restricts the number of nutrients, and causes microenvironmental acidity, which inhibits the activity of immune cells [10]. In addition, metabolic reprogramming is essential for the maintenance of immune cell stability and body balance [11]. Presently, increasing studies have pointed out that metabolic reprogramming occurs in the mechanism of immune cell growth, development, and functional activity, which is

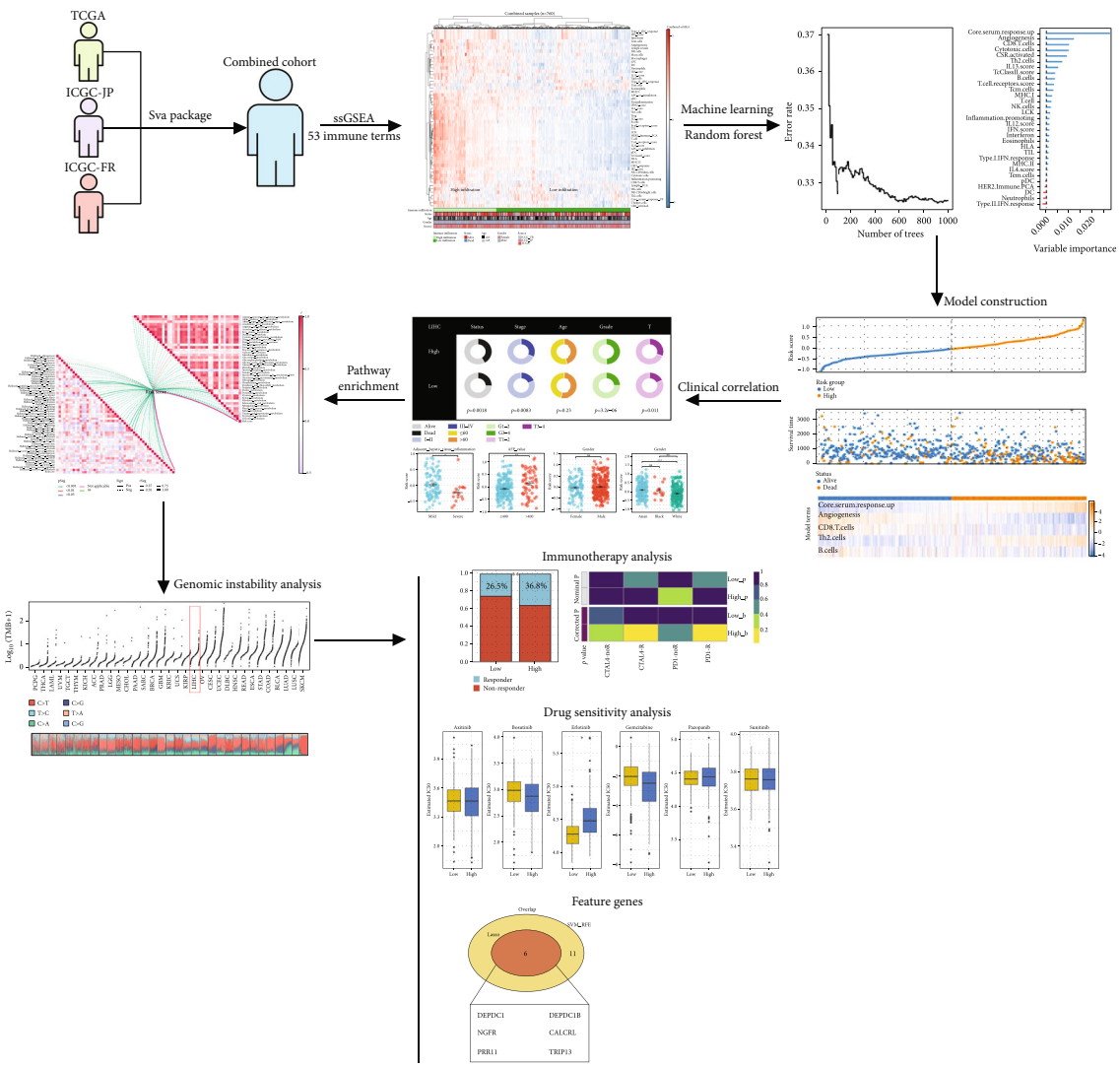


Figure 1: The flowchart of the whole study.

crucial for immune reaction [12]. Nowadays, immunotherapy presents promising therapeutic effects in specific cancer populations. Cancer immune status can also influence the intensity and time of the anticancer reaction of immunotherapy [13]. Therefore, it is meaningful to explore the immune microenvironment in liver cancer.

Here, we firstly quantified 53 immune terms in a combined cohort with large populations using the ssGSEA algorithm. Then, a prognostic modeling was created based on five immunological elements (CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Th2.T.cells, and B.cells), which showed great prognosis prediction efficiency. Analysis of clinical correlations revealed that high-risk patients may exhibit higher clinical characteristics progression. Subsequently, a pathway enrichment analysis was conducted to investigate intrinsic biological variations between high- and low-risk patients. Assessment of genomic instability revealed that high-risk individuals could possess a greater TMB score. The risk score showed a high positive correlation with tumor stemness index. In addition, the Tumor Immune Dysfunction and Exclusion (TIDE) outcome indicated that high-risk patients may show higher responsiveness to immunotherapy, whereas low-risk patients may show higher responsiveness to Erlotinib. Finally, six characteristic genes DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13 were identified for risk group prediction.

2. Methods

2.1. Data Acquisition. Open access data retrieval was carried out based on The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov/), Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/gds), and International Cancer Genome Consortium (ICGC; https://dcc.icgc.org/) databases (Access time: 2022/06/19). For the TCGA database, the expression profile information was obtained in a manner of "STAR-Counts" and then collated using the author's code. Clinical information was obtained in the "xml" form. For the GEO database, GSE14520 and GSE76427 were first included in the study for complete prognosis and transcriptional

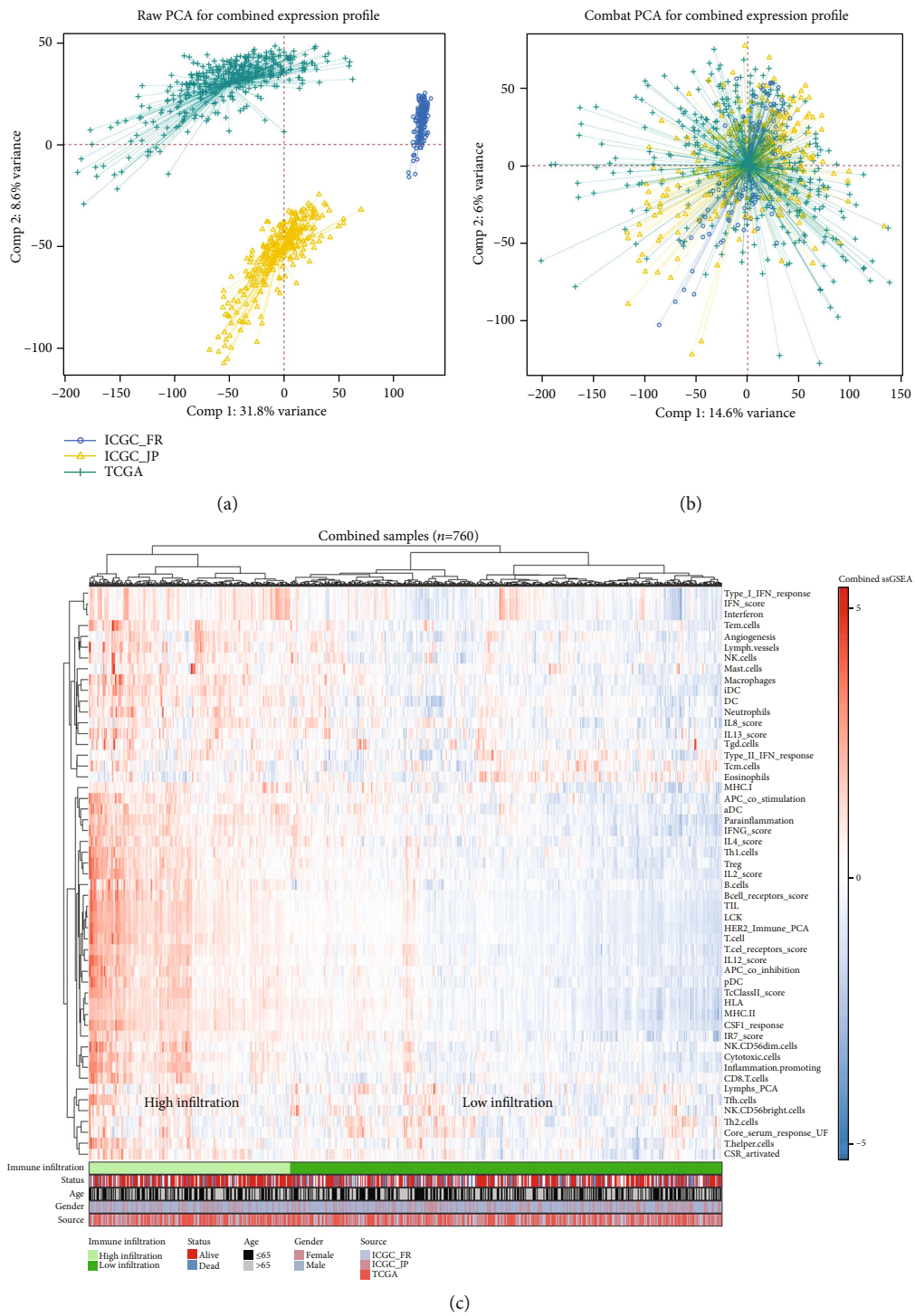


Figure 2: Quantification of 53 immune terms based on the ssGSEA algorithm. Notes: (a) three LIHC cohorts were selected for our analysis, including TCGA-LIHC, ICGC-FR, and ICGC-JP cohorts; (b) the combat function in sva package was utilized for data integration and batch difference reduction; and (c) 53 immune elements were measured using ssGSEA.

profiling data. After data quality evaluation, the GSE14520 was eliminated for the reason that the number of probes is less than 20,000; the GSE76427 was eliminated for a large number of missing values (NaN). For the ICGC, the ICGC-JP and ICGC-FR projects were included in the study. Sva package was applied for the data combination with batch effect reduction. First, by taking the logarithm, the order of magnitude of the data (TCGA and ICGC) reached the same range. Then, the combat function in sva package was used to detect and reduce the batch effect between different cohorts. Baseline information of the included patients was shown in Tables 1–3.

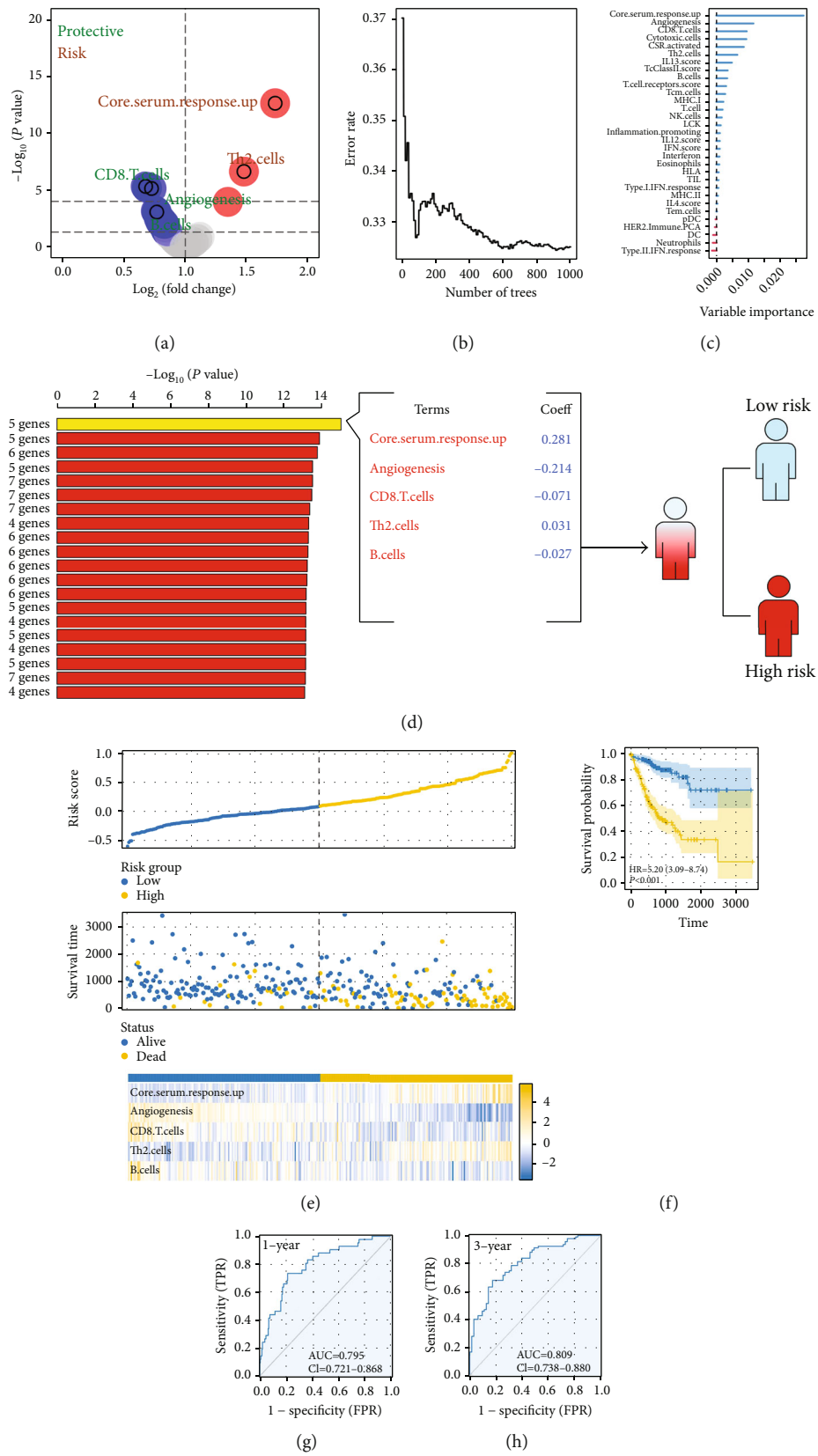


Figure 3: Continued.

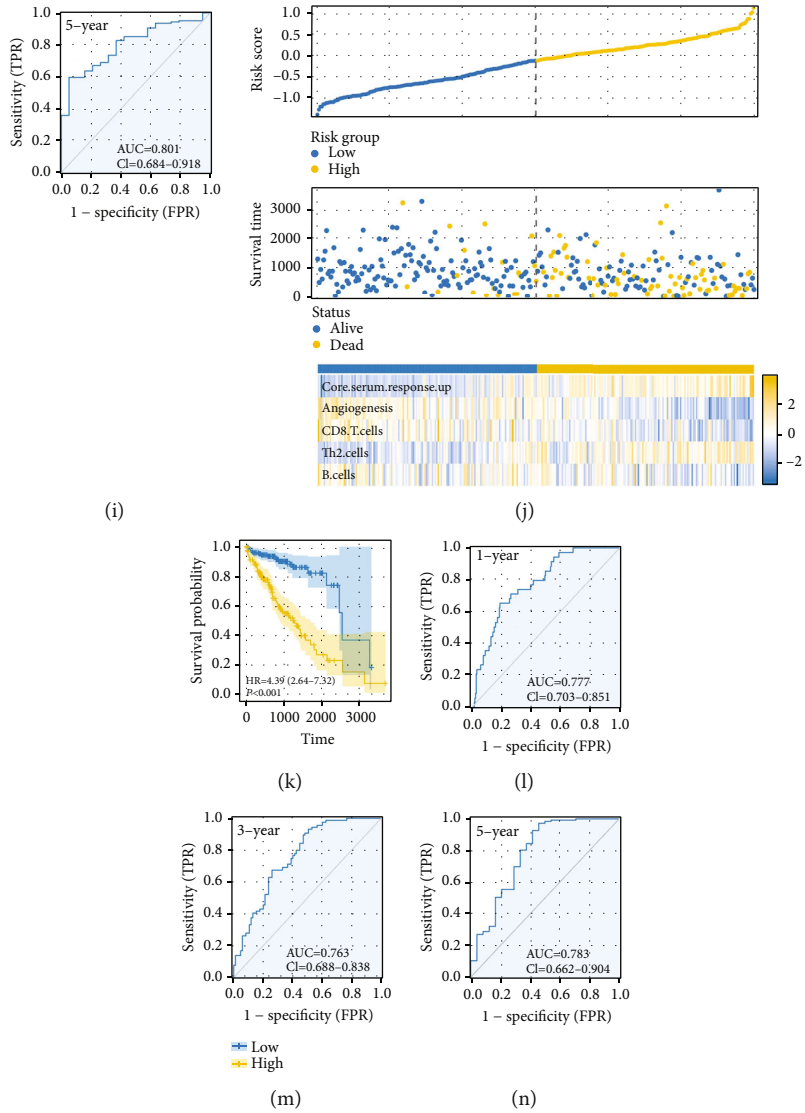


Figure 3: Prognosis model construction. Notes: (a) univariate Cox regression analysis was utilized to find prognosis-related elements with a significance level of $P < 0.05$; (b, c) the random forest technique was used for dimensionality reduction; (d) five elements, including CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Th2.cells, and B.cells, were selected for model creation; (e) the summary of our model in the training cohort; (f) the KM curves indicated that high-risk patients may have a poorer outcome (training cohort); (g–i) ROC curves for one, three, and five years (training cohort); (j) the summary of our model in the validation cohorts; (k) the KM curves indicated that high-risk patients may have a poorer outcome (validation cohort); and (l–n) ROC curves for one, three, and five years (validation cohort).

2.2. Immune Term Quantification. Immune elements were quantified using the single sample gene set enrichment analysis (ssGSEA) technique, which is an in-built algorithm of Gene Set Variation Analysis (GSVA) [14]. The reference immune terms set was obtained from the previous study, which was used to quantify the enrichment score of 53 immune terms [15]. The advantage of ssGSEA algorithm is the high freedom, in which you can quantify the enrichment score according to the given gene set. However, considering that gene sets can be freely defined, potential quality bias is inevitable.

2.3. Prognosis Analysis and Model Construction. Based on the assessment of 53 immunological elements from the ssGSEA algorithm, a univariate Cox regression analysis was conducted to discover the prognosis-related terms ($P < 0.05$). Afterward, the random survival forest variable hunting (RSFVH) method was processed to reduce the number of dimensions and filter genes. Finally, a multivariate Cox regression analysis was conducted for the development of a prognostic model. Kaplan-Meier (KM) analysis was employed to detect the best gene combination or final signature by analyzing log-rank P values. Each enrolled patient with complete prognosis and expression profile data was assigned a risk score with the formula of "Risk score = Terms A $*$ Coef A + Terms B $*$ Coef B + $\cdots$+ Terms N $*$ Coef N" [15]. If the risk score was greater than the median value, the patients were classified into the high-risk or low-risk group, accordingly. ROC and KM survival curves

Table 4: The terms significantly related to prognosis based on the univariate Cox regression.

Gene	HR	*P* value	Lower	Upper
CORE.SERUM.RESPONSE.UP	1.733	<0.001	1.496	2.007
Th2.cells	1.478	<0.001	1.274	1.714
CD8.T.cells	0.676	<0.001	0.572	0.800
Cytotoxic.cells	0.687	<0.001	0.583	0.809
Angiogenesis	0.724	<0.001	0.629	0.834
CSR.Activated	1.347	<0.001	1.158	1.567
IL13.score	0.759	0.001	0.648	0.891
B.cells	0.767	0.001	0.656	0.896
T.cell	0.783	0.003	0.667	0.920
Type.I.IFN.Reponse	0.804	0.005	0.691	0.935
TcClassII.score	0.797	0.005	0.680	0.933
TIL	0.795	0.006	0.676	0.936
LCK	0.803	0.007	0.684	0.942
NK.cells	0.815	0.009	0.698	0.951
Type.II.IFN.Reponse	0.828	0.009	0.718	0.955
MHC.I	0.829	0.011	0.717	0.959
Neutrophils	0.823	0.014	0.704	0.961
T.cell.receptors.score	0.834	0.018	0.718	0.969
DC	0.832	0.018	0.715	0.970
MHC.II	0.833	0.019	0.716	0.971
Tcm.cells	0.837	0.019	0.720	0.972
IFN.score	0.835	0.021	0.717	0.974
HLA	0.840	0.027	0.720	0.980
Inflammation.promoting	0.838	0.029	0.715	0.983
Interferon	0.845	0.031	0.726	0.985
Eosinophils	0.851	0.033	0.733	0.987
IL12.score	0.844	0.038	0.719	0.990
Tem.cells	0.852	0.040	0.731	0.992
pDC	0.853	0.043	0.731	0.995
HER2.Immune.PCA	0.849	0.045	0.724	0.997
IL4.score	0.857	0.049	0.735	0.999

were utilized to assess the accuracy of our model's prognostic projections. Univariate and multivariate Cox regression were also employed to validate the independence of the prediction model. The 1-, 3-, and 5-year survival can reflect the short-, medium-, and long-term prognosis of patients, and therefore, were selected as the time node in prognosis analysis.

2.4. Pathway Enrichment and Genomic Analysis. The GSVA and GSEA algorithms were utilized for pathway enrichment analysis [16]. Hallmark was used as the standard gene set for the GSVA algorithms, whereas metabolism-related gene sets (41 metabolism terms) were acquired from the website https://www.gsea-msigdb.org/. Standard gene sets for the GSEA algorithm were c2.cp.kegg.v7.5.1.symbols.gmt and c5.go.v7.5.1.symbols.gmt. The TCGA database was accessed using genomic mutation information, including the tumor mutation burden (TMB) and microsatellite instability (MSI) score. Based on the expression profile and utilizing the one-class logistic regression machine learning (OCLR) machine-learning technique, the tumor stemness index was computed [17].

2.5. Immunotherapy and Drug Sensitivity Analysis. Using the TIDE methodology [18], patients were evaluated for immunotherapy sensitivity. The parameter of "Cancer type" was set as "Other." The parameter of "Previous immunotherapy" was set as "No." The analysis of drug responsiveness was conducted using the database of Genomics of Drug Responsiveness in Cancer [19].

2.6. Feature Gene Identification. To identify the feature genes for the risk group, the LASSO regression and SVM-RFE (support vector machine recursive feature elimination) algorithm were applied to find the best variable [20].

2.7. Western Blot. Total proteins were extracted using a total protein extraction kit (Beyotime, China). Western blot was conducted based on the standardized process (10% SDS-PAGE gel). The primary antibody of CALCRL (1 : 2000) and GAPDH (1 : 50000) was purchased from Proteintech.

2.8. Statistical Analysis. This research was analyzed using R software version 4.2.1. Two-sided P values < 0.05 were considered statistically significant. For continuous variables with normal distributions, an independent t-test was applied, and for continuous variables with skewed distributions, a Wilcoxon rank-sum test was conducted. The study of differentially expressed genes (DEGs) was conducted using the limma program with the criteria $|\log FC| > 1$ and $P < 0.05$.

3. Results

3.1. Immune Term Quantification. Figure 1 displays the flowchart of the entire investigation. Three distinct liver cancer cohorts, TCGA-LIHC, ICGC-JP, and ICGC-FR, were chosen for our research (Figure 2(a)). The sva package was employed to combine data and decrease the batch effect. Then, a significant batch impact decline was noted (Figure 2(b)). The ssGSEA method was employed in the pooled cohorts to quantify 53 immunological elements (Figure 2(c)).

3.2. Prognosis Model Construction. First, the patients were randomly divided into training and validation cohorts according to the 1 : 1 ratio. Based on the 53 immunological elements, a univariate Cox regression analysis was conducted to find the prognosis-related variables with $P < 0.05$ (Figure 3(a) and Table 4). The random forest approach was then used to reduce the dimensionality, and the top ten important terms were CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Cytotoxic.cells, CSR.activated, Th2.cells, IL13.score, TcClassII.score, B.cells, and T.cells.receptors.score (Figures 3(b) and 3(c)). Through multivariate Cox regression analysis and permutations, the five immune terms were used for prognosis model construction, including CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Th2.cells, and B.cells (Figure 3(d)). The risk score was calculated with the formula of "Risk score = CORE.SERUM.RESPONSE.UP * 0.281 + angiogenesis *

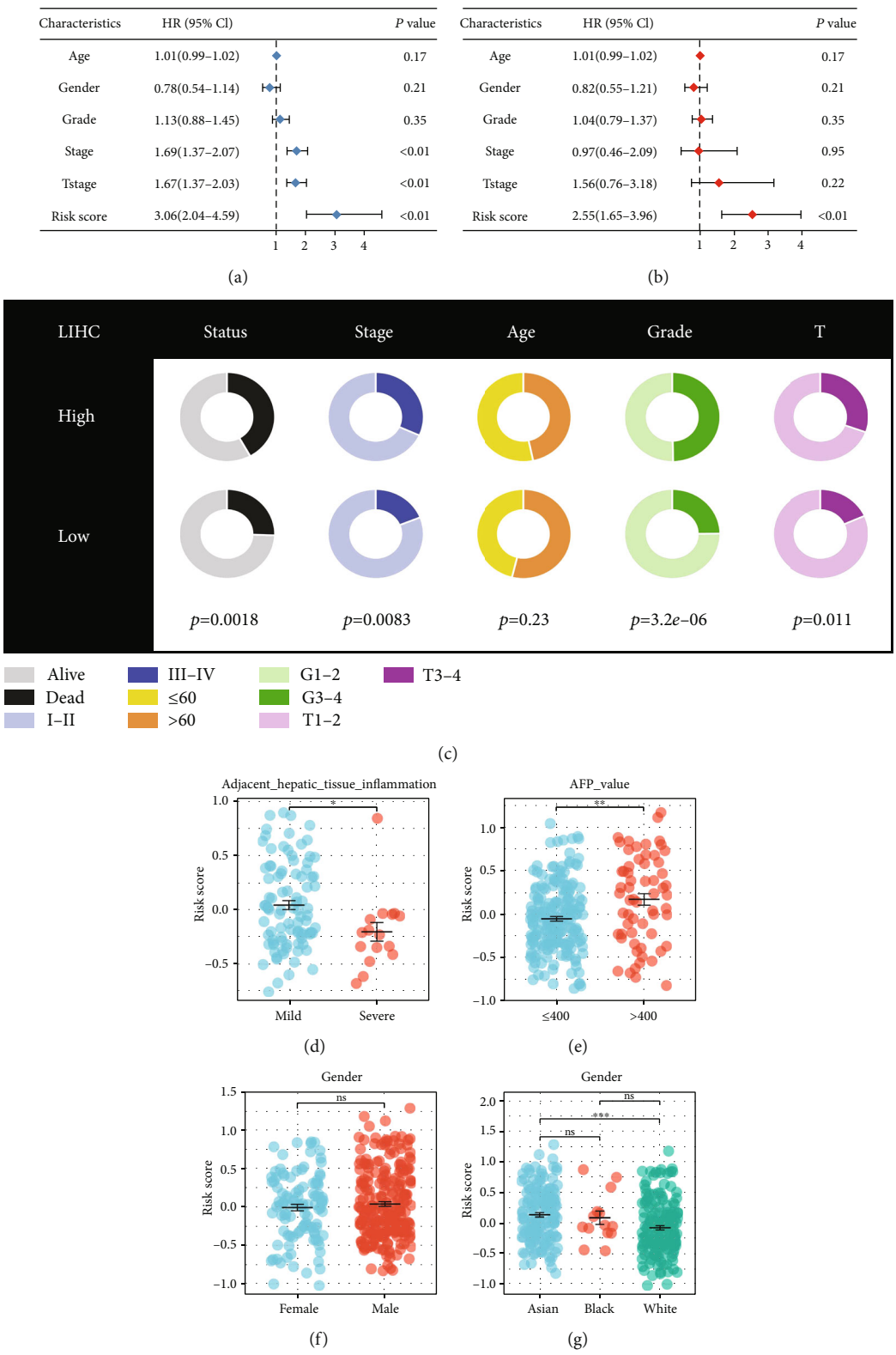


Figure 4: Continued.

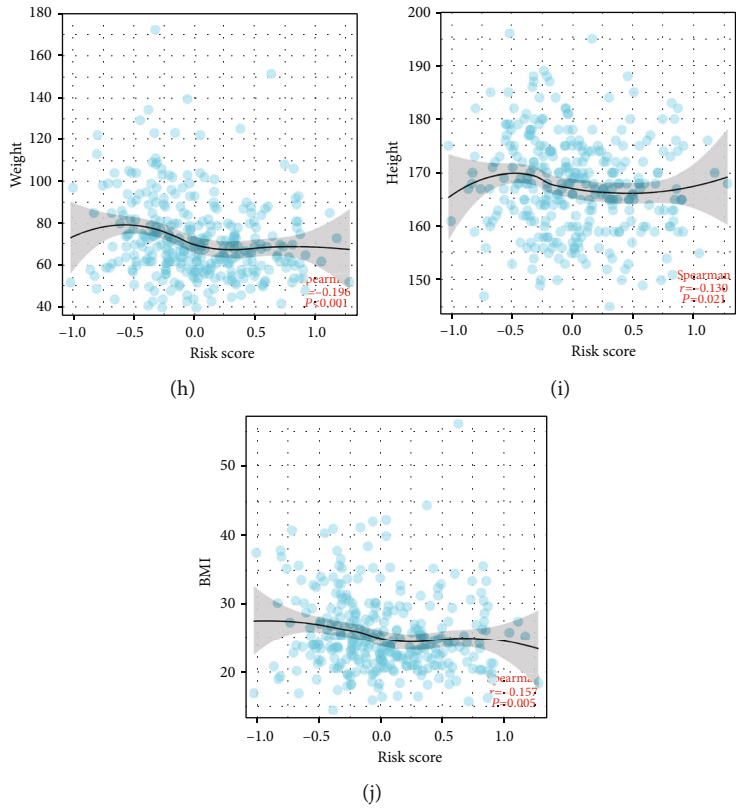


Figure 4: Clinical correlation of the riskscore. (a) Univariate Cox analysis of the riskscore as well as other clinical variables; (b) multivariate Cox analysis of riskscore and other clinical characteristics; (c) variations in clinical findings between high and low risk patients; (d–g) the riskscore difference in specific patients; and (h–j) the correlation of riskscore with patients height, weight, and BMI.

-0.214 + CD8.T.cells * -0.071 + Th2.cells * 0.031 + B.cells * -0.027". The overview of our model (training group) was shown in Figure 3(e), whereas a greater number of fatalities were seen in the high-risk group. The KM survival curve revealed that high-risk individuals tend to have a poorer outcome (Figure 3(f), HR = 5.20, $P < 0.001$, and concordance index = 0.893). ROC curves for 1-, 3-, and 5-year patients demonstrated a high predictive accuracy (Figures 3(g)–3(i), 1-year: AUC = 0.795, 3-year: AUC = 0.809, and 5-year: AUC = 0.801). Also, in the validation group, the same trend was observed (Figure 3(j)). KM survival curves indicated that the high-risk patients might have a worse prognosis performance (Figure 3(k), HR = 4.39, $P < 0.001$, and concordance index = 0.714). Meanwhile, the performances of ROC curves for 1-, 3-, and 5-year-old patients are still satisfactory (Figures 3(l)–3(n), 1-year: AUC = 0.777, 3-year: AUC = 0.763, and 5-year: AUC = 0.783).

3.3. Clinical Correlation Analysis. Furthermore, we explore the clinical correlation of our model. Univariate Cox regression and multivariate Cox regression analysis revealed because our model is not related to other clinical characteristics (Figures 4(a) and 4(b), univariate: HR = 3.06, $P < 0.01$; multivariate: HR = 2.55, $P < 0.01$). Clinical correlation analysis showed that high-risk patients could possess more aggressive clinical characteristics, such as clinical stage, grade, and T classification (Figure 4(c)). Interestingly, we found that the patients with mild adjacent hepatic tissue inflammation might have a higher risk score compared to the severe group (Figure 4(d)). Moreover, the patients with AFP > 400 ng/ml had a higher risk score than those with AFP < 400 ng/ml (Figure 4(e)). No significant difference was found in different gender patients. Asian populations might have a higher risk score than White populations (Figure 4(f)). A negative correlation was found between height, weight, and BMI (Figures 4(h)–4(j), weight, $R = -0.196$, $P < 0.001$, height, $R = -0.130$, $P = 0.021$, BMI, $R = -0.157$, $P = 0.005$).

3.4. Pathway Enrichment Analysis. We investigated the biological differences between high-and low-risk patients. For Hallmark pathways, we found that risk score was positively correlated with DNA repair, E2F targets, G2M checkpoints, HEDGEHOG signaling, mTORC1 signaling, and MYC targets (Figure 5(a)). For metabolism-related pathways, we observed that riskscore was positively correlated with purine and pyrimidine metabolism (Figure 5(a)). The GSEA method illustrated that in the high-risk group, the terms of DNA-dependent DNA replication, chromosome, cell-cell junction assembly, nuclear chromosome, leukocyte transendothelial migration, ubiquitin-mediated proteolysis, tight junction, actin cytoskeleton regulation, and MAPK signaling were upregulated (Figures 5(b) and 5(c)). Also, based on the DEGs discovered between high- and low-risk patients, a network of protein-protein linkages was developed (Figure 5(d)). ClueGO analysis

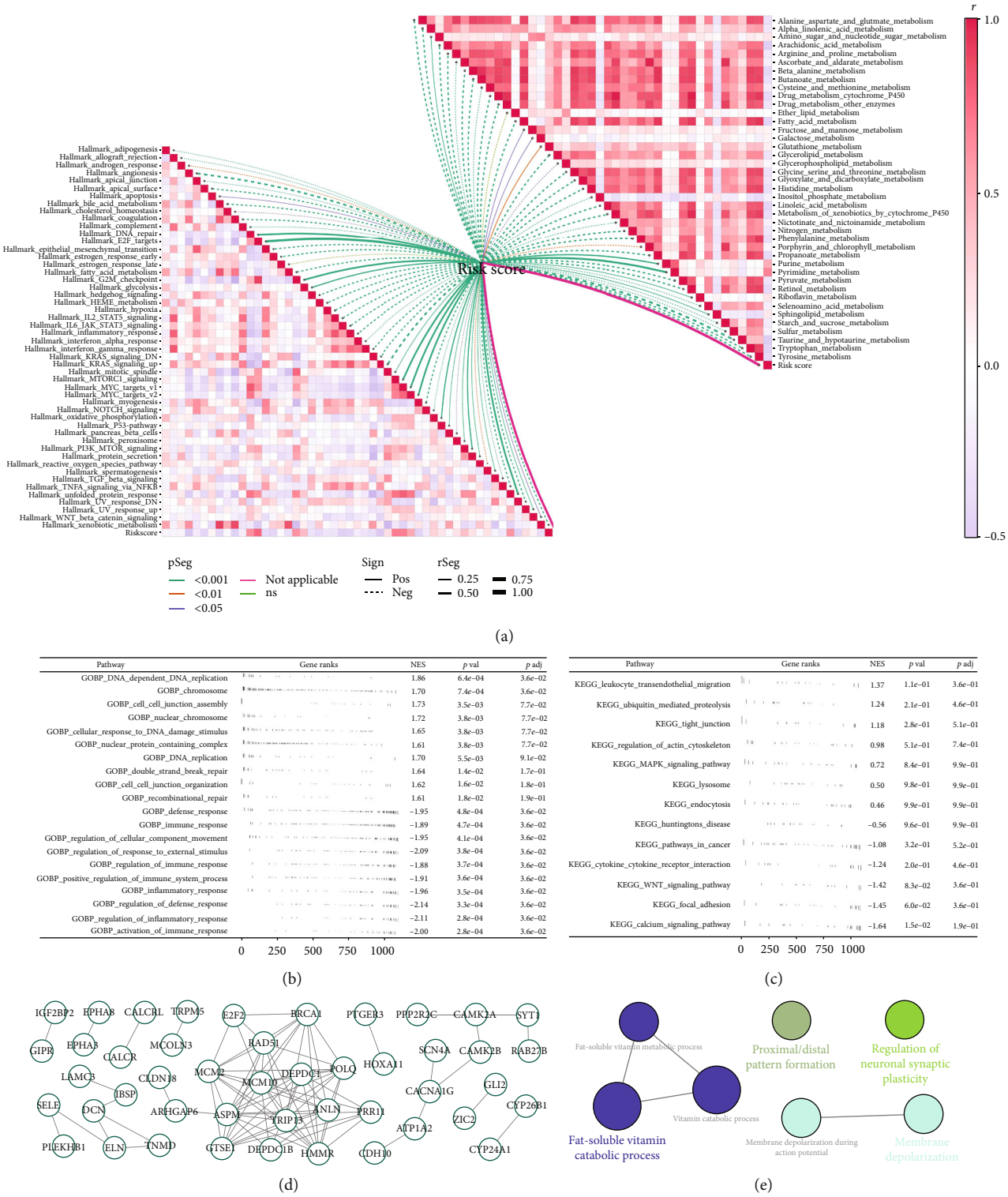


Figure 5: Pathway enrichment analysis. (a) GSVA analysis indicated that the correlation between riskscore and Hallmark and metabolism pathways; (b, c) GSEA analysis of GO and KEGG pathways; (d) the PPI network of the DEGs between high- and low-risk group; and (e) ClueGO analysis of the nodes.

revealed that these nodes were mainly enriched in fat-soluble vitamin catabolic process, proximal/distal pattern formation, regulation of neuronal synaptic plasticity, and membrane depolarization (Figure 5(e)).

3.5. Genomic Instability Analysis. The genomic feature can also affect the tumor biological process. Genetic variations are between high- and low-risk patients. Mutation information was obtained from the TCGA database (Figure 6(a)).

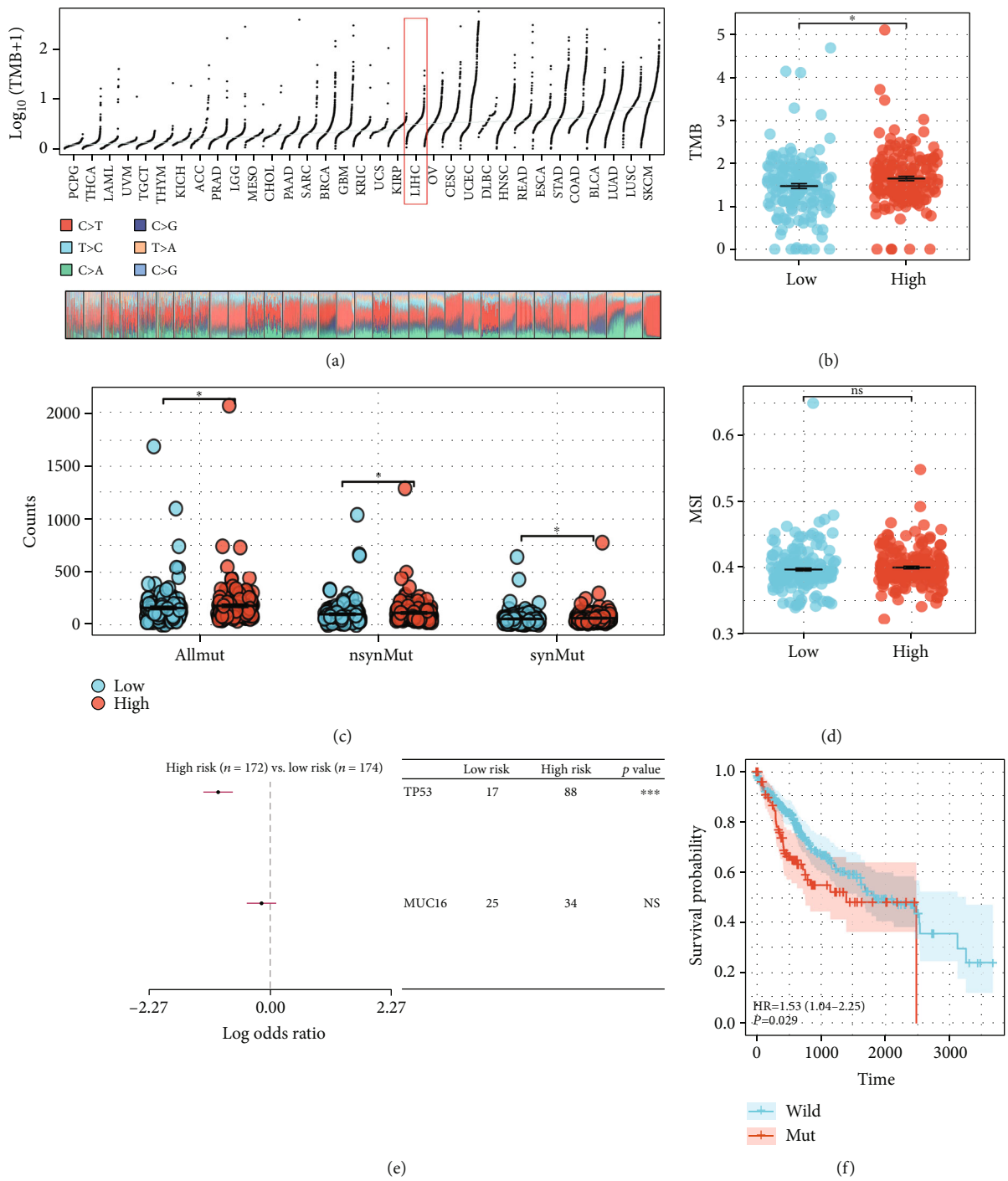


Figure 6: Continued.

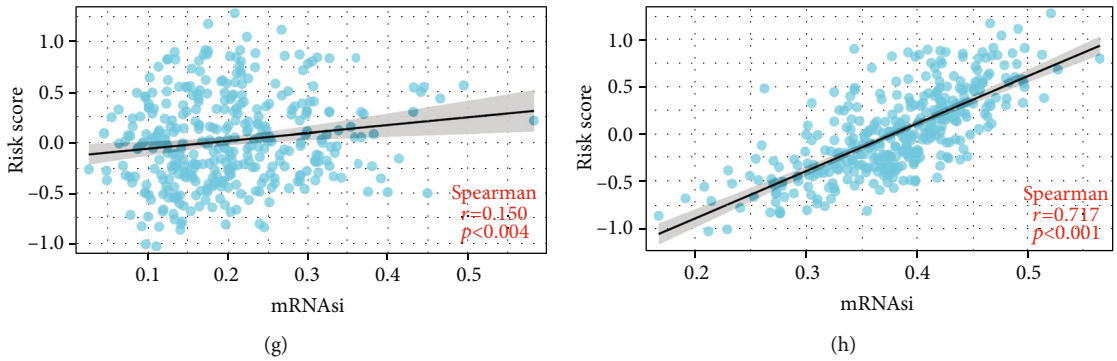


Figure 6: Genomic instability analysis. (a) A summary of the distribution of TMB in TCGA pan-cancer; (b) comparison of TMB between high-risk and low-risk patients; (c) the disparity between high-risk and low-risk patients' mutant count, nonsynonymous mutation count, and synonymous mutation count; (d) the MSI distinction between patients at high and low risk; (e) TP53 and MUC16 were the most frequently mutated genes comparing patients at high and low risk; (f) KM survival curve of the TP53 wild and mut patients; and (g, h) the correlation of riskscore with tumor stemness index.

The outcome indicated that high-risk patients may possess a greater TMB score (Figure 6(b)). Meanwhile, we found that all mutant counts, synonymous mutation counts, and non-synonymous mutation counts were elevated in patients at high risk (Figure 6(c)). No significant variation in the MSI score was observed between high- and low-risk patients (Figure 6(d)). Moreover, we discovered that TP53 was the gene greatest substantially mutated across patients at high and low risk (Figure 6(e)). The KM curve indicated that patients with TP53 mutations may suffer a poorer outcome (Figure 6(f)). In addition, a positive significant correlation was identified between risk score and tumor stemness index (Figures 6(g) and 6(h), mDNAsi, $R = 0.150$, $P = 0.004$; mRNAsi, $R = 0.71$, $P < 0.001$).

3.6. Immunotherapy and Drug Sensitivity Analysis. Immune checkpoint modules contribute substantially to the course of cancer; hence, we analyzed the correlation between the risk score and numerous checkpoint modules. Correlation analysis indicated significant variations between the high- and low-risk groups for various immune checkpoint modules, such as PD-L1 and CTLA-4 (Figures 7(a)–7(e)). TIDE research revealed that high-risk patients may show a lower TIDE score and a greater probability of immunotherapy responses (Figures 7(f) and 7(g)). The submap algorithm suggested that high-risk patients could be highly responsive to PD-1 and CTLA-4 therapy (Figure 7(h)). Analysis of drug responsiveness revealed that low-risk patients could be higher responsive to Erlotinib (Figure 7(i)).

3.7. Identification of the Characteristic Genes of the Risk Group. Considering the prognosis and therapy sensitivity difference between high- and low-risk patients, we try to identify the characteristic gene that could robustly indicate the risk group. Utilizing LASSO regression and the SVM-RFE technique, distinctive genes were identified (Figures 8(a) and 8(b)). The intersection of these two algorithms identified six genes, including DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13 (Figure 8(c) LASSO logistic regression: DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13; SVM-RFE: CALCRL, MCM10, SCN4A, NGFR, DEPDC1, PRR11, DEPDC1B, RAD51, ANLN, PYGM, CD5L, DCN, IYD, GLI2, TRIP13, TNMD, and SLC12A1). Among these genes, DEPDC1, DEPDC1B, PRR11, and TRIP13 were increased in the high-risk group, while NGFR and CALCRL were decreased (Figure 8(d)). All these characteristic genes showed great prediction efficiency in the patients' risk group (Figure 8(e), training cohort, DEPDC1, AUC = 0.792; DEPDC1B, AUC = 0.782; TRIP13, AUC = 0.814; CALCRL, AUC = 0.760; PRR11, AUC = 0.804; NGFR, AUC = 0.754). Logistic regression was performed to combine these genes with the formula of "$0.3245 + 0.3252 * \text{DEPDC1} + 0.3248 * \text{DEPDC1B} + - 0.8162 * \text{NGFR} + - 1.388 * \text{CALCRL} + 0.3101 * \text{PRR11} + 1.8485 * \text{TRIP13}$", which showed extremely great prediction efficiency in patients risk group (Figure 8(f), AUC = 0.932). In the validation cohort, these characteristic genes also showed great prediction efficiency in the patients' risk group (Figure S1A–F, DEPDC1, AUC = 0.816; DEPDC1B, AUC = 0.766; TRIP13, AUC = 0.833; CALCRL, AUC = 0.711; PRR11, AUC = 0.820; NGFR, AUC = 0.716), as well as the logistic regression model (Figure S1G, AUC = 0.896). Among these characteristic genes, DEPDC1, DEPDC1B, PRR11, and TRIP13 were risk factors for liver cancer patients (Figure 8(g)). Meanwhile, DEPDC1, DEPDC1B, CALCRL, PRR11, and TRIP13 were significantly upregulated in liver cancer tissue, yet NGFR was downregulated (Figure 8(h)). In liver cancer, the genes DEPDC1, DEPDC1B, PRR11, and TRIP13 have been explored in previous studies [21–24]. Therefore, we selected CALCRL for further validation. Western blot indicated that the protein level of CALCRL was upregulated in liver cancer tumor tissue (Figure S2).

4. Discussion

Liver cancer is still a serious public health concern worldwide [25]. Recently, researchers have focused on the immune microenvironment in liver cancer, which can significantly affect the progression of the disease. Thus, a

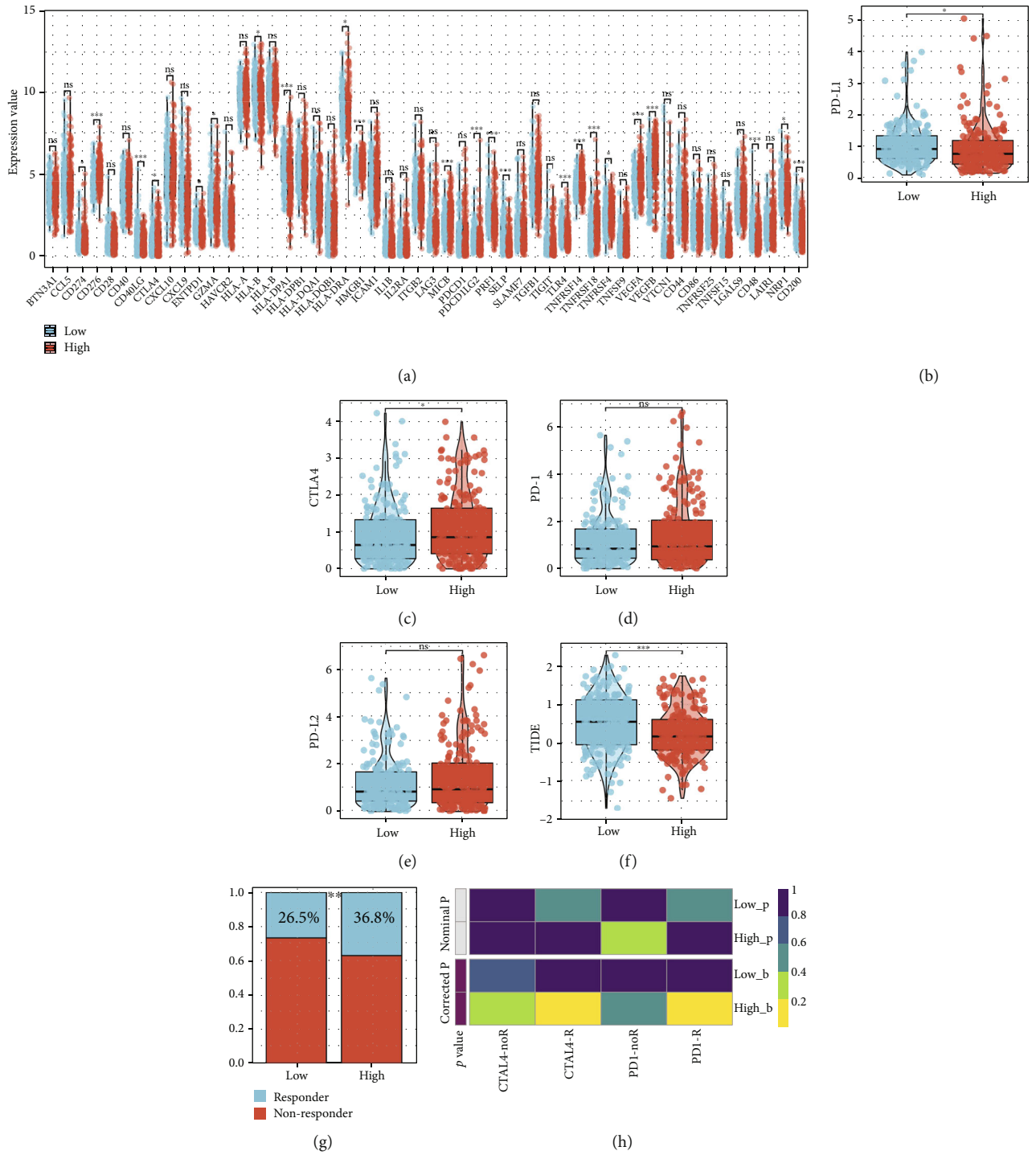


Figure 7: Continued.

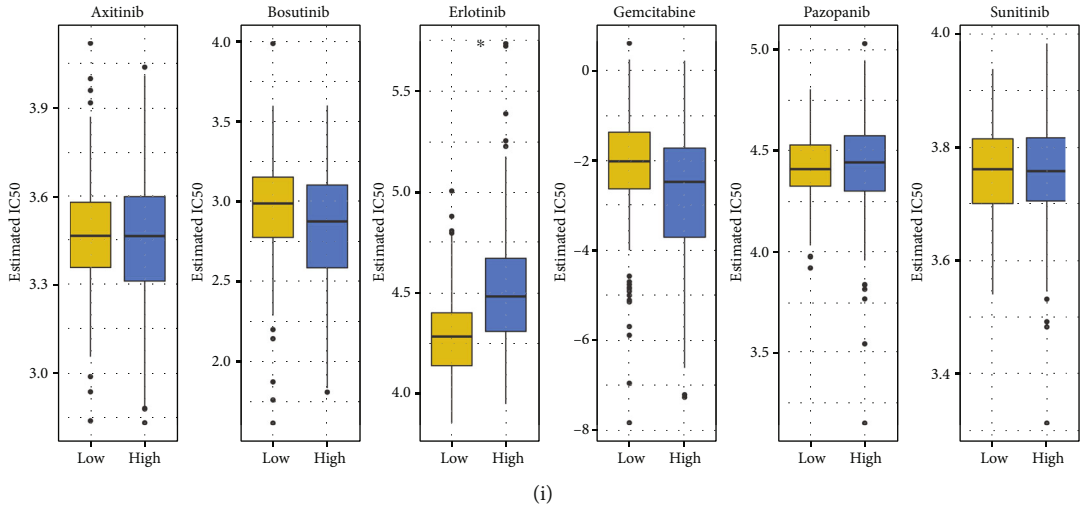


Figure 7: The risk score was associated with the sensibility to immunotherapy and chemotherapy. (a) Several disparities in immunological checkpoints between high- and low-risk groups; (b–e) variations in PD-L1, CTLA4, PD-1, and PD-L2 between high- and low-risk patients; (f, g) High-risk patients may have a lower TIDE score and a higher percentage of immunotherapy responders; (h) high-risk individuals may respond better to CTLA4 and PD-1 treatment; and (i) low-risk patients may respond better to Erlotinib.

comprehensive investigation of the liver cancer immune microenvironment can contribute to its diagnosis and therapy options.

In our study, we firstly quantified 53 immune terms in a combined cohort with large populations using the ssGSEA algorithm. Afterward, a prognostic model based on five immune elements (CORE.SERUM.RESPONSE.UP, angiogenesis, CD8.T.cells, Th2.cells, and B.cells) was established, which showed great prognosis prediction efficiency in both training and validation cohorts. Analysis of clinical correlations revealed that high-risk patients could possess higher clinical progression characteristics. Angiogenesis plays an important role in tumor metastasis and progression. From the comprehensive review conducted by Morse et al., abnormal angiogenesis in liver cancer often leads to poor prognosis and facilitated progression [26]. Wolf et al. found that the cross-talk between intrahepatic CD8+ T cells and NKT cells contributes to nonalcoholic steatohepatitis and liver cancer [27]. Xu et al. found that the Th2 response could affect liver fibrosis, which is a risk factor for liver cancer [28]. Garnelo et al. indicated that the interaction between tumor-infiltrating B cells and T cells significantly affects the progression of liver cancer [29].

Next, a pathway enrichment analysis was conducted to investigate the intrinsic biological variations between high- and low-risk patients. Analysis of genomic instability revealed that high-risk patients could possess a higher TMB score. The risk score showed a high positive correlation with the tumor stemness index. In addition, the TIDE outcome indicated that high-risk patients could be better responsive to immunotherapy, whereas low-risk patients could be better responsive to Erlotinib. Finally, six characteristic genes DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13 were identified for risk group prediction.

Pathway enrichment analysis showed that in high-risk patients, the Hallmark pathway of DNA repair, E2F targets, G2M checkpoints, Hedgehog signaling, mTORC1 signaling, and MYC targets were elevated. Malignant tumor cells have higher genomic damage compared to normal cells. Meanwhile, abnormal DNA repair procedures are difficult to meet the needs of DNA damage, therefore, leading to a more aggressive phenotype [30]. G2/M checkpoint is a critical phase in the cell cycle, which might directly affect cell proliferation [31]. Hedgehog signaling has been reported and extensively included in liver cancer progression. For instance, Gu et al. revealed that circular RNA circIPO11 could drive self-renewal and cancer progression of liver cancer through Hedgehog signaling, which might be an underlying therapeutic target [32]. In addition, Wu et al. discovered that CK2 might enhance stemness and chemotherapy resistance via the Hedgehog signaling pathway in liver cancer [33]. These findings suggested that patients in the high-risk category may have a higher level of activity in the aforementioned pathways, thus resulting in a poor prognosis.

Genomic analysis revealed that high-risk patients might have a greater level of genomic instability. In cancer, genomic instability is a prominent feature, which is also a key marker for separating cancerous cells from normal ones [34]. Genomic instability is defined as the increase in mutation frequency in the genome and its potential source is the sum of defects in the DNA damage and repair pathway, which could lead to uncontrolled proliferation of cancer cells [34]. In the pathway enrichment analysis, we have found the abnormal activity of DNA repair in high-risk patients. Therefore, higher genomic instability in high-risk patients seems to be reasonable. Moreover, a strong correlation was observed between the risk score and tumor stemness. Liver cancer has a high recurrence rate, making it one of the most highly deadly malignant tumors, partly due to the existence of cancer stem cells (CSC) [35]. Tumor stemness could also affect the immune microenvironment of liver cancer. For example, M2 TAMs could promote the expression of stemness proteins [36]. Another aspect, the CXCL11, a cytokine

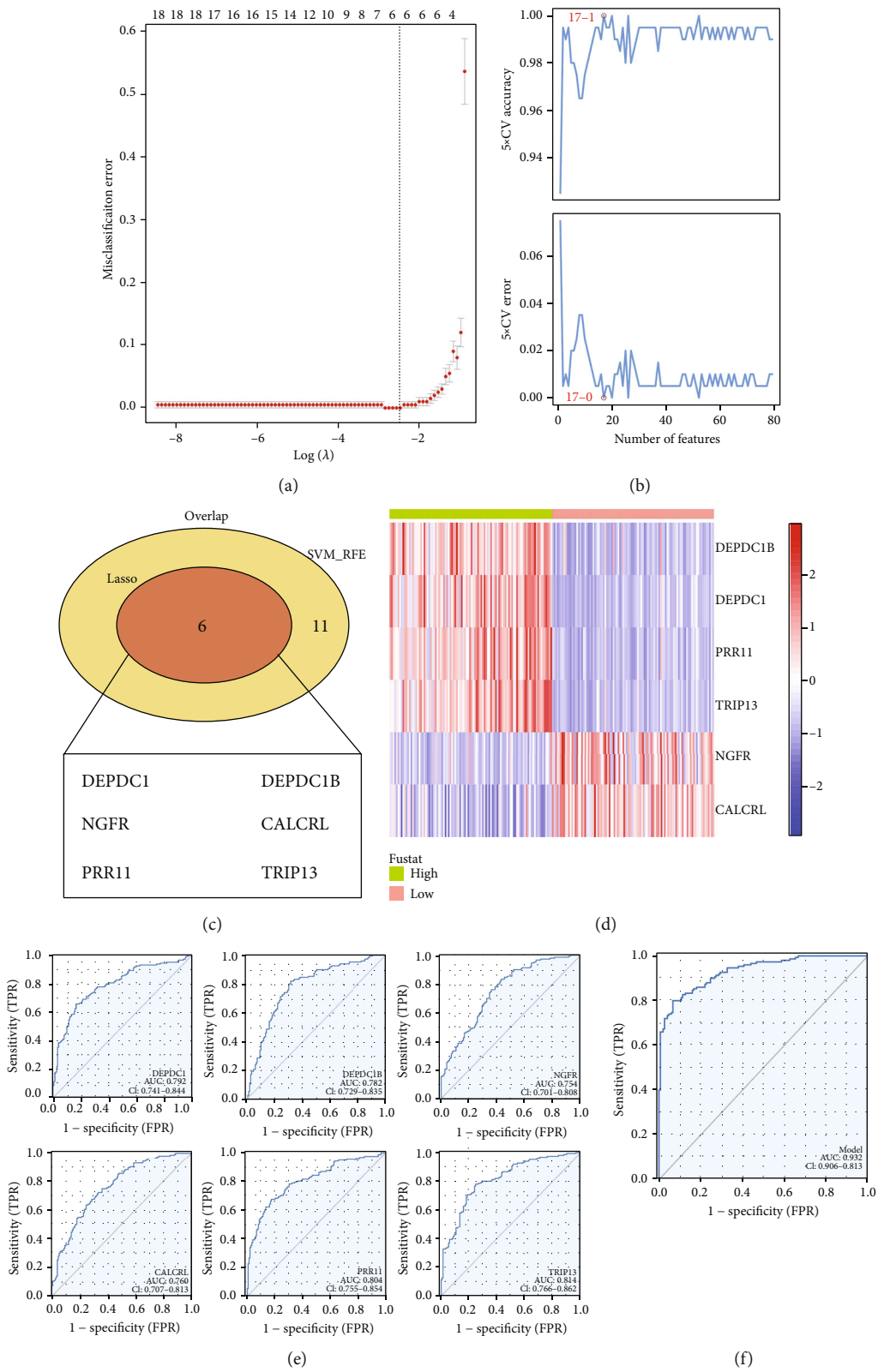
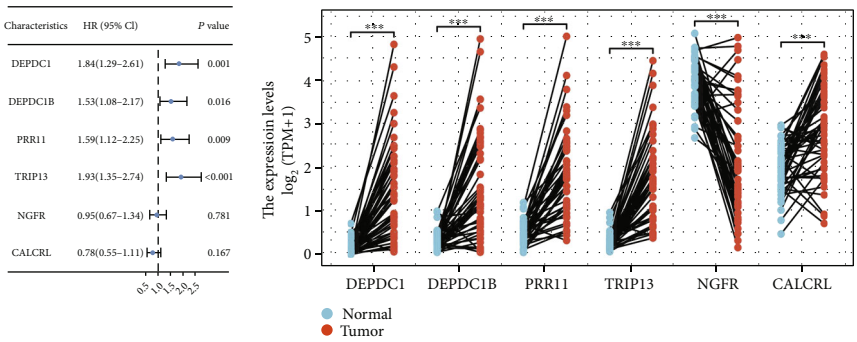


Figure 8: Continued.



(g) (h)

Figure 8: Identification of the characteristic genes of the risk group. Notes: (a) the LASSO regression algorithm was employed to determine the distinctive genes; (b) the SVM-RFE algorithm was utilized to determine the genes features; (c, d) DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13 were identified as the characteristic genes; (e) ROC curves were utilized to assess the prediction accuracy of characteristic genes on risk group (training cohort); (f) the logistic regression model showed an extremely high prediction efficiency in the risk group (training cohort); (g) the prognosis effect of DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11, and TRIP13; and (h) expression level of these six genes in tumor and normal tissue.

involved in the recruitment of activated T cells to inflammatory sites, was found significantly affect the stemness genes, sphere formation, and tumorigenicity in liver cancer [37].

Immunotherapy and chemotherapy are the best choices for the advanced liver cancer stage [4]. Low-risk patients could be higher responsive to Erlotinib, whereas high-risk patients could be more responsive to PD-1 and CTLA-4 treatment. According to the LASSO regression and SVM-RFE algorithm, we identified six characteristic genes, including DEPDC1, DEPDC1B, NGFR, CALCRL, PRR11 and TRIP13. The logistic regression model showed an extremely great prediction efficiency in the patients risk group. In the clinical application, measuring the expression level of these six genes might accurately identify the risk group of a patient, which might have the potential to guide the prognosis and therapy options of liver cancer patients.

Although our study was performed based on high-quality bioinformatics analysis, some limitations should be noticed. First, the populations included in our study were mainly White populations. Therefore, the underlying race bias might reduce the credibility of our conclusion. Second, some specific clinical features, for example, laboratory examination, lifestyle, and other basic diseases were not provided in most patients. However, our approach proved an excellent pattern for identifying liver cancer patients' survival time as well as their immunotherapy and chemotherapy responsiveness.

Data Availability

Based on reasonable request, data are available from the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Supplementary Materials

Supplementary 1. Figure S1: the prediction accuracy of characteristic genes on risk group in validation cohort. Notes: (a–g): ROC curves of characteristic genes.

Supplementary 2. Figure S2: Western blot of CALCRL.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] A. Marengo, C. Rosso, and E. Bugianesi, "Liver cancer: connections with obesity, fatty liver, and cirrhosis," *Annual Review of Medicine*, vol. 67, no. 1, pp. 103–117, 2016.

[3] A. Kanazawa, T. Tsukamoto, S. Shimizu, S. Yamamoto, A. Murata, and S. Kubo, "Laparoscopic hepatectomy for liver cancer," *Digestive Diseases*, vol. 33, no. 5, pp. 691–698, 2015.

[4] D. Anwanwan, S. K. Singh, S. Singh, V. Saikam, and R. Singh, "Challenges in liver cancer and possible treatment approaches," *Biochimica et Biophysica Acta Reviews on Cancer*, vol. 1873, no. 1, article 188314, 2020.

[5] G. Wang, Q. Wang, N. Liang et al., "Oncogenic driver genes and tumor microenvironment determine the type of liver cancer," *Cell Death & Disease*, vol. 11, no. 5, p. 313, 2020.

[6] X. Mao, J. Xu, W. Wang et al., "Crosstalk between cancer-associated fibroblasts and immune cells in the tumor microenvironment: new findings and future perspectives," *Molecular Cancer*, vol. 20, no. 1, p. 131, 2021.

[7] P. McGuirk and K. H. Mills, "Pathogen-specific regulatory T cells provoke a shift in the Th1/Th2 paradigm in immunity to infectious diseases," *Trends in Immunology*, vol. 23, no. 9, pp. 450–455, 2002.

[8] J. Kim and J. S. Bae, "Tumor-associated macrophages and neutrophils in tumor microenvironment," *Mediators of Inflammation*, vol. 2016, Article ID 6058147, 11 pages, 2016.

[9] S. Zhao, Y. Mi, B. Guan et al., "Tumor-derived exosomal mir-934 induces macrophage M2 polarization to promote liver metastasis of colorectal cancer," *Journal of Hematology & Oncology*, vol. 13, no. 1, p. 156, 2020.

[10] L. Xia, L. Oyang, J. Lin et al., "The cancer metabolic reprogramming and immune response," *Molecular Cancer*, vol. 20, no. 1, p. 28, 2021.

[11] H. J. Hurley, H. Dewald, Z. S. Rothkopf et al., "Frontline science: Ampk regulates metabolic reprogramming necessary for interferon production in human plasmacytoid dendritic cells," *Journal of Leukocyte Biology*, vol. 109, no. 2, pp. 299–308, 2021.

[12] S. Ricciardi, N. Manfrini, R. Alfieri et al., "The translational machinery of human $CD4^+$ T cells is poised for activation and controls the switch from quiescence to metabolic remodeling," *Cell Metabolism*, vol. 28, no. 6, pp. 895–906.e5, 2018.

[13] D. S. Chen and I. Mellman, "Elements of cancer immunity and the cancer-immune set point," *Nature*, vol. 541, no. 7637, pp. 321–330, 2017.

[14] S. Hänzelmann, R. Castelo, and J. Guinney, "GSVA: gene set variation analysis for microarray and Rna-Seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.

[15] X. Ren, X. Chen, X. Zhang et al., "Immune microenvironment and response in prostate cancer using large population cohorts," *Frontiers in Immunology*, vol. 12, article 686809, 2021.

[16] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.

[17] T. M. Malta, A. Sokolov, A. J. Gentles et al., "Machine learning identifies stemness features associated with oncogenic dedifferentiation," *Cell*, vol. 173, no. 2, pp. 338–54.e15, 2018.

[18] J. Fu, K. Li, W. Zhang et al., "Large-scale public data reuse to model immunotherapy response and resistance," *Genome Medicine*, vol. 12, no. 1, p. 21, 2020.

[19] W. Yang, J. Soares, P. Greninger et al., "Genomics of drug sensitivity in cancer (Gdsc): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Research*, vol. 41, pp. D955–D961, 2013.

[20] R. Y. Choi, A. S. Coyner, J. Kalpathy-Cramer, M. F. Chiang, and J. P. Campbell, "Introduction to machine learning, neural networks, and deep learning," *Translational Vision Science & Technology*, vol. 9, no. 2, p. 14, 2020.

[21] C. Tian, M. Abudoureyimu, X. Lin, X. Chu, and R. Wang, "Linc-Ror facilitates progression and angiogenesis of hepatocellular carcinoma by modulating Depdc1 expression," *Cell Death & Disease*, vol. 12, no. 11, p. 1047, 2021.

[22] X. W. Dang, Q. Pan, Z. H. Lin et al., "Overexpressed Depdc1b contributes to the progression of hepatocellular carcinoma by Cdk1," *Aging*, vol. 13, no. 16, pp. 20094–20115, 2021.

[23] W. Qiao, H. Wang, X. Zhang, and K. Luo, "Proline-rich protein 11 silencing inhibits hepatocellular carcinoma growth and epithelial-mesenchymal transition through *B*-catenin signaling," *Gene*, vol. 681, pp. 7–14, 2019.

[24] M. X. Zhu, C. Y. Wei, P. F. Zhang et al., "Elevated Trip13 drives the Akt/Mtor pathway to induce the progression of hepatocellular carcinoma via interacting with Actn4," *Journal of Experimental & Clinical Cancer Research: CR*, vol. 38, no. 1, p. 409, 2019.

[25] C. Ayuso, J. Rimola, R. Vilana et al., "Diagnosis and staging of hepatocellular carcinoma (Hcc): current guidelines," *European Journal of Radiology*, vol. 101, pp. 72–81, 2018.

[26] M. A. Morse, W. Sun, R. Kim et al., "The role of angiogenesis in hepatocellular carcinoma," *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, vol. 25, no. 3, pp. 912–920, 2019.

[27] M. J. Wolf, A. Adili, K. Piotrowitz et al., "Metabolic activation of intrahepatic Cd8+ T cells and Nkt cells causes nonalcoholic steatohepatitis and liver cancer via cross-talk with hepatocytes," *Cancer Cell*, vol. 26, no. 4, pp. 549–564, 2014.

[28] F. Xu, R. Cheng, S. Miao et al., "Prior Toxoplasma gondii infection ameliorates liver fibrosis induced by Schistosoma japonicum through inhibiting Th2 response and improving balance of intestinal flora in mice," *International Journal of Molecular Sciences*, vol. 21, no. 8, 2020.

[29] M. Garnelo, A. Tan, Z. Her et al., "Interaction between tumour-infiltrating B cells and T cells controls the progression of hepatocellular carcinoma," *Gut*, vol. 66, no. 2, pp. 342–351, 2017.

[30] A. K. Basu, "DNA damage, mutagenesis and cancer," *International Journal of Molecular Sciences*, vol. 19, no. 4, 2018.

[31] L. L. Liu, J. M. Zhu, X. N. Yu et al., "Ube2t promotes proliferation via G2/M checkpoint in hepatocellular carcinoma," *Cancer Management and Research*, vol. 11, pp. 8359–8370, 2019.

[32] Y. Gu, Y. Wang, L. He et al., "Circular Rna Circipo11 drives self-renewal of liver cancer initiating cells via Hedgehog signaling," *Molecular Cancer*, vol. 20, no. 1, p. 132, 2021.

[33] D. Wu, Y. Q. Yin, Y. Li, L. Zhang, Y. H. Jiang, and Z. Wang, "Ck2α causes stemness and chemotherapy resistance in liver cancer through the Hedgehog signaling pathway," *Hepatobiliary & Pancreatic Diseases International*, vol. 20, 2021.

[34] H. Li, S. E. Zimmerman, and U. Weyemi, "Genomic instability and metabolism in cancer," *International Review of Cell and Molecular Biology*, vol. 364, pp. 241–265, 2021.

[35] M. Huch and L. Dollé, "The plastic cellular states of liver cells: are Epcam and Lgr5 fit for purpose?," *Hepatology*, vol. 64, no. 2, pp. 652–662, 2016.

[36] F. Zhu, X. Li, S. Chen, Q. Zeng, Y. Zhao, and F. Luo, "Tumor-associated macrophage or chemokine ligand Ccl17 positively regulates the tumorigenesis of hepatocellular carcinoma," *Medical Oncology*, vol. 33, no. 2, p. 17, 2016.

[37] Y. Zhang, W. Zhao, S. Li et al., "CXCL11 promotes self-renewal and tumorigenicity of $\alpha 2\delta 1^+$ liver tumor-initiating cells through CXCR3/ERK1/2 signaling," *Cancer Letters*, vol. 449, pp. 163–171, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2023, Article ID 1553408, 14 pages
https://doi.org/10.1155/2023/1553408

Research Article

Prognostic Factors of Gliosarcoma in the Real World: A Retrospective Cohort Study

Ziye Yu,[1,2,3,4,5] Zhirui Zhou,[6] Ming Xu,[7] Kun Song,[1,2,3,4,5] Jingjing Shen,[7] Wenhao Zhu,[7] Liqun Wei,[7] and Hongzhi Xu[1,2,3,4,5]

[1]*Department of Neurosurgery, Huashan Hospital, Shanghai Medical College, Fudan University, 200040, China*
[2]*National Center for Neurological Disorders, 200040, China*
[3]*Shanghai Key Laboratory of Brain Function and Restoration and Neural Regeneration, 200040, China*
[4]*Neurosurgical Institute of Fudan University, 200040, China*
[5]*Shanghai Clinical Medical Center of Neurosurgery, 200040, China*
[6]*Department of Radiotherapy, Huashan Hospital, Fudan University, 200040, China*
[7]*Department of Anesthesiology, Huashan Hospital, Fudan University, 200040, China*

Correspondence should be addressed to Hongzhi Xu; xuhongzhi95@sina.com

Received 20 September 2022; Revised 25 October 2022; Accepted 24 November 2022; Published 30 January 2023

Academic Editor: Jincheng Wang

Purpose. Gliosarcoma is a histopathological variant of glioblastoma, which is characterized by a biphasic growth pattern consisting of glial and sarcoma components. Owing to its scarcity, data regarding the impact of available treatments on the clinical outcomes of gliosarcoma are inadequate. The purpose of this retrospective cohort study was to analyze the prognostic factors of gliosarcoma. *Methods*. By screening the clinical database of neurosurgical cases at a single center, patients with gliosarcoma diagnosed histologically from 2013 to 2021 were identified. Clinical, pathological, and molecular data were gathered founded on medical records and follow-up interviews. Prognostic factors were derived using the Cox proportional hazards model with backward stepwise regression analysis. *Results*. Forty-five GSM patients were included. Median overall survival was 25.6 months (95% CI 8.0–43.1), and median relapse-free survival was 15.2 months (95% CI 9.7–20.8). In multivariable analysis, total resection ($p = 0.023$, HR = 0.192, 95% CI 0.046–0.797) indicated an improved prognosis. And low expression of Ki-67 ($p = 0.059$, HR = 2.803, 95% CI 0.963–8.162) would be likely to show statistical significance. However, there might be no statistically significant survival benefit from radiotherapy with concurrent temozolomide ($n = 33$, 73.3%, log-rank $p = 0.99$) or adjuvant temozolomide ($n = 32$, 71.1%, log-rank $p = 0.74$). *Conclusion*. This single-center retrospective study with a limited cohort size has demonstrated the treatment of gross total resection and low expression of Ki-67 which are beneficial for patients with GSM, while radiotherapy or temozolomide is not.

1. Introduction

Gliosarcoma (GSM) is a rare malignant tumor of the central nervous system (CNS) and has been classified as glioblastoma (GBM) since the publication of 2016 World Health Organization Classification of Tumors of the Central Nervous System. It is considered as a histopathological variant of GBM, as well as epitheliod GBM and giant cell GBM, accounting for about 2-4% of all cases [1, 2]. GSM presents unique histopathological manifestations characterized by a biphasic growth model of glial and sarcomatous elements. The glioma components often exhibit the typical characteristics of glioblastoma and have different degrees of anaplasia and glial fibrillary acidic protein (GFAP) expression. Meanwhile, the sarcomatoid area appears microscopically which shows dense long-spindle cells arranged within a fishbone fibrosarcoma structure and occasionally malignant fibrous histiocytomas. The key points of diagnosis are the reticular fibres in sarcoma and GFAP expression in glioma [3]. GSM has similar radiological and clinical representations to GBM but a comparatively poorer prognosis [4–7]. Several researches have reported the median overall survival (OS) of

gliosarcoma ranging from 6.6 to 18.5 months [8–11]. At present, special therapies for GSM are virgin, and treatments still cannot exceed the limits of GBM guidelines, including maximum surgical resection, radiotherapy (RT), and temozolomide (TMZ) [12].

Data regarding the impact of excision extension and postoperative adjuvant therapy on GSM outcomes are insufficient. Although gross total resection (GTR) of GBM independent of adjuvant therapy has been associated with progression-free survival (PFS) improvement, this has not yet been confirmed in GSM [11]. Some studies have elucidated a potential a possible benefit of RT and TMZ for GSM [10, 13]. However, there is still considerable controversy due to the lack of forward-looking clinical evidence [14, 15]. Distinctive histopathological properties may be associated with the differential therapeutic susceptibility between GSM and GBM. In addition, age, tumor size [16], and the diagnosis of primary or secondary [17] have also been reported being associated with GSM prognosis. There are still questions: what type of patient features indicates a better prognosis, and what kind of treatments can benefit GSM patients.

Here, we performed a retrospective cohort study of 45 GSM patients at single center. The purposes of our study were to analyze the independent prognostic factors of GSM and to understand the efficacy of classical treatments on survival outcomes.

2. Materials and Methods

2.1. Patients. From November 2013 until May 2021, patients with pathologically authenticated GSM were scanned from the medical record database of our hospital. Below are other criteria for inclusion. Patients were treated with at least one surgical resection of GSM, including the primary operation and reoperation. In accordance with the safety principle, the focus and surrounding brain tissue should be removed as much as possible. Patients had undergone at least one postoperative cranial magnetic resonance (MR) or computerized tomography (CT) examination, which was compared with the preoperative imaging data to clarify the location and integrity of resection (Figures 1(a) and 1(b)). The GSM was diagnosed according to the biphasic growth pattern of hematoxylin-eosin (HE) staining as well as GFAP staining demonstrating GFAP-positive glioma components and GFAP-negative sarcoma components containing tumor spindle cells (Figures 1(c)–1(f)). The pathological specimens before the release of 2016 World Health Organization Classification of Tumors of the Central Nervous System were reviewed and confirmed by the pathologist. Radiologic and pathologic findings were read and examined by specialists who did not know patient information following the principle of blindness.

During the course of clinical data collection, we found that data collected since 2013 was relatively intact due to our electoral medical record system. Patients were routinely followed up every three months until May 2022. Meanwhile, there was no notable change in the treatment of GSM during this period. We excluded the patients with uncertain pathological diagnosis or severely incomplete clinical data. Following the identification of the cohort, the clinical records of patients in the cohort were reviewed carefully, including disease-specific demographic, clinical, and treatment characteristics.

2.2. Data Compilation. Clinical or treatment-related variables included age, sex, Karnofsky performance score (KPS), diagnosis of primary or secondary, single or multiple lesions, tumor location, preoperative size, epileptic seizure, intracranial hypertension, Ki-67 level, genetic mutation (e.g., IDH1, p53, PTEN, MGMT, 1p/19q codeletion, TERT, BRAF, PIK3CA, ATRX, and EGFR), and surgical resection extension, adjuvant therapies (RT, concurrent TMZ, adjuvant TMZ, and targeted therapy). The conditions of postoperative complications were difficult to collect and were not included in the analysis due to the high scarcity and the poor reliability. Tumor-specific characteristics were acquired from the initial MR or CT examinations. Typically, early imaging examinations provided a generalized diagnosis such as glioma instead of GSM. The final diagnosis was pathologically determined during subsequent surgical treatment.

2.3. Outcomes. We defined the first diagnosis date as the day of the first neurosurgery for GSM confirmed by postoperative pathology, and the OS was calculated from this day as the starting point.

In 8 cases of secondary GSM, all of these patients had explicit primary and secondary pathological diagnosis. The primary diagnosis was GBM (WHO IV) in 5 cases and non-GBM (WHO II-III) in 3 cases. The location of the secondary GSM was identical to the primary tumor, and the time interval between two onsets was no more than two years. According to 2016 World Health Organization Classification of Tumors of the Central Nervous System, these patients can be clearly defined as secondary GSM. Before the first diagnosis of GSM, all secondary GSM patients in this cohort had received neurosurgery and at least one adjuvant treatment. The diagnosis of primary and secondary is worth being included in the Cox proportional hazards model.

Another survival analysis index was relapse-free survival (RFS). For either primary or secondary GSM after tumor resection, relapse of the disease was defined as the date of radiographic recurrence necessitating either secondary surgical intervention or adjuvant therapy. RFS is defined from the date of the first resection of GSM lesions until the date of the first indication of tumor recurrence during the imaging examination. Clinical manifestations, multiple imaging examinations, and secondary postoperative pathology identified and excluded the possibility of tumor pseudoprogression in recurrent patients.

2.4. Statistical Analysis. The Cox regression model was used for univariable analysis, and the hazard ratio (HR) and its 95% confidence interval (95% CI) were estimated. Kaplan-Meier survival curves were plotted according to different prognostic factors and were analyzed using log-rank test. Then, multivariate analysis recruiting proper variables was performed by the backward regression method to confirm

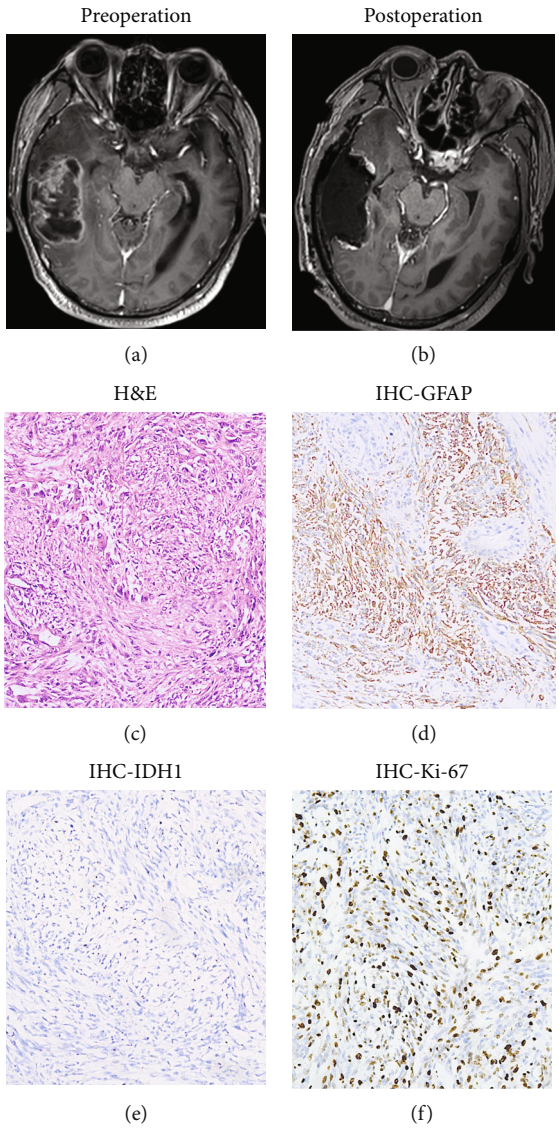


Figure 1: (a, b) Horizontal T1-weighted contrast-enhanced MRI imaging of GSM located in the temporal lobe. (c–f) Representative images of GSM stained using HE and immunolabeled using antibodies to GFAP, IDH1, and Ki-67. Through HE staining (c) and GFAP (d) staining, it can be observed in the biphasic growth pattern of GFAP-positive glioma components and GFAP-negative sarcoma components.

the statistical significance of prognostic factors. Calculations and graphical representations were performed using the *R* software package (version 4.0.3, The *R* Foundation for Statistical Computing).

3. Results

3.1. Cohort Characteristics. A total of 45 patients with histopathologically confirmed GSM between 2013 and 2021 were included (Table 1). As of the last follow-up, 30 patients (66.7%) had relapsed from the first GSM resection operation. Among them, 24 patients (53.3%) had passed away. In 15 patients (33.3%), no tumor recurrence was observed during postoperative imaging reexamination. In the cohort, 26 cases were medial and young (age<60, 57.8%), and 32 cases were males (71.1%). Most patients were diagnosed with primary GSM ($n = 37$, 82.2%), with the remainder being secondary GSM transformed from other types of glioma ($n = 8$, 17.8%). Median preoperative KPS was 80 (range 20-100). Preoperatively, 5 patients (11.1%) had epileptic seizures, and 36 patients (80%) suffered from ventricular enlargement and intracranial hypertension caused by tumor compression. Almost all but one patient had a single lesion ($n = 44$, 97.8%), primarily in temporal ($n = 19$, 42.2%) and frontal ($n = 12$, 26.7%) lobes. Other tumor locations included parietal lobe ($n = 5$, 11.1%), basal ganglia ($n = 4$, 8.9%), callosum ($n = 3$, 6.7%), thalamus ($n = 1$, 2.2%), and brainstem ($n = 1$, 2.2%). In terms of laterality, the left side ($n = 20$, 44.4%) was a little more than the right side ($n = 16$, 35.6%), and others were located deep in the brain ($n = 9$, 20.0%). Most tumors had a maximum diameter greater than 3 centimeters ($n = 32$, 71.1%).

The majority successfully underwent GTR ($n = 41$, 91.1%), and the remaining patients underwent subtotal

Table 1: Clinical characteristics of 45 GSM cases.

(a)

Characteristics	Number of patients ($n = 45$)	Proportion (%)
Relapse	30	66.7
Death	24	53.3
Age		
Average age, years	53.2	
Median age, years	56	
<60	26	57.8
≥60	19	42.2
Sex		
Male	32	71.1
Female	13	28.9
KPS		
≤70	15	33.3
>70	30	66.7
Primary or secondary		
Primary GSM	37	82.2
Secondary GSM	8	17.8
GBM-origined	5	11.1
Non-GBM-origined	3	6.6
Number of lesions		
Single	44	97.8
Multiple	1	2.2
Size, diameter		
≤3 cm	13	28.9
>3 cm	32	71.1
Location		
Temporal	19	42.2
Frontal	12	26.7
Parietal	5	11.1
Basal ganglia	4	8.9
Callosum	3	6.7
Thalamus	1	2.2
Brainstem	1	2.2
Laterality		
Left	20	44.4
Right	16	35.6
Profound	9	20.0
Epileptic seizure		
No	40	88.9
Yes	5	11.1
Intracranial hypertension		
No	9	20.0
Yes	36	80.0
Level of Ki-67		
≤15%	15	33.3
>15%	30	66.7
Resection range		
GTR	41	91.1
STR/NTR	4	8.9
Biopsy	0	0
Radiotherapy		
No	11	24.4
Yes	33	73.3
Missing	1	2.2
Dose, mean (SD)	57.97 (2.75)	
Concurrent TMZ		
No	11	24.4
Yes	33	73.3
Missing	1	2.2
Adjuvant TMZ		
No	12	26.7
Yes	32	71.1
Missing	1	2.2

(b)

Gene locus	Wild-type (%)	Mutated (%)	Missing
IDH1	40 (100.0)	0 (0.0)	5
p53	14 (33.3)	28 (66.7)	3
PTEN	5 (55.6)	4 (44.4)	36
MGMT	11 (64.7)	6 (35.3)	28
1p/19q codeletion	10 (66.7)	5 (33.3)	30
TERT	5 (27.8)	13 (72.2)	27
BRAF	19 (90.5)	2 (9.5)	24
PIK3CA	9 (81.8)	2 (18.2)	34
ATRX	8 (21.6)	29 (78.4)	8
EGFR	8 (80.0)	2 (20.0)	35

resection (STR) and near total resection (NTR), based on the detailed surgical records and postoperative imaging data of each patient. Benefit from the progress of surgical technology and strictly observing surgical indications, the complete resection rate of malignant brain tumors increases gradually in our medical center, and some deep brain tumors can also be totally removed.

Except for one perioperative death, all cases had received at least one adjuvant therapy within the specified time, mainly RT ($n = 33$, 73.3%), concurrent TMZ ($n = 33$, 73.3%), and adjuvant TMZ ($n = 32$, 71.1%). This is the reliable information from clinical data. Some patients ($n = 11$) only accomplished adjuvant TMZ treatment because of terrible postoperative physical condition or poor tolerance to radiotherapy. Coincidentally, all patients receiving RT in this study took TMZ simultaneously, and part of them ($n = 21$) also received adjuvant TMZ. A few patients were also treated with nimustine ($n = 4$, 8.8%) and bevacizumab ($n = 2$, 4.4%). Some recidivist patients received reoperation ($n = 12$, 40%) or secondary

radiotherapy ($n = 7$, 23.3%), and five of them received both of them.

In some patients, gene mutations were detected through next generation sequencing (NGS) and/or immunohistochemistry (IHC). The results are listed in Table 1.

3.2. Median OS and RFS. After discharge, patients were systematically reevaluated and reexamined every three months. Most patients were given a long-term follow-up of more than 2 years, and the minimum follow-up period was 4 quarters. Survival data for 45 patients was clarified by April 2022, and the date of death and recurrence was precisely obtained. The median OS and the median RFS were calculated via the Kaplan-Meier curve. In this cohort, the median OS was 25.6 months (95% CI 8.0–43.1) (Figure 2(a)), and the median RFS was 15.2 months (95% CI 9.7–20.8) (Figure 2(b)).

3.3. Characteristics Associated with OS. Potential prognostic factors were analyzed using the Cox proportional hazards model (Table 2). In backward regression multivariable analysis excluding the variables of all gene mutations but p53 due to their incomplete data, total resection ($p = 0.023$, HR = 0.192, 95% CI 0.046–0.797) was significantly associated with longer OS, indicating a better prognosis of GSM patients. Moreover, expression level of Ki-67 ($p = 0.059$, HR = 2.803, 95% CI 0.963–8.162) would probably be statistically significant if the sample size increased (Table 2).

However, none of RT with concurrent TMZ ($p = 0.991$, HR = 0.995, 95% CI 0.389–2.546) and adjuvant TMZ ($p = 0.742$, HR = 0.860, 95% CI 0.351–2.110) indicated a statistical significance either in univariable and multivariable analysis (Table 2).

All patients with secondary GSM ($n = 8$) received postoperative RT or TMZ after the first operation. Of them, 4 patients received renewedly radiotherapy and concurrent TMZ after the second operation, and others also accepted adjuvant TMZ treatment. Compared with patients with primary GSM, there was no significant difference in patients with secondary GSM ($p = 0.749$, HR = 0.850, 95% CI 0.315–2.298), whether for GBM-origined ($p = 0.317$, HR = 1.742, 95% CI 0.588–5.166) or non-GBM-origined ($p = 0.507$, HR = 0.503, 95% CI 0.066–3.827) (Table 2).

Kaplan-Meier curves have been drawn, respectively, for interested factors, respectively. None of them demonstrated a significant difference. Median OS was 29.7 months (95% CI 11.5–47.9) in the total resection group ($n = 41$, 91.1%) and 4.8 months (95% CI 0.0–18.1) in the nontotal resection group ($n = 4$, 8.9%), $p = 0.11$ (Figure 2(c)). Median OS was 17.0 months (95% CI 8.0–43.1) in the Ki-67 high expression group ($Ki - 67 > 15\%$, $n = 30$, 66.7%) and cannot be calculated in the Ki-67 low expression group ($\text{Ki} - 67 \leq 15\%$, $n = 15$, 33.3%) because the median OS is unreached, $p = 0.12$ (Figure 2(d)). Median OS was 25.6 months (95% CI 9.2–42.0) in the RT with the concurrent TMZ group ($n = 33$, 73.3%) and 41.0 months (95% CI 0.0–82.1) in the non-RT with the concurrent TMZ group ($n = 11$, 24.4%), $p = 0.99$ (Figure 3(a)) and 29.7 months (95% CI 4.1–55.3) in the adjuvant TMZ group ($n = 32$, 71.1%) and 25.6 months (95% CI 0.0–54.3) in the nonadjuvant TMZ group ($n = 12$, 26.7%), $p = 0.74$ (Figure 3(b)).

Median OS was 25.6 months (95% CI 9.2–42.0) in the RT with the concurrent TMZ group ($n = 33$, 73.3%) and 41.0 months (95% CI 0.0–82.1) in the non-RT with the concurrent TMZ group ($n = 11$, 24.4%), $p = 0.99$ (Figure 3(a)) and 29.7 months (95% CI 4.1–55.3) in the adjuvant TMZ group ($n = 32$, 71.1%) and 25.6 months (95% CI 0.0–54.3) in the nonadjuvant TMZ group ($n = 12$, 26.7%), $p = 0.74$ (Figure 3(b)).

In addition, according to the results of NGS and/or immunohistochemistry (IHC), the fact whether some gene loci were mutated was used as a prognostic factor for survival analysis, including p53 (negative, $n = 14$, 33.3% vs. positive, $n = 28$, 66.7%, log-rank $p = 0.81$), ATRX (negative, $n = 8$, 21.6% vs. positive, $n = 29$, 88.4%, log-rank $p = 0.16$), BRAF (negative, $n = 19$, 90.5% vs. positive, $n = 2$, 9.5%, log-rank $p = 0.27$), TERT (negative, $n = 5$, 27.8% vs. positive, $n = 13$, 72.2%, log-rank $p = 0.025$), MGMT methylated (negative, $n = 11$, 64.7% vs. positive, $n = 6$, 35.3%, log-rank $p = 0.95$), 1p/19q codeletion (negative, $n = 10$, 66.7% vs. positive, $n = 5$, 33.3%, log-rank $p = 0.37$), PIK3CA (negative, $n = 9$, 81.8% vs. positive, $n = 2$, 18.2%, log-rank $p = 0.47$), EGFR (negative, $n = 8$, 80.0% vs. positive, $n = 2$, 20.0%, log-rank $p = 0.97$), and PTEN (negative, $n = 5$, 21.6% vs. positive, $n = 4$, 88.4%, log-rank $p = 0.81$).

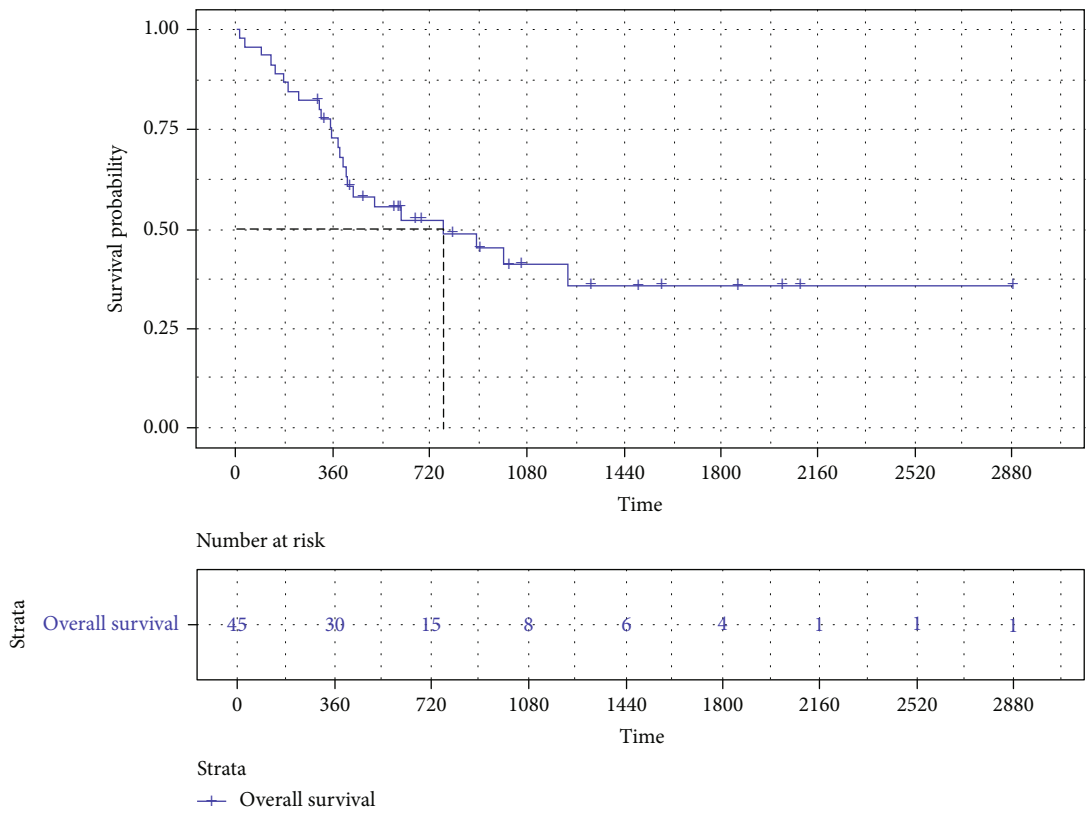
3.4. Characteristics Associated with RFS. We carried out multivariable analysis to calculate RFS using the same method and variables. Similar to OS benefits, the results showed that the variable of total resection ($p = 0.022$, HR = 0.181, 95% CI 0.042–0.782) and low expression of Ki-67 ($p = 0.052$, HR = 2.996, 95% CI 0.992–9.053) were helpful to improve RFS (Table 2).

Similarly, RT with concurrent TMZ ($p = 0.861$, HR = 0.919, 95% CI 0.360–2.348) and adjuvant TMZ ($p = 0.631$, HR = 0.804, 95% CI 0.330–1.959) did not demonstrate RFS benefits either in univariable and multivariable analysis (Table 2).

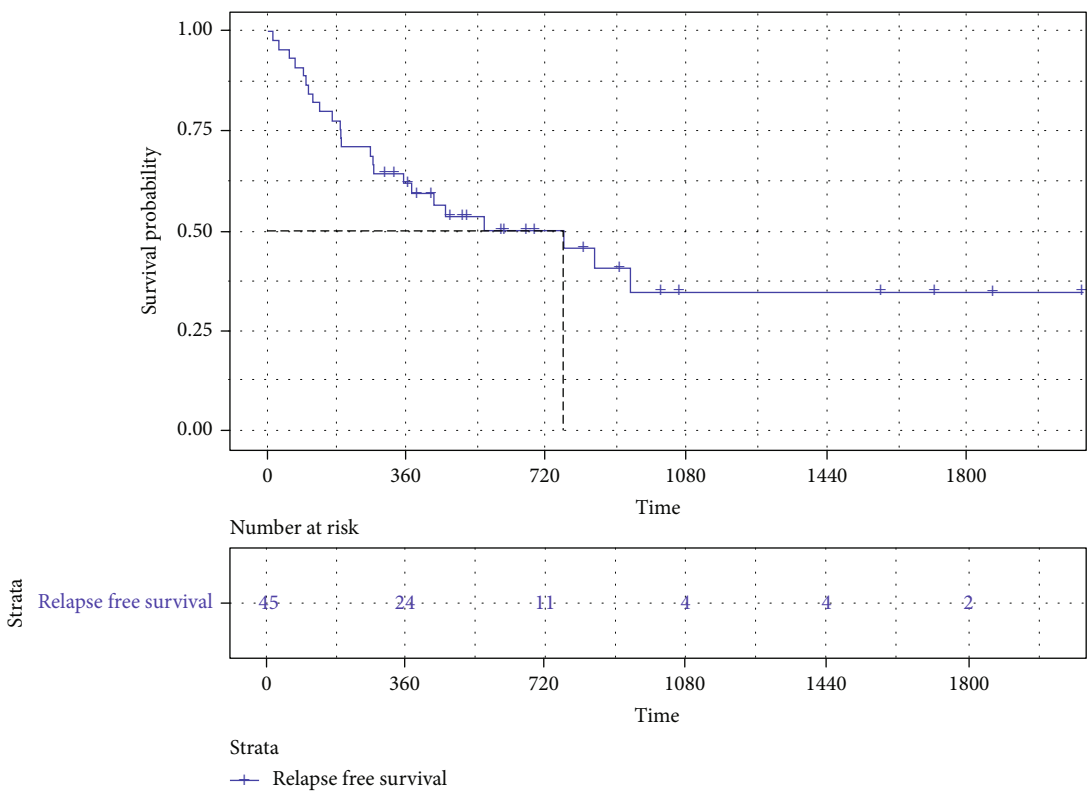
There was no distinct difference between the primary group and the secondary group ($p = 0.715$, HR = 1.202, 95% CI 0.448–3.228), whether for GBM-origined ($p = 0.417$, HR = 1.566, 95% CI 0.529–4.636) or non-GBM-origined ($p = 0.646$, HR = 0.624, 95% CI 0.083–4.684) (Table 2).

Kaplan-Meier curves were drawn, respectively, for interested factors, and none of them were considered significant that could affect RFS. Median RFS was 25.3 months (95% CI 9.5–41.1) in the total resection group ($n = 41$, 91.1%) and 2.3 months (95% CI 0.0–13.9) in the nontotal resection group ($n = 4$, 8.9%), $p = 0.12$ (Figure 2(e)). Median RFS was 18.5 months (95% CI 6.0–31.0) in the Ki-67 high expression group ($\text{Ki} - 67 > 15\%$, $n = 30$, 66.7%) and cannot be counted in the Ki-67 low expression group ($\text{Ki} - 67 \leq 15\%$, $n = 15$, 33.3%) because the death toll is less than half, $p = 0.12$ (Figure 2(f)).

Median RFS was 25.3 months (95% CI 10.4–40.2) in the RT with the concurrent TMZ group ($n = 33$, 73.3%) and 31.1 months (95% CI 0.0–64.1) in the non-RT with the concurrent TMZ group ($n = 11$, 24.4%), $p = 0.86$ (Figure 3(c)) and

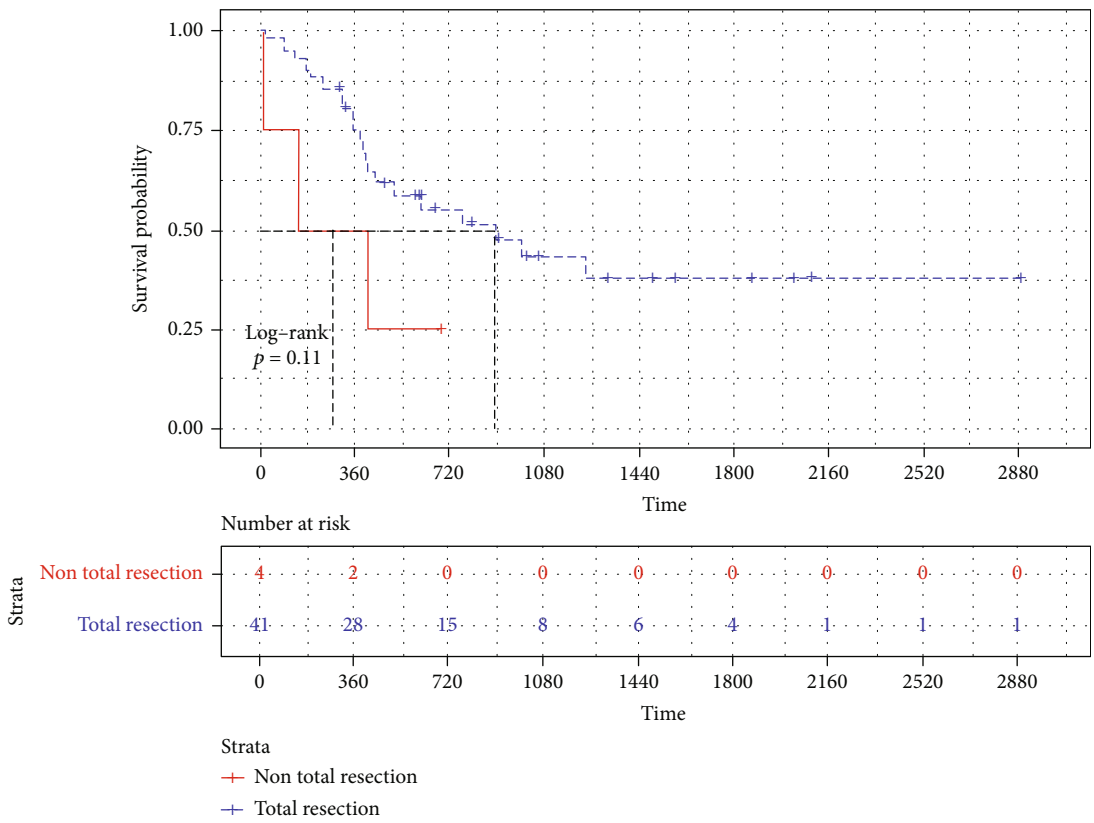


(a)

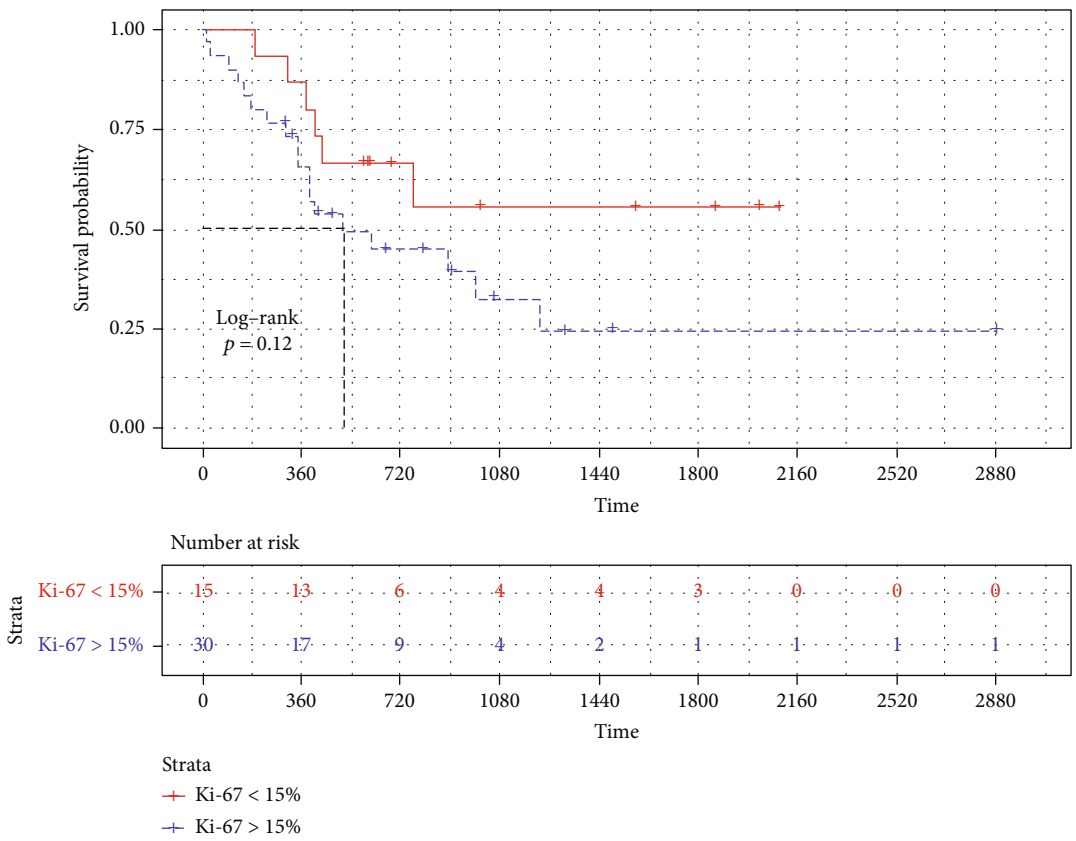


(b)

Figure 2: Continued.

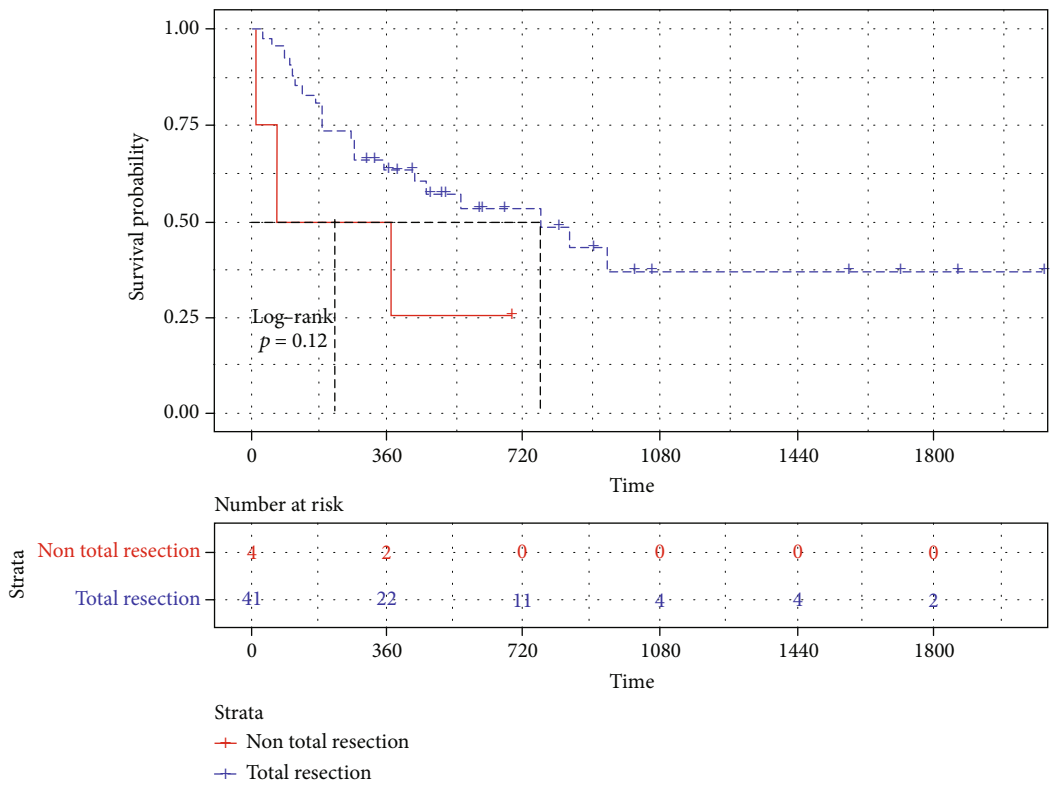


(c)

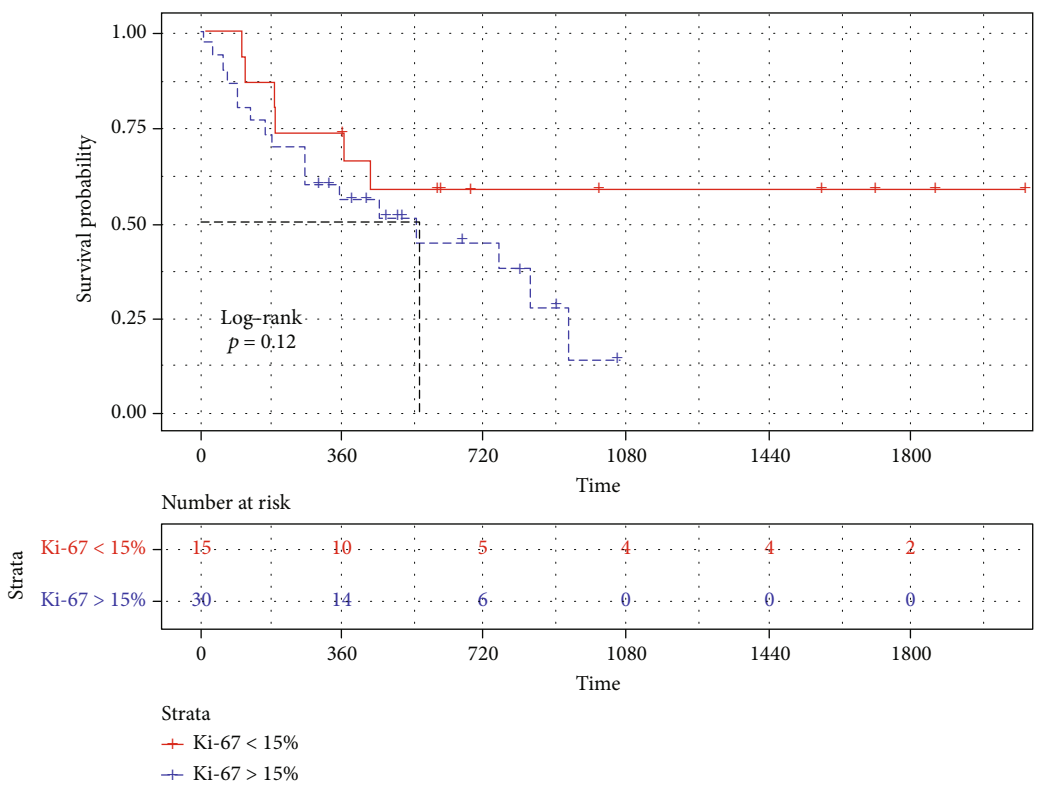


(d)

Figure 2: Continued.



(e)



(f)

Figure 2: Kaplan-Meier curves for OS (a) and RFS (b) of all patients included in our cohort. (c) Kaplan-Meier curves and log-rank tests according to OS were used for analysis of prognostic factors. According to resection range, they were divided into total resection group and nontotal resection group. (d) According to the expression level of Ki-67, they were divided into high expression group and low expression group. (e, f) Kaplan-Meier curves and log-rank tests according to RFS were used for analysis of prognostic factors.

Table 2: Univariable and multivariable analysis according to OS and RFS.

(a)

Variable name	Univariable analysis, OS		Multivariable analysis, OS	
	p value	HR (95% CI)	*p* value	HR (95% CI)
Age	0.732	1.151 (0.515-2.572)		
Sex	0.358	0.678 (0.296-1.553)		
Primary or secondary	0.749	0.850 (0.315-2.298)		
Primary vs. GBM-origined	0.317	1.742 (0.588–5.166)		
Primary vs. non-GBM-origined	0.507	0.503 (0.066-3.827)		
KPS	0.847	0.919 (0.392–2.157)		
Location	0.548			
Frontal vs. deep	0.195	0.452 (0.136-1.500)		
Parietal vs. deep	0.989	0.990 (0.244-4.024)		
Temporal vs. deep	0.364	0.624 (0.226-1.728)		
Size	0.334	0.657 (0.280-1.542)		
Number of lesions	0.036	0.095 (0.011-0.853)		
Epileptic seizure	0.684	0.739 (0.172-3.170)		
Intracranial hypertension	0.377	0.639 (0.237-1.725)		
Ki-67 level	0.132	2.039 (0.807-5.151)	0.059	2.803 (0.963-8.162)
Resection range	0.125	0.382 (0.111-1.306)	0.023	0.192 (0.046-0.797)
Radiotherapy	0.991	0.995 (0.389-2.546)		
Concurrent TMZ	0.991	0.995 (0.389-2.546)		
Adjuvant TMZ	0.742	0.860 (0.351-2.110)		
Gene mutation				
IDH1	—	—		
p53	0.807	0.893 (0.360-2.216)		
PTEN	0.809	1.221 (0.243-6.122)		
MGMT	0.953	0.950 (0.173-5.229)		
1p/19q codeletion	0.381	2.057 (0.410-10.318)		
TERT	0.209	47.014 (0.116-18994)		
BRAF	0.288	2.436 (0.472-12.569)		
PIK3CA	0.483	2.262 (0.232-22.095)		
ATRX	0.165	0.483 (0.173-1.349)		
EGFR	0.975	0.966 (0.111-8.415)		

(b)

Variable name	Univariable analysis, RFS		Multivariable analysis, RFS	
	p value	HR (95% CI)	*p* value	HR (95% CI)
Age	0.700	1.171 (0.524-2.616)		
Sex	0.568	0.784 (0.341-1.805)		
Primary or secondary	0.715	1.202 (0.448-3.228)		
Primary vs. GBM-origined	0.417	1.566 (0.529-4.636)		
Primary vs. non-GBM-origined	0.646	0.624 (0.083-4.684)		
KPS	0.800	0.896 (0.383-2.096)		
Location	0.708			
Frontal vs. deep	0.367	0.578 (0.176-1.900)		
Parietal vs. deep	0.995	1.004 (0.250-4.041)		
Temporal vs. deep	0.424	0.660 (0.238-1.831)		
Size	0.426	0.706 (0.300-1.663)		

Table 2: Continued.

Variable name	Univariable analysis, RFS		Multivariable analysis, RFS	
	p value	HR (95% CI)	*p* value	HR (95% CI)
Number of lesions	0.022	0.071 (0.007-0.679)		
Epileptic seizure	0.857	0.875 (0.205-3.731)		
Intracranial hypertension	0.513	0.714 (0.260-1.958)		
Ki-67 level	0.130	2.066 (0.809-5.278)	0.052	2.996 (0.992-9.053)
Resection range	0.137	0.395 (0.116-1.345)	0.022	0.181 (0.042-0.782)
Radiotherapy	0.861	0.919 (0.360-2.348)		
Concurrent TMZ	0.861	0.919 (0.360-2.348)		
Adjuvant TMZ	0.631	0.804 (0.330-1.959)		
Gene mutation				
IDH1	—	—		
p53	0.995	0.997 (0.402-2.474)		
PTEN	0.745	1.307 (0.261-6.550)		
MGMT	0.992	1.009 (0.184-5.527)		
1p/19q codeletion	0.433	1.900 (0.382-9.440)		
TERT	0.202	49.123 (0.125-19.370)		
BRAF	0.270	2.529 (0.487-13.134)		
PIK3CA	0.838	1.252 (0.145-10.817)		
ATRX	0.298	0.581 (0.209-1.617)		
EGFR	0.951	1.070 (0.124-9.249)		

25.3 months (95% CI 9.7–40.9) in the adjuvant TMZ group ($n = 32$, 71.1%) and 14.3 months (95% CI 0.0–35.2) in the nonadjuvant TMZ group ($n = 12$, 26.7%), $p = 0.63$ (Figure 3(d)).

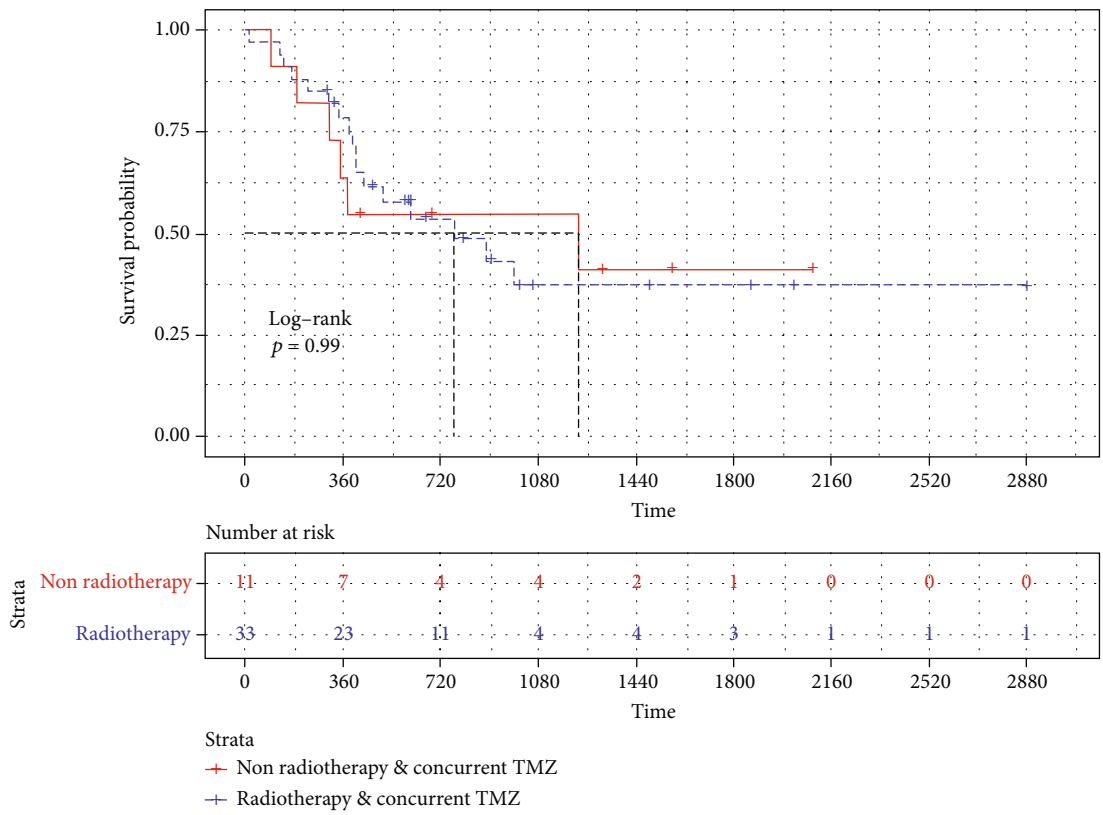
4. Discussion

As a histological variant of GBM, GSM has some different prognostic factors. Reviewing the clinical data of the cohort, the people with high incidence of GSM are concentrated around the age of 60 ($50 \leq \text{age} < 70$, $n = 29$, 64.4%), mainly male ($n = 32$, 71.1%) and showing a predilection for temporal lobe ($n = 19$, 42.2%). The situation is generally consistent with the previous research results [17–22]. In our cohort, the proportion of patients with secondary GSM is 17.8%, which is still consistent with existing reports [22, 23].

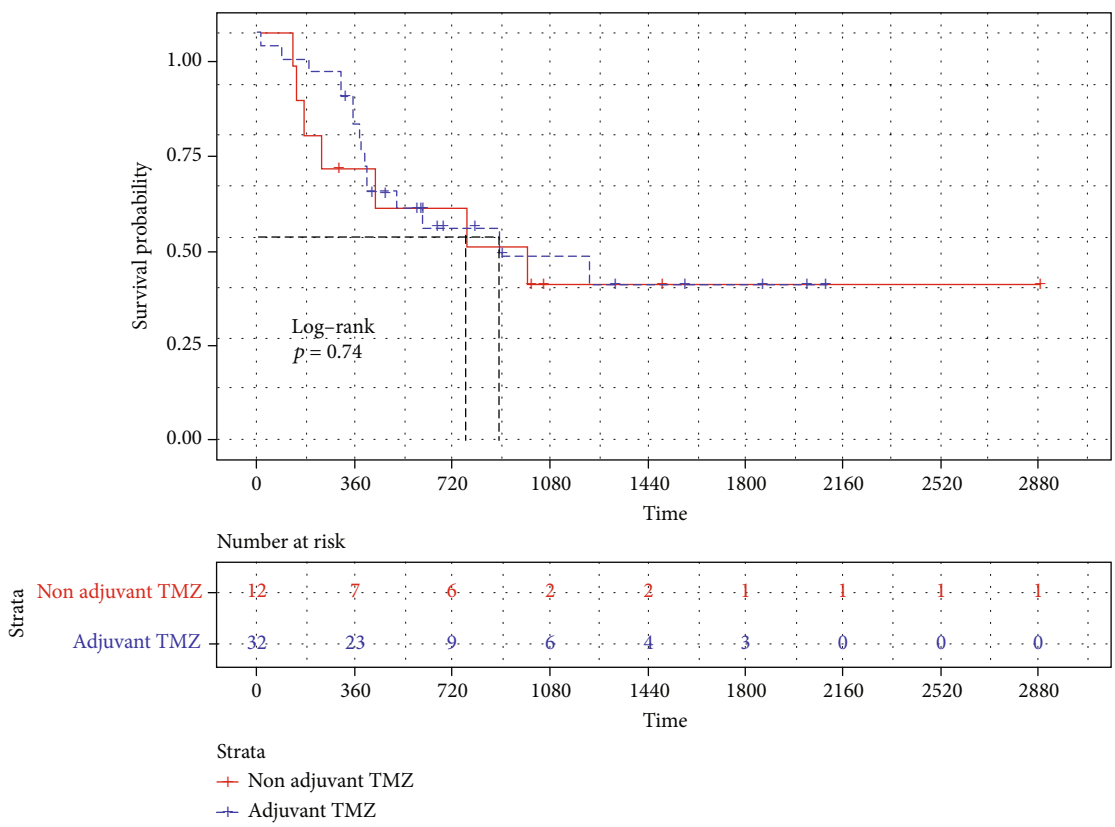
The median OS in our study was 25.6 months and longer than those in most other studies, which range from 5.7 months to 24.7 months [24, 25]. The results of our study suggested that total resection of GSM was associated with prolonged OS and RFS. Although some retrospective studies have reached the opposite conclusions [9, 11, 24], the resection range is a well-accepted independent prognostic factor for improved OS in GSM [4, 5, 17, 26, 27]. Cachia et al. reported that patients of primary GSM undergoing GTR tended to have a greater OS (median 24.7 months) than those having subtotal resection (median 10.1 months) [25]. And with advances in medical technology, more thorough surgeries were accepted by patients over the past six years. In our cohort, even if the lesions were deep in 9 patients, 41 patients had undergone GTR, and no patients received biopsy. Most patients were in relatively good physical condition at the beginning of treatment (median KPS was 80) and were more inclined to tolerate the treatments with large physiological burden.

Currently, almost GSM patients are guided to accept the Stupp protocol in clinical practice, RT, and chemotherapy following surgery as same as GBM [19]. Radiotherapy has been suggested to improve the outcomes of patients, because it may increase OS by 2-4 months [28]. TMZ has been proved as the most effective chemotherapeutic drug for high-grade gliomas [9, 29]. But due to the lack of prospective studies or large-scale multicenter retrospective studies, the debate on the therapeutic value of RT and TMZ in GSM remains unresolved [5]. The scholars from Stanford University Medical Center provided that a significantly improved PFS (median 32.97 months) and OS (median 56.73 months) occurred in patients receiving surgical resection followed by RT and concurrent TMZ [17]. Different from their results, RT, concurrent TMZ, and adjuvant TMZ did not show any survival benefits. Our study is inclined to support the view that RT or TMZ is ineffective for GSM patients. The histopathological differences between GSM and classical GBM cannot be overlooked. GSM is rich in sarcoma components, which were insensitive to RT and chemotherapy [30], which is related to minimal statistical effectiveness. In addition, lots of patients in this cohort have not not yet arrived at the terminal of the life train, which inevitably influences the judgment of the effect of RT and chemotherapy.

The lower Ki-67 expression is parallel to a lower rate of tumor cell division [31], which indicates a longer RFS. This inference has been supported in the study of GSM [32], and our results enhance the reliability.



(a)



(b)

Figure 3: Continued.

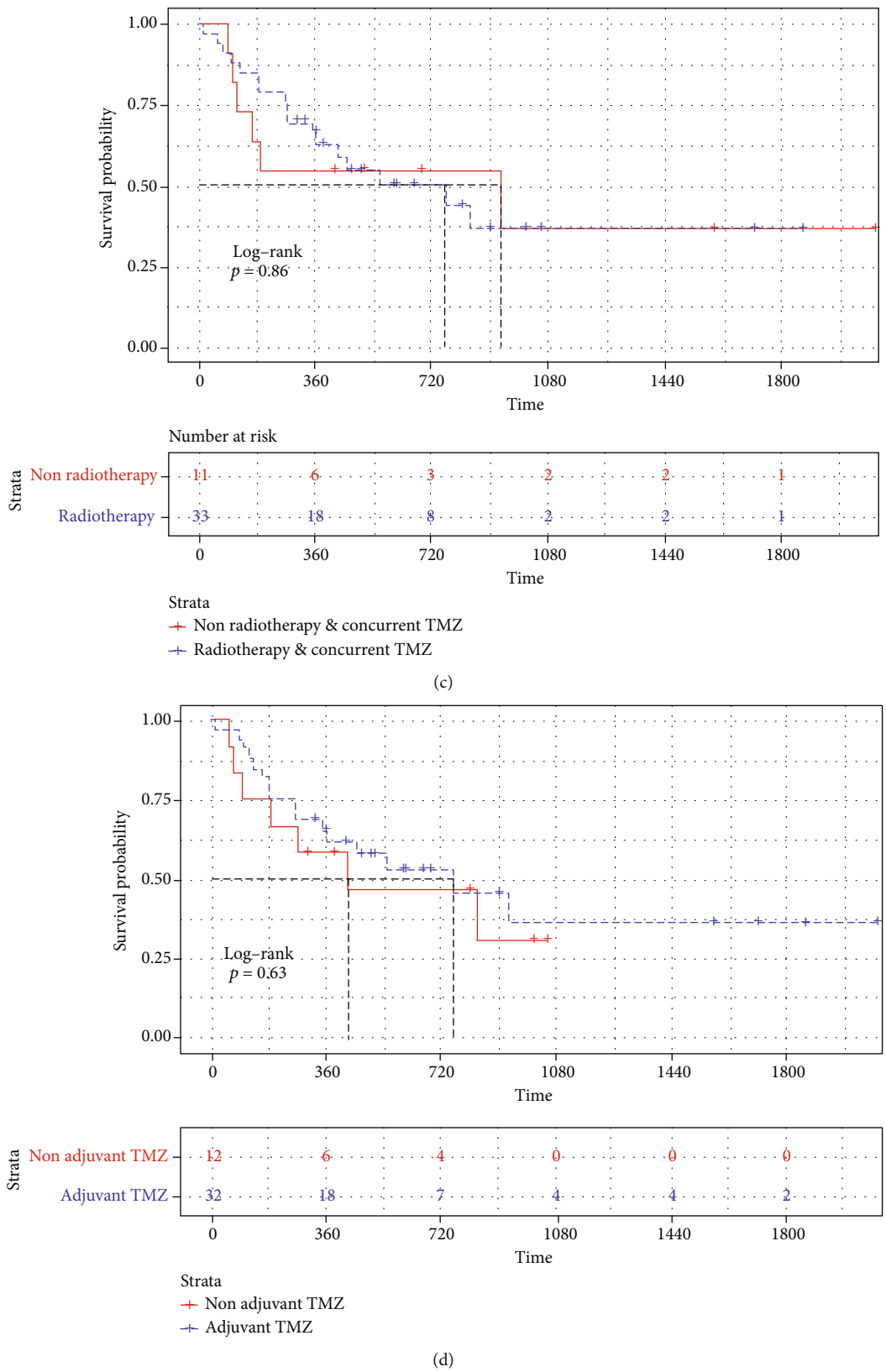


Figure 3: The results of Kaplan-Meier curves and log-rank tests showed that the effect of radiotherapy and concurrent TMZ (a) or adjuvant TMZ (b) on the OS benefits of GSM was not statistically significant. The results of Kaplan-Meier curves and log-rank tests according to RFS indicated that there was no statistically significant impact of radiotherapy and concurrent TMZ (c) or adjuvant TMZ (d).

In univariable analysis, no statistically significant parameters were found ($p \leq 0.05$) but lesion number, but there is only one case with multiple lesions. The enormous sample size between groups makes the results lack credibility and be unable to participate in multivariable analysis. Existing reports suggest that TERT [33] and PIK3CA [34] mutations are frequent in GSM, but there is still a lack of evidence of prognostic relevance. In our study, the p values of TERT are below 0.05 in the log-rank test. The small sample size determines that the univariable analysis of these transgenations has no practical reference value. The results are just listed.

A series of other favorable prognostic factors have been reported previously, including MGMT methylation [5, 14], tumor size [21, 35], temporal tumor location [36], and younger age [5, 13, 21, 32, 35]. These conditions are rarely repeated because of small sample size. No such association was seen in our study.

Compared to prospective studies or large multicenter randomized controlled trials (RCTs), there are many limitations to the single-center retrospective cohort studies for rare diseases. The limited cohort size limits the analysis of potential subset analysis and may hinder our ability to adequately adjust for confounding covariates. All patients had obtained postoperative pathological diagnosis when they were selected into the cohort, which might result in inevitable selection bias. We should be cautious about the conclusions of this study, but we are confident to persist in the clinical value of our results for further prospective study or RCTs.

5. Conclusions

This study is a retrospective cohort study with limited cohort size, our results demonstrate that GTR and low expression of Ki-67 are of great significance to the prognosis of GSM, while the survival benefits of adjuvant therapies including RT or TMZ are limited.

Data Availability

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

Disclosure

In addition, a preprint has previously been published in Research Square [37].

Conflicts of Interest

The authors have no conflicts of interest to declare.

Authors' Contributions

ZY and HX conceived the project. ZY and ZZ obtained the requisite data, conducted the statistical analysis, interpreted the data, drafted the manuscript, generated tables and figures, and received the final manuscript submission. MX, KS, JS, WZ, and LW critically revised the manuscript draft. Ziye Yu and Zhirui Zhou contribute equally to the work.

References

[1] P. Diamandis and K. Aldape, "World Health Organization 2016 Classification of Central Nervous System Tumors," *Neurologic Clinics*, vol. 36, no. 3, pp. 439–447, 2018.

[2] D. N. Louis, A. Perry, G. Reifenberger et al., "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," *Acta Neuropathologica*, vol. 131, no. 6, pp. 803–820, 2016.

[3] K. M. Kiang, A. A. Chan, and G. K. Leung, "Secondary gliosarcoma: the clinicopathological features and the development of a patient-derived xenograft model of gliosarcoma," *BMC Cancer*, vol. 21, no. 1, p. 265, 2021.

[4] K. R. Kozak, A. Mahadevan, and J. S. Moody, "Adult gliosarcoma: epidemiology, natural history, and factors associated with outcome," *Neuro-Oncology*, vol. 11, no. 2, pp. 183–191, 2009.

[5] J. Frandsen, A. Orton, R. Jensen et al., "Patterns of care and outcomes in gliosarcoma: an analysis of the National Cancer Database," *Journal of Neurosurgery*, vol. 128, no. 4, pp. 1133–1138, 2018.

[6] D. N. Louis, A. Perry, P. Wesseling et al., "The 2021 WHO Classification of Tumors of the Central Nervous System: a summary," *Neuro-Oncology*, vol. 23, no. 8, pp. 1231–1251, 2021.

[7] X. Yi, H. Cao, H. Tang et al., "Gliosarcoma: a clinical and radiological analysis of 48 cases," *European Radiology*, vol. 29, no. 1, pp. 429–438, 2019.

[8] R. Ma, D. M. Alexe, and E. A. Pereira, "Primary gliosarcoma: epidemiology, clinical presentation, management, and survival," *Journal of Neurosurgical Sciences*, vol. 64, no. 4, pp. 341–346, 2020.

[9] S. Adeberg, D. Bernhardt, S. B. Harrabi et al., "Radiotherapy plus concomitant temozolomide in primary gliosarcoma," *Journal of Neuro-Oncology*, vol. 128, no. 2, pp. 341–348, 2016.

[10] G. V. Walker, M. R. Gilbert, S. S. Prabhu, P. D. Brown, and M. F. McAleer, "Temozolomide use in adult patients with gliosarcoma: an evolving clinical practice," *Journal of Neuro-Oncology*, vol. 112, no. 1, pp. 83–89, 2013.

[11] G. Singh, S. Mallick, V. Sharma et al., "A study of clinicopathological parameters and O6 – methylguanine DNA methyltransferase (MGMT) promoter methylation status in the prognostication of gliosarcoma," *Neuropathology*, vol. 32, no. 5, pp. 534–542, 2012.

[12] H. Gittleman, A. Boscia, Q. T. Ostrom et al., "Survivorship in adults with malignant brain and other central nervous system tumor from 2000-2014," *Neuro-Oncology*, vol. 20, p. vii6, 2018.

[13] Q. Huang, F. Li, Y. Chen, F. Hong, H. Wang, and J. Chen, "Prognostic factors and clinical outcomes in adult primary gliosarcoma patients: a surveillance, epidemiology, and end results (SEER) analysis from 2004 to 2015," *British Journal of Neurosurgery*, vol. 34, no. 2, pp. 161–167, 2020.

[14] D. R. Smith, C. C. Wu, H. J. Saadatmand et al., "Clinical and molecular characteristics of gliosarcoma and modern prognostic significance relative to conventional glioblastoma," *Journal of Neuro-Oncology*, vol. 137, no. 2, pp. 303–311, 2018.

[15] H. Srivastava, A. Dewan, S. K. Sharma et al., "Adjuvant radiation therapy and temozolomide in gliosarcoma: is it enough?

Case series of seven patients," *Asian Journal of Neurosurgery*, vol. 13, no. 2, pp. 297–301, 2018.

[16] B. Chen, B. Liu, C. Wu, and Z. Wang, "Prognostic factors among single primary gliosarcoma cases: a study using surveillance, epidemiology, and end results data from 1973–2013," *Cancer Medicine*, vol. 8, no. 14, pp. 6233–6242, 2019.

[17] M. C. Jin, E. K. Liu, S. Shi et al., "Evaluating surgical resection extent and adjuvant therapy in the management of gliosarcoma," *Frontiers in Oncology*, vol. 10, p. 337, 2020.

[18] S. J. Han, I. Yang, T. Tihan, M. D. Prados, and A. T. Parsa, "Primary gliosarcoma: key clinical and pathologic distinctions from glioblastoma with implications as a unique oncologic entity," *Journal of Neuro-Oncology*, vol. 96, no. 3, pp. 313–320, 2010.

[19] E. Galanis, J. C. Buckner, R. P. Dinapoli et al., "Clinical outcome of gliosarcoma compared with glioblastoma multiforme: North Central Cancer Treatment Group results," *Journal of Neurosurgery*, vol. 89, no. 3, pp. 425–430, 1998.

[20] G. Singh, K. K. Das, P. Sharma et al., "Cerebral gliosarcoma: analysis of 16 patients and review of literature," *Asian Journal of Neurosurgery*, vol. 10, no. 3, pp. 195–202, 2015.

[21] D. Pierscianek, Y. Ahmadipour, A. Michel et al., "Demographic, radiographic, molecular and clinical characteristics of primary gliosarcoma and differences to glioblastoma," *Clinical Neurology and Neurosurgery*, vol. 200, article 106348, 2021.

[22] S. J. Han, I. Yang, T. Tihan, S. M. Chang, and A. T. Parsa, "Secondary gliosarcoma: a review of clinical features and pathological diagnosis," *Journal of Neurosurgery*, vol. 112, no. 1, pp. 26–32, 2010.

[23] J. R. Perry, L. C. Ang, J. M. Bilbao, and P. J. Muller, "Clinicopathologic features of primary and postirradiation cerebral gliosarcoma," *Cancer*, vol. 75, no. 12, pp. 2910–2918, 1995.

[24] C. S. Dejonckheere, A. M. C. Böhner, D. Koch et al., "Chasing a rarity: a retrospective single-center evaluation of prognostic factors in primary gliosarcoma," *Strahlentherapie und Onkologie*, vol. 198, no. 5, pp. 468–474, 2022.

[25] D. Cachia, C. Kamiya-Matsuoka, J. J. Mandel et al., "Primary and secondary gliosarcomas: clinical, molecular and survival characteristics," *Journal of Neuro-Oncology*, vol. 125, no. 2, pp. 401–410, 2015.

[26] T. J. Brown, M. C. Brennan, M. Li et al., "Association of the extent of resection with survival in glioblastoma: a systematic review and meta-analysis," *JAMA Oncology*, vol. 2, no. 11, pp. 1460–1469, 2016.

[27] G. Zhang, S. Huang, J. Zhang, Z. Wu, S. Lin, and Y. Wang, "Clinical outcome of gliosarcoma compared with glioblastoma multiforme: a clinical study in Chinese patients," *Journal of Neuro-Oncology*, vol. 127, no. 2, pp. 355–362, 2016.

[28] X. Wang, J. Jiang, M. Liu, and C. You, "Treatments of gliosarcoma of the brain: a systematic review and meta-analysis," *Acta Neurologica Belgica*, vol. 121, no. 6, pp. 1789–1797, 2021.

[29] L. B. Nabors, J. Portnow, M. Ammirati et al., "NCCN guidelines insights: central nervous system cancers, version 1.2017," *Journal of the National Comprehensive Cancer Network*, vol. 15, no. 11, pp. 1331–1345, 2017.

[30] Y. Zhang, J. P. Ma, J. C. Weng et al., "The clinical, radiological, and immunohistochemical characteristics and outcomes of primary intracranial gliosarcoma: a retrospective single-centre study," *Neurosurgical Review*, vol. 44, no. 2, pp. 1003–1015, 2021.

[31] M. G. Davey, S. O. Hynes, M. J. Kerin, N. Miller, and A. J. Lowery, "Ki-67 as a prognostic biomarker in invasive breast cancer," *Cancers*, vol. 13, no. 17, p. 4455, 2021.

[32] F. I. Ahmed, K. G. Abdullah, J. Durgin, R. D. Salinas, D. M. O'Rourke, and S. Brem, "Evaluating the association between the extent of resection and survival in gliosarcoma," *Cureus*, vol. 11, no. 4, article e4374, 2019.

[33] J. E. Oh, T. Ohta, N. Nonoguchi et al., "Genetic alterations in gliosarcoma and giant cell glioblastoma," *Brain Pathology*, vol. 26, no. 4, pp. 517–522, 2016.

[34] C. B. Knobbe and G. Reifenberger, "Genetic alterations and aberrant expression of genes related to the phosphatidylinositol-3′-kinase/protein kinase B (Akt) signal transduction pathway in glioblastomas," *Brain Pathology*, vol. 13, no. 4, pp. 507–518, 2003.

[35] S. S. Feng, H. B. Li, F. Fan et al., "Clinical characteristics and disease-specific prognostic nomogram for primary gliosarcoma: a SEER population-based analysis," *Scientific Reports*, vol. 9, no. 1, article 10744, 2019.

[36] S. Frandsen, H. Broholm, V. A. Larsen et al., "Clinical characteristics of gliosarcoma and outcomes from standardized treatment relative to conventional glioblastoma," *Frontiers in Oncology*, vol. 9, p. 1425, 2019.

[37] Z. Y. Yu, Z. R. Zhou, M. Xu et al., *Prognostic factors of gliosarcoma in the real world: a retrospective cohort study*, Research Square, 2022.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 9130958, 14 pages
https://doi.org/10.1155/2022/9130958

Research Article

Comprehensive Genomic Analysis for Identifying FZD6 as a Novel Diagnostic Biomarker for Acute Myeloid Leukemia

Li Yang,[1,2] **Deyu Ma**,[1,3,4,5] **Shi Tang**,[1,3,4,5] **Tingting Jiang**,[3,4,5,6] **Jie Yu**,[3,4,5,6] **Li Wang**,[2] **and Lin Zou**[1,7,8]

[1]*Center for Clinical Molecular Laboratory Medicine of Children's Hospital of Chongqing Medical University, Chongqing, China 400014*
[2]*Department of Hematology, The First Affiliated Hospital of Chongqing Medical University, Chongqing, China 400016*
[3]*National Clinical Research Center for Child Health and Disorders (Chongqing), Chongqing, China 400014*
[4]*Ministry of Education Key Laboratory of Child Development and Disorders, Chongqing, China 400014*
[5]*China International Science and Technology Cooperation Base of Child Development and Critical Disorders, Chongqing, China 400014*
[6]*Center for Pediatric Hematology Diseases of Children's Hospital of Chongqing Medical University, Chongqing, China 400014*
[7]*Clinical Research Unit, Children's Hospital of Shanghai Jiao Tong University, Shanghai, China 200062*
[8]*Institute of Pediatric Infection, Immunity and Critical Care Medicine, Shanghai Jiao Tong University School of Medicine, Shanghai, China 200062*

Correspondence should be addressed to Lin Zou; cmmc_sm@163.com

Received 26 August 2022; Revised 26 September 2022; Accepted 7 October 2022; Published 18 November 2022

Academic Editor: Jincheng Wang

As a family of G protein-coupled receptors (GPCRs) with a seven-span transmembrane structure, frizzled class receptors (FZDs) play crucial roles in regulating multiple biological functions. However, their transcriptional expression profile and prognostic significance in acute myeloid leukemia (AML) are unclear. In AML, the role of FZDs was explored by performing the comprehensive analysis on the relationship between clinical characteristics and mRNA expression profiles from public databases including cBioPortal for Cancer Genomics, Gene Expression Profile Interactive Analysis (GEPIA), and Cancer Cell Line Encyclopedia (CCLE). We identified that in the majority of 27 AML cell lines, frizzled class receptor 6 (FZD6) was high-expressed. A significantly higher expression of FZD6 in AML patients was observed when compared to normal controls ($P < 0.01$). Compared with intermediate and poor/adverse risk group patients, FZD6 expressed much lower in cytogenetic favorable risk group patients ($P < 0.0001$). Patients with higher-expressed FZD6 were associated with shorter overall survival (OS) ($P = 0.0089$) rather than progression-free survival (PFS). However, the predictive effect of FZD6 on OS could be reversed by hematopoietic stem cell transplantation (HSCT). The data of gene set enrichment analysis (GSEA) demonstrated that 4 gene sets, including MYC targets, HEME metabolism, E2F targets, and UV response, were differentially enriched in the high-expression FZD6 group. To conclude, the study suggested that high expression of FZD6 might be a novel poor prognostic biomarker for AML treatment.

1. Introduction

As a malignant hematological disease, acute myeloid leukemia (AML) is resulted from the clonal expansion of abnormally differentiated blasts, manifesting as severe infections, anemia, hemorrhage, and organ infiltration [1]. The overall prognosis of AML is poor, about 40% of adult patients aged 60 years old or younger and 10% of patients aged over 60 years could achieve long-term survival [2]. The majority of AML patients could achieve complete remission through 1 to 2 courses of induction chemotherapy, but AML shows heterogeneity. It is essential to accurately assess the prognosis and assign postremission therapies (hematopoietic stem cell transplantation (HSCT) or consolidation chemotherapy)

for patients at first remission [3]. Risk stratification is often determined by consensus guidelines. However, even low-risk patients may develop a poor prognosis, suggesting that urgent requirements for novel markers can more accurately indicate the prognosis.

Class frizzled (FZD) is one of the GPCR subfamilies, including 10 FZD isoforms denoted FZD1–10 and smoothened (SMO) in mammals. Each FZD receptor is encoded by a separate gene and contains a cysteine-rich domain (CRD), which is crucial for binding to secreted Wingless/Integrated (WNT) proteins [4]. As a main receptor for receiving WNT signals, FZDs can activate the canonical or non-canonical WNT signaling pathway through the interaction with Dishevelled (DVL) protein [5, 6], fulfilling a critical function in stem cell maintenance, cell proliferation, organ formation, cell migration, damage repair, and occurrence of human diseases [7–9]. Activated canonical WNT pathway results in β-catenin accumulation and nuclear translocation, subsequently promoting the transcription of WNT-related genes [10]. Non-canonical pathway refers to the pathways such as the planar cell polarity (PCP) pathway and WNT/$Ca2^{+-}$ pathway that rely on WNT signal transduction but do not cause changes in soluble β-catenin [11]. Dysregulated expression or mutation of FZD genes and the prominent role of these molecules in cancer have been observed in various human malignant diseases [12–15].

Some studies explored the role of FZDs in hematological malignant diseases. FZD7 and FZD8 are expressed in most acute lymphoblastic leukemia (ALL) cells, while FZD3, FZD4, and FZD9 are occasionally detected. Wnt3a activates related receptors to promote the proliferation of ALL cells [16]. Meanwhile, targeting FZD7 and FZD8 can increase the drug sensitivity of multidrug-resistant ALL [17]. The relative mRNA expression levels of FZD4, FZD5, and FZD7 are upregulated in drug-mediated apoptotic chronic myeloid leukemia (CML) cells, suggesting a correlation with programmed cell death [18]. FZD4 knockdown inhibits CML progenitor growth and might increase the sensitivity of CML to Tyrosine kinase inhibitors (TKI) [19]. The expression of FZD1, especially FZD6, is progressively upregulated in the transformed chronic lymphocytic leukemia (CLL) cells, and this reveals a key role of leukemogenesis [20]. CLL shows a higher FZD3 expression than normal B cells, which indicates a less favorable clinical prognosis [21, 22]. The mRNA expression of FZD1 and its ligand Wnt3 is upregulated in mantle cell lymphoma-initiating cells, and this is possibly related to chemical resistance [23]. High-expressed FZD4 in AML blasts enhances β-catenin stability in myeloid progenitor cells induced by Wnt3a to regulate cell apoptosis [24]. miR-212-5p with FZD5 as a functional target has been found to be low-expressed in AML cases and cells, and it served as a tumor-suppressor gene. miR-212-5p/FZD5 is likely to become a new therapeutic target for AML [25]. In patients experiencing a relapse, FZD1 expression in blasts is significantly higher. Overexpression of FZD1 may cause drug resistance in leukemia cells, whereas silencing FZD1 may reverse multidrug resistance [26].

Abnormal expression of the FZDs is common in AML. However, the correlation between clinical characteristics of AML and the FZDs mRNA expression profile has not been well studied. We therefore studied the transcriptional expression profiles of FZDs in AML patients using Gene Expression Profiling Interactive Analysis (GEPIA) online databases and in AML cell lines using Cancer Cell Line Encyclopedia (CCLE). Between normal control individuals and AML patients, the cBioPortal TCGA database was applied to comprehensively analyze the differences in clinical prognostic significance. The potential underlying mechanisms and biological functions of FZD6 in AML were explored by protein-protein interaction (PPI) network and gene set enrichment analysis (GSEA). Flow work was shown in Figure 1.

2. Materials and Methods

2.1. CCLE Database. The mRNA expression profiles of FZDs in AML cell lines were downloaded directly from the CCLE website (https://portals.broadinstitute.org/ccle/home) [27], which stores tumor genomics information of 1019 cell lines from individuals of various ethnicities. The mRNA expression of the ten FZDs in 27 AML cell lines was analyzed in detail.

2.2. GEPIA Database. GEPIA (http://gepia.cancer-pku.cn/) supports the analysis on standardized expression data of 8,587 normal samples and 9,736 tumors from GTEx and TCGA databases [28]. The GEPIA dataset was applied to analyze the differences in FZDs' transcriptional expression between normal control individuals and AML patients, and the connection between the expression level of FZDs and OS.

2.3. cBioPortal for Cancer Genomics Database. cBioPortal for Cancer Genomics (https://www.cbioportal.org/) supports the exploration, visualization, and analysis of multilayer clinical and cancer genome data [29]. We downloaded the mRNA expression profiles, clinical characteristics, laboratory features, and survival data of 173/200 new AML patients from the TCGA dataset on the cBioPortal website.

2.4. STRING and GeneMANIA Database. STRING (https://string-db.org/) could calculate and predict PPI information to generate an objective and comprehensive interaction network, including indirect (functional) and direct (physical) interactions [30]. GeneMANIA (http://www.genemania.org) predicts gene function, analyzes gene lists, and prioritizes genes for functional assays [31]. STRING and GeneMANIA datasets were used to explore the PPI network of FZD6.

2.5. GSEA (Gene Set Enrichment Analysis). The underlying biological mechanism of FZD6 in AML was explored. GSEA identified the potential statistically significant differences between low FZD6 expression and high FZD6 expression groups. GSEA was performed on predefined gene sets, hallmark gene set (h.all.v7.0.symbols.gmt) from MsigDB.

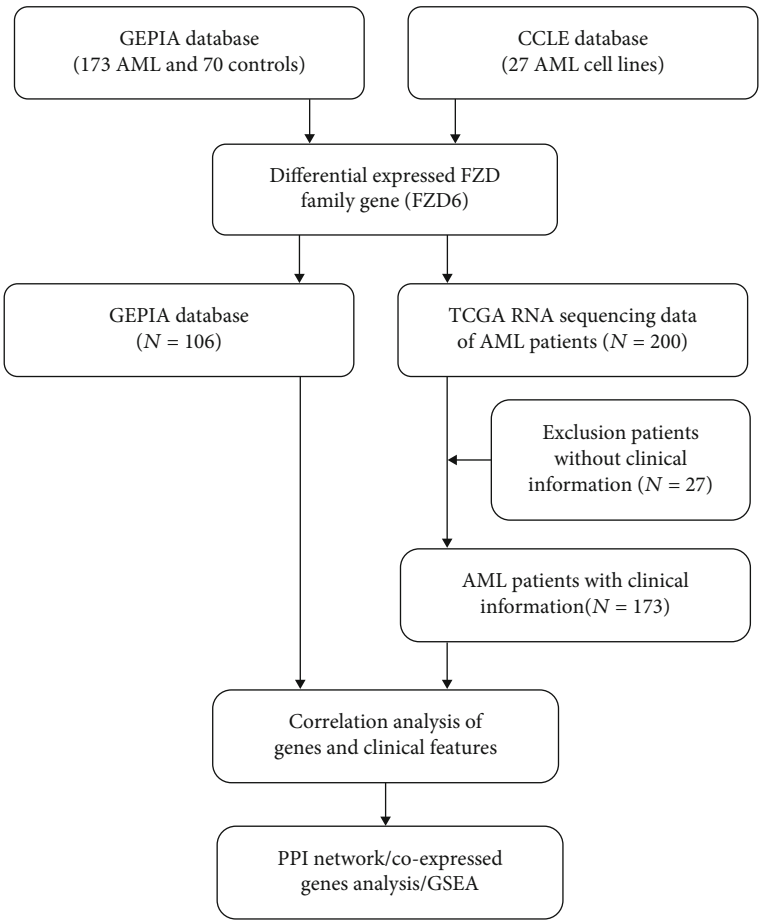


Figure 1: Flow chart of the study design.

Statistically significant was considered when the false discovery rate (FDR) <0.05 and adjusted P value <0.05.

2.6. Statistical Analysis. Based on the median FZD6 expression level, the patients were grouped into two groups according to low- and high-expression of FZD6. All statistical analyses were carried out in SPSS23.0 software (SPSS Inc, Chicago, IL, USA) and GraphPad Prism5 (GraphPad Software Inc, La Jolla, CA, USA). The relationship between FZD6 expression and clinical features was analyzed by the chi-squared test or Fisher's exact test. Survival curves were plotted by the Kaplan-Meier method, and the survival difference between the two groups of patients was analyzed by log-rank test. Multivariate analyses were conducted with Cox proportional hazard model. P value <0.05 was defined as of statistical significance.

3. Results

3.1. mRNA Expression Profiles of FZDs in AML Cell Lines and AML Patients. The mRNA expression of FZDs was determined in 27 AML cell lines from the CCLE database. High mRNA expression of FZD1, FZD2, FZD5, FZD6, and FZD7 were detected in AML cell lines. FZD6 had the highest expression, and FZD10 had the lowest expression in the cells. FZD4 showed a polarized expression (Figure 2(a)). To further investigate the mRNA expression profiles of FZDs in AML patients, we explored their mRNA expression using the GEPIA online database. It has been found that FZD6 was remarkably higher-expressed ($P < 0.01$) in AML patients compared with that in normal controls (Figure 2(b)). FZD4 was significantly lower-expressed ($P < 0.01$) in AML patients compared with that in normal individuals (Figure 2(c)). Other FZDs did not exhibit obvious differences in their mRNA expression levels ($P > 0.05$) (Figure S1). High-expressed FZD6 in AML patients was consistent with that in the cell lines. However, low-expressed FZD4 was different from that in cell lines. Therefore, only FZD6 was included in the follow-up correlation analysis with clinical features.

3.2. High FZD6 Expression Is Related to Poorer Risk Classification in AML Patients. The correlation of clinical characteristics with FZD6 mRNA expression in AML patients was analyzed, as shown in the workflow (Figure S2). From the cBioPortal TCGA database, clinical characteristics of AML patients and the mRNA expression data of FZD6 were downloaded. Based on the median FZD6 expression level, all patients ($n = 173$) with clinical information and RNA sequencing data were grouped into the low FZD6 expression group ($n = 87$) and high FZD6 expression group ($n = 86$) [32]. No significant differences

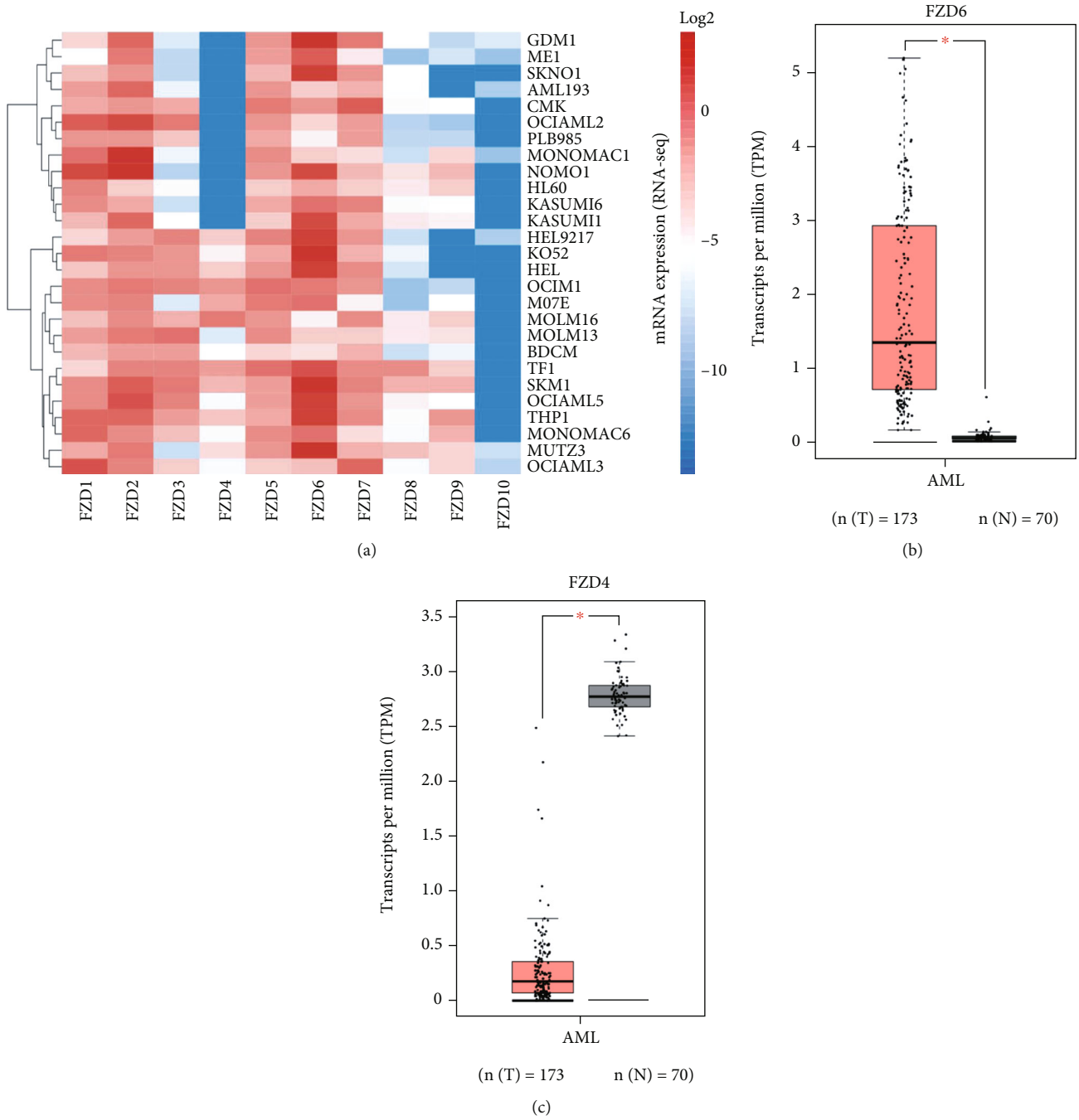


Figure 2: mRNA expression profiles of FZDs in AML cell lines and AML patients. (a) Heat map of FZD1 to FZD10 expression in 27 AML cell lines. (b) Expression of FZD6 in AML patients ($n = 173$) compared to normal samples ($n = 70$) in TCGA and GTEx dataset. (c) Expression of FZD4 in AML patients ($n = 173$) compared to normal samples ($n = 70$) in TCGA and GTEx dataset ($^*P < 0.01$).

were detected between the two groups in terms of sex, age, peripheral blood (PB) blasts, white blood cell (WBC) count, and induction chemotherapy regimens (intensive regimens or nonintensive regimens) ($P > 0.05$) (Table 1, Figures S3(a)–S3(c). Our data showed that low FZD6 expression patients had a higher percentage of bone marrow (BM) blasts ($P = 0.0014$) (Table 1). According to the cytogenetic risk stratification (CRC) and European Leukemia Net (ELN) risk stratification, the proportion of patients with intermediate-risk and poor-/adverse-risk was higher in the high FZD6 expression group than in the low FZD6 expression group ($P < 0.001$) (Table 1). Meanwhile, FZD6 expression was significantly upregulated and was accompanied by elevated CRC and ELN risk stratification (Figures 3(a), 3(b)). Additionally, more patients in the high FZD6 expression group were treated by HSCT ($P = 0.003$) (Table 1). Correspondingly, patients who received HSCT showed much higher FZD6 expression compared with those who received chemotherapy or no treatment ($P < 0.0001$) (Figure 3(c)). Further exploration showed that

Table 1: Correlation of FZD6 expression with clinical characteristics in AML patients.

	FZD6 high ($n = 86$)	FZD6 low ($n = 87$)	P
Sex, n (%)			0.490
Female	38 (44.2)	43 (49.4)	
Male	48 (55.8)	44 (50.6)	
Age, years			0.818
Median (range)	58 (21-88)	58 (18-82)	
WBC ($\times 10^9$/L)			0.985
Median (range)	14.7 (0.5-223.8)	14.5 (0.4-297.4)	
BM blasts (%)			0.001
Median (range)	67 (30-100)	78 (32-100)	
PB blasts (%)			0.099
Median (range)	48 (0-97)	25 (0-98)	
CRC, n (%)*			<0.001
Favorable	9 (10.7)	23 (26.7)	
Intermediate	44 (52.4)	57 (66.3)	
Poor	31 (36.9)	6 (7.0)	
ELN risk stratification, n (%) *			<0.001
Favorable	9 (10.7)	24 (27.9)	
Intermediate	41 (48.8)	51 (59.3)	
Adverse	34 (39.5)	11 (13.0)	
Induction therapy, n (%) #			0.910
Intensive	66 (80.5)	69 (81.2)	
Nonintensive	16 (19.5)	16 (18.8)	
HSCT, n (%)			0.003
Yes	46 (53.5)	27 (31.0)	
No	40 (46.5)	60 (69.0)	
Gene mutation			
FLT3, n (%)	20 (23.3)	29 (33.3)	0.141
TP53, n (%)	13 (15.1)	1 (1.2)	<0.001
DNMT3A, n (%)	23 (26.7)	19 (21.8)	0.451
NPM1, n (%)	9 (10.5)	37 (42.5)	<0.001
RUNX1, n (%)	10 (11.6)	5 (5.8)	0.188
ASXL1, n (%)	3 (3.49)	0 (0)	0.121
IDH1, n (%)	9 (10.5)	8 (9.2)	0.804
IDH2, n (%)	9 (10.5)	7 (8.1)	0.611

*The total patient number is 170 for three patients are lacking evaluable cytogenetic or molecular Information. #Total patient number is 167 for six patients who did not receive any treatment after diagnosis. Intensive treatment means the induction therapy regimen is abased on 7 + 3 regimens. Nonintensive treatment means epigenetic therapy and low-intensive treatment. Abbreviations: WBC, white blood cell; BM, bone marrow; PB, peripheral blood; CRC, cytogenetic risk classification; ELN, European Leukemia Net; HSCT, hematopoietic stem cell transplantation.

in patients receiving different types of allogeneic hematopoietic stem cell transplantation (allo-HSCT), there were no differences in FZD6 expression level, but the level was higher than in patients with chemotherapy or without treatment (Figure S3(d)). According to FAB (French-American-British) classification, the expressions of FZD6 in AML-M3 patients (a highly curable subtype of AML) were significantly lower than in other types (Figure S3(e)).

Furthermore, we explored the relationship between common gene mutations in AML and FZD6 expression. The incidence analysis of FLT3, TP53, DNMT3A, NPM1, RUNX1, ASXL1, IDH1, and IDH2 mutations in the FZD6 high-expression group and the FZD6 low-expression group showed that the FZD6 low-expression group had a greater possibility of NPM1 mutation ($P < 0.001$) (Table 1). TP53 mutations, which independently affect prognosis, were higher in the high-expression group ($P < 0.001$) (Table 1). The expression level of FZD6 in the FLT3 mutation group ($P = 0.0186$) and NPM1 mutation group ($P < 0.0001$) was lower (Figure 3(d)–3(e)), while that of FZD6 in the TP53 mutation group was higher ($P = 0.0001$) (Figure 3(f)). The relationship between FZD6 expression and other gene mutations was not statistically significant (Table 1) (Figures S3(f)–S3(j)). Given that the FLT3 mutation rate and the

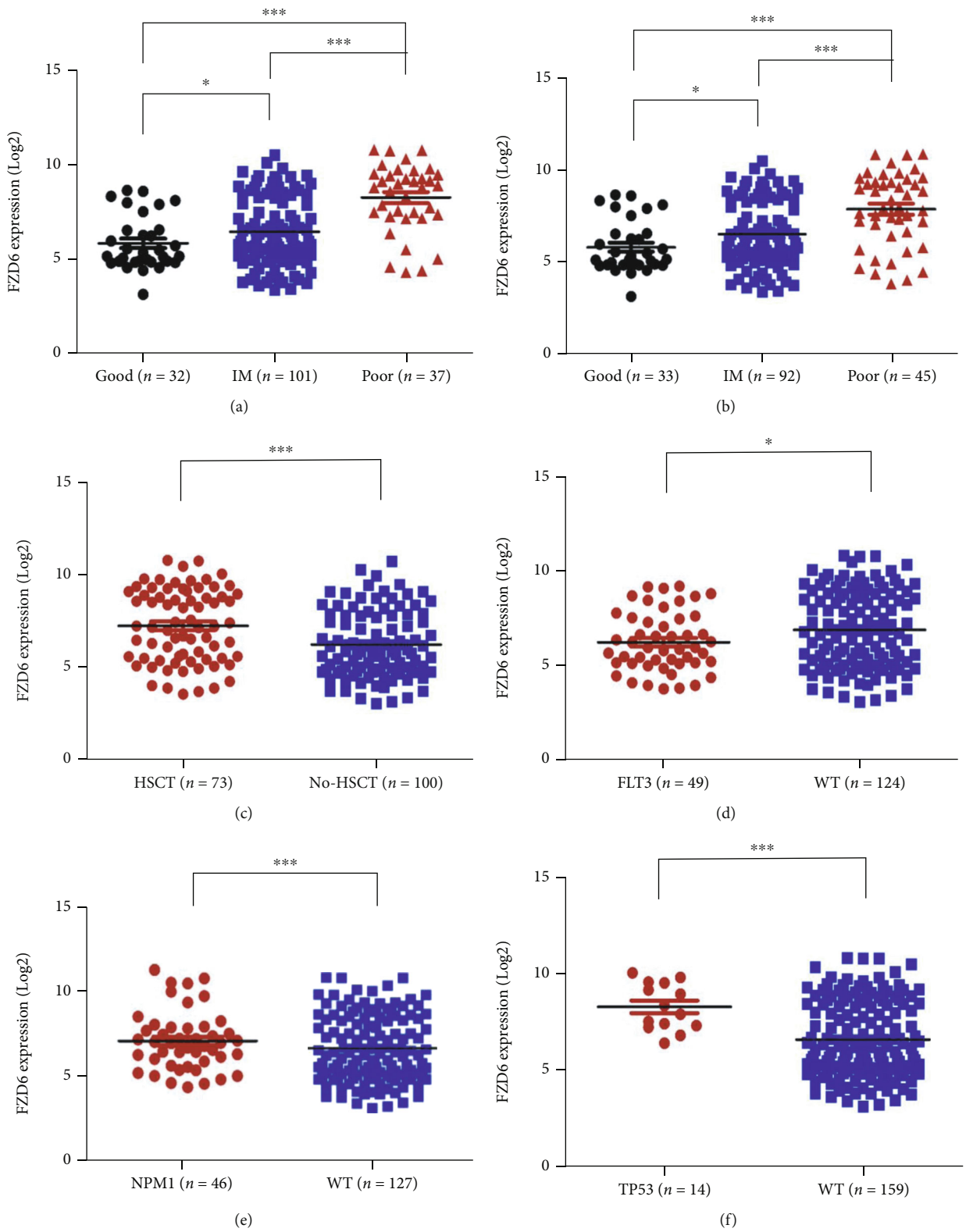


Figure 3: High FZD6 expression is related to poorer risk classification in AML patients. (a) FZD6 expression differences among low risk ($n = 32$), intermediate (IM) risk ($n = 101$), and poor risk ($n = 37$), according to CRC. (b) FZD6 expression differences among good risk ($n = 33$), intermediate (IM) risk ($n = 92$), and adverse risk ($n = 45$), according to ELN risk stratification. (c) FZD6 expression differences between patients received HSCT ($n = 73$) or did not received HSCT (no-HSCT) ($n = 100$). (d) FZD6 expression differences between patients had FLT3 mutation ($n = 49$) or not ($n = 124$). (e) FZD6 expression differences between patients had NPM1 mutation ($n = 46$) or not ($n = 127$). (f) FZD6 expression differences between patients had TP53 mutation ($n = 14$) or not ($n = 159$). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

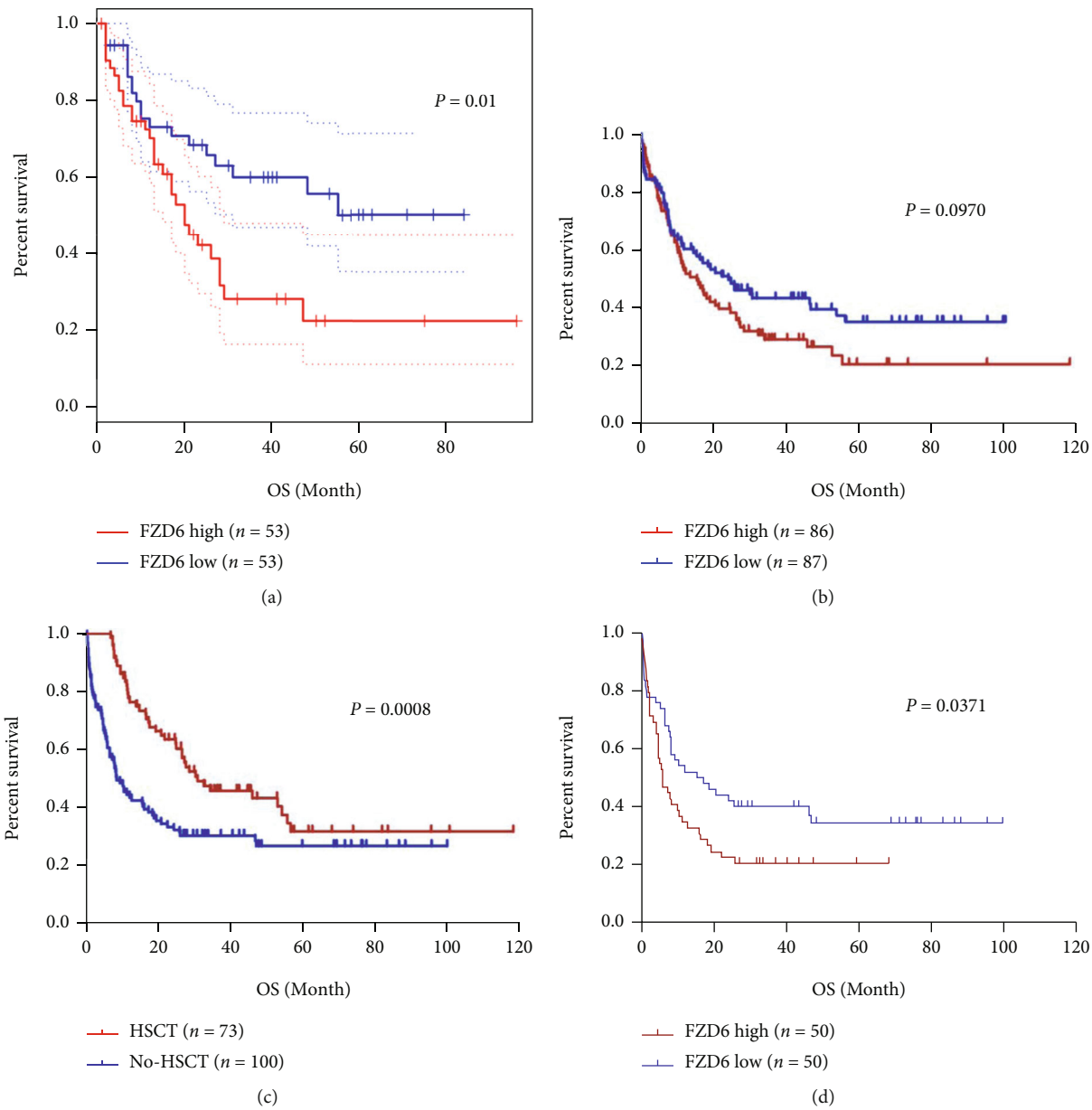


Figure 4: Continued.

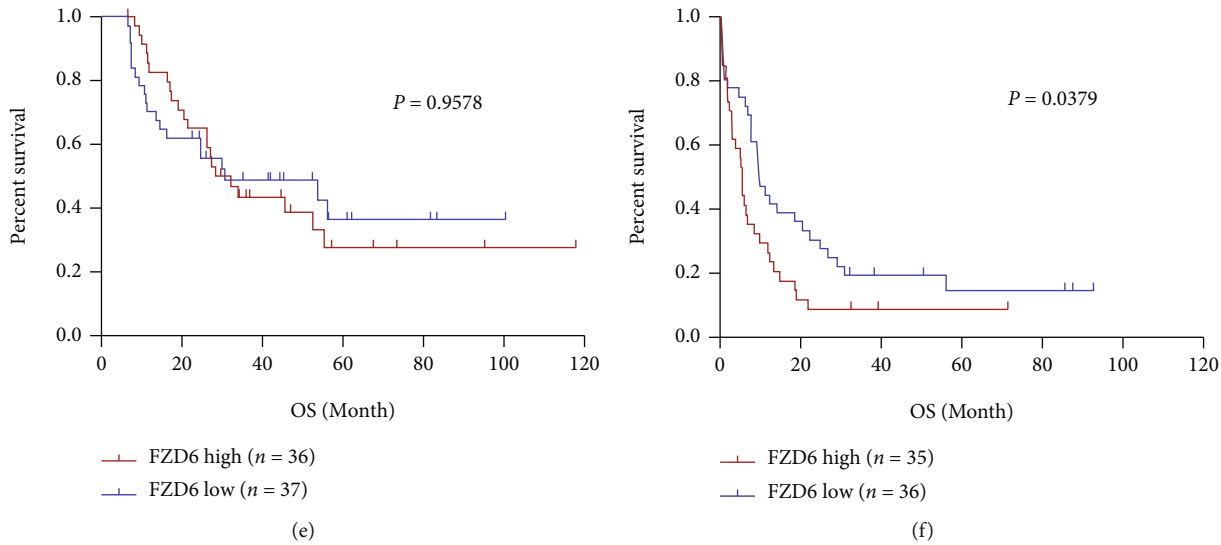


Figure 4: High FZD6 expression is related to poor OS in AML patients. (a) The association between FZD6 and OS of AML patients on the GEPIA website. (b) The association between FZD6 and OS of AML patients based on the cBioPortal TCGA database. (c) The association between HSCT and OS of AML patients. (d) The association between FZD6 and OS of AML patients did not receive HSCT. (e) The association between FZD6 and OS of AML patients received HSCT. (f) The association between FZD6 and OS of AML patients with intermediate/adverse risk did not receive HSCT.

Table 2: Cox proportional hazards model for OS and PFS in AML patients.

Variables	OS		PFS	
	HR (95% CI)	*P*	HR (95% CI)	*P*
Age	1.018 (1.001-1.035)	0.040	1.008 (0.991-1.026)	0.335
Sex (male vs. female)	0.811 (0.547-1.203)	0.298	0.752 (0.480-1.177)	0.212
WBC counts	1.007 (1.002-1.011)	0.002	1.007 (1.002-1.011)	0.006
BM blasts percentage	1.009 (0.998-1.020)	0.127	1.002 (0.989-1.016)	0.769
ELN risk stratification		0.006		0.018
(intermediate/adverse vs. favorable)	2.720 (1.341-5.518)		2.521 (1.173-5.419)	
Induction therapy		0.001		0.474
(intensive vs. nonintensive)	0.371 (0.211-0.654)		0.736 (0.318-1.702)	
HSCT (yes vs. no)	0.508 (0.304-0.848)	0.010	1.149 (0.644-2.052)	0.637
FZD6 expression	1.264 (1.025-1.559)	0.028	1.081 (0.853-1.370)	0.518

Abbreviations: OS, overall survival; PFS, progression free survival; HR, hazard ratio; CI, confidence interval.

coexistence of NPM1 mutations could affect prognosis, the expression of FZD6 is closely related to TP53 mutation, which suggests a poor prognosis.

3.3. High FZD6 Expression Is Related to Poor OS in AML Patients. The association of the prognosis of AML patients with FZD6 was explored. The GEPIA was applied to analyze the Kaplan-Meier survival curve of AML patients' OS. A shorter OS of AML patients with higher FZD6 expression than those with lower-expressed FZD6 was observed ($P = 0.0089$) (Figure 4(a)). Then, we analyzed the association of FZD6 expression with OS and progression-free survival (PFS) based on data downloaded from the cBioPortal TCGA database. Patients in the high FZD6 expression group showed a shorter OS but without statistical significance (median: 15.80 vs. 24.80 months; $P = 0.097$) (Figure 4(b)). As HSCT is an important treatment option, we divided 173 AML patients into HSCT group ($n = 73$) and no-HSCT group ($n = 100$), and then analyzed the relationship between FZD6 level and prognosis of patients in each group. The results showed that the OS in patients without HSCT treatment was significantly reduced (median: 30.60 vs. 8.20 months; $P = 0.0008$) (Figure 4(c)). In the group of patients without HSCT treatment, higher FZD6 expression patients had a shorter OS compared with those with lower FZD6 expression (median: 5.70 vs. 16.20 months; $P = 0.037$) (Figure 4(d)). However, no statistical differences were detected in the patient group who received HSCT (median: 32.30 vs. 30.60 months; $P = 0.9578$) (Figure 4(e)). Except for low-risk patients who generally did not need HSCT, among the intermediate-/adverse-risk patients with HSCT, a high FZD6 expression group showed a significantly shorter

OS than that of low FZD6 expression group (median: 4.55 vs. 8.05 months; $P = 0.0379$), indicating that high FZD6 expression was correlated with unfavorable prognosis (Figure 4(f)). No similar results on the relationship between FZD6 expression and PFS were observed (Figure S4(a)–(d)).

In addition to FZD6 expression, as shown by Cox regression analyses, other clinical factors may also affect the prognosis of AML. Age (hazard ratio (HR) = 1.018, 95% CI: 1.001-1.035, $P = 0.040$), WBC counts (HR = 1.007, 95% CI: 1.002-1.011, $P = 0.040$), ELN risk stratification (HR = 2.720 , 95% CI: 1.341-5.518, P = 0.006), induction therapy (HR = 0.371, 95% CI: 0.211-0.654, $P = 0.001$), HSCT (HR = 0.508, 95% CI: 0.304-0.848, $P = 0.010$), and FZD6 expression (HR = 1.264, 95% CI: 1.025-1.559, $P = 0.028$) were the independent prognosis factors of OS in AML patients. However, only WBC counts (HR = 1.007, 95% CI: 1.002-1.011, $P = 0.006$) and ELN risk stratification (HR = 2.521, 95% CI: 1.173-5.419, $P = 0.018$) could predict the PFS of AML patients (Table 2).

3.4. Potential Role of FZD6 in AML. To explore the core regulatory genes and underlying mechanisms, we used the STRING and GeneMANIA databases to construct a PPI network. The results demonstrated that FZD6 interacted with the protein related to canonical and noncanonical WNT signaling pathways, which was in line with its physiological function (Figures 5(a) and 5(b)). Based on the cBioPortal dataset, 266 coexpressed genes are shown in Table S1 (Spearman's correlation>0.4 and P value <0.05), and the top 10 positively coexpressed genes and negatively coexpressed genes are shown in Table 3.

Furthermore, to explore underlying mechanisms, we compared the transcriptomes of the high FZD6 expression group and the low FZD6 expression groups based on the cBioPortal TCGA database. A total of 1152 genes were identified, including 386 upregulated and 766 downregulated. Statistical differences between high FZD6 and low FZD6 groups were found ($P \leq 0.05$, $|\log 2FC| \geq 1$) (Figure S5(a)). The Gene Ontology (GO) analysis and the top 20 enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are shown in Figures S5(b) and S5(c). GSEA was used to analyze the differences between high and low FZD6 expression data sets for screening gene sets involved in AML. Our result demonstrated that 4 gene sets, including HEME metabolism (NES = 1.75, FDR = 0.003, $P < 0.001$), UV response (NES = 1.52, FDR = 0.035, $P = 0.002$), E2F targets (NES = 1.74, FDR = 0.002, $P < 0.001$), MYC targets (NES = 1.78, FDR = 0.004, $P < 0.001$), were differentially enriched in the high FZD6 expression group. All the 4 sets were critical for AML development and progression [33–36]. The results pointed to a potential function of FZD6 in AML development (Figure 5(c)).

4. Discussion

The clinical outcome of AML is highly heterogeneous. In particular, the prognosis and postremission treatment of AML is based on risk stratification. It is a great challenge for clinicians that patients with the same risk stratification may have completely different prognoses. Thus, finding novel biomarkers for prognostic evaluation is an urgent task. Combined analysis using the databases of CCLE, GEPIA, cBioPortal, STRING, and GeneMANIA databases manifested high-expressed FZD6 in AML cell lines and patients, which was positively correlated with the risk factors, OS and HSCT, and that 4 gene sets (MYC targets, HEME metabolism, E2F targets, and UV response) were differentially enriched in the high FZD6 expression group.

FZDs serve as receptors for secreted WNT ligands in the WNT signaling pathway and play crucial roles in regulating multiple biological functions [37]. Their transcriptional expression profile and prognostic significance in AML remain relatively unknown. Although we measured all members of the FZD family in the AML cell lines, only the expression of FZD1, FZD2, FZD5, FZD6, and FZD7 were higher, with FZD6 being the highest. Therefore, we further explored the correlation of the clinical characteristics of AML with FZD6 expression.

High FZD6 expression has been identified in various cancer cells, including in glioblastoma, oral squamous cell carcinoma, and pancreatic adenocarcinoma, showing a strong correlation with tumor malignancy and prognosis [38–40]. At the same time, we also used SangerBox (http://vip.sangerbox.com) to evaluate the relationship between the expression and prognosis of FZD6 in pan-cancer, and the results were consistent with the previous literature (Figure S6) [39, 40]. In hematopoietic malignancies, WNT signaling activated by WNT10B/FZD6 promotes intracellular effectors and leukemic expansion in WNT10BR-positive T-ALL cells [41]. WNT/FZD (especially FZD6) self-renewal signals are pathologically reactivated in the neoplastic transformation of mature B cells in CLL [20]. lncRNA prostate cancer-associated transcript-1(PCAT-1) interacts with FZD6 to activate WNT/β-catenin signaling and may exert a crucial effect on AML pathogenesis [42]. Currently, no relevant studies were conducted to explore the association between FZD6 and the clinical features of AML.

We detected that high FZD6 expression was associated with poor cytogenetic risk classification, adverse ELN risk stratification, TP53 mutation, and short survival of AML patients, according to several powerful publicly available datasets. No statistically significant difference has been found between FZD6 expression level and OS of AML patients. However, when excluding AML patients who received HSCT, patients with a higher FZD6 expression had a shorter OS. This was possibly due to the fact that HSCT as a potentially curative treatment option for AML was the most effective prognosis protection factor, which may partially interfere with the effect of FZD6 expression on prognosis. The results indicated that the predictive effect of FZD6 on OS could be reversed by HSCT, and that intermediate/adverse risk patients with higher FZD6 showed a more urgent need for HSCT. Multivariate survival analysis demonstrated that older age, higher WBC counts, intermediate/adverse risk, and higher FZD6 expression are independent prognostic factors for poor OS, while intensive chemotherapy and HSCT are protective factors for the

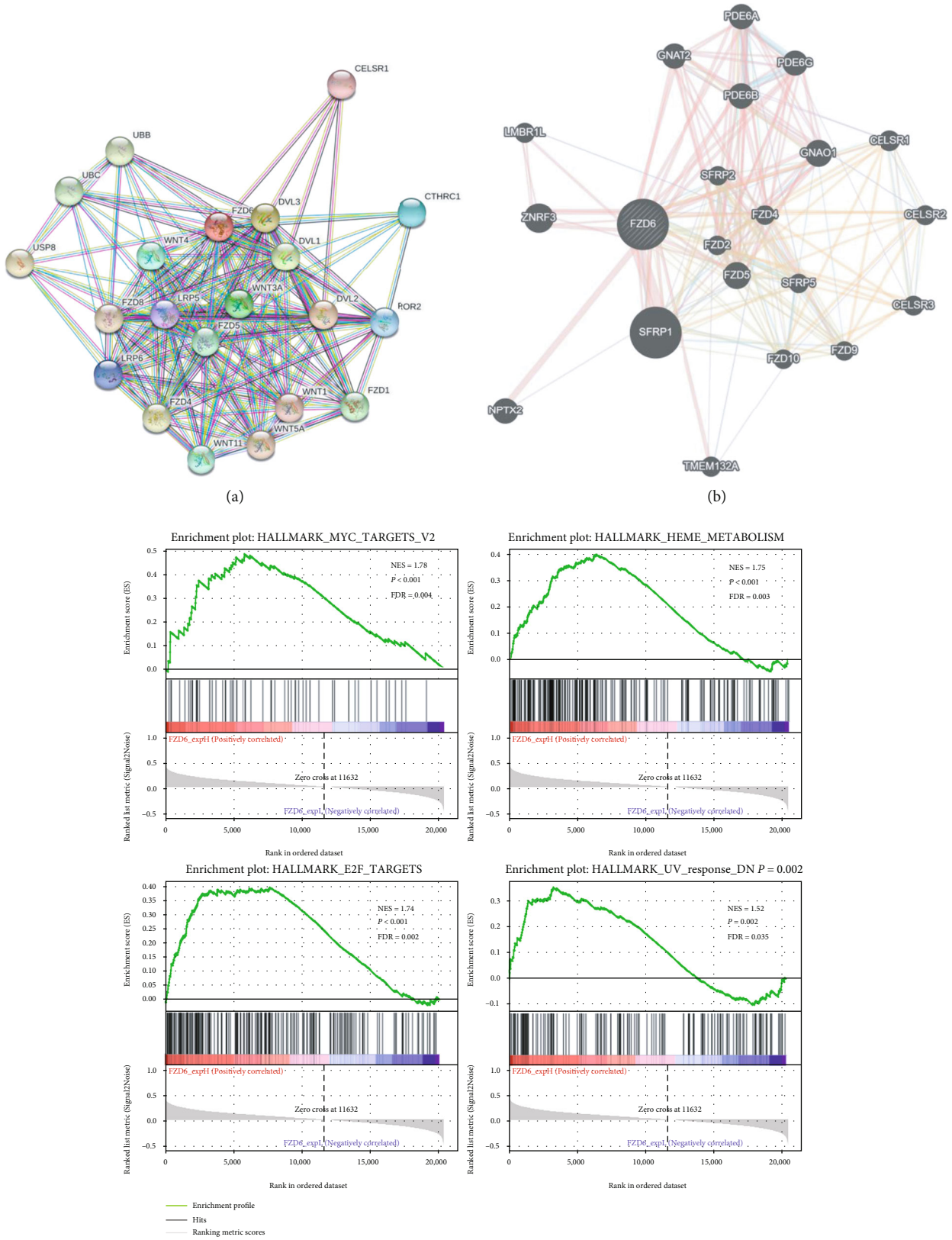


Figure 5: Potential role of FZD6 in AML. (a) PPI network of FZD6 analyzed by STRING. (b) PPI network of FZD6 analyzed by GeneMANIA. (c) GSEA analysis of AML patients based on FZD6 expression.

Table 3: Top 10 positively and negatively coexpressed genes of FZD6.

Correlated gene	Cytoband	Co-ex relationship	Spearman's correlation	P	q value
MLLT3	9p21.3	Positively	0.576	$1.17E-16$	$1.42E-12$
SPAG16	2q34	Positively	0.574	$1.44E-16$	$1.42E-12$
AKAP6	14q12	Positively	0.558	$1.50E-15$	$9.88E-12$
PRKCH	14q23.1	Positively	0.549	$4.92E-15$	$2.43E-11$
MREG	2q35	Positively	0.535	$3.51E-14$	$1.39E-10$
ATP8A1	4p13	Positively	0.528	$7.94E-14$	$2.61E-10$
ARMCX5	Xq22.1	Positively	0.525	$1.21E-13$	$3.09E-10$
KLHL6	3q27.1	Positively	0.525	$1.25E-13$	$3.09E-10$
RAB39B	Xq28	Positively	0.522	$1.74E-13$	$3.81E-10$
TRAF5	1q32.3	Positively	0.521	$2.10E-13$	$4.06E-10$
PARL	3q27.1	Negatively	-0.492	$6.47E-12$	$3.65E-09$
BTG1	12q21.33	Negatively	-0.49	$8.02E-12$	$4.05E-09$
CFD	19p13.3	Negatively	-0.485	$1.35E-11$	$6.12E-09$
RPA4	Xq21.33	Negatively	-0.481	$2.08E-11$	$8.73E-09$
HOMER3	19p13.11	Negatively	-0.475	$4.07E-11$	$1.54E-08$
RAC1	7p22.1	Negatively	-0.462	$1.57E-10$	$4.24E-08$
PPP1R27	17q25.3	Negatively	-0.46	$2.00E-10$	$4.92E-08$
RNASE2	14q11.2	Negatively	-0.453	$4.06E-10$	$8.90E-08$
ZNHIT1	7q22.1	Negatively	-0.451	$4.96E-10$	$1.05E-07$
CST3	20p11.21	Negatively	-0.449	$5.87E-10$	$1.21E-07$

prognosis [43, 44]. Taken together, FZD6 may become a new prognostic biomarker for AML.

Moreover, mediator of the WNT canonical ligands binding to FZD6 promotes β-catenin accumulation and nuclear translocation and activates downstream target genes, such as Cyclin D1, c-MYC, Survivin, and MMP, thereby regulating cell proliferation [45]. However, most reports indicate a prevalent role in the noncanonical pathway [46]. PCP signaling pathway sequentially activates Rac and Rho small GTPases and JNK and promotes actin polymerization and cytoskeleton modification. WNT/Ca^{2+} pathway activates calcineurin and the nuclear factor of activated T cells (NFAT) transcription factors, which regulate transcriptional programs involved in cell fate and cell migration [47]. To further understand the functional roles of FZD6, we analyzed the FZD6 PPI network and generated a gene network using STRING and GeneMANIA databases. The results showed that FZD6 mainly interacted with proteins involved in the WNT signaling pathway, such as WNT proteins, DVL proteins, secreted frizzled-related protein 1 (SFRP1), neuronal pentraxin 2 (NPTX2), and zinc and ring finger 3 (ZNRF3), and so on. Moreover, the effect of FZD6 on AML resulted from the activation of the WNT signaling pathway, therefore showing a prognostic significance in AML [48]. Then, we explored the FZD6 coexpression network, and the data suggested that FZD6 may participate in other signaling pathways and activate its underlying molecular mechanisms to exert its effect on AML.

There were several differentially expressed genes in the high FZD6 expression group and the FZD6 low-expression group. GO analysis and KEGG enrichment analysis of these genes do not show specific function and pathway changes. We speculated that FZD6 exerted its biological function through the gene set analyzed by GSEA. High-expressed FZD6 was mainly involved in 4 gene sets, including MYC targets, HEME metabolism, E2F targets, and UV response, as shown by GSEA analysis. This provided a potential direction to explore its biological functions in AML patients. MYC target gene network is estimated to account for about 15% of all human genes, involving in metabolism, mitochondrial function cell cycle regulation, protein synthesis, and ribosome biogenesis [49]. Previous research had shown that higher MYC expressions are related to poorer survival outcomes and contribute to the chemo-resistance of AML [33]. Consistently, in our study, high-expression of FZD6 can activate the WNT signaling pathway and promote the expression of MYC gene, resulting in a poor prognosis. Crosstalk between metabolic and survival pathways is critical for cellular homeostasis. Heme biosynthesis, which could reduce apoptosis through electron transport chain (ETC) activation, has been seen as an apoptotic modulator in AML [34]. As a survival-enhancing molecule of AML, Heme oxygenase-1 (HO-1) promotes tumor progression, carcinogenesis, and chemical resistance [50]. At the same time, E2F expression and/or elevated E2F target expression in tumors could cause uncontrolled proliferation and were

linked to a poor prognosis [35]. UV response mainly reflects DNA damage [51]. Recent findings highlighted that increased DNA damage and abnormal DNA damage response (DDR) were the key features of AML blasts [36], which would affect susceptibility, disease progression, and resistance to standard chemotherapy [52]. These may explain the effect of FZD6 on AML, but the specific mechanisms required further exploration.

The limitations of this study need to be equally acknowledged. Firstly, our study mainly explored the transcriptional levels of FZDs in AML and their association with prognosis. However, as GPCRs, the protein levels of FZDs are closely related to their function, which requires further exploration. This study did not verify the biological function and related mechanism of FZD6 in vitro and in vivo, and in-depth research is needed. In addition, AML patients show heterogeneity, and the one-sidedness of a single index can only provide a certain clinical reference, which requires comprehensive evaluation.

5. Conclusions

To conclude, we comprehensively analyzed the mRNA expression profiles of FZDs in AML patients and 27 AML cell lines through different online resources. Our results indicated that FZD6 was the only overexpressed molecule in FZDs of AML patients, and that high-expressed FZD6 was associated with poor/adverse risk stratification. Furthermore, FZD6 was a potential independent adverse survival factor in patients, but the predictive effect on OS could be reversed by HSCT. Overall, here we showed that the effect of FZD6 on AML may be derived from the activation of the WNT signaling pathway, and the underlying molecular mechanism demands further illustration. Our findings helped clinicians better understand the role of FZDs in AML and highlighted FZD6 as a candidate gene for AML prognosis. FZD6 is also expected to become a new therapeutic target, but this should be confirmed in future studies before direct application.

Data Availability

The data used to support the findings of this study have been deposited in the CCLE (https://portals.broadinstitute.org/ccle/home), CEPIA (http://gepia.cancer-pku.cn/), cBioPortal for Cancer Genomics (https://www.cbioportal.org/), STRING (https://string-db.org/),and GeneMANIA (http://www.genemania.org) repositories.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

This work was supported in part by research grants from the National Natural Science Foundation of China (Nos. 81870126 and 82070167).

Supplementary Materials

Supplementary 1. Supplementary Figure 1: expression of FZD1 to FZD10 in AML patients ($n = 173$) compared to normal samples ($n = 70$) in TCGA and GTEx dataset.

Supplementary 2. Supplementary Figure 2: flowchart for extraction of database.

Supplementary 3. Supplementary Figure 3: (a) FZD6 expression differences between patients ≥60 years old and < 60 years old; (b) FZD6 expression differences between males and females; (c) FZD6 expression differences between WBC $< 10 \times 10^9$/L and WBC $\geq 10 \times 10^9$/L; (d) FZD6 expression differences in patients who received different types of HSCT and no-HSCT; (e) FZD6 expression difference in FAB classifications; (f) FZD6 expression differences between patients had DMNT3A mutation ($n = 42$) or not ($n = 131$); (g) FZD6 expression differences between patients had RUNX1 mutation ($n = 15$) or not ($n = 158$). (h) FZD6 expression differences between patients had ASXL1 mutation ($n = 3$) or not ($n = 170$). (i) FZD6 expression differences between patients had IDH1 mutation ($n = 17$) or not ($n = 156$). (j) FZD6 expression differences between patients had DMNT3A mutation ($n = 16$) or not ($n = 157$). $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$.

Supplementary 4. Supplementary Figure 4: (a) the association between FZD6 and EFS of AML patients; (b) the association between HSCT and EFS of AML patients; (c) the association between FZD6 and EFS of AML patients did not receive HSCT; (d) the association between FZD6 and EFS of AML patients who received HSCT.

Supplementary 5. Supplementary Figure 5: (a) volcano plot of different gene-expression profiles between the high FZD6 and low FZD6 groups; (b) GO analysis associated with FZD6 expression; (c) top 20 KEGG pathway associated with FZD6 expression.

Supplementary 6. Supplementary Figure 6: the relationship between the expression of FZD6 and prognosis in pan-cancer.

Supplementary 7. Supplementary Table 1: coexpressed genes of FZD6 based on the cBioPortal dataset.

References

[1] N. J. Short, M. E. Rytting, and J. E. Cortes, "Acute myeloid leukaemia," *Lancet*, vol. 392, no. 10147, pp. 593–606, 2018.

[2] H. Dohner, D. J. Weisdorf, and C. D. Bloomfield, "Acute myeloid leukemia," *The New England Journal of Medicine*, vol. 373, no. 12, pp. 1136–1152, 2015.

[3] J. J. Cornelissen and D. Blaise, "Hematopoietic stem cell transplantation for patients with AML in first complete remission," *Blood*, vol. 127, no. 1, pp. 62–70, 2016.

[4] G. Schulte, "International Union of Basic and Clinical Pharmacology. LXXX. The class frizzled receptors," *Pharmacological Reviews*, vol. 62, no. 4, pp. 632–667, 2010.

[5] P. Bhanot, M. Brink, C. H. Samos et al., "A new member of the _frizzled_ family from _Drosophila_ functions as a wingless receptor," *Nature*, vol. 382, no. 6588, pp. 225–230, 1996.

[6] B. T. Mac Donald and X. He, "Frizzled and LRP5/6 receptors for Wnt/beta-catenin signaling," *Cold Spring Harbor Perspectives in Biology*, vol. 4, no. 12, article a007880, 2012.

[7] Q. Ren, J. Chen, and Y. Liu, "LRP5 and LRP6 in Wnt signaling: similarity and divergence," *Frontiers in Cell and Development Biology*, vol. 9, article 670960, 2021.

[8] S. Suthon, R. S. Perkins, V. Bryja, G. A. Miranda-Carboni, and S. A. Krum, "WNT5B in physiology and disease," *Frontiers in Cell and Development Biology*, vol. 9, article 667581, 2021.

[9] M. K. Yum, S. Han, J. Fink et al., "Tracing oncogene-driven remodelling of the intestinal stem cell niche," *Nature*, vol. 594, no. 7863, pp. 442–447, 2021.

[10] R. Nusse and H. Clevers, "Wnt/β-catenin signaling, disease, and emerging therapeutic modalities," *Cell*, vol. 169, no. 6, pp. 985–999, 2017.

[11] W. J. Chae and A. L. M. Bothwell, "Canonical and non-canonical Wnt signaling in immune cells," *Trends in Immunology*, vol. 39, no. 10, pp. 830–847, 2018.

[12] D. J. Flanagan, N. Barker, N. S. D. Costanzo et al., "Frizzled-7Is required for Wnt signaling in gastric tumors with and without Apc mutations," *Cancer Research*, vol. 79, no. 5, pp. 970–981, 2019.

[13] Q. Li, L. Ye, X. Zhang et al., "FZD8, a target of p53, promotes bone metastasis in prostate cancer by activating canonical Wnt/β-catenin signaling," *Cancer Letters*, vol. 402, pp. 166–176, 2017.

[14] P. Ruenraroengsak, D. Kiryushko, I. G. Theodorou et al., "Frizzled-7-targeted delivery of zinc oxide nanoparticles to drug-resistant breast cancer cells," *Nanoscale*, vol. 11, no. 27, pp. 12858–12870, 2019.

[15] S. Nagayama, C. Fukukawa, T. Katagiri et al., "Therapeutic potential of antibodies against FZD10, a cell-surface protein, for synovial sarcomas," *Oncogene*, vol. 24, no. 41, pp. 6201–6212, 2005.

[16] N. I. Khan, K. F. Bradstock, and L. J. Bendall, "Activation of Wnt/β-catenin pathway mediates growth and survival in B-cell progenitor acute lymphoblastic leukaemia," *British Journal of Haematology*, vol. 138, no. 3, pp. 338–348, 2007.

[17] S. Hamdoun, E. Fleischer, A. Klinger, and T. Efferth, "Lawsone derivatives target the Wnt/β-catenin signaling pathway in multidrug- resistant acute lymphoblastic leukemia cells," *Biochemical Pharmacology*, vol. 146, pp. 63–73, 2017.

[18] H. O. Sercan, M. Pehlivan, O. Simsek, H. Ates, and Z. Sercan, "Induction of apoptosis increases expression of non-canonical WNT genes in myeloid leukemia cell lines," *Oncology Reports*, vol. 18, no. 6, pp. 1563–1569, 2007.

[19] P. Agarwal, B. Zhang, Y. Ho et al., "Enhanced targeting of CML stem and progenitor cells by inhibition of porcupine acyltransferase in combination with TKI," *Blood*, vol. 129, no. 8, pp. 1008–1020, 2017.

[20] Q. L. Wu, C. Zierold, and E. A. Ranheim, "Dysregulation of frizzled 6 is a critical component of B-cell leukemogenesis in a mouse model of chronic lymphocytic leukemia," *Blood*, vol. 113, no. 13, pp. 3031–3039, 2009.

[21] D. Lu, Y. Zhao, R. Tawatao et al., "Activation of the Wnt signaling pathway in chronic lymphocytic leukemia," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 9, pp. 3118–3123, 2004.

[22] M. Kaucká, K. Plevová, Š. Pavlová et al., "The planar cell polarity pathway drives pathogenesis of chronic lymphocytic leukemia by the regulation of B-lymphocyte migration," *Cancer Research*, vol. 73, no. 5, pp. 1491–1501, 2013.

[23] R. Mathur, L. Sehgal, F. K. Braun et al., "Targeting Wnt pathway in mantle cell lymphoma-initiating cells," *Journal of Hematology & Oncology*, vol. 8, no. 1, p. 63, 2015.

[24] L. Tickenbrock, S. Hehn, B. Sargin et al., "Activation of Wnt signalling in acute myeloid leukemia by induction of frizzled-4," *International Journal of Oncology*, vol. 33, no. 6, pp. 1215–1221, 2008.

[25] J. F. Lin, H. Zeng, and J. Q. Zhao, "miR-212-5p regulates the proliferation and apoptosis of AML cells through targeting FZD5," *European Review for Medical and Pharmacological Sciences*, vol. 22, no. 23, pp. 8415–8422, 2018.

[26] Y. H. Wang, Y. Imai, M. Shiseki, J. Tanaka, and T. Motoji, "Knockdown of the Wnt receptor Frizzled-1 (FZD1) reduces _MDR1_ /P-glycoprotein expression in multidrug resistant leukemic cells and inhibits leukemic cell proliferation," *Leukemia Research*, vol. 67, pp. 99–108, 2018.

[27] M. Ghandi, F. W. Huang, J. Jané-Valbuena et al., "Next-generation characterization of the cancer cell line encyclopedia," *Nature*, vol. 569, no. 7757, pp. 503–508, 2019.

[28] Z. Tang, C. Li, B. Kang, G. Gao, C. Li, and Z. Zhang, "GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses," *Nucleic Acids Research*, vol. 45, no. W1, pp. W98–W102, 2017.

[29] E. Cerami, J. Gao, U. Dogrusoz et al., "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data," *Cancer Discovery*, vol. 2, no. 5, pp. 401–404, 2012.

[30] D. Szklarczyk, A. L. Gable, K. C. Nastou et al., "The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets," *Nucleic Acids Research*, vol. 49, no. D1, pp. D605–D612, 2021.

[31] D. Warde-Farley, S. L. Donaldson, O. Comes et al., "The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function," *Nucleic Acids Research*, vol. 38, Supplement 2, pp. W214–W220, 2010.

[32] J. Li and Z. Ge, "High HSPA8 expression predicts adverse outcomes of acute myeloid leukemia," *BMC Cancer*, vol. 21, no. 1, p. 475, 2021.

[33] M. Ohanian, U. Rozovski, R. Kanagal-Shamanna et al., "MYC protein expression is an important prognostic factor in acute myeloid leukemia," *Leukemia & Lymphoma*, vol. 60, no. 1, pp. 37–48, 2019.

[34] K. H. Lin, A. Xie, J. C. Rutter et al., "Systematic dissection of the metabolic-apoptotic Interface in AML reveals Heme biosynthesis to be a regulator of drug sensitivity," *Cell Metabolism*, vol. 29, no. 5, pp. 1217–1231.e7, 2019.

[35] L. N. Kent and G. Leone, "The broken cycle: E2F dysfunction in cancer," *Nature Reviews. Cancer*, vol. 19, no. 6, pp. 326–338, 2019.

[36] R. Pennisi, J. Albanesi, P. Ascenzi, C. Nervi, and A. di Masi, "Are DNA damage response kinases a target for the differentiation treatment of acute myeloid leukemia?," *IUBMB Life*, vol. 70, no. 11, pp. 1057–1066, 2018.

[37] C. M. Zeng, Z. Chen, and L. Fu, "Frizzled receptors as potential therapeutic targets in human cancers," *International Journal of Molecular Sciences*, vol. 19, no. 5, p. 1543, 2018.

[38] S. Sun, J. Wang, J. Liu et al., "MiR-302b suppresses tumor metastasis by targeting frizzled 6in OSCC," *Journal of Dental Research*, vol. 100, no. 7, pp. 739–745, 2021.

[39] D. Zhang, S. Ma, C. Zhang et al., "microRNA-935 directly targets FZD6 to inhibit the proliferation of human glioblastoma and correlate to glioma malignancy and prognosis," *Frontiers in Oncology*, vol. 11, article 566492, 2021.

[40] Y. Li, Z. Liu, and Y. Zhang, "Expression and prognostic impact of FZDs in pancreatic adenocarcinoma," *BMC Gastroenterology*, vol. 21, no. 1, p. 79, 2021.

[41] A. Cassaro, G. Grillo, M. Notaro et al., "FZD6 triggers Wnt–signalling driven by WNT10BIVS1 expression and highlights new targets in T-cell acute lymphoblastic leukemia," *Hematological Oncology*, vol. 39, no. 3, pp. 364–379, 2021.

[42] Y. Yuan, Q. Wang, S. L. Ma et al., "lncRNA PCAT-1 interacting with FZD6 contributes to the malignancy of acute myeloid leukemia cells through activating Wnt/beta-catenin signaling pathway," *American Journal of Translational Research*, vol. 11, no. 11, pp. 7104–7114, 2019.

[43] M. Stelljes, U. Krug, D. W. Beelen et al., "Allogeneic transplantation versus chemotherapy as postremission therapy for acute myeloid leukemia: a prospective matched pairs analysis," *Journal of Clinical Oncology*, vol. 32, no. 4, pp. 288–296, 2014.

[44] X. Xin, H. Zhu, Z. Chang et al., "Risk factors and prognosis analysis of acute myeloid leukemia in children," *Journal of BUON*, vol. 26, no. 1, pp. 166–172, 2021.

[45] K. M. Cadigan and M. L. Waterman, "TCF/LEFs and Wnt signaling in the nucleus," *Cold Spring Harbor Perspectives in Biology*, vol. 4, no. 11, article a007906, 2012.

[46] G. Corda and A. Sala, "Non-canonical WNT/PCP signalling in cancer: Fzd6 takes centre stage," *Oncogene*, vol. 6, no. 7, article e364, 2017.

[47] M. T. Butler and J. B. Wallingford, "Planar cell polarity in development and disease," *Nature Reviews. Molecular Cell Biology*, vol. 18, no. 6, pp. 375–388, 2017.

[48] A. M. Gruszka, D. Valli, and M. Alcalay, "Wnt signalling in acute myeloid leukaemia," *Cells*, vol. 8, no. 11, p. 1403, 2019.

[49] C. V. Dang, K. A. O'Donnell, K. I. Zeller, T. Nguyen, R. C. Osthus, and F. Li, "The c-Myc target gene network," *Seminars in Cancer Biology*, vol. 16, no. 4, pp. 253–264, 2006.

[50] L. Salerno, G. Romeo, M. N. Modica et al., "Heme oxygenase-1: a new druggable target in the management of chronic and acute myeloid leukemia," *European Journal of Medicinal Chemistry*, vol. 142, pp. 163–178, 2017.

[51] M. Kciuk, B. Marciniak, M. Mojzych, and R. Kontek, "Focus on UV-induced DNA damage and repair-disease relevance and protective strategies," *International Journal of Molecular Sciences*, vol. 21, no. 19, p. 7264, 2020.

[52] M. T. Esposito and C. W. So, "DNA damage accumulation and repair defects in acute myeloid leukemia: implications for pathogenesis, disease progression, and chemotherapy resistance," *Chromosoma*, vol. 123, no. 6, pp. 545–561, 2014.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 7397250, 16 pages
https://doi.org/10.1155/2022/7397250

Research Article

PDP1 Promotes Cell Malignant Behavior and Is Associated with Worse Clinical Features in Ovarian Cancer Patients: Evidence from Bioinformatics and In Vitro Level

Yan Song,[1] Juan Zhang,[1] Lei Zhang,[2] Suxia Zhang,[1] and Chengcheng Shen[1]

[1]*Jinan Maternity and Child Care Hospital Affiliated to Shandong First Medical University, Jinan, Shandong, China*
[2]*Liaocheng People's Hospital, Liaocheng, Shandong, China*

Correspondence should be addressed to Chengcheng Shen; chengcheng20170324@163.com

Received 21 July 2022; Accepted 6 September 2022; Published 14 October 2022

Academic Editor: Jincheng Wang

PDP1 has been reported in multiple diseases. However, it has not been fully explored in ovarian cancer (OC). The public data was downloaded from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. Differentially expressed gene analysis was conducted out using the limma package. Prognosis analysis was performed using the survival package. Gene Set Enrichment Analysis (GSEA) was performed using the fgsea package. Immune infiltration analysis was performed based on the CIBERSORT algorithm. CCK8 assay was used to evaluate the cell proliferation ability of cancer cells. Transwell assay was used for the invasion and migration ability. Our result showed that PDP1 was overexpressed in OC tissue in RNA and protein level based on multiple databases (TCGA, GSE18520, GSE27651, and GSE54388). At the same time, we found PDP1 was correlated with poor prognosis and worse clinical parameters. In vitro experiment showed that PDP1 could significantly promote proliferation, invasion, and migration ability of OC cells. GSEA analysis showed that in the OC patients with high PDP1 expression, the pathway of IL6/JAK/STAT3 signaling, interferon-alpha response, apoptosis, adipogenesis, KRAS signaling, and IL2/STAT5 signaling was activated, which might be responsible for its oncogenic effect in OC. Immune infiltration analysis indicated that PDP1 was positively correlated with activated myeloid dendritic cells, resting CD4 memory T cells, neutrophil, and M1 and M2 macrophages, yet negatively correlated with M0 macrophages, plasma B cells, $\gamma\delta$T cells, and activated CD4 memory T cells. Drug sensitivity analysis showed a negative correlation between PDP1 expression and the IC50 of bleomycin and gemcitabine, yet a positive correlation of cisplatin, indicating that the OC patients with high PDP1 expression might be more sensitive to bleomycin and gemcitabine and more resistant to cisplatin. PDP1 could facilitate OC progression and is associated with patient prognosis and chemosensitivity, making it an underlying biomarker of OC.

1. Introduction

Ovarian cancer (OC) is a common malignancy with a rising trend of morbidity in females, making it a major health concern worldwide [1]. As a heterogeneous disease, OC has characteristics of high incidence and high mortality rate. During the past decades, the five-year survival rate of OC remained below 50%, and no significant improvement was observed along with the advancement of medical conditions [2]. Meanwhile, due to the insidious symptoms, the early deletion of OC is difficult to complete, and a significant number of patients presented advanced stage at diagnosis [3]. Therefore, it is meaningful to identify novel and effective biomarkers for OV early detection and treatment.

Pyruvate dehydrogenase phosphatase catalytic subunit 1 (PDP1) encodes the protein that is one of the three components (E1, E2, and E3) of the large pyruvate dehydrogenase complex [4]. PDP1 plays an important role in protein phosphorylation

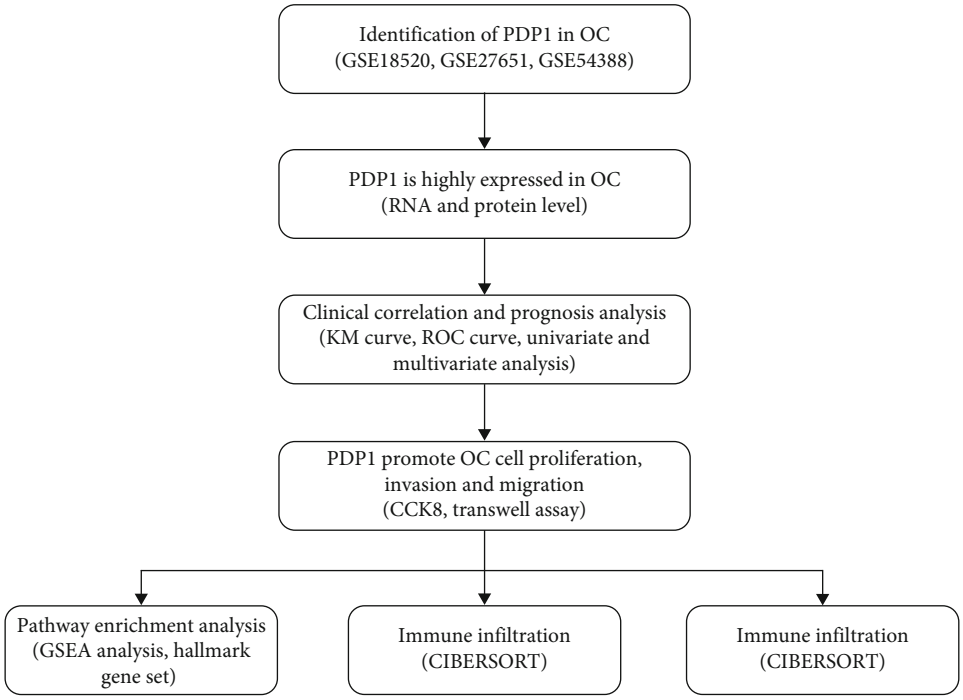


Figure 1: The flow chart of the whole study.

[4, 5]. The biological role of PDP1 was also observed in multiple diseases. Shi and colleagues demonstrated that the PDP1-PDH-histone acetylation retrograde signaling activated by mitochondrial dysfunction could significantly increase the radioresistance in colorectal cancer [6]. Feng and colleagues found that miR-18a-3p could improve cartilage matrix remodeling and inhibit inflammation in osteoarthritis by suppressing PDP1 [7]. In pancreatic cancer, Li and colleagues found that PDP1 promotes pancreatic cancer proliferation and invasion through regulating the MAPK/mTOR signaling pathway [8]. Shan and colleagues revealed that phosphorylation of the Tyr-94 site could hamper PDP1 expression and facilitate the growth of leukemia cells [9]. A comprehensive review conducted by Jeoung concluded that pyruvate dehydrogenase kinases (PDKs) and pyruvate dehydrogenase phosphatases (PDPs) have unique tissue-specific expression, kinetic properties, and sensitivity to regulatory molecules, which might be therapeutic targets for diabetes and cancers [10]. Chen and colleagues PDP1 was associated with the prognosis of breast cancer patients and their immune microenvironment [11]. Moreover, Chen and colleagues found that PDP1 was overexpressed in prostate cancer and could sustain prostate tumorigenesis by controlling lipid biosynthesis [12]. However, the underlying role of PDP1 in OC has been thoroughly studied.

Our study systematically investigated the role of PDP1 in OC. PDP1 was upregulated in OC tissue and cell lines, which was associated with worse clinical features and poor prognosis. In vitro experiment showed that PDP1 could promote proliferation, invasion, and migration of OC cells. GSEA and immune infiltration analysis were performed to explore the underlying pathway and immune microenvironment difference between low and high PDP1 patients. Moreover, we found that PDP1 could affect the sensitivity of chemotherapy, making it a potential therapeutic target of OC patients.

2. Results

2.1. Identification of the Common Differentially Expressed Genes (DEGs) between OC and Normal Tissue. The flow chart of the whole study was shown in Figure 1. Firstly, we performed DEG analysis between the OC and normal ovarian tissue based on the data of GSE18520, GSE27651, and GSE54388. GSE18520 identified 1009 upregulated and 1554 downregulated genes (Figure 2(a)); GSE27651 identified 1966 upregulated and 2553 downregulated genes (Figure 2(b)); GSE54388 identified 625 upregulated and 743 downregulated genes (Figure 2(c)). Meanwhile, we performed univariate Cox regression analysis to identify the prognosis-related genes. Commonly upregulated genes intersect the genes with $HR > 1$ and finally identified five genes, KLHL14, SLC4A11, S100A2, PDP1, and TMC4 (Figure 2(d)). Commonly downregulated genes intersect the genes with $HR < 1$ and finally identified six genes, ITLN1, HSD17B2, FGF13, ME1, TNFSF13B, and AADAC (Figure 2(e)).

2.2. Exploration of the Expression Pattern of PDP1 in OC. Pan-cancer analysis based on TCGA and GTEx data showed that PDP1 has an aberrant expression analysis in most cancers, including OC (Figure 3(a)). Also, we observed a higher expression level of PDP1 in OC tumor tissue in multiple cohorts, consisting of TCGA + GTEx, GSE18520, GSE27651, and GSE54388 (Figures 3(b)–3(e)). Further, we evaluate the protein level of PDP1 through the IHC images obtained from the HPA database. The result indicated that

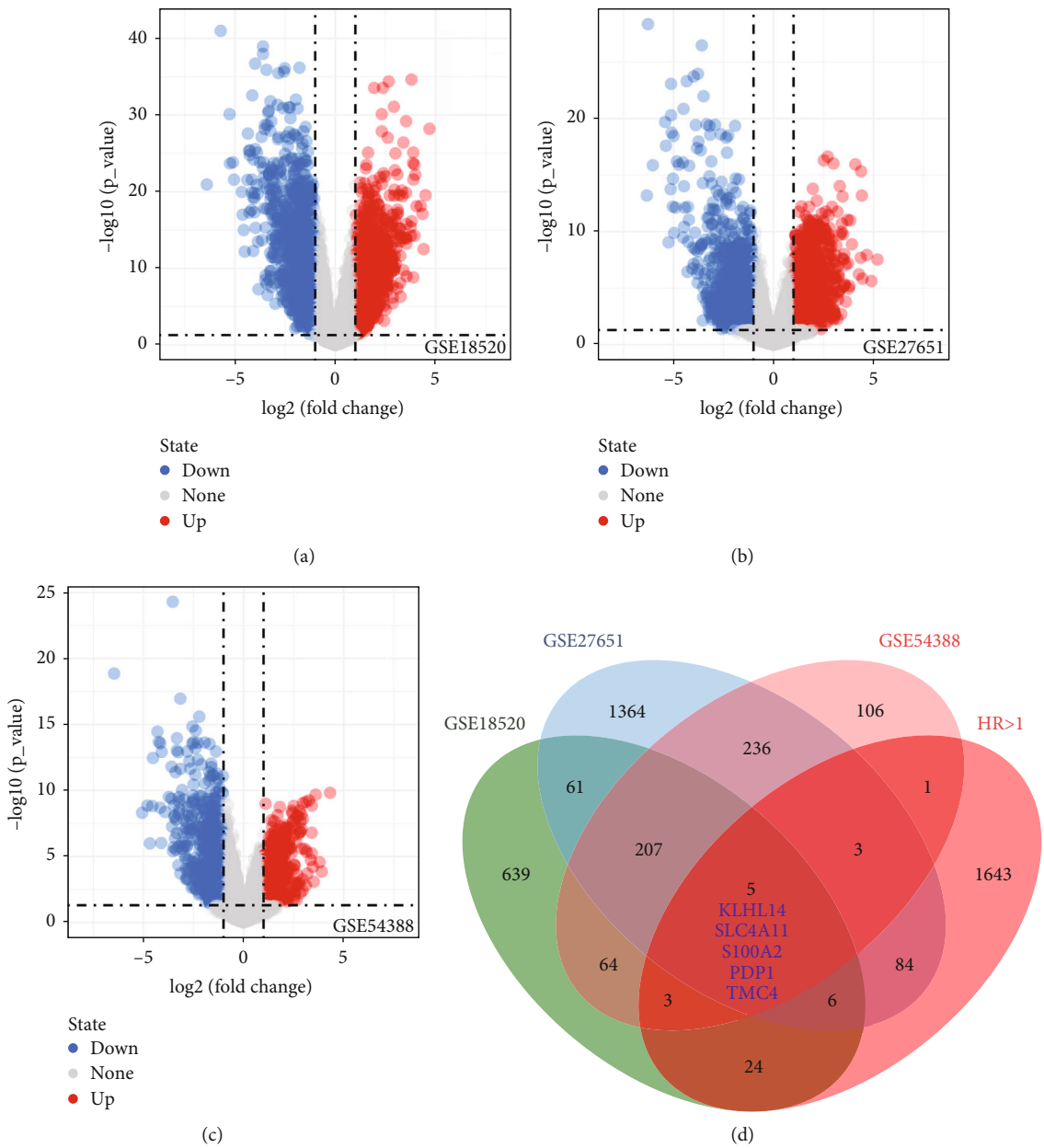
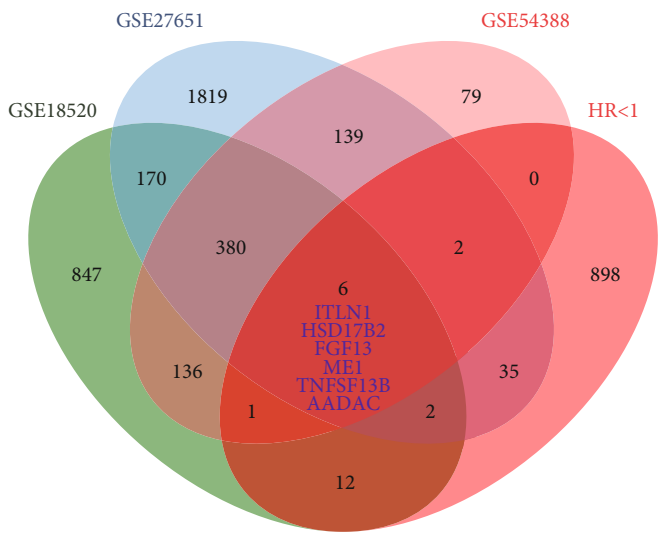


Figure 2: Continued.



(e)

Figure 2: Identification of differentially expressed genes between ovarian cancer and normal tissue. Notes: (a) differentially expressed gene analysis was performed under the threshold of $|\log FC > 1|$ and P value < 0.05 in GSE18520 database; (b) differentially expressed gene analysis was performed under the threshold of $|\log FC > 1|$ and P value < 0.05 in GSE27651 database; (c) differentially expressed gene analysis was performed under the threshold of $|\log FC > 1|$ and P value < 0.05 in GSE54388 database; (d) five genes were commonly upregulated in GSE18520, GSE27651, and GSE54388 databases, which were also risk factors; (e) six genes were commonly downregulated in GSE18520, GSE27651, and GSE54388 databases, which were also protective factors.

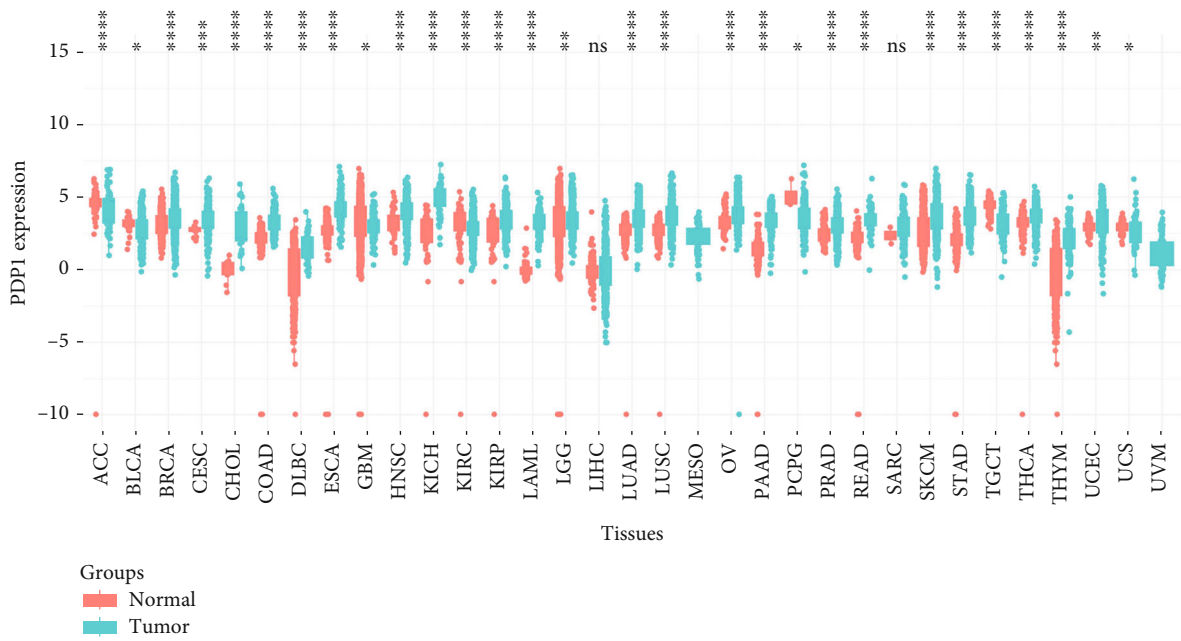
PDP1 also had a higher protein level in OC tumor tissue compared with the normal tissue (Figures 3(f) and 3(g)).

2.3. PDP1 Might Be Prognosis and Clinical Biomarker of OC. Next, we evaluate the PDP1 expression in OC patients with diverse clinical features. We found that the OC patients with lymphatic invasion, venous invasion, and worse clinical-stage tend to have a higher PDP1 expression (Figures 4(a)–4(c)). Although the statistical was not significant, considering the small sample bias, we still considered that the PDP1 was associated with worse clinical features. Moreover, we found that the OC patients with higher PDP1 level have a poor OS, DSS, and PFI (Figures 4(d)–4(f); overall survival, HR = 1.32, $P = 0.039$; disease-specific survival, HR = 1.27, $P = 0.1$; progress-free interval, HR = 1.24, $P = 0.071$). Furthermore, a receiver-operating characteristic (ROC) curve was used for quantifying the predictive ability of patient prognosis. The result showed that PDP1 might have a satisfactory prognosis prediction ability for the OS, DSS, and PFI of OC patients (Figure 4(g), OS, 1-year AUC: 0.580, 3-year AUC: 0.725, 5-year AUC: 0.700; Figure 4(h), DSS, 1-year AUC: 0.558, 3-year AUC: 0.723, 5-year AUC: 0.697; Figure 4(i), 1-year AUC: 0.627, 3-year AUC: 0.760, 5-year AUC: 0.787). Univariate and multivariate analysis indicated that PDP1 is a risk factor independent of other clinical features (Figures 4(j) and 4(k)).

2.4. PDP1 Promotes Proliferation, Invasion, and Migration of OC Cells. To explore the underlying biological role of PDP1 in OC, we performed in vitro experiments to assess the influence of PDP1 on OC cell malignant behaviors. The qRT-PCR result showed a higher PDP1 level in OC cell lines than the normal IOSE80 cells (Figure 5(a)). An effective knockdown of PDP1 was also observed in selected A2780 and SKOV3 cell lines (Figures 5(b) and 5(c)). CCK8 assay showed that the inhibition of PDP1 could significantly suppress the proliferation ability of OC cells (Figures 5(d) and 5(e)). Moreover, the Transwell assay showed that the knockdown the PDP1 could remarkably inhibit the invasion and migration of OV cells (Figure 5(f)).

2.5. PDP1 Activated Multiple Oncogenic Signaling in OC. GSEA analysis was performed to explore the potential biological difference between low and high PDP1 OC patients. The result showed that in the OC patients with high PDP1 expression, the pathway of IL6/JAK/STAT3 signaling, interferon-alpha response, apoptosis, adipogenesis, KRAS signaling, and IL2/STAT5 signaling was activated, which might be responsible for its oncogenic effect in OC (Figure 6(a)). Gene Ontology (GO) analysis in the patients with high PDP1 expression, the terms of cellular response to prostaglandin E stimulus, negative regulation of activated T cell proliferation, positive regulation of macrophage differentiation, regulation of extracellular matrix assembly, and regulation of membrane invagination were significantly enriched (Figure 6(b)). Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis showed that in the patients with high PDP1 expression, terms of glycosphingolipid biosynthesis lacto and neolacto series, allograft rejection, graft versus host disease, and colorectal cancer were mainly enriched in (Figure 6(c)).

2.6. PDP1 Acted as an Immune-Related Gene and Was Associated with Chemosensitivity. The tumor immune microenvironment plays an important role in tumorigenesis and tumor development. Therefore, we explored the effect of



(a)

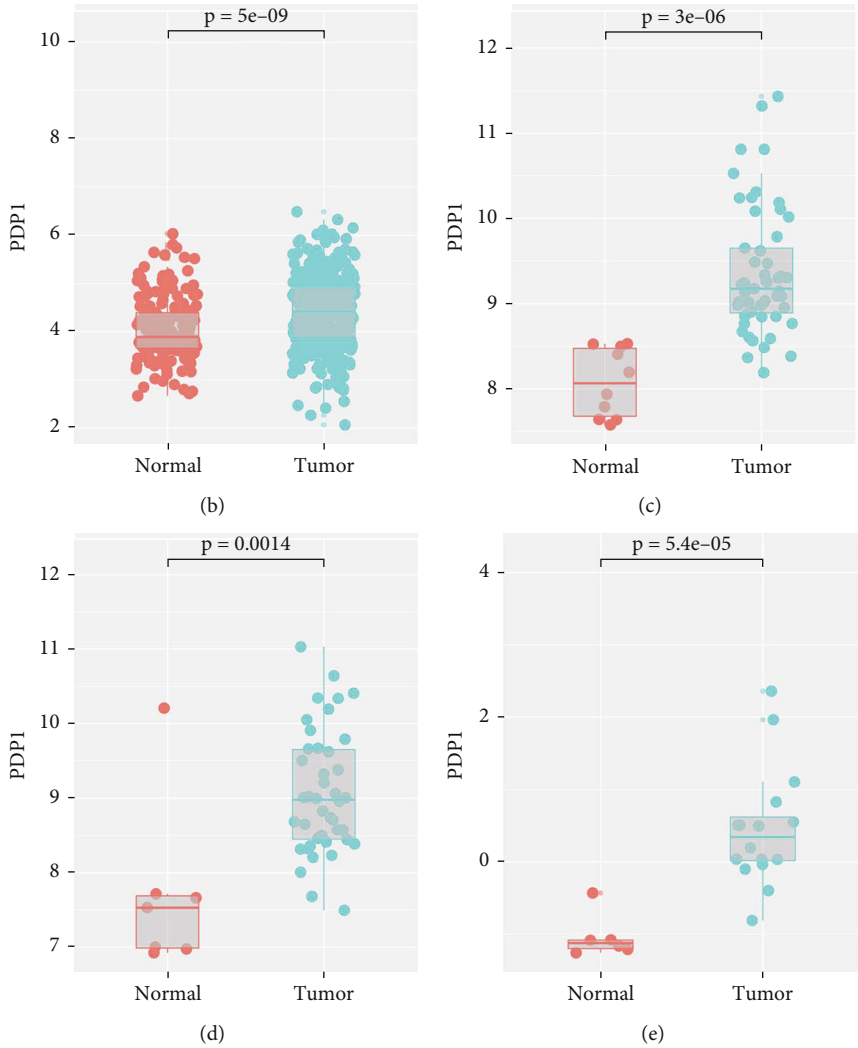
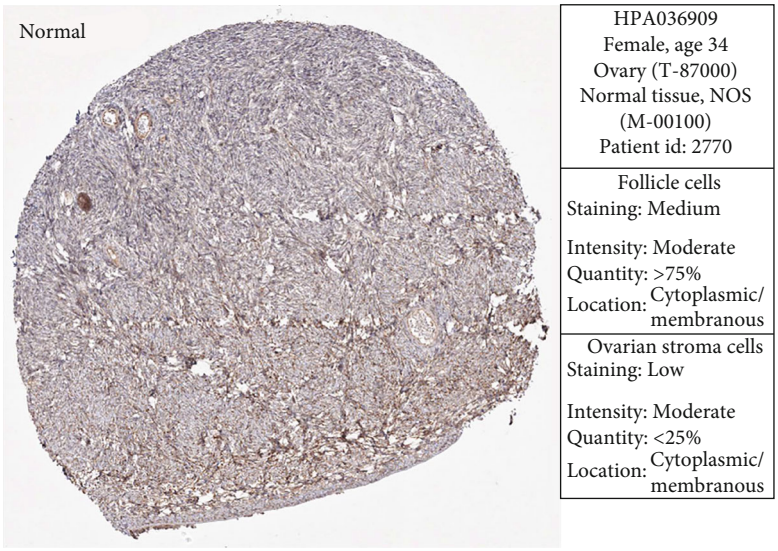
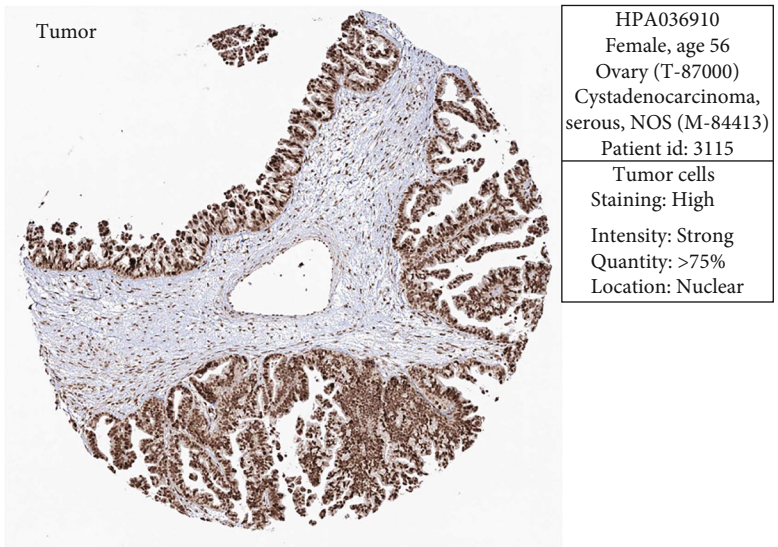


Figure 3: Continued.



(f)



(g)

Figure 3: PDP1 was overexpressed in ovarian cancer tissue. Notes: (a): the expression pattern of PDP1 in pan-cancer; (b–e): PDP1 was overexpressed in ovarian cancer tissue compared with the normal tissue in TCGA+GTEx, GSE18520, GSE27651, and GSE54388 databases; (f, g): the representative immunohistochemistry image of PDP1 in ovarian cancer and normal tissue.

PDP1 on the tumor immune microenvironment. CIBERSORT algorithm was used for immune cell quantification. The result showed that PDP1 was positively correlated with activated myeloid dendritic cells, resting CD4 memory T cells, neutrophil, and M1 and M2 macrophages, yet negatively correlated with M0 macrophages, plasma B cells, $\gamma\delta$T cells, and activated CD4 memory T cells (Figure 7(a)). Bleomycin, gemcitabine, cisplatin, and paclitaxel were the frequent chemotherapy regimen of OC. We further investigated the influence of PDP1 on the chemosensitivity of these drugs based on the GDSC database. Drug sensitivity analysis showed a negative correlation between PDP1 expression and the IC50 of bleomycin and gemcitabine, indicating that the OC patients with high PDP1 expression might be more sensitive to bleomycin and gemcitabine (Figures 7(b) and 7(c); bleomycin: $R = -0.23$, $P < 0.001$; gemcitabine: $R = -0.10$, $P = 0.048$). In contrast, cisplatin showed opposite results (Figure 7(d); cisplatin: $R = 0.35$, $P < 0.001$). Meanwhile, no significant effect was observed in paclitaxel (Figure 7(e)).

3. Discussion

As one of the most threatening malignant tumors in females, OC still has high incidence and mortality despite the advancement of medicine [13]. For advanced OC, the therapeutic effects are not satisfactory currently. Targeted therapy is a promising therapeutic strategy for progressive OC. Therefore, it is necessary to identify the novel target with potential for clinical translation.

To the best of our knowledge, this study was the first study comprehensively exploring the role of PDP1 in OC. Here, we found that PDP1 was overexpressed in OC tissue

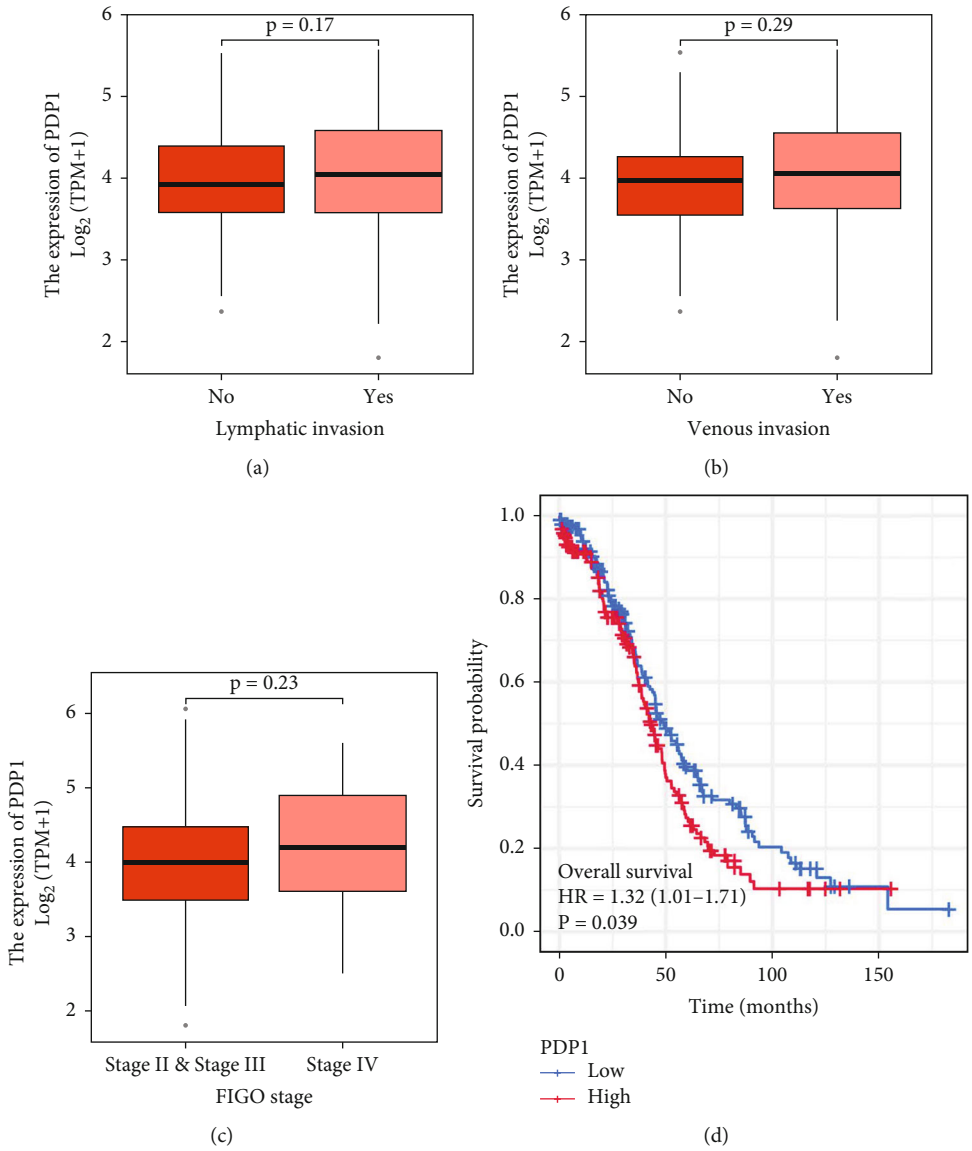


Figure 4: Continued.

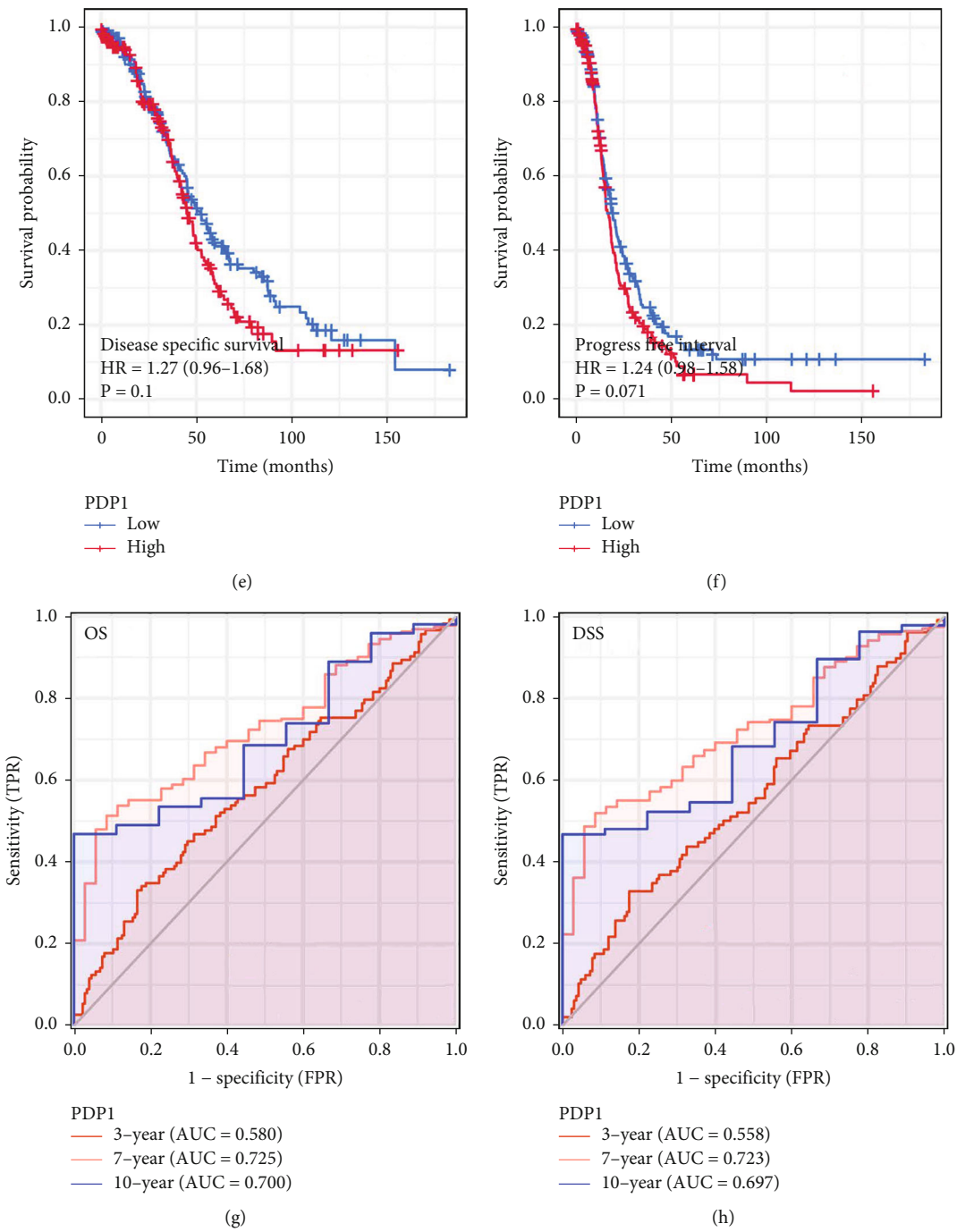


Figure 4: Continued.

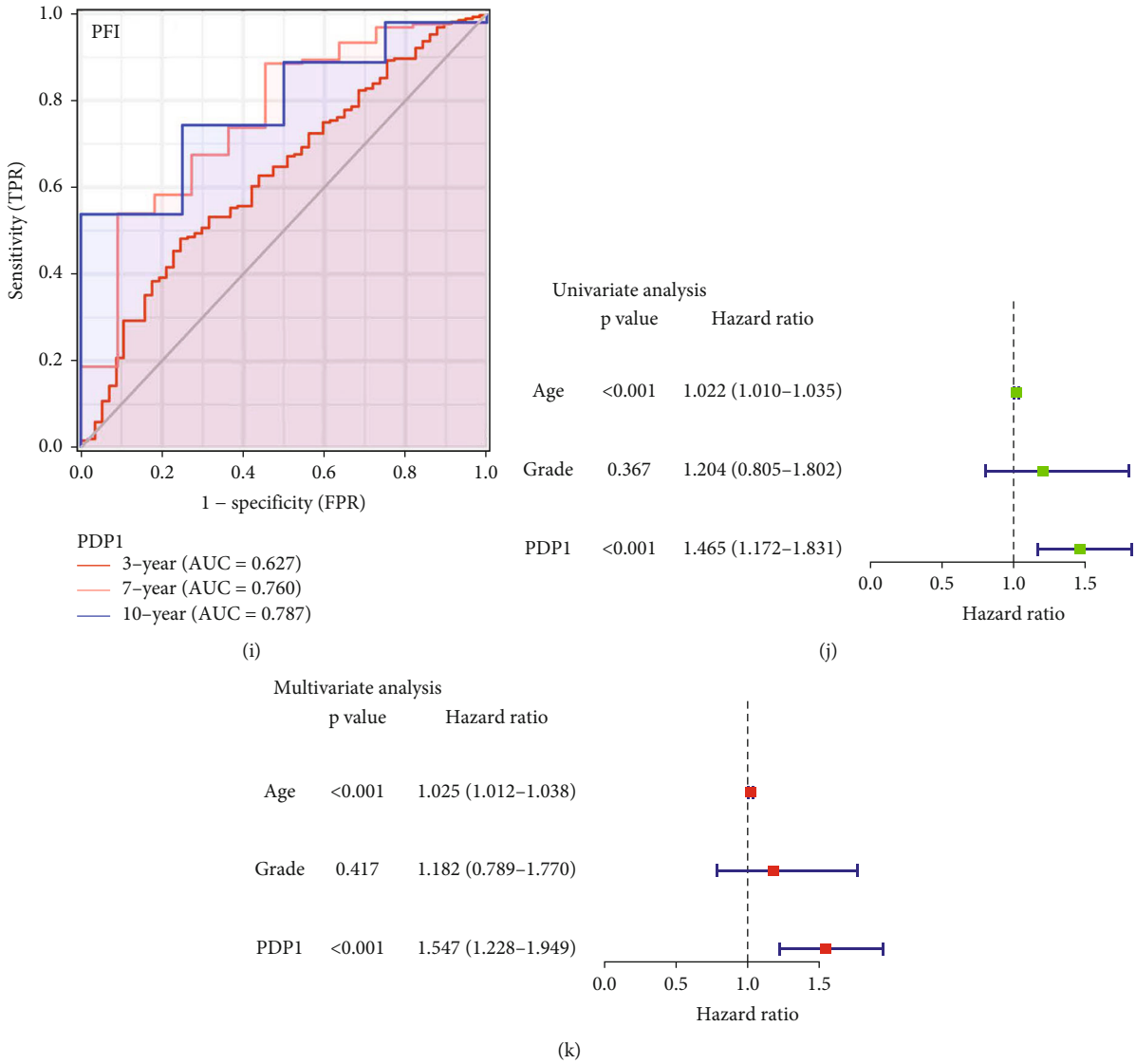


Figure 4: Clinical and prognosis correlation of PDP1 in ovarian cancer. Notes: (a) the expression level of PDP1 in patients with or without lymphatic invasion; (b) the expression level of PDP1 in patients with or without venous invasion; (c) the expression level of PDP1 in stage II/III and stage IV patients; (d–f) the prognosis difference between PDP1 low and high ovarian cancer patients (OS, DSS, and PFI); (g–i) the prognosis prediction efficiency of PDP1 in OS, DSS, and PFI; (j) univariate analysis of PDP1; (k) multivariate analysis of PDP1.

and cells, which was associated with more progressive clinical features and poor prognosis. Moreover, the knockdown of PDP1 could significantly suppress the cell malignant behavior of OC cells. Multiple oncogenic pathways were abnormally activated in high PDP1 patients. Meanwhile, PDP1 was remarkably correlated with several immune cells and might affect the chemosensitivity of OC patients.

PDP1 is a key regulator of pyruvate dehydrogenases complex (PDC) and can positively regulate the catalytic activity of PDC through mediating the dephosphorylation from the serine sites on E1α of the complex [14]. After being activated by PDP1, PDC could catalyze the oxidative decarboxylation of pyruvate to acetyl-CoA, which is a major energy-producing substrate [10]. Aberrant PDP1 level might lead to cell energy metabolism disorders, affecting cell malignant biological behaviors. The cancer-promoting effect of PDP1 on cancer cells was observed in multiple cancers, including colon cancer, prostate cancer, and nonsolid tumor [6, 9, 12]. Our study fills the gap of PDP1 in OC, and the result showed that PDP1 could promote OC progression, making it an underlying therapeutic target.

GSEA result showed that the pathway of IL6/JAK/STAT3 signaling, interferon-alpha response, apoptosis, adipogenesis, KRAS signaling, and IL2/STAT5 signaling was activated in high PDP1 patients. IL6/JAK/STAT3 signaling was reported to facilitate cancer progression in multiple cancer. Ni and colleagues found that miR-515-5p could suppress the migration and invasion of liver cancer cells through targeting IL6/JAK/STAT3 pathway [15]. In gastric cancer, Zhao and colleagues indicated that IL6 could promote proliferation, invasion, and lymphangiogenesis by regulating the JAK/STAT3/VEGF-C signaling [16]. Moreover, Salimian and colleagues found that the omental adipose stromal cells could increase the nitric oxide level in OC cells,

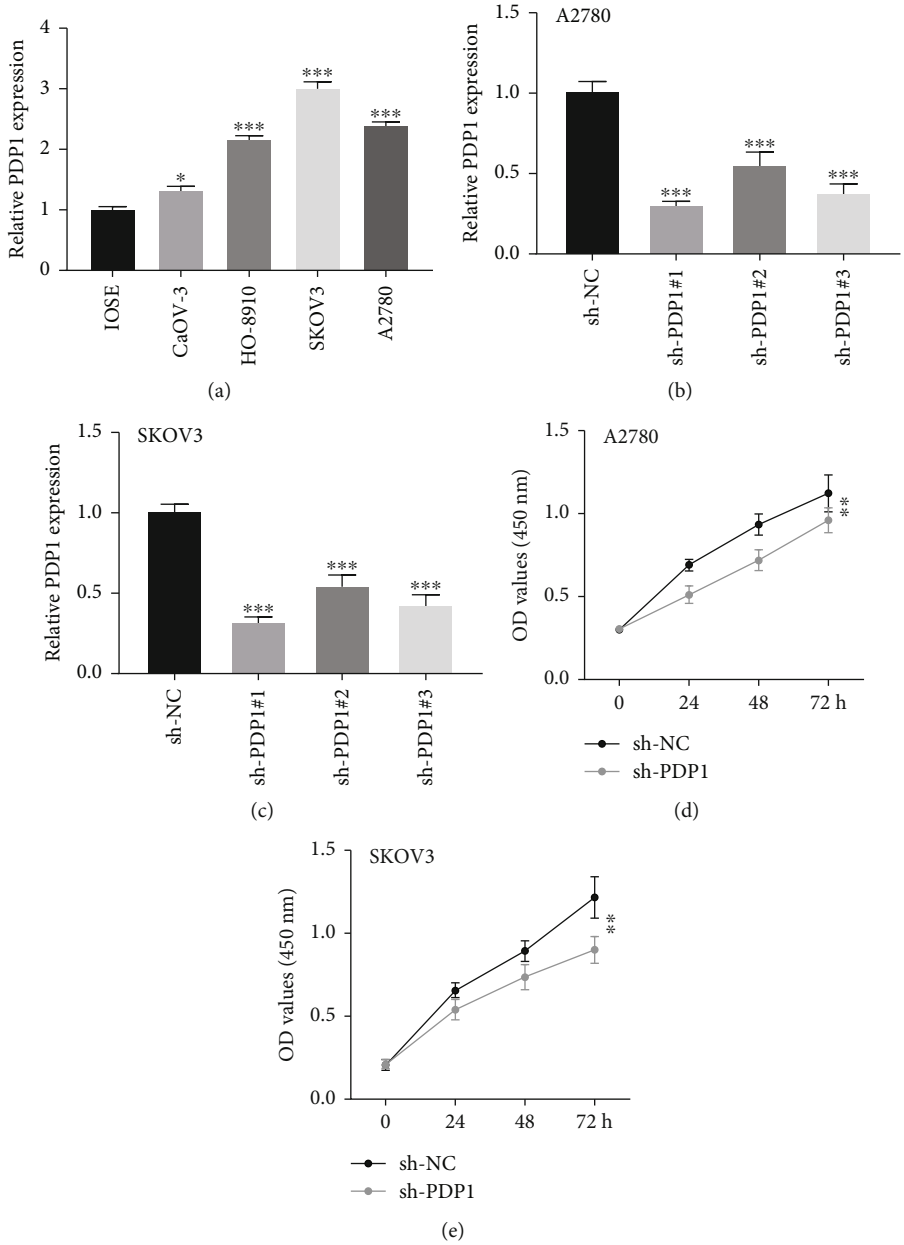
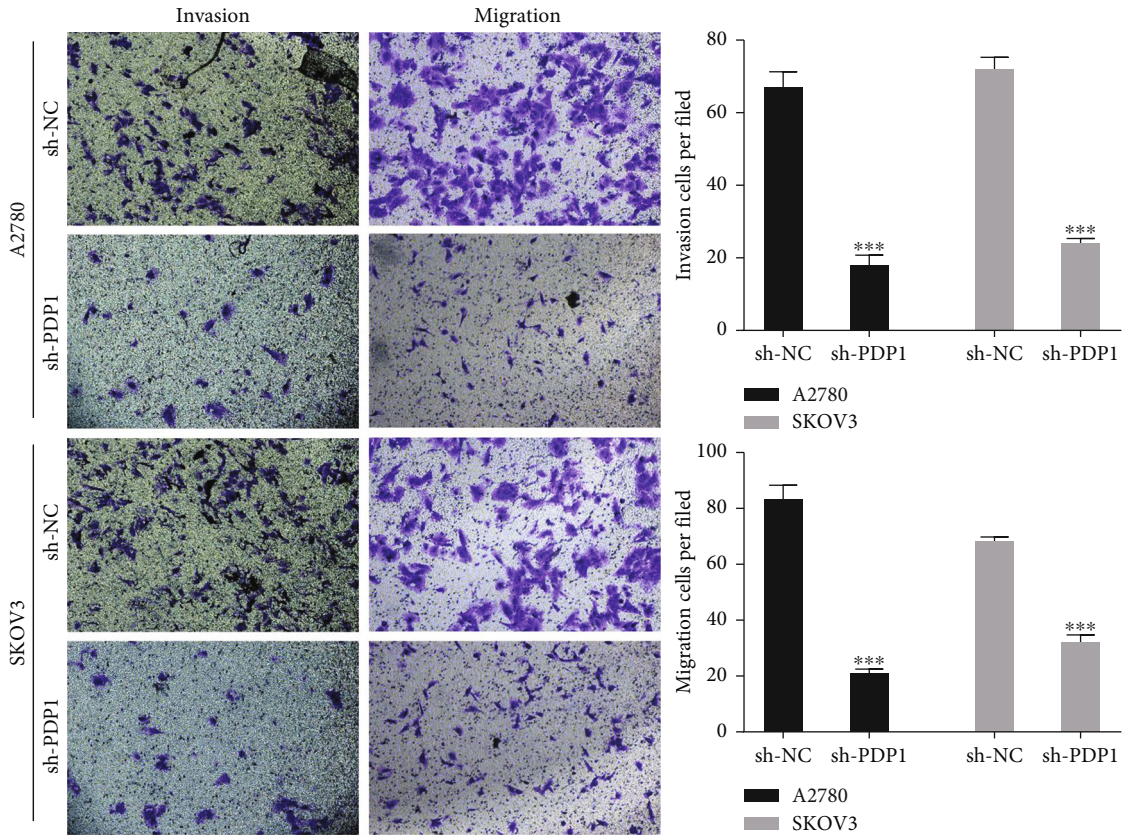


Figure 5: Continued.



(f)

Figure 5: PDP1 promotes cell proliferation, invasion, and migration of ovarian cancer. Notes: (a) the expression level of PDP1 in ovarian cancer cell lines; (b, c) the knockdown efficiency of PDP1 in A2780 and SKOV3 cell lines; (d, e) CCk8 assay was performed to evaluate the cell proliferation ability; (f) Transwell assay was performed to evaluate the invasion and migration ability of cancer cells.

leading to the decrease of mitochondrial and enhanced cell malignant behavior [17]. KRAS mutation is the most frequently mutated RAS, which could activate downstream signaling and promote tumorigenesis [18]. Rahman and colleagues revealed that the gene amplification of KRAS and MAPK1 was essential for type II ovarian carcinomas' growth [19]. Wang and colleagues indicated that the MGP protein could directly bind to the p-STAT5 in the nucleus and activate JAK2/STAT5 signaling in gastric cancer, further facilitating tumor progression [20]. Our result showed that the cancer-promoting effect of PDP1 in OC might be dependent on these oncogenetic pathways.

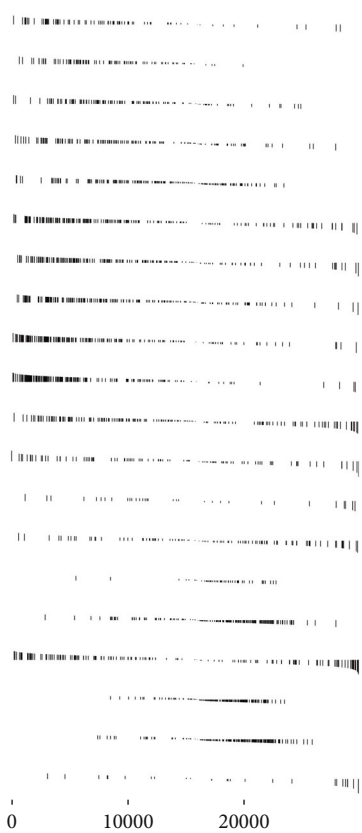
Recently, increased attention has been paid to the effect of the tumor microenvironment on cancer cells [21]. The immune infiltration result showed that PDP1 was positively correlated with activated myeloid dendritic cells, resting CD4 memory T cells, neutrophil, and M1 and M2 macrophages, yet negatively correlated with M0 macrophages, plasma B cells, $\gamma\delta$T cells, and activated CD4 memory T cells. Neutrophils are increasingly acknowledged to contribute to tumor development [22]. Xiao and colleagues found that the CTSC could promote lung metastasis of breast cancer through mediating recruitment of neutrophils and formation of neutrophil extracellular traps based on the CTSC/PR3/IL-1β axis [23]. In OC, Lee and colleagues revealed that neutrophils could facilitate OC premetastatic niche formation in the omentum, therefore resulting in an environment favorable for cancer cells to implant [24]. Meanwhile, Zeng and colleagues found that M2-like tumor-associated macrophages-secreted EGF could facilitate epithelial OC metastasis by activating EGFR-ERK signaling and suppressing lncRNA LIMT expression [25]. Therefore, the underlying interaction between PDP1 and these immune cells might be partly responsible for the cancer-promoting effect of PDP1 in OC. Drug sensitivity analysis showed the OC patients with high PDP1 expression might be more sensitive to bleomycin and gemcitabine, which increase its potential for clinical application.

Although our study was based on reliable analysis, some limitations should be noticed. Firstly, the populations enrolled in our analysis were mainly the Western world individuals, and the race bias was inevitable. Secondly, the clinical information of OC patients in TCGA was incomplete, only including the grade and clinical stage. If the clinical information was more complete, the conclusion of our study would be more credible.

4. Conclusions

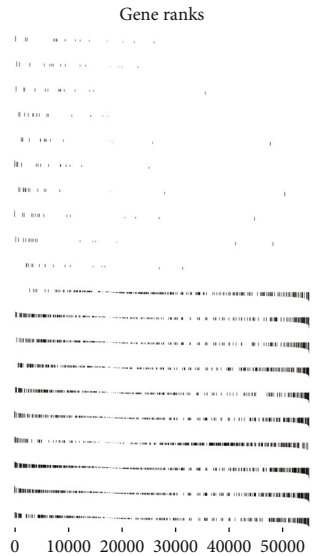
Based on the high-quality bioinformatics analysis and in vitro experiments, we found that PDP1 was overexpressed in OC tissue and cell lines. Moreover, PDP1 could significantly promote

Pathway	NES	p val	p adj
HALLMARK_IL6_JAK_STAT3_SIGNALING	2.05	3.2e–04	1.6e–03
HALLMARK_INTERFERON_ALPHA_RESPONSE	2.27	3.4e–04	1.6e–03
HALLMARK_UV_RESPONSE_DN	1.88	3.8e–04	1.6e–03
HALLMARK_APOPTOSIS	1.94	4.0e–04	1.6e–03
HALLMARK_ADIPOGENESIS	1.56	4.4e–04	1.6e–03
HALLMARK_KRAS_SIGNALING_UP	1.69	4.5e–04	1.6e–03
HALLMARK_COMPLEMENT	1.54	4.6e–04	1.6e–03
HALLMARK_IL2_STAT5_SIGNALING	1.88	4.6e–04	1.6e–03
HALLMARK ALLOGRAFT REJECTION	2.52	4.6e–04	1.6e–03
HALLMARK_INFLAMMATORY_RESPONSE	2.77	4.6e–04	1.6e–03
HALLMARK_MYOGENESIS	–0.88	7.5e–01	9.9e–01
HALLMARK_BILE_ACID_METABOLISM	–0.89	6.9e–01	9.6e–01
HALLMARK_HEDGEHOG_SIGNALING	–1.34	9.7e–02	1.7e–01
HALLMARK_SPERMATOGENESIS	–1.29	8.0e–02	1.4e–01
HALLMARK_MYC_TARGETS_V2	–1.36	7.3e–02	1.4e–01
HALLMARK_G2M_CHECKPOINT	–1.38	1.8e–02	4.4e–02
HALLMARK_KRAS_SIGNALING_DN	–1.54	4.0e–03	1.1e–02
HALLMARK_MYC_TARGETS_V1	–1.58	1.0e–03	3.2e–03
HALLMARK_E2F_TARGETS	–1.78	2.6e–04	1.6e–03
HALLMARK_PANCREAS_BETA_CELLS	–2.13	1.6e–04	1.6e–03



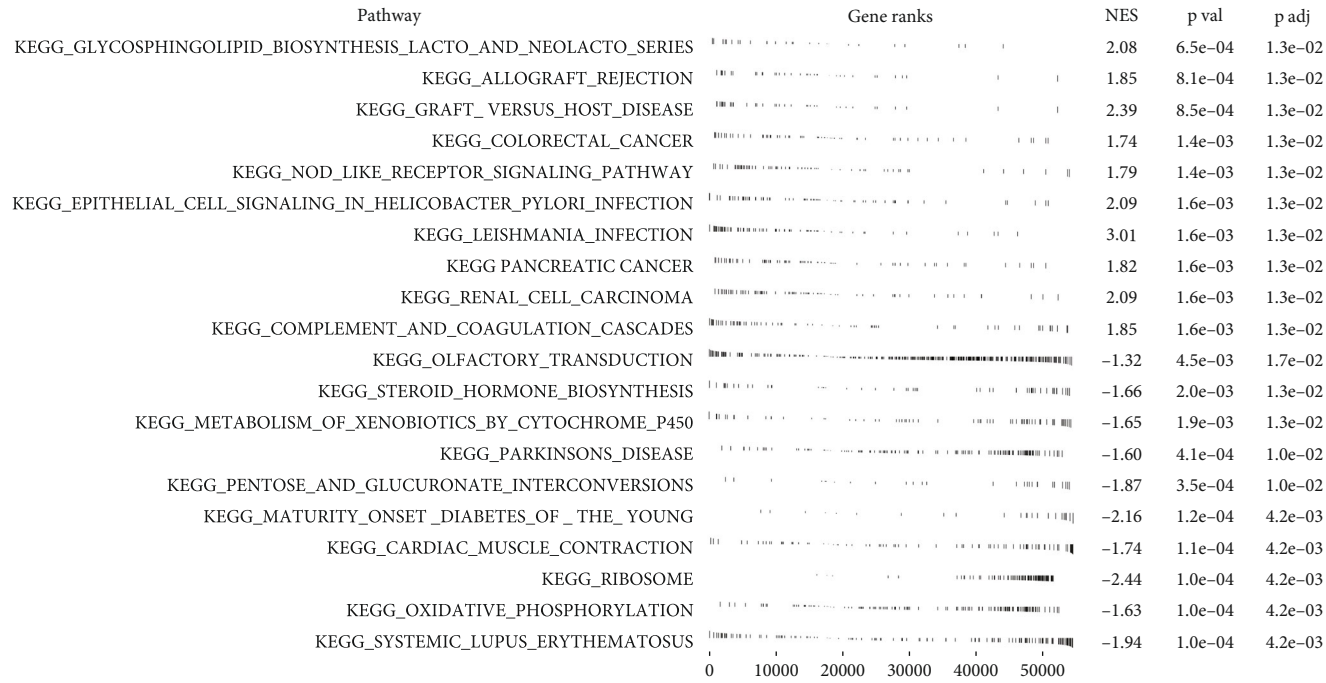
(a)

Pathway	NES	p val	p adj
GOBP_CELLULAR_RESPONSE_TO_PROSTAGLANDIN_E_STIMULUS	2.24	4.6e–04	1.2e–02
GOBP_NEGATIVE_REGULATION_OF_ACTIVATED_T_CELL_PROLIFERATION	2.40	4.6e–04	1.2e–02
GOBP_POSITIVE_REGULATION_OF_MACROPHAGE_DIFFERENTIATION	2.32	4.6e–04	1.2e–02
GOBP_REGULATION_OF_EXTRACELLULAR_MATRIX_ASSEMBLY	2.66	4.6e–04	1.2e–02
GOBP_REGULATION_OF_MEMBRANE_INVAGINATION	2.30	4.6e–04	1.2e–02
GOBP_TRANSFORMING_GROWTH_FACTOR_BETA1_PRODUCTION	2.54	4.6e–04	1.2e–02
GOMF_LOW_DENSITY_LIPOPROTEIN_PARTICLE_RECEPTOR_ACTIVITY	2.40	4.6e–04	1.2e–02
GOBP_GLOMERULAR_MESANGIUM_DEVELOPMENT	2.44	4.7e–04	1.2e–02
GOBP_HEAT_GENERATION	2.49	4.7e–04	1.2e–02
GOBP_NEUTROPHIL_HOMEOSTASIS	2.28	4.7e–04	1.2e–02
GOBP_RNA_SPLICING_VIA_TRANSESTERIFICATION_REACTIONS	–1.53	1.0e–04	8.5e–03
GOBP_HORMONE_TRANSPORT	–1.56	1.0e–04	8.5e–03
GOBP_AMIDE_TRANSPORT	–1.52	1.0e–04	8.5e–03
GOMF_PROTEIN_HETERODIMERIZATION_ACTIVITY	–1.53	1.0e–04	8.5e–03
GOBP_CAMERA_TYPE_EYE_DEVELOPMENT	–1.46	1.0e–04	8.5e–03
GOBP_REGIONALIZATION	–1.74	1.0e–04	8.5e–03
GOMF_OLFACTORY_RECEPTOR_ACTIVITY	–1.44	1.0e–04	8.5e–03
GOBP_REGULATION_OF_HORMONE_LEVELS	–1.52	1.0e–04	8.5e–03
GOBP_PATTERN_SPECIFICATION_PROCESS	–1.51	1.0e–04	8.5e–03
GOBP_FOREBRAIN_DEVELOPMENT	–1.62	1.0e–04	8.5e–03



(b)

Figure 6: Continued.



(c)

Figure 6: Biological enrichment of PDP1. Notes: (a) GSEA analysis of PDP1 based on Hallmark gene set; (b) GO analysis of PDP1; (c) KEGG analysis of PDP1.

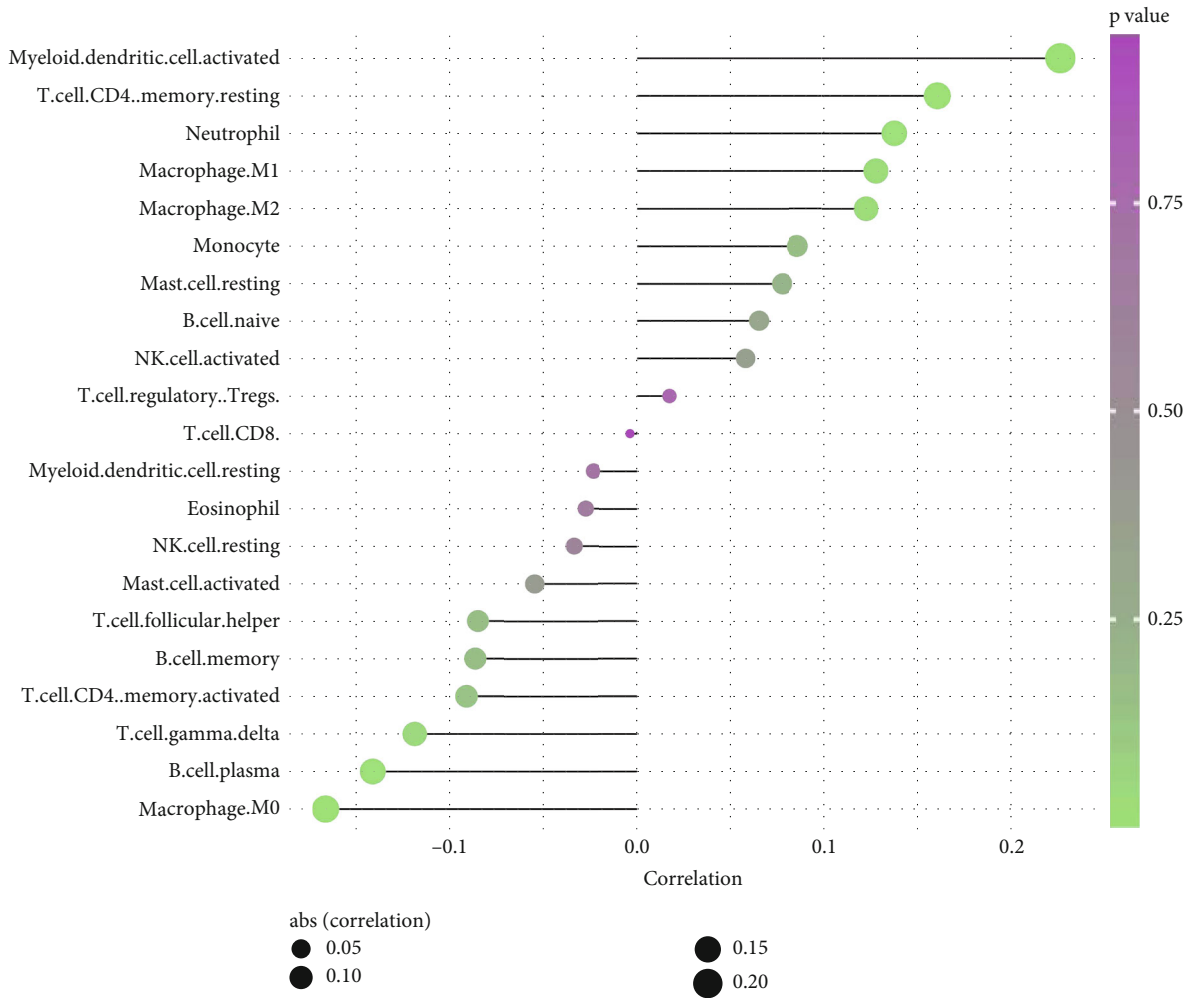
cancer cell proliferation, invasion, and migration. GSEA analysis showed that PDP1 could activate several oncogenic pathways, like IL6/JAK/STAT3 signaling and KRAS signaling. Immune infiltration analysis showed that PDP1 was positively correlated with activated myeloid dendritic cells, resting CD4 memory T cells, neutrophil, and M1 and M2 macrophages, yet negatively correlated with M0 macrophages, plasma B cells, $\gamma\delta$T cells, and activated CD4 memory T cells. Also, PDP1 was associated with the chemosensitivity of bleomycin, gemcitabine, and cisplatin.

5. Methods

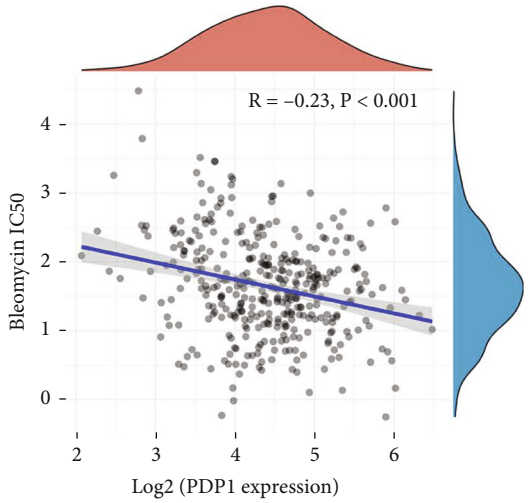
5.1. Public Data Acquisition and Processing. All the data used for analysis were downloaded from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. For the TCGA database, the transcriptomic profiling (https://portal.gdc.cancer.gov/, TCGA-OV, date category: transcriptome profiling, data type: gene expression quantification, workflow type: FPKM) and clinical information (TCGA-OV, date category: clinical, data format: bcr xml) of OC patients were obtained from the TCGA-GDC server. The FPKM file was then converted into TPM form for further analysis. Data was collated and preprocessed using the R software. For the GEO database, the datasets GSE18520 (platform: GPL570), GSE27651 (platform: GPL570), and GSE54388 (platform: GPL570) were downloaded for analysis. The patients with complete gene expression profile and clinical information were included in our analysis. Limma package was used for differential expressed gene (DEG) analysis with the threshold of $|\log FC > 1|$ and P value < 0.05 [26]. The immunohistochemistry (IHC) pictures of PDP1 in OC tumor and normal tissue was obtained from The Human Protein Atlas (HPA) database. Survival analysis of PDP1 was performed using the survival package in R software. Gene set enrichment analysis was performed to explore the biological pathway difference between low and high PDP1 OC patients and the reference file was Hallmark, c2.cp.kegg.v7.5.1.symbols, and c5.go.v7.5.1.symbols pathway set. Immune infiltration analysis was performed using the CIBERSORT algorithm [27]. Drug sensitivity analysis was conducted out based on the Genomics of Drug Sensitivity in Cancer (GDSC) database.

5.2. Cell Culture and Transfection. The human normal ovarian epithelial cell line (IOSE80) and human epithelial ovarian cancer cell lines (CaOV-3, HO-8910, SKOV3, and A2780) were laboratory stocks. Cells were cultured in the conventional incubator atmosphere containing 5% CO2 at 37°C and routinely passaged three times weekly. Cell transfection was performed using the Lipofectamine 2000 according to standard protocol. PDP1 knockdown and control plasmids were purchased from Shanghai Jikai Gene Chemical Co., Ltd., whose target sequence were as follows: shRNA1: 5′-GCCTGTTTAATGGATATGTTT-3′; shRNA2: 5′-CTGTTAAAGTTTGTCAATTAA-3′; shRNA3: 5′-GACGATAAAGTGTTTTTAGTA-3′.

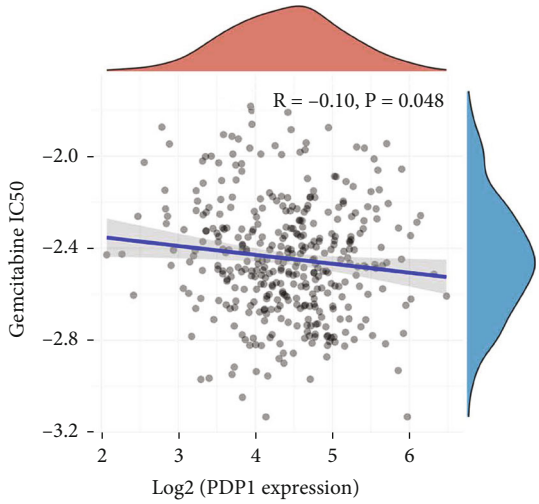
5.3. Quantitative Real-Time PCR (qRT-PCR). A TRIzol RNA extraction kit (TaKaRa) was used to extract the total RNA following the protocol. Total RNA was then reverse-transcribed to cDNA. qRT-PCR was performed based on



(a)



(b)



(c)

Figure 7: Continued.

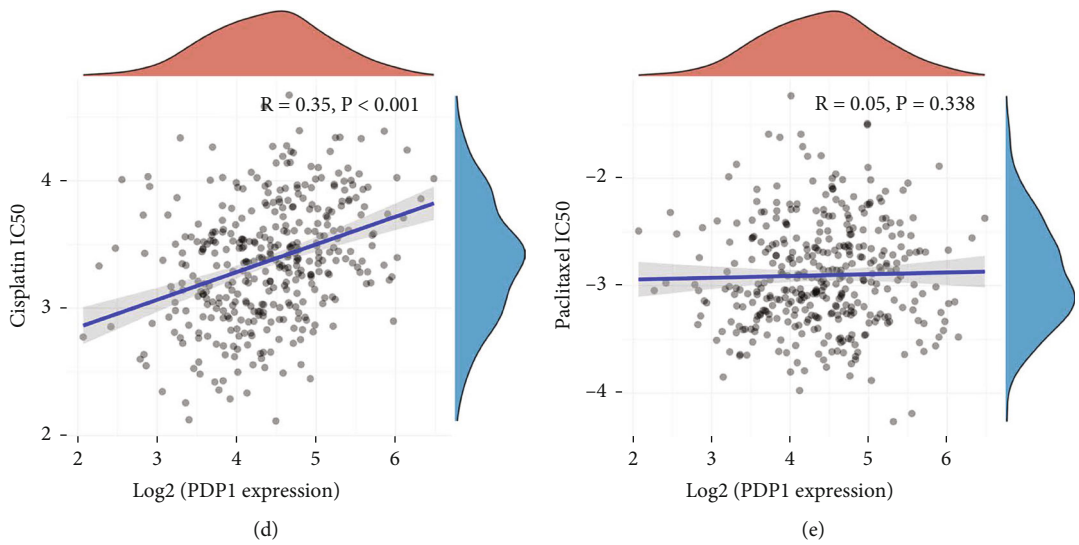


Figure 7: Immune infiltration and drug sensitivity analysis. Notes: (a) CIBERSORT algorithm was performed to explore association between PDP1 and immune cell infiltration; (b–e) the association between PDP1 and drug chemosensitivity.

the SyBr Green system. The primer used was as follows: PDP1, forward, 5′-GTCCTTCCCATTCTGCAACC-3′; reverse, 5′-GAAACAGAGGAGGACCAAACA-3′, GAPDH, forward, 5′-GCAAATTCCATGGCACCGT-3′; reverse, 5′-TCGCCCCACTTGATTTTGG-3′.

5.4. CCK8 Assay. CCK8 assay was performed using a CCK8 Kit (Dojindo, Shanghai, China) according to the protocol. Cells were plated in a 96-well plate and added with a CCK8 solution. After that, cells were incubated for 1 h at 37°C, and then, the absorbance was detected at A450 nm for 0 h, 24 h, 48 h, and 72 h.

5.5. Transwell Assay. Transwell assay was performed using the Transwell chamber (8 μm pore size, Corning). The Transwell chamber was added into the 24-well plates and divided into the compartment into the upper and lower chambers. The upper chamber was added with cells with 500 μl in serum-free medium, and the lower chamber was added with 600 μl complete media. After that for 24 h, cells were fixed with paraformaldehyde and stained with crystal violet.

5.6. Statistical Analysis. Statistical analysis was performed using the R software and GraphPad Prism 8. P value less than 0.05 was regarded as statistically significant. Student's T test was used for variables with normal distribution, and Wilcoxon test was used for variables with nonnormal distribution. All the experiments were repeated at least three times.

Abbreviations

OC: Ovarian cancer
TCGA: The Cancer Genome Atlas
GEO: Gene Expression Omnibus
GSEA: Gene Set Enrichment Analysis
PDP1: Pyruvate dehydrogenase phosphatase catalytic subunit 1
PDKs: Pyruvate dehydrogenase kinases
PDPs: Pyruvate dehydrogenase phosphatases.

Data Availability

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

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

Y.S performed the bioinformatic analysis. J.Z and L.Z performed the experiments. C.CS designed the work. Yan Song and Juan Zhang contributed equally to this work.

References

[1] U. Menon, C. Karpinskyj, and A. Gentry-Maharaj, "Ovarian cancer prevention and screening," *Obstetrics and Gynecology*, vol. 131, no. 5, pp. 909–927, 2018.

[2] S. Narod, "Can advanced-stage ovarian cancer be cured?," *Nature Reviews Clinical Oncology*, vol. 13, no. 4, pp. 255–261, 2016.

[3] L. A. Torre, B. Trabert, C. E. DeSantis et al., "Ovarian cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 4, pp. 284–296, 2018.

[4] J. Fan, C. Shan, H. B. Kang et al., "Tyr phosphorylation of PDP1 toggles recruitment between ACAT1 and SIRT3 to regulate the pyruvate dehydrogenase complex," *Molecular Cell*, vol. 53, no. 4, pp. 534–548, 2014.

[5] S. Miriyala, T. Subramanian, M. Panchatcharam et al., "Functional characterization of the atypical integral membrane lipid

phosphatase PDP1/PPAPDC2 identifies a pathway for interconversion of isoprenols and isoprenoid phosphates in mammalian cells," *The Journal of Biological Chemistry*, vol. 285, no. 18, pp. 13918–13929, 2010.

[6] Y. Shi, Y. Wang, H. Jiang et al., "Mitochondrial dysfunction induces radioresistance in colorectal cancer by activating $[Ca^{2+}]_m$-PDP1-PDH-histone acetylation retrograde signaling," *Cell Death & Disease*, vol. 12, no. 9, 2021.

[7] X. Feng, J. Lu, Y. Wu, and H. Xu, "MiR-18a-3p improves cartilage matrix remodeling and inhibits inflammation in osteoarthritis by suppressing PDP1," *The Journal of Physiological Sciences : JPS*, vol. 72, no. 1, 2022.

[8] Y. Li, J. Shen, C. S. Cheng, H. Gao, J. Zhao, and L. Chen, "Overexpression of pyruvate dehydrogenase phosphatase 1 promotes the progression of pancreatic adenocarcinoma by regulating energy-related AMPK/mTOR signaling," *Cell & Bioscience*, vol. 10, no. 1, p. 95, 2020.

[9] C. Shan, H. B. Kang, S. Elf et al., "Tyr-94 phosphorylation inhibits pyruvate dehydrogenase phosphatase 1 and promotes tumor growth," *The Journal of Biological Chemistry*, vol. 289, no. 31, pp. 21413–21422, 2014.

[10] N. H. Jeoung, "Pyruvate dehydrogenase kinases: therapeutic targets for diabetes and cancers," *Diabetes & Metabolism Journal*, vol. 39, no. 3, pp. 188–197, 2015.

[11] X. Chen, Y. Wang, Y. Li, G. Liu, K. Liao, and F. Song, "Identification of immune-related cells and genes in the breast invasive carcinoma microenvironment," *Aging*, vol. 14, no. 3, pp. 1374–1388, 2022.

[12] J. Chen, I. Guccini, D. Di Mitri et al., "Compartmentalized activities of the pyruvate dehydrogenase complex sustain lipogenesis in prostate cancer," *Nature Genetics*, vol. 50, no. 2, pp. 219–228, 2018.

[13] P. M. Webb and S. J. Jordan, "Epidemiology of epithelial ovarian cancer," *Best Practice & Research Clinical Obstetrics & Gynaecology*, vol. 41, pp. 3–14, 2017.

[14] P. W. Stacpoole, "Therapeutic targeting of the pyruvate dehydrogenase complex/pyruvate dehydrogenase kinase (PDC/PDK) axis in cancer," *Journal of the National Cancer Institute*, vol. 109, no. 11, 2017.

[15] J. S. Ni, H. Zheng, Y. L. Ou et al., "miR-515-5p suppresses HCC migration and invasion via targeting IL6/JAK/STAT3 pathway," *Surgical Oncology*, vol. 34, pp. 113–120, 2020.

[16] G. Zhao, G. Zhu, Y. Huang et al., "IL-6 mediates the signal pathway of JAK-STAT3-VEGF-C promoting growth, invasion and lymphangiogenesis in gastric cancer," *Oncology Reports*, vol. 35, no. 3, pp. 1787–1795, 2016.

[17] B. Salimian Rizi, C. Caneba, A. Nowicka et al., "Nitric oxide mediates metabolic coupling of omentum-derived adipose stroma to ovarian and endometrial cancer cells," *Cancer Research*, vol. 75, no. 2, pp. 456–471, 2015.

[18] D. Uprety and A. A. Adjei, "KRAS: from undruggable to a druggable cancer target," *Cancer Treatment Reviews*, vol. 89, article 102070, 2020.

[19] M. T. Rahman, K. Nakayama, M. Rahman et al., "KRAS and MAPK1 gene amplification in type II ovarian carcinomas," *International Journal of Molecular Sciences*, vol. 14, no. 7, pp. 13748–13762, 2013.

[20] M. Wang, L. Chen, Y. Chen et al., "Intracellular matrix Gla protein promotes tumor progression by activating JAK2/STAT5 signaling in gastric cancer," *Molecular Oncology*, vol. 14, no. 5, pp. 1045–1058, 2020.

[21] T. Wu and Y. Dai, "Tumor microenvironment and therapeutic response," *Cancer Letters*, vol. 28, no. 387, pp. 61–68, 2017.

[22] F. Mollinedo, "Neutrophil degranulation, plasticity, and cancer metastasis," *Trends in Immunology*, vol. 40, no. 3, pp. 228–242, 2019.

[23] Y. Xiao, M. Cong, J. Li et al., "Cathepsin C promotes breast cancer lung metastasis by modulating neutrophil infiltration and neutrophil extracellular trap formation," *Cancer Cell*, vol. 39, no. 3, 2021.

[24] W. Lee, S. Y. Ko, M. S. Mohamed, H. A. Kenny, E. Lengyel, and H. Naora, "Neutrophils facilitate ovarian cancer premetastatic niche formation in the omentum," *The Journal of Experimental Medicine*, vol. 216, no. 1, pp. 176–194, 2019.

[25] X. Y. Zeng, H. Xie, J. Yuan et al., "M2-like tumor-associated macrophages-secreted EGF promotes epithelial ovarian cancer metastasis via activating EGFR-ERK signaling and suppressing lncRNA LIMT expression," *Cancer Biology & Therapy*, vol. 20, no. 7, pp. 956–966, 2019.

[26] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, 2015.

[27] B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, and A. A. Alizadeh, "Profiling tumor infiltrating immune cells with CIBERSORT," *Methods in Molecular Biology*, vol. 1711, pp. 243–259, 2018.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 1968829, 20 pages
https://doi.org/10.1155/2022/1968829

Research Article

Identification of the Characteristic Genes and their Roles in Lung Adenocarcinoma Lymph Node Metastasis through Machine Learning Algorithm

Qian Zhou, Xianghui Wang, Haiyun Qian, Shengwei Ma, Chenggang Lei, and Fenghe Cui

Department of Cardiothoracic Surgery, Jingzhou Central Hospital, Jingzhou Hospital Affiliated to Yangtze University, Jingzhou, Hubei, China

Correspondence should be addressed to Fenghe Cui; jzszxyyxxwk@163.com

Received 23 July 2022; Revised 1 September 2022; Accepted 9 September 2022; Published 12 October 2022

Guest Editor: Xiaoye Xu

Background. Lymph node metastasis is an important route of lung cancer metastasis and can significantly affect the survival of lung cancer. *Methods*. All the analysis was conducted out in the R software. Expression profile and clinical information of lung adenocarcinoma (LUAD) patients were downloaded from The Cancer Genome Atlas database. *Results*. In our study, we firstly identified the characteristic genes of lymph node metastasis in LUAD through two machine learning algorithms, least absolute shrinkage and selection operator (LASSO) logistic regression, and SVM-RFE algorithms. Ten characteristic genes were finally identified, including *CRHR2*, *ITIH1*, *PRSS48*, *MAS1L*, *CYP4Z1*, *LMO1*, *TCP10L2*, *KRT78*, *IGFBP1*, and *PITX3*. Next, we performed univariate Cox regression, LASSO regression, and multivariate Cox regression sequentially to construct a prognosis model based on *MAS1L*, *TCP10L2*, and *CRHR2*, which had a good prognosis prediction efficiency in both training and validation cohorts. Univariate and multivariate analysis indicated that our model is a risk factor independent of other clinical features. Pathway enrichment analysis showed that in the high-risk patients, the pathway of MYC target, unfolded protein response, interferon alpha response, DNA repair, reactive oxygen species pathway, and glycolysis were significantly enriched. Among three model genes, *MAS1L* aroused our interest and therefore was selected for further analysis. KM survival curves showed that the patients with higher *MAS1L* might have better disease-free survival and progression-free survival. Further, pathway enrichment, genomic instability, immune infiltration, and drug sensitivity analysis were performed to in-deep explore the role of *MAS1L* in LUAD. *Conclusions*. Results showed that the signature based on *MAS1L*, *TCP10L2*, and *CRHR2* is a useful tool to predict prognosis and lung cancer lymph node metastasis.

1. Introduction

Lung cancer is a leading cause of cancer-related deaths all over the world, in which non-small-cell lung cancer (NSCLC) is the most common pathological subtype and accounts for approximately 85% of cases [1]. Many factors may contribute to the occurrence and development of lung cancer, including smoking, genetic susceptibility, environmental exposure, and so on [2]. Surgery can significantly improve the prognosis of early-stage and resectable lung cancer patients. Unfortunately, inadequate screening plans and hidden clinical symptoms have resulted in most patients being diagnosed with advanced disease at the time of their first presentation [3]. However, for those patients in the advanced stage, the prognosis is still unsatisfactory [3]. Therefore, it is meaningful to identify novel molecules associated with patient's prognosis and with the potential to guide therapy options.

Lymph node metastasis is an important feature of lung cancer and is associated with poor prognosis [4]. In the absence of distant metastasis, an accurate assessment of lymph node involvement is a crucial step for NSCLC patients [4]. Throughout the last decades, radical lobectomy has been universally accepted as a standard procedure for lung cancer patients [5]. In recent years, intrathoracic reassessments during thoracotomies for lung cancer have

Table 1: Baseline information of enrolled patients.

Features		Numbers	Percentage (%)
Age	<= 65	241	46.2%
	> 65	262	50.2%
	Unknown	19	3.6%
Gender	Female	280	53.6%
	Male	242	46.4%
Stage	Stage I	279	53.4%
	Stage II	124	23.8%
	Stage III	85	16.3%
	Stage IV	26	5.0%
	Unknown	8	1.5%
T-stage	T1	172	33.0%
	T2	281	53.8%
	T3	47	9.0%
	T4	19	3.6%
	Unknown	3	0.6%
M-stage	M0	353	67.6%
	M1	25	4.8%
	Unknown	144	27.6%
N-stage	N0	335	64.2%
	N1	98	18.8%
	N2	75	14.4%
	N3	2	0.4%
	Unknown	12	2.3%

evolved into detailed and complex assessments, and the core of which is to evaluate the involvement of lymph nodes at the mediastinal and hilar levels [5]. This technique is called "systematic lymph node dissection", which has been accepted as an important part of lung cancer surgery. At the genome level, researchers have focused on the molecules involved in lung cancer lymph node metastasis that might be the underlying therapy target. For instance, Liu et al. found that in small cell lung cancer, patients with high *CCL19* expression had poorer outcomes and more LN metastasis, as well as impaired CD8+ T cell function [6]. Moreover, Bi et al. revealed that *CXCR4* and *VEGF-C* are correlated with lymph node metastasis and might synergistically promote lung cancer progression [7]. Thus, exploring the potential mechanism of lymph node metastasis of lung cancer from a molecular perspective is helpful for the diagnosis and treatment of lung cancer.

In our study, we firstly identified the characteristic genes of lymph node metastasis in LUAD through least absolute shrinkage and selection operator (LASSO) logistic regression and SVM-RFE algorithms. A prognosis model based on *MAS1L*, *TCP10L2*, and *CRHR2* was then established, which had a good prognosis prediction efficiency in both training and validation cohorts. Next, pathway enrichment analysis was performed to explore the underlying biological differences between high- and low-risk patients. Among three model genes, *MAS1L* aroused our interest and therefore was selected for further analysis. KM survival curves showed that the patients with higher *MAS1L* might have better disease-free survival and progression-free survival. Further, pathway enrichment, genomic instability, immune infiltration, and drug sensitivity analysis were performed to in-deep explore the role of *MAS1L* in LUAD.

2. Methods

2.1. Data Acquisition. The open-accessed transcriptional profiling data and clinical information of lung adenocarcinoma (LUAD) were downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/; 522 patients, age: 65.33 ± 10.02) database. The original expression profile files were "STAR-counts" form and further sorted using the author's R code (tpm_unstranded). Clinical information was collated using Perl code. Genomic reference file GRCh38.gtf was used for probe annotation. The baseline information of patients was shown in Table 1. For the enrolled patients, the patients with N0 stage was regarded as the population without lymph node metastasis, while the N1-3 stage patients were regarded as the population with lymph node metastasis.

2.2. Machine Learning Algorithm. LASSO logistic regression and support vector machine recursive feature elimination (SVM-RFE) algorithms were used for characteristic genes

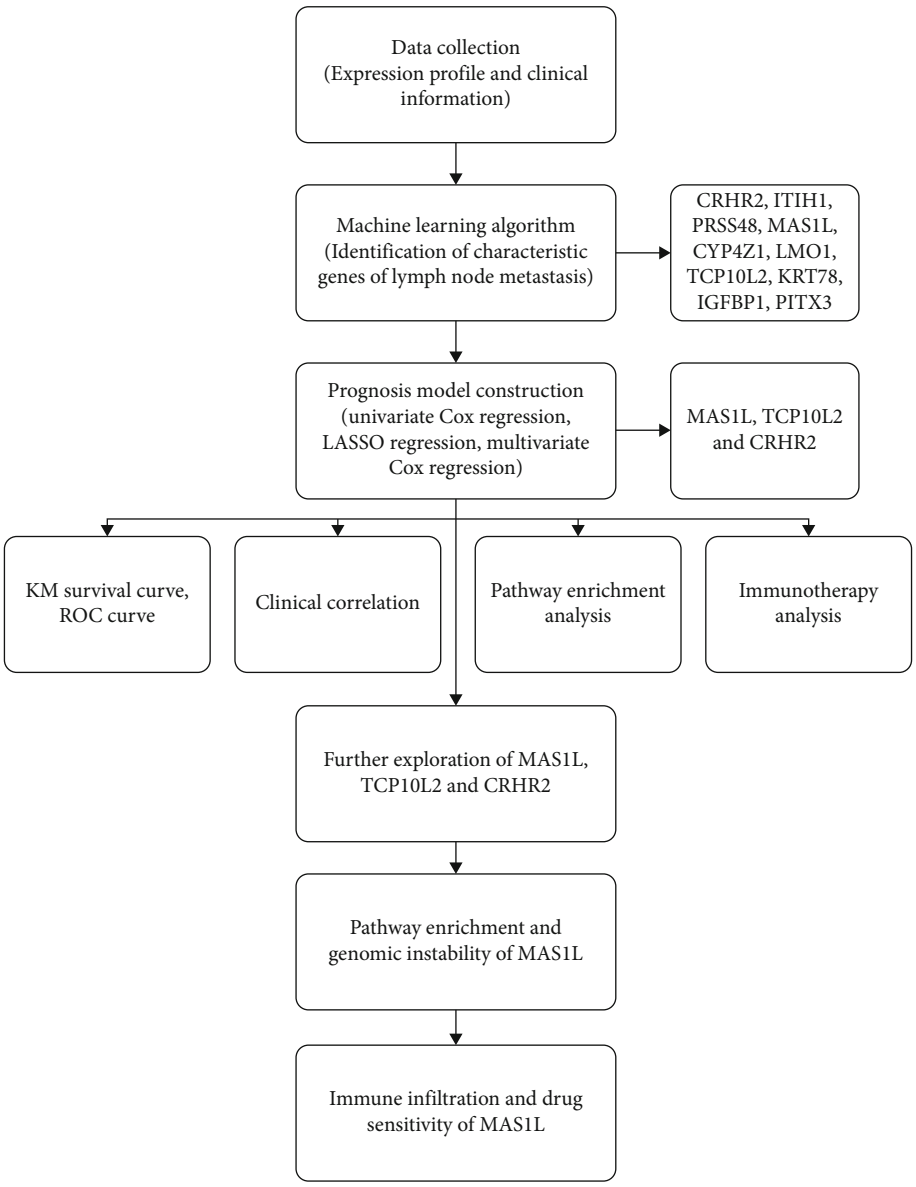


Figure 1: The flow chart of the whole study.

screening. LASSO logistic regression was performed based on the glmnet package [8]. SVM is a machine learning method that can find the best variables by deleting the feature vectors generated by SVM [9].

2.3. Establishment of Prognosis Model. Patients were randomly assigned to training and validation cohorts with a 1 : 1 ratio. Based on the characteristic genes identified by LASSO logistic regression and SVM-RFE algorithms, univariate Cox regression analysis was firstly performed to determine prognosis-related genes with the threshold of $P < 0.05$. Next, LASSO regression was used for dimensionality reduction [10]. Finally, multivariate Cox regression analysis was utilized for prognosis model construction with the formula of "Riskscore = Gene A*Coef A + Gene B * Coef B + … + Gene N * Coef N". Each patient would be assigned a riskscore according to the above formula [11].

2.4. Model Evaluation. According to the calculated riskscore, the patients with riskscore higher than the median were defined as high-risk group, otherwise, the low-risk group. Kaplan-Meier (KM) survival curve was used to evaluate the prognosis difference between high- and low-risk patients. The receiver operating characteristic (ROC) curve was used to evaluate the prediction efficacy of our model in a specific time node. The independence of our prognosis model was validated by univariate and multivariate Cox regression analysis [12].

2.5. Pathway Enrichment and Genomic Instability Analysis. Underlying biological differences can lead to different outcomes in patients. Pathway enrichment analysis was performed based on the GSEA algorithm. The reference gene set was set as Hallmark, c2.cp.kegg.v7.5.1.symbols, and c5.go.v7.5.1.symbols gene set. ClueGO plug-in in cytoscape

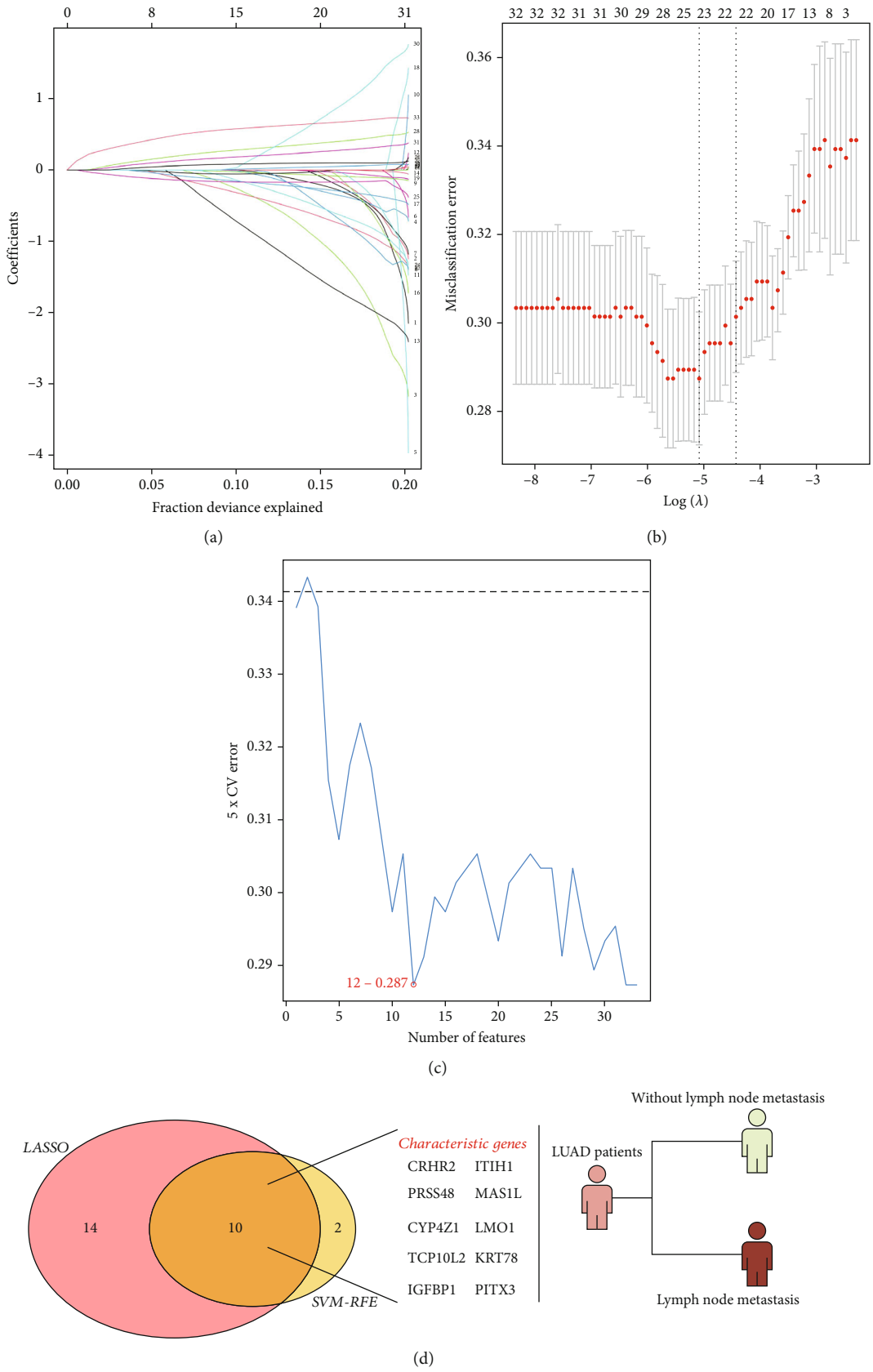


Figure 2: Identification of the characteristic genes of lung cancer lymph node metastasis. Notes: (a–b) LASSO logistic regression; (c) SVM-RFE algorithm; (d) ten characteristic genes were identified based on LASSO logistic regression and SVM-RFE algorithms, including *CRHR2*, *ITIH1*, *PRSS48*, *MAS1L*, *CYP4Z1*, *LMO1*, *TCP10L2*, *KRT78*, *IGFBP1*, and *PITX3*.

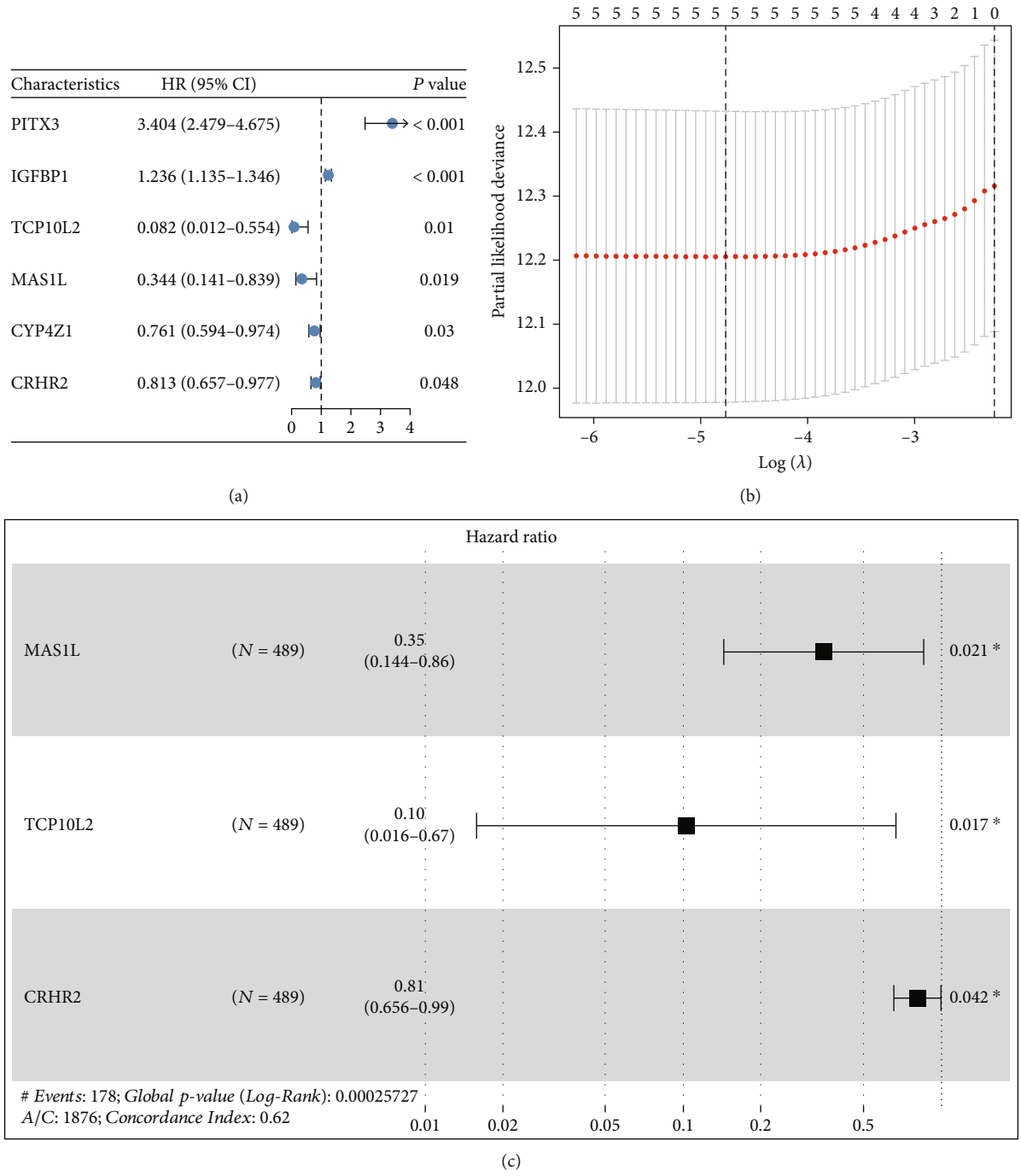


Figure 3: Continued.

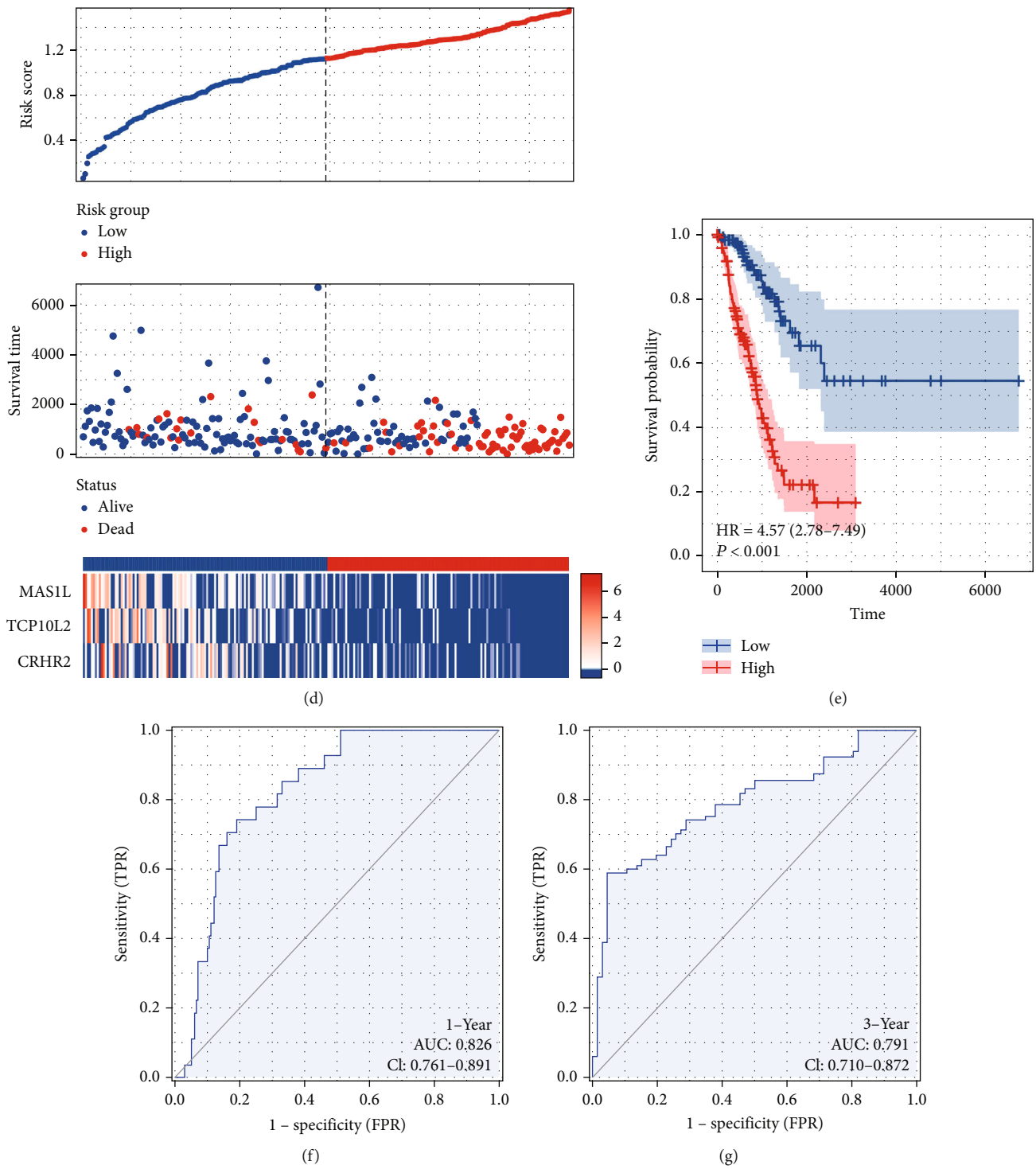


Figure 3: Continued.

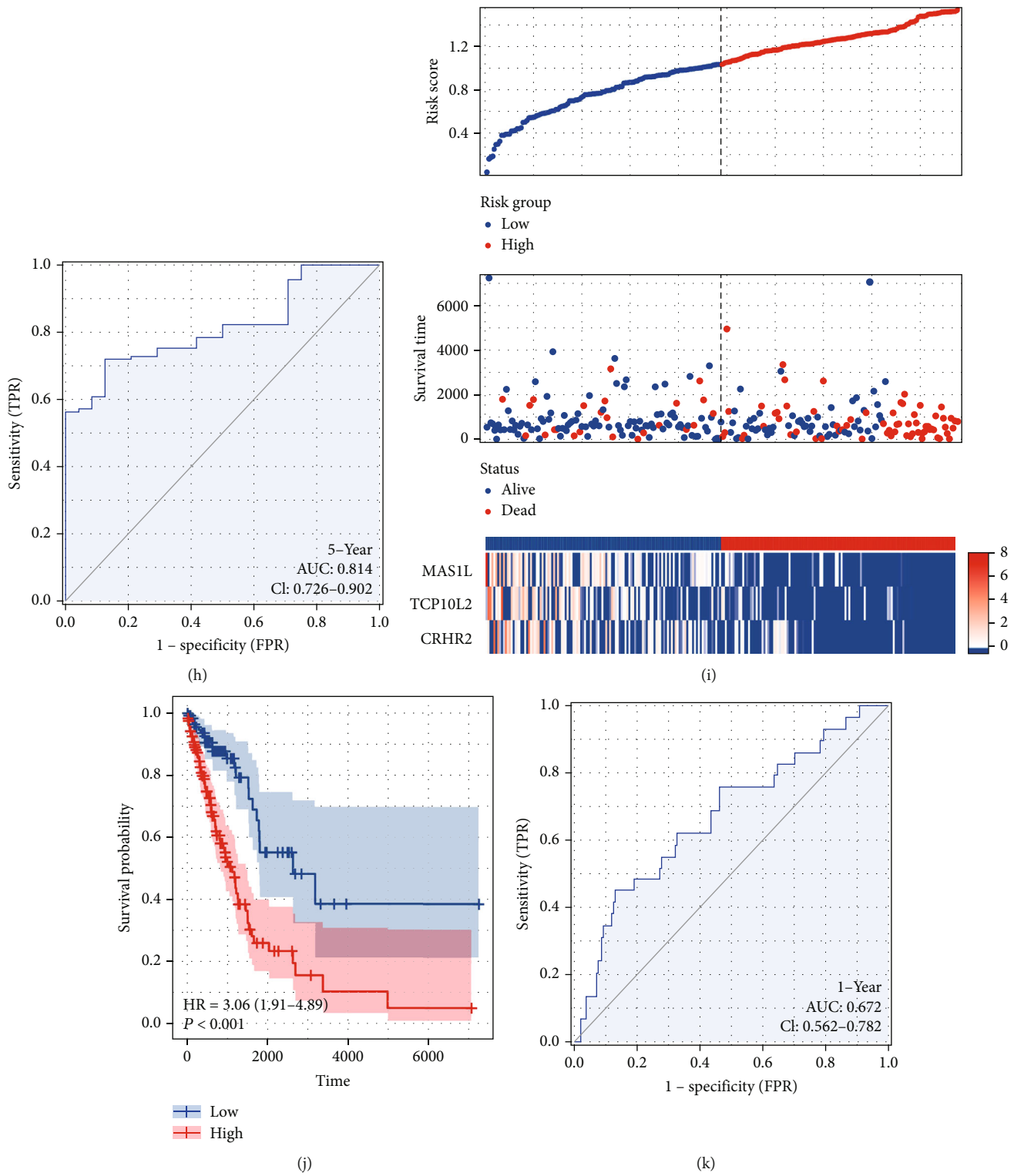


Figure 3: Continued.

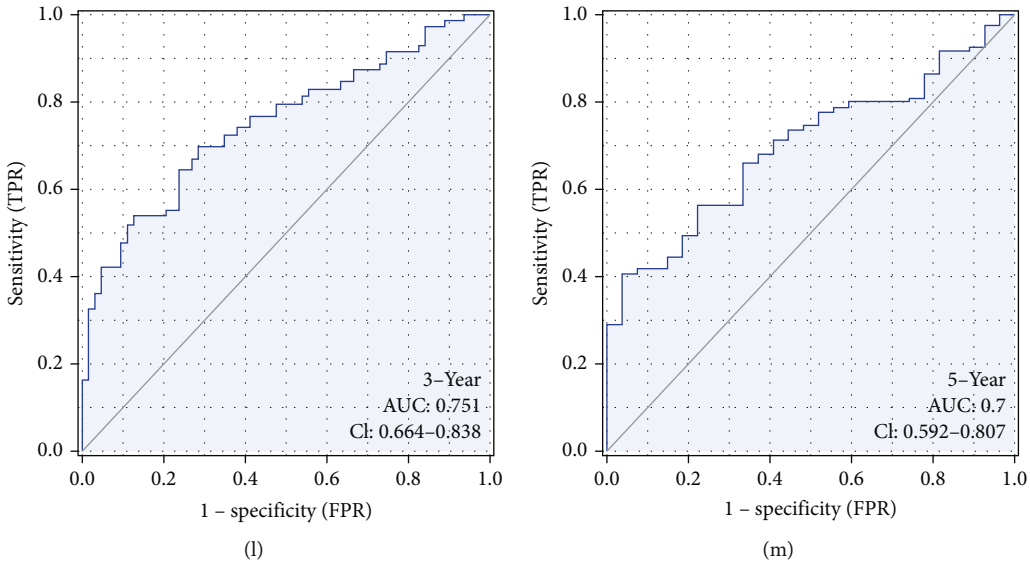


Figure 3: Prognosis model construction. Notes: (a) univariate Cox regression analysis of the identified characteristic genes to select prognosis-related genes; (b) LASSO regression was used for dimensionality reduction; (c) multivariate Cox regression analysis identified three genes *MAS1L*, *TCP10L2*, and *CRHR2* for model construction; (d) overview of the riskscore in the training model; (e) KM survival curve of high- and low-risk patients in training cohort; (f–h) ROC curves of 1-, 3-, and 5-years survival in the training cohort; (i) overview of the riskscore in the validation model; (j) KM survival curve of high- and low-risk patients in validation cohort; (k–m) ROC curves of 1-, 3-, and 5-years survival in the validation cohort.

software was used for the gene pathways enrichment [13]. The data of tumor mutational burden (TMB) and microsatellite instability (MSI) were downloaded from the TCGA database. The tumor stemness index (mRNAsi and EREG-mRNAsi) of TCGA-LUAD patients was obtained from the previous study [14].

2.6. Immune-Related and Drug Sensitivity Analysis. Immune infiltration analysis was conducted using the single sample gene set enrichment analysis (ssGSEA) algorithm [15]. Evaluation of the sensitity on immunotherapy was performed using the tumor immune dysfunction and exclusion (TIDE) analysis, in which the TIDE score < 0 was defined as the immunotherapy responders, and > 0 was regarded as the nonresponders [16]. Drug sensitivity analysis was performed based on the Genomics of Drug Sensitivity in Cancer (GDSC) database [17].

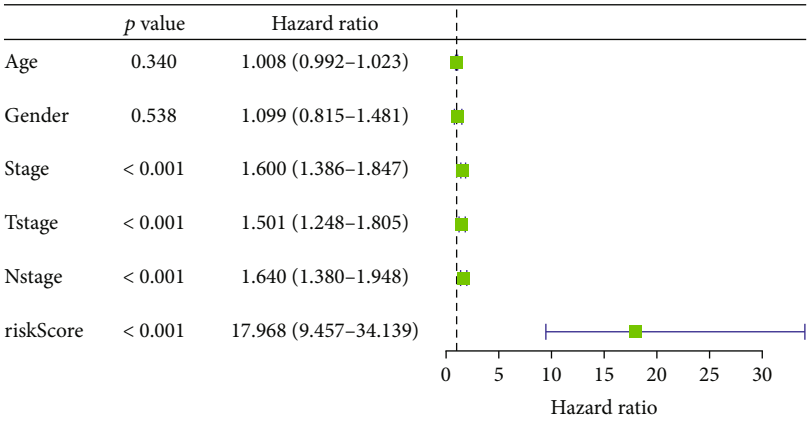
2.7. Statistical Analysis. All the analysis was performed using the R software version 4.0.0. Values of $P < 0.05$ were considered statistically significant. Normal distribution was tested by the Student t-test. Nonnormally distributed variables were compared using the Mann–Whitney U test.

3. Result

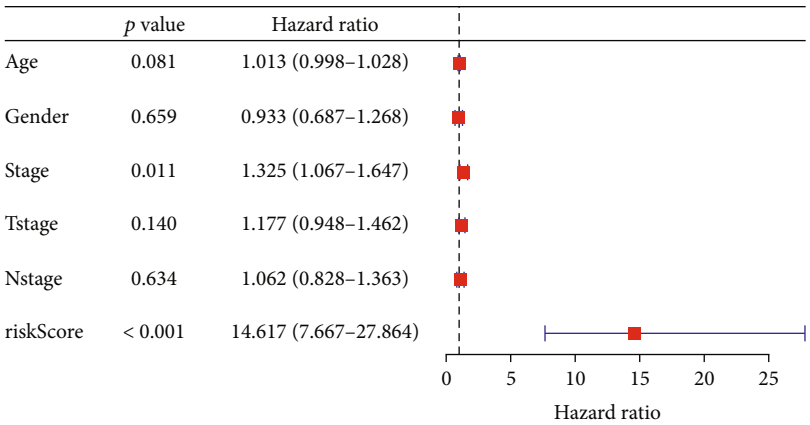
3.1. Identification of the Characteristic Genes of Lymph Node Metastasis. The flow chart of the whole study was shown in Figure 1. For the obtained data of TCGA-LUAD patients, we divided then into lymph node metastasis (N1-3) and non-lymph node metastasis (N0) group. LASSO logistic regression and SVM-RFE algorithms were used to identify the characteristic gene of lymph node metastasis (Figures 2(a–c)). Finally, LASSO logistic regression and SVM-RFE algorithms intersected ten characteristic genes, including *CRHR2*, *ITIH1*, *PRSS48*, *MAS1L*, *CYP4Z1*, *LMO1*, *TCP10L2*, *KRT78*, *IGFBP1*, and *PITX3* (Figure 2(d)).

3.2. Prognosis Model Construction. Next, univariate Cox regression analysis was performed to identify the prognosis-related characteristic genes. The result showed that among all these ten genes, *IGFBP1*, *TCP10L2*, *MAS1L*, *CYP4Z1*, and *CRHR2* were the protective factors, while *PITX3* was the risk factor (Figure 3(a)). LASSO regression was then used for data dimensionality reduction (Figure 3(b)). Multivariate Cox regression analysis identified three genes for prognosis model construction, including *MAS1L*, *TCP10L2*, and *CRHR2* (Figure 3(c)). In the training cohort, a higher proportion of dead cases were observed in the high-risk group (Figure 3(d)). KM survival curve showed that the high-risk patient might have a poor prognosis compared to the patients in low-risk group (Figure 3(e)). ROC curves showed a great prediction efficiency of patients 1-, 3-, and 5-year survival (Figures 3(f–h), 1-year AUC = 0.826, 3-year AUC = 0.791, and 5-year AUC = 0.814). Meanwhile, the same trend was also found in the validation group (Figure 3(i)). KM survival curve showed that in the validation group, the high-risk patient might have a wose prognosis (Figure 3(j)). Also, ROC curves showed a good prediction efficiency of patients' 1-, 3-, and 5-year survival in validation group (Figures 3(k–m), 1-year AUC = 0.672, 3-year AUC = 0.751, and 5-year AUC = 0.7).

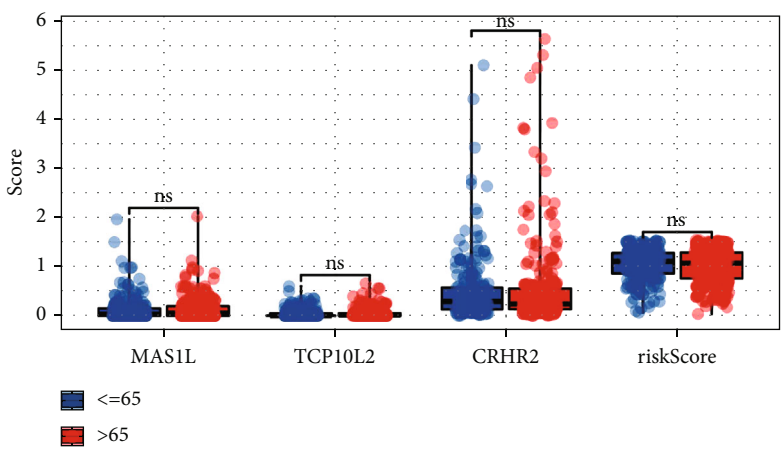
3.3. Clinical Correlation. Univariate and multivariate analysis showed that our model is a risk factor independent of other clinical features (Figures 4(a) and 4(b)). Then, we



(a)

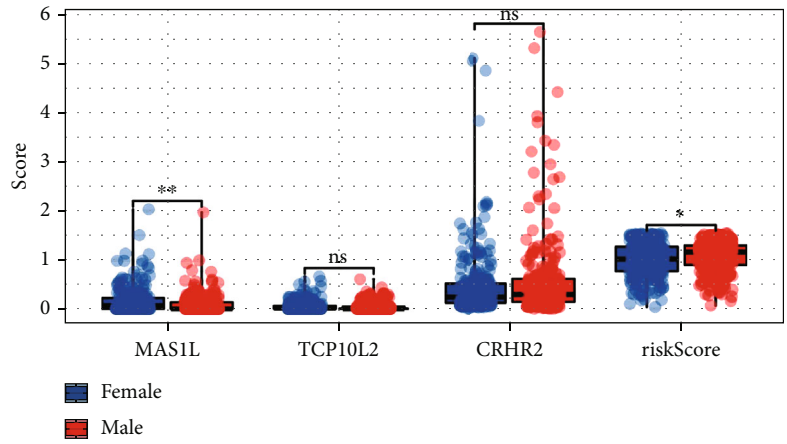


(b)

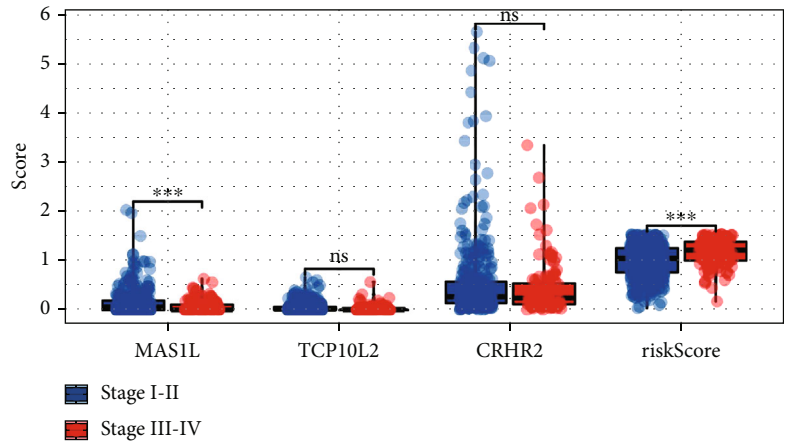


(c)

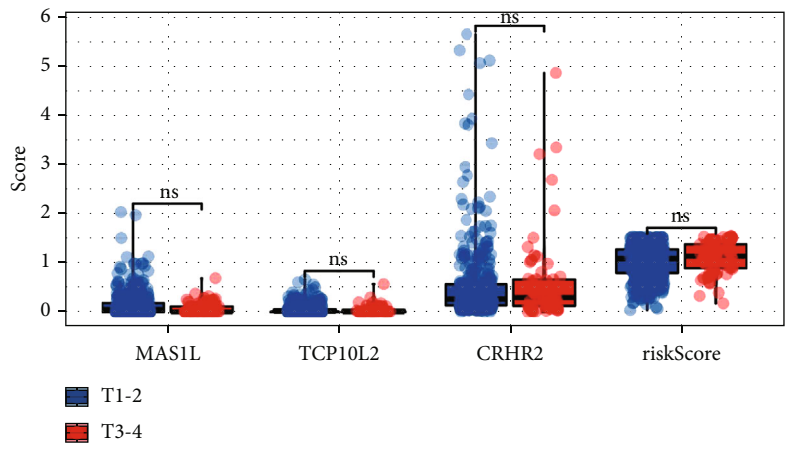
Figure 4: Continued.



(d)

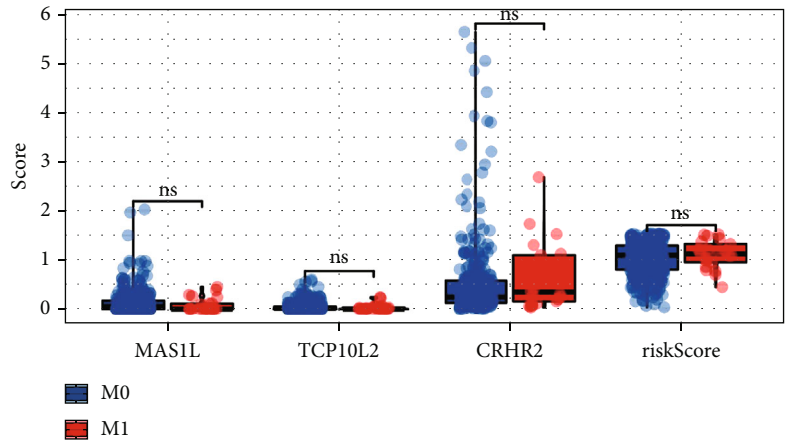


(e)

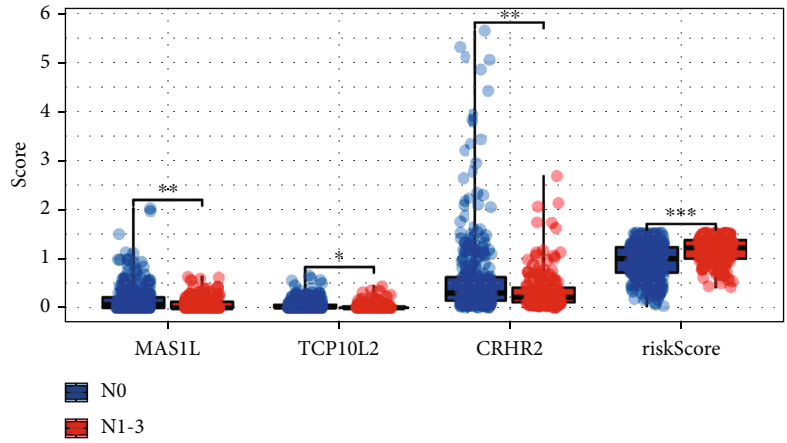


(f)

Figure 4: Continued.



(g)



(h)

Figure 4: Clinical correlation of our model. Notes: (a–b) univariate and multivariate analysis showed that the model is a risk factor independent of other clinical features; (c) the expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different age groups, ns = $P > 0.05$; (d) The expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different gender groups, ns = $P > 0.05$, $^{*} = P < 0.05$, and $^{**} = P < 0.01$; (e) the expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different stage groups, ns = $P > 0.05$ and $*** = P < 0.001$; (f) the expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different T-stage groups, ns = $P > 0.05$; (g) the expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different M-stage groups, ns = $P > 0.05$; (h) the expression level of *MAS1L*, *TCP10L2*, *CRHR2*, and riskscore in patients with different N-stage groups, $^{*} = P < 0.05$, $^{**} = P < 0.01$, and $^{***} = P < 0.001$.

performed the clinical correlation of our model, as well as the model genes. No significant difference was found in model genes and riskscore between <= 65 and > 65 patients (Figure 4(c)); MAS1L was higher expressed in female patients compared with the male patients (Figure 4(d)); MAS1L was higher expressed in stage I-II patients compared with the stage III-IV patients (Figure 4(e)); no significant difference was found in model genes and riskscore between different T-stage patients (Figure 4(f)); no significant difference was found in model genes and riskscore between different M-stage patients (Figure 4(g)); N1-3 patients had a lower *MAS1L*, *TCP10L2*, and *CRHR2* expression, but a higher riskscore (Figure 4(h)).

3.4. Pathway Enrichment Analysis and Immunotherapy Analysis. We next explored the underlying biological differences between high- and low-risk patients. GSEA analysis showed that in the high-risk patients, the pathway of *MYC* target, unfolded protein response, interferon alpha response, DNA repair, reactive oxygen species pathway, and glycolysis were significantly enriched in. (Figure 5(a)). ClueGO analysis showed that our model was mainly enriched in cell proliferation in external granule layer, proximal/distal pattern formation, dorsal/ventral pattern formation, response to immobilization stress, and negative regulation of gene expression and epigenetic (Figure 5(b)). Gene ontology (GO) analysis showed that in the high-risk patients, the terms of DNA replication checkpoint signaling, DNA strand elongation involved in DNA replication, positive regulation of telomerase RNA localization to cajal body, DNA replication origin binding, anaphase promoting complex dependent catabolic process, establishment of protein localization to telomere, and kinetochore assembly were significantly enriched (Figure S2A). Kyoto Encyclopedia of Genes and

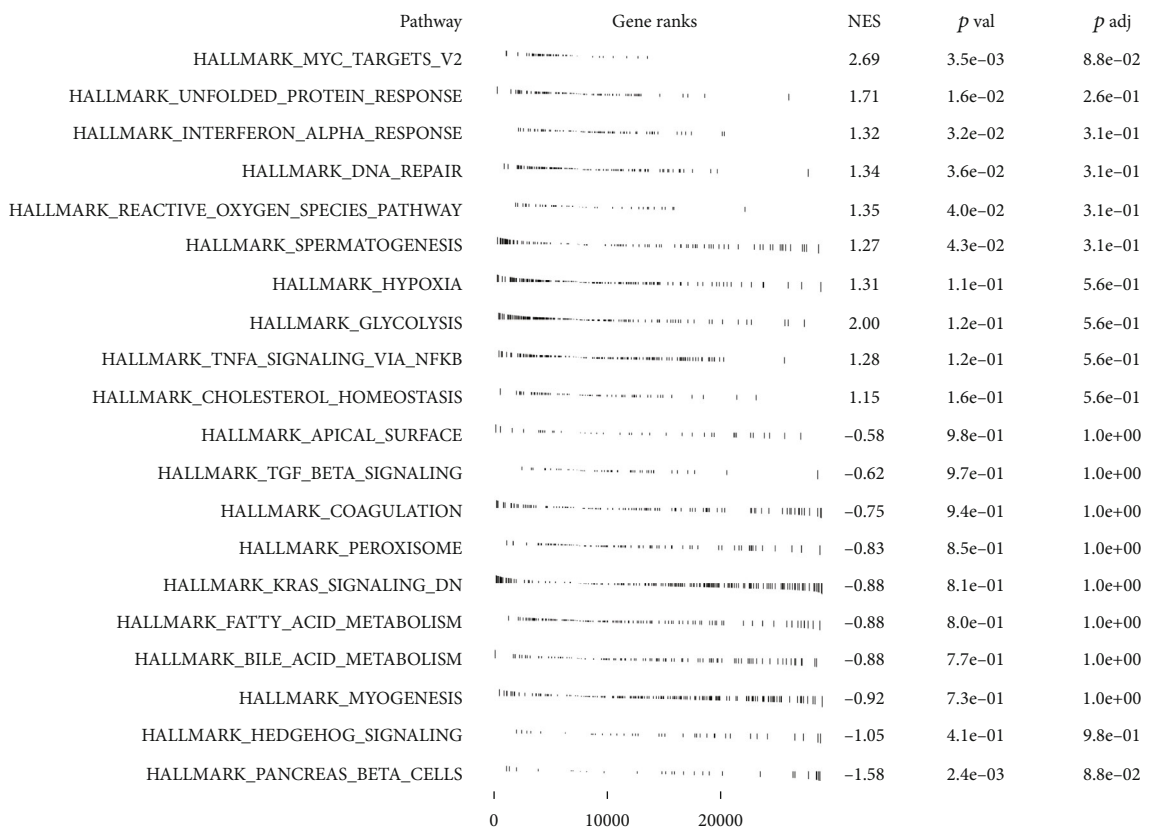


Figure 5: Pathway enrichment analysis. Notes: (a) GSEA analysis of high-risk based on the Hallmark gene set; (b) ClueGO analysis of the high-risk.

Genomes (KEGG) analysis indicated that mismatch repair, citrate cycle TCA cycle, homologous recombination, DNA replication, proteasome, and ribosome were significantly enriched (Figure S2B). A positive correlation was found between TIDE and riskscore (Figure S1A, Correlation = 0.193, $P < 0.001$). Meanwhile, a higher TIDE score was found in the high-risk patients, indicating a lower percentage of immunotherapy responders in high-risk group (Figure S1B-C, 23.1% vs. 43.4%).

3.5. Further Exploration of MAS1L, TCP10L2, and CRHR2. Furthermore, we tried to compare the expression level of *MAS1L*, *TCP10L2*, and *CRHR2* in normal and LUAD samples (Figures 6(a–c)). The result showed that *MAS1L* was

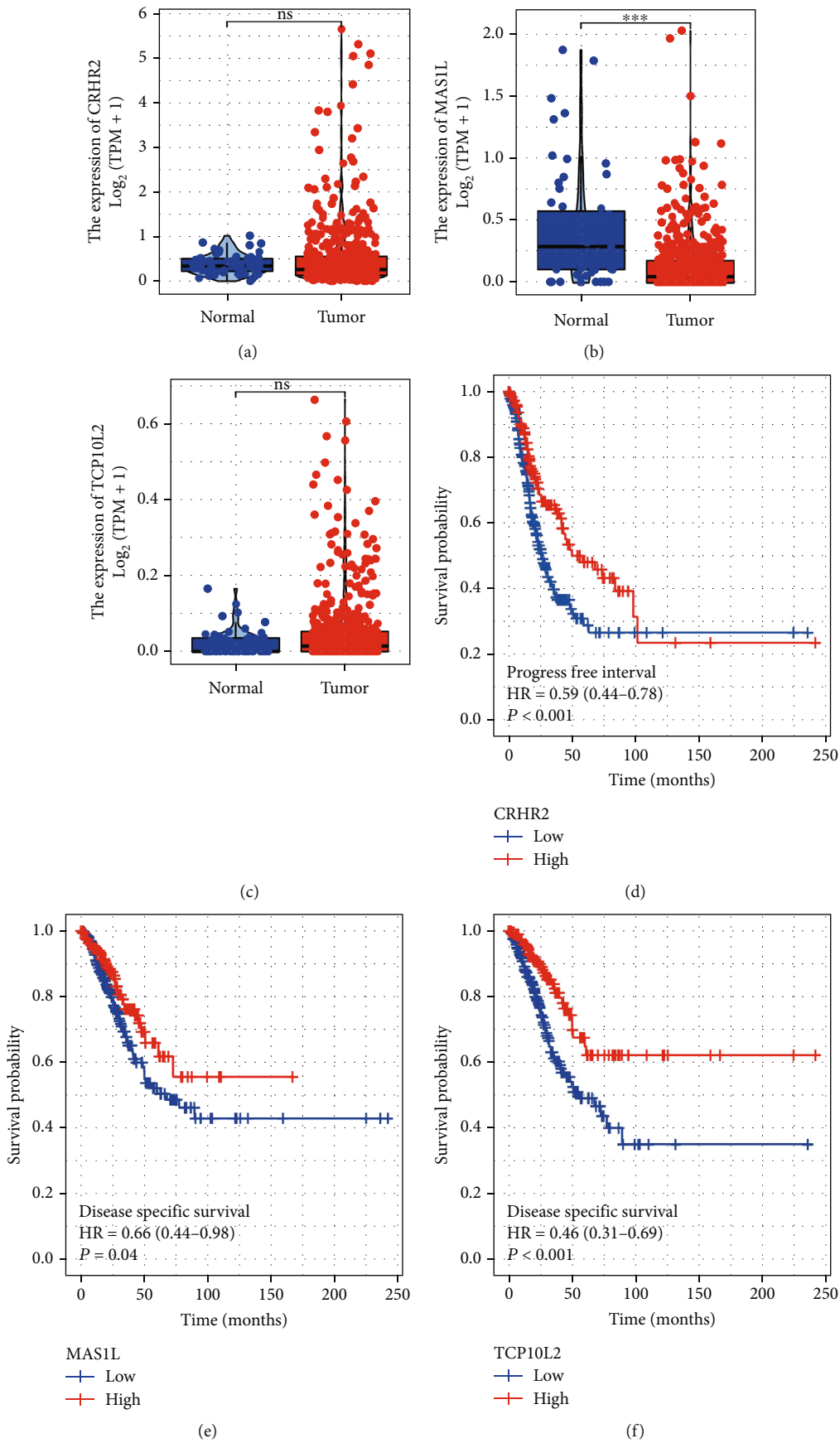


Figure 6: Continued.

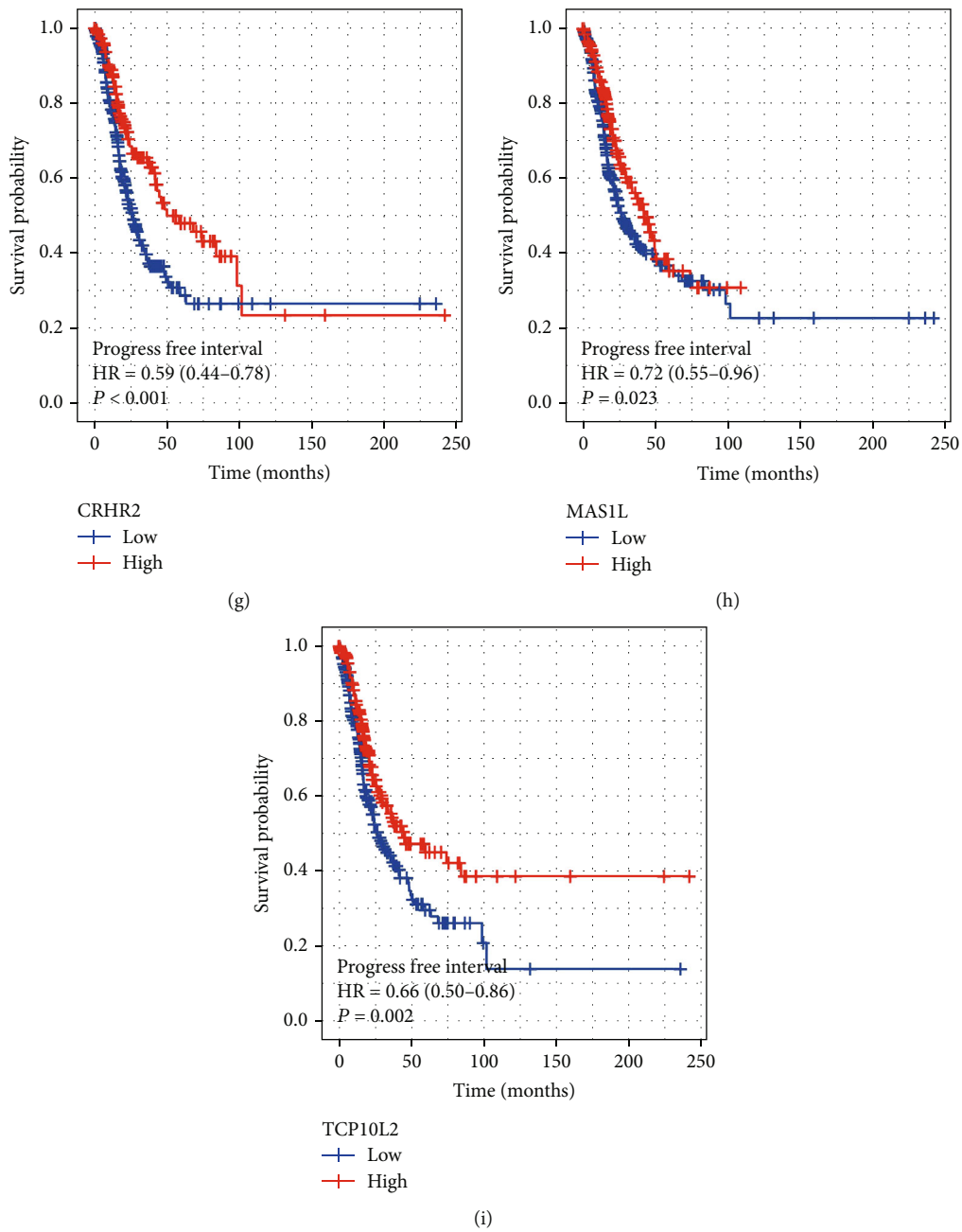


Figure 6: Further exploration of *MAS1L*, *TCP10L2*, and *CRHR2*. Notes: (a–c) the expression level of *MAS1L*, *TCP10L2*, and *CRHR2* in normal and tumor tissue; (d–f) disease-specific survival of *MAS1L*, *TCP10L2*, and *CRHR2*; (g–i) progression-free survival of *MAS1L*, *TCP10L2*, and *CRHR2*.

significantly downregulated in LUAD samples (Figure 6(b)). Moreover, KM survival curves showed that the patients with higher *MAS1L*, *TCP10L2*, and *CRHR2* might have better disease-free survival and progression-free survival (Figures 6(d–i)). *MAS1L* aroused our interest and therefore selected for further analysis. Pathway enrichment analysis showed that in the patients with high MAS1L expression, the pathway of apical surface, TGF-β signaling, coagulation, peroxisome, KRAS signaling, fatty acid metabolism, and bile acid metabolism hedgehog signaling were downregulated (Figure 7(a)). GO analysis showed that the terms of RNA binding involved in posttranscriptional gene silencing, T cell receptor complex, plasma membrane signaling receptor complex, bitter taste receptor activity, spliceosomal tri snrnp complex assembly, and cajal body were significantly enriched in patients with high *MAS1L* level (Figure S3A). KEGG analysis showed that the terms of intestinal immune network for iga production, asthma, allograft rejection, hematopoietic cell lineage, viral myocarditis, and autoimmune thyroid disease were significantly enriched in the patients with high *MAS1L* level (Figure S3B). Next, we explored the correlation between *MAS1L* and genomic

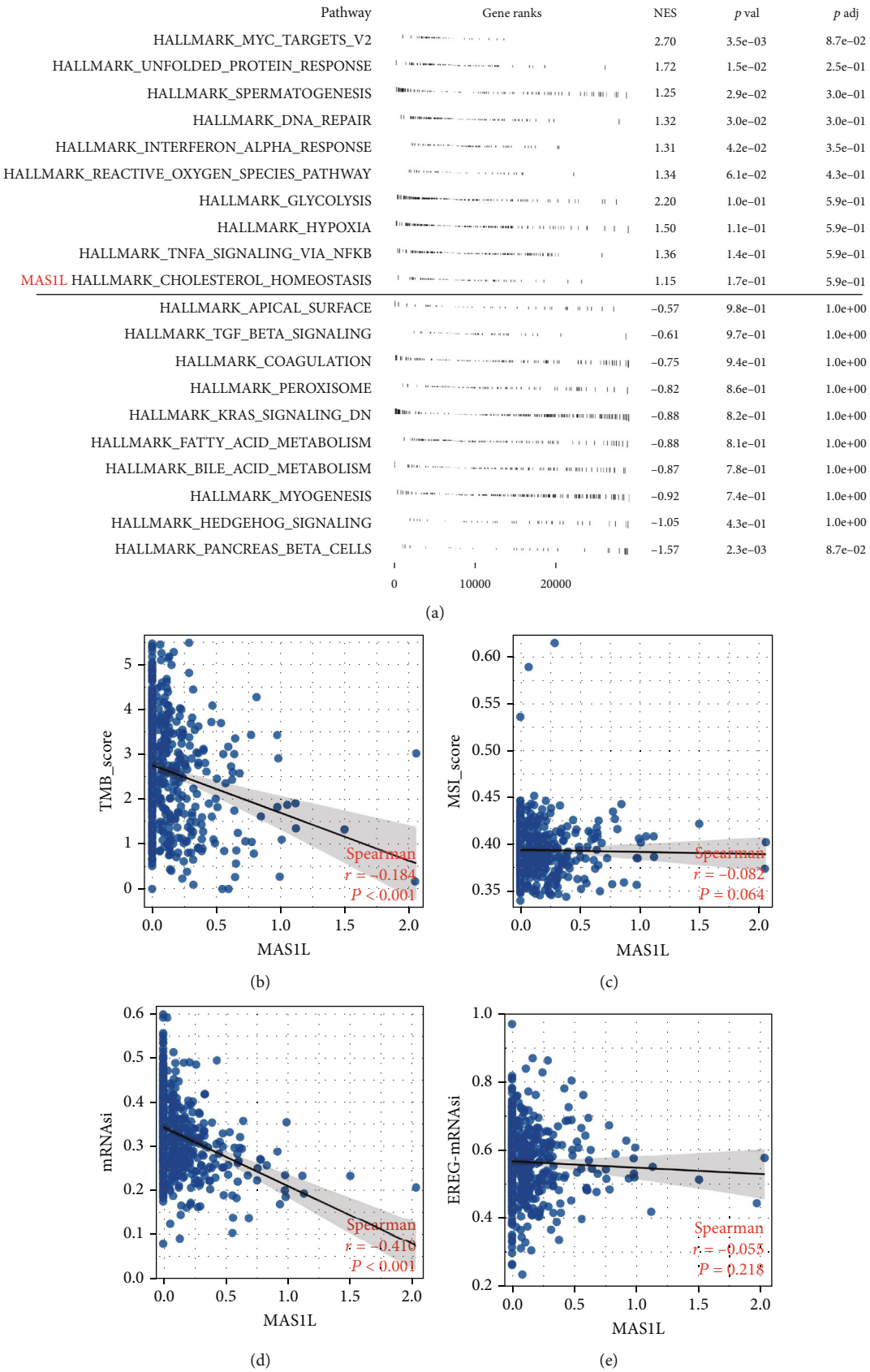


Figure 7: Pathway enrichment and genomic instability of MAS1L. Notes: (a) GSEA analysis of high and low MAS1L based on the Hallmark gene set; (b–e) the correlation between riskscore and TMB, MSI, mRNAsi, and EREG-mRNAsi.

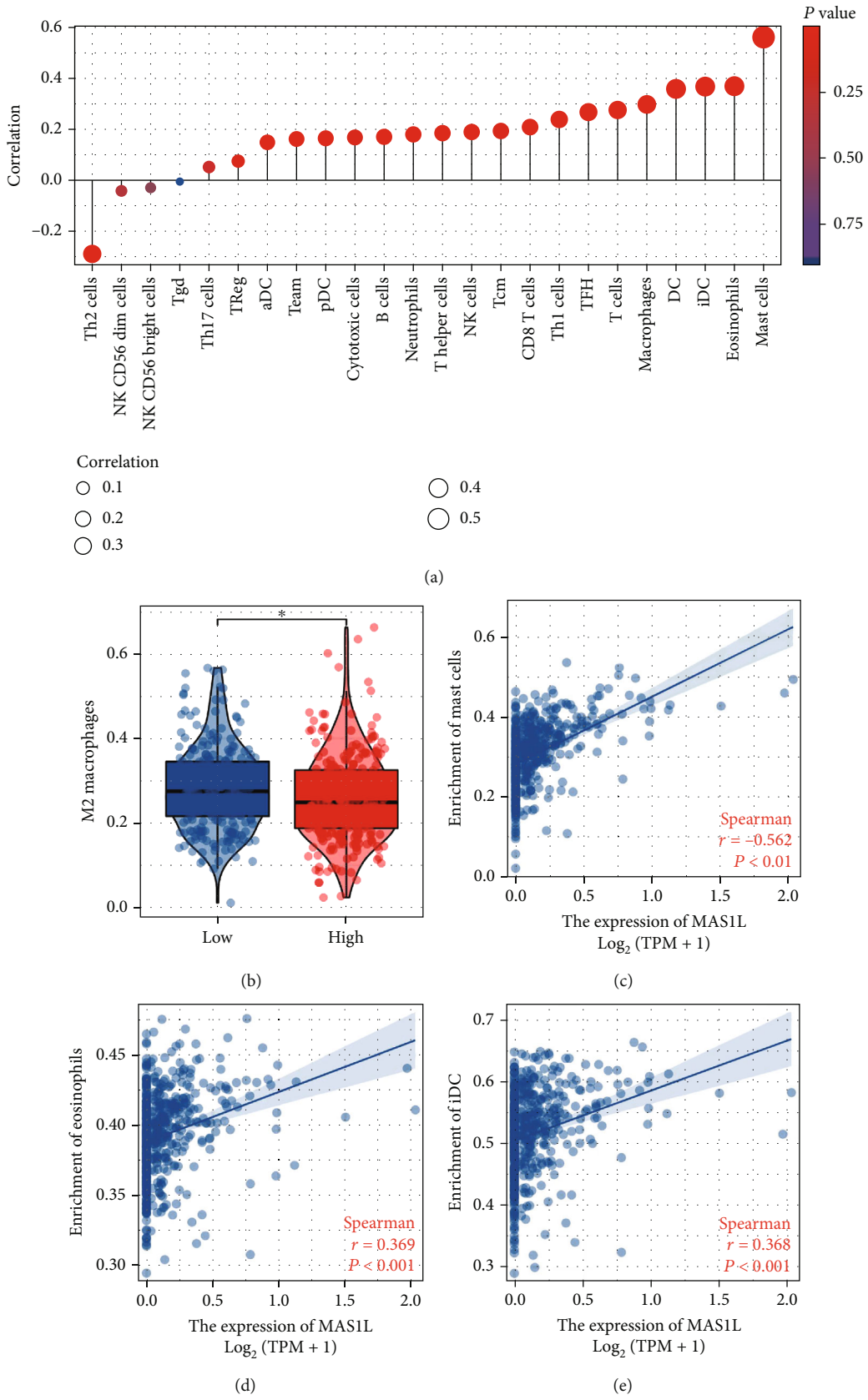


Figure 8: Continued.

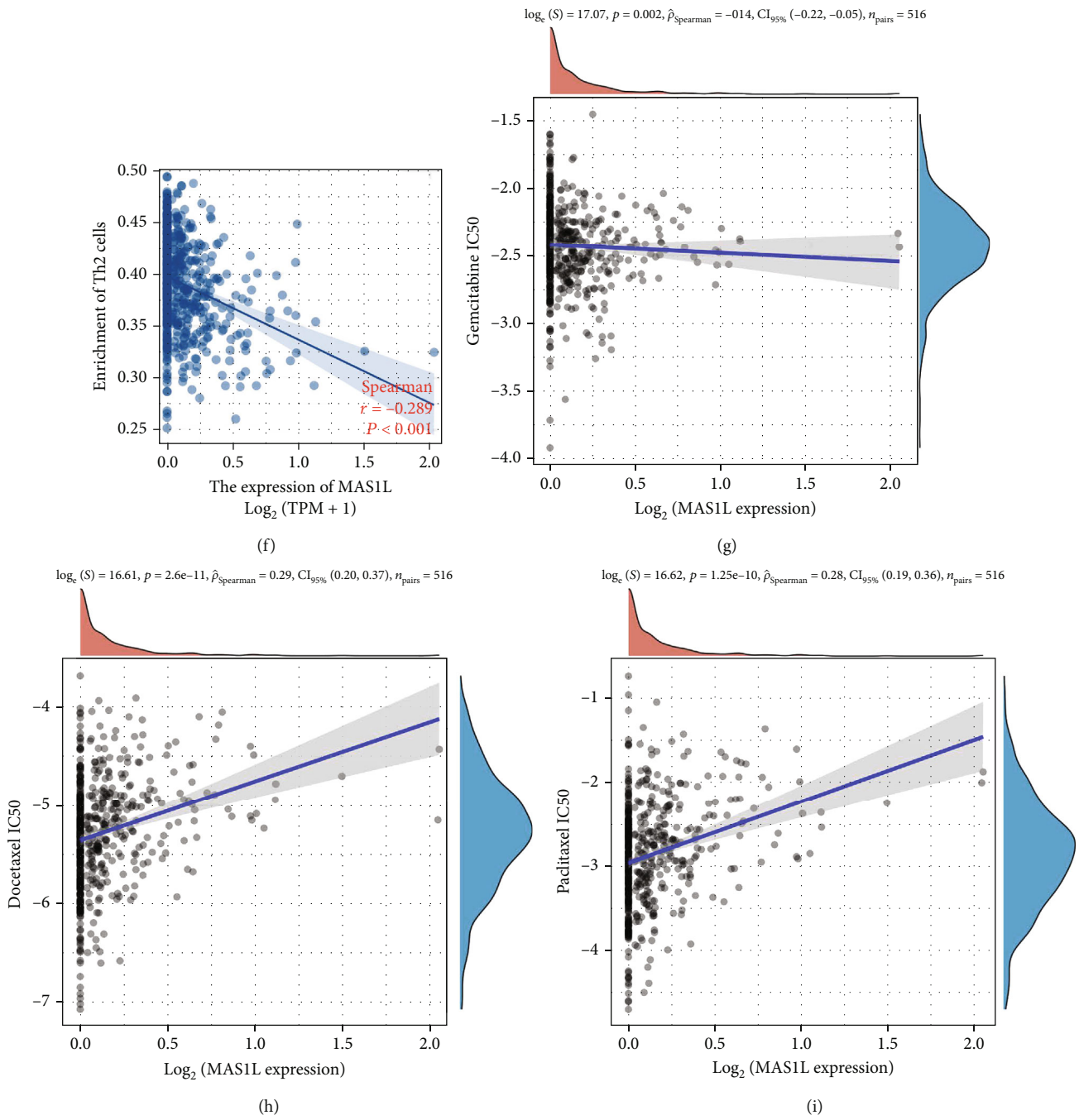
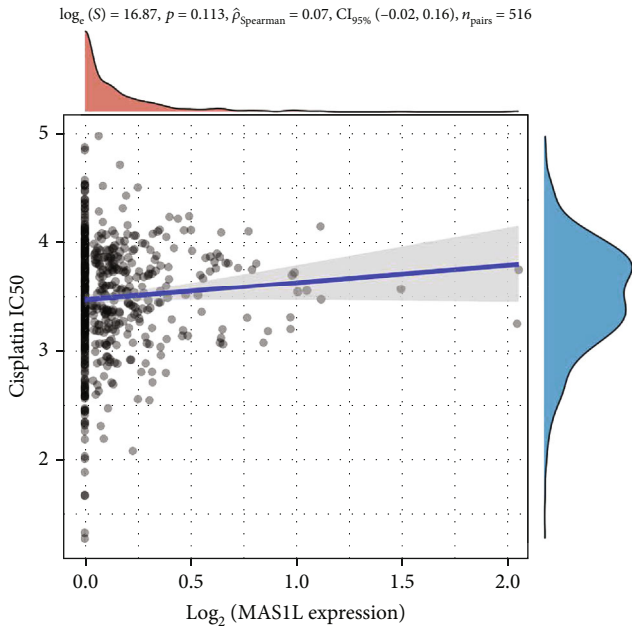


Figure 8: Continued.



(j)

Figure 8: Immune infiltration and drug sensitivity analysis of *MAS1L*. Notes: (a) the CIBERSORT algorithm was used to quantify the immune infiltration of cancer tissue; (b) the patients with high *MAS1L* might have a lower M2 macrophages infiltration; (c–f) the correlation of riskscore and mast cells, eosinophil, iDC, and Th2 cells; (g–j) the correlation of riskscore and the IC50 of Gemcitabine, Docetaxel, Paclitaxel, and Cisplatin.

instability (Figures 7(b–e)). A negative correlation was found between *MAS1L* and TMB score and mRNAsi (Figures 7(b) and 7(r) = -0.184, $P < 0.001$; Figures 7(d) and 7(r) = -0.416, $P < 0.001$). However, no remarkable correlation was observed between the *MAS1L* and MSI score and EREG-mRNAsi (Figures 7(c) and 7(e)).

3.6. Immune Infiltration and Drug Sensitivity of MAS1L. Immune microenvironment played an important role in tumor development. Immune infiltration analysis showed that riskscore was positively correlated with the mast cells, eosinophils, iDC, DC, and macrophages, while negatively correlated with Th2 cells (Figures 8(a) and 8(c–f)). Also, we found that the patients with high *MAS1L* expression might have a higher M2 macrophages infiltration (Figure 8(b)). Moreover, we performed drug sensitivity analysis to explore the underlying effect of MAS1L on the chemotherapeutic drugs of lung cancer (Figures 8(g–j)). The result indicated that the patients with high *MAS1L* expression might have a lower sensitivity to docetaxel and paclitaxel (Figures 8(h) and 8(i)).

4. Discussion

Lung cancer is a serious public health concern worldwide [1]. Lymph node metastasis is common in lung cancer and regarded as an independent prognosis factor [18]. However, the underlying biological mechanisms affecting the lymph node metastasis of lung cancer have not been fully explored.

In our study, we firstly identified the characteristic genes of lymph node metastasis in LUAD through two machine learning algorithms, LASSO logistic regression and SVM-RFE algorithms. Ten characteristic genes were finally identified, including *CRHR2*, *ITIH1*, *PRSS48*, *MAS1L*, *CYP4Z1*, *LMO1*, *TCP10L2*, *KRT78*, *IGFBP1*, and *PITX3*. In the clinical practice, detecting the relative expression of these genes can indicate the risk group of patients, as well as their response to chemotherapy and immunotherapy.

Next, we performed univariate Cox regression, LASSO regression, and multivariate Cox regression sequentially to construct a prognosis model based on *MAS1L*, *TCP10L2*, and *CRHR2*, which had a good prognosis prediction efficiency in both training and validation cohorts. Univariate and multivariate analysis indicated that our model is a risk factor independent of other clinical features. Next, pathway enrichment analysis was performed to explore the underlying biological differences between high- and low-risk patients. Among three model genes, *MAS1L* aroused our interest and therefore was selected for further analysis. KM survival curves showed that the patients with higher MAS1L might have better disease-free survival and progression-free survival. Further, pathway enrichment, genomic instability, immune infiltration, and drug sensitivity analysis were performed to in-deep explore the role of *MAS1L* in LUAD.

Our result showed that the pathway of MYC target, unfolded protein response, interferon alpha response, DNA repair, reactive oxygen species pathway, and glycolysis were significantly enriched in the high-risk patients. A break in the balance of DNA damage and repair would lead to the accumulation of oncogenes in tumor cells, leading to

genomic instability and malignant progression [19]. In lung cancer, Tian et al. found that targeting UHRF1-dependent DNA repair could selectively sensitize *KRAS* mutant lung cancer to chemotherapy [20]. Glycolysis is widely involved in the development of lung cancer. Hua et al. revealed that LINC01123 could facilitate growth and aerobic glycolysis of lung cancer through the miR-199a-5p/c-Myc axis [21]. Wiel et al. found that *BACH1* can activate the transcription of hexokinase 2 and *GAPDH* and increases glucose uptake, glycolysis rate, and lactate secretion, thereby stimulating glycolysis-dependent metastasis of lung cancer cells [22]. Zhou et al. found indicated that CircRNA-ENO1 could promote glycolysis and tumor progression in LUAD through upregulating its host gene *ENO1* [23]. The difference in prognosis between high-risk and low-risk patients may be the result of the interaction of these pathways.

Underlying genomic burden lead to the diverse performance of patients. Therefore, we found that the *MAS1L* was negatively correlated with the TMB and mRNAsi. In brief, TMB is the number of mutations in tumors, which can reflect the instability of the genome to some extent [24]. Generally, a higher TMB level in the tumor microenvironment can increase the intratumoral heterogeneity, making cancer cells more aggressive [25]. Tumor stemness index, like mRNAsi, is an index to evaluate the similarity between tumor cells and stem cells, which is associated with the active biological processes in stem cells and a higher degree of tumor dedifferentiation [26]. In lung cancer, Hong et al. found that the circular RNA circ_CPA4 could promote lung cancer proliferation, stemness, drug resistance, and immune evasion through the miR-let-7/PD-L1 axis [27]. Interestingly, Schaal et al. revealed that nicotine and electronic cigarettes could promote self-renewal of stem-like side-population cells, implicated in the dormancy, metastasis, and drug resistance in lung cancer [28].

Recently, the microenvironment of tumor cells located in has gained increasing attention from researchers. Immune cells are an important component of tumor microenvironment. Our result showed that *MAS1L* was positively correlated with the mast cells, eosinophils, iDC, DC, and M1 macrophages, while negatively correlated with Th2 cells. Eosinophils are rare multifunctional granulocytes and have been reported to play an antitumor role in cancer. Through manipulating eosinophil-related cytokines *CCL11* and *IL-5*, Simson et al. found a negative correlation between tumor growth and eosinophil infiltration [29]. The activation of activated eosinophils promotes tumor rejection through recruitment, activation, and maturation of several immune cells in addition to its direct cytotoxic actions on cancer cells [30]. Carretero et al. indicated that eosinophils could recruit cytotoxic CD8+ T cells to promote tumor rejection [31].

Some limitations should be noticed. Firstly, the patients enrolled in our study were predominantly western populations, which might lead to underlying race bias and reduce the credibility of our conclusions. Secondly, the location of lymph node metastasis is not fully provided. If the relevant data is further improved in the future, this will increase the stability of our conclusions.

Data Availability

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

Conflicts of Interest

The authors declare that they have no competing interests.

Supplementary Materials

Figure S1: immunotherapy difference between high- and low-risk patients. Notes: (a) correlation between the risk-score and TIDE; (b) TIDE difference between high- and low-risk patients; (c–d) the percentage of immunotherapy responders and nonresponders in high- and low-risk patients. Figure S2: biological enrichment of high- and low-risk patients. Notes: (a) GO analysis; (b) KEGG analysis. Figure S3: biological enrichment of patients with high and low MAS1L expression. Notes: (a) GO analysis; (b) KEGG analysis. *(Supplementary Materials)*

References

[1] F. Nasim, B. F. Sabath, and G. A. Eapen, "Lung cancer," *The Medical Clinics of North America*, vol. 103, no. 3, pp. 463–473, 2019.

[2] L. G. Collins, C. Haines, R. Perkel, and R. E. Enck, "Lung cancer: diagnosis and management," *American Family Physician*, vol. 75, no. 1, pp. 56–63, 2007.

[3] C. Gridelli, A. Rossi, D. P. Carbone et al., "Non-small-cell lung cancer," *Nature Reviews. Disease Primers*, vol. 1, no. 1, p. 15009, 2015.

[4] K. G. Tournoy, J. T. Annema, M. Krasnik, F. J. Herth, and J. P. van Meerbeeck, "Endoscopic and endobronchial ultrasonography according to the proposed lymph node map definition in the seventh edition of the tumor, node, metastasis classification for lung cancer," *Journal of Thoracic Oncology*, vol. 4, no. 12, pp. 1576–1584, 2009.

[5] S. Watanabe and H. Asamura, "Lymph node dissection for lung cancer: significance, strategy, and technique," *Journal of Thoracic Oncology*, vol. 4, no. 5, pp. 652–657, 2009.

[6] Q. Liu, M. Qiao, Z. Lohinai et al., "CCL19 associates with lymph node metastasis and inferior prognosis in patients with small cell lung cancer," *Lung Cancer*, vol. 162, pp. 194–202, 2021.

[7] M. M. Bi, B. Shang, Z. Wang, and G. Chen, "Expression of CXCR4 and VEGF-C is correlated with lymph node metastasis in non-small cell lung cancer," *Thoracic Cancer*, vol. 8, no. 6, pp. 634–641, 2017.

[8] J. Friedman, T. Hastie, and R. Tibshirani, "Regularization paths for generalized linear models via coordinate descent," *Journal of Statistical Software*, vol. 33, no. 1, pp. 1–22, 2010.

[9] M. L. Huang, Y. H. Hung, W. M. Lee, R. K. Li, and B. R. Jiang, "Svm-Rfe based feature selection and Taguchi parameters optimization for multiclass Svm classifier," *The Scientific World Journal*, vol. 2014, Article ID 795624, 10 pages, 2014.

[10] R. Tibshirani, "The Lasso method for variable selection in the cox model," *Statistics in Medicine*, vol. 16, no. 4, pp. 385–395, 1997.

[11] C. Peng, L. Li, G. Luo, S. Tan, R. Xia, and L. Zeng, "Integrated analysis of the M2 macrophage-related signature associated with prognosis in ovarian cancer," *Frontiers in Oncology*, vol. 12, p. 986885, 2022.

[12] X. Ren, X. Chen, K. Fang et al., "COL5A2 promotes proliferation and invasion in prostate cancer and is one of seven Gleason-related genes that predict recurrence-free survival," *Frontiers in Oncology*, vol. 11, p. 583083, 2021.

[13] G. Bindea, B. Mlecnik, H. Hackl et al., "Cluego: a Ccytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks," *Bioinformatics*, vol. 25, no. 8, pp. 1091–1093, 2009.

[14] T. M. Malta, A. Sokolov, A. J. Gentles et al., "Machine learning identifies stemness features associated with oncogenic dedifferentiation," *Cell*, vol. 173, no. 2, pp. 338–354.e15, 2018.

[15] S. Hänzelmann, R. Castelo, and J. Guinney, "Gsva: gene set variation analysis for microarray and Rna-Seq data," *BMC Bioinformatics*, vol. 14, no. 1, p. 7, 2013.

[16] J. Fu, K. Li, W. Zhang et al., "Large-scale public data reuse to model immunotherapy response and resistance," *Genome Medicine*, vol. 12, no. 1, p. 21, 2020.

[17] W. Yang, J. Soares, P. Greninger et al., "Genomics of drug sensitivity in cancer (Gdsc): a resource for therapeutic biomarker discovery in cancer cells," *Nucleic Acids Research*, vol. 41, no. - Database issue, pp. D955–D961, 2013.

[18] T. Sato, Y. Shimada, T. Mimae et al., "The impact of pathological lymph node metastasis with lymphatic invasion on the survival of patients with clinically node-negative non-small cell lung cancer: a multicenter study," *Lung Cancer*, vol. 158, pp. 9–14, 2021.

[19] N. Chatterjee and G. C. Walker, "Mechanisms of DNA damage, repair, and mutagenesis," *Environmental and Molecular Mutagenesis*, vol. 58, no. 5, pp. 235–263, 2017.

[20] D. Tian, J. Tang, X. Geng et al., "Targeting UHRF1-dependent DNA repair selectively sensitizes Kras mutant lung cancer to chemotherapy," *Cancer Letters*, vol. 493, pp. 80–90, 2020.

[21] Q. Hua, M. Jin, B. Mi et al., "Linc 01123, a C-Myc-activated long non-coding Rna, promotes proliferation and aerobic glycolysis of non-small cell lung cancer through Mir-199a-5p/C-Myc Axis," *Journal of Hematology & Oncology*, vol. 12, no. 1, p. 91, 2019.

[22] C. Wiel, K. le Gal, M. X. Ibrahim et al., "BACH1 stabilization by antioxidants stimulates lung cancer metastasis," *Cell*, vol. 178, no. 2, pp. 330–345.e22, 2019.

[23] J. Zhou, S. Zhang, Z. Chen, Z. He, Y. Xu, and Z. Li, "CircRNA-ENO1 promoted glycolysis and tumor progression in lung adenocarcinoma through upregulating its host gene ENO1," *Cell Death & Disease*, vol. 10, no. 12, p. 885, 2019.

[24] D. M. Merino, L. M. McShane, D. Fabrizio et al., "Establishing guidelines to harmonize tumor mutational burden (Tmb): in silico assessment of variation in Tmb quantification across diagnostic platforms: phase I of the friends of cancer Rresearch Tmb harmonization Project," *Journal for Immunotherapy of Cancer*, vol. 8, no. 1, article e000147, 2020.

[25] T. A. Chan, M. Yarchoan, E. Jaffee et al., "Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic," *Annals of Oncology*, vol. 30, no. 1, pp. 44–56, 2019.

[26] P. R. Prasetyanti and J. P. Medema, "Intra-tumor heterogeneity from a cancer stem cell perspective," *Molecular Cancer*, vol. 16, no. 1, p. 41, 2017.

[27] W. Hong, M. Xue, J. Jiang, Y. Zhang, and X. Gao, "Circular Rna circ-CPA4/ let-7 Mirna/Pd-L1 Axis regulates cell growth, stemness, drug resistance and immune evasion in non-small cell lung cancer (Nsclc)," *The Journal of Experimental & Clinical Cancer Research*, vol. 39, no. 1, p. 149, 2020.

[28] C. M. Schaal, N. Bora-Singhal, D. M. Kumar, and S. P. Chellappan, "Regulation of sox 2 and stemness by nicotine and electronic-cigarettes in non-small cell lung cancer," *Molecular Cancer*, vol. 17, no. 1, p. 149, 2018.

[29] L. Simson, J. I. Ellyard, L. A. Dent et al., "Regulation of carcinogenesis by Il-5 and Ccl 11: a potential role for eosinophils in tumor immune surveillance," *Journal of Immunology (Baltimore, Md. : 1950)*, vol. 178, no. 7, pp. 4222–4229, 2007.

[30] S. M. O'Flaherty, K. Sutummaporn, W. L. Häggtoft et al., "Tlr-stimulated eosinophils mediate recruitment and activation of Nk cells in vivo," *Scandinavian Journal of Immunology*, vol. 85, no. 6, pp. 417–424, 2017.

[31] R. Carretero, I. M. Sektioglu, N. Garbi, O. C. Salgado, P. Beckhove, and G. J. Hämmerling, "Eosinophils orchestrate cancer rejection by normalizing tumor vessels and enhancing infiltration of $CD8^+$ T cells," *Nature Immunology*, vol. 16, no. 6, pp. 609–617, 2015.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 5741437, 8 pages
https://doi.org/10.1155/2022/5741437

Research Article

The Evaluation of Clinical Status of Endoscopic Retrograde Cholangiography for the Placement of Metal and Plastic Stents in Cholangiocarcinoma Therapy

Min Gong, Qiang Li, You Xu, and Yunhui Fu

Department of Gastroenterology, Jiangxi Pingxiang People's Hospital, Pingxiang 337055, Jiangxi, China

Correspondence should be addressed to Min Gong; gongminxh0799@126.com

Received 20 July 2022; Revised 7 September 2022; Accepted 20 September 2022; Published 11 October 2022

Academic Editor: Jincheng Wang

Objective. Cholangiocarcinoma is a common malignant tumor that occurs in the bile duct system, which can be treated by using the endoscopic retrograde cholangiography (ERCP). This study was aimed at exploring the therapeutic effect of ERCP with metal stent and plastic stent for cholangiocarcinoma. *Methods.* The clinical data of 71 patients with cholangiocarcinoma treated by ERCP in our hospital from June 2020 to October 2021 were retrospectively analyzed. According to different stent types, the patients were divided into plastic stent group ($n = 43$) and metal stent group ($n = 28$). Patients in the plastic stent group and metal stent group were received with plastic stent and metal stent, respectively. The indexes of liver function (serum alkaline phosphatase (ALT), direct bilirubin (DBIL), glutamic oxaloacetic transaminase (AST), alkaline phosphatase (ALP), and total bilirubin (TBIL)), postoperative complications, success rate of stent implantation, and survival time of patients in the two groups were determined. Logistic multivariate regression analysis was used to analyze the prognostic factors of postoperative cholangiocarcinoma. *Results.* The liver function indexes of the two groups were significantly improved after treatment with the stent, in which the ameliorative effect in the metal stent group was better than that in the plastic stent group ($P < 0.05$). The incidence of postoperative complications in the plastic stent group and the metal stent group was 53.49% and 14.29%, respectively, and the success rate of stent placement was 60.47% and 96.43%, respectively. The incidence of complications in the metal stent group was lower than that in the plastic stent group, and the success rate of stent placement was higher than that in the plastic stent group ($P < 0.05$). The median survival time of patients in the plastic stent group and the metal stent group was 8.15 and 11.83 months, respectively. The survival time of patients in the metal stent group was longer than that of the plastic stent group. The median survival time of patients with types I, II, III, and IV was 12.73, 11.54, 10.57, and 9.36 months, respectively. The survival time of patients with stage I was significantly higher than that of patients with types II, III, and IV. There was an inverse relationship between the disease type and the survival time of patients. Logistic multivariate regression analysis showed that tumor diameter ≥ 5 cm, portal vein invasion, lymph node metastasis, and classification of hilar cholangiocarcinoma were the risk factors ($P < 0.05$) and metal stent type was the protective factor ($P < 0.05$). *Conclusion.* In the clinical treatment of patients with cholangiocarcinoma, the placement of metal stent and plastic stent under ERCP plays an important role. The placement of the metal stent under ERCP has a higher success rate and better prognosis and can prolong the survival time of patients to a greater extent, but the price of the metal stent is relatively expensive. For patients with an expected survival period of more than 4-6 months, the metal stent should be considered; otherwise, the plastic stent can be used to maintain cost-effectiveness. Therefore, it is necessary to comprehensively analyze the patient's economic affordability, expected survival time, stent drainage time, and personal needs and then select an appropriate treatment method.

1. Introduction

Cholangiocarcinoma is the malignant tumor originating from the extrahepatic bile duct, including the bile duct from the hilar area to the lower segment of the common bile duct. As a common malignant tumor disease, cholangiocarcinoma occurs in the bile duct system with incompletely defined etiology. It is generally believed that the cause of the disease is

related to diseases such as bile duct stones and primary cirrhotic cholangitis. In addition, smoking and drinking are the main causes. Cholangiocarcinoma usually occurs in people aged 50-70 years with slight preference in males. Patients have no special clinical symptoms in the early stage of the disease. With the prolongation of the disease time, the patients will have clinical symptoms such as fever, abdominal pain, fatigue, jaundice, and loss of appetite [1]. At present, cholangiocarcinoma is mainly treated by nutritional support, surgery, radiotherapy and chemotherapy, and interventional therapy [2]. Palliative biliary drainage is commonly used in patients with unresectable cholangiocarcinoma to relieve obstructive jaundice, pruritus, or pain to prolong survival. In recent years, with the development of endoscopic technology, endoscopic retrograde cholangiography (ERCP) stent placement has been gradually applied in clinical practice. ERCP stent placement is first used for clinical treatment in 1979. ERCP stent placement has replaced some surgical operations in the treatment of biliary and pancreatic diseases. It is a minimally invasive interventional treatment and has become the main method for the treatment of cholangiocarcinoma [3].

In this study, the placement of metal and plastic stents in ERCP can effectively relieve the symptoms of obstructive jaundice in patients with obvious advantages. In addition, patients with the placement of stents in ERCP display high postoperative survival rate with no obvious damage, lower pain degree than traditional treatment, and safer treatment, which make it being with broad prospect of clinical application. Biliary stents can be divided into metal stents and plastic stents according to the different materials of the implanted biliary stent. Each has its own advantages and disadvantages. When choosing the type of stent placement, the extent and degree of the lesion should be considered, and the patient survival time and stent drainage time should be fully estimated to obtain a higher benefit ratio. The choice of the number of stents to be placed in the biliary tract is still controversial. In this study, metal stents and plastic stents were placed under ERCP to analyze the safety of the two different stents and their impact on patient survival and long-term prognosis. The reports were as follows.

2. Materials and Methods

2.1. General Information. The clinical data of 71 patients with cholangiocarcinoma treated by ERCP in our hospital from June 2020 to October 2021 were retrospectively analyzed. According to different stent types, the patients were divided into plastic stent group ($n = 43$) and metal stent group ($n = 28$). This study has been approved by the Ethics Committee of our hospital. The following are the inclusion criteria: (1) the patients were diagnosed as cholangiocarcinoma by B-ultrasound, magnetic resonance, or CT; (2) the patients with distant metastasis or late symptoms confirmed by pathology; (3) the patients could be successfully placed with the stent under the guidance of ERCP; and (4) the patients with the complete clinical medical records, and the patients and their families gave informed consent to participate in this study. The following are the exclusion criteria: (1) the patients with severe impairment of liver, kidney, and spleen function; (2) the patients with systemic tumor disease; (3) the patients with mental or cognitive impairment; (4) the patients who do not meet the surgical indications; and (5) the patient with incomplete clinical data or the researcher who withdrawn halfway. There were 43 cases in the control group, 21 women and 22 men, aged 40-92 years, with an average age of 68.34 ± 5.48, while 28 cases in the study group, 13 women and 15 men, aged 40-92 years, with an average age of 68.58 ± 5.26. There was no significant difference in general information between the two groups ($P > 0.05$). The general information of the two groups is shown in Table 1. The process of general data selection is shown in Figure 1.

2.2. Methods. Chest X-ray (X-Ray systems, CGR500, France), cardiopulmonary function (K4b~2, COSME, Italy), coagulation function (CX9ALX, Beckman, USA), blood routine (CX9ALX, Beckman), ECG (MAC5500, GE, USA), and liver function (CX9ALX, Beckman) were examined before operation. If the patient was complicated with other diseases, it needed to consult with other departments to accurately evaluate the operation risk. Patients were forbidden to drink six hours before operation. Half an hour before operation, the patient was anesthetized by giving 50 mg pethidine, 10 mg diazepam, and 10 mg anisodamine. The patient was placed in the left prone position, connected to oxygen inhalation, ECG monitoring, and electrode pads. Under endoscope, the hydrophilic guide wire was inserted into the bile duct until the obstruction site, and the obstruction site, scope, and severity were observed. The length and location of the stenosis were determined by injecting nonionic contrast agents such as iodophor. After dilating the stenosis, gradually advance to the appropriate bile duct position through the catheter, and place the bile duct metal stent and plastic stent on the bile duct obstruction site, respectively, ensuring that both ends of the stent need to exceed the obstruction site by about 1 cm. After the operation, the patients were forbidden to drink and eat. ECG monitoring was performed for 12 hours. The vital signs and abdominal conditions of the patients were observed, and the patients were treated with rehydration and liver protection. Antibiotics were given if biliary tract infection occurred. The patients were followed up for two years after the operation to understand the recovery status of the patient's condition.

2.3. Outcome Measures. (1) Before treatment and 24 hours after treatment, 3 mL of fasting venous blood was drawn in the morning. Then, the fasting venous blood was centrifuged at 3000 r/min for 10 min to collect the sera. The levels of serum alkaline phosphatase (ALT), direct bilirubin (DBIL), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and total bilirubin (TBIL) were detected using a 7080 automatic biochemical analyzer [4] (CX9ALX, Beckman) in strict accordance with the instructions.

(2) The number of patients with complications such as stent displacement, intraoperative expansion, postoperative pancreatitis, cholangitis, no decrease in total bilirubin, liver

Table 1: Analysis of general information of two groups.

General information	The plastic stent group ($n = 43$)	The metal stent group ($n = 28$)	χ^2/t	P
Gender (female/male)	21/22	13/15	0.039	0.843
Age (year)	68.34 ± 5.48	68.58 ± 5.26	0.183	0.855
Biliary stricture length (cm)	3.24 ± 0.87	3.22 ± 0.92	0.093	0.927
Albumin (<35/≥35 g/L)	15/28	9/19	0.057	0.811
Preoperative GGT (U/mL)	475.32 ± 29.54	476.51 ± 29.63	0.166	0.869
Total bilirubin (ummol/L)				
≥427	9 (20.93)	6 (21.43)	0.470	0.791
342-427	26 (60.47)	15 (53.57)		
205-342	8 (18.60)	7 (25.00)		

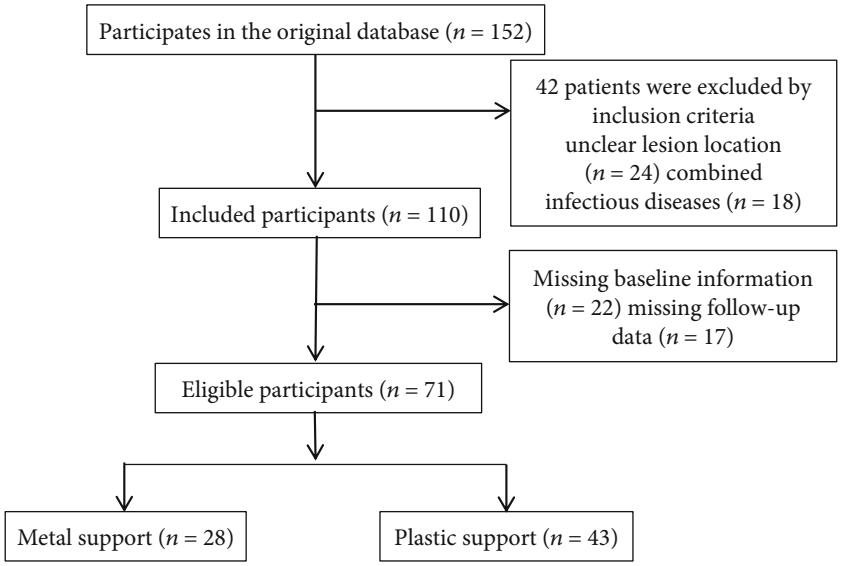


Figure 1: Process of general data selection.

Table 2: Comparison of liver function between the two groups ($\bar{x} \pm s$).

Groups		ALT (U/L)	DBIL (μmol/L)	AST (U/L)	ALP (U/L)	TBIL (μmol/L)
The plastic stent group ($n = 43$)	Before treatment	105.24 ± 4.26	165.24 ± 12.85	67.52 ± 6.58	438.52 ± 27.41	164.52 ± 12.73
	After treatment	59.65 ± 3.25[a]	53.24 ± 3.57[a]	34.84 ± 4.28[a]	135.41 ± 11.54[a]	53.42 ± 2.57[a]
The metal stent group ($n = 28$)	Before treatment	105.18 ± 4.23	165.38 ± 12.87	67.54 ± 6.43	438.41 ± 27.44	163.92 ± 12.84
	After treatment	56.42 ± 3.17[ab]	50.37 ± 3.16[ab]	30.21 ± 3.16[ab]	110.39 ± 11.51[ab]	48.46 ± 2.13[ab]

Note: [a]$P < 0.05$ compared with the same group before treatment. [b]$P < 0.05$ compared with the plastic stent group after treatment.

abscess, and bleeding was counted 24 hours after treatment, and the total incidence was calculated.

(3) The patients with cholangiocarcinoma were classified by the Bismuth-Corlette classification method. Type I: the tumor was located below the bifurcation of the common hepatic duct with the communication between the left and right hepatic ducts. Type II: the tumor occupied the confluence of the left and right hepatic ducts without no channel between them. Type III: the tumor invaded one hepatic duct. Therein, the patients with the tumor involved the right liver ducts were defined as type IIIa, while those with tumor involved the left hepatic ducts were defined as type IIIb. Type IV: type IV tumors involved bilateral hepatic ducts. The success rate of stent implantation in patients with different types was counted.

(4) The Mann-Whitney U test was used to test the stent patency time in the plastic stent group and the metal stent group. The life table method was used to calculate the overall survival rate of patients, and the Kaplan-Meier method was used to represent the survival curve.

Table 3: Comparison of postoperative complications between the two groups (cases, %).

Groups	The plastic stent group ($n = 43$)	The metal stent group ($n = 28$)	χ^2	P
Stent displacement	6 (13.95)	1 (3.57)	2.057	0.152
Intraoperative expansion	4 (9.30)	1 (3.57)	0.851	0.356
Postoperative pancreatitis	5 (11.63)	1 (3.57)	1.423	0.233
Cholangitis	3 (13.95)	1 (3.57)	2.057	0.152
No decrease in total bilirubin	2 (0.00)	0 (0.00)	1.340	0.247
Liver abscess	1 (0.00)	0 (0.00)	0.661	0.416
Bleeding	2 (4.65)	0 (0.00)	1.340	0.247
Total complication rate	23 (53.49)	4 (14.29)	11.059	<0.001

Table 4: Comparison of stent implantation success rate between the two groups (cases, %).

Type	The plastic stent group ($n = 43$)		The metal stent group ($n = 28$)		χ^2	P
	Cases	Success rate	Cases	Success rate		
Type I	7	6 (85.71)	5	5 (100.00)	0.197	0.657
Type II	15	10 (66.67)	9	9 (100.00)	0.683	0.408
Type III	11	6 (54.55)	8	8 (100.00)	2.289	0.130
Type IV	10	4 (40.00)	6	5 (83.33)	1.121	0.290
Total success rete	43	26 (60.47)	28	27 (96.43)	11.589	<0.001

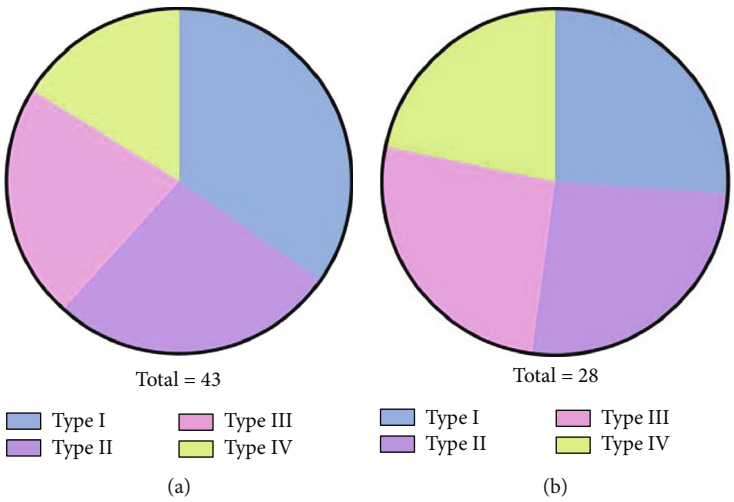


Figure 2: Comparison of stent implantation success rate between the two groups. (a) The plastic stent group. (b) The metal stent group. The success rates of stent implantation in the metal stent group were higher than that in the plastic stent group ($P < 0.05$).

(5) The following prognostic factors were analyzed: age, gender (male and female), smoking history, drinking history, family tumor history, tumor diameter (<5 cm and ≥5 cm), tumor differentiation degree (well differentiated, moderately differentiated, and poorly differentiated), portal vein invasion, lymph node metastasis, and hilar cholangiocarcinoma classification (Bismuth-Corlette classification, including type I, type II, type III, and type IV) were collected from all patients.

2.4. Statistical Analysis. The test results were analyzed by SPSS 20.0 software. The measurement data were tested by normal distribution. The data conforming to the normal distribution were expressed as ($\bar{x} \pm s$) and analyzed through independent sample t test. The enumeration data were expressed as % and analyzed through Fisher or χ^2 test. The Kaplan-Meier analysis was used to determine the survival time of patients with different types of stents. Log-rank analysis tested the survival curve. Logistic multivariate regression analysis examined the influencing factors for postoperative prognosis of cholangiocarcinoma. $P < 0.05$ indicated the statistical significance.

3. Results

3.1. Comparison of Liver Function between the Two Groups. There was no significant difference in liver function between the two groups before the treatment ($P > 0.05$). After the

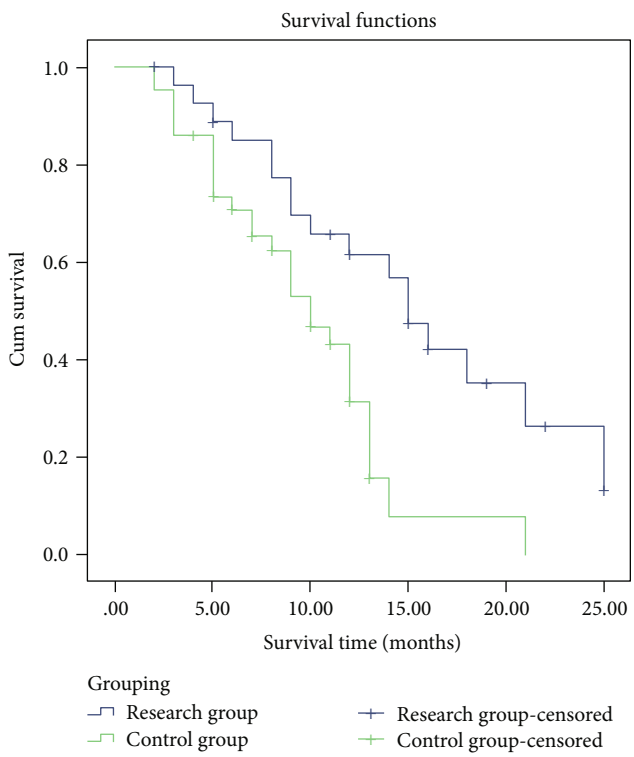


Figure 3: Comparison of survival time between the two groups.

treatment, the liver function indexes of the two groups were both significantly improved, among which the ameliorative effect of the liver function indexes in the metal stent group were notably better than that in the plastic stent group ($P < 0.05$), as shown in Table 2.

3.2. Comparison of Postoperative Complications between the Two Groups. The incidence of postoperative complications in the plastic stent group and the metal stent group was 53.49% and 14.29%, respectively. The incidence of postoperative complications in the metal stent group was lower than that in the plastic stent group ($P < 0.05$, Table 3).

3.3. Comparison of Stent Implantation Success Rate between the Two Groups. The success rates of stent implantation in the plastic stent group and the metal stent group were 60.47% and 96.43%, respectively (Table 4 and Figure 2).

3.4. Comparison of Survival Time between the Two Groups. The median survival time of the plastic stent group and the metal stent group was 8.15 and 11.83 months, respectively. The survival time of the metal stent group was longer than that of the plastic stent group ($P < 0.05$, Figure 3).

3.5. Univariate Analysis of Prognostic Factors in Patients with Cholangiocarcinoma. A total of 71 patients with cholangiocarcinoma were divided into the poor prognosis group with 33 cases and into the good prognosis group with 38 cases according to the recurrence and metastasis within two years after operation. Univariate analysis showed that there were significant differences between the poor prognosis group and the good prognosis group in the proportion of family tumor history, tumor diameter, degree of tumor differentiation, portal vein infiltration, lymph node metastasis, classification of hilar cholangiocarcinoma, and stent type ($P < 0.05$, Table 5).

3.6. Logistic Multivariate Regression Analysis. The factors with statistical significance in Table 5 were taken as independent variables, and the recurrence and metastasis of patients with cholangiocarcinoma two years after operation were taken as dependent variables for logistic multivariate regression analysis. Logistic multiple regression analysis showed that tumor diameter ≥ 5 cm, portal vein infiltration, lymph node metastasis, and classification of hilar cholangiocarcinoma were all risk factors ($P < 0.05$), while metal stent type was protective factor ($P < 0.05$, Table 6).

4. Discussion

Patients with cholangiocarcinoma are prone to malignant obstruction of the biliary tract. If not treated in time, it will threaten the patient's life [5]. The anatomical location of cholangiocarcinoma is relatively special, and the location and symptoms of cholangiocarcinoma are dormant [6]. Clinical treatment by placing bile duct stents under ERCP endoscopy can be consistent with human physiological characteristics, reduce the sense of compression on the bile duct, dredge the bile duct, and improve the quality of life of patients after operation [7]. Metal stent and plastic stent are the main materials of biliary stent, and the polyethylene plastic stent is often used at present. The plastic support has the characteristics of convenient replacement, low price, and easy acceptance by patients. However, the replacement cycle of plastic stents is relatively short, and they need to be replaced every 3-6 months, which is prone to bile sludge adhesion and high bacterial infection rate [8, 9]. In this study, metal stents and plastic stents were placed under ERCP to analyze the safety and the effects on survival cycle and long-term prognosis.

In general, metal stents have obvious advantages, such as convenient operation and high histocompatibility. At present, the clinical application of metal stent is more and more extensive, because it is not easy to block and fall off and has good patency [10]. Nevertheless, compared with plastic stents, metal stents are relatively more expensive and less selective in grass-roots hospitals. At the same time, the advantage of metal stents is not obvious for patients with short expected survival cycle, poor drainage effect, and obvious bile duct invasion [11]. Therefore, it is necessary to consider the situation of patients in all aspects when selecting stent materials [12]. This study analyzed the effect of placing bile duct metal and plastic stents under ERCP on liver function and long-term prognosis of patients with cholangiocarcinoma. The results showed that the liver function indexes of the two groups were both significantly improved after treatment, and the improvement of liver function in the metal stent group was greater than that in the plastic stent group. Placement of bile duct metal stent under ERCP can avoid the impact of bile stasis on patients' liver function, improve the blood flow in patients' liver, reduce the level of serum

Table 5: Univariate analysis of prognostic factors in patients with cholangiocarcinoma.

Prognostic factors		The poor prognosis group ($n = 33$)	The good prognosis group ($n = 38$)	χ^2	P
Age		65.88 ± 10.16	67.26 ± 8.90		
Gender	Male	19 (57.58)	15 (39.47)	2.319	0.128
	Female	14 (42.42)	23 (60.53)		
Smoking history		8 (24.24)	10 (26.32)	0.040	0.841
Drinking history		6 (18.18)	9 (23.68)	1.592	0.207
Family tumor history		6 (18.18)	1 (2.63)	4.806	0.028
Tumor diameter	<5 cm	13 (39.39)	25 (65.79)	6.180	0.013
	≥5 cm	20 (60.61)	13 (34.21)		
Degree of tumor differentiation	Well differentiated	7 (21.21)	3 (7.89)	6.697	0.035
	Moderately differentiated	8 (24.24)	20 (52.63)		
	Poorly differentiated	18 (54.55)	15 (39.47)		
Portal vein infiltration	Yes	19 (57.58)	12 (31.58)	4.853	0.028
	No	14 (42.42)	26 (68.42)		
Lymph node metastasis	Yes	25 (75.76)	15 (39.47)	9.453	0.002
	No	8 (24.24)	23 (60.53)		
Classification of hilar cholangiocarcinoma	Type I	5 (15.15)	14 (36.84)	10.702	0.013
	Type II	9 (27.27)	16 (42.11)		
	Type III	12 (36.36)	6 (15.79)		
	Type IV	7 (21.21)	2 (5.26)		
Stent type	Metal stent	7 (21.21)	21 (55.26)	8.574	0.003
	Plastic stent	26 (78.79)	17 (44.74)		

Table 6: Logistic multivariate regression analysis.

Factors	B	SE	Wald	P	OR	95% CI
Tumor diameter ≥ 5 cm	0.871	0.290	9.105	0.003	2.382	1.352~4.129
Portal vein infiltration	1.496	0.302	24.355	0.0001	4.460	2.452~8.075
Lymph node metastasis	1.110	0.356	10.213	0.001	3.026	1.526~5.972
Classification of hilar cholangiocarcinoma	1.050	0.370	8.084	0.004	2.850	1.372~5.852
Metal stent type	-1.861	0.692	7.238	0.007	0.155	0.040~0.603

total bilirubin, and enhance liver metabolic function; thus, it is of great significance for relieving bile duct pressure [13]. The results from the present study showed that metal stent was more effective than plastic stent in improving liver function. Endoscopic operation is the retrograde operation. The hydrophilic guide wire can pass through the narrow part more smoothly, which improves the flexibility of the guide wire and avoids the blindness of puncture [14]. The success rates of stent implantation in the plastic stent group and the metal stent group were 60.47% and 96.43%, respectively, in the present study. The results showed that ERCP endoscopic operation could greatly improve the success rate of stent implantation. However, the pancreatic duct will be compressed during stent implantation, which will lead to intestinal fluid reflux [15]. In addition, biliary tract infection complications may occur due to excessive injection of contrast medium during operation. Therefore, the position of stent placement should be controlled within the range of duodenal papilla to ensure the normal function of sphincter and reduce intestinal fluid reflux and the compression on pancreatic duct [16]. In this study, the incidence of postoperative complications in the plastic stent group and the metal stent group was 53.49% and 14.29%, respectively. The results confirmed that ERCP endoscopic metal stent implantation was more feasible and worthy of palliative treatment.

In recent years, new interventional devices and metal stents have been greatly developed and applied. Metal stents can make up for the disadvantages of plastic support, which mainly includes three types, namely, partially covered self-expanding metal support, self-expanding metal support, and fully covered self-expanding metal support [17]. The

metal stent material has high finish and large lumen, which is not easy to adhere to bile sludge and bacteria, and will not contact with bacteria in a large area. The outer layer of the metal stent is covered by bile duct mucosal epithelial cells, which reduces the rate of bacterial infection and the probability of bile duct collateral obstruction and is convenient for placement [18]. This study compared the median survival time of patients in two groups, and the median survival time of the plastic stent group and the metal stent group was 8.15 and 11.83 months, respectively. The results showed that there was an inverse relationship between the stage and survival time. With the continuous progress of the disease, the survival time shows a downward trend, so it is necessary to choose an appropriate way for early treatment to improve the quality of life of patients as much as possible. Logistic multiple regression analysis revealed that tumor diameter ≥ 5 cm, portal vein infiltration, lymph node metastasis, and classification of hilar cholangiocarcinoma were all risk factors ($P < 0.05$), while the metal stent type was protective factor ($P < 0.05$), further confirming the clinical application value of metal stents. Plastic stents are inexpensive and easy to maneuver endoscopically; however, plastic stents have a small inner diameter and are often blocked within the stent due to biliary deposits and bacterial infection. Compared with the plastic stents, the metal stents have wider inner diameter and smoother surface. After placing plastic stents and metal stents in 100 patients with bile hilar cholangiocarcinoma, the median patency period of the metal stent group is 5.56 months, which is significantly higher than that of the plastic stent group of 1.86, and the reintervention rate of the metal stents is significantly higher than that of the plastic stents [19]. Metal stents or plastic stents can still be placed in the original stent even if the metal stent is blocked. However, once placed, metal stents cannot be removed and cost more than plastic stents. Therefore, metal stents should be considered for patients with expected survival beyond 4-6 months; otherwise, plastic stents can be used to maintain cost-effectiveness [20]. This study comprehensively analyzed the therapeutic effect of ERCP endoscopic implantation of stents with different materials on cholangiocarcinoma, which has good clinical significance. At the same time, small sample experiments have been done before this study, which could ensure the smooth progress of the study and provide a good foundation for the treatment of patients.

In conclusion, the placement of bile duct metal stent and plastic stent under ERCP plays an important role in the clinical treatment of patients with cholangiocarcinoma. ERCP placement of bile duct metal stent has higher success rate and better prognosis and can prolong the survival time of patients to a greater extent, but the price is relatively expensive. Therefore, it is necessary to comprehensively analyze the patient's economic affordability, expected survival time, stent drainage time, and personal needs and then select an appropriate treatment method. However, this present study still has some limitations. Research time and sample size both affect the research results. In addition, the subjective operation of the operator and the detailed medical history of the patient might induce the result deviation. Therefore, more qualified samples in the following study will be included and the study time and follow-up time will be appropriately extend to avoid the impact of subjective operation on the accuracy of results as much as possible.

Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

Conflicts of Interest

The author(s) declare(s) that they have no conflicts of interest.

References

[1] R. Paz-Fumagalli, J. Core, C. Padula et al., "Safety and initial efficacy of ablative radioembolization for the treatment of unresectable intrahepatic cholangiocarcinoma," *Oncotarget*, vol. 12, no. 20, pp. 2075–2088, 2021.

[2] Z. Fu, Y. Fan, C. Wu et al., "Clinical efficacy and mechanism for focused ultrasound (FUS) in the management of cervical intraepithelial neoplasia 1 (CIN1)," *International Journal of Hyperthermia*, vol. 37, no. 1, pp. 339–345, 2020.

[3] L. Monino, P. H. Deprez, and T. G. Moreels, "Percutaneous cholangioscopy with short Spyscope combined with endoscopic retrograde cholangiography in case of difficult intrahepatic bile duct stone," *Digestive Endoscopy*, vol. 33, no. 4, pp. e65–e66, 2021.

[4] T. Powles, M. Kockx, A. Rodriguez-Vida et al., "Clinical efficacy and biomarker analysis of neoadjuvant atezolizumab in operable urothelial carcinoma in the ABACUS trial," *Nature Medicine*, vol. 25, no. 11, pp. 1706–1714, 2019.

[5] K. Hnaris, C. Taylor, L. Hookey, and R. Bechara, "Fistulotomy-first approach is safe and effective in endoscopic retrograde cholangiography: results of a prospective case series," *Journal of the Canadian Association of Gastroenterology*, vol. 4, no. 4, pp. 173–178, 2021.

[6] R. Singh, M. Fischer, and S. Sagi, "S0690 clinical efficacy of vedolizumab in chronic pouchitis and CD of the pouch: a systematic review," *The American Journal of Gastroenterology*, vol. 115, no. 1, pp. S345–S346, 2020.

[7] M. Goetz, J. Fisch, J. Hetzel, and G. Grözinger, "Reverse rendezvous with endoscopic retrograde cholangiography and percutaneous transhepatic cholangio drainage: who meets whom?," *Endoscopy*, vol. 51, no. 3, pp. E47–E48, 2019.

[8] K. A. Morgan, E. H. Mann, A. R. Young, and C. M. Hawrylowicz, "ASTHMA-comparing the impact of vitamin DversusUVR on clinical and immune parameters," *Photochemical & Photobiological Sciences*, vol. 16, no. 3, pp. 399–410, 2017.

[9] F. Iida and J. Kusama, "Surgical evaluation of endoscopic retrograde cholangiography for biliary tract diseases," *The Japanese Journal of Surgery*, vol. 12, no. 4, pp. 257–261, 2019.

[10] B. Sarah, B. Giovanna, K. Emanuela, N. Nadi, V. Josè, and P. Alberto, "Clinical efficacy of the enzyme replacement therapy in patients with late-onset Pompe disease: a systematic review and a meta-analysis," *Journal of Neurology*, vol. 16, no. 15, pp. 1–9, 2021.

[11] A. R. Hatfield, A. Smithies, R. Wilkins, and A. J. Levi, "Assessment of endoscopic retrograde cholangio-pancreatography

(ERCP) and pure pancreatic juice cytology in patients with pancreatic disease," *Gut*, vol. 17, no. 1, pp. 14–21, 2019.

[12] N. Paracha, A. Reyes, V. Diéras, I. Krop, X. Pivot, and A. Urruticoechea, "Evaluating the clinical effectiveness and safety of various HER2-targeted regimens after prior taxane/trastuzumab in patients with previously treated, unresectable, or metastatic HER2-positive breast cancer: a systematic review and network meta-analysis," *Breast Cancer Research and Treatment*, vol. 180, no. 3, pp. 597–609, 2020.

[13] A. R. Sondhi, C. J. Sonnenday, N. D. Parikh, and R. Law, "EUS-guided gastrojejunal anastomosis to facilitate endoscopic retrograde cholangiography in a patient with a right lobe liver transplant and Roux-en-Y anatomy," *VideoGIE*, vol. 5, no. 10, pp. 473–475, 2020.

[14] L. W. Zhang, X. Fei, and Y. Song, "The clinical efficacy of novel vacuum suction ureteroscopic lithotripsy in the treatment of upper ureteral calculi," *World Journal of Urology*, vol. 39, no. 11, pp. 4261–4265, 2021.

[15] M. Sabbah, A. Nakhli, N. Bellil et al., "Predictors of failure of endoscopic retrograde pancreatocholangiography during common bile duct stones," *Heliyon*, vol. 6, no. 11, article e05515, 2020.

[16] S. Marta, G. Elena, and G. M. Pilar, "The pharmacology and clinical efficacy of matricaria recutita L.: a systematic review of in vitro, in vivo studies and clinical trials," *Food Reviews International*, vol. 2, no. 1, pp. 1–35, 2020.

[17] R. Nishio, H. Kawashima, M. Nakamura et al., "Double-balloon endoscopic retrograde cholangiopancreatography for patients who underwent liver operation: a retrospective study," *World Journal of Gastroenterology*, vol. 26, no. 10, pp. 1056–1066, 2020.

[18] A. Malli, C. Durkin, J. R. Groce, A. Hinton, D. L. Conwell, and S. G. Krishna, "Unavailability of endoscopic retrograde cholangiography adversely impacts hospital outcomes of acute biliary pancreatitis: a national survey and propensity-matched analysis[J]," *Pancreas*, vol. 49, no. 1, pp. 39–45, 2020.

[19] E. Elshimi and W. Morad, "Cost analysis of biliary drainage using metal versus plastic stents in hepatocellular carcinoma patients with obstructive jaundice," *Gastrointest Tumors*, vol. 7, no. 1-2, pp. 1–10, 2020.

[20] M. X. Xia, Y. L. Pan, X. B. Cai et al., "Comparison of endoscopic bilateral metal stent drainage with plastic stents in the palliation of unresectable hilar biliary malignant strictures: large multicenter study," *Digestive Endoscopy*, vol. 33, no. 1, pp. 179–189, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 2492488, 11 pages
https://doi.org/10.1155/2022/2492488

Research Article

The Association of Waist Circumference with the Prevalence and Survival of Digestive Tract Cancer in US Adults: A Population Study Based on Machine Learning Methods

Xingyu Jiang,[1] **Qi Liang**,[1] **Huanhuan Xu**,[1] **Shouyong Gu**,[2] **and Lingxiang Liu**[1]

[1]*Department of Oncology, The First Affiliated Hospital of Nanjing Medical University, 300 Guangzhou Road, Nanjing 210029, China*
[2]*Geriatric Institute, Jiangsu Province Geriatric Hospital, Nanjing Medical University Affiliated Geriatric Hospital, Nanjing 210029, China*

Correspondence should be addressed to Shouyong Gu; gushouyong@jspgh.com and Lingxiang Liu; llxlau@163.com

Received 5 August 2022; Accepted 21 September 2022; Published 6 October 2022

Academic Editor: Jincheng Wang

Aims. This paper aims to investigate the relationship of waist circumference (WC) with digestive tract cancer morbidity and mortality. *Methods*. Based on the data from a nationally representative US population survey, we summarized the prevalence of digestive tract cancer and all-cause mortality of cancer patients across WC quartiles. Adjusted logistic regression and restricted spline curve were used to analyze WC and the prevalence of digestive tract cancer. Moreover, Cox regression and the Kaplan-Meier curve were applied to investigate the association of WC with all-cause mortality. We also attempted to make a model to predict cancer happening. *Results*. This paper included a total of 34,041 participants, with digestive tract cancer observed in 265 (0.7%) individuals. WC was positively associated with digestive tract cancer morbidity after full adjustment of covariates (OR: 1.72 and 95% CI: 1.41-2.10). Also, individuals in the highest WC group had a higher risk of digestive tract cancer (Q4, OR: 2.71 and 95% CI: 1.48-5.00). Moreover, no significant association was observed in upper digestive cancer, and WC was associated with a longer survival time once diagnosed (hazard ratio (HR): 0.50 and 95% CI: 0.28-0.92). Finally, the model we made proved to be effective. *Conclusion*. High WC is a risk factor for digestive tract cancer with or without adjusting for body mass index, especially those located in the lower digestive tract. However, once digestive tract cancer has been diagnosed, patients with higher WC showed better survival outcomes. Moreover, machine learning methods can be used to predict digestive tract cancer risk in the future.

1. Introduction

Over the past decade, obesity has become a growing health threat worldwide, which is estimated to contribute to about 11.9% of cancer in males and 13.1% in females [1, 2]. Accumulating research had certificated that obese individuals were at an elevated of multiple digestive system cancer, including esophagus, gastric, colon, and rectal cancer [3–7].

Body mass index (BMI) is currently the foremost anthropometric index to evaluate the fat distribution of the body or the related health risk in most studies [8, 9]. However, obesity is a heterogeneous metabolic condition: Abdominal fat accumulation results in a more adverse obesity phenotype associated with worse metabolic profiles than subcutaneous accumulation. BMI alone is unable to capture the distribution of body fat or distinguish between adipose and muscle tissue [10]. A mildly elevated BMI (25-30 kg/m^2) was reported to improve the survival of certain digestive tract cancer [11, 12]. Therefore, BMI is insufficient to fully understand obesity-related digestive tract cancer risk.

Among the several body measures, WC strongly correlates with abdominal fat distribution [13], and self-measurement of WC can be obtained easily. A recent consensus statement emphasized the importance of WC in clinical practice, given the advantages in stratifying obesity-related health risks than BMI [10, 14]. There used to be

several studies revealing that WC had an adverse implication on all-cause death and cancer-related mortality [15–18]. Interestingly, after adjusting WC, BMI seemed to be a protective or neutral factor [10]. Instead, the clinical significance of WC can be fully demonstrated only after adjusting BMI [10, 19]. However, most evidence was from cardiovascular or metabolic diseases, and few studies investigated the association between WC and cancer prevalence with BMI adjusted for [20–24]. It remains unclear whether high WC could elevate the risk of digestive tract cancer.

Therefore, this article aims to analyze the relationship of WC with the prevalence and prognosis of digestive tract cancer.

2. Materials and Methods

2.1. Data Source. National Health and Nutrition Examination Survey (NHANES) is a publicly available database recording fit and nutrition conditions of adults and teenagers in the US (https://www.cdc.gov/nchs/nhanes/irba98.htm). The continuous NHANES data were collected in a 2-year cycle via face-to-face conversation and physical or laboratory examination. Our research used the 7 continuous NHANES data cycles, including (2001-2002, 2003-2004, 2005-2006, 2007-2008, 2009-2010, 2011-2012, and 2013-2014). Demographic characteristics, dietary, lifestyle factors, education levels, and medical conditions were collected via the conversation. Physical examinations, including weight, height, and WC, were measured in the mobile examination centers [25]. Moreover, the National Death Index (NDI) was used to investigate the association of all-cause mortality with digestive tract cancer patients. This database collects information about death or censoring by a medical examination (December 31, 2015).

All individuals aged ≥18 years with matched survival data were included, while those who were (1) without weight, height, or WC information ($n = 2326$) and (2) pregnant ($n = 1081$) were excluded from further analysis. NHANES followed the Health and Human Services policy and was approved by the National Center for Health Statistics. Sample participants were fully informed of the process and consented to participate in this survey.

2.2. Measurement of BMI and WC. BMI is calculated as the following equation: $BMI = weight\,(kilograms)/height\,(meters\,squared)$. Next, BMI was divided into groups following the categories of World Health Organization (WHO): <18.5 kg/m^2 is underweight, 18.5-24.9 kg/m^2 is normal weight, overweight is defined as 25.0-29.9 kg/m^2, class I obesity means BMI between 30.0 and 34.9 kg/m^2, class II obesity means 35.0 and 39.9 kg/m^2, and class III obesity means ≥40 kg/m^2. Waist circumference was measured at the end of normal expiration in a standing position. It was measured above the uppermost lateral border of the right ilium, to the nearest 0.1 cm with the tape snug but not compressing the skin. The classification standard of WHO is between 94 and 101.9 cm for men and 80.0 and 87.9 cm for women, based on which abdominal obesity was defined as a WC greater than 102 in men and 88 in women [10]. More details are recorded in the Anthropometry Procedures Manual of NHANES (https://www.cdc.gov/nchs/data/nhanes/nhanes_07_08/manual_an.pdf).

2.3. Definition of Digestive Tract Cancer. In the face-to-face conversation, participators were asked the following two questions: (1) "whether they were ever told that they had cancer or malignancy" and (2) "which kind of cancer or malignancy they suffered by the health professionals." In this paper, individuals with malignant tumors in the esophagus ($n = 17$, 0.04%), stomach ($n = 26$, 0.07%), colon ($n = 217$, 0.64%), and rectum ($n = 12$, 0.04%) were defined as patients with digestive tract cancer. Among them, small bowel cancer was not present in the dataset due to its low prevalence. Moreover, we stratified the digestive tract cancer into the upper and lower. Esophagus cancer and stomach cancer belong to the upper digestive tract cancer, and the rest belong to the other. These definitions were consistent with several epidemiological original types of research and one meta-analysis [26–28].

2.4. Covariates. Covariates were included by taking reference to the previous studies: age, gender (male or female), race (non- Hispanic Whites, non-Hispanic Blacks, Mexican Americans, other Hispanic, and other races), economic status (described as poverty-income ratio), and different education levels (not attended high school, high school, college, or above) [29, 30]. Diabetes history (diabetes, borderline diabetes, or nondiabetic) and whether they were smokers or alcoholics were collected using a health questionnaire during the conversation. Participants who smoked 100 cigarettes or above in their lifetime were recorded as smokers; participants who consumed 12 drinks or above per year were recorded as alcoholics [31]. Hemoglobin A1c (HbA1c) (mmol/L) and fasting plasma glucose (FPG) (mmol/L) were obtained in the laboratory, whereas the estimated glomerular filtration rate (eGFR) was calculated according to chronic kidney disease-epidemiology collaboration [32].

2.5. Statistical Analysis. First, we preprocessed the dataset via multiple imputations to fill missing values (except the outcome variable) to maximize statistical power and minimize deviation [33, 34]. We used the Kolmogorov-Smirnov test to test whether the continuous data is the normal distribution or skewed distribution. Normally distributed variables were presented as $mean \pm standard\,deviation$, whereas skewed distributed variables were presented as median with Q1-Q3. Categorical variables were presented as percentages. Characteristics between those with or without digestive tract cancer risk were compared by one-way ANOVA analysis (normally distributed variables), Kruskal-Wallis test (non-normal distributed variables), and chi-square test (categorical variables) as appropriate. Most cancer patients are elderly, so we separated the age as <80 and ≥80 to make the underlying association clearer. We used the adjusted logistic regression models to analyze the association between WC and digestive tract cancer risk, and the results were shown as OR with a 95% CI. In the minimally adjusted model, only old age, gender, smoking, drinking, and HbA1c

were in adjustment. The BMI was additionally adjusted for the fully adjusted model. WC was also analyzed as four categories grouped by interquartile range in the regression model, setting the lowest group (55.5 to 86.8 cm) as the reference. In contrast, we performed a similar logistic regression analysis on BMI and the prevalence of digestive tract cancer. Moreover, a restricted cubic spline with 5 knots (5%, 25%, 50%, 72.5%, and 95%) was used to illustrate the relationship between WC and digestive tract cancer. The median WC (97.0 cm) was set as a reference point according to the guidance [35]. We also modelled the underlying relationship between upper digestive and lower digestive using restricted cubic spline, respectively. Additionally, sensitivity analyses were employed to investigate this association in different subgroups, involving cancer subtypes (upper digestive and lower digestive), BMI (<25, 25-30, 30-35, and ≥35), WC (> median and ≤ median), and sex (male and female) categories. For a more in-depth study of the relationship between WC and digestive tract cancer, we further divided WC into <80 or ≥80 and divided WC according to the presence or absence of abdominal obesity (male: ≥102 and female: ≥88).

Furthermore, we adopted multivariate Cox regression analysis to assess the relationship of WC and BMI with all-cause mortality, and the associations were shown by hazard ratio (HR) with 95% CI. Apart from old age, gender, smoking, drinking, and HbA1c, WC and BMI were also mutually adjusted for. The impact of WC (≤ median and > median) on overall survival for those with digestive tract cancer was illustrated by a Kaplan-Meier curve. We also conducted the same curve by dividing WC into abdominal obesity or not. Statistical significance was defined as a value of $P < 0.05$. R software performed all statistical analyses (version 4.1.2; binary for macOS 11, Big Sur).

2.6. Prediction Model. Moreover, on the basis of the analysis we conducted above, a prediction model was accomplished to evaluate the risk of digestive tract cancer. The predictive factors were identified in the logistic regression analysis or in line with references, including waist circumference, body mass index, age, race, gender, education, PIR levels (poverty-income ratio), FPG, smoking, drinking, diabetes, and HbA1c. After eliminating missing values for each variable, 476 samples were eligible. 70% of the samples were used in constructing the model, and the other 30% group served as a validation set. The area under the receiver operating characteristic curve (AUC-ROC) was used to assess the discrimination ability of the predictive model.

3. Results

3.1. Characteristics of Baseline Data. Of the 37448 individuals aged between ≥18 in the NHANES population, 2326 were excluded due to the lack of a valid BMI or WC and 1081 due to pregnancy. This study eventually involved 34,041 participants, and 265 (0.7%) individuals were diagnosed with digestive tract cancer (Figure 1). Among participants, 50% were female, the median age at inclusion was 46 years (interquartile range, IQR 33–63), and the median WC was 97 cm (IQR 86.8–108). Moreover, the number of participants with abdominal obesity based on WC (male: ≥102 and female: ≥88) was 18359 (53.9%), with 164 of which had digestive tract cancer; while 15658 participants did not have abdominal obesity, the cancer case in which was only 77, showing that abdominal obesity was associated with digestive tract cancer risk (chi-square test, $P < 0.01$). After a median follow-up of 67 months, 97 cases of all-cause death were observed in those with digestive tract cancer, including 31 who died because of cancer-related diseases.

Table 1 summarizes the baseline characteristics of all participants by WC qualities. Compared with the low WC group, the high WC group was older, more male, and less educated. Digestive tract cancer was more observed in the participators with a high WC. Participants with a high WC also had a more possibility to have elevated HbA1c, high FPG, high eGFR, more self-reported diabetes, and more self-reported diseases in the cardiovascular system. Besides, more current smokers and alcoholics were in the low WC and normal WC category than those in the high WC category.

3.2. Association of WC with Digestive Tract Cancer. WC was significantly correlated with the risk of digestive tract cancer (Table 2). In the non-adjusted, minimally adjusted, or fully adjusted models, the ORs with 95% CI were 1.15 (1.07-1.23), 1.13 (1.04-1.22), and 1.72 (1.41-2.10), respectively. After adjusting fully for old age, gender, smoking, drinking, HbA1c, and BMI, individuals with high WC (Q4) had a 2.71-fold increased risk of digestive tract cancer compared to the lowest quartile (Q1). In Figure 2(a), the restricted cubic spline shows a consistent and significant positive association between WC and digestive tract cancer morbidity. However, no significant association was observed between WC and upper digestive cancer (Figure 2(b)). In contrast, lower digestive cancer remained significantly associated with WC (Figure 2(c)).

3.3. Association of BMI with Digestive Tract Cancer. In the non-adjusted and minimally adjusted models, there was no significant association between BMI and digestive tract cancer morbidity (OR: 1.01, 95% CI = 0.92-1.10; OR: 1.05; and 95% CI = 0.95-1.15). However, after fully adjusting the model, the BMI was found to be a protective factor for digestive tract cancer (OR: 0.57 and 95% CI: 0.44-0.73). Moreover, when calculated as a categorical valuable, the OR for BMI were 0.60 (0.41-0.87), 0.47 (0.28-0.79), 0.26 (0.12-0.55), and 0.17 (0.06-0.47) in the overweight, class I obesity, class II obesity, and class III obesity groups. This association was not found in the patients under normal weight (OR: 1.34 and 95% CI = 0.40-3.37) (Table 3).

3.4. Sensitivity Analysis. Figure 3 shows the robust association across cancer subtype (lower digestive), BMI (<25, 25-30, 30-35, and ≥35 kg/m^2), WC (above median: >97 cm; and below median: ≤97 cm), and sex (male and female). Importantly, when the cut-off WC was set smaller relatively, although a consistent trend remained, WC was not associated with digestive tract cancer risk significantly in the

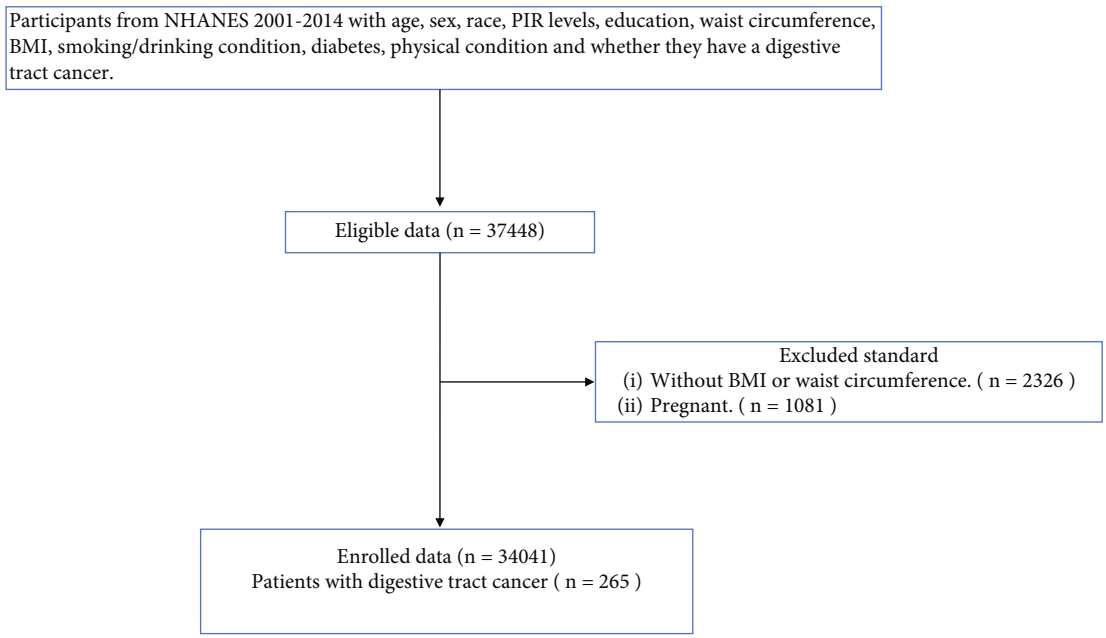


Figure 1: Flow chart of selection of eligible participants. NHANES: National Health and Nutrition Examination Survey.

subgroup of WC <80 cm (OR: 1.14 and 95% CI: 0.68-1.91) and participants without abdominal obesity (men: <102 and women: <88) (OR: 1.22 and 95% CI: 0.83-1.80) (Figure S1).

3.5. Survival Analysis. After adjusting the BMI, WC was associated with a decrement in all-cause mortality (HR: 0.50 and 95% CI: 0.28-0.92). On the contrary, BMI did not show a statistical association once adjusting for the WC (HR: 0.45 and 95% CI: 0.13-1.58) (Table 4). The Kaplan-Meier curve revealed that the survival time was longer among higher WC categories ($P = 0.0097$). Patients whose WC was below or equal to the median (≤97 cm) showed statistically poorer 5-year survival compared with those with a WC above the median (>97 cm) (Figure 4). More importantly, the same conclusion is still met when dividing the WC value according to whether it is consistent with abdominal obesity (abdominal obesity, male: ≥102, and female: ≥88) (Figure S2).

3.6. Prediction Model. As to the model we build, the AUC-ROC of the multivariate logistic regression model to predict the probability of digestive tract cancer was 0.71 (95% CI and 0.46-0.94). Moreover, the specificity of the proposed nomogram was 0.708, the sensitivity was 0.750, and the accuracy was 0.972 (Figure 5).

4. Discussion

Accumulating evidence indicates that fat distribution is a primary cause of obesity heterogeneity [36, 37], and abdominal obesity has been recognized as a more serious health problem worldwide, surpassing even obesity defined by BMI [38, 39]. Studies showed that the mean WC of China increased to a greater extent among men and women separately after an adjustment of BMI over 1993-2011 [40]. A similar trend was also observed in the US, England, Mexico, and Canada [40, 41]. Currently, WC is adopted more frequently than BMI or waist-to-hip ratio as the preferred body metric for assessing abdominal fat accumulation, suggesting a more robust association with absolute visceral fat mass [10, 42]. Visceral fat is the underlying culprit for health problems but can only be measured by using expensive instruments directly [43]. In summary, abdominal obesity due to WC significantly increased the adverse health risk, regardless of BMI adjustment.

However, the association between WC and digestive tract cancer now is still vague, and only a few researchers focus on the relationship between WC and digestive tract cancer risk. A recent cross-sectional study reported high WC associated with increased colorectal cancer incidents based on 63057 South Korean population with normal weight [14]. The sample size and selection bias (normal-weight individuals from the health checkup program) made it difficult to generalizable to populations in different settings. In another meta-analysis of prospective studies, Du et al. analyzed the association of WC with total gastroesophageal cancer. It was reported that WC was associated with gastric cancer and esophageal cancer (for gastric cancer, relative risk, $RR = 1.48$ and $95\%CI = 1.20 - 1.83$; for esophageal cancer, $RR = 2.13$ and $95\%CI = 1.07 - 4.22$), while waist-to-hip ratio associated with gastric cancer only (for gastric cancer, $RR = 1.40$ and $95\%CI = 1.08 - 1.82$; for esophageal cancer, $RR = 2.30$ and $95\%CI = 0.86 - 6.17$) [24]. Similarly, Dong et al. [23] performed a meta-analysis on 134,356,0 participants to clarify the association between abdominal obesity and the incidence of colorectal cancer; the relative risk for total colorectal cancer was of more significance in those with greater WC, compared to the low category of WC ($RR = 1.42$ and $95\%CI = 1.30 - 1.55$) [23].

Table 1: Baseline characteristics divided by quartile of waist circumference.

	Q1 (55.5, 86.8)	Q2 (86.8, 97)	Q3 (97, 108)	Q4 (108, 178)	*P*
N	8520	8591	8443	8487	
Digestive tract cancer (%)	45 (0.5%)	64 (0.7%)	70 (0.8%)	86 (1.0%)	0.004
Age (years)	37.0 (25.0, 53.0)	47.0 (34.0, 62.0)	53.0 (39.0, 65.0)	53.0 (39.0, 65.0)	<0.001
Gender (female, %)	5312 (62.3%)	4263 (49.6%)	3690 (43.7%)	3787 (44.6%)	<0.001
Race (%)					<0.001
Non-Hispanic White	3813 (52.1%)	3735 (47.7%)	3988 (49.8%)	4306 (52.4%)	
Non-Hispanic Black	1758 (24.0%)	1661 (21.2%)	1669 (20.8%)	2073 (25.2%)	
Mexican American	1116 (15.2%)	1672 (21.3%)	1685 (21.0%)	1294 (15.8%)	
Other Hispanic	632 (8.6%)	770 (9.8%)	664 (8.3%)	542 (6.6%)	
PIR level (%)					<0.001
<1.33	2550 (29.9%)	2433 (28.3%)	2401 (28.4%)	2627 (31.0%)	
1.33-3.50	2751 (32.3%)	2938 (34.2%)	2895 (34.3%)	2942 (34.7%)	
≥3.50	3219 (37.8%)	3220 (37.5%)	3147 (37.3%)	2918 (34.4%)	
Education (%)					<0.001
Below high school	1906 (22.4%)	2435 (28.4%)	2451 (29.0%)	2349 (27.7%)	
High school	1841 (21.6%)	1914 (22.3%)	2044 (24.2%)	2140 (25.2%)	
Above high school	4760 (56.0%)	4233 (49.3%)	3943 (46.7%)	3995 (47.1%)	
Waist circumference (cm)	80.1 (75.6, 83.6)	92.2 (89.6, 94.6)	102.0 (99.4, 104.8)	116.3 (111.4, 124.3)	<0.001
BMI (kg/m^2)	22.2 (20.4, 23.9)	26.1 (24.4, 27.8)	29.3 (27.5, 31.4)	35.4 (32.3, 39.7)	<0.001
Esophagus cancer	4	4	6	3	0.765
Stomach cancer	9	7	6	4	0.579
Colon cancer	29	53	57	78	<0.001
Rectal cancer	4	1	3	4	0.566
HbA1c (mmol/L)	5.3 (5.1, 5.5)	5.4 (5.2, 5.7)	5.5 (5.3, 5.9)	5.7 (5.4, 6.2)	<0.001
FPG (mmol/L)	87.0 (81.0, 94.0)	91.0 (85.0, 100.0)	94.0 (87.0, 105.0)	98.0 (89.0, 115.0)	<0.001
Smoking (yes, %)	3508 (41.2%)	3807 (44.3%)	4059 (48.1%)	4352 (51.3%)	<0.001
Drinking (yes, %)	1087 (12.8%)	1099 (12.8%)	1086 (12.9%)	1301 (15.3%)	<0.001
Diabetes (yes, %)	413 (4.8%)	985 (11.5%)	1561 (18.5%)	2574 (30.3%)	<0.001
eGFR (ml/min/1.73m^2)	95.7 [76.2, 114.3]	102.8 [78.1, 126.0]	108.4 [83.4, 137.1]	133.7 [100.8, 172.5]	<0.001

BMI: body mass index; HbA1c: hemoglobin A1c; FPG: fasting plasma glucose; eGFR: estimated glomerular filtration rate.

Table 2: Association of waist circumference with digestive cancer using logistic regression models.

Items	Non-adjusted model		Minimally adjusted model		Fully adjusted model	
	Odds ratio	*P*	Odds ratio	*P*	Odds ratio	*P*
Waist circumference (per 10 cm)	1.15 (1.07-1.23)	<0.001	1.13 (1.04-1.22)	0.002	1.72 (1.41-2.10)	<0.001
Categories						
Q1 (55.5,86.8)	Reference		Reference		Reference	
Q2 (86.8,97)	1.41 (0.97-2.08)	0.076	1.26 (0.86-1.87)	0.239	1.47 (0.98-2.25)	0.082
Q3 (97,108)	1.57 (1.09-2.31)	0.018	1.30 (0.89-1.92)	0.182	1.72 (1.08-2.77)	0.025
Q4 (108,178)	1.93 (1.35-2.79)	<0.001	1.65 (1.14-2.41)	0.009	2.71 (1.48-5.00)	0.001

Minimally adjust model: adjusted for age (<80 and ≥80), gender, smoking, drinking, and HbA1c. Fully adjust model: adjusted for age (<80 and ≥80), gender, smoking, drinking, HbA1c, and BMI.

However, few of the included studies in these two meta-analyses studies conducted further adjustments between WC and BMI to clarify their independent role. A consensus statement clearly stated that the robustness of WC and all-cause morbidity can be fully recognized only after adjustment for BMI [10]. Besides, these two meta-analyses did not provide a thorough review of digestive tract cancer.

Consistent with previous studies, we revealed that WC was positively correlated with digestive tract cancer. More importantly, our result ulteriorly showed that the adverse

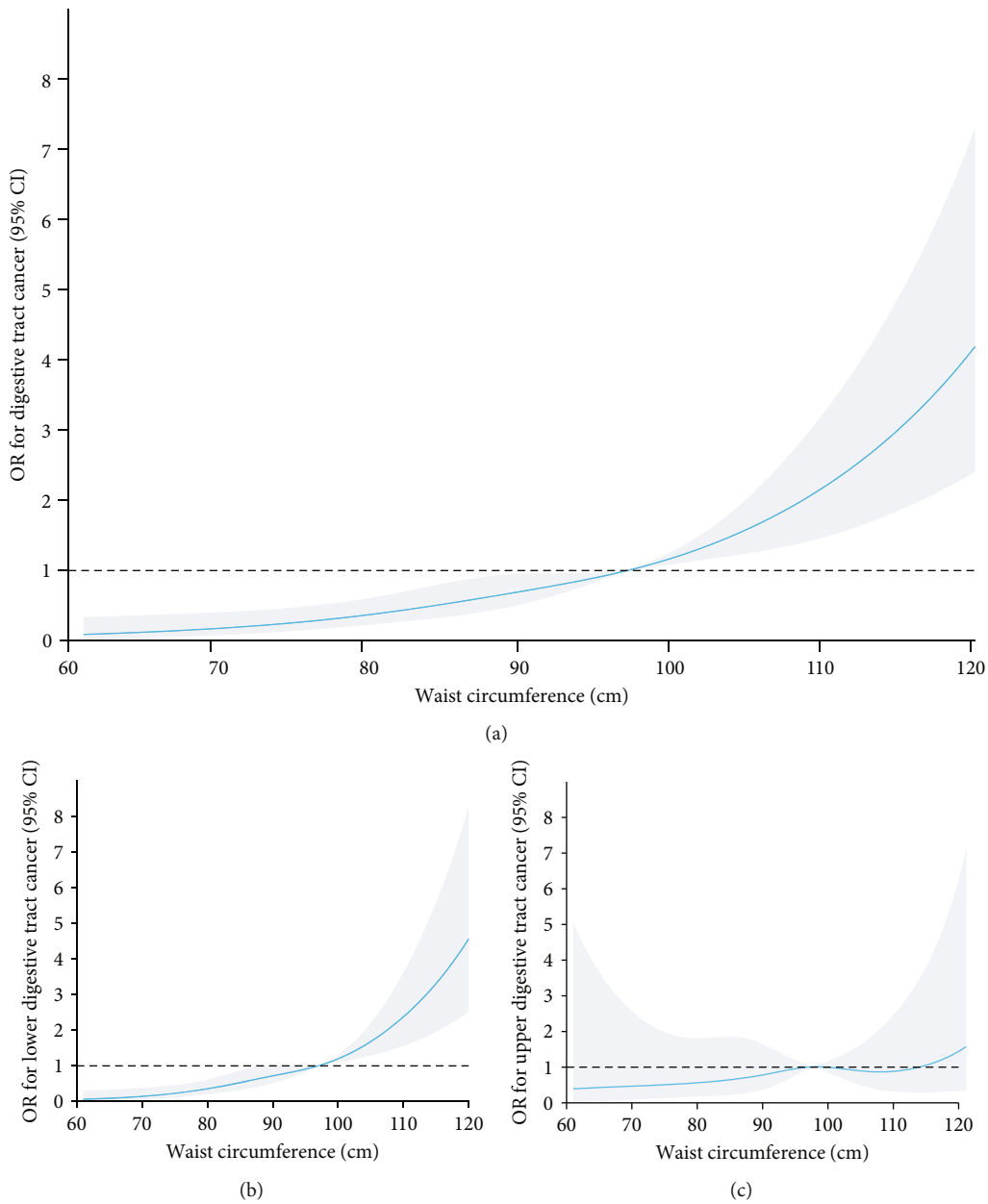


Figure 2: (a) Restricted cubic spline plots of the association between waist circumference and digestive tract cancer. The association was adjusted for age (<80 and ≥80), gender, smoking, drinking, HbA1c, and BMI. The median of the waist circumference was set as the reference for this figure. (b) Restricted cubic spline plots of the association between waist circumference and lower digestive tract cancer. (c) Restricted cubic spline plots of the association between waist circumference and upper digestive tract cancer. BMI: body mass index; CI: Confidence interval.

effects of WC persisted even after the adjustment of BMI. Therefore, WC is a strong anthropometric cancer biomarker under any total weight. Those with similar BMI but higher WC represent a phenotype with increased deposition of visceral adipose tissue, rather than the subcutaneous adipose tissue beneath the skin [44]. Visceral adipose tissue is now considered metabolic tissue, playing an important role in immunological, metabolic, and endocrine functions. The pro-inflammatory cytokines (such as tumor necrosis factor-α, interleukin-6, and interleukin-1β) from visceral adipose tissue contribute to a chronic inflammatory state in the whole body, thus creating a general environment better suitable for tumor growth [45]. As a measurement, WC can be an alternative to visceral adipose tissue [46]. A recent study compared five indicators for measuring visceral fat tissue, demonstrating that WC was plausible to reflect visceral fat accumulation. Unlike WC, BMI is only a weak support measurement of visceral adipose tissue [47]. A preceding study was convinced that abdominal obesity could predict advanced cancer better than BMI [48]. Besides, the WC threshold is currently designed to replace BMI as a proxy for the anthropometric index of obesity without considering its unique advantages in estimating cancer risk. Our findings offered the possibility of establishing a new WC grading

Table 3: Association of BMI with digestive cancer using logistic regression.

Items	Non-adjusted model		Minimally adjusted model		Fully adjusted model	
	Odds ratio	*P*	Odds ratio	*P*	Odds ratio	*P*
BMI (per 5 kg/m^2)	1.01 (0.92-1.10)	0.835	1.05 (0.95-1.15)	0.359	0.57 (0.44-0.73)	<0.001
BMI categories						
Normal weight	Reference		Reference		Reference	
Underweight	0.87 (0.26-2.10)	0.785	0.79 (0.24-1.92)	0.646	1.34 (0.40-3.37)	0.581
Overweight	1.01 (0.74-1.37)	0.963	0.98 (0.72-1.34)	0.892	0.60 (0.41-0.87)	0.007
Class I obesity	1.14 (0.81-1.60)	0.438	1.17 (0.83-1.66)	0.362	0.47 (0.28-0.79)	0.004
Class II obesity	0.89 (0.53-1.43)	0.642	0.99 (0.58-1.60)	0.961	0.26 (0.12-0.55)	<0.001
Class III obesity	1.03 (0.58-1.71)	0.926	1.22 (0.68-2.08)	0.476	0.17 (0.06-0.47)	<0.001

Minimally adjust model: adjusted for age (<80 and ≥80), gender, smoking, drinking, and HbA1c. Fully adjust model: adjusted for age (<80 and ≥80), gender, smoking, drinking, HbA1c, and waist circumference.

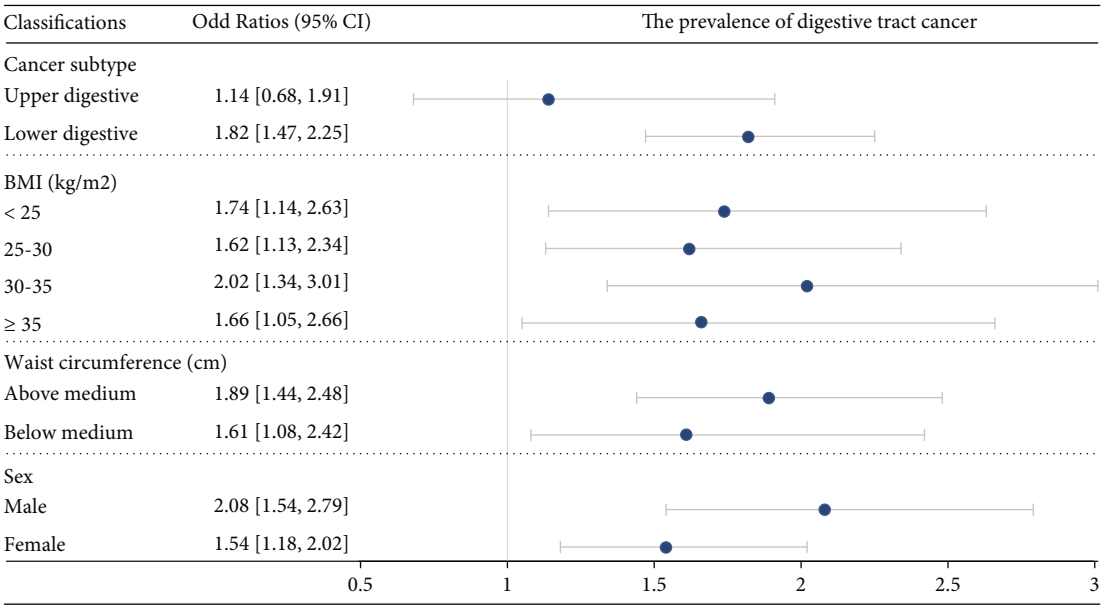


Figure 3: Sensitivity analysis on the association between waist circumference and digestive tract cancer risk based on logistic regression analysis. The association was adjusted for BMI, age, gender, smoking, drinking, and HbA1c. CI: confidence interval; BMI: body mass index.

Table 4: Association of waist circumference with all-cause mortality of digestive cancer using cox regression model.

Items	Hazard ratio (95% CI)	*P* value
Waist circumference (per 10 cm)	0.50 (0.28-0.92)	0.025
Body mass index (kg/m^2)	0.45 (0.13-1.58)	0.211

Cox regression model adjusted for age (<80 and ≥80), sex, smoking condition, drinking condition, HbA1c, and interaction between waist circumference and BMI.

scale consistent with BMI, from the perspective of cancer development. Further studies are expected to make this vision a reality.

Also, subgroup analysis in our study showed that WC and the upper digestive tract cancer risk association were not statistically significant. Refined restricted cubic splines of our study showed the same result. The upper digestive tract difference with WC has been reported in previous studies [24, 49, 50]. However, a recent study based on the UK Biobank database conducted a more detailed survey and came to a different result. It reported that WC associated with morbidity of esophageal adenocarcinoma (highest vs lowest category: $HR = 2.30$ and $95\% CI = 1.47 - 3.57$) and with gastric cardia cancer only in men ($HR = 2.21$ and $95\% CI = 1.27 - 3.84$). Besides, it reported that there was no statistically significant association of WC with other tissue types of upper digestive tract cancer, which is in line with the result our article observed [51], reminding us that the histological type of cancer may have an impact on causality.

Moreover, we found that when WC was set to a lower cut-off value in sensitivity analysis, the association between WC and digestive tract cancer became vague. Although these relationships were still approximately positive, their 95% CI continue to widen with the decreasing of cut-off values (Figure 3, Figure s1). Also, Wei et al. conducted a

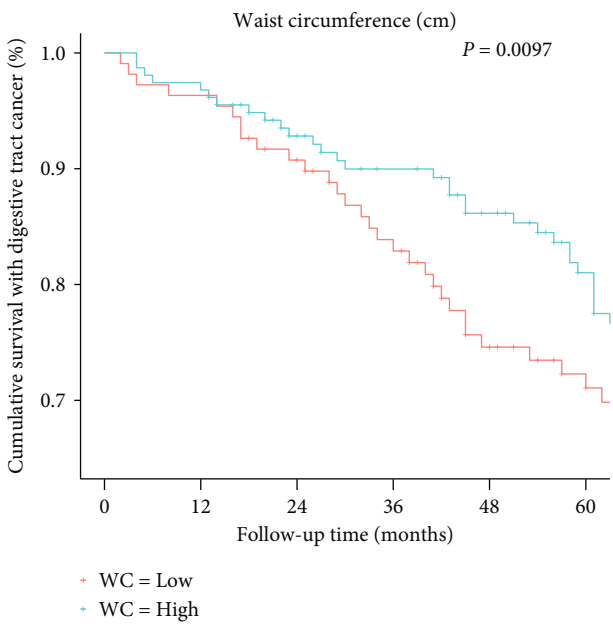


Figure 4: A Kaplan-Meier curve of the association between waist circumference and all-cause mortality of digestive tract cancer in the following 60 months (5 years). The waist circumference was divided by median into two groups (≤97 cm and >97 cm), and the survival comparison among groups was adjusted by the Bonferroni-Holm method. WC: waist circumference.

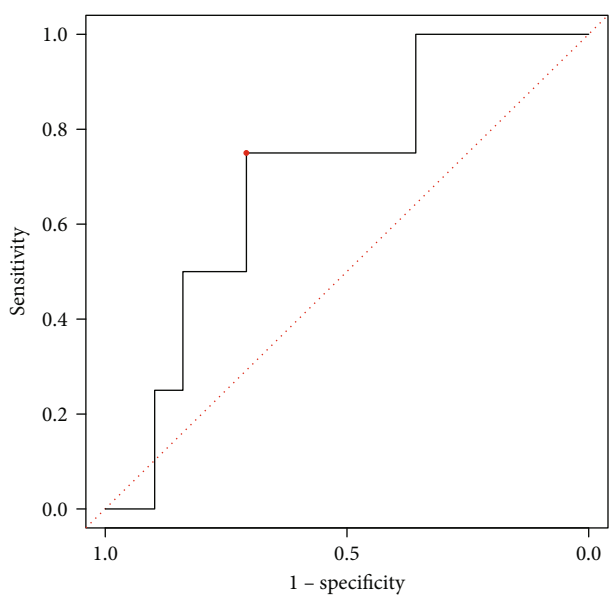


Figure 5: The receiver operating characteristic curve of the identification ability of the proposed nomogram. The nomogram showed relatively strong identification ability with an area under the curve of 0.71 (95% confidence interval, 0.46-0.94), specificity of 0.708, sensitivity of 0.750, and accuracy of 0.972.

cohort study on 104,825 males in China to clarify the association between WC and primary liver cancer. After a median of 8.9 years of follow-up, they found a U-shaped association between WC and the prevalence of cancer using a restricted cubic spline model (P-non-linear = 0.017) [52]. This article reminds us that when the WC is in a low range, the result comes differently. However, all participants in our study came from an American database. Different from the Asian populations, the overall distribution of WC in our study was mostly concentrated in a large value (Figure S3). It is difficult to analyze these changes at lower WCs due to the limited sample in NHANES.

Interestingly, once digestive tract cancer had developed, WC became a protective factor against all-cause mortality in patients ($HR = 0.50$ and $95\%CI = 0.28 - 0.92$). Several studies had similar findings. For example, Lo et al. [53] found that BMI (OR : 0.95 and $95\%CI = 0.87$ to 1.03) and elevated WC (OR : 0.82 and $95\%CI = 0.67$ to 0.99) had an inverse association with cancer-caused mortality. A Cox regression model on 3976 African-American participants in the Jackson Heart Study (JHS) showed a J-shape relationship between WC and overall mortality, after adjusting age, sex, and smoking [54]. Besides, a study investigating the obesity paradox showed that overweight and class I obese (BMI 25-35 kg/m^2) patients always have a low risk of all-cause mortality after cancer has been diagnosed [55]. However, a recent meta-analysis study found that in women, obesity was associated with higher all-cause and cancer-related mortality in breast cancer [56]. It is of great difference from the findings summarized above, where a high WC appears to be of survival benefit. In summary, during a subset rather than the general population, WC was associated with the decline of all-cause mortality in patients. This result is partially consistent with our research. As a cross-sectional analysis and limited by the NHANES database, we only accumulate data from participants at a certain point in time, ignoring the reverse causation that might exist between WC and the length of survival. Since individuals with cancer have more possibility of a long-time weight loss and decrement of the WC, this limitation may lead to a difference in findings.

Finally, as an extension of previous work, we build a model to evaluate the risk of digestive tract cancers, showing that machine learning algorithms might be a potential tool to predict the occurrence of digestive tract cancers universally, with the factors like waist circumference.

The advantages of this study should be highlighted. Firstly, the database came from the NHANES, a large, constantly updated database. There had a lot of studies using this database to analyze questions, involving cardiovascular diseases, cancers, and metabolic diseases [57–59]. Secondly, previous articles tended to prefer analyzing single cancer and physical indicators such as BMI or WC [60–62]. This article summarized the relationship between cancers that were prone to occur in the entire digestive tract and WC. Thirdly, the association between WC and digestive tract cancer was investigated after adjusting the BMI. This result suggests that WC, independent of BMI, can be a more relevant anthropometric factor than BMI per se. Last but not the least, we made a model to predict cancer happening, showing that machine learning is a tendency in the future to predict digestive tract cancer. There are still some limitations in this study. First, chronic diseases, such as cancer, can cause

weight loss [63]. However, due to insufficient data, this article failed to consider the loss of weight caused by long-term chronic diseases. Second, despite the huge size of the NHANES, the amount of upper digestive tract cancer patients in this study was quite small (esophagus cancer: 17, gastric cancer: 26, and total: 43), which may skew the results of this study. The future direction of research will expand the proportion of cancer cases in the total sample size and investigate the relationship between cancer histological types and high WC. Also, most of the cancer data in this article came from the self-reports of patients. There was a possibility that the patient himself was not informed of the condition by the doctor out of humanitarianism. Thirdly, the cross-sectional study of cancer was difficult to judge the causal link. Cohort studies may have a good display of causality; a population-based cohort study that investigated colorectal cancer strongly demonstrated the causality between WC and colorectal cancer [64]. In the next phase of the task, we suggest conducting a cohort study of WC and cancer with a larger database. Finally, previous studies showed that male and female obesity differed in physiology and indicators [65, 66]. Due to the limitation of sample size, this article had not investigated the difference in cancer risk of gender exhaustively and only accrued a subgroup analysis of gender in the sensitivity analysis. Further research is expected. Last but not least, the role of ageing in the association between WC and digestive tract cancer should be further investigated in the following research.

5. Conclusions

Measuring WC provides an additional opportunity to improve the estimation of digestive tract cancer, especially cancer located in the lower digestive tract. However, patients with higher WC showed better survival once digestive tract cancer have developed. Further studies should reveal the association between WC and digestive tract cancer and use machine learning methods to predict cancer happening.

Data Availability

The data used to support the findings of this study comes from National Health and Nutrition Examination Survey (NHANES), a publicly available database in the U.S (https://www.cdc.gov/nchs/nhanes/irba98.htm).

Conflicts of Interest

The authors declared no conflict of interest.

Authors' Contributions

Xingyu Jiang, Qi Liang, Shouyong Gu, and Lingxiang Liu conceived and designed the study. Xingyu Jiang, Qi Liang, and Huanhuan Xu analyzed the data. Xingyu Jiang and Qi Liang wrote the paper. All authors provided critical revisions of the manuscript and approved the final manuscript. Xingyu Jiang and Qi Liang contributed equally to this work and Lingxiang Liu is the primary corresponding author..

Acknowledgments

The authors sincerely acknowledge the NHANES research group and all the people who contributed to the NHANES data we studied. This study was supported by the National Natural Science Foundation of China (81472782) and the National Key Research and Development Program: the key technology of palliative care and nursing for cancer patients (2017YFC1309201). Grant sponsor: Yili Institute of Clinical Medicine Grant ID: yl2021ms0.

Supplementary Materials

Supplementary 1. Figure S1. A supplemental sensitivity analysis. The WC was subgrouped into <80 or ≥80 and was subgrouped according to the presence or absence of abdominal obesity (male: ≥102, female: ≥88).

Supplementary 2. Figure S2. A supplemental Kaplan-Meier curve of the association between waist circumference and all-cause mortality of digestive tract cancer in the following 60 months (5 years). The waist circumference was divided by having abdominal obesity or not (male <102 cm, female <88 cm and male ≥102 cm, female ≥88 cm).

Supplementary 3. Figure S3. The overall distribution of waist circumference.

References

[1] C. Andolfi and P. M. Fisichella, "Epidemiology of obesity and associated comorbidities," *Journal of Laparoendoscopic & Advanced Surgical Techniques. Part A*, vol. 28, no. 8, pp. 919–924, 2018.

[2] K. I. Avgerinos, N. Spyrou, C. S. Mantzoros, and M. Dalamaga, "Obesity and cancer risk: emerging biological mechanisms and perspectives," *Metabolism*, vol. 92, pp. 121–135, 2019.

[3] B. R. Alsop and P. Sharma, "Esophageal cancer," *Gastroenterology Clinics of North America*, vol. 45, no. 3, pp. 399–412, 2016.

[4] J. Garai, R. B. Uddo, M. C. Mohler et al., "At the crossroad between obesity and gastric cancer," *Methods in Molecular Biology*, vol. 1238, pp. 689–707, 2015.

[5] M. Bardou, A. N. Barkun, and M. Martel, "Obesity and colorectal cancer," *Gut*, vol. 62, no. 6, pp. 933–947, 2013.

[6] K. R. Lee, M. H. Seo, K. Do Han, J. Jung, and I. C. Hwang, "Waist circumference and risk of 23 site-specific cancers: a population-based cohort study of Korean adults," *British Journal of Cancer*, vol. 119, no. 8, pp. 1018–1027, 2018.

[7] N. Murphy, A. J. Cross, M. Abubakar et al., "A nested case-control study of metabolically defined body size phenotypes and risk of colorectal cancer in the European prospective investigation into cancer and nutrition (EPIC)," *PLoS Medicine*, vol. 13, no. 4, article e1001988, 2016.

[8] M. Recalde, V. Davila-Batista, Y. Díaz et al., "Body mass index and waist circumference in relation to the risk of 26 types of cancer: a prospective cohort study of 3.5 million adults in Spain," *BMC Medicine*, vol. 19, no. 1, p. 10, 2021.

[9] J. Han, Y. Zhou, Y. Zheng et al., "Positive effect of higher adult body mass index on overall survival of digestive system cancers except pancreatic cancer: a systematic review and meta-

analysis," *BioMed Research International*, vol. 2017, Article ID 1049602, 15 pages, 2017.

[10] R. Ross, I. J. Neeland, S. Yamashita et al., "Waist circumference as a vital sign in clinical practice: a consensus statement from the IAS and ICCR working group on visceral obesity," *Nature Reviews. Endocrinology*, vol. 16, no. 3, pp. 177–189, 2020.

[11] H. Lennon, M. Sperrin, E. Badrick, and A. G. Renehan, "The obesity paradox in cancer: a review," *Current Oncology Reports*, vol. 18, no. 9, p. 56, 2016.

[12] E. M. Cespedes Feliciano, C. H. Kroenke, and B. J. Caan, "The obesity paradox in cancer: how important is muscle?," *Annual Review of Nutrition*, vol. 38, no. 1, pp. 357–379, 2018.

[13] H. Fang, E. Berg, X. Cheng, and W. Shen, "How to best assess abdominal obesity," *Current Opinion in Clinical Nutrition and Metabolic Care*, vol. 21, no. 5, pp. 360–365, 2018.

[14] Y. S. Jung, N. H. Kim, H. J. Yang et al., "Association between waist circumference and risk of colorectal neoplasia in normal-weight adults," *Journal of Gastroenterology and Hepatology*, vol. 35, no. 1, pp. 43–49, 2020.

[15] D. F. Quail and A. J. Dannenberg, "The obese adipose tissue microenvironment in cancer development and progression," *Nature Reviews. Endocrinology*, vol. 15, no. 3, pp. 139–154, 2019.

[16] T. Pischon, H. Boeing, K. Hoffmann et al., "General and abdominal adiposity and risk of death in Europe," *The New England Journal of Medicine*, vol. 359, no. 20, pp. 2105–2120, 2008.

[17] C. Zhang, K. M. Rexrode, R. M. van Dam, T. Y. Li, and F. B. Hu, "Abdominal obesity and the risk of all-cause, cardiovascular, and cancer mortality," *Circulation*, vol. 117, no. 13, pp. 1658–1667, 2008.

[18] J. R. Cerhan, S. C. Moore, E. J. Jacobs et al., "A pooled analysis of waist circumference and mortality in 650,000 adults," *Mayo Clinic Proceedings*, vol. 89, no. 3, pp. 335–345, 2014.

[19] M. G. O'Doherty, N. D. Freedman, A. R. Hollenbeck, A. Schatzkin, and C. C. Abnet, "A prospective cohort study of obesity and risk of oesophageal and gastric adenocarcinoma in the NIH-AARP diet and health study," *Gut*, vol. 61, no. 9, pp. 1261–1268, 2012.

[20] R. Huxley, S. Mendis, E. Zheleznyakov, S. Reddy, and J. Chan, "Body mass index, waist circumference and waist:hip ratio as predictors of cardiovascular risk—a review of the literature," *European Journal of Clinical Nutrition*, vol. 64, no. 1, pp. 16–22, 2010.

[21] Y. J. Lee, S. K. Myung, B. Cho et al., "Adiposity and the risk of colorectal adenomatous polyps: a meta-analysis," *Cancer Causes & Control*, vol. 22, no. 7, pp. 1021–1035, 2011.

[22] Y. Ma, Y. Yang, F. Wang et al., "Obesity and risk of colorectal cancer: a systematic review of prospective studies," *PLoS One*, vol. 8, no. 1, article e53916, 2013.

[23] Y. Dong, J. Zhou, Y. Zhu et al., "Abdominal obesity and colorectal cancer risk: systematic review and meta-analysis of prospective studies," *Bioscience Reports*, vol. 37, no. 6, 2017.

[24] X. Du, K. Hidayat, and B. M. Shi, "Abdominal obesity and gastroesophageal cancer risk: systematic review and meta-analysis of prospective studies," *Bioscience Reports*, vol. 37, no. 3, 2017.

[25] G. Zipf, M. Chiappa, K. S. Porter, Y. Ostchega, B. G. Lewis, and J. Dostal, *National Health and nutrition examination, survey: plan and operations NCfHS*, 2013.

[26] X. F. Zhang, X. K. Wang, Y. J. Tang et al., "Association of whole grains intake and the risk of digestive tract cancer: a systematic review and meta-analysis," *Nutrition Journal*, vol. 19, no. 1, p. 52, 2020.

[27] F. Islami, A. Goding Sauer, K. D. Miller et al., "Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 31–54, 2018.

[28] P. Anand, A. B. Kunnumakara, C. Sundaram et al., "Cancer is a preventable disease that requires major lifestyle changes," *Pharmaceutical Research*, vol. 25, no. 9, pp. 2097–2116, 2008.

[29] E. Oliveros, V. K. Somers, O. Sochor, K. Goel, and F. Lopez-Jimenez, "The concept of normal weight obesity," *Progress in Cardiovascular Diseases*, vol. 56, no. 4, pp. 426–433, 2014.

[30] E. Giovannucci, A. Ascherio, E. B. Rimm, G. A. Colditz, M. J. Stampfer, and W. C. Willett, "Physical activity, obesity, and risk for colon cancer and adenoma in men," *Annals of Internal Medicine*, vol. 122, no. 5, pp. 327–334, 1995.

[31] X. Chen, G. Wei, T. Jalili et al., "The associations of plant protein intake with all-cause mortality in CKD," *American Journal of Kidney Diseases*, vol. 67, no. 3, pp. 423–430, 2016.

[32] A. S. Levey, L. A. Stevens, C. H. Schmid et al., "A new equation to estimate glomerular filtration rate," *Annals of Internal Medicine*, vol. 150, no. 9, pp. 604–612, 2009.

[33] J. A. Sterne, I. R. White, J. B. Carlin et al., "Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls," *BMJ*, vol. 338, no. jun29 1, article b2393, 2009.

[34] J. C. Jakobsen, C. Gluud, J. Wetterslev, and P. Winkel, "When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts," *BMC Medical Research Methodology*, vol. 17, no. 1, p. 162, 2017.

[35] L. Desquilbet and F. Mariotti, "Dose-response analyses using restricted cubic spline functions in public health research," *Statistics in Medicine*, vol. 29, no. 9, pp. 1037–1057, 2010.

[36] A. Tchernof and J. P. Després, "Pathophysiology of human visceral obesity: an update," *Physiological Reviews*, vol. 93, no. 1, pp. 359–404, 2013.

[37] C. L. Donohoe, S. L. Doyle, and J. V. Reynolds, "Visceral adiposity, insulin resistance and cancer risk," *Diabetology and Metabolic Syndrome*, vol. 3, no. 1, p. 12, 2011.

[38] A. Kubo, M. B. Cook, N. J. Shaheen et al., "Sex-specific associations between body mass index, waist circumference and the risk of Barrett's oesophagus: a pooled analysis from the international BEACON consortium," *Gut*, vol. 62, no. 12, pp. 1684–1691, 2013.

[39] E. E. Frezza, M. S. Wachtel, and M. Chiriva-Internati, "Influence of obesity on the risk of developing colon cancer," *Gut*, vol. 55, no. 2, pp. 285–291, 2006.

[40] S. S. Albrecht, P. Gordon-Larsen, D. Stern, and B. M. Popkin, "Is waist circumference per body mass index rising differentially across the United States, England, China and Mexico?," *European Journal of Clinical Nutrition*, vol. 69, no. 12, pp. 1306–1312, 2015.

[41] I. Janssen, M. Shields, C. L. Craig, and M. S. Tremblay, "Changes in the obesity phenotype within Canadian children and adults, 1981 to 2007-2009," *Obesity (Silver Spring)*, vol. 20, no. 4, pp. 916–919, 2012.

[42] M. E. Lean, T. S. Han, and C. E. Morrison, "Waist circumference as a measure for indicating need for weight management," *BMJ*, vol. 311, no. 6998, pp. 158–161, 1995.

[43] T. Gadekar, P. Dudeja, I. Basu, S. Vashisht, and S. Mukherji, "Correlation of visceral body fat with waist-hip ratio, waist circumference and body mass index in healthy adults: a cross sectional study," *Medical Journal, Armed Forces India*, vol. 76, no. 1, pp. 41–46, 2020.

[44] C. P. S. Pinho, A. D. S. Diniz, I. K. G. Arruda, A. P. D. L. Leite, M. M. V. Petribu, and I. G. Rodrigues, "Waist circumference measurement sites and their association with visceral and subcutaneous fat and cardiometabolic abnormalities," *Archives of endocrinology and metabolism*, vol. 62, no. 4, pp. 416–423, 2018.

[45] S. L. Doyle, C. L. Donohoe, J. Lysaght, and J. V. Reynolds, "Visceral obesity, metabolic syndrome, insulin resistance and cancer," *The Proceedings of the Nutrition Society*, vol. 71, no. 1, pp. 181–189, 2012.

[46] I. S. Jung, C. M. Shin, S. J. Park et al., "Association of visceral adiposity and insulin resistance with colorectal adenoma and colorectal cancer," *Intestinal research*, vol. 17, no. 3, pp. 404–412, 2019.

[47] M. G. Swainson, A. M. Batterham, C. Tsakirides, Z. H. Rutherford, and K. Hind, "Prediction of whole-body fat percentage and visceral adipose tissue mass from five anthropometric variables," *PLoS One*, vol. 12, no. 5, article e0177175, 2017.

[48] N. H. Kim, Y. S. Jung, J. H. Park, D. I. Park, and C. I. Sohn, "Abdominal obesity is more predictive of advanced colorectal neoplasia risk than overall obesity in men: a Cross-sectional study," *Journal of Clinical Gastroenterology*, vol. 53, no. 7, pp. e284–e290, 2019.

[49] H. Sanikini, D. C. Muller, M. Sophiea et al., "Anthropometric and reproductive factors and risk of esophageal and gastric cancer by subtype and subsite: results from the European prospective investigation into cancer and nutrition (EPIC) cohort," *International Journal of Cancer*, vol. 146, no. 4, pp. 929–942, 2020.

[50] Y. Chen, L. Liu, X. Wang et al., "Body mass index and risk of gastric cancer: a meta-analysis of a population with more than ten million from 24 prospective studies," *Cancer Epidemiology, Biomarkers & Prevention*, vol. 22, no. 8, pp. 1395–1408, 2013.

[51] H. Sanikini, D. C. Muller, M. Chadeau-Hyam, N. Murphy, M. J. Gunter, and A. J. Cross, "Anthropometry, body fat composition and reproductive factors and risk of oesophageal and gastric cancer by subtype and subsite in the UK biobank cohort," *PLoS One*, vol. 15, no. 10, article e0240413, 2020.

[52] L. Wei, N. Li, G. Wang et al., "Waist circumference might be a predictor of primary liver cancer: a population-based cohort study," *Frontiers in Oncology*, vol. 8, p. 607, 2018.

[53] K. Lo, Y. Q. Huang, G. Shen et al., "Effects of waist to height ratio, waist circumference, body mass index on the risk of chronic diseases, all-cause, cardiovascular and cancer mortality," *Postgraduate Medical Journal*, vol. 97, no. 1147, pp. 306–311, 2021.

[54] Y. I. Min, Y. Gao, P. Anugu, A. Anugu, and A. Correa, "Obesity and overall mortality: findings from the Jackson heart study," *BMC Public Health*, vol. 21, no. 1, p. 50, 2021.

[55] B. J. Caan, E. M. Cespedes Feliciano, and C. H. Kroenke, "The importance of body composition in explaining the overweight paradox in cancer-counterpoint," *Cancer Research*, vol. 78, no. 8, pp. 1906–1912, 2018.

[56] D. S. M. Chan, A. R. Vieira, D. Aune et al., "Body mass index and survival in women with breast cancer–systematic literature review and meta-analysis of 82 follow-up studies," *Annals of Oncology*, vol. 25, no. 10, pp. 1901–1914, 2014.

[57] Q. Yang, Z. Zhang, E. W. Gregg, W. D. Flanders, R. Merritt, and F. B. Hu, "Added sugar intake and cardiovascular diseases mortality among US adults," *JAMA Internal Medicine*, vol. 174, no. 4, pp. 516–524, 2014.

[58] A. J. McEligot, V. Poynor, R. Sharma, and A. Panangadan, "Logistic LASSO regression for dietary intakes and breast cancer," *Nutrients*, vol. 12, no. 9, p. 2652, 2020.

[59] E. Lau, J. S. Neves, M. Ferreira-Magalhães, D. Carvalho, and P. Freitas, "Probiotic ingestion, obesity, and metabolic-related disorders: results from NHANES, 1999-2014," *Nutrients*, vol. 11, no. 7, p. 1482, 2019.

[60] H. Li, D. Boakye, X. Chen et al., "Associations of body mass index at different ages with early-onset colorectal cancer," *Gastroenterology*, vol. 162, no. 4, pp. 1088–1097.e3, 2022.

[61] X. Zhang, E. Theodoratou, X. Li et al., "Genetically predicted physical activity levels are associated with lower colorectal cancer risk: a Mendelian randomisation study," *British Journal of Cancer*, vol. 124, no. 7, pp. 1330–1338, 2021.

[62] S. C. Houghton, H. Eliassen, R. M. Tamimi, W. C. Willett, B. A. Rosner, and S. E. Hankinson, "Central adiposity and subsequent risk of breast cancer by menopause status," *Journal of the National Cancer Institute*, vol. 113, no. 7, pp. 900–908, 2021.

[63] A. C. Sylvetsky and K. I. Rother, "Nonnutritive sweeteners in weight management and chronic disease: a review," *Obesity (Silver Spring)*, vol. 26, no. 4, pp. 635–640, 2018.

[64] D. B. Kim, K. M. Lee, J. M. Lee, S. H. Ko, K. D. Han, and Y. G. Park, "Waist circumference, body mass index, and colorectal cancer risk according to diabetes status: a Korean nationwide population-based cohort study," *Journal of Gastroenterology and Hepatology*, vol. 36, no. 2, pp. 397–405, 2021.

[65] A. Linauskas, K. Overvad, D. Symmons, M. B. Johansen, K. Stengaard-Pedersen, and A. Thurah, "Body fat percentage, waist circumference, and obesity as risk factors for rheumatoid arthritis: a Danish cohort study," *Arthritis Care & Research (Hoboken)*, vol. 71, no. 6, pp. 777–786, 2019.

[66] T. L. Berentzen, M. U. Jakobsen, J. Halkjaer, A. Tjønneland, K. Overvad, and T. I. A. Sørensen, "Changes in waist circumference and mortality in middle-aged men and women," *PLoS One*, vol. 5, no. 9, p. e13097, 2010.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 5676570, 8 pages
https://doi.org/10.1155/2022/5676570

Research Article

Establishment and Validation of a Machine Learning Prediction Model Based on Big Data for Predicting the Risk of Bone Metastasis in Renal Cell Carcinoma Patients

Chan Xu,[1,2] Wencai Liu,[3] Chengliang Yin,[4] Wanying Li,[2] Jingjing Liu,[5] Wanli Sheng,[6] Haotong Tang,[4] Wenle Li,[7] and Qingqing Zhang[8]

[1]*Department of Dermatology, Xianyang Central Hospital, Xianyang 712000, China*
[2]*Department of Clinical Medical Research Center, Xianyang Central Hospital, Xianyang 712000, China*
[3]*Department of Orthopaedic Surgery, The First Affiliated Hospital of Nanchang University, Nanchang 330006, China*
[4]*Faculty of Medicine, Macau University of Science and Technology, Macau 999078, China*
[5]*Department of Shanghai National Engineering Research Center of Biochip, Shanghai 201203, China*
[6]*Hohhot Technical Center of Hohhot Customs District, Hohhot 010020, China*
[7]*Molecular Imaging and Translational Medicine Research Center, State Key Laboratory of Molecular Vaccinology and Molecular Diagnostics, Xiamen University, Xiamen 361005, China*
[8]*Department of Otolaryngology-Head and Neck Surgery, The Second Affiliated Hospital of Xi'an Jiao Tong University, Xi'an 710004, China*

Correspondence should be addressed to Haotong Tang; 2009853ghp20006@student.must.edu.mo, Wenle Li; drlee0910@163.com, and Qingqing Zhang; zhangqingqingxly@163.com

Received 20 June 2022; Accepted 28 July 2022; Published 3 October 2022

Academic Editor: Xiaohan Ren

Purpose. Since the prognosis of renal cell carcinoma (RCC) patients with bone metastasis (BM) is poor, this study is aimed at using big data to build a machine learning (ML) model to predict the risk of BM in RCC patients. *Methods*. A retrospective study was conducted on 40,355 RCC patients in the SEER database from 2010 to 2017. LASSO regression and multivariate logistic regression analysis was performed to determine independent risk factors of RCC-BM. Six ML algorithm models, including LR, GBM, XGB, RF, DT, and NBC, were used to establish risk models for predicting RCC-BM. The prediction performance of ML models was weighed by 10-fold cross-validation. *Results*. The study investigated 40,355 patients diagnosed with RCC in the SEER database, where 1,811 (4.5%) were BM patients. Independent risk factors for BM were tumor grade, T stage, N stage, liver metastasis, lung metastasis, and brain metastasis. Among the RCC-BM risk prediction models established by six ML algorithms, the XGB model showed the best prediction performance (AUC = 0.891). Therefore, a network calculator based on the XGB model was established to individually assess the risk of BM in patients with RCC. *Conclusion*. The XGB risk prediction model based on the ML algorithm performed a good prediction effect on BM in RCC patients.

1. Introduction

Renal cell carcinoma (RCC), a renal space-occupying lesion originating from renal tubular epithelial cells, accounts for 4% of all malignant tumors [1]. As the most common type of renal cancer, RCC took up 85%-90% of renal malignant tumors in adults [2]. Since the early clinical manifestations of RCC are hidden, 20%-40% of patients will eventually suffer from metastatic RCC (mRCC) [3]. At present, surgical resection is the first-line treatment for RCC. However, in some RCC patients, distant metastasis after radical nephrectomy still reappeared or even occurred [4]. In addition, most RCC will become resistant to chemotherapy and radiotherapy once they develop into relapse or metastasis [5].

Bone metastasis (BM) is one of the most frequent sites of solid malignant tumors, the occurrence of which reveals the

poor prognosis of tumor patients. As RCC has obvious osteotaxis in distant metastasis, bone is the second common distant metastasis site of RCC except the lung [6, 7]. Studies have shown that in about 20-35% of patients with RCC disease progression, renal cell carcinoma bone metastasis (RCC-BM) will emerge, such as the pelvis, spine, and ribs [3]. BM of RCC mainly leads to osteolytic destructive changes such as skeletal-related events (SER) [8], including pathological fracture, spinal cord and/or nerve root compression, and bone pain [9]. More than 70% of BM-RCC patients have experienced at least one SER during their survival, which severely reduced their quality of life and survival time [10]. Previous studies indicated that the prognosis of patients with RCC-BM was poor. Their median overall survival time (OST) was only 12-28 months, and the 5-year overall survival rate was only 11%, while the median OST of mRCC patients without BM was prolonged to 31 months and the 5-year overall survival rate increased to 47% [10, 11]. Therefore, bone metastasis is an important cause of death in patients with advanced RCC and it is crucial to predict the risk of RCC-BM.

Several previous studies reported the risk factors and prognostic factors of BM in RCC patients [12–15], and some developed traditional nomograms to predict the risk of RCC-BM [5]. However, there are few studies using machine learning (ML) method to construct the risk prediction model of RCC-BM based on big data. This study is aimed at (1) setting up the RCC-BM risk prediction model through ML and verifying the validity of the model with external data and (2) constructing a network calculator to facilitate clinicians to choose more reasonable diagnosis and treatment for RCC patients.

2. Methods

2.1. Study Population Selection. The training group data analyzed in this study are from the SEER database (http://seer.cancer.gov/about/), where the analysis of anonymous data is exempt from medical ethics review and does not require informed consent from patients.

2.2. Data Collection. All RCC data in the retrospective cohort study from 2010 to 2017 were extracted and subsumed as training group data with SEER∗Stat (version 8.3.6) software. According to the inclusion/exclusion criteria, 40,355 patients were selected into the training group. The inclusion/exclusion criteria were as follows: (1) RCC was the first or primary tumor; (2) patients with RCC diagnosed by pathology (the validation group was diagnosed by at least two pathologists blindly); (3) patients with complete clinicopathological features, demographic data, and follow-up data; and (4) patients with RCC proved by autopsy or death were excluded from this study.

Based on the specific information of RCC patients from the SEER database, 17 variables were selected to determine the independent risk factors of BM in RCC patients, including marital status, age, race, serial number, survival time, survival status, gender, primary location, grade, side, pathological stage, T stage, N stage, tumor size, bone metastasis, brain metastasis, and liver metastasis. The risk prediction models were framed using data of the training group.

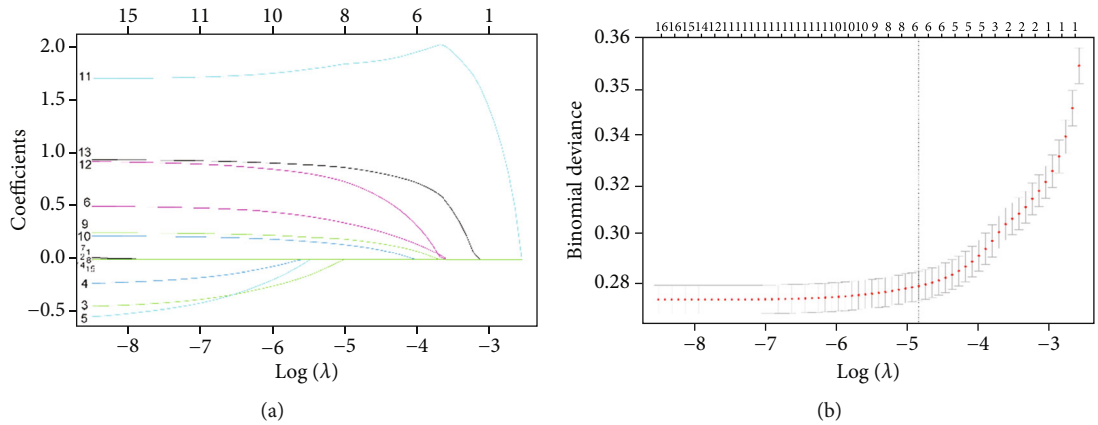
2.3. Establishment and Verification of Prediction Models. Six ML models, including logistic regression (LR), gradient boosting machine (GBM), extreme gradient boosting (XGB), Random Forest (RF), Decision Tree (DT), and Naive Bayesian model (NBC), were used to build prediction models, the performance of which was compared by 10-fold cross-validation method [16–19]. The model with the greatest AUC value was regarded as the preferred prediction model, whose corresponding network calculator is designed to individually assess the risk of BM in patients with RCC [20–23].

2.4. Statistical Analysis. The measurement data is expressed in mean (SD), and the counting data is expressed in frequency (percentage). Independent samples t-test, chi-square test, LASSO regression analysis, univariate and multivariate logistic regression analysis, 10-fold cross-validation, and other statistical analysis were performed by SPSS 26.0 (SPSS Inc., Chicago, USA) software. P values < 0.05 were considered statistically significant. R software (version 4.0.5, https://www.r-project.org/) was applied for drawing the correlation heat map and ROC curve and developing a predictive model which used the "shiny" package to establish a web calculator.

3. Results

3.1. Characteristics of the Study Population. A total of 40,355 RCC patients from the SEER database were included in this study to establish the training group. 4.5% (1811 cases) of RCC patients progressed to BM. Then, six risk factors for predicting BM in RCC patients were screened by LASSO regression, including tumor grade, T stage, N stage, liver metastasis, lung metastasis, and brain metastasis (Figure 1), which were viewed as predictors in the correlation heat map (Figure 2).

3.2. Independent Risk Factors of BM in RCC Patients. Univariate and multivariate logistic regression analyses were carried out to value the independent risk factors of BM in RCC patients. Univariate logistic regression analysis displayed that brain metastasis, liver metastasis, lung metastasis, tumor grade, T stage, and N stage were importantly associated with BM in patients with RCC ($P < 0.05$). Further multivariate logistic regression analysis indicated that brain metastasis (OR = 2.46, 95%CI = 1.98 − 3.05), liver metastasis (OR = 2.37, 95%CI = 2.01 − 2.8), lung metastasis (OR = 5.2, 95%CI = 4.58 − 5.89), tumor grade (poorly differentiated: OR = 3.08, 95%CI = 1.87 − 5.08; undifferentiated: OR = 4.47, 95%CI = 2.69 − 7.42; undifferentiated: OR = 7.97, 95% CI = 4.9 − 12.97), T stage (T2 stage: OR = 2.13, 95%CI = 1.81 − 2.5; T3 stage: OR = 1.84, 95%CI = 1.59 − 2.13; T4 stage: OR = 2.08, 95%CI = 1.68 − 2.59; and TX stage: OR = 3.11, 95%CI = 2.51 − 3.86), and N stage (N1 stage: OR = 2.18, 95%CI = 1.9 − 2.51; NX stage: OR = 1.64, 95%CI = 1.34 − 2.01) were independent risk factors for BM in RCC patients ($P < 0.001$, Table 1).



(a) (b)

Figure 1: The plot of the LASSO model: (a) optimal parameter (λ) selection in the LASSO model, with the optimal tuning parameter log (λ) in the horizontal coordinate and the regression coefficients in the vertical coordinate; (b) distribution of LASSO coefficients about the clinical factors, with the optimal tuning parameter log (λ) in the horizontal coordinate and the binomial deviation in the vertical coordinate.

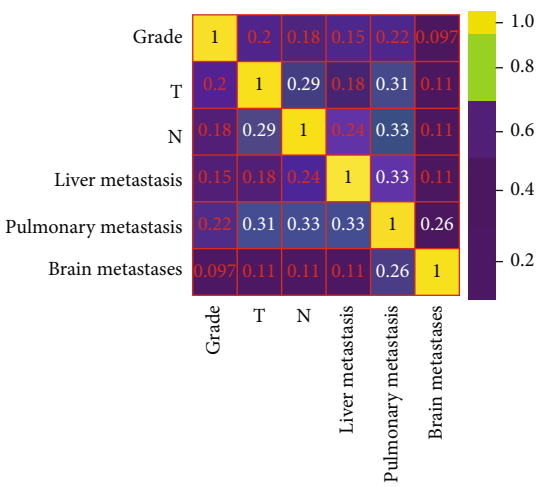


Figure 2: The correlation heat map of risk factors.

3.3. Selection and Verification of the Prediction Models. The prediction performance of six ML algorithm models (LR, NBC, DT, RF, GBM, and XGB) was compared by 10-fold cross-validation method, which indicates that the prediction value of all models above was great (AUC > 0.850). In descending order, the predictive ability of models is XGB, RF, GBM, NBC, LR, and DT, of which XGB is the best in predicting RCC-BM (average AUC = 0.891, Figure 3). Therefore, the XGB model is selected as the optimal prediction model finally.

The importance of each risk factor is not identical in different ML prediction models. Among them, lung metastasis is the most important clinical feature in the six models, while brain metastasis is of the least significance feature in RF, GBM, and XGB models, familiar as tumor grade in NBC and DT models and N stage in the LR model. In the XGB model, the independent risk factors are arranged according to their importance, which are lung metastasis, T stage, liver metastasis, tumor grade, N stage, and brain metastasis. The value of risk factors in other models is shown in Figure 4.

3.4. Construction of the Web Calculator. Based on the GBM model possessing the best performance, a risk web calculator was designed in this study (https://share.streamlit.io/liuwencai5/renal_bone/main/renal_bone.py). By inputting the relevant clinicopathological variables of RCC patients, clinicians could predict the risk of BM in patients with RCC (Figure 5).

4. Discussion

As an important marker of poor prognosis in patients with RCC, early detection and intervention for BM are urgently needed. Guo et al. [14] analyzed the data of 45,824 RCC patients recorded in the SEER database from 2010 to 2014 and found that 3.29% patients were diagnosed with BM at the initial diagnosis. In our study, 4.5% RCC patients in the training group developed into BM, while Zekri et al. [24] reported that 30-40% of advanced RCC patients turned into BM. Therefore, the incidence rate of BM may be underestimated since the patients showed no symptoms when the initial diagnosis was made or BM appeared at the advanced stage of RCC disease which was not recorded in the SEER database. At present, the guidelines only recommend bone imaging for patients with uncomfortable symptoms or abnormal alkaline phosphatase level [25]. Thus, consequently, patients with asymptomatic BM could not be treated timely and effectively. At present, bone metastasis has been recognized as one of the adverse prognostic factors of RCC patients [8]. In addition, the resection of whole spinal BM can prolong the survival time potentially for the patients with isolated spinal BM and no visceral metastasis [26], but the prognosis of BM in RCC patients is still poor compared with that of other tumors such as lung cancer. With the improved bone-targeted therapy of BM patients, the average OST of BM patients is 12.0-31.8 months [27–29]. The lack of effective chemotherapeutic drugs may be the main reason for the poor outcome of RCC patients with BM [11]. Due to the occult onset and poor prognosis of BM, it is necessary to study the risk factors of BM in patients with RCC. Additionally, early identification and

Table 1: Univariate and multivariate logistic regression for the risk of bone metastasis in patients with renal cancer.

Characteristics	Univariate logistic regression			Multivariable logistic regression		
	OR	CI	*P*	OR	CI	*P*
Brain metastases						
No	Ref	Ref	Ref	Ref	Ref	Ref
Yes	14.72	12.24-17.7	<0.001	2.46	1.98-3.05	<0.001
Grade						
Well differentiated						
Moderately differentiated	1.89	1.14-3.13	0.014	1.62	0.97-2.69	0.064
Poorly differentiated	5.84	3.57-9.56	<0.001	3.08	1.87-5.08	<0.001
Undifferentiated; anaplastic	13.92	8.48-22.84	<0.001	4.47	2.69-7.42	<0.001
Unknown	21.09	13.05-34.09	<0.001	7.97	4.9-12.97	<0.001
Liver metastasis						
No	Ref	Ref	Ref	Ref	Ref	Ref
Yes	15.57	13.54-17.9	<0.001	2.37	2.01-2.8	<0.001
N						
N0	Ref	Ref	Ref	Ref	Ref	Ref
N1	10.08	8.99-11.3	<0.001	2.18	1.9-2.51	<0.001
N2	4.47	2.88-6.94	<0.001	1.58	0.97-2.58	0.067
NX	4.88	4.14-5.75	<0.001	1.64	1.34-2.01	<0.001
Pulmonary metastasis						
No	Ref	Ref	Ref	Ref	Ref	Ref
Yes	18.6	16.8-20.61	<0.001	5.2	4.58-5.89	<0.001
T						
T1	Ref	Ref	Ref	Ref	Ref	Ref
T2	4.53	3.93-5.23	<0.001	2.13	1.81-2.5	<0.001
T3	3.78	3.34-4.27	<0.001	1.84	1.59-2.13	<0.001
T4	10.76	9-12.88	<0.001	2.08	1.68-2.59	<0.001
TX	18.08	15.14-21.59	<0.001	3.11	2.51-3.86	<0.001

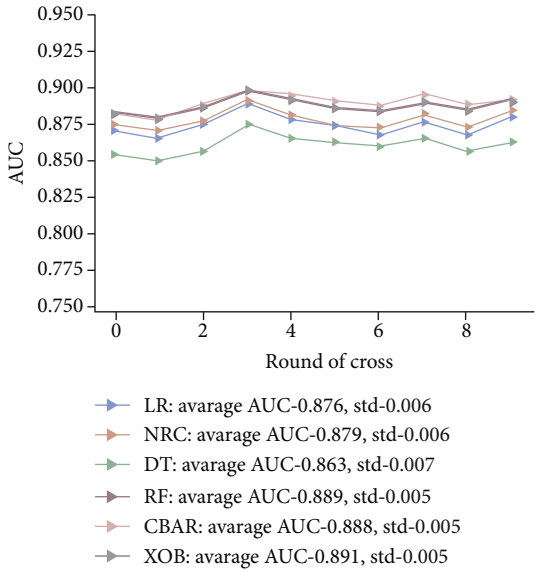


Figure 3: The plot of 10-fold cross-validation. LR: logistic regression; GBM: gradient boosting machine; XGB: extreme gradient boosting; RF: Random Forest; DT: Decision Tree; NBC: Naïve Bayesian model.

evaluation of BM are of great significance to improve the precision of the diagnosis, determine the treatment plans, and prevent RCC complications such as SER in patients with symptomless BM.

In this study, multivariate logistic regression analysis revealed that brain metastasis, liver metastasis, lung metastasis, poor tumor differentiation, high T stage, and N stage were independent risk factors for BM in RCC patients. Similarly, Guo et al. [14] found that male gender; higher T stage; lymph node involvement; poor tumor differentiation; presence of lung, liver, and brain metastasis; and the collecting duct type of RCC were positively associated with BM occurrence. Furthermore, Fan et al. [30], using nomogram to quantify the risk of RCC-BM patients, found that the independent factors of RCC complicated with BM include grade, histological type, N stage, operation, brain metastasis, and lung metastasis, which was basically consistent with our research results. Additionally, through a retrospective analysis of 372 RCC patients, Chen et al. [31] discovered that the concentrations of ALP, calcium, and Hb were potential risk factors for bone metastasis in patients with RCC. ALP > 105.5 U/L,calcium > 2.615 mmol/L, andHb < 111.5 g/L in newly diagnosed RCC patients suggest that BM is more likely to occur in these patients; hence, close monitoring

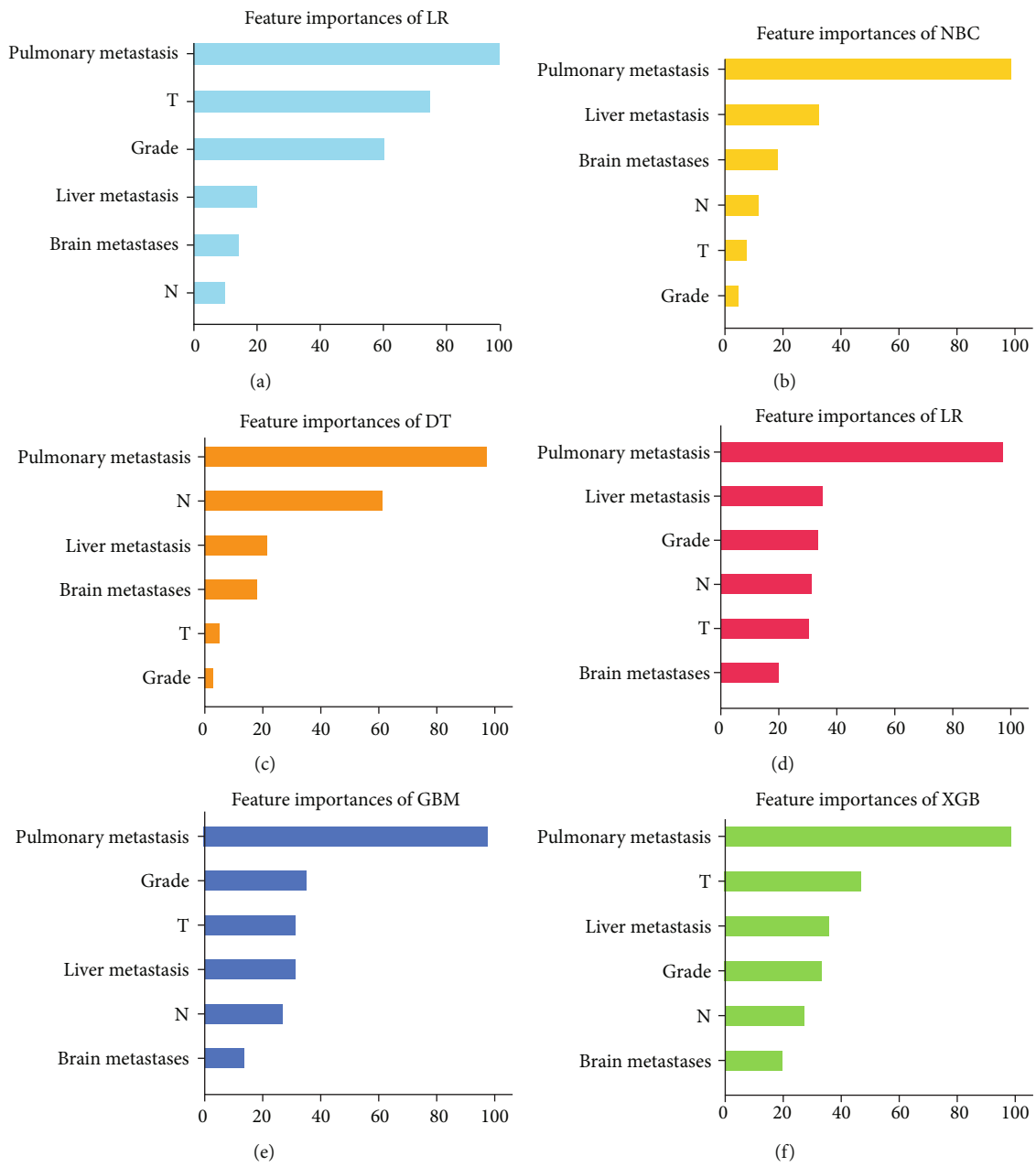


Figure 4: Feature importance distribution map of ML models.

and active bone scanning should be carried out to determine whether bone metastasis has happened. With the in-depth study of RCC-BM, more and more prognostic factors of bone metastasis also have been discovered. Yoshiyama et al. [32] considered that patients' age, ECOG performance, histology, MSKCC prognosis score, concomitant metastasis, and the time from nephrectomy to bone metastasis were important factors related to the prognosis of RCC-BM. Subsequently, Ruatta et al. [15] tracked 1750 RCC patients and found that MSKCC score, BMs number, and radical resection were essential prognostic factors for RCC patients with BMs.

With the development of TNM staging system and pathological classification criteria of renal cell carcinoma, a variety of RCC prognosis analysis systems have appeared. But they have several limitations. TNM staging system depends on three pathological indexes while it ignores other risk factors, which reduces the accuracy of prognosis prediction of RCC patients. MSKCC model (Memorial Sloan Kettering Cancer Center-based poor-risk groups) and IMDC model (International Metastatic Renal Cell Carcinoma Database Consortium-based poor-risk groups) [27] lack the function of comprehensive analysis for patients. GRCC model (Gustave Roussy Cancer Campus) [33] is more accurate and convenient than the MSKCC and IMDC prognostic analysis model, but it is not designed for RCC-BM patients specifically. The B-FOM scoring system (Fujimoto–Owari–Miyake bone score) is characterized by bone metastasis specificity, yet poor tissue source specificity is its short board [34]. The traditional Cox regression or logistic regression

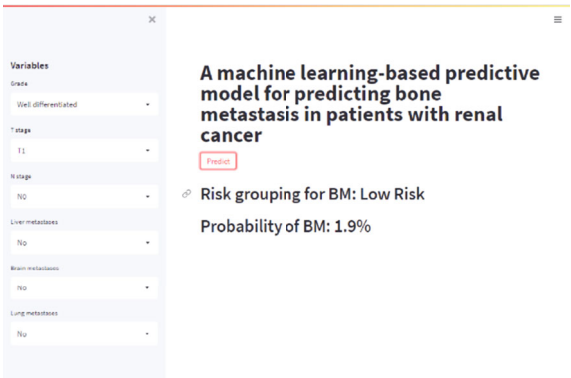


Figure 5: The risk web calculator was designed based on the GBM model.

analysis is visualized in this article; thus, clinicians can easily calculate the probability of BM in RCC patients without understanding the complex underlying mathematical formula. Distinguished from prediction analysis system or model mentioned above, this study creatively developed and verified six machine learning algorithms models, which were specially performed to estimate the risk of BM in RCC patients. The XGB model with the best prediction performance was selected through 10-fold verification methods, and an online calculator was established to evaluate the individual probability of RCC-BM. The ML-based model can be used to guide clinical treatment decisions, help clinicians better predict the BM risk, and take necessary interventions to improve the survival time and life quality of RCC patients.

The limitations of our study could not be ignored though. Firstly, as a retrospective cohort study, the inevitable selection bias may affect the results considerably. Since the SEER database only collects the initial diagnosis results, BM arising in the advanced stage of RCC may be omitted. Secondly, the deficiency of external validation using the data from the local validation group patients could not assess the accuracy of the selected XGB models in diagnosing BM in RCC patients and therefore establish the other clinical utility analysis such as probability density functions (PDF) and clinical utility curves (CUC). Additionally, we were unable to obtain some effective indicators for predicting RCC-BM from SEER, such as transforming growth factor-β (TGF-β) [35], fibroblast growth factor (FGF) [1], insulin-like growth factor [36], bone morphogenetic protein [37], AFP [38], CA-199 [31], and Fuhrman nuclear grade [39]. Future studies are needed to incorporate tumor characteristics, laboratory results, and treatment regimens to establish a higher dimensional predictive model.

5. Conclusion

This study retrospectively analyzed the independent risk factors of BM in renal cell carcinoma based on the SEER database, including tumor grade, T stage, N stage, liver metastasis, lung metastasis, and brain metastasis. On the foundation of the SEER dataset, we constructed and validated six machine learning models including LR, GBM, XGB, RF, DT, and NBC and subsequently selected XGB as the optimal prediction model. The network calculator designed on the basis of XGB provided important support for clinicians to make accurate treatment decisions.

Abbreviations

BM:	Bone metastasis
RCC:	Renal cell carcinoma
ML:	Machine learning
SEER:	Surveillance, Epidemiology, and End Results
LR:	Logistic regression
GBM:	Gradient boosting machine
XG:	Extreme gradient boosting
RF:	Random Forest
DT:	Decision Tree
NBC:	Naive Bayesian model
AUC:	Area under the curve
mRCC:	Metastatic RCC
SER:	Skeletal-related events
OST:	Overall survival time
PDF:	Probability density functions
CUC:	Clinical utility curves
MSKCC model:	Memorial Sloan Kettering Cancer Center-based poor-risk groups
IMDC model:	International Metastatic Renal Cell Carcinoma Database Consortium-based poor-risk groups
GRCC model:	Gustave Roussy Cancer Campus
B-FOM scoring system:	Fujimoto–Owari–Miyake bone score
TGF-β:	Transforming growth factor-β
FGF:	Fibroblast growth factor.

Data Availability

The training group data analyzed in this study are from the SEER database (http://seer.cancer.gov/about/), where the analysis of anonymous data is exempt from medical ethics review and does not require informed consent from patients.

Ethical Approval

This study is based on the SEER database and does not require ethical approval.

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflict of interest.

Authors' Contributions

CX, WCL, and CLY are all considered as the first authors. QQZ, WLL, and HTT completed the study design. CX and WLL performed the study and collected and analyzed the data. QQZ, STD, and CX drafted the manuscript. WYL, XFX, XMQ, and JJL provided the expert consultations and

suggestions. WCL and WLS conceived of the study, participated in its design and coordination, and helped to embellish language. All authors reviewed the final version of the manuscript. Chan Xu, Wencai Liu, and Chengliang Yin contributed equally to this work.

Acknowledgments

This study was supported by Shaanxi Provincial Health and Health Research Fund Project (2022E006).

References

[1] Y. Adachi, H. Kamiyama, K. Ichikawa et al., "Inhibition of FGFR reactivates IFNγ signaling in tumor cells to enhance the combined antitumor activity of lenvatinib with anti-PD-1 antibodies," *Cancer Research*, vol. 82, no. 2, pp. 292–306, 2022.

[2] Y. Wang, Y. Zhang, P. Wang, X. Fu, and W. Lin, "Circular RNAs in renal cell carcinoma: implications for tumorigenesis, diagnosis, and therapy," *Molecular Cancer*, vol. 19, no. 1, p. 149, 2020.

[3] S. V. L. Vig, E. Zan, and S. K. Kang, "Imaging for metastatic renal cell carcinoma," *The Urologic Clinics of North America*, vol. 47, no. 3, pp. 281–291, 2020.

[4] T. Klatte, S. H. Rossi, and G. D. Stewart, "Prognostic factors and prognostic models for renal cell carcinoma: a literature review," *World Journal of Urology*, vol. 36, no. 12, pp. 1943–1952, 2018.

[5] A. Brozovich, B. Garmezy, T. Pan, L. Wang, M. C. Farach-Carson, and R. L. Satcher, "All bone metastases are not created equal: revisiting treatment resistance in renal cell carcinoma," *Journal of Bone Oncology*, vol. 31, article 100399, 2021.

[6] J. Wang, H. You, J. Qi, C. Yang, Y. Ren, and H. Cheng, "Autocrine and paracrine STIP1 signaling promote osteolytic bone metastasis in renal cell carcinoma," *Oncotarget*, vol. 8, no. 10, pp. 17012–17026, 2017.

[7] W. Li, S. Dong, H. Wang et al., "Risk analysis of pulmonary metastasis of chondrosarcoma by establishing and validating a new clinical prediction model: a clinical study based on SEER database," *BMC Musculoskeletal Disorders*, vol. 22, no. 1, p. 529, 2021.

[8] M. Umer, Y. Mohib, M. Atif, and M. Nazim, "Skeletal metastasis in renal cell carcinoma: a review," *Annals of Medicine and Surgery*, vol. 27, pp. 9–16, 2018.

[9] T. Negishi, N. Furubayashi, D. Takamatsu et al., "Radiographical efficacy of systemic treatment for bone metastasis from renal cell carcinoma," *Oncology Letters*, vol. 20, no. 5, p. 1, 2020.

[10] S. C. Chen and P. L. Kuo, "Bone metastasis from renal cell carcinoma," *International Journal of Molecular Sciences*, vol. 17, no. 6, p. 987, 2016.

[11] E. Woodward, S. Jagdev, L. McParland et al., "Skeletal complications and survival in renal cancer patients with bone metastases," *Bone*, vol. 48, no. 1, pp. 160–166, 2011.

[12] C. R. Goodwin, A. K. Ahmed, C. Boone et al., "The challenges of renal cell carcinoma metastatic to the spine: a systematic review of survival and treatment," *Global Spine Journal*, vol. 8, no. 5, pp. 517–526, 2018.

[13] A. Guida, B. Escudier, and L. Albiges, "Treating patients with renal cell carcinoma and bone metastases," *Expert Review of Anticancer Therapy*, vol. 18, no. 11, pp. 1135–1143, 2018.

[14] Q. Guo, C. Zhang, X. Guo et al., "Incidence of bone metastasis and factors contributing to its development and prognosis in newly diagnosed renal cell carcinoma: a population-based study," *Cancer Management and Research*, vol. 10, pp. 2935–2944, 2018.

[15] F. Ruatta, L. Derosa, B. Escudier et al., "Prognosis of renal cell carcinoma with bone metastases: experience from a large cancer centre," *European Journal of Cancer*, vol. 107, pp. 79–85, 2019.

[16] W. Li, Y. Liu, W. Liu et al., "Machine learning-based prediction of lymph node metastasis among osteosarcoma patients," *Frontiers in Oncology*, vol. 12, article 797103, 2022.

[17] W. Li, W. Liu, M. F. Hussain et al., "An external-validated prediction model to predict lung metastasis among osteosarcoma: a multicenter analysis based on machine learning," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 2220527, 10 pages, 2022.

[18] S. Jiang, X. Ren, S. Liu et al., "Integrated analysis of the prognosis-associated RNA-binding protein genes and candidate drugs in renal papillary cell carcinoma," *Frontiers in Genetics*, vol. 12, article 627508, 2021.

[19] X. Ren, X. Chen, X. Zhang et al., "Immune microenvironment and response in prostate cancer using large population cohorts," *Frontiers in Immunology*, vol. 12, article 686809, 2021.

[20] W. Li, H. Wang, S. Dong et al., "Establishment and validation of a nomogram and web calculator for the risk of new vertebral compression fractures and cement leakage after percutaneous vertebroplasty in patients with osteoporotic vertebral compression fractures," *European Spine Journal*, vol. 31, no. 5, pp. 1108–1121, 2022.

[21] W. Li, J. Wang, W. Liu et al., "Machine learning applications for the prediction of bone cement leakage in percutaneous vertebroplasty," *Frontiers in Public Health*, vol. 9, article 812023, 2021.

[22] J. C. Wang, R. Fu, X. W. Tao et al., "A radiomics-based model on non-contrast CT for predicting cirrhosis: make the most of image data," *Biomarker Research*, vol. 8, no. 1, p. 47, 2020.

[23] J. Wang, S. Tang, Y. Mao et al., "Radiomics analysis of contrast-enhanced CT for staging liver fibrosis: an update for image biomarker," *Hepatology International*, vol. 16, no. 3, pp. 627–639, 2022.

[24] J. Zekri, N. Ahmed, R. E. Coleman, and B. W. Hancock, "The skeletal metastatic complications of renal cell carcinoma," *International Journal of Oncology*, vol. 2001, pp. 379–382, 2001.

[25] R. J. Motzer, E. Jonasch, N. Agarwal et al., "Kidney cancer, version 2.2017, NCCN clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 15, no. 6, pp. 804–834, 2017.

[26] J. Zhai, N. Liu, H. Wang, G. Huang, and L. Man, "Clinical characteristics and prognosis of renal cell carcinoma with spinal bone metastases," *Frontiers in Oncology*, vol. 11, article 659779, 2021.

[27] Y. Dong, Z. Wang, X. Lu et al., "Clinical outcomes of 168 Chinese patients after local surgery for bone metastases arising from advanced renal cell carcinoma," *Cancer*, vol. 126, Supplement9, pp. 2079–2085, 2020.

[28] Y. Du, S. Pahernik, B. Hadaschik et al., "Survival and prognostic factors of patients with renal cell cancer with bone metastasis in the era of targeted therapy: a single-institution analysis," *Urologic Oncology*, vol. 34, no. 10, pp. 433.e1–433.e8, 2016.

[29] M. Santoni, A. Conti, G. Procopio et al., "Bone metastases in patients with metastatic renal cell carcinoma: are they always associated with poor prognosis?," *Journal of Experimental & Clinical Cancer Research*, vol. 34, no. 1, 2015.

[30] Z. Fan, Z. Huang, and X. Huang, "Bone metastasis in renal cell carcinoma patients: risk and prognostic factors and nomograms," *Journal of Oncology*, vol. 2021, Article ID 5575295, 17 pages, 2021.

[31] X. Y. Chen, M. Lan, Y. Zhou et al., "Risk factors for bone metastasis from renal cell cancer," *Journal of Bone Oncology*, vol. 9, pp. 29–33, 2017.

[32] A. Yoshiyama, T. Morii, M. Susa et al., "Preoperative evaluation of renal cell carcinoma patients with bone metastases on risks for blood loss, performance status and lethal event," *Journal of Orthopaedic Science*, vol. 22, no. 5, pp. 924–930, 2017.

[33] L. Derosa, M. A. Bayar, L. Albiges, G. Le Teuff, and B. Escudier, "A new prognostic model for survival in second line for metastatic renal cell carcinoma: development and external validation," *Angiogenesis*, vol. 22, no. 3, pp. 383–395, 2019.

[34] T. Owari, M. Miyake, Y. Nakai et al., "A genitourinary cancer-specific scoring system for the prediction of survival in patients with bone metastasis: a retrospective analysis of prostate cancer, renal cell carcinoma, and urothelial carcinoma," *Anticancer Research*, vol. 38, no. 5, pp. 3097–3103, 2018.

[35] Y. Hong, L. Gong, B. Yu, and Y. Dong, "PPM1A suppresses the proliferation and invasiveness of RCC cells via Smad2/3 signaling inhibition," *Journal of Receptor and Signal Transduction Research*, vol. 41, no. 3, pp. 245–254, 2021.

[36] W. Solarek, M. Koper, S. Lewicki, C. Szczylik, and A. M. Czarnecka, "Insulin and insulin-like growth factors act as renal cell cancer intratumoral regulators," *Journal of Cell Communication and Signaling*, vol. 13, no. 3, pp. 381–394, 2019.

[37] V. Brand, C. Lehmann, C. Umkehrer et al., "Impact of selective anti-BMP9 treatment on tumor cells and tumor angiogenesis," *Molecular Oncology*, vol. 10, no. 10, pp. 1603–1620, 2016.

[38] H. Z. Zhang and S. Y. Wang, "AFP-producing Xp11 translocation renal cell carcinoma: case report and review of the literature," *Current Problems in Cancer*, vol. 45, no. 3, article 100689, 2021.

[39] T. F. Lin, W. R. Lin, M. Chen et al., "Compare Fuhrman nuclear and chromophobe tumor grade on chromophobe RCC," *Open Medicine*, vol. 14, pp. 336–342, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8216339, 8 pages
https://doi.org/10.1155/2022/8216339

Research Article

The Efficacy of Anterior Cervical Corpectomy and Fusion and Posterior Total Laminectomy on Cervical Spinal Cord Injury and Quality of Life

Yanlin Yin,[1] Xinming Yang,[1] Ye Tian,[1] Ying Zhang,[1] Peinan Zhang,[1] Yongli Jia,[1] Yao Yao,[1] Xiuyu Du,[2] Tianmin Li,[3] and Xiaodong Li[4]

[1]*Department of Orthopedics, The First Affiliated Hospital of Hebei North University, Qiaoxi District, Zhangjiakou, Hebei, China*
[2]*Department of Trauma, The First Affiliated Hospital of Hebei North University, Qiaoxi District, Zhangjiakou, Hebei, China*
[3]*Department of General Medicine, The First Affiliated Hospital of Hebei North University, Qiaoxi District, Zhangjiakou, Hebei, China*
[4]*Department of Imaging, The First Affiliated Hospital of Hebei North University, Qiaoxi District, Zhangjiakou, Hebei, China*

Correspondence should be addressed to Xinming Yang; yangxinming@stu-edu.cn

Received 7 July 2022; Revised 14 August 2022; Accepted 18 August 2022; Published 29 September 2022

Academic Editor: Jincheng Wang

This study is aimed at investigating the efficacy of anterior cervical corpectomy and fusion and posterior total laminectomy in the treatment of cervical spinal cord injury and assessing the impact of the two approaches on cervical spine function and patient quality of life. Retrospectively analyze the clinical data from 180 patients with cervical spinal cord injury who were admitted to the First Affiliated Hospital of Hebei North University from June 2019 to June 2021. The patients were divided into an anterior approach group ($n = 89$, treated with anterior cervical corpectomy and fusion) and a posterior approach group ($n = 91$, treated with posterior total laminectomy). The amount of blood loss in the posterior approach group was larger compared to the anterior approach group. Patients in the posterior approach group had higher wound diameters and operation times compared to the anterior approach group, as well as the operation cost. The visual analogue scale (VAS) scores of patients in the posterior approach group were significantly higher than in the anterior approach group one month after operation. The Japanese Orthopaedic Association (JOA), neck disability index (NDI), and American Spinal Injury Association (ASIA) scores of patients in both groups at 1, 6, and 9 months after surgery were higher compared to those before surgery, yet no significant differences were observed between the two groups. Also, no significant difference was observed in the incidence of complication and the quality of life between the two groups before and after treatment. Anterior cervical corpectomy and fusion and posterior total laminectomy can effectively restore the cervical nerve function in the treatment of cervical spinal cord injury. However, anterior subtotal vertebral resection is associated with improved perioperative indicators compared to posterior total laminectomy. Clinically, surgical methods can be selected according to imaging findings, the general condition of patients, and individual economic status.

1. Introduction

Cervical spinal cord injury is a common clinical indication and one of the spinal cord injuries [1]. Spinal cord injury has major impacts on the structure and function of the spinal cord that can be caused by various factors and results in autonomic, motor, and sensory nerve dysfunction. Severe cervical spine injury may result in lifelong paralysis or even death [2–4]. Recently, the incidence of clinical traumatic cervical spinal cord injury has continued to increase due to the frequent occurrence of road traffic accidents, high altitude falls and other serious accidents [5].

In cases of cervical spinal cord injury, the primary treatment goal is to reverse the nerve injury, avoid secondary injury, and restore the integrity of the spine [6]. Dehydration, nerve nutrition, prevention of nerve edema, and

traction are generally adopted in the early stage of treatment of cervical spinal cord injury [7, 8]. Also, glucocorticoids are used to prevent further aggravation of cervical spinal cord injury [9].

Surgery has an important role in the management of patients with cervical spinal cord injury that is based on the imaging data and physical condition of individual patients. Due to the various mechanisms of cervical spinal cord injury, a range of surgical methods can be used. Anterior [10] or posterior approach decompression resection [11] is a common surgical method for the treatment of cervical spondylosis and has also been gradually applied in the treatment of cervical spinal cord injury. Decompression surgery can release pressure on the spinal cord with edema after injury, and the timing of surgery is particularly important to optimize recovery. The incorrect timing of surgery can lead to a slower recovery of function and even failure of the recovery of spinal cord function [12].

Anterior cervical corpectomy and fusion is a commonly used anterior approach, whilst posterior total laminectomy is a commonly used posterior surgical approach [13, 14]. These surgical methods have respective advantages and disadvantages, yet few reports have compared the efficacy of these approaches in the treatment of cervical spinal cord injury. Thus, the motivation and novelty of this study is to evaluate the efficacy of the two surgical approach on cervical spinal cord injury from multiple aspects. In this study, we retrospectively analyzed the clinical data of 180 patients with cervical spinal cord injury admitted to our hospital from June 2019 to June 2021. The curative effects of anterior subtotal vertebrectomy and posterior total laminectomy were compared, and a comprehensive comparison was made from the level of multiple postoperative indicators to provide a reference for the selection of clinical treatments and surgical methods.

2. Materials and Methods

2.1. Subjects. The clinical data from 180 patients with cervical spinal cord injury admitted to our hospital from June 2019 to June 2021 were retrospectively analyzed. The patients were divided into an anterior approach group ($n = 89$) and a posterior approach group ($n = 91$) based on the treatment method. The inclusion criteria were as follows: (1) cervical spinal cord injury based on the history of trauma, clinical manifestations and signs of patients, cervical dislocation in combination with cervical X-ray and CT examination, and MRI showing changes in spinal cord signals at different segments of the cervical spinal cord; (2) no history of cervical spine or neck surgery; (3) compete clinical data; (4) cases with follow-up times longer than 6 months. The exclusion criteria were as follows: (1) patients with severe head and chest trauma combined with other fractures; (2) patients with unstable vital signs; (3) patients with other serious diseases complicated with tumors; (4) pregnant or lactating patients; (5) patients with incomplete follow-up information.

In the anterior approach group, there were 53 males and 36 females that had an average age of 41.3 ± 6.3 years and included 42 traffic accident injuries, 28 tumbling injuries, 12 fall injuries, and 7 crush injuries. In the posterior approach group, there were 54 males and 37 females that had an average age of 42.2 ± 7.6 years and included 47 traffic accident injuries, 23 tumbling injuries, 11 fall injuries, and 10 crush injuries. The general characteristics of the two groups of patients were not significantly different, as gender ($\chi^2 = 0.0008$, $P = 0.9771$), age ($t = 0.8640$, $P = 0.3888$), and classification ($\chi^2 = 1.3223$, $P = 0.7239$).

2.2. Therapeutic Methods. Upon admission to hospital, all patients were given a neck brace for neck immobilization with symptomatic supportive treatment using drugs such as glucocorticoid and mannitol. Patients in the anterior approach group were treated with anterior subtotal vertebral resection, decompression, and internal fixation. The patients were placed in the supine position, and the surgical incision was made on the right side of the anterior part of the neck. The skin and tissues were cut layer by layer until the front of the vertebral body was reached to fully expose the responsible segment. The surgical segments were again determined based on the intraoperative C-arm X-ray results (Philips Medical Systems Nederland B.V. Medical IT), and the decompression range was again determined based on the patient's preoperative cervical CT or cervical MRI findings. After localization, the vertebral body was resected and the corresponding injured disc was resected completely. Autologous bone was implanted with titanium mesh and installed in the decompression area. The vertebral body in the decompression area was fixed with a locking internal fixation plate. A drainage tube was placed, and the incision was sutured after confirmation.

Patients in the posterior approach group were treated with posterior total laminectomy, decompression, and internal fixation. The patients were placed in the prone position, and continuous traction was performed on the skull to maintain the neutral position of the neck. In the absence of compression of both eyes, a posterior median cervical incision was made layer by layer to fully expose the vertebral plate and lateral masses. After localization, a lateral mass screw was inserted and the connecting rod was fixed. The ligamentum flavum and all vertebral plates of the corresponding segment were resected. A drainage tube was placed, and the incision was sutured after confirmation.

Patients in both groups were returned to the ward with a neck brace after surgery and underwent routine ECG monitoring, atomization of sputum, nasogastric gastrointestinal nutrition, oxygen inhalation, and sputum aspiration. Changes in the vital signs of patients were closely monitored, and particular attention was given to spinal cord function and incision drainage. In cases in which the patient had dyspnea or asphyxia, or the muscle strength of the limbs was significantly decreased or paralyzed compared to before surgery, the incision surface dressing was opened immediately to assess swelling and cyanosis on the skin around the incision and to squeeze the incision to ensure the drainage fluid was smooth.

2.3. Observational Indicators

(1) The perioperative indicators were compared between the two groups, including operation times, blood

loss, wound lengths, the length of hospital stay, and surgical costs

(2) The pain degree of patients in the two groups was compared. The visual analogue scale (VAS) [15] was used to score the degree of pain in the upper limb and neck before and after 1, 6, and 9 months of follow-up. On a scale of 0 to 10, the patients were asked to express the degree of pain according to how they felt using the following scoring criteria (0: painless; <3: mild pain, tolerable; 4-6: moderate pain, increasing but tolerable, 7-10: severe pain, increasing and unbearable)

(3) Cervical spinal cord function was evaluated using the JOA [16] scoring system before surgery and at 1, 6, and 9 months of postoperative follow-up. The items included daily activity restriction, and clinical and subjective symptoms that scored a total of 17 with lower scores indicating higher levels of patient dysfunction

(4) Improvements in cervical function were evaluated by NDI [17] before surgery and at 1, 6, and 9 months of follow-up and included 10 items (pain degree, living conditions, lifting, reading, headache, concentration, work, driving, sleep, and entertainment). Each item was allocated 5 points and the total score was 50 points with higher scores indicating higher levels of cervical spine dysfunction

(5) The ASIA spinal nerve function score [18] was used to evaluate the motor and sensory cervical nerve function before surgery and at 1, 6, and 9 months after the surgery. The score is divided into 5 grades as shown below. Grade A (1 point): complete spinal cord injury with no motor or sensory function in the sacral segment; grade B (2 points): incomplete spinal cord injury with sensory function and without motor function; grade C (3 points): incomplete spinal cord injury with sensory and motor functions and muscle strength below grade 3; grade D (4 points): Incomplete spinal cord injury with sensory and motor function, and muscle strength above grade 3; grade E (5 points): normal sensory and motor function, normal muscle strength

(6) Postoperative complications were counted including incision and lung infections, esophageal injury, hoarseness, and loosening of internal fixation

(7) Before surgery and at 9 months after surgery, the SF-36 scale [19] was used to evaluate the quality of life of patients including physiological and physical function, pain, social function, energy, emotional function, and mental and overall health. The score of each dimension was 100 points with a higher score indicating a better quality of life

2.4. Statistical Analysis. The approximate normal distribution method in PASS 15.0 software was used to calculate the sample size. A preliminary analysis of the data was performed with a two-sided test level with an α value of 0.05 and an odds ratio (OR) of 2.0. Assuming that the rate of surgical complication is 22% and 80% power is required, the sample size of each group should be at least 68 cases with 136 cases in total. Based on a 20% loss rate, each group, therefore, needed to include at least 85 cases totaling 170 cases.

SPSS 25.0 software (SPSS, Inc., Chicago, USA) was used for statistical analysis. Counting data were represented as n (%), and a chi-square test or Fisher's precise test was used for comparison between groups. Measurement data were expressed as the mean ± SD. For measurement data, a paired sample t-test was used for intragroup comparison, and a repeated ANOVA was used for intergroup comparisons. Bonferroni correction was used for post hoc multiple comparisons. P values of <0.05 were considered statistically significant.

3. Results

3.1. Comparison of Perioperative Indicators. According to the comparisons of the perioperative indicators of the two groups, there was no significant difference in the length of hospital stay, between the two groups. In contrast, the intraoperative blood loss was higher in the posterior approach group compared to the anterior approach group. Also, the wound diameters and operation times were longer, and the operation cost was higher in the posterior group than in the anterior approach group (all $P < 0.05$, Table 1).

3.2. Comparison of the Degrees of Pain. The pain degree in the two groups was compared using the VAS score. The results showed that the VAS scores in both groups had a decreasing trend. One month after surgery, the VAS score in the anterior approach group was significantly lower than that in the posterior approach group ($P < 0.05$). The VAS scores before surgery and at 6 and 9 months after surgery showed were not significantly different between the two groups (all $P > 0.05$, Figure 1).

3.3. Comparison of JOA Scores. The JOA scores of the patients in the two groups were compared before surgery and at 1, 6, and 9 months after surgery. The results showed that the JOA scores of both groups at 1, 6, and 9 months after surgery were significantly higher than before surgery (all $P < 0.05$); however, no significant difference was observed in the JOA scores between the two groups at each time point ($P > 0.05$, Table 2).

3.4. Comparison of NDI Scores. The NDI scores of the two groups were compared before surgery and at 1, 6, and 9 months after surgery. The results showed that the NDI scores at 1, 6, and 9 months after surgery in both groups were significantly lower than before surgery (all $P < 0.05$); however, no significant difference was observed in the NDI score between the two groups at any time point ($P > 0.05$, Table 3).

3.5. Comparison of ASIA Scores. The ASIA scores of the two groups of patients were compared before surgery and at 1, 6, and 9 months after surgery. No significant differences in the

Table 1: Comparison of perioperative indexes between the two groups.

	Intraoperative blood loss (mL)	Time of operation (min)	Wound diameter (mm)	Length of stay (d)	Surgery expense (CNY)
Anterior group ($n = 89$)	165.48 ± 18.93	108.18 ± 10.53	32.21 ± 2.02	9.36 ± 2.01	6908.21 ± 468.08
Posterior group ($n = 91$)	324.88 ± 44.45	115.07 ± 11.98	33.74 ± 2.84	10.00 ± 2.65	7393.50 ± 556.73
χ^2/t	31.1770	4.0949	4.1569	0.3511	76.7487
P	**<0.0001**	**0.0001**	**<0.0001**	0.0700	**<0.0001**

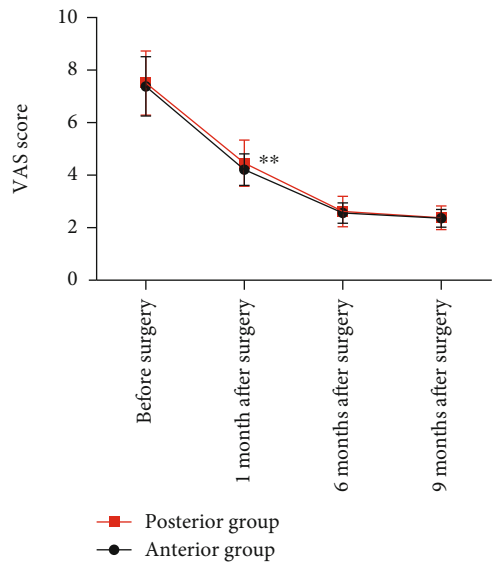


Figure 1: Comparison of the degree of pain between the two groups. (** $P < 0.01$).

ASIA score were observed between the two groups of patients at each time point ($P > 0.05$, Table 4).

3.6. Comparison of Complications. In the anterior approach group, there were 3 cases of infected incisions, 2 cases of lung infection, 2 cases of esophageal injury, 10 cases of hoarseness, and 2 cases of internal fixation loosening, with a total incidence of 21.2%. In the posterior approach group, there were 5 cases of infected incisions, 2 cases of lung infection, 5 cases of esophageal injury, 10 cases of hoarseness, and 3 cases of internal fixation loosening, with a total incidence of 27.4%. There were no significant differences in the incidence of complications between the two groups (all $P > 0.05$, Table 5).

3.7. Comparison of Quality of Life. The pre- and postoperative quality of life of the patients was scored based on physiological and physical function, bodily pain, social function, energy, emotional function, and mental and general health. The results showed no significant differences in the scores of quality of life indicators between the two groups before and at 9 months after surgery (both $P > 0.05$, Table 6).

4. Discussion

Cervical spinal cord injury is a form of central nervous system injury that involves permanent complete or partial loss of sensory functions and is mainly caused by direct or indirect injury [20]. In the case of high-energy injuries, the spinal cord is squeezed and deformed after traumatic force resulting in a series of pathological changes [21]. Currently, the treatment of cervical spinal cord injury involves surgical decompression or drug therapy. Whilst surgery cannot change the degree of the primary spinal cord injury, it can restore the stability of the cervical spine and relieve spinal cord compression to alleviate secondary spinal cord injury. This study analyzed the differences in the curative effects of anterior cervical corpectomy and fusion and posterior total laminectomy in the treatment of cervical spinal cord injury.

Our results showed no significant differences in the length of hospital stay between the two groups, yet the amount of blood loss in the posterior approach group was larger than in the anterior approach group. Also, the wound diameters and operation times in the posterior approach group were longer than in the anterior approach group. The surgical cost was higher in the posterior approach group than in the anterior approach group. In a meta-analysis to evaluate the efficacy of anterior vertebrectomy and posterior laminoplasty in the treatment of posterior longitudinal ligament calcification, it is found that in terms of operation time and bleeding volume, meta-analysis of studies showed that the operation time of posterior 1aminoplasty was shorter than that of anterior cervical corpectomy and fusion, and there was no significant difference in intraoperative bleeding volume between the two groups [22, 23].

Decompression and internal fixation of subtotal vertebral body resection is a common anterior approach operation for cervical spinal cord injury. Anterior approach surgery has the advantages of providing immediate stability, easy exposure of the surgical site, and a large range of exposure. Herniated discs and anterior margins of fractured vertebrae can be removed under direct vision, and hanging fragments can be removed without injury to the spinal cord and nerve roots. In the anterior approach surgery, the path is mostly muscle space involving the medial space of the sternocleidomastoid muscle and the medial space of carotid sheath [24] with almost no fat and loose anterior tissue. For skilled surgeons, loose fascia and space undoubtedly reduce the time required for exposure. In addition, the

Table 2: Comparison of JOA scores between the two groups.

	Before surgery	1 month after surgery	6 months after surgery	9 months after surgery
Anterior group ($n = 89$)	8.89 ± 1.03	9.92 ± 0.75*	11.35 ± 1.11*	12.55 ± 1.45*
Posterior group ($n = 91$)	9.05 ± 1.25	10.09 ± 0.99*	11.17 ± 1.26*	12.38 ± 1.41*
χ^2/t	0.9361	1.2965	1.0162	0.7975
P	0.3505	0.1965	0.3109	0.4263

Note: *$P < 0.05$ vs. the same group before surgery.

Table 3: Comparison of NDI scores between the two groups.

	Before surgery	1 month after surgery	6 months after surgery	9 months after surgery
Anterior group ($n = 89$)	25.19 ± 2.90	19.23 ± 2.26*	15.35 ± 2.14*	12.62 ± 3.29*
Posterior group ($n = 91$)	24.96 ± 3.12	18.67 ± 2.64*	15.76 ± 2.25*	13.23 ± 3.89*
χ^2/t	0.5120	1.5273	1.2522	1.1348
P	0.6093	0.1285	0.2121	0.2580

Note: *$P < 0.05$ vs. the same group before surgery.

Table 4: The ASIA score.

	Before surgery	1 month after surgery	6 months after surgery	9 months after surgery
Anterior group ($n = 89$)	2.45 ± 0.24	3.16 ± 0.31*	3.96 ± 0.25*	4.33 ± 0.44*
Posterior group ($n = 91$)	2.51 ± 0.31	3.21 ± 0.32*	4.03 ± 0.29*	4.40 ± 0.44*
t	1.4498	1.0819	1.7329	1.0672
P	0.1489	0.2808	0.0848	0.2873

Note: *$P < 0.05$ vs. the same group before surgery.

Table 5: Comparison of the incidence of complications between the two groups.

	Infected incisions	Lung infection	Esophageal injury	Hoarseness	Internal fixation loosening	Total incidence rate
Anterior group ($n = 89$)	3 (3.4)	2 (2.2)	2 (2.2)	10 (11.2)	2 (2.2)	19 (21.2)
Posterior group ($n = 91$)	5 (5.5)	2 (2.2)	5 (5.5)	10 (10.9)	3 (3.3)	25 (27.4)
χ^2/t						0.9137
P						0.3391

anterior approach has less room for exposure and, accordingly, minimal bleeding without damaging large vessels. For the posterior approach, the muscle tissue behind the vertebral body is more developed, the structure is close, the blood supply is more abundant, and there are more small vessels [11]. Although the structure is simple, it takes more time to better expose the operation site and control the amount of blood loss.

The VAS scores in patients in the posterior approach group were significantly higher than the anterior approach group one month after surgery. These data may be because the wound diameter of the posterior approach group was larger than the anterior approach group and the range of osteotomy and grooving during surgery was often larger. Subsequently, the recovery of the cervical spine and neurological function in the two groups after surgery were compared. The JOA, NDI, and ASIA scores of both groups at 1, 6, and 9 months after surgery were higher than before surgery but the differences were not statistically significant. These results indicate that both surgical methods can quickly relieve the symptoms of spinal cord compression, promote the rapid recovery of spinal cord function, and improve the cervical spine and neurological function recovery of patients.

Patients with cervical spinal cord injury are most concerned with the recovery and improvement of limb function after treatment particularly the recovery of hand function as it is closely related to life and work. It has been reported that patients with cervical spinal cord injury can improve hand function using functional exercise equipment [25] and by changing hand activity patterns after treatment [26]. In this study, no significant differences in the complication rate and

Table 6: Comparison of quality of life scores between the two groups.

	Physiological function		Physical function		Bodily pain		Social function	
	Before surgery	9 months after surgery	Before surgery	9 months after surgery	Before surgery	9 months after surgery	Before surgery	9 months after surgery
Anterior group ($n = 89$)	64.51 ± 6.29	81.33 ± 5.81*	61.12 ± 6.61	81.26 ± 6.52*	60.63 ± 5.30	80.49 ± 5.05*	64.81 ± 7.04	84.54 ± 5.95*
Posterior group ($n = 91$)	64.89 ± 7.29	82.82 ± 6.15*	62.24 ± 5.38	81.70 ± 7.85*	61.14 ± 6.35	80.53 ± 4.20*	64.03 ± 6.72	84.19 ± 5.31*
χ^2/t	0.3741	1.6701	1.3480	0.4086	0.5843	0.0578	0.7605	0.4166
P	0.7088	0.9667	0.2136	0.6833	0.5597	0.9539	0.4479	0.6775
	Energy		Emotional function		Mental health		General health	
	Before surgery	9 months after surgery	Before surgery	9 months after surgery	Before surgery	9 months after surgery	Before surgery	9 months after surgery
Anterior group ($n = 89$)	66.83 ± 6.40	83.56 ± 3.79*	69.42 ± 6.17	84.49 ± 6.96*	68.75 ± 4.33	85.26 ± 5.94*	64.20 ± 5.21	85.02 ± 4.70*
Posterior group ($n = 91$)	66.98 ± 6.40	83.82 ± 4.79*	68.35 ± 5.15	84.12 ± 5.83*	68.52 ± 5.12	85.40 ± 5.41*	63.87 ± 6.70	85.26 ± 4.24*
χ^2/t	0.1572	0.4033	1.2642	0.3870	0.3251	0.1654	0.3683	0.3599
P	0.8753	0.6872	0.2078	0.6992	0.7455	0.8688	0.7131	0.7194

Note: *$P < 0.05$ vs. the same group before surgery.

quality of life score were observed between the two groups. We also found that hoarseness occurred in a large number of cases caused mainly by edema and short-term ischemia from excessive intraoperative traction of the esophagus, trachea, and nerves. As this is a short-term injury, in most cases, the associated symptom cases will be relieved after a period of aerosol inhalation and symptomatic treatment with few permanent injuries.

Whilst our data are robust, this study had several limitations. In addition to the influence of the surgical approach on cervical spine function recovery, other physical function recoveries such as hand function were not been observed. Also, our study used a relatively small sample size. Although the observed differences are statistically significant, the long-term efficacy was not determined due to a short follow-up period. Further validation in a larger patient cohort with long-term clinical follow-up is necessary.

5. Conclusion

In conclusion, both anterior cervical corpectomy and fusion and posterior total laminectomy can effectively restore neurological function and achieve ideal therapeutic efficacy in the treatment of cervical spinal cord injury. However, the perioperative indicators of the former are better than that of the latter, and the surgical cost of the former is lower than that of the latter. Therefore, surgical methods should be selected according to the imaging findings, the general condition of patients, and individual economic status.

Data Availability

The data used for this study have been included in the manuscript.

Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work was supported by the 2020 Hebei Provincial Science and Technology Plan (20477707D). The First Affiliated Hospital of Hebei North University provides data support.

References

[1] A. M. Tarawneh, D. D'Aquino, A. Hilis, A. Eisa, and N. A. Quraishi, "Can MRI findings predict the outcome of cervical spinal cord injury? A systematic review," *European Spine Journal*, vol. 29, no. 10, pp. 2457–2464, 2020.

[2] F. Mneimneh, C. Moussalem, N. Ghaddar, K. Aboughali, and I. Omeis, "Influence of cervical spinal cord injury on thermoregulatory and cardiovascular responses in the human body: literature review," *Journal of Clinical Neuroscience*, vol. 69, pp. 7–14, 2019.

[3] M. Seif, G. David, E. Huber, K. Vallotton, A. Curt, and P. Freund, "Cervical cord neurodegeneration in traumatic and non-traumatic spinal cord injury," *Journal of Neurotrauma*, vol. 37, no. 6, pp. 860–867, 2020.

[4] T. G. Eskesen, J. S. Baekgaard, T. Peponis et al., "Cervical spinal cord injury after blunt assault: just a pain in the neck?," *The American Journal of Surgery*, vol. 217, no. 4, pp. 648–652, 2019.

[5] B. Aarabi, J. S. Albrecht, J. M. Simard et al., "Trends in demographics and markers of injury severity in traumatic cervical spinal cord injury," *Journal of Neurotrauma*, vol. 38, no. 6, pp. 756–764, 2021.

[6] C. Lo, Y. Tran, K. Anderson, A. Craig, and J. Middleton, "Functional priorities in persons with spinal cord injury: using discrete choice experiments to determine preferences," *Journal of Neurotrauma*, vol. 33, no. 21, pp. 1958–1968, 2016.

[7] W. Yisheng, Z. Fuying, W. Limin, L. Junwei, P. Guofu, and W. Weidong, "First aid and treatment for cervical spinal cord injury with fracture and dislocation," *Indian Journal of Orthopaedics*, vol. 41, no. 4, pp. 300–304, 2007.

[8] W. K. Rogers and M. Todd, "Acute spinal cord injury," *Best Practice & Research Clinical Anaesthesiology*, vol. 30, no. 1, pp. 27–39, 2016.

[9] J. H. Badhiwala, J. R. Wilson, C. D. Witiw et al., "The influence of timing of surgical decompression for acute spinal cord injury: a pooled analysis of individual patient data," *The Lancet Neurology*, vol. 20, no. 2, pp. 117–126, 2021.

[10] Z. Ma, X. Ma, H. Yang, X. Guan, and X. Li, "Anterior cervical discectomy and fusion versus cervical arthroplasty for the management of cervical spondylosis: a meta-analysis," *European Spine Journal*, vol. 26, no. 4, pp. 998–1008, 2017.

[11] Y. Zhang, H. Zhu, Z. Zhou et al., "Deep learning-based cervical spine posterior percutaneous endoscopic disc nucleus resection for the treatment of cervical spondylotic radiculopathy," *Journal of Healthcare Engineering*, vol. 2021, Article ID 7245566, 12 pages, 2021.

[12] M. G. Fehlings, A. Vaccaro, J. R. Wilson et al., "Early versus delayed decompression for traumatic cervical spinal cord injury: results of the surgical timing in acute spinal cord injury study (stascis)," *PLoS One*, vol. 7, no. 2, article e32037, 2012.

[13] C. Tatter, O. Persson, G. Burström, E. Edström, and A. Elmi-Terander, "Anterior cervical corpectomy and fusion for degenerative and traumatic spine disorders, single-center experience of a case series of 119 patients," *Operative Neurosurgery*, vol. 20, no. 1, pp. 8–17, 2021.

[14] G. M. Overdevest, W. Jacobs, C. Vleggeert-Lankamp et al., "Effectiveness of posterior decompression techniques compared with conventional laminectomy for lumbar stenosis," *Cochrane Database of Systematic Reviews*, vol. 3, article CD010036, 2015.

[15] Y.-T. Sung and J.-S. Wu, "The visual analogue scale for rating, ranking and paired-comparison (VAS-RRP): a new technique for psychological measurement," *Behavior Research Methods*, vol. 50, no. 4, pp. 1694–1715, 2018.

[16] E. Wada, M. Fukui, K. Takahashi et al., "Japanese orthopaedic association cervical myelopathy evaluation questionnaire (JOACMEQ): part 5. Determination of responsiveness," *Journal of Orthopaedic Science*, vol. 24, no. 1, pp. 57–61, 2019.

[17] M. Sterling and T. Rebbeck, "The neck disability index (NDI)," *The Australian Journal of Physiotherapy*, vol. 51, no. 4, p. 271, 2005.

[18] T. T. Roberts, G. R. Leonard, and D. J. Cepela, *Classifications in Brief: American Spinal Injury Association (ASIA) Impairment Scale*, Springer, 2017.

[19] J. T. King Jr., K. A. McGinnis, and M. S. Roberts, "Quality of life assessment with the medical outcomes study short form-36 among patients with cervical spondylotic myelopathy," *Neurosurgery*, vol. 52, no. 1, pp. 113–121, 2003.

[20] A. Nouri, J. Montejo, X. Sun et al., "Cervical cord-canal mismatch: a new method for identifying predisposition to spinal cord injury," *World Neurosurgery*, vol. 108, pp. 112–117, 2017.

[21] S. Sikka, L. Callender, S. Driver et al., "Healthcare utilization following spinal cord injury: objective findings from a regional hospital registry," *The Journal of Spinal Cord Medicine*, vol. 42, no. 2, pp. 194–200, 2019.

[22] K. Sakai, A. Okawa, M. Takahashi et al., "Five-year follow-up evaluation of surgical treatment for cervical myelopathy caused by ossification of the posterior longitudinal ligament: a prospective comparative study of anterior decompression and fusion with floating method: versus: laminoplasty," *Spine*, vol. 37, no. 5, pp. 367–376, 2012.

[23] M. Iwasaki, S. Y. Okuda, A. Miyauchi et al., "Surgical strategy for cervical myelopathy due to ossification of the posterior longitudinal ligament," *Spine*, vol. 32, no. 6, pp. 654–660, 2007.

[24] J. M. Rhee and K. L. Ju, "Anterior cervical discectomy and fusion," *JBJS Essential Surgical Techniques*, vol. 6, no. 4, p. e37, 2016.

[25] H. Hoffman, T. Sierro, T. Niu et al., "Rehabilitation of hand function after spinal cord injury using a novel handgrip device: a pilot study," *Journal of Neuroengineering and Rehabilitation*, vol. 14, no. 1, pp. 1–7, 2017.

[26] L. Britten, R. Coats, R. Ichiyama, W. Raza, F. Jamil, and S. Astill, "Bimanual reach to grasp movements after cervical spinal cord injury," *PLoS One*, vol. 12, no. 4, article e0175457, 2017.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 7099476, 13 pages
https://doi.org/10.1155/2022/7099476

Research Article

The Application Value of MRI T2*WI Radiomics Nomogram in Discriminating Hepatocellular Carcinoma from Intrahepatic Cholangiocarcinoma

Feng Huang,[1] Xiaoyun Liu,[1] Peng Liu,[1] Dan Xu,[1] Zeda Li,[1] Huashan Lin,[2] and An Xie[1]

[1]*Department of Radiology, Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University), Changsha 410005, China*
[2]*Department of Pharmaceutical Diagnosis, GE Healthcare, Changsha 410005, China*

Correspondence should be addressed to An Xie; xiean@hunnu.edu.cn

Received 14 August 2022; Accepted 12 September 2022; Published 27 September 2022

Academic Editor: Jincheng Wang

Objective. To establish and validate an MRI T2*WI-based radiomics nomogram model and to discriminate hepatocellular carcinoma (HCC) from intrahepatic cholangiocarcinoma (ICCA). *Methods*. 174 patients were retrospectively collected, who were diagnosed with primary hepatic carcinoma by surgery or puncture pathology and received preoperative MRI scans including T2*WI scans. There were 113 cases of HCC and 61 cases of mass-type ICCA. T2*WI was used for feature extraction, the extent of the lesions was manually outlined at the largest lesions layer of the T2*WI, and the feature dimension reduction was performed by the mRMR and LASSO to obtain the optimal feature set. The radiomics features and clinical risk factors were combined to establish the radiomics nomogram model. In both training and validation groups, calibration curves and ROC curves were applied to validate the efficacy of the established model. Finally, calibration curves were applied to assess the degree of fitting and DCA to assess the clinical utility of the established model. *Results*. The radiomics model had the AUC of 0.90 (95% CI, 0.85–0.96) and 0.91 (95% CI, 0.83–0.99) in the training and validation groups, respectively; the AUC of the radiomics nomogram was 0.97 (95% CI, 0.94–0.99) in the training group and 0.95 (95% CI, 0.95–0.99) in the validation group. DCA suggested the clinical application value of the nomogram model. *Conclusion*. Radiomics nomogram model based on MRI T2*WI scan without enhancement can be used to discriminate HCC from ICCA.

1. Introduction

Primary hepatic carcinoma ranked first among the malignant tumor of the liver in China. It can be classified into hepatocellular carcinoma (HCC), intrahepatic cholangiocarcinoma (ICCA), and mixed type of hepatocarcinoma. HCC has the second-highest mortality among tumors in China [1], and ICCA, a subtype of cholangiocarcinoma, is second only to HCC, accounting for 5%–10% of primary hepatic carcinoma [2]. Different types of primary hepatic carcinoma vary greatly in prognosis and responses to adjuvant therapies. Their symptoms, serum tumor markers, and imaging manifestations have differences as well as similarities. Although it has been reported that radiologic diagnostic models based on CT and MRI images can rapidly differentiate HCC from ICCA [3, 4], especially noncontrast MRI which is difficult for radiologists to correctly distinguish them accordingly, so far, there is no relevant literature report on noncontrast MRI based on T2*WI to differentiate ICCA from HCC. Therefore, it would be of great benefit to both doctors and patients to discover a method to identify different types of primary hepatic carcinoma on non-contrast-enhanced MRI scans.

MRI, with multiparametric and arbitrary layer imaging, has become one of the routine examinations for occupying liver lesions. T2*WI is a novel noninvasive functional imaging technique. It can obtain magnetic sensitivity variability between tissues and thus is used to evaluate the biological properties of tumor tissue [5]. Radiomics make it possible to extract enormous image features and transform them into

data that can quantitatively characterize tumor biology. A great deal of clinical data can be integrated to develop models in favor of clinical decision-making and tumor heterogeneity quantification, which enables noninvasive, comprehensive, and dynamic accurate treatment and prognosis prediction of diseases [6–8]. Comprehensive analysis of multiple features tends to be the greatest encouraging approach. Nomograms can parallelly study multiple features and transform complex regression equations into visualized tables that present the quantitative assignment of different predictors, which can be helpful for clinic treatment [9–11]. This research attempts to develop and validate the radiomics nomogram model based on MRI T2*WI sequences and clinical risk factors, and it is aimed at comparing its ability to discriminate HCC from ICCA.

2. Materials and Methods

2.1. Patients. Patients who were hospitalized in Hunan Provincial People's Hospital from October 2019 to December 2021 were retrospectively included. Inclusion criteria were as follows: (1) patients with HCC or ICCA confirmed pathologically by surgery or puncture biopsy; (2) MRI scans were performed 3 weeks before surgery or puncture biopsy; (3) all patients underwent a preoperative MRI scans including T2* WI sequence on the same MRI device; (4) patients who did not receive tumor-related treatment before the examination; and (5) the maximum diameter of the lesion was ≥10 mm. Exclusion criteria were as follows: (1) MR image quality could not meet the diagnostic requirements; (2) clinical information was incomplete; and (3) patients had existing intrahepatic or distant metastases. Finally, 174 patients were included in the study. 113 patients with HCC consisted of 102 males and 11 females, aged 29–80 years, with a mean age of 54.9 ± 11.4 years. 61 patients with ICCA consisted of 36 males and 25 females, aged 29–84 years, with a mean age of 58.4 ± 10.4 years. According to the ratio of 7 : 3, all cases were randomly assigned, with 123 and 51 cases in the training and validation group, respectively. The flowchart for the screening of patients is presented in Figure 1.

2.2. MRI Image Acquisition. All images were obtained by a Philips Ingenia DNA 3.0 T MR scanner with a 32-channel phased-array coil. The liver MRI scan and mDixon-Quant quantitative imaging were performed. mDixon-Quant generated T2*WI. mDixon-Quant used breath-hold 3D volume interpolated spoiled gradient recalled echo sequence, with a TR of 9.1 ms, a TE of 1.33 ms, an echo spacing of 1.3 ms, the layer thickness of 5 mm, an interval of 0 mm, a FOV of 400 mm × 350mm, and a matrix of 224 × 170.

2.3. Image Segmentation and Radiomics Feature Extraction. T2*WI were selected with the largest lesion layer in DICOM format and imported into the open-source ITK-SNAP software (version 3.8.0). The lesion extent was manually segmented by 2 physicians with 10 years of experience in diagnostic abdominal MRI (Figures 2(a)–2(d)). Thirty-two liver lesions were selected randomly to compute intraobserver and interobserver consistency on feature extraction. Two weeks later, a physician unaware of the results of tumor pathology extracted radiomics features using the same method and calculated the intraclass correlation coefficient (ICC) to assess the reproducibility of radiomics feature extraction. ICC > 0.80 indicated good consistency, and only features with ICC > 0.8 were included in feature screening and modeling. AK (Artificial Kit, GE Healthcare) software was used to extract features.

2.4. Development of Radiomics Features. Max-relevance and min-redundancy (mRMR) algorithms were used to remove redundant and irrelevant features from the radiomics features extracted from the T2*WI and retain the 20 most meaningful features. Then, the least absolute shrinkage and selection operator algorithm (LASSO) was used to further reduce the dimensionality of the 20 features to obtain the optimal feature set [12]. The linear combination of selected features was used to calculate the radiomics score for discriminating HCC from ICCA by weighting their respective coefficients. The receiver operating characteristic (ROC) curves were performed to evaluate the discriminative efficacy of the radiomics features of the two sets.

2.5. Development of Radiomics Nomogram. Clinical risk factors included gender, age, hepatitis B virus (HBV), alpha-fetoprotein (AFP), carcinoma embryonic antigen (CEA), and carbohydrate antigen 199 (CA199). Univariate and multivariate logistic regression was performed to select risk factors, and nomograms were established based on multivariate regression in the training group.

2.6. Validation and Performance Evaluation of Radiomics Nomogram. The predictive efficacy of radiomics nomogram in the training group was evaluated using the ROC curve, and the consistency between the observed and predicted results was evaluated by the calibration curve. Classification of primary hepatic carcinoma cases was in good consistency with the prediction accuracy based on radiomics nomogram when the calibration curve got close to the diagonal line. The Hosmer-Lemeshow test was performed to identify the degree of fitting of radiomics nomogram. Both the training and validation groups used the same method to validate the radiomics nomogram.

Decision curve analysis (DCA) was used to assess the clinical utility of radiomics nomogram in the validation group. ROC curve was used to calculate the area under curve (AUC), and this only considers the specificity and sensitivity of the method, while DCA quantifies the net benefits at different risk thresholds in the validation group and determines the clinical benefits of radiomics nomogram [13].

2.7. Statistical Analysis. R language (version 4.1.0) software was used to perform statistical analysis. Continuous variables were presented by median ± range interquartile. Quantitative data were compared by independent sample t-test or U-test. Categorical variables were presented by numbers and percentages and compared by the χ^2 test; ROC curve, calibration curve, and DCA were performed to assess the discrimination efficiency of the radiomics nomogram. A

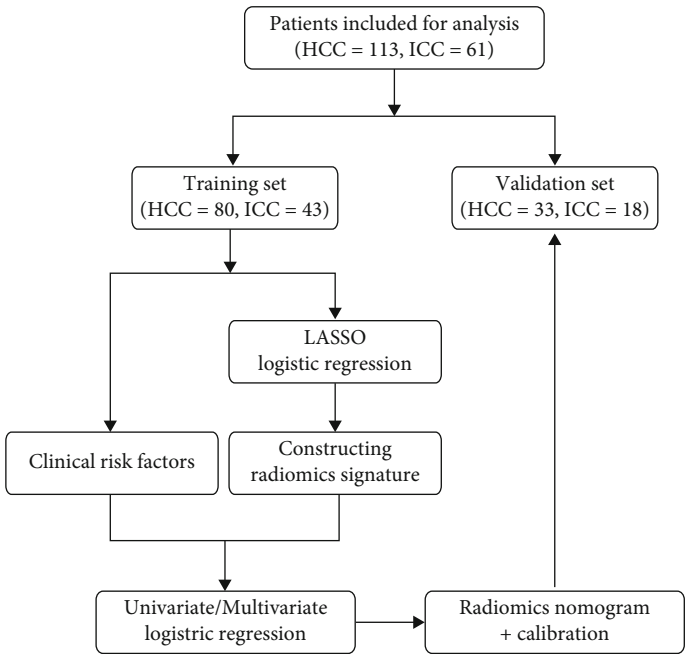


Figure 1: Workflow of the study.

two-sided $P < 0.05$ was considered that the differences were statistically significant.

3. Results

3.1. Patient Characteristics. There were 123 cases in the training group, including 80 cases of HCC and 43 cases of ICCA. There were 51 cases in the validation group, including 33 cases of HCC and 18 cases of ICCA. No significant difference was shown in the age of onset, gender, HBV, AFP, CEA, and CA199 between the cases in the training and validation groups ($P > 0.05$). This proved the reasonableness of the case assignment between the training and validation groups. There were statistically significant differences between HCC and ICCA in terms of gender, HBV, AFP, and CEA within the training and validation groups ($P < 0.05$), as detailed in Table 1. The proportion of males between HCC and ICCA showed a statistical difference (90.3% vs. 59%, $P < 0.0001$). AFP positivity rate was 60.2% in HCC and 26.2% in ICCA, with statistically significant difference ($P < 0.0001$). Only 1 of 113 HCC patients was CEA positive, and the positive rate of ICCA was 32.8%, which was statistically significant ($P < 0.0001$). The results of univariate logistic regression of HCC and ICCA are shown in Table 2.

3.2. Radiomics Feature Screening, Model Development, and Validation. Among the 932 total features, 861 had ICC > 0.8, accounting for 92.4%, with good intra- and interobserver consistency. The optimized model was calculated by the LASSO regression algorithm and the tenfold cross-validation method with the parameter λ of 0.0005 (Figure 3(a)). Eight first-order statistic features, 3 gray-level size zone matrix (GLSZM), 4 gray-level cooccurrence matrix (GLCM), and 1 gray-level run-length matrix (GLRLM) were screened, with a total of 16 nonzero coefficients of the features (Figure 3(b)). The weighting coefficients of each feature are shown in Figure 3(c). A significant difference was shown between the radiomics scores of HCC and ICCA in the training and validation groups ($P < 0.001$, Figure 3(d)). Radiomics model had AUCs of 0.90 (95% CI, 0.85–0.96) and 0.91 (95% CI, 0.83–0.99) in the training and validation groups, respectively (Figures 4(a) and 4(b)).

3.3. Development and Validation of the Radiomics Nomogram Model. The results of the multivariate logistic regression of the clinicopathological influence factors on the discrimination of HCC and ICCA are shown in Table 2. The nomogram model included radiomics features and clinical risk factors (Figure 5(a)). Radiomics nomogram exhibited an AUC of 0.97 (95% CI, 0.94–0.99) in the training group and 0.95 (95% CI, 0.91–1.00) in the validation group (Figures 4(a) and 4(b)). Figures 5(b) and 5(c) present the calibration curves. The diagonal line is the ideal model curve, the red line is the actual prediction of the model, and the better fit of the red line to the diagonal line represents the better fit of the model to the actual situation. The accuracy, sensitivity, and specificity of the different discriminative prediction models are shown in Table 3. Calibration curves using the Hosmer-Lemeshow test displayed good consistency between the true classification of HCC and ICCA and the prediction probabilities based on the radiomics nomogram model ($P = 0.983$). DCA showed a high net benefit of the model in the threshold range of 0.1–1.0 (Figure 5(d)).

4. Discussion

The surgical methods and prognosis of primary hepatic carcinoma are completely different for different pathological

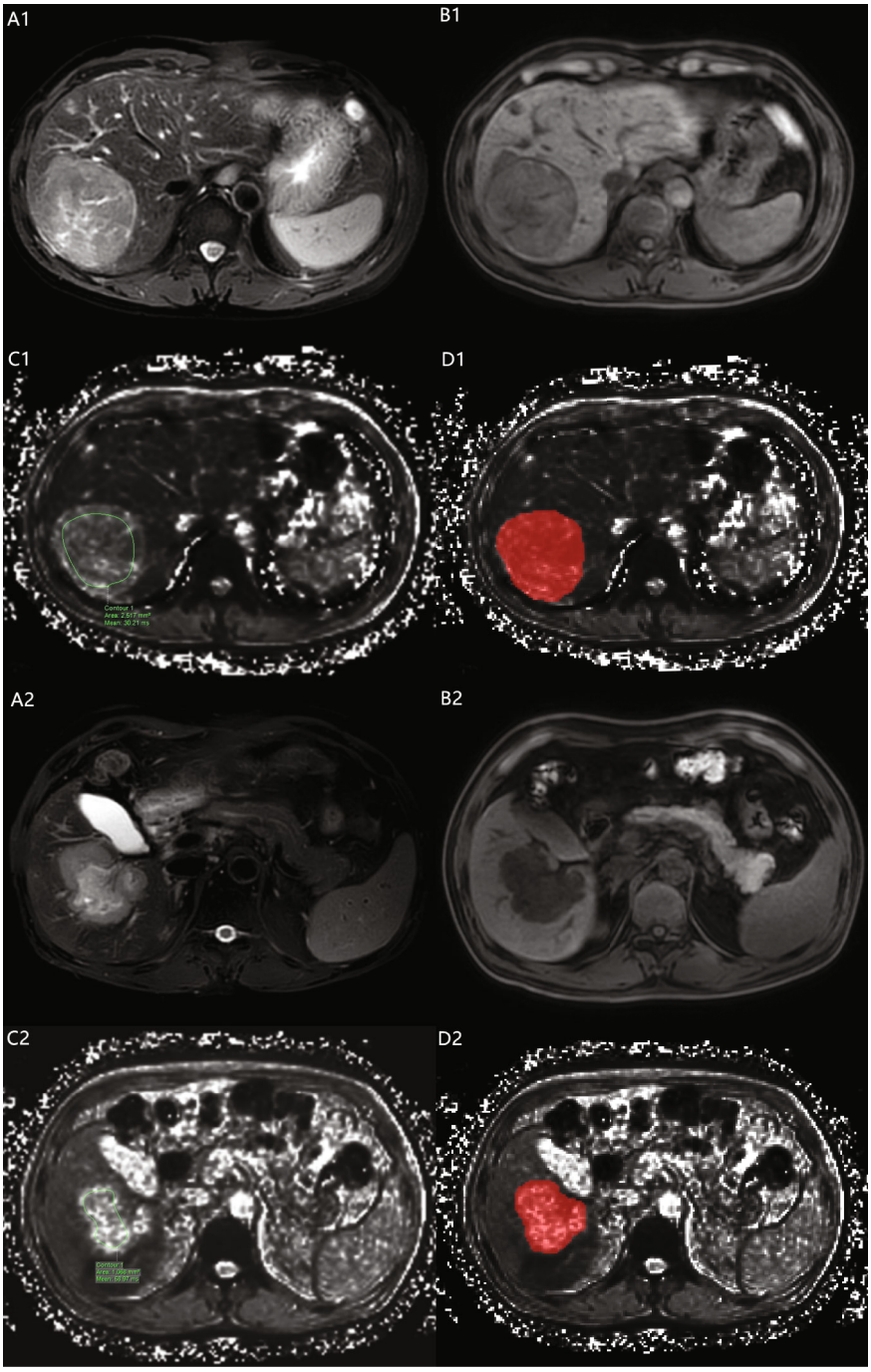


Figure 2: A1-D1 and A2-D2 show the imaging of T1WI, T2WI, T2*WI, and ROI segmentation on T2*WI in the case of HCC and ICCA, respectively. T2* value in HCC and ICCA patient is 30.21 ms and 58.97 ms.

types [14]. Kianmanesh et al. [15] reported a 5-year survival rate of 50% after surgical resection for HCC compared to 39% after radical resection for ICCA [16]. Currently, the main examination methods to discriminate between HCC and ICCA are CT and MRI [14]. However, both methods are subjective to the observer and still produce many misdiagnoses for HCC and ICCA that are smaller in size or have atypical enhancement; especially, the diagnostic accuracy is even lower when contrast enhancement is not performed. Literature has reported that the misdiagnosis rate can reach 50% in discriminating intrahepatic tumor types based only on preoperative CT versus MRI [17]. Therefore, methods with higher diagnostic efficacy need to be explored. The general data of the present study showed a significantly higher incidence of HCC in males than in females, which was consistent with the study of Pinheiro et al. [18] and might be related to the following genetic aspects: the higher incidence of hepatitis B infection in males, the gender specificity of estrogen and its receptors with HCC [19], and the higher adiponectin in females compared to males, which show

Table 1: Clinical features.

Variable	Training set ($n = 123$)			Validation set ($n = 51$)			Combined		
	HCC ($n = 80$)	ICCA ($n = 43$)	P	HCC ($n = 33$)	ICCA ($n = 18$)	P	Training	Validation	P
Gender (n (%))									
Female	8 (10.0)	16 (37.2)		3 (9.1)	9 (50.0)		24 (19.5)	12 (23.5)	
Male	72 (90.0)	27 (62.8)	0.0007	30 (90.0)	9 (50.0)	0.0032	99 (80.5)	39 (76.5)	0.6966
Age (years)									
Mean (SD)	54.1 (12.3)	59.3 (10)	0.0174	56.8 (8.8)	56.4 (11.6)	0.8818	55.9 (11.8)	56.7 (9.8)	0.6990
HBV (n (%))									
No	9 (11.2)	12 (27.9)		1 (3.0)	6 (33.3)		21 (17.1)	7 (13.7)	
Yes	71 (88.8)	31 (72.1)	0.0366	32 (97.0)	12 (66.7)	0.0098	102 (82.9)	44 (86.3)	0.7487
AFP (ng/ml)									
<20	31 (38.8)	31 (72.1)		14 (42.4)	14 (77.8)		62 (50.4)	28 (54.9)	
≥20	49 (61.2)	12 (27.9)	0.0008	19 (57.6)	4 (22.2)	0.0331	61 (49.6)	23 (45.1)	0.7088
CEA (ng/ml)									
<5	79 (98.8)	27 (62.8)		33 (100.0)	14 (77.8)		106 (86.2)	47 (92.2)	
≥5	1 (1.2)	16 (37.2)	0.0001	0 (0.0)	4 (22.2)	0.0228	17 (13.8)	4 (7.8)	0.3974
CA199 (U/ml)									
<35	57 (71.2)	24 (55.8)		20 (60.6)	14 (77.8)		81 (65.9)	34 (66.7)	
≥35	23 (28.8)	19 (44.2)	0.1280	13 (39.4)	4 (22.2)	0.3511	42 (34.1)	17 (33.3)	1.0000

Data are feature's numbers or means, with percentage in parentheses.

Table 2: Risk factors.

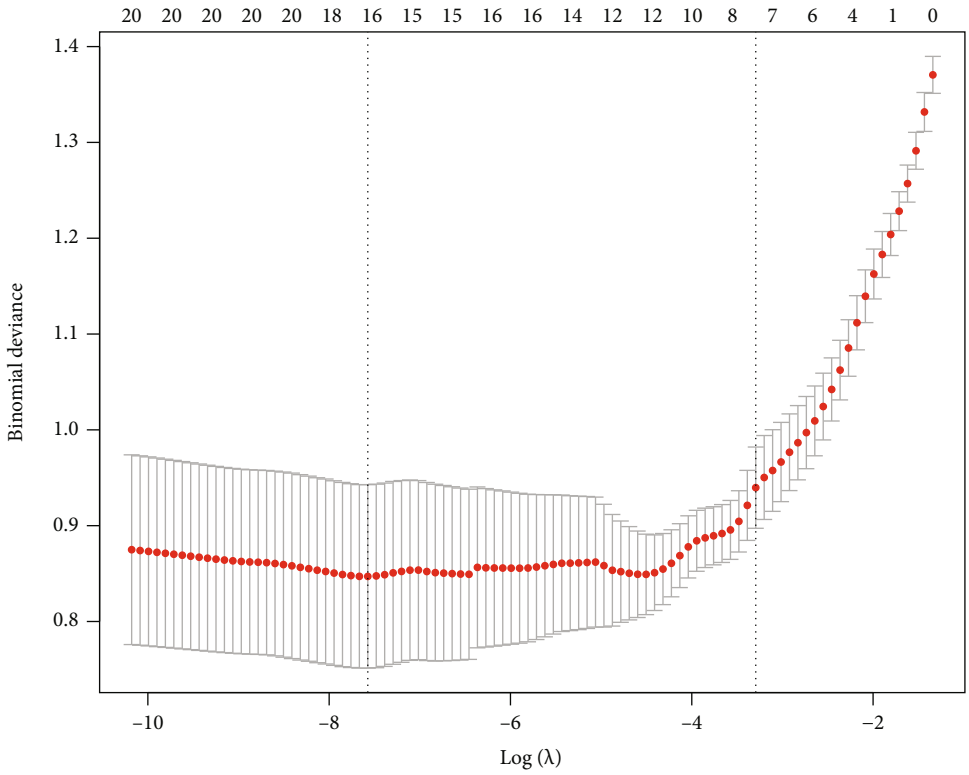
Variable	Univariate logistic regression		Multivariate logistic regression	
	OR (95% CI)	P	OR (95% CI)	P
Gender	0.1875 (0.072; 0.4882)	0.0006	0.20 (0.06; 0.64)	0.0072
Age	1.0407 (1.0058; 1.0769)	0.0219	NA	NA
HBV	0.3275 (0.1252; 0.8567)	0.0229	NA	NA
AFP	0.2449 (0.1096; 0.5472)	0.0006	0.15 (0.05-0.46)	0.0008
CEA	46.815 (5.925; 369.876)	0.0002	85.97 (9.13; 809.83)	<0.0001
CA199	1.9619 (0.9062; 4.2477)	0.0873	2.17 (0.79; 5.94)	0.1327

OR: odds ratio; NA: not available.

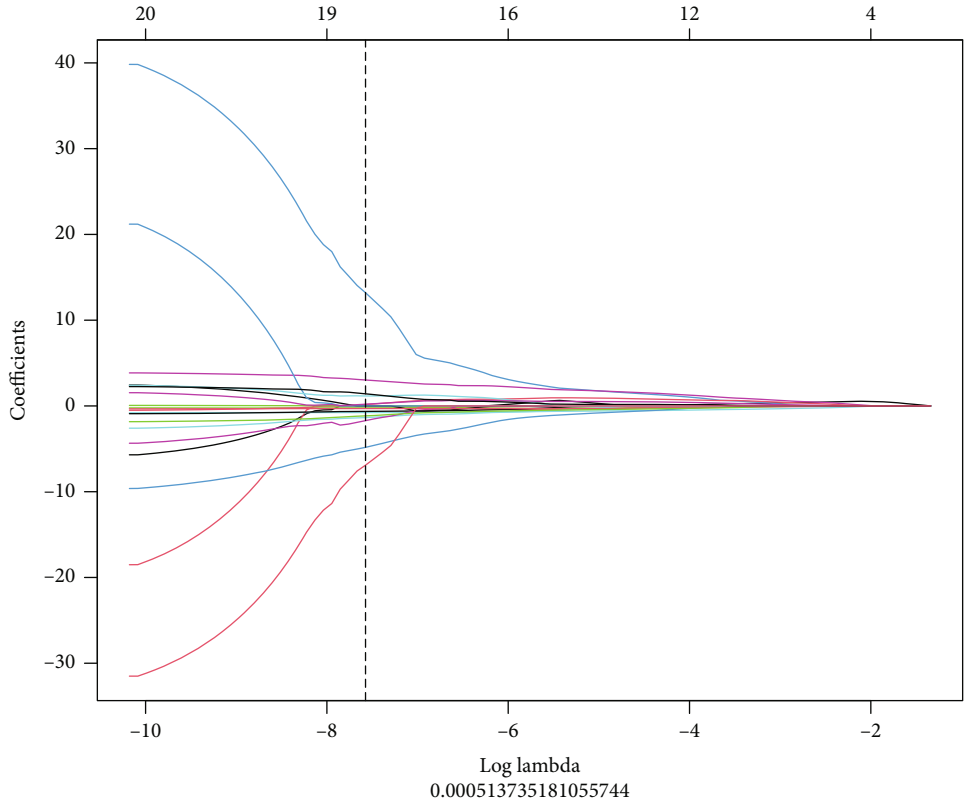
anti-HCC effect [20]. Besides, alcohol abuse and social stress in males are also associated. When pregnancy, active liver disease, and gastrointestinal tumors are excluded, AFP ≥ 400 μg/l is highly suggestive of HCC [21]. The AFP positivity rate in this study was 60.2% for HCC and 26.2% for ICCA ($P < 0.0001$), and the odds ratio (OR) and 95% CI for AFP and CEA positivity were 0.2449 and 46.815, respectively ($P < 0.006$ and 0.002). Although these tumor serum markers were statistically different, the specificity was not high [21]. The AUC for discriminating HCC with ICCA using clinical features was only 0.88 (95% CI, 0.81–0.95) in the training group and 0.83 (95% CI, 0.72–0.94) in the validation group.

The sequence of a multiecho gradient recalled echo (GRE) T2*WI is a relatively new MRI technique. It can detect the smallest changes in uniformity in the magnetic field and can improve the rate of small lesion detection. In addition, the T2* value can indirectly reflect changes in tissue biochemical components, such as iron deposits and microbleeds. Moreover, it can be used for the early diagnosis and quantitative diagnosis of some diseases [22]. HCC cells were arranged in strips separated by blood sinuses with less fibrous interstitium [14], whereas ICCA was predominantly adenocarcinoma originating from the lining epithelium of the intrahepatic bile duct and its branches to the interlobular fine bile duct tree, with cuboidal or columnar cancer cells and abundant fibrous tissue around the cancer cells, often accompanied by a dilatation of the fine bile ducts [2]. Compared to HCC, the large amount of sparse fibrous interstitium and dilated interlobular fine bile ducts in the center of ICCA promote the diffusive movement of water molecules [23–25]. This results in less random fluctuations in the interproton surrounding the magnetic environment and slower proton out of phase, prolonging the tissue T2* values. Figures 2(C1) and 2(C2) show that the ICCA signal is significantly higher than that of HCC on T2*WI.

Sun et al. [26] reported that magnetic resonance blood oxygen level parameters R2* and T2* and their associated measurements were correlated with the clinical and

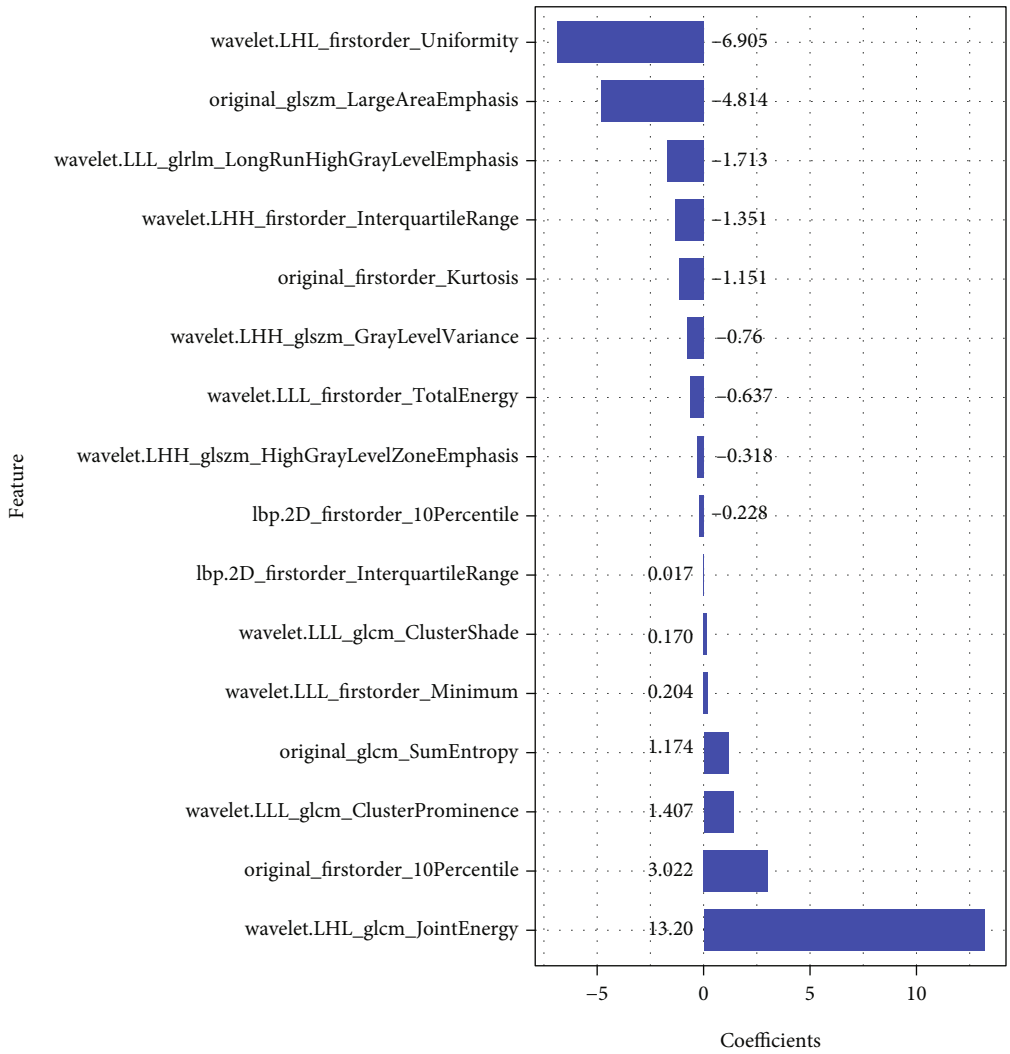


(a)



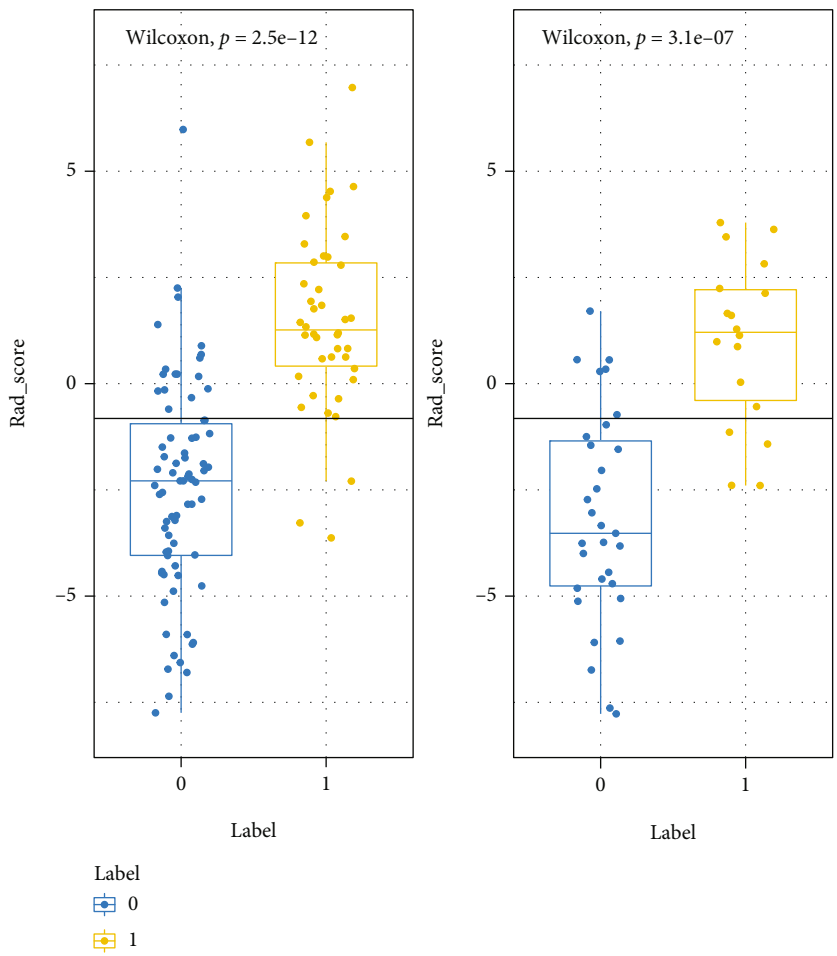
(b)

Figure 3: Continued.



(c)

Figure 3: Continued.



(d)

Figure 3: LASSO algorithm for radiomics feature selection. (a) Mean square error path using 10-fold cross-validation; (b) LASSO coefficient profiles of the radiomics features. (c) Rad-score was calculated by summing the selected features weighted by their coefficients. (d) 0 HCC and 1 ICC, rad scores from class 0 and class 1 on the training set and validation set, respectively.

pathological features of HCC. Zhang et al. [27] investigated the significance of mean platelet volume (MPV) in discriminating HCC from ICCA and found that MPV correlation discrimination had the AUC, sensitivity, and specificity of 0.698, 80.2%, and 54.1%, respectively. After the combination of sex, AFP, CA19-9, HBsAg, and MPV, the diagnostic efficiency was improved, with an AUC of 0.907, a sensitivity of 85.4%, and a specificity of 82%. The study of Zhang et al. demonstrated that the combination of biomarkers based on MPV was sufficiently accurate in differentiating HCC from ICCA. This study proposed T2*WI radiomics nomogram based on MRI plain scan and showed an AUC of 0.95, sensitivity of 81%, and specificity of 96.7% in the validation group, which exhibited better performance than the utility of mean platelet volume. The radiomics and clinical features used in the study can be acquired in a noninvasive approach before surgery, thus demonstrating good utility.

The discrimination and calibration performance of the radiomics nomogram is not representative of the clinical application value, and DCA can effectively assess the ability of the model to discriminate HCC from ICCA in clinical work. In the threshold range of 0.1–1.0, radiomics nomogram offers more net benefit than all diagnostic results for HCC or ICCA.

Based on the MRI T2*WI radiomics features and clinical features, this study comprehensively considered the clinical features and imaging features of different pathological types of primary hepatic carcinoma and used their respective weights to develop the nomogram model for discriminating HCC from ICCA. The T2* values of liver tissues are influenced not only by the tissue microenvironment but also by iron deposition, tumor tissue cystic degeneration, hemorrhage, and necrosis. Though there are significant differences in T2*7 values between HCC and ICCA, T2* values alone may be biased by the choice of observer region of interest (ROI). The use of tissue features extracted by outlining all lesion areas at the largest tumor tissue layer can better reflect the specific differences between HCC and ICCA tissues and help to accurately identify HCC and ICCA.

The present study has some limitations. First, patients came from a single center and the sample size was limited, especially the mixed type of hepatocarcinoma was not included in the study population, and a large sample multicenter study is still needed. Second, due to the low resolution

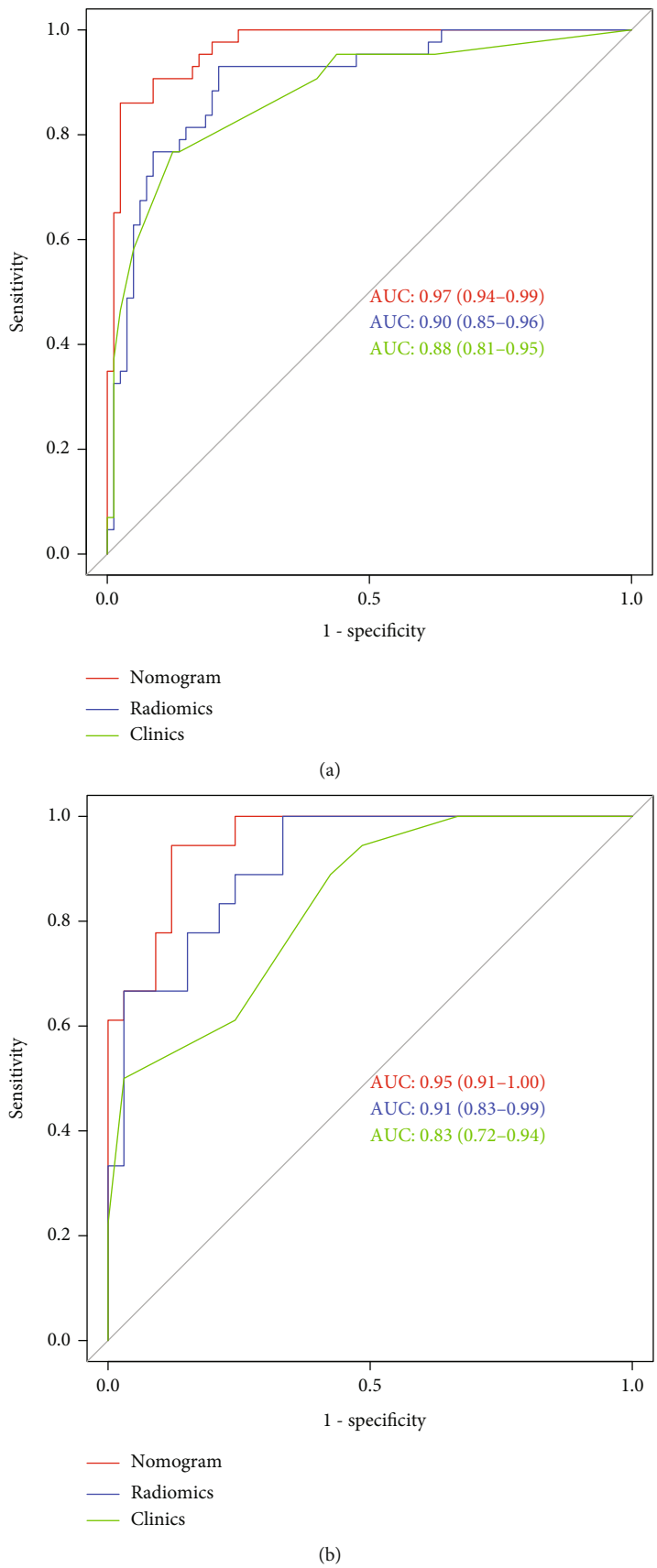


Figure 4: ROC curves of the radiomics clinics and nomogram in the training and validation sets: (a) training set; (b) validation set.

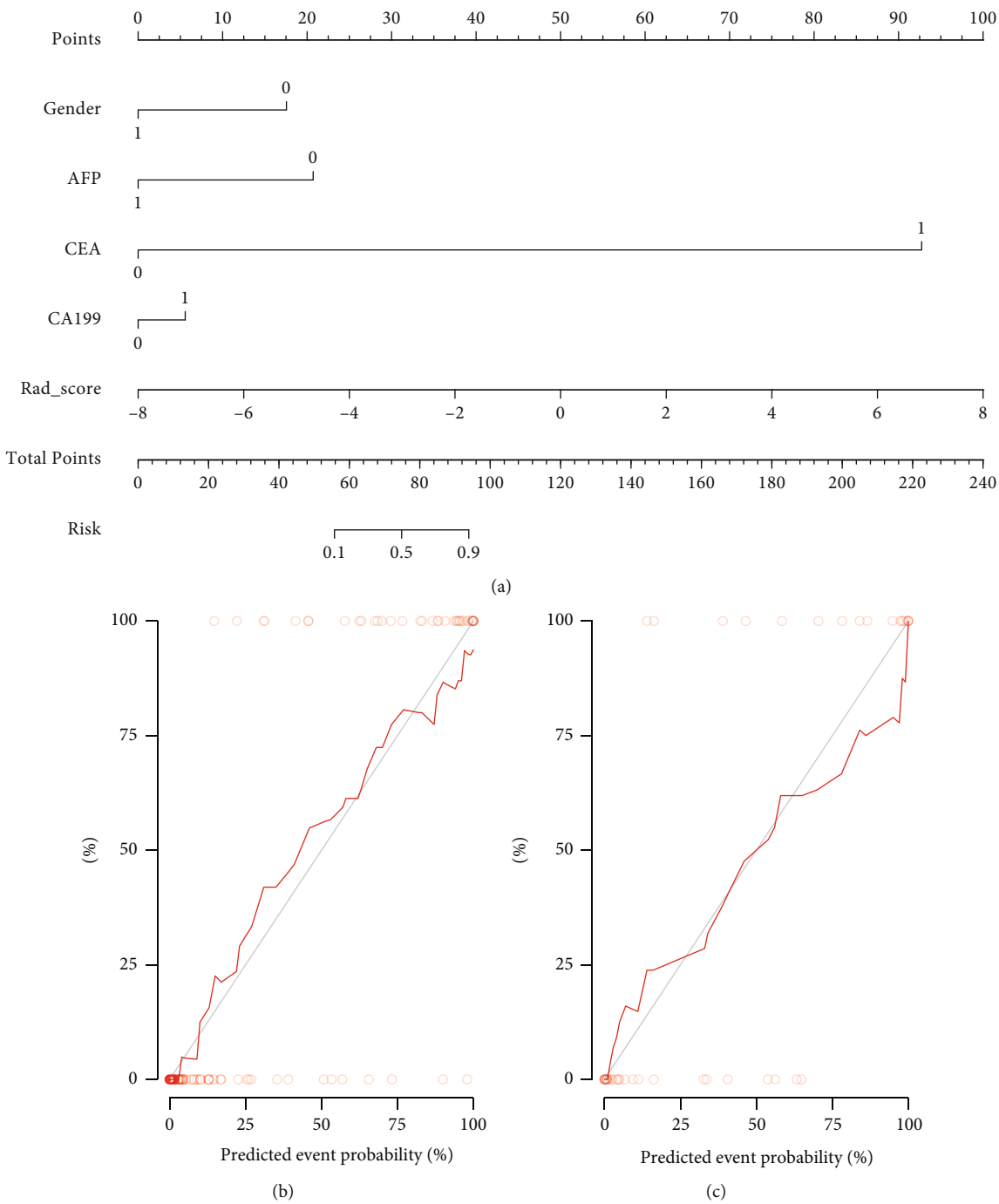
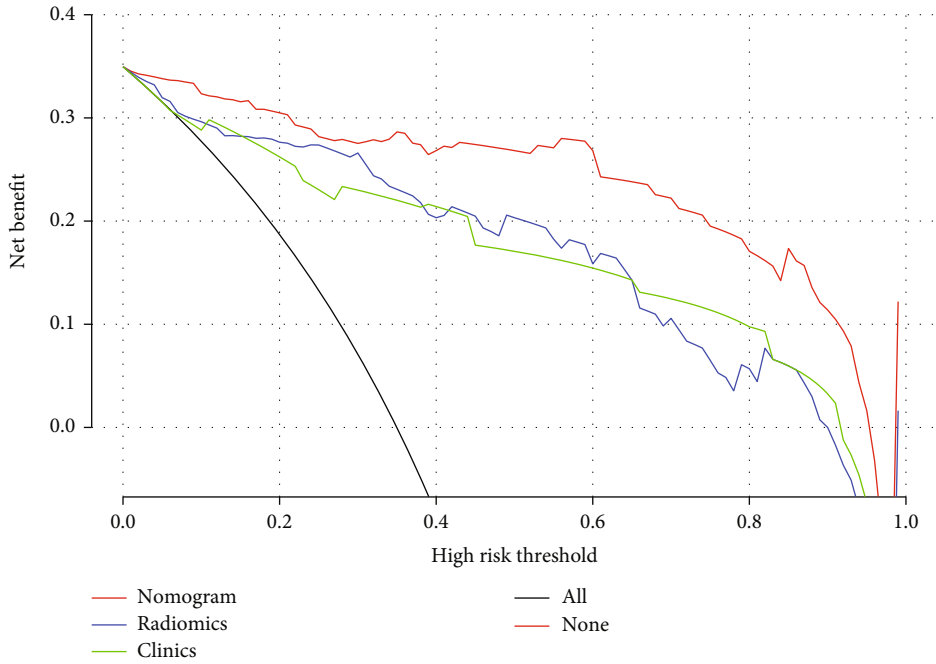


Figure 5: Continued.



(d)

Figure 5: The evaluation of the degree of fitting for the combined model and comparison of clinical utility of three models: (a) radiomics nomogram with radiomics signature and clinical factors; (b) calibration curves of the radiomics nomogram in the training; (c) validation; (d) DCA of the radiomics nomogram.

Table 3: Accuracy and predictive value between three models.

Model	Accuracy	95% CI	Sensitivity	Specificity	PPV	NPV	Cutoff
Training							
Radiomics	0.837	0.760-0.898	0.930	0.788	0.702	0.955	NA
Clinics	0.837	0.760-0.898	0.767	0.875	0.767	0.875	NA
Nomogram	0.935	0.876-0.972	0.860	0.975	0.949	0.929	NA
Validation							
Radiomics	0.804	0.668-0.902	0.778	0.818	0.700	0.871	-0.819
Clinics	0.706	0.562-0.825	0.611	0.757	0.579	0.781	-0.345
Nomogram	0.902	0.786-0.967	0.810	0.967	0.944	0.879	-0.900

CI: confidence interval; PPV: positive-predictive value; NPV: negative-predictive value.

of T2*WI, T2WI are sometimes needed to refer to when outlining the tumor tissue, which however does not affect the results of this study. To comprehensively assess the discriminative ability of HCC and ICCA in the MRI plain scans, more other imaging features can be added in future studies. Third, this study is a retrospective study, and double-blind prospective studies should be used in the future to overcome selective bias.

5. Conclusion

The nomogram model based on MRI T2*WI radiomics and clinical risk factors showed good discriminative performance in discriminating HCC from ICCA. This model uses no contrast injection and can better predict the pathological type of primary hepatic carcinoma preoperatively and help clinicians to choose the best treatment plan, which has high clinical application value. MRI T2*WI can be used to differentiate HCC from ICCA when patients have certain contraindications that do not allow enhanced MR.

Data Availability

No additional data are available.

Ethical Approval

This study was reviewed and approved by the Ethics Committee of the Hunan Provincial People's Hospital (The First Affiliated Hospital of Hunan Normal University).

Consent

Informed consent was waived due to the retrospective nature of this study. We state that we will keep patient data confidential and obey the Declaration of Helsinki.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Huang Feng and Xie An prepared the study plan, processed the data, and drafted the manuscript; Li Zeda, Liu Xiaoyun, and Xu Dan collected cases; Liu Peng reviewed and provided technical support and guidance for the intellectual content of the manuscript; Lin Huashan performed data processing. All the authors contributed to the manuscript.

Acknowledgments

We specially thanked Professor Liu Yu from the Pathology Department for his pathological support. This study was funded by the Science and Technology Project of Hunan Province, China, No. 2020SK50925, and Scientific Research Program of Hunan Provincial Health Commission, China, No. 202109011194.

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *Cancer Journal for Clinicians*, vol. 69, no. 1, pp. 7–34, 2019.

[2] J. Bridgewater, P. R. Galle, S. A. Khan et al., "Guidelines for the diagnosis and management of intrahepatic cholangiocarcinoma," *Journal of Hepatology*, vol. 60, no. 6, pp. 1268–1289, 2014.

[3] X. Xu, Y. Mao, Y. Tang et al., "Classification of hepatocellular carcinoma and intrahepatic cholangiocarcinoma based on radiomic analysis," *Computational and Mathematical Methods in Medicine*, vol. 2022, Article ID 5334095, 9 pages, 2022.

[4] X. Wang, S. Wang, X. Yin, and Y. Zheng, "MRI-based radiomics distinguish different pathological types of hepatocellular carcinoma," *Computers in Biology and Medicine*, vol. 141, article 105058, 2022.

[5] X. Han, M. Sun, M. Wang et al., "The enhanced T2 star weighted angiography (ESWAN) value for differentiating borderline from malignant epithelial ovarian tumors," *European Journal of Radiology*, vol. 118, pp. 187–193, 2019.

[6] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.

[7] P. Lambin, R. Leijenaar, T. M. Deist et al., "Radiomics: the bridge between medical imaging and personalized medicine," *Nature Reviews. Clinical Oncology*, vol. 14, no. 12, pp. 749–762, 2017.

[8] J. Wang, S. Tang, Y. Mao et al., "Radiomics analysis of contrast-enhanced CT for staging liver fibrosis: an update for image biomarker," *Hepatology International*, vol. 16, no. 3, pp. 627–639, 2022.

[9] N. Mao, Y. Dai, F. Lin et al., "Radiomics nomogram of DCE-MRI for the prediction of axillary lymph node metastasis in breast cancer," *Frontiers in Oncology*, vol. 27, no. 10, article 541849, 2020.

[10] Q. Pei, X. Yi, C. Chen et al., "Pre-treatment CT-based radiomics nomogram for predicting microsatellite instability status in colorectal cancer," *European Radiology*, vol. 32, no. 1, pp. 714–724, 2022.

[11] T. Y. Tang, X. Li, Q. Zhang et al., "Development of a novel multiparametric MRI radiomic nomogram for preoperative evaluation of early recurrence in resectable pancreatic cancer," *Journal of Magnetic Resonance Imaging*, vol. 52, no. 1, pp. 231–245, 2020.

[12] J. C. Wang, R. Fu, X. W. Tao et al., "A radiomics-based model on non-contrast CT for predicting cirrhosis: make the most of image data," *Biomarker Research*, vol. 8, no. 1, pp. 1–12, 2020.

[13] K. F. Kerr, M. D. Brown, K. Zhu, and H. Janes, "Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use," *Journal of Clinical Oncology*, vol. 34, no. 21, pp. 2534–2540, 2016.

[14] A. B. Benson, M. I. D'Angelica, D. E. Abbott et al., "Hepatobiliary cancers, version 2.2021, NCCN clinical practice guidelines in oncology," *Journal of the National Comprehensive Cancer Network*, vol. 19, no. 5, pp. 541–565, 2021.

[15] R. Kianmanesh, J. M. Regimbeau, and J. Belghiti, "Selective approach to major hepatic resection for hepatocellular carcinoma in chronic liver disease," *Surgical Oncology Clinics of North America*, vol. 12, no. 1, pp. 51–63, 2003.

[16] D. H. Kim, D. W. Choi, S. H. Choi, J. S. Heo, and A. W.-C. Kow, "Is there a role for systematic hepatic pedicle lymphadenectomy in intrahepatic cholangiocarcinoma? A review of 17 years of experience in a tertiary institution," *Surgery*, vol. 157, no. 4, pp. 666–675, 2015.

[17] T. A. Potretzke, B. R. Tan, M. B. Doyle, E. M. Brunt, J. P. Heiken, and K. J. Fowler, "Imaging features of biphenotypic primary liver carcinoma (hepatocholangiocarcinoma) and the potential to mimic hepatocellular carcinoma: LI-RADS analysis of CT and MRI features in 61 cases," *American Journal of Roentgenology*, vol. 207, no. 1, pp. 25–31, 2016.

[18] P. S. Pinheiro, K. E. Callahan, P. D. Jones et al., "Liver cancer: a leading cause of cancer death in the United States and the role of the 1945-1965 birth cohort by ethnicity," *JHEP Reports*, vol. 1, no. 3, pp. 162–169, 2019.

[19] E. Kim and P. Viatour, "Hepatocellular carcinoma: old friends and new tricks," *Experimental & Molecular Medicine*, vol. 52, no. 12, pp. 1898–1907, 2020.

[20] L. Chen, F. R. Zeng, L. Yao et al., "Nomogram based on inflammatory indices for differentiating intrahepatic cholangiocarcinoma from hepatocellular carcinoma," *Cancer Medicine*, vol. 9, no. 4, pp. 1451–1461, 2020.

[21] N. Fujiwara, S. L. Friedman, N. Goossens, and Y. Hoshida, "Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine," *Journal of Hepatology*, vol. 68, no. 3, pp. 526–549, 2018.

[22] M. Y. Tang, T. W. Chen, X. M. Zhang, and X. H. Huang, "GRE T2*-weighted MRI: principles and clinical applications," *BioMed research international*, vol. 2014, Article ID 312142, 12 pages, 2014.

[23] H. El Fattach, A. Dohan, Y. Guerrache et al., "Intrahepatic and hilar mass-forming cholangiocarcinoma: qualitative and quantitative evaluation with diffusion-weighted MR imaging,"

European Journal of Radiology, vol. 84, no. 8, pp. 1444–1451, 2015.

[24] A. J. Patterson, A. N. Priest, D. J. Bowden et al., "Quantitative BOLD imaging at 3T: temporal changes in hepatocellular carcinoma and fibrosis following oxygen challenge," *Journal of Magnetic Resonance Imaging*, vol. 44, no. 3, pp. 739–744, 2016.

[25] J. Y. Xin, S. S. Gao, J. G. Liu et al., "The value of ESWAN in diagnosis and differential diagnosis of prostate cancer: preliminary study," *Magnetic Resonance Imaging*, vol. 44, pp. 26–31, 2017.

[26] M. Y. Sun, S. Wang, Q. Song et al., "Utility of R2* obtained from T2*-weighted imaging in differentiating hepatocellular carcinomas from cavernous hemangiomas of the liver," *PLoS One*, vol. 9, no. 3, article e91751, 2014.

[27] X. Zhang, W. J. Huang, M. L. Zhang et al., "Utility of mean platelet volume in differentiating intrahepatic cholangiocarcinoma from hepatocellular carcinoma," *BMC Gastroenterology*, vol. 22, no. 1, p. 288, 2022.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 4457696, 7 pages
https://doi.org/10.1155/2022/4457696

Research Article

Efficacy and Safety Analysis of Submucosal Tunnel Endoscopic Resection for Submucosal Masses in Esophageal Muscularis Propria

Qianyi Liu, Weishan Ruan, Zhishang Liu, Jiefeng Li, and Jiayan Li

Department of Gastroenterology, Zhongshan City People's Hospital, Zhongshan 528400, Guangdong Province, China

Correspondence should be addressed to Weishan Ruan; ruanweishanrws@21cn.com

Received 13 July 2022; Revised 25 August 2022; Accepted 29 August 2022; Published 26 September 2022

Academic Editor: Jincheng Wang

Objective. To analyze the efficacy and safety of submucosal tunnel endoscopic resection (STER) for the treatment of submucosal masses in esophageal muscularis propria. *Method*. A total of 272 patients with submucosal masses in esophageal muscularis propria diagnosed and treated in our hospital from February 2019 to January 2022 were randomly selected for the study and then were randomly divided into the STER group ($n = 136$) and the endoscopic mucosal dissection (ESD) group ($n = 136$) according to the random number table method. Patients in the STER and ESD groups were treated with STER and ESD, respectively. The clinical data of patients from the two groups were collected and compared. The clinical effects and the changes of surgery-related indexes of patients after ESD and STER treatment were observed. The safety of ESD and STER was compared. The factors influencing the efficacy of STER treatment for submucosal masses in esophageal muscularis propria were analyzed. *Result*. There were significant differences between the STER group and the ESD group in terms of tumor size, lesion level, adhesion and surgical approaches ($P < 0.05$). The effective rates of ESD treatment and STER treatment were 98.53% and 88.97%, respectively. Meanwhile, the effective rates of STER treatment were significantly higher than those in the control group ($P < 0.05$). In addition, the patients in the STER group had longer operation time, less blood loss, and shorter hospital stay compared with those in the ESD group ($P < 0.05$). Adverse reactions occurred during ESD treatment and STER treatment included delayed bleeding, adhesion, perforation, and pleural effusion with the total incidence of adverse reactions of 4.41% and 13.97%, respectively. The adverse reactions in STER group were prominently less than these in the ESD group ($P < 0.05$). Logistic multivariate regression analysis showed that independent risk factors, including tumor size, lesion level, adhesion, and surgical approaches, affected the efficacy of STER in the treatment of submucosal masses in esophageal muscularis propria ($P < 0.05$). *Conclusion*. STER is an effective method for the treatment of submucosal masses in esophageal muscularis propria, which can exhibit a good effect with faster postoperative recovery and higher safety, thereby being worthy of clinical application and promotion. Tumor size, lesion level, adhesion, and surgical approaches are all related factors affecting the effect of STER treatment.

1. Introduction

Submucosal masses in esophageal muscularis propria is a kind of tumor that originated from the mucosal layer of the esophageal proper muscularis, which usually does not cause clinical symptoms. Most of them are benign lesions, and a few are gastrointestinal stromal tumors with malignant potential. As the tumor continues to grow, it will compress the surrounding organs, causing difficulty in breathing. Additionally, if the tumor has malignant features such as ulcers and erosions, it will manifest as gastrointestinal bleeding, hematemesis, and melena. With the continuous progress of medical technology, the diagnosis and diagnosis rate of diseases are gradually increasing. Due to the submucosal masses located in the esophageal muscularis propria, it is easily overlooked, thus delaying the treatment time. Surgery is the main method for the treatment of submucosal masses in esophageal muscularis propria and the main approach for diagnosing whether the disease is malignant [1, 2]. The traditional open surgery used in the past clinically

not only causes greater trauma to the patient but also has a high recurrence rate after surgery, which seriously affects the patient's life [3].

With the rapid development of minimally invasive technology and endoscopic treatment technology, endoscopic treatment has gradually been widely used in clinical practice and has been recognized by people due to its advantages of less trauma to patients and shorter postoperative recovery time. Submucosal tunneling endoscopic resection (STER) is a resection treatment mainly used for the upper gastrointestinal mucosa originating from the muscularis propria. By establishing a submucosal tunnel to remove the lesions, the integrity of the mucosa can be effectively maintained and the incidence of perforation can be reduced. During the STER process, the vascular network of the wound can be clearly displayed, which greatly reduces the bleeding caused by accidental cutting of blood vessels. Additionally, the operation of STER in the esophageal cavity can avoid the free operation of the peripheral mediastinum structure of the esophagus and reduce the possibility of damage to the tissue structure [4]. It has been reported that STER has achieved good results in the treatment of rectal carcinoid, and there is no residual disease or disease recurrence after surgery and no serious complications with a good effect [5]. However, its effect on the submucosal masses in esophageal muscularis propria remains unclear.

In this study, patients with submucosal masses in esophageal muscularis propria who were diagnosed and treated in our hospital from 2019 to 2022 were selected as the research subjects, and they were randomly divided into two groups treated with ESD and STER, respectively. The study was aimed at investigating the efficacy and safety of STER in the treatment of diseases.

2. Materials and Methods

2.1. General Materials. A total of 272 patients with submucosal masses in esophageal muscularis propria diagnosed and treated in our hospital from February 2019 to January 2022 were selected as the research subjects by a randomized, double-blind method. There were 150 males and 122 females, aged 26-72 years, with an average age of 51.58 ± 5.29 years. The following are the inclusion criteria: (1) patients with esophageal submucosal masses found through imaging detection, (2) patients with benign lesion verified through pathological examination, (3) patients with intact mucosa, (4) patients who had signed the informed consent form and actively participated in the study, (5) patients with complete clinical and pathological data and could cooperate with the study, and (6) patients over 20 years old. The following are the exclusion criteria: (1) patients with malignant lesions or distant metastasis of tumors, (2) patients with obviously abnormal liver and kidney function or heart function, (3) patients with abnormal coagulation function, (4) patients who were intolerant to surgery, and (5) patients in pregnant or lactating stage. A total of 272 patients with submucosal masses in esophageal muscularis propria were randomly divided into the STER group and the ESD group according to the random number table method, with 136 patients in each group. According to whether the treatment was effective or not, the patients were divided into an effective group and an ineffective group, including 255 cases in the effective group and 17 cases in the ineffective group. All the studies were approved by the Clinical Research and Laboratory Animal Ethics Committee of Zhongshan People's Hospital (approval number: K2020-107). The general data selection is displayed in Figure 1.

2.2. Methods

2.2.1. The ESD Group. The patients in the ESD group were treated with ESD. Blood routine examination and electrocardiogram were performed before operation. The patient was placed in lateral position and given intravenous anesthesia. The size and location of the lesion were determined by indigo carmine staining, and the lesion was marked at the edge of the lesion. 2.5 mL of indigo carmine was mixed with an appropriate amount of adrenaline in normal saline and injected at the lesion and edge marking points with the multipoint injection to make the mucosa fully uplift. The mucosa outside of the mark was completely cut open with a hook knife and dissection to ensure complete excision of the lesion with the use of coagulation or heat forceps for hemostasis. The patients underwent postoperative fasting and infection prevention and were closely monitored with vital signs.

2.2.2. The STER Group. The patients in the STER group were treated with STER. The size and location of lesions were confirmed by using endoscopic ultrasonography before surgery, and the patients were fasting before surgery. The patients were placed in a lateral position and administered with intravenous anesthesia. The esophagus was flushed with normal saline, and a mixture of indigo carmine and adrenaline in normal saline was injected for labeling. The upper and lower mucosal layers were separated, the mucosal layer was incised, a submucosal tunnel was established at a distance of about 1.5 cm from the tumor until the tumor was completely exposed, and the lesions were excavated. The tumor was separated from the muscularis propria of the esophagus, the tumor was taken out, the tunnel was flushed with sterile saline, the bleeding was stopped using electrocoagulation or thermal forceps, and the tunnel was clamped with titanium clips. The patients underwent postoperative fasting and were strictly monitored with vital signs and with postoperative antibiotics for anti-infection (Figure 2).

2.3. Clinical Pathological Data Collection. The clinical pathological data of patients were collected, including the age (≤45 years old, >45 years old), gender (male, female), tumor size (≤6 cm^3, >6 cm^3), pathological type (lipoma, leiomyoma, granulosa cell tumor, and stromal tumor), pathological level (nonintrinsic muscle layer, intrinsic muscle layer), location of onset (upper 1/3 of esophagus, middle 1/3 of esophagus, and lower 1/3 of esophagus), adhesion (yes, no), and surgical method (STER, ESD).

2.4. Outcome Measures

2.4.1. Efficacy Analysis. Endoscopy review was performed 6 months after treatment [6], including deterioration, stability, partial remission, and complete remission. Among them,

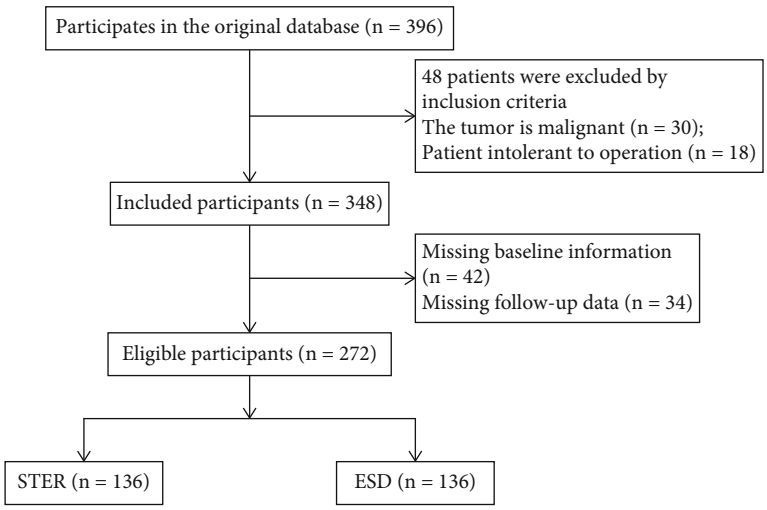


Figure 1: The process of general data selection.

manifestations that the volume of the lesions increased by more than 30% or new lesions were appeared after treatment was regarded as deterioration; manifestations that the volume of lesions decreased by less than 30% after treatment and did not reach the degree of deterioration was regarded as stable; manifestations that the volume of lesions decreased by more than 30% after treatment was regarded as partial remission; manifestations that lesions were complete obliteration was regarded as complete remission. Total effective rate of treatment = (stability + partial remission + complete remission)/total number of cases × 100%.

2.4.2. Detection of Operation-Related Indexes. The changes of indexes, including the operation time, intraoperative blood loss, and hospitalization time of ESD and STER in the treatment of submucosal masses in esophageal muscularis propria were recorded.

2.4.3. Safety Analysis. Adverse reactions, including delayed bleeding, adhesion, perforation, and pleural effusion, were recorded in the treatment of submucosal masses in esophageal muscularis propria, to compare the safety of ESD and STER treatment.

2.4.4. Analysis of Influencing Factors. The two groups of patients were divided into an effective group and an ineffective group according to the treatment effect, and the influencing factors of the effect of STER in the treatment of submucosal masses in esophageal muscularis propria were analyzed.

2.5. Statistical Analysis. The data were analyzed using SPSS 21.0 statistical software. Enumeration data, such as the univariate analysis of the efficacy and safety of STER in the treatment of submucosal masses in esophageal muscularis propria, were expressed as cases (%) and compared with the χ^2 test or Fisher's exact test. Measurement data, such as operation time, blood loss, and hospitalization time, were tested by normal distribution, which were in line with normal distribution. Measurement data were expressed as $\bar{x} \pm s$ and compared using a t-test. $P < 0.05$ indicated that the difference was statistically significant.

3. Results

3.1. Comparative Analysis of Clinical Data. There were no significant differences in age, gender, pathological type, and location of the disease between the STER group and the ESD group ($P > 0.05$), while significant differences were discovered in tumor size, lesion level, adhesion, and surgical approaches ($P < 0.05$) (Table 1).

3.2. Analysis of the Curative Effect of STER in the Treatment of Submucosal Masses in Esophageal Muscularis Propria. The effective rates of ESD treatment and STER treatment were 98.53% and 88.97%, respectively, The effective rates in the STER group were significantly higher than those in the ESD group ($P < 0.05$). (Table 2).

3.3. Comparison of Surgical Related Indexes in the Treatment of Submucosal Masses in Esophageal Muscularis Propria. Compared with those in the ESD group, the operation time was significantly prolonged, the bleeding volume was observably reduced, and the hospitalization time was notably shortened in the STER group ($P < 0.05$) (Table 3).

3.4. Safety Analysis of STER in the Treatment of Submucosal Masses in Esophageal Muscularis Propria. Adverse reactions that occurred during ESD and STER treatment included delayed bleeding, adhesions, perforation, and pleural effusion. The incidences of adverse reactions in the ESD groups were 0.74%, 3.68%, 2.21%, and 5.15%, severally, while these in the STER groups were 0.74%, 2.21%, 0.74%, and 2.94%, respectively. The total incidence of adverse reactions in the two groups was 4.41% and 13.97% separately, which was significantly lower in the STER group than in the ESD group ($P < 0.05$). (Table 4).

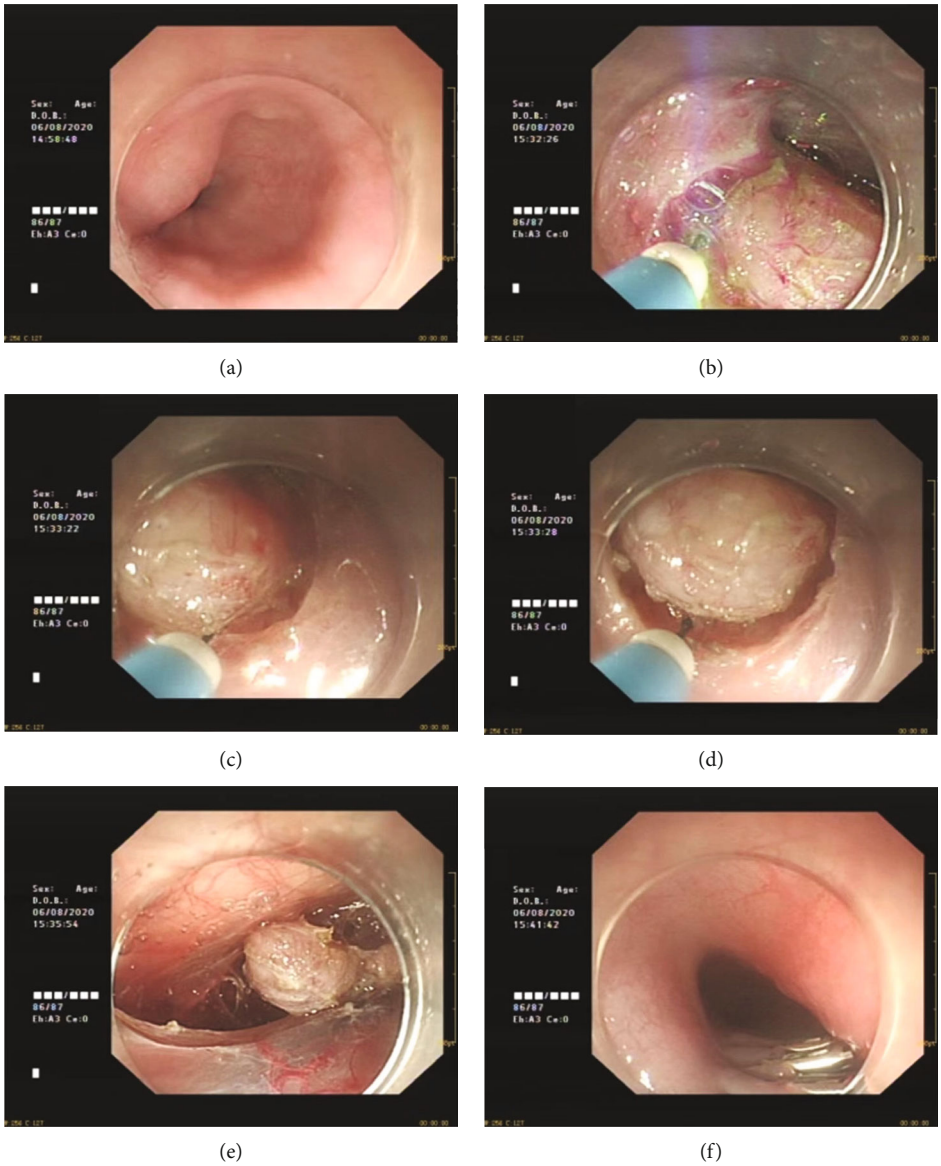


Figure 2: STER procedure and the follow-up. (a) Large tumor found under esophageal mucosa; (b) incision of esophageal mucosa; (c) tumor exposure; (d) gradual stripping; (e) complete stripping; (f) postoperative follow-up.

3.5. Multivariate Analysis of the Effect of STER in the Treatment of Submucosal Masses in Esophageal Muscularis Propria. The indicators with significant differences in univariate analysis were selected for logistic multivariate regression analysis. The results showed that tumor size, lesion level, adhesion, and surgical approaches were all independent risk factors affecting the efficacy of STER in the treatment of submucosal masses in esophageal muscularis propria ($P < 0.05$) (Table 5).

4. Discussion

Submucosal masses in esophageal muscularis propria is a relatively rare esophageal tumor in clinic. According to relevant statistics, submucosal masses in esophageal muscularis propria accounts for only about 1% of all esophageal tumors. However, most lesions in submucosal masses in esophageal muscularis propria have no obvious clinical symptoms with the manifestation of dysphagia, which leads to missed diagnosis, misdiagnosis, and delayed treatment [7, 8]. The pathological types of submucosal masses in esophageal muscularis propria include lipoma, leiomyoma, granulosa cell tumor, and stromal tumor. Among them, leiomyoma and stromal tumor are more common. Moreover, stromal tumors have a certain potential of malignant advance. The larger the volume is, the higher the degree of malignancy and the worse the prognosis is. Therefore, early diagnosis and timely treatment are of great significance. In the past, the clinical diagnosis is mainly based on imaging, endoscopic, and pathological examinations to observe the shape,

Table 1: Comparative analysis of clinical data between the STER group and the ESD group [cases (%)].

	The STER group ($n = 136$)	The ESD group ($n = 136$)	$\frac{\chi^2}{t}$	P
Age (year)			1.075	0.300
≤45	40 (29.41)	48 (35.29)		
>45	96 (70.59)	88 (64.71)		
Gender (%)			0.375	0.540
Male	75 (55.15)	80 (58.82)		
Female	61 (44.85)	56 (41.18)		
Size of the tumor (cm^3)			66.489	<0.001
≤6	114 (83.82)	48 (35.29)		
>6	22 (16.18)	88 (64.71)		
Pathological type (%)			5.895	0.117
Lipoma	1 (0.74)	0 (0.00)		
Leiomyoma	132 (97.06)	128 (94.12)		
Granulosa cell tumor	2 (1.47)	1 (0.74)		
Stromal tumor	1 (0.74)	7 (5.15)		
Lesion level (%)			77.305	<0.001
Non-intrinsic muscle layer	96 (70.59)	24 (17.65)		
Intrinsic muscle layer	40 (29.41)	112 (82.35)		
Location of onset (%)			4.517	0.104
Upper 1/3 of esophagus	16 (11.76)	8 (5.88)		
Middle 1/3 of esophagus	64 (47.06)	58 (42.65)		
Lower 1/3 of esophagus	56 (41.18)	70 (51.47)		
Adhesions (%)			70.588	<0.001
Yes	4 (2.94)	64 (47.06)		
No	132 (97.06)	72 (52.94)		
Surgical method (%)			66.104	<0.001
STER	37 (27.21)	104 (76.47)		
ESD	99 (72.79)	32 (23.53)		

Table 2: Efficacy analysis of STER in the treatment of submucosal masses in esophageal muscularis propria [cases (%)].

Groups	Cases	Complete remission	Partial remission	Stability	Deterioration	Effective rate
The STER group	136	56 (41.18)	70 (51.47)	8 (5.88)	2 (1.47)	134 (98.53)
The ESD group	136	49 (36.03)	59 (43.38)	13 (9.56)	15 (11.03)	121 (88.97)
χ^2		0.760	1.784	1.290	10.604	10.604
P		0.383	0.182	0.256	0.001	0.001

Table 3: Comparison of surgical related indexes for the treatment of submucosal masses in esophageal muscularis propria ($\bar{x} \pm s$).

Groups	Cases	Operative time (min)	Operation speed (mm^2/min)	Bleeding volume (mL)	Length of stay (d)
The STER group	136	66.38 ± 18.13	2.73 ± 0.68	2.63 ± 0.52	3.74 ± 0.42
The ESD group	136	46.24 ± 15.47	3.95 ± 0.71	4.78 ± 1.26	6.50 ± 0.74
t		9.855	14.472	18.394	37.858
P		<0.001	<0.001	<0.001	<0.001

color, and mobility of the lesions. However, the diagnosis is difficult, because the lesions are located in the submucosal mucosa [9].

With the continuous development of minimally invasive technology, resection of submucosal masses in esophageal muscularis propria by thoracoscopy is gradually applied in

Table 4: Safety analysis of STER in the treatment of submucosal masses in esophageal muscularis propria [cases (%)].

Groups	Cases	Delayed bleeding	Adhesion	Perforation	Pleural effusion	Total cases
The STER group	136	1 (0.74)	3 (2.21)	1 (0.74)	1 (0.74)	6 (4.41)
The ESD group	136	5 (3.68)	7 (5.15)	3 (2.21)	4 (2.94)	19 (13.97)
χ^2		2.727	1.661	1.015	1.834	7.444
P		0.099	0.197	0.314	0.176	0.006

Table 5: Multivariate analysis of the effect of STER in the treatment of submucosal masses in esophageal muscularis propria.

Factors	Regression coefficient	Standard error	Wald value	P value	Odds ratio	95% confidence interval (CI)	
						Lower limit	Upper limit
Size of the tumor	1.287	0.531	5.150	0.023	3.571	1.414	9.505
Lesion level	1.839	0.514	10.056	<0.001	7.245	2.048	12.015
Adhesions	1.319	0.526	6.468	0.008	4.867	1.368	10.016
Surgical approach	1.505	0.615	5.591	0.018	4.023	1.268	11.520

clinical practice. However, the application of thoracoscopy in large-area lesions is still immature, because it cannot completely remove the lesions, resulting in a high recurrence rate. ESD is a technology developed on the basis of endoscopic mucosal resection, which has a high success rate in large-area lesion resection and significantly reduces the postoperative recurrence rate. However, ESD is always accompanied with a high perforation and postoperative bleeding rate [10, 11]. The effectiveness and safety of clinical treatment are important factors affecting the promotion and application. In recent years, STER has been gradually developed, which is a method of tumor resection under the submucosal tunnel constructed between the mucosal layer and the muscularis propria. STER is developed based on ESD, but it can maintain the integrity of mucosal layer to the greatest extent, reduce the risk of infection, and prevent the outflow of postoperative fluid [12]. Studies have found that STER has shorter hospitalization time and faster postoperative recovery compared with other endoscopic resection and traditional open surgery [13]. In the study of gastric submucosal tumors, it was found that the combined estimated values of total resection and complete resection of STER were about 95.12% and 97.86%, respectively, and the combined estimated values of gas-related complications, mucosal tear, and delayed bleeding were about 8.72%, 4.20%, and 2.10%, respectively [14]. STER is a safe and effective resection of gastric submucosal tumors with few complications. In the study of gastroesophageal tumors by domestic scholars [15], by analyzing the treatment of STER and endoscopic mucosal tumor excision (ESE), it was shown that STER has a shorter operative time, less pain, less intraoperative blood loss, and faster postoperative recovery compared with ESE. STER is an effective method for the treatment of gastroesophageal muscularis propria tumors. In this study, the efficacy, surgical related indicators, and adverse reactions of ESD and STER in the treatment of submucosal masses in esophageal muscularis propria were compared. It was found that STER treatment had a higher clinical effect, a lower incidence of adverse reactions, an apparently reduced amount of bleeding, and an obviously shorter length of stay, but a longer operation time and a slower operation speed. The reason is that by establishing a tunnel, STER can effectively reduce the wound surface, protect the mucosa at the perforation, and avoid perforation of the pipe wall. Moreover, the muscle layer on the surface of the tumor can be completely stripped to fully expose through the tunnel, so as to quickly and completely peel off and avoid obvious bleeding. However, since STER needs to establish a tunnel and descend slowly, the closer the mucosal surface is, the richer the submucosal blood vessels is, resulting in a longer operation time and slower operation speed [16, 17]. Taken together, STER provides a safe and effective method for the treatment of submucosal masses in esophageal muscularis propria.

In addition, through the analysis about the relevant factors affecting the efficacy of submucosal masses in esophageal muscularis propria, it showed that the size of tumor, lesion level, adhesion, and surgical approaches were all the factors affecting the efficacy of STER in the treatment of submucosal masses in esophageal muscularis propria, and all of them were independent risk factors. Larger tumors generally tend to adopt the traditional treatment method with low treatment cost. Tumors with large volume or growing out of the cavity are difficult to be completely removed from the body, which is a high-risk factor for postoperative adverse reactions of patients [18]. The occurrence of adhesion may increase the pain of patients during the operation and affect the treatment effect. In contrast, the STER tunnel has a large operating space and a clear operating field of view. After the mucosal opening, the submucosa and the muscularis propria are separated, and a tunnel structure is formed between the submucosa and the muscularis propria. The tumor is removed under the tunnel, which can effectively prevent the injury caused by the operation during the operation [19].

In general, STER is an effective method for the treatment of submucosal masses in esophageal muscularis propria,

which can have a good effect with faster postoperative patients recover and higher safety, which is worthy of clinical application and promotion. The size of the tumor, the level of the lesion, adhesion, and surgical approaches are all related factors that affect the therapeutic effect of STER. Nevertheless, there are still some limitations in this study. Compared with ESD, the STER technology is not yet fully mature; thus, the early treatment time is relatively long. Additionally, the sample size of the study is small, and the time is still short, which may affect the results. Therefore, more cases and comparative studies confirm the safety and efficacy of this procedure in the following study.

Data Availability

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

Ethical Approval

This research was approved by the Ethics Review Committees of Affiliated Hospital of Zhongshan City People's Hospital (approval number: K2020-107).

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

Qianyi Liu and Weishan Ruan conceived and designed the project. Zhishang Liu and Jiefeng Li performed the experiments. Qianyi Liu and Jiayan Li analyzed the data. Qianyi Liu drafted the manuscript. Weishan Ruan edited and revised the manuscript. All authors have read and approved the final manuscript.

References

[1] M. Onimaru, H. Inoue, R. Bechara et al., "Clinical outcomes of per-oral endoscopic tumor resection for submucosal tumors in the esophagus and gastric cardia," *Digestive Endoscopy*, vol. 32, no. 4, pp. 328–336, 2019.

[2] C. Du, N. L. Chai, E. Q. Ling-Hu et al., "Submucosal tunneling endoscopic resection: an effective and safe therapy for upper gastrointestinal submucosal tumors originating from the muscularis propria layer," *World Journal of Gastroenterology*, vol. 25, no. 2, pp. 245–257, 2019.

[3] M. Ebi, K. Sakamoto, S. Inoue et al., "esophageal leiomyosarcoma diagnosed by endoscopic ultrasound-guided fine-needle aspiration biopsy and cured with surgical resection," *Internal Medicine*, vol. 58, no. 17, pp. 2479–2483, 2019.

[4] D. B. Hao, J. Hardwick, and J. J. Boonstra, "Endoscopic intermuscular dissection with intermuscular tunneling for local resection of rectal cancer with deep submucosal invasion," *VideoGIE*, vol. 52, no. 8, pp. E269–E270, 2022.

[5] G. X. Wang, G. Yu, Y. L. Xiang, Y. D. Miu, H. G. Wang, and M. D. Xu, "Submucosal tunneling endoscopic resection for large symptomatic submucosal tumors of the esophagus: a clinical analysis of 24 cases," *The Turkish Journal of Gastroenterology*, vol. 31, no. 1, pp. 42–48, 2020.

[6] F. Z. Xia and D. W. Lu, "Clinical effect of endoscopic tunnel technique in the treatment of esophageal-derived muscularis propria submucosal tumors," *Zhejiang Trauma Surgery*, vol. 26, no. 4, pp. 714-715, 2021.

[7] H. Wang, Y. Tan, J. Huo, and D. Liu, "Submucosal 1-tunnel endoscopic resection for treating upper gastrointestinal multiple submucosal tumor originating from the muscularis propria layer," *Medicine*, vol. 98, no. 6, pp. e14484–e14484, 2019.

[8] D. Allison and D. Rao, "Granular cell tumor confined to the muscularis propria of the appendix: a case report," *American Journal of Clinical Pathology*, vol. 154, Supplement_1, pp. S57–S58, 2020.

[9] B. Youssef, D. Asberry, and R. Mohamed, "Malignant gastrointestinal neuroectodermal tumor: a case report and a review of the literature," *American Journal of Clinical Pathology*, vol. 156, Supplement_1, pp. S66–S67, 2021.

[10] S. Tanaka, H. Kashida, Y. Saito et al., "Japan Gastroenterological Endoscopy Society guidelines for colorectal endoscopic submucosal dissection/endoscopic mucosal resection," *Digestive Endoscopy*, vol. 32, no. 2, pp. 219–239, 2020.

[11] A. Podboy, K. S. Kolahi, S. Friedland, and C. Y. Louie, "Endoscopic submucosal dissection is associated with less pathologic uncertainty than endoscopic mucosal resection in diagnosing and staging Barrett's-related neoplasia," *Digestive Endoscopy*, vol. 32, no. 3, pp. 346–354, 2020.

[12] V. Wadhwa, F. X. Franco, and T. Erim, "Submucosal tunneling endoscopic resection," *Surgical Clinics of North America*, vol. 100, no. 6, pp. 1201–1214, 2020.

[13] Y. Kim, H. Y. Jung, D. H. Kim et al., "Feasibility of endoscopic resection in gastric gastrointestinal stromal tumor," *Journal of Clinical Oncology*, vol. 38, 4_supplement, pp. 822–822, 2020.

[14] B. Cao, J. X. Lu, and Y. Y. Tan, "Efficacy and safety of submucosal tunneling endoscopic resection for gastric submucosal tumors: a systematic review and meta-analysis," *Revista Espanola de Enfermadades Digestivas*, vol. 113, no. 1, pp. 52–59, 2021.

[15] J. P. Yue and N. Hui, "Comparison of the effect of ESE and STER in the treatment of tumors of the gastroesophageal junction of the muscularis propria," *International Journal of Digestive Diseases*, vol. 40, no. 5, pp. 335–338, 2020.

[16] A. Carbajo, J. Katz, I. Andalib, R. Alkhiari, and M. Kahaleh, "Submucosal tunneling endoscopic resection of a gastric lesion: a double-sided approach," *Endoscopy*, vol. 52, no. 7, pp. E259–E260, 2020.

[17] Z. H. Geng, Y. Zhu, W. F. Chen, Q. L. Li, and P. H. Zhou, "P-O06A scoring system to support surgical decision-making for cardial submucosal tumors," *British Journal of Surgery*, vol. 10, no. 4, pp. E468–E478, 2022.

[18] V. H. Bhagat, M. Kim, and M. Ahaleh, "A review of endoscopic full-thickness resection, submucosal tunneling endoscopic resection, and endoscopic submucosal dissection for resection of subepithelial lesions," *Journal of Clinical Gastroenterology*, vol. 55, no. 4, pp. 309–315, 2021.

[19] J. Cao, Y. Chu, T. Chen et al., "Sa2049 endoscopic resection for leiomyoma of submucosal tumors originating from the muscularis propria layer in the esophagogastric junction," *Gastrointestinal Endoscopy*, vol. 91, no. 6, 2020.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8304405, 8 pages
https://doi.org/10.1155/2022/8304405

Research Article

Effects of Perinatal Cognitive Behavioral Therapy on Delivery Mode, Fetal Outcome, and Postpartum Depression and Anxiety in Women

Xiuqin Guo,[1] Xiuling Guo,[2] Ruijun Wang,[2] and Yuan Zhang[3]

[1]*College of Humanities Education, Inner Mongolia Medical University, Hohhot, Inner Mongolia 010059, China*
[2]*Obstetrics and Gynecology Department, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot, Inner Mongolia 010059, China*
[3]*School of Health Management, Inner Mongolia Medical University, Hohhot, Inner Mongolia 010110, China*

Correspondence should be addressed to Yuan Zhang; zhangyuanhoh@21cn.com

Received 15 July 2022; Revised 17 August 2022; Accepted 23 August 2022; Published 26 September 2022

Academic Editor: Hongye Wang

Objective. To explore the effects of perinatal cognitive-behavioral therapy on delivery mode, fetal outcome, and postpartum depression and anxiety in women. *Methods*. The clinical data of 88 perinatal pregnant women who came to our hospital from May 2020 to May 2021 were retrospectively analyzed and grouped into the routine group and the cognitive behavioral intervention group according to different perinatal nursing methods, with 44 cases in the cognitive behavioral intervention group received by cognitive behavioral therapy, and 44 cases in the routine group obtained by routine obstetric care during the perinatal period. The anxiety of pregnant women was evaluated by the Hamilton Anxiety Scale (HAMA). The positive cooperation and negative response of the perinatal pregnant women in two groups were recorded. The trial delivery rate, mode of delivery, amount of intrapartum bleeding, neonatal Apgar score, and visual pain simulation score (VAS) within 48 hours after delivery of the pregnant women in the two groups were also recorded and compared. The incidence of depression of pregnant and lying-in women in the two groups was recorded on the 5th and 42nd day after delivery. *Results*. After the intervention, the anxiety score of pregnant women in the cognitive behavioral intervention group was significantly lower than that in the routine group ($P < 0.05$). Following the intervention, the positive cooperation score of pregnant women in the cognitive behavioral intervention group was prominently higher than that in the routine group, and the negative cooperation score was observably lower than that in the routine group ($P < 0.05$). The rate of spontaneous delivery in the cognitive behavioral intervention group was significantly higher than that in the conventional group ($P < 0.05$), while the VAS score and blood loss in the cognitive behavioral intervention group were notably lower than those in the routine group ($P < 0.05$). The proportion of women with EPDS score <9 points, i.e., no postpartum depression both on the 5th and 42nd day after delivery, were significantly higher than those in the routine group ($P < 0.05$), whereas the proportion of patients with postpartum depression symptoms scored 9–13 points were markedly lower than those of the routine group ($P < 0.05$). *Conclusion*. The cognitive behavioral therapy can improve the adverse physiological and psychological reactions of pregnant women with perinatal anxiety disorder, enhance the natural delivery rate and postoperative recovery, reduce the risk of neonatal asphyxia, and ensure the safety of mothers and infants in the perinatal period. Compared with routine nursing, this intervention method is more targeted and scientific, and is worthy of clinical promotion.

1. Introduction

The period from 28 weeks of pregnancy to one week after delivery is called the perinatal period. As the delivery period gets closer and closer, pregnant women will have a series of physiological and psychological reactions, which are physiologically accompanied by the changes of related hormone levels and the reduction of immune function. Psychologically, pregnant women worry too much about the fetus and the outcome of delivery as the delivery period

approaches, resulting in tension, anxiety, and other adverse emotions. The adverse state may lead to endocrine mechanism disorders, and further affects multiple organ functions. In recent years, the anxiety and depression of pregnant and lying-in woman in China are increasing day by day. It has been shown that prenatal anxiety and depression increase the rate of cesarean section, lead to prolonged labor, increase the risk of postpartum hemorrhage, and even affect the development of fetal brain [1]. Therefore, it is of great significance to understand the psychological status of pregnant women and formulate relevant intervention measures for maternal and infant outcomes.

Cognitive behavioral therapy is a combination of both cognition and behavior, during which the process of cognition determines the influence of behavior, and the change of behavior can also change cognition [2]. Based on the theory of human cognitive process, cognitive behavioral therapy changes bad cognition by changing thinking, belief, and behavior, thus solving a series of physiological and psychological problems of patients. At present, cognitive behavioral therapy has been widely used in the treatment of psychological diseases [3, 4]. Some scholars have conducted a comparative study on cognitive behavioral therapy for patients with generalized anxiety disorder, and found that cognitive behavioral therapy is conducive to controlling and ameliorating anxiety disorder, improving patients' negative emotions, and the quality of life [5, 6].

This study mainly analyzed the efficacy of cognitive behavioral therapy in the treatment of perinatal anxiety symptoms, which could provide evidence-based medical proof for the clinical application of cognitive behavioral therapy.

2. Materials and Methods

2.1. General Materials. This study was approved by the Ethics Committee of our hospital. A total of 88 pregnant women with perinatal anxiety disorder who came to our hospital from May 2020 to May 2021 were selected as the research objects. The inclusion criteria are as follows: (1) all pregnant women met the diagnosis of anxiety disorder according to the mental disorder classification and diagnosis table [7], with Hamilton Anxiety Rating Scale (HAMA) [8] of ≥14 points; (2) pregnant women with no other pathological changes of body organs; (3) pregnant women were able to understand and complete the contents of the treatment without language communication barriers; (4) the pregnant women and their families agreed and signed after comprehending the study; (5) pregnant women at or above 28 weeks of gestation; and (6) first-time mothers with no history of depression at the time of enrollment. The exclusion criteria are as follows: (1) pregnant women with severe pregnancy reaction; (2) pregnant women with serious pregnancy complications and ectopic pregnancy; (3) pregnant women at less than 37 weeks of gestation; (4) pregnant women with mental disease and central nervous system disease; and (5) pregnant women who were not willing to participate in the project or could not adhere to regular postpartum follow-up. According to different perinatal nursing methods, 44 cases in the cognitive behavioral intervention group were received by cognitive behavioral therapy, while 44 cases in the routine group were obtained by routine obstetric care during the perinatal period. In the cognitive behavioral intervention group, the age of participants was between 22 and 35 years old with the average of 28.65 ± 3.91 years old, and the gestational weeks ranged from 28 weeks to one week after delivery with the average of 37.04 ± 2.25 weeks. In the routine group, the age of participants was between 22 and 37 years old with the average of 29.19 ± 4.26 years old, and the gestational weeks ranged from 29 weeks to one week after delivery with the average of 37.56 ± 3.87 weeks. No significant differences existed in age and gestational age between the two groups ($P > 0.05$). The process of general data selection is exhibited in Figure 1.

2.2. Methods. The participators in the routine group received routine obstetric nursing. According to the traditional nursing experience, we should have an intervention with the pregnant women, including explaining the delivery process, mechanism, and precautions to the pregnant women. To solve the delivery problems raised by the pregnant women, we need to record the sleep status of the pregnant women, conduct psychological counseling, and complete the corresponding psychological questionnaire.

The participators in the cognitive behavioral intervention group received cognitive behavioral therapy on the basis of the interventions as the routine group. Before nursing, an in-depth communication with pregnant and lying-in women was conducted, each pregnant and lying-in woman's cognition and existing adverse emotions were evaluated, the real psychological characteristics of pregnant women were understood, and their existing adverse thinking was supported and understood, thus providing a basis for formulating cognitive behavioral intervention plans. The specific steps were as follows.

(1) Cognitive remodeling: the face-to-face communication was done with pregnant women to understand their most concerned problems and the causes of anxiety, with full affirmation of the positive sexual behavior of pregnant women, and the negative behavior and cognition of pregnant women were corrected. Some mothers exaggerated the risks of pregnancy complications and childbirth complications, and caused fear and anxiety due to excessive worry. Health education knowledge manuals, slide shows and other easy-to-accept formats should be used to help pregnant women establish correct cognition and guide pregnant women to vent their emotions and encourage them to talk, so as to reduce the uncertainty and anxiety level of pregnant women. In addition, the nursing staff should explain a series of physiological changes in the body once childbirth begins, and describe the hazards of negative emotions, such as affecting the function of the body's air supply, resulting in the inability of immune mechanisms to fully exert their effects, and increasing the risk of postpartum hemorrhage. The nursing staff should also instruct pregnant women how to

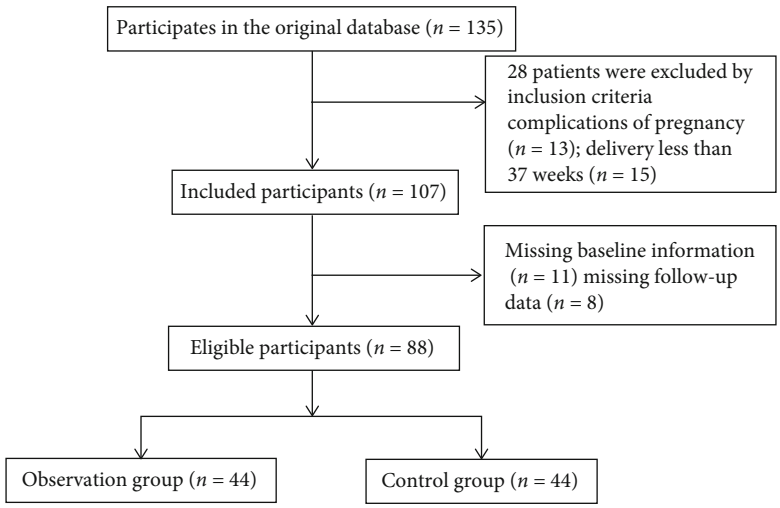


Figure 1: Process of general data selection.

self-regulate their emotions, and at the same time mobilize the maternal social support system to give full play to its support role in the stress state of childbirth, so as to improve the safety of childbirth

(2) Behavior training: a one-on-one approach was used, including attention shifting, music therapy, and progressive muscle relaxation. During training, the room should be quiet. Attention shifting method: to guide pregnant women to divert their attention through chatting, watching dramas, and other recreational activities in case of anxiety, depression, and other adverse situations, so as to achieve the purpose of alleviating anxiety. Music therapy: to guide pregnant women to play some relaxing and soothing music in case of anxiety, depression and other adverse emotions. During this process, pregnant women could be instructed to master the correct breathing method. They could sit or stand, breathe deeply to make the air reach the lungs, and then slowly exhale the gas with 10 times repetition for 5 minutes each time to relieve tension and fear, so that the pregnant and lying-in women could relax and breathe correctly after they enter the labor process.

(3) Progressive muscle relaxation: the relaxation training could be conducted using the portable skin temperature biofeedback instrument combined with progressive muscle relaxation guidance recording. The contraction and relaxation of head, neck, shoulder, chest, hands, abdomen, lower limbs, and feet were completed in the order of voice prompts, 30 minutes each time, once in the morning and once in the evening every day.

2.3. Outcome Measures

(1) The trained therapist distributed the questionnaire to the pregnant women and introduced the precautions when filling in the form, so as to ensure that the pregnant women could check their actual situation in the recent week after understanding the meaning of each item. After the survey was completed, the general data of all pregnant women could be collected, including the age, gestational week, education background, permanent residence, life satisfaction, abortion history, pregnancy preparation, gender expectation, payment method, childbirth preparation, expected childbirth method, and postpartum care

(2) The anxiety of pregnant and lying-in women was evaluated according to the HAMA scale which consisted of 2 items. Each item included 7 factors, a total of 14 factors. Each factor adopted the 5-level scoring method of 0~4 points, that was, 0 points for asymptomatic cases, 1 point for mild degree, 2 points for moderate, 3 points for severity, and 4 points for extreme weight. A score of ≥14 points was judged as mild anxiety, and a score of ≥29 points was judged as severe anxiety

(3) Before the intervention, the positive cooperation and negative response of the pregnant women in the two groups were recorded, and the cooperation of pregnant women in two groups after the intervention was compared. According to the investigation regulations, the score of "not using" intervention measures was 0, the score of "occasionally using" was 1, the score of "sometimes using" was 2, and the score of "often using" was 3. The total score of 0~32 points was positive cooperation with the investigation, and 0~26 points were a negative response

(4) Clinical data indicators including the rate of production trial from vagina, delivery mode, amount of intrapartum hemorrhage, neonatal Apgar score, and the score of maternal pain within 48 hours after delivery assessed using Visual analog scale (VAS) in the two groups were collected

(5) Evaluation of postpartum depression: 88 pregnant and lying-in women were followed up to 42 days

postpartum. The Edinburgh Postpartum Depression Scale (EPDS) with a full score of the scale of 30 points [9] was used to evaluate maternal depression at 5 days postpartum and 42 days postpartum, respectively. A score of <9 points indicated postpartum depression, a score of 9–13 points indicated postpartum depression symptoms, and a score of ≥13 points indicated postpartum depression

2.4. Statistical Analysis. The data were analyzed using SPSS 22.0 statistical software. The measurement data were expressed in ($\bar{x} \pm s$), and the enumeration data were expressed in (%). The comparison of normally distributed measurement data between the two groups was performed by independent samples t test, and the comparison before and after intervention was performed by paired t test. Wilcoxon rank sum test was used to compare the differences of measurement data that did not conform to the normal distribution. Enumeration data were analyzed using the χ^2 test or Fisher's exact test (two-tailed). $P < 0.05$ meant that the difference was statistically significant.

3. Results

3.1. Comparison of Clinical Data between the Two Groups. No significant differences existed in general data between the cognitive behavioral intervention group and the routine group ($P > 0.05$) (Table 1).

3.2. Comparison of Anxiety Scores between the Two Groups before and after Intervention. Before the intervention, there was no significant difference in perinatal anxiety scores between the two groups ($P > 0.05$). After the intervention, the pregnant women in the cognitive behavioral intervention group had lower anxiety scores than the women in the routine group ($P < 0.05$) (Table 2).

3.3. Comparison of Coordination Degree between the Two Groups before and after Intervention. No significant differences were observed in the scores of positive cooperation and negative cooperation between the two groups before intervention ($P > 0.05$). After the intervention, the pregnant women in the cognitive behavioral intervention group had a higher score of positive cooperation and a lower score of negative cooperation than the women in the routine group (both $P < 0.05$) (Table 3).

3.4. Comparison of Delivery Modes between the Two Groups. After the cognitive behavioral therapy intervention, the mode of delivery of pregnant women in the two groups was compared. The women in the cognitive behavioral intervention group had a higher spontaneous delivery rate than the women in the routine group ($P < 0.05$) (Table 4).

3.5. Comparison of Apgar Scores of Newborns between the Two Groups. After the intervention, Apgar scores of newborns in the two groups were compared. The results showed that the cognitive behavioral intervention group had a much higher proportion of newborns with Apgar scores of 8~10 than those in the routine group ($P < 0.05$) (Table 5).

3.6. Comparison of Intraoperative Bleeding Volume and Pain Score between the Two Groups. VAS and bleeding volume during delivery were compared between natural parturient in the two groups. The women in the cognitive behavioral intervention group had lower VAS score and bleeding volume than the women in the routine group ($P < 0.05$) (Table 6).

3.7. The Incidence of Depression in the Two Groups on the 5th Day after Delivery. The proportion of women with EPDS score of <9 points, i.e., no postpartum depression on the 5th day after delivery, were significantly higher than that in the routine group ($P < 0.05$), whereas the proportion of patients with postpartum depression symptoms who scored 9–13 points were markedly lower than those of the routine group ($P < 0.05$) (Table 7).

3.8. The Incidence of Depression in the Two Groups on the 42nd Day after Delivery. The proportion of women with EPDS score of <9 points on the 5th day after delivery were significantly higher than those in the routine group ($P < 0.05$), while the proportion of patients with EPDS score of 9–13 points were markedly lower than those of the routine group ($P < 0.05$) (Table 8).

4. Discussion

The perinatal period is the period from the 28th week of pregnancy to the first week of postpartum, which is a very important period before and after delivery. With the obviously increasing frequency of perinatal anxiety in pregnant women, more and more attention is paid to perinatal anxiety [10, 11]. Mood swings during pregnancy can lead to insomnia in pregnant women. Lack of sleep not only affects maternal physiological function but also leads to fetal growth retardation and even fetal deformity. Pain during childbirth is a key stressor, which can easily lead to increased cesarean section or midwifery rate, induce postpartum hemorrhage, and affect postpartum lactation. Studies have shown that the anxiety of a cesarean section is higher than that of vaginal delivery. Lack of sleep and high psychological pressure are more likely to lead to premature birth [12]. Foreign studies indicate that the incidence of antenatal anxiety is between 5.0% and 16.5%, and is still gradually increasing in recent years. Systematic and scientific perinatal nursing intervention is of great significance to promote the emotional stability of pregnant women. Therefore, it is essential to implement the necessary cognitive behavioral intervention for pregnant women with anxiety disorder.

Cognitive behavioral therapy changes people's cognition from psychological, physiological, and spiritual levels by changing people's thoughts and beliefs, so as to eliminate people's negative emotions and behaviors [14, 15]. Cognitive behavioral therapy is based on positive psychology. Clinical studies have confirmed that cognitive behavioral therapy can fully ease the anxious state of hemodialysis patients, cancer patients, and patients with generalized anxiety disorder, enhance their self-confidence, and improve treatment compliance and quality of life [16]. Cognitive behavioral therapy

Table 1: Comparison of clinical data between the two groups [n (%)].

General data	Classification	The cognitive behavioral intervention group ($n = 44$)	The routine group ($n = 44$)	χ^2	P
Age (years)	<30	15 (34.09)	16 (36.37)	0.054	0.974
	30~40	24 (54.55)	23 (52.27)		
	>40	5 (11.36)	5 (11.36)		
Gestational week (weeks)	28~33	9 (20.45)	10 (22.73)	0.322	0.852
	34~42	26 (59.10)	27 (61.36)		
	One week after delivery	9 (20.45)	7 (15.91)		
Education background	Junior high school and below	5 (11.36)	6 (13.64)	0.558	0.906
	High school and technical secondary school	12 (27.27)	14 (31.82)		
	College and undergraduate	19 (43.18)	18 (40.91)		
	Graduate and above	8 (18.18)	6(13.64)		
Permanent residence	City	22 (50.00)	21 (47.73)	0.232	0.890
	Town	16 (36.36)	18 (40.91)		
	Village	6 (13.64)	5 (11.36)		
Life satisfaction	Satisfaction	25 (56.82)	26 (59.09)		0.863
	Ordinary	16 (36.36)	14 (31.82)		
	Dissatisfaction	3 (6.82)	4 (9.09)		
Abortion history	0	28 (63.64)	25 (56.82)		0.548
	1~2	15 (34.09)	16 (36.36)		
	≥3	1 (2.27)	3 (6.82)		
Pregnancy preparation	Sufficient	19 (43.18)	18 (40.91)	0.050	0.975
	General	21 (47.73)	22 (50.00)		
	Not ready	4 (9.09)	4 (9.09)		
Gender expectation,	Boy	8 (18.18)	9 (20.45)	0.076	0.963
	Girl	6 (13.64)	6 (13.64)		
	All can be	30 (68.18)	29 (65.91)		
Payment method	At his own expense	24 (54.55)	25 (56.82)	0.046	0.830
	Medical insurance	20 (45.46)	19 (43.18)		
Childbirth preparation	Sufficient	12 (27.27)	13 (29.55)	1.863	0.393
	General	28 (63.64)	23 (52.27)		
	Not ready	4 (9.09)	8 (18.18)		
Expected childbirth method	Spontaneous labor	38 (86.36)	33 (75.00)	1.823	0.177
	Cesarean section	6 (13.64)	11 (25.00)		
Postpartum care	Husband's family	24 (54.55)	25 (56.81)		0.844
	Mother's family	18 (40.91)	16 (36.36)		
	Maternity matron	2 (4.55)	3 (6.82)		

Table 2: Comparison of anxiety scores between the two groups before and after intervention ($\bar{x} \pm s$).

Groups	Before intervention	After intervention	t	P
The cognitive behavioral intervention group ($n = 44$)	22.35 ± 4.51	10.42 ± 2.38	15.518	<0.001
The routine group ($n = 44$)	23.74 ± 2.25	18.14 ± 4.13	7.898	<0.001
t	1.829	10.743		
P	0.071	<0.001		

Table 3: Comparison of coordination degree between the two groups before and after intervention ($\bar{x} \pm s$).

Groups	Positive cooperation (score)		Negative cooperation (score)	
	Before Intervention	After Intervention	Before Intervention	After Intervention
The cognitive behavioral intervention group ($n = 44$)	2.11 ± 0.52	2.32 ± 0.47	2.21 ± 0.48	1.98 ± 0.37
The routine group ($n = 44$)	2.13 ± 0.51	1.98 ± 0.62	2.18 ± 0.53	2.25 ± 0.48
t	0.182	2.899	0.278	2.955
P	0.856	0.004	0.781	0.004

Table 4: Comparison of delivery modes between the two groups [n (%)].

Groups	Delivery modes		
	Spontaneous labor	Cesarean section group due to medical factors	Cesarean section group due to nonmedical factors
The cognitive behavioral intervention group (n = 44)	31 (70.45)	8 (18.18)	5 (11.36)
The routine group (n = 44)	23 (59.09)	6 (13.64)	15 (34.09)
χ^2		6.471	
P		0.039	

Table 5: Comparison of Apgar scores of newborns between the two groups [n (%)].

Groups	Scores of 0~3	Scores of 4~7	Scores of 8~10
The cognitive behavioral intervention group ($n = 44$)	0 (0.00)	2 (9.09)	42 (90.91)
The routine group ($n = 44$)	0 (0.00)	8 (18.18)	36 (81.82)
χ^2		4.062	
P		0.044	

Table 6: Comparison of intraoperative bleeding volume and pain score between the two groups ($\bar{x} \pm s$).

Groups	VAS (score)	Bleeding volume (mL)
The cognitive behavioral intervention group (31 cases of spontaneous labor)	5.56 ± 2.13	200.36 ± 54.21
The routine group (23 cases of spontaneous labor)	8.19 ± 1.89	289.91 ± 50.18
t	11.468	3.551
P	<0.001	0.001

Table 7: The incidence of depression in the two groups on the 5th day after delivery [n (%)].

Groups	Cases	<9	9~13	≥13
The cognitive behavioral intervention group	44	38 (86.36)	6 (13.64)	0 (0.00)
The routine group	44	29 (65.91)	14 (31.82)	1 (2.27)
χ^2		5.066	4.141	1.012
P		0.024	0.042	0.315

application helps pregnant women get rid of negative emotions in the perinatal period. The medical staffs gradually help the pregnant women correct the wrong understanding during the long-term treatment process by communicating step-by-step. Thus, the pregnant women can recognize their current problems and establish a new understanding. At the same time, pregnant women should be given timely encouragement to help them rebuild confidence and broad cognitive concepts [17]. Relevant foreign studies point out that cognitive behavioral nursing intervention improves maternal immune mechanism and enhances maternal and infant immunity by alleviating anxiety and tension [18]. The present study showed that the pregnant women in the cognitive behavioral intervention group had a much lower anxiety

Table 8: The incidence of depression in the two groups on the 42nd day after delivery [n (%)].

Groups	Cases	<9	9~13	≥13
The cognitive behavioral intervention group	44	36 (81.82)	7 (15.91)	1 (2.27)
The routine group	44	23 (52.27)	13 (29.55)	8 (18.18)
χ^2		8.692	2.329	6.065
P		0.003	0.127	0.014

score than the women in the routine group after the intervention, indicating that cognitive behavior intervention therapy could alleviate the anxiety of pregnant women. The reason might be that cognitive behavioral therapy helped the pregnant women fully understood the delivery knowledge and delivery complications during the perinatal period. At the same time, targeted strategies for eliminating the negative thinking of the pregnant women were conducted to make them master emotion regulation method and improve the quality of delivery.

The present study showed that the pregnant women in the cognitive behavioral intervention group had a higher score of positive cooperation and a lower score of negative cooperation than the women in the routine group after the intervention. It can be seen that cognitive behavioral therapy has greatly mobilized the enthusiasm of pregnant women to actively participate in the delivery process and actively deal with the delivery pain and uncertain delivery outcome. Some scholars have found that cognitive behavioral therapy can not only ease maternal anxiety and depression, but also improve the rate of spontaneous delivery and breast-feeding and reduce the incidence of neonatal asphyxia during delivery [19]. The results of this study showed that the participants in the cognitive behavioral intervention group had a much higher natural spontaneous labor rate and proportion of neonatal Apgar scores of 8~10 than the routine group, suggesting that cognitive behavioral therapy could improve the rate of spontaneous labor and reduce the risk of neonatal asphyxia. The reason may be that cognitive behavioral therapy not only lay emphasis on the health education of maternal delivery knowledge but also emphasize the benefits of delivery mode on breast milk and postpartum, which could both promote the health of newborns [20]. This present study also showed that women in the cognitive behavioral intervention group had much lower VAS score and bleeding volume than in the routine group. It may be that the cerebral cortex is suppressed by anxiety, resulting in reduced oxytocin secretion, weakening uterine contraction, and increased bleeding during delivery [21]. The results of this study also showed that the proportion of patients without postpartum depression at 5 and 42 days postpartum in the cognitive behavioral intervention group were significantly higher than those in the routine group, and the proportion of pregnant women diagnosed with depression at 42 days postpartum were significantly lower than that in the routine group, further illustrating that cognitive behavioral therapy can reduce the incidence of postpartum depression while improving maternal prenatal anxiety.

In conclusion, cognitive behavioral therapy can improve the adverse physiological and psychological reactions of pregnant and lying-in women with perinatal anxiety disorder, improve the natural delivery rate and postoperative recovery, reduce the risk of neonatal asphyxia, ensure the safety of mothers and infants in the perinatal period, and be consistent with routine nursing. In comparison with routine nursing, this intervention method is more targeted and scientific, and is worthy of clinical promotion. In this study, the prenatal anxiety status of pregnant women was screened according to the definition of the perinatal period, but the pregnancy period was divided into early, middle, and late pregnancy. Therefore, the anxiety status of the whole prenatal period in this study does not represent the anxiety status of pregnant women in rectifying the prenatal period. A large sample, multicenter, long-term follow-up study is needed throughout pregnancy.

Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

Ethical Approval

This research was approved by the Ethics Review Committees of Inner Mongolia Medical University and conducted according to the Declaration of Helsinki.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

Xiuqin Guo and Yuan Zhang conceived and designed the project. Xiuqin Guo, Xiuling Guo, and Ruijun Wang performed the experiments. Xiuqin Guo and Xiuling Guo analyzed the data. Xiuqin Guo drafted the manuscript. Yuan Zhang edited and revised the manuscript. All authors have read and approved the final manuscript.

Acknowledgments

This study was funded by The Inner Mongolia Autonomous Region Higher Education Science and Technology Research Project (NJSY21623) and The 2021 Inner Mongolia Autonomous Region Higher Education Scientific Research Project (NJZY21603).

References

[1] S. M. Green, E. Donegan, R. E. McCabe et al., "Cognitive behavior therapy for women with generalized anxiety disorder in the perinatal period: impact on problematic behaviors," *Behavior Therapy*, vol. 52, no. 4, pp. 907–916, 2021.

[2] A. Peipert, N. Rodriguez-Quintana, and L. Lorenzo-Luaces, "Outcomes of student trainee-delivered cognitive behavioral therapy (CBT) on internalizing symptoms, CBT skills, and life satisfaction," *International Journal of Cognitive Therapy*, vol. 15, no. 1, pp. 94–113, 2022.

[3] A. Pozza and D. Dèttore, "Modular cognitive-behavioral therapy for affective symptoms in young individuals at ultra-high risk of first episode of psychosis: randomized controlled trial," *Journal of Clinical Psychology*, vol. 76, no. 3, pp. 392–405, 2020.

[4] J. Reavell, M. Hopkinson, D. Clarkesmith, and D. A. Lane, "Effectiveness of cognitive behavioral therapy for depression and anxiety in patients with cardiovascular disease: a systematic review and meta-analysis," *Psychosomatic Medicine*, vol. 80, no. 8, pp. 742–753, 2018.

[5] L. R. Conklin, A. J. Curreri, T. J. Farchione, and D. H. Barlow, "Homework compliance and quality in cognitive behavioral therapies for anxiety disorders and obsessive-compulsive disorder," *Behavior Therapy*, vol. 52, no. 4, pp. 1008–1018, 2021.

[6] D. David, C. Cotet, S. Matu, C. Mogoase, and S. Stefan, "50 years of rational-emotive and cognitive-behavioral therapy: a systematic review and meta-analysis," *Journal of Clinical Psychology*, vol. 74, no. 3, pp. 304–318, 2018.

[7] D. A. Kalmbach, P. Cheng, L. M. O'Brien et al., "0470 A randomized controlled trial of digital cognitive behavioral therapy for insomnia in pregnant women," *Sleep medicine*, vol. 43, Supplement_1, p. A180, 2020.

[8] R. M. Butler, E. B. O'Day, and R. G. Heimberg, "The benefits of a longer course of cognitive behavioral therapy for some patients with social anxiety disorder," *Cognitive Behaviour Therapy*, vol. 50, no. 3, pp. 351–365, 2020.

[9] J. R. Carl, C. B. Miller, A. L. Henry et al., "Efficacy of digital cognitive behavioral therapy for moderate-to-severe symptoms of generalized anxiety disorder: a randomized controlled trial," *Depression and Anxiety*, vol. 37, no. 12, pp. 1168–1178, 2020.

[10] S. Pearl and P. J. Norton, "Transdiagnostic cognitive behavioural therapy for individuals with anxiety disorders: an open trial," *Behaviour Change*, vol. 38, no. 2, pp. 1–11, 2020.

[11] B. R. Marriott, E. Cho, S. K. Tugendrajch et al., "Role-play assessment of therapist adherence and skill in implementation of trauma-focused cognitive-behavioral therapy," *Administration and Policy in Mental Health and Mental Health Services Research*, vol. 49, no. 3, pp. 374–384, 2022.

[12] K. Cohen, L. Leibovich, R. Markin, and S. Zilcha-Mano, "The process of change in psychotherapy with a pregnant patient following perinatal losses: an analysis of a case study," *Journal of Clinical Psychology*, vol. 75, no. 5, pp. 874–885, 2019.

[13] Y. Cao, J. Liu, Y. Zhang, Y. Li, Z. Chen, and J. Lu, "Pregnant women's psychological state and influence factors: anxiety, and depression during COVID-19 outbreak," *Journal of Perinatal Medicine*, vol. 49, no. 6, pp. 664–673, 2021.

[14] S. Masoume, Z. Zahra, and B. Alireza, "Efficacy of mindfulness-based cognitive therapy and cognitive behavioral therapy for anxiety, depression, and fatigue in cancer patients: a randomized clinical trial," *Iranian Journal of Psychiatry*, vol. 16, no. 3, pp. 271–280, 2021.

[15] Y. N. Li, N. Buys, S. Ferguson, Z. J. Li, and J. Sun, "Effectiveness of cognitive behavioral therapy-based interventions on health outcomes in patients with coronary heart disease: a meta-analysis," *World Journal of Psychiatry*, vol. 11, no. 11, pp. 1147–1166, 2021.

[16] M. Kashani Lotfabadi, S. S. Aboozarian, M. Farnosh, A. Mohammadi, B. M. Hashemi, and A. Salarhaji, "The effectiveness of cognitive-behavioral group therapy on symptoms of depression and anxiety in women with chronic schizophrenia admitted to Hejazi hospital in Mashhad," *Avicenna Journal of Nursing and Midwifery Care*, vol. 29, no. 2, pp. 146–159, 2021.

[17] L. Giani, M. Caputi, B. Forresi, G. Michelini, and S. Scaini, "Evaluation of cognitive-behavioral therapy efficacy in the treatment of separation anxiety disorder in childhood and adolescence: a systematic review of randomized controlled trials," *International Journal of Cognitive Therapy*, vol. 15, no. 1, pp. 57–80, 2022.

[18] D. H. Rosmarin, E. S. Bocanegra, G. Hoffnung, and M. Appel, "Effectiveness of cognitive behavioral therapy for anxiety and depression among Orthodox Jews," *Cognitive and Behavioral Practice*, vol. 26, no. 4, pp. 676–687, 2019.

[19] S. Khatibi, M. Kazemi, S. Zenoozian, M. M. Vakili, and R. Kharaghani, "Cognitive behavioral counseling and sleep quality in pregnant women: a randomized controlled Trial," *Current Sleep Medicine Reports*, vol. 7, no. 3, pp. 120–128, 2021.

[20] J. Apolinário-Hagen, M. Drüge, and L. Fritsche, "Cognitive behavioral therapy, mindfulness-based cognitive therapy and acceptance commitment therapy for anxiety disorders: integrating traditional with digital treatment approaches," *Advances in Experimental Medicine and Biology*, vol. 1191, pp. 291–329, 2020.

[21] S. Stefan, I. A. Cristea, A. Szentagotai Tatar, and D. David, "Cognitive-behavioral therapy (CBT) for generalized anxiety disorder: contrasting various CBT approaches in a randomized clinical trial," *Journal of Clinical Psychology*, vol. 75, no. 7, pp. 1188–1202, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 5005195, 7 pages
https://doi.org/10.1155/2022/5005195

Research Article

Predictive Models for Knee Pain in Middle-Aged and Elderly Individuals Based on Machine Learning Methods

Lu Liu,[1,2] Min-min Zhu,[1,2] Lin-lin Cai,[1,2] and Xiao Zhang[1,2]

[1]*Department of Anesthesiology, The Affiliated Wuxi NO.2 People's Hospital of Nanjing Medical University, Wuxi, Jiangsu, China*
[2]*Department of Anesthesiology, Wuxi NO.2 People's Hospital, Wuxi, Jiangsu, China*

Correspondence should be addressed to Xiao Zhang; xiaozhang0810@njmu.edu.cn

Received 26 July 2022; Revised 1 September 2022; Accepted 3 September 2022; Published 26 September 2022

Academic Editor: Xiaoye Qu

Aim. This study used machine learning methods to develop a prediction model for knee pain in middle-aged and elderly individuals. *Methods*. A total of 5386 individuals above 45 years old were obtained from the National Health and Nutrition Examination Survey. Participants were randomly divided into a training set and a test set at a 7 : 3 ratio. The training set was used to create a prediction model, whereas the test set was used to validate the proposed model. We constructed multiple predictive models based on three machine learning methods: logistic regression, random forest, and Extreme Gradient Boosting. The model performance was evaluated by areas under the receiver (AUC), sensitivity, specificity, positive predictive value, and negative predictive value. Additionally, we created a simplified nomogram based on logistic regression for better clinical application. *Results*. About 31.4% (1690) individuals were with self-reported knee pain. The logistic regression showed that female gender (odds ratio [OR] = 1.28), pain elsewhere (OR = 4.64), and body mass index (OR = 1.05) were significantly associated with increased risk of knee pain. In the test set, the logistic regression (AUC = 0.71) showed similar but slightly higher accuracy than the random forest (AUC = 0.70), while the performance of the Extreme Gradient Boosting model was less reliable (AUC = 0.59). Based on mean decrease accuracy, the most important first five predictions were pain elsewhere, waist circumference, body mass index, age, and gender. Additionally, the most important first five predictions with the highest mean decrease Gini index were pain elsewhere, body mass index, waist circumference, triglycerides, and age. The nomogram model showed good discrimination ability with an AUC of 0.75 (0.73-0.77), a sensitivity of 0.72, specificity of 0.71, a positive predictive value of 0.45, and a negative predictive value of 0.88. *Conclusion*. This study proposed a convenient nomogram tool to evaluate the risk of knee pain for the middle-aged and elderly US population in primary care. All the input variables can be easily obtained in a clinical setting, and no additional radiologic assessments were required.

1. Introduction

Knee pain is estimated to affect about 35% of men and 62% of women over 40 years old, constituting a significant health threat worldwide [1]. Patients with knee pain usually experience reduced physical ability and poor life quality [2–4]. The disease burden of knee pain is increasing due to the aging population and limited preventive strategies. For middle-aged and elderly, knee osteoarthritis is the primary cause of knee pain [5]. Importantly, most knee joint diseases usually progress slowly but would eventually result in joint failure with pain and disability.

However, there lacks a close association between radiological alteration and the occurrence of pain, and it remains unclear at which stage the disease would cause knee pain [6]. Considering the significant individual and socioeconomic burden, attention has been paid to the early detection and prevention of knee pain [7]. Many studies have revealed risk factors for knee pain, such as elder age, female gender, and obesity. A better understanding of the risk factor could provide an insightful and cost-effective tool for identifying those with an increased risk of knee pain [8]. When high-risk individuals are identified, clinicians can offer them preventive strategies and change their lifestyles [9].

Table 1: The characteristics of the study population.

	Knee pain	No knee pain	*P*
N	1690	3696	
Age	64.0 (54.0, 74.0)	64.0 (53.0, 74.0)	0.202
Gender (female), *n* (%)	962 (56.9%)	1758 (47.6%)	< 0.001
Race, *n* (%)			0.383
Non-Hispanic white	977 (57.8%)	2150 (58.2%)	
Non-Hispanic black	320 (18.9%)	643 (17.4%)	
Mexican American	296 (17.5%)	658 (17.8%)	
Others	97 (5.7%)	245 (6.6%)	
Education, *n* (%)			0.029
Below high school	591 (35.0%)	1223 (33.1%)	
High school	424 (25.1%)	855 (23.1%)	
Above high school	675 (39.9%)	1618 (43.8%)	
Hypertension (yes), *n* (%)	917 (54.3%)	1648 (44.6%)	< 0.001
Diabetes (yes), *n* (%)	305 (18.0%)	549 (14.9%)	0.003
Pain elsewhere (yes), *n* (%)	1192 (70.5%)	1247 (33.7%)	< 0.001
Moderate activity (yes), *n* (%)	683 (40.4%)	1663 (45.0%)	0.002
Vigorous activity (yes), *n* (%)	255 (15.1%)	734 (19.9%)	< 0.001
Smoking (yes), *n* (%)	920 (54.4%)	2014 (54.5%)	0.994
Drinking (yes), *n* (%)	315 (18.6%)	638 (17.3%)	0.234
BMI (kg/m^2)	29.1 (25.8, 33.3)	27.1 (24.1, 30.6)	< 0.001
Waist circumference (cm)	102.5 (93.5, 111.8)	98.2 (89.4, 107.3)	< 0.001
Albumin (g/L)	42.0 (40.0, 44.0)	42.0 (40.0, 44.0)	0.002
Phosphorus (mg/dL)	3.7 (3.4, 4.1)	3.7 (3.4, 4.0)	0.075
Total calcium (mg/dL)	9.5 (9.2, 9.7)	9.5 (9.2, 9.7)	0.590
Triglycerides (mg/dL)	128.0 (89.0, 184.8)	123.0 (85.0, 179.0)	0.018
Cholesterol (mg/dL)	205.0 (179.0, 234.8)	205.0 (180.0, 233.0)	0.608
Vitamin D (nmol/L)	55.7 (39.7, 70.6)	58.1 (43.4, 72.9)	< 0.001
eGFR (ml/min/1.73m^2)	77.9 (20.6)	78.4 (20.6)	0.410

BMI: body mass index; eGFR: estimated glomerular filtration rate.

Although previous studies have proposed prediction models for joint pain or osteoarthritis [10–12], the small sample size and the clinical inapplicability make it difficult to apply to clinical practice. Therefore, this study sought to develop a risk prediction model for knee pain based on easily obtained demographics and laboratory biomarkers. Multiple machine learning methods (logistic regression, random forest, and Extreme Gradient Boosting) were applied, and we visualized the logistic regression model using a nomogram.

2. Methods

2.1. Study Population. The National Health and Nutrition Examination Survey (NHANES) is a program survey for the health and nutritional status of the US population. We analyzed two continuous NHANES surveys, including the 2001-2002 and the 2003-2004 surveys. Participants with demographic, anthropometry, and laboratory records were included in this study. We excluded those below 45 years old or without records of joint pain ($N = 4366$). A total of 5386 individuals were analyzed in this study.

2.2. Knee Pain Assessment. In the continuous NHANES survey, self-reported knee pain was obtained by questionnaires (detailed descriptions were provided at: http://wwwn.cdc.gov/Nchs/Nhanes/2001-2002/MPQ_B.htm). Participants were first asked the screening question: "Joint pain/aching/stiffness in past year?". For those who answered 'Yes', they would be subsequently asked the following questions: 'Right/left shoulder affected?', 'Right/left elbow affected?', 'Right/left hip affected?', 'Right/left wrist affected?', 'Right/left knee affected?', 'Right/left ankle affected?', 'Right toes affected?', and 'Right fingers/thumb affected?'. Based on the questionnaires, we identified patients with joint pain and the affected joints. We defined patients with knee pain as those who responded 'Yes' to 'Right/left knee affected?'. Additionally, pain in other areas (shoulder, elbow, hip, wrist, ankle, toes, or fingers) was defined as pain elsewhere.

2.3. Predictive Biomarkers. This study selected multiple predictive biomarkers associated with knee pain based on literature review and expert recommendations [13, 14]. All selected biomarkers can be easily obtained by inquiry, body

Table 2: Multivariable logistic regression.

	Odds ratio	95% CI	*P*
Age	1.01	1.00-1.02	0.105
Gender (female)	1.28	1.06-1.55	0.011
Race			
Non-Hispanic white	*Ref*		
Non-Hispanic black	1.15	0.91-1.45	0.249
Mexican American	1.02	0.81-1.28	0.889
Others	0.91	0.65-1.26	0.563
Education			
Below high school	*Ref*		
High school	1.06	0.86-1.31	0.588
Above high school	0.95	0.78-1.15	0.598
Hypertension (yes)	1.11	0.95-1.31	0.198
Diabetes (yes)	0.91	0.73-1.12	0.376
Pain elsewhere (yes)	4.64	3.98-5.43	< 0.001
Moderate activity (yes)	1.04	0.89-1.22	0.617
Vigorous activity (yes)	0.90	0.72-1.12	0.338
Smoking (yes)	1.12	0.96-1.31	0.158
Drinking (yes)	0.86	0.70-1.05	0.152
BMI (kg/m^2)	1.05	1.02-1.08	< 0.001
Waist circumference	1.00	0.99-1.02	0.471
Albumin (g/L)	1.01	0.98-1.04	0.584
Phosphorus (mg/dL)	0.99	0.85-1.15	0.850
Total calcium (mg/dL)	1.01	0.82-1.24	0.927
Triglycerides (mg/dL)	1.00	1.00-1.00	0.470
Cholesterol (mg/dL)	1.00	1.00-1.00	0.674
Vitamin D (nmol/L)	1.00	0.99-1.00	0.438
eGFR (ml/min/1.73m^2)	1.00	1.00-1.01	0.457

BMI: body mass index; eGFR: estimated glomerular filtration rate.

measure, and blood test. Age (years), gender (male or female), race (Non-Hispanic White, Non-Hispanic Black, Mexican American, and others), education (Below high school, high school, and above high school), hypertension (yes or no), diabetes (yes or no), pain elsewhere (shoulder, elbow, hip, wrist, ankle, toes, or fingers/thumb), moderate activity (yes or no), vigorous activity (yes or no), smoking (yes or no), and drinking (yes or no) were collected by questionnaires. Body mass index (BMI) and waist circumference were obtained by body measure. Plasm levels of albumin (g/L), phosphorus (mg/dL), total calcium (mg/dL), triglycerides (mg/dL), cholesterol (mg/dL), and vitamin D (nmol/L), were also examined. Additionally, we calculated the estimated glomerular filtration rate (eGFR) by the Chronic Kidney Disease Epidemiology Collaboration equation [15].

2.4. Development and Validation of Prediction Model. Participants were randomly divided into a training set and a test set at a 7 : 3 ratio. The training set was used to create a prediction model, whereas the test set was used to validate the proposed model. Patients' characteristics in the training and testing set are shown in Table S1. The upmentioned variables were used as inputs, and we set the prevalence of knee pain as the outcome. We constructed multiple predictive models based on multiple machine learning methods, including logistic regression, random forest, and Extreme Gradient Boosting (XGBoost). Random forest is an ensemble learning method for data regression and classification based on a multitude of decision trees [16], whereas XGBoost is a scalable end-to-end tree boosting system. We showed the model performance by receiver operating characteristic curve and calculated the areas under the receiver (AUC) of the three models. Sensitivity, specificity, positive predictive value, and negative predictive value were also provided.

Moreover, we used gender, age, hypertension, diabetes, vitamin D, pain elsewhere, total calcium, waist circumference, and BMI to create a logistic regression-based prediction model. The prediction model was then visualized by a nomogram, which is more practical for clinical application. Each variable of the nomogram was assigned a preliminary score, and the total score could be accordingly calculated. Eventually, the total score would be converted to the probability of knee pain (0-100%).

2.5. Statistical Analysis. The missing variables were filled by the multivariate multiple imputation method. Continuous variables were presented as median (Q1, Q3) and compared by the Kruskal-Wallis test between groups, whereas the categorical variables were presented as percentages and compared by the chi-square test. We performed multivariate regression to investigate the association between the biomarkers and knee pain. Analyses were performed by R software (version 3.6.1). $P < 0.05$ as considered statistically significant.

3. Results

3.1. Study Population. Table 1 describes the characteristics of the study population. Among the 5386 individuals, 1690 (about 31.4%) had self-reported knee pain. Compared with the normal group, knee pain patients were more with gender sex, lower education, hypertension, diabetes, pain elsewhere, moderate activity, and vigorous activity. Also, knee pain patients showed higher BMI, waist circumference, and triglycerides but low vitamin D levels.

3.2. Multivariable Logistic Regression. We performed logistic regression on the included biomarkers. The results showed that female gender (OR = 1.28, 95% CI = 1.06 – 1.55), pain elsewhere (OR = 4.64, 95% CI = 3.98 – 5.43), and BMI (OR = 1.05, 95% CI = 1.02 – 1.08) were significantly associated with increased risk of knee pain (Table 2).

3.3. Performance of the Prediction Models. We used age, gender, race, education, hypertension, diabetes, pain elsewhere, moderate activity, vigorous activity, smoking, drinking, BMI, waist circumference, albumin, phosphorus, total calcium, triglycerides, cholesterol, vitamin D, and eGFR as input variables. Three different models based on logistic regression, random forest, and XGBoost were created using the training set, respectively. Among the three models, logistic regression (AUC = 0.71, 95% CI = 0.68 – 0.74) showed

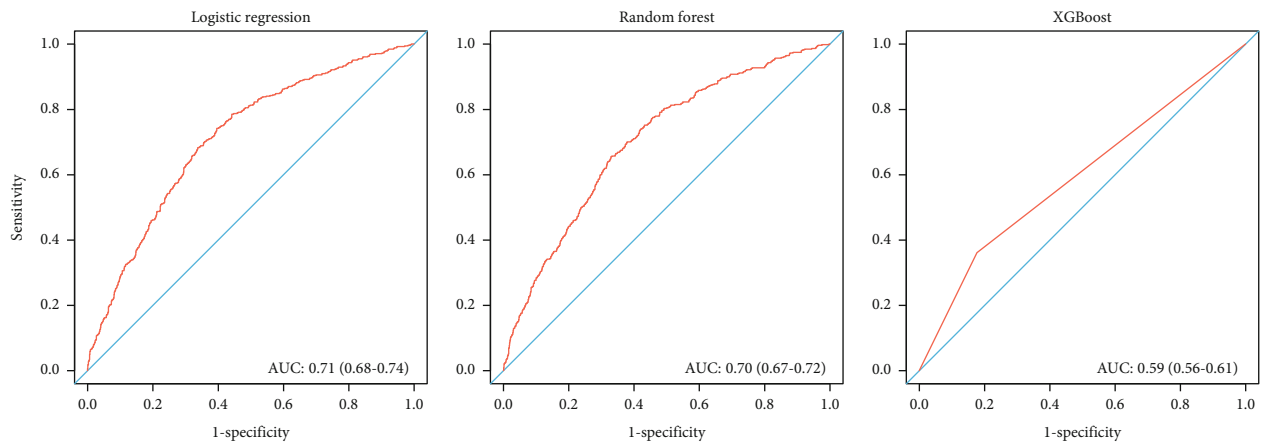


Figure 1: The receiver operating characteristic curves of the logistic regression, random forest, and XGBoost based on the test set.

similar but slightly higher accuracy than random forest (AUC = 0.70, 95% CI = 0.67 – 0.72), while the performance of the XGBoost model was less reliable (Figure 1).

The random forest model shows a sensitivity of 0.72, specificity of 0.61, a positive predictive value of 0.46, and a negative predictive value of 0.83. The variable importance of the random forest model is illustrated in Figure 2. The higher mean decrease accuracy and decrease Gini index suggested the more important role of a variable in knee pain. Based on mean decrease accuracy, the most important first five predictions were pain elsewhere, waist circumference, body mass index, age, and gender. Additionally, the most important first five predictions with the highest mean decrease Gini index were pain elsewhere, body mass index, waist circumference, triglycerides, and age. The logistic regression model shows a sensitivity of 0.71, a specificity of 0.64, a positive predictive value of 0.47, and a negative predictive value of 0.83.

Moreover, the prediction model based on logistic regression was visualized by a nomogram (Figure 3). Gender, age, hypertension, diabetes, vitamin D, pain elsewhere, total calcium, waist circumference, and BMI were input into the simplified nomogram. In the testing set, the nomogram model showed good discrimination ability with an AUC of 0.75 (0.73-0.77), a sensitivity of 0.72, specificity of 0.71, a positive predictive value of 0.45, and a negative predictive value of 0.88.

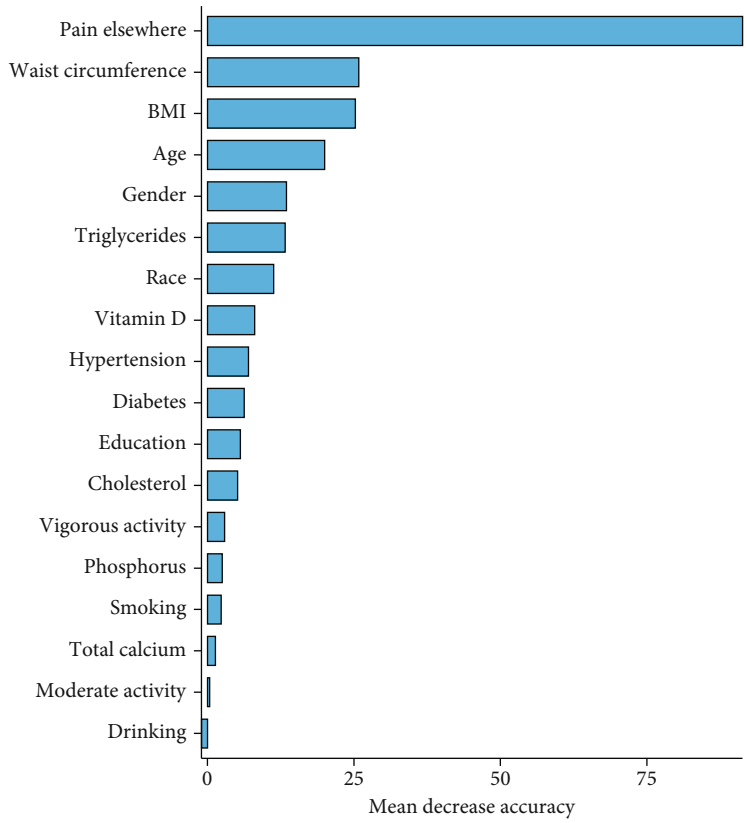
4. Discussion

Knee pain is closely associated with the middle-aged and elderly population and has become a major reason for early retirement. The NHANES I survey indicated that about 14.6% population reported knee pain [17]. It was reported that knee pain made about 20% of individuals with knee osteoarthritis retire earlier by eight years [18]. This study developed machine learning models to evaluate the risk of knee pain for the general US population. In the test set, random forest showed similar but slightly higher accuracy than logistic regression, while the performance of the XGBoost model was less reliable.

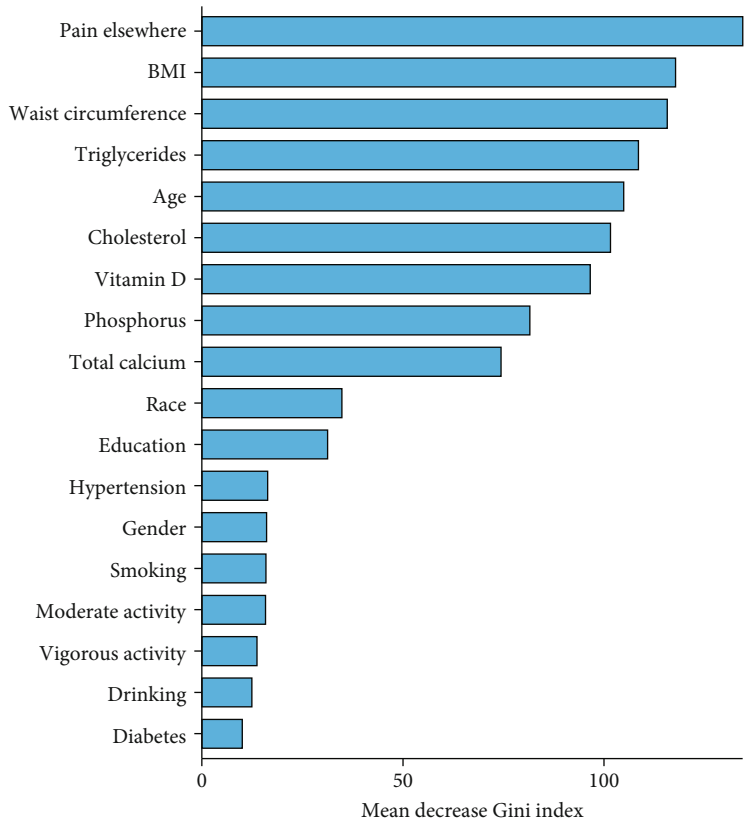
Many risk factors have been proposed for knee pain, such as age, female gender, obesity, and pain elsewhere. [8, 14, 19]. However, individuals with one or some risk factors might not experience knee pain, and a single risk factor alone was insufficient to evaluate the disease risk comprehensively. Therefore, several conventional risk factors were used as the input of the prediction models in this study. All the input variables can be measured easily in clinical practice, and no radiologic assessment was required in these models. Using easily available biomarkers without additional laboratory or radiologic examinations makes the prediction model simple to use and cost-effective. The nomogram model has a high negative predictive value (0.84) but a lower positive predictive value (0.47). Therefore, this prediction model is more suitable for identifying individuals with low knee pain risk.

The aging-induced joint pain is a multifactorial process involving numerous factors, such as cartilage thinning, muscle weakening, and proprioception reduction. Aging would also decline the capability of maintaining tissue homeostasis, thus causing an inadequate response to joint injury. We also observed that BMI and waist circumference were positively associated with a higher risk of knee pain [20]. Due to the population aging and the elevated obesity prevalence, knee pain was expected to be a growing health problem.

Besides aging and obesity, other biomarkers were also involved in the prediction models [8]. Pain elsewhere (shoulder, elbow, hip, wrist, ankle, toes, or fingers/thumb) was a significant biomarker for knee pain with an OR of 4.64 (95% CI = 3.98 – 5.43), which was consistent with previous studies [14, 18]. The association between joint pain and pain elsewhere might be attributed to the shared pathology or the progress of chronic pain syndrome [21]. Fernandes et al. [18] analyzed 1822 participants at risk for knee pain from the Nottingham community and followed the participants for 12 years. The results showed that pain elsewhere led to a 2.49-fold risk of knee pain [18]. In another prospective cohort study of 2982 people, the baseline pain other than the knee increased the risk of the new onset of knee pain but not for the progression from mild to severe [14].



(a)



(b)

Figure 2: Variable importance of the random forest model evaluated by (a) mean decrease accuracy and (b) mean decrease Gini index.

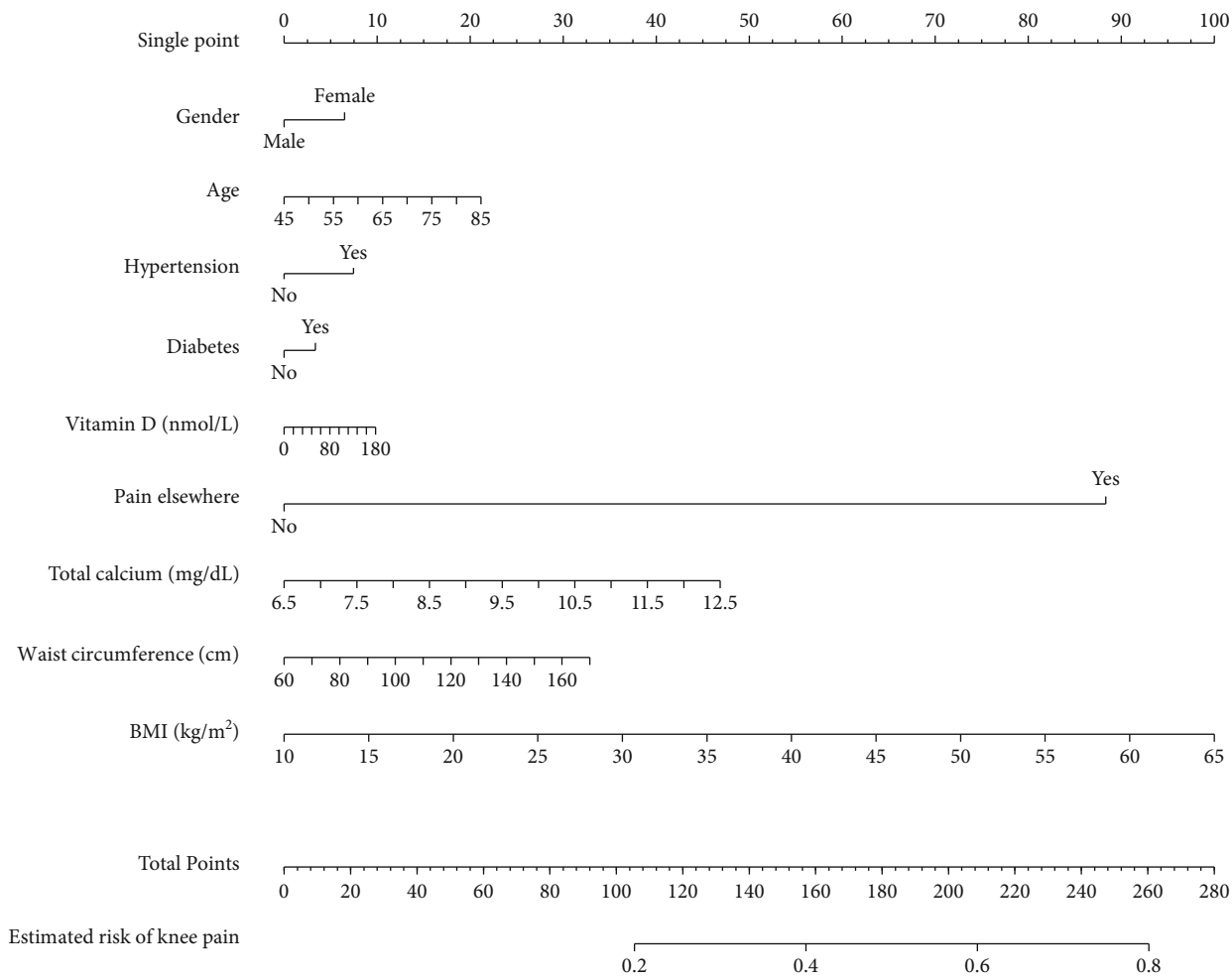


Figure 3: The prediction model based on logistic regression visualized by a nomogram.

Previous studies also proposed prediction models for joint pain or osteoarthritis [10–12]. Zhang et al. created a prediction model for radiographic knee osteoarthritis based on a 12-year retrospective community cohort (UK Nottingham cohort). Age, gender, BMI, family history, and joint injury were included in the prediction model (AUC = 0.70) [11]. Similarly, Kerkhof et al. [10] used age, gender, BMI, questionnaire variables, genetic scores, and radiographic signs to develop a prediction model for radiographic knee osteoarthritis based on Netherlands individuals aged 55 years and over (AUC = 0.79). Compared with previous prediction models, our model showed similar accuracy but was based on easily available biomarkers without additional laboratory or radiologic assessments. These advantages make it a simple-to-use and cost-effective tool suitable for primary care.

Still, some limitations should be motioned. First, the definition of knee pain was based on self-reported knee pain. A proportion of self-reported knee pain might be referred to as hips/spine pain instead of pain from the knee. Second, we tested the model performance in the internal set. However, we are unsure if the proposed knee pain prediction tool can be applied to other populations, such as the Chinese or European population. Third, the NHANES was a cross-sectional design which induces the inherent bias. Further validation and improvement were required in the following research. Fourth, although we input multiple variables in the model, many risk factors potentially remain. The investigation of additional biomarkers would improve the model performance.

5. Conclusion

This study proposed a convenient tool to evaluate the risk of knee pain for the middle-aged and elderly US population in primary care. All the input variables can be easily obtained in a clinical setting, and no additional radiologic assessments were required. In the internal validation, the nomogram model showed reliable performance with an AUC of 0.72.

Data Availability

The data in this study can be obtained from https://www.cdc.gov/nchs/nhanes/index.htm

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Lu Liu and Min-min Zhu contributed equally to this work.

Acknowledgments

Wuxi Municipal Health Commission Major Scientific Research Project (No. Z202112); National Outstanding Youth Science Fund Project of National Natural Science Foundation of China (No. 81800554); Wuxi Municipal Health Commission Achievement Technology Promotion Project (No. T201915).

Supplementary Materials

Table S1. Characteristic between the training and testing set. *(Supplementary Materials)*

References

[1] H. Breivik, B. Collett, V. Ventafridda, R. Cohen, and D. Gallacher, "Survey of chronic pain in Europe: prevalence, impact on daily life, and treatment," *European Journal of Pain*, vol. 10, no. 4, pp. 287–333, 2006.

[2] L. T. Ho-Pham, T. Q. Lai, L. D. Mai, M. C. Doan, H. N. Pham, and T. V. Nguyen, "Prevalence of radiographic osteoarthritis of the knee and its relationship to self-reported pain," *PLoS One*, vol. 9, no. 4, article e94563, 2014.

[3] C. Jinks, K. Jordan, and P. Croft, "Measuring the population impact of knee pain and disability with the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC)," *Pain*, vol. 100, no. 1, pp. 55–64, 2002.

[4] M. Grotle, K. B. Hagen, B. Natvig, F. A. Dahl, and T. K. Kvien, "Prevalence and burden of osteoarthritis: results from a population survey in Norway," *The Journal of Rheumatology*, vol. 35, no. 4, pp. 677–684, 2008.

[5] U. S. Nguyen, Y. Zhang, Y. Zhu, J. Niu, B. Zhang, and D. T. Felson, "Increasing prevalence of knee pain and symptomatic knee osteoarthritis: survey and cohort data," *Annals of Internal Medicine*, vol. 155, no. 11, pp. 725–732, 2011.

[6] H. G. Schaible, F. Richter, A. Ebersberger et al., "Joint pain," *Experimental Brain Research*, vol. 196, no. 1, pp. 153–162, 2009.

[7] D. J. Hunter, D. Schofield, and E. Callander, "The individual and socioeconomic impact of osteoarthritis," *Nature Reviews Rheumatology*, vol. 10, no. 7, pp. 437–441, 2014.

[8] M. Blagojevic, C. Jinks, A. Jeffery, and K. P. Jordan, "Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis," *Osteoarthritis and Cartilage*, vol. 18, no. 1, pp. 24–33, 2010.

[9] P. Sheeran, P. R. Harris, and T. Epton, "Does heightening risk appraisals change people's intentions and behavior? A meta-analysis of experimental studies," *Psychological Bulletin*, vol. 140, no. 2, pp. 511–543, 2014.

[10] H. J. Kerkhof, S. M. A. Bierma-Zeinstra, N. K. Arden et al., "Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors," *Annals of the Rheumatic Diseases*, vol. 73, no. 12, pp. 2116–2121, 2014.

[11] W. Zhang, D. F. McWilliams, S. L. Ingham et al., "Nottingham knee osteoarthritis risk prediction models," *Annals of the Rheumatic Diseases*, vol. 70, no. 9, pp. 1599–1604, 2011.

[12] L. Sharma, M. Hochberg, M. Nevitt et al., "Knee tissue lesions and prediction of incident knee osteoarthritis over 7 years in a cohort of persons at higher risk," *Osteoarthritis and Cartilage*, vol. 25, no. 7, pp. 1068–1075, 2017.

[13] V. Silverwood, M. Blagojevic-Bucknall, C. Jinks, J. L. Jordan, J. Protheroe, and K. P. Jordan, "Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis," *Osteoarthritis and Cartilage*, vol. 23, no. 4, pp. 507–515, 2015.

[14] C. Jinks, K. P. Jordan, M. Blagojevic, and P. Croft, "Predictors of onset and progression of knee pain in adults living in the community. a prospective study," *A Prospective Study. Rheumatology*, vol. 47, no. 3, pp. 368–374, 2008.

[15] A. S. Levey, L. A. Stevens, C. H. Schmid et al., "A new equation to estimate glomerular filtration rate," *Annals of Internal Medicine*, vol. 150, no. 9, pp. 604–612, 2009.

[16] C. Liu, J. Liao, Y. Zheng, Y. Chen, H. Liu, and X. Shi, "Random forest algorithm-enhanced dual-emission molecularly imprinted fluorescence sensing method for rapid detection of pretilachlor in fish and water samples," *Journal of Hazardous Materials*, vol. 439, article 129591, 2022.

[17] T. M. Link, L. S. Steinbach, S. Ghosh et al., "Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings," *Radiology*, vol. 226, no. 2, pp. 373–381, 2003.

[18] G. S. Fernandes, A. Bhattacharya, D. F. McWilliams, S. L. Ingham, M. Doherty, and W. Zhang, "Risk prediction model for knee pain in the Nottingham community: a Bayesian modelling approach," *Arthritis Research & Therapy*, vol. 19, no. 1, p. 59, 2017.

[19] H. Frilander, E. Viikari-Juntura, M. Heliövaara, P. Mutanen, V. M. Mattila, and S. Solovieva, "Obesity in early adulthood predicts knee pain and walking difficulties among men: a life course study," *European Journal of Pain*, vol. 20, no. 8, pp. 1278–1287, 2016.

[20] R. Marks, "Obesity profiles with knee osteoarthritis: correlation with pain, disability, disease progression," *Obesity*, vol. 15, no. 7, pp. 1867–1874, 2007.

[21] P. Croft, K. Jordan, and C. Jinks, ""Pain elsewhere" and the impact of knee pain in older people," *Arthritis and Rheumatism*, vol. 52, no. 8, pp. 2350–2354, 2005.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 4133663, 13 pages
https://doi.org/10.1155/2022/4133663

Research Article

Optimization Method of an Antibreast Cancer Drug Candidate Based on Machine Learning

Zhibai Huang, Shengji Jiang, and Weiqiang Xiao

East China Institute of Computing Technology, Shanghai, China

Correspondence should be addressed to Zhibai Huang; huangzhibainudt4@nudt.edu.cn

Received 19 July 2022; Revised 17 August 2022; Accepted 22 August 2022; Published 5 September 2022

Academic Editor: Jincheng Wang

Breast cancer is a common but serious and even lethal disease. Fortunately, compared with other cancers, breast cancer treatments currently are relatively well developed. The use of specific drugs is typically essential in the majority of breast cancer treatment strategies. Given the aforementioned factors, it is important to continue researching effective antibreast cancer drug design. Machine learning-based computer-aided drug design is currently a common practice in both drug industries and academic institutes. According to the characteristics of breast cancer, we selected multiple candidate compounds; based on the corresponding molecular descriptors, biological activities, and pharmacokinetic properties, a dataset of inhibition potency and pharmacokinetic properties paired with multiple features of compounds was constructed. On this basis, the random forest method was utilized to choose greater-influenced feature embeddings; thus, 224 main operating variables were selected for further analysis; we then employed the efficient MobileNetV3 deep neural network as the backbone to establish the prediction models for the inhibition potency and pharmacokinetic properties of the compounds. After data preprocessing, the weights are obtained by training on the refined dataset. Finally, we define an optimization problem to discover compounds with the best properties. The problem is solved using the genetic algorithm with the acquired prediction model, and the solution value for the corresponding operating variables with the best clinical properties in theory is then obtained. Analysis demonstrates that our approach could be used to aid the screening process of antibreast cancer drug candidates.

1. Introduction

Breast cancer currently ranks among the most prevalent cancers worldwide [1] and has a high fatality rate. Estrogen receptors are connected with the development of breast cancer [2, 3]. Given that the estrogen receptor α subtype (estrogen receptor alpha (ERα)) is present in roughly 70% of breast cancer cells [4], it has been widely considered in the diagnosis of breast cancer [2]. Studies on mice with ERα gene modifications have demonstrated that ERα does, in fact, play a crucial role in the development of the uterus and mammary glands [4, 5]. Consequently, as ERα is considered a key target for the treatment of breast cancer, substances that can suppress ERα activity might be proper candidates for use as therapeutics [2].

For a long time, the gold standard for the endocrine treatment of several breast cancer types was tamoxifen [6], a common drug with estrogen-like actions. Since tamoxifen was found to be effective in treating breast cancer, numerous studies have been conducted to highlight the significance of hormone therapy for the disease [7, 8]. In addition to hormone therapy, other medicines for breast cancer include chemotherapy and immunological therapy [9]. As the leading drugs for the aforementioned therapies, cyclophosphamide, docetaxel, pertuzumab, and trastuzumab are currently commonly used to treat breast cancer [10–13].

Building an inhibitory potency prediction model can be used to screen candidate compounds during the conventional drug design process to save time and money [14]. The precise procedure is as follows: first, for a biological target associated with a disease, gather data on a number of compounds that affect the target and their biological activity. Next, build a quantitative structure-activity relationship (QSAR) model of candidates using a number of molecular descriptors as independent variables and the biological activity value of the compound as the dependent variable. Finally,

the model is employed to forecast how a molecule might seem when having sufficient biological activity or to direct the structural improvement of already existing compounds [15].

In addition to having significant biological activity, a chemical should also have appropriate pharmacokinetic and safety qualities in the human body to be employed as a new medicine. Similar to that, it can be evaluated using an established QSAR model [16]. There are numerous ways to create prediction models at the time, whereas methods based on artificial neural networks have received more attention from academic communities than other alternatives [15, 17, 18]. Computer-aided drug design techniques have been applied extensively in many aspects of drug design after years of development [19, 20]. Deep neural networks have been extensively used since the dawn of the big data age, and numerous research in drug design have been conducted [21, 22]. A few researchers have attempted to use the effective convolutional neural network MobileNetV3 [23] in the field of medicine [24, 25].

In this research, based on the current knowledge, we created a dataset of potential antibreast cancer drug candidates, and then, we refined the dataset by applying the random forest algorithm. We explore adopting MobileNetV3 to create a QSAR model on the refined dataset to construct the qualitative model of pharmacokinetic performance and the quantitative prediction model of inhibitory potency. Finally, a problem for optimizing the attributes of the chemical is created based on the model that was obtained and the genetic algorithm is utilized to solve the problem.

2. Dataset Construction and Preprocessing

2.1. Dataset Construction. According to prior knowledge and experience, we selected a total of 1974 compounds, calculated the corresponding 2-dimensional and 3-dimensional molecular descriptors by computing software [26], and subsequently marked their biological activity values and ADMET properties to complete the construction of the dataset. We choose IC50, pIC50 as the index for biological activity, Caco-2 for A (absorption), CYP3A4 for D (distribution), hERG for M (metabolism), HOB for E (excretion), and MN for T (toxicity). Specifically, Caco-2 is the permeability of small intestinal epithelial cells, which can measure the ability of the compound to be absorbed by the human body. CYP3A4 is the cytochrome P450 enzyme 3A4 isoform, which is the main metabolic enzyme in the human body, which can measure the compound. hERG is the cardiac safety evaluation of the compound, which can measure the cardiotoxicity of the compound. HOB is the oral bioavailability of the human body, which can measure the proportion of the drug absorbed into the human blood circulation after entering the human body. MN is the micronucleus test and is a method to detect whether a compound is genotoxic.

Based on the dataset (including 1974 compound samples, each with 1361 molecular descriptor variables, e.g., electrotopological state atom type descriptor, ring count descriptor, WHIM descriptor etc., 2 biological activity data, and 5 ADMET property data), we then built a quantitative prediction model for compound biological activity and a categorical prediction model for ADMET properties.

2.2. Data Preprocessing

2.2.1. Data Cleaning. Due to some problems in the collected raw data, to ensure the data analysis quality, the raw data should be cleaned in a certain level. The overall data processing method is as follows:

(1) There is dimensionless normalization of molecular descriptor data in all samples

(2) For data columns with most of the data being 0, delete them directly

(3) Only the pIC50 array was selected as the biological activity label

Considering the large difference in raw values between different molecular descriptors, to improve the model accuracy, we first use the min-max normalization method [27] to perform dimensionless normalization on the molecular descriptor data.

On this basis, we double check the normalized samples and delete most of the data columns with 0 values (dimensionless normalized molecular descriptors) to reduce data redundancy. Make datasets more compact and efficient without losing too much information. By excluding some factors with low impact on biological activity in advance, the convergence speed of subsequent selection of main features should be accelerated.

Finally, we chose pIC50 as the only numerical annotation for biological activity. Since the pIC50 value is distributed in the $[0, 10]$ interval, it is more friendly to the deep network model. Considering that the IC50 and pIC50 can be equivalently transformed through numerical calculation, dropping the IC50 label should not ignore valuable information. Only pIC50 is selected as the biological activity numerical labeling instead of the IC50 and pIC50 binary label group; we believe that the sole existence of pIC50 should make the dataset more "compact," thus leading to a more efficient and accurate prediction model.

2.2.2. Selecting Main Features. Considering the large number of data columns in the dataset, it is necessary to further compress the number of data columns; we chose to use the random forest algorithm [28] to select features to further compress the dataset. The importance of each feature can be obtained by performing certain operations on the result of the sample classification. The smaller the result is, the smaller impact that this feature affects the prediction result. According to the variable contribution ranking obtained by random forest algorithm, we select a total of 224 data columns (which are processed molecular descriptors) in order of contribution, as shown in Figure 1.

Among them, XlogP is the lipid-water partition coefficient, which reflects the absorption effect of molecules through the cell membrane. TopoPSA is the topological polar surface area, reflecting factors such as molecular size and solubility. From the statistical results of the categories

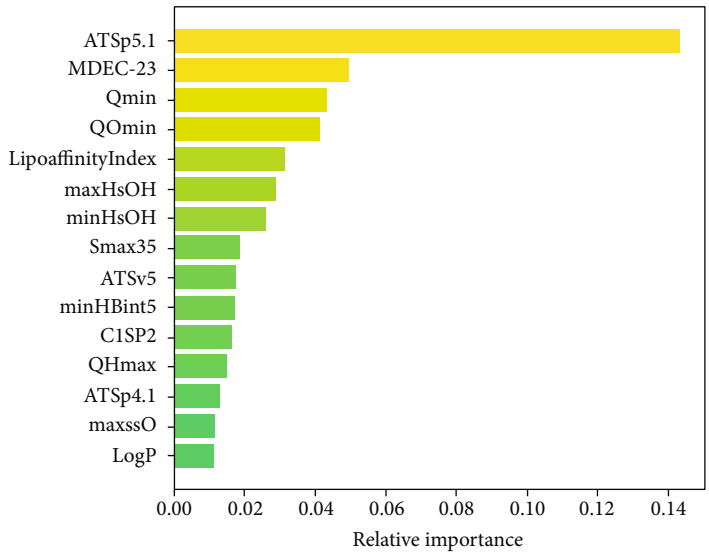


Figure 1: Visualization of the top 15 variables of contribution.

to which each variable belongs, it can be seen that the extracted variables include a certain level of comprehensive types of compound fingerprints. The variance is relatively minimal, which suggests that the extracted variables place a balanced emphasis on each category, according to the distribution of the number of variables contained in each category. The final selected normalized molecular descriptors are shown in Table 1. After the compressing process, we have done all data preprocessing for the deep neural network.

3. Compound Property Prediction Model

3.1. Quantitative Prediction Model for Biological Activity

3.1.1. Model Design. Artificial neural networks are currently employed and widely applied in the field of computer-aided drug design. The most often used neural network is the BP (back propagation) neural network, a multivariate feedforward neural network trained via error back propagation. Deep neural networks, a version of BP neural networks, have drawn considerable attention in many fields of academia and industry. Convolutional neural networks among them have significant advantages in performance and are especially well liked in the field of computer vision. However, it is important to keep in mind that the majority of the current popular convolutional neural networks have complex structures, thus containing a lot of parameters, combined with the neural network's data-hungry nature making model training very challenging. Special consideration should be given when training these networks on relatively small-amount biological activity datasets.

On the other hand, it is crucial to properly design the number of neurons in the hidden layer during the whole network construction process. The workload required to make the network function will significantly arise if the hidden layer contains too many neurons, which can quickly result in an undesired overfitting issue. Conversely, if the hidden layer contains too few neurons, which will also negatively affect the network's quality, thus resulting in poor prediction accuracy. The total number of neurons in a neural network's hidden layer is directly correlated with the difficulty of the task, the number of neurons in the input and output layers, and the expected bias settings of those neurons.

Considering the mentioned problems, we chose the MobileNetV3 deep convolutional neural network as the backbone to construct a quantitative prediction model for the biological activity of compounds. As one of the representatives of lightweight models, compared to the classic convolutional network VGG16 [29], MobileNetV3 greatly reduces the number of parameters but is more efficient and easier to train while ensuring similar performance. The schematic diagram of the network structure that we use is shown in Figure 2.

3.1.2. Model Training. We first divide the 1974 group of data in the dataset into training set data (around 80% in amount), test set data (around 15% in amount), and validation set (around 5% in amount) according to the proportions of 80%, 15%, and 5%, respectively. After the division is completed, 224 main variables in the dataset and pIC50 annotations were constructed as a pair; then, randomly sample 10 data pairs as a batch for model input.

As an important part of model optimization, the loss function needs to be carefully considered. Considering that the quantitative prediction problem can be summarized as a regression problem, we choose MSELoss (mean square error loss), the most commonly used one in the regression task, as the loss function.

Deep learning tasks will produce varying results depending on the optimizers used. We first identified the SGD (stochastic gradient descent) [30] and Adam optimizer (adaptive moment estimation optimizer) [31] as alternatives based on the properties of the MobileNetV3 network itself and the properties of the dataset; we then compared the performance in the experimental training, and Adam was ultimately selected as the optimizer.

Table 1: The contribution of the top 224 important molecular descriptors (from low to high).

Descriptor	Importance
Smax11	0.00065
MATSp7	0.00065
CIC4	0.000652
MDEC-14	0.000656
S17	0.000659
minHCsats	0.00066
Smin34	0.000661
SHaaCH	0.000664
SIC3	0.000664
SHCsats	0.000681
Smin	0.000681
CrippenLogP	0.000682
maxsOH	0.000684
ATSc1	0.000684
bcutm13	0.000684
phi	0.000686
MATSm3	0.000688
CIC3	0.000688
VSAEstate7	0.000691
SPC-4	0.000695
EstateVSA7	0.000702
Smin8	0.000704
WTPT-5	0.000706
TPSA1	0.000708
naccr	0.00071
MATSm7	0.000712
maxdsN	0.000712
CIC1	0.000713
Smin35	0.000714
ATSe5	0.000716
minHCsatu	0.000725
GATSp3	0.000726
GATSm5	0.000727
ALogp2	0.000729
GATSp7	0.00073
EstateVSA1	0.000737
IDE	0.000741
mindO	0.000744
mChi1	0.000745
SaasC	0.00076
bcute9	0.000761
nAtomLAC	0.000762
maxdssC	0.000771
GATSe7	0.000775
Smax	0.000781
ETA_Epsilon_1	0.000787
MATSv5	0.000789
bcutp5	0.000793

Table 1: Continued.

Descriptor	Importance
IC1	0.000796
maxHBint7	0.000797
QCss	0.000823
CIC6	0.000823
ALogP	0.000826
bcutm3	0.00083
SsOH	0.000847
BertzCT	0.000851
EstateVSA4	0.000851
SdssC	0.000855
bcutm2	0.000866
MAXDN	0.000868
PC6	0.000872
MATSm6	0.000891
SHBint5	0.000897
SaaCH	0.0009
MATSp5	0.000901
MRVSA6	0.000904
slogPVSA1	0.000904
MATSm5	0.000922
bcute12	0.000926
J	0.000927
GATSm4	0.000927
MRVSA5	0.000933
MATSm1	0.000934
GATSm8	0.000939
Smin12	0.000946
hmin	0.00095
VC-4	0.000962
MATSe5	0.000963
MATSp4	0.000964
PEOEVSA5	0.000967
minHBd	0.000971
GATSv3	0.000974
bcutm9	0.000979
PEOEVSA8	0.00098
ECCEN	0.000987
MATSm8	0.000988
IC2	0.000995
BCUTp-1l	0.001004
minssCH2	0.001017
QHss	0.001019
Smax16	0.00102
bcutm12	0.001026
ETA_EtaP_F	0.001026
ETA_dEpsilon_D	0.001038
bcute4	0.001038
WTPT-3	0.001042
MAXDP2	0.001042

Table 1: Continued.

Descriptor	Importance
knotpv	0.001043
MDEO-11	0.001043
maxHCsats	0.00105
Chiv5ch	0.001063
GATSe5	0.001072
VPC-5	0.001081
MATSv8	0.00109
maxsF	0.001096
QNmin	0.001109
ETA_BetaP_s	0.001109
Chiv6ch	0.00111
IC3	0.001117
VPC-6	0.001119
VSAEstate2	0.001121
MATSp3	0.001137
slogPVSA2	0.00114
WTPT-4	0.001162
gmin	0.001163
minHBint6	0.001176
minHBint7	0.001195
Smax24	0.001225
MATSp6	0.001229
PEOEVSA1	0.001234
SIC2	0.001237
S34	0.001257
bcute1	0.001278
MATSv3	0.001281
SC-5	0.001283
dchi0	0.00129
SIC1	0.001291
maxHBd	0.001307
PEOEVSA7	0.001342
MDEC-24	0.001345
SCH-7	0.001347
SHBd	0.001349
MATSe8	0.001375
MATSv1	0.001375
SHCsatu	0.001386
Smin15	0.001398
BCUTp-1h	0.00141
GATSm3	0.001461
bcutp12	0.001465
MLFER_BH	0.001485
GATSv1	0.001559
QOmax	0.001589
slogPVSA0	0.001592
bcute10	0.001605
Smin24	0.001609
MATSp1	0.001616

Table 1: Continued.

Descriptor	Importance
Chiv3	0.001662
QNmax	0.001663
bcutv4	0.00168
VCH-5	0.001717
VSAEstate4	0.001785
ATSc5	0.001813
C3SP2	0.001831
mindssC	0.001846
ATSc2	0.001859
minHBint10	0.001866
ATSc3	0.001892
MDEC-22	0.001909
MAXDP	0.001935
knotp	0.001942
GATSm1	0.002
GATSp4	0.002026
maxsssCH	0.002031
S25	0.002032
bcutp1	0.002049
ETA_Shape_Y	0.002124
bcutp9	0.002186
XLogP	0.002237
ATSc4	0.002298
maxHBint5	0.002321
maxHBint8	0.002379
minsOH	0.002424
GATSm2	0.002443
SPC-6	0.00248
MATSe3	0.002549
MLFER_S	0.002597
SHBint6	0.002733
ndssC	0.002743
bcutv1	0.002787
VCH-7	0.002879
BCUTc-1l	0.002923
QCmax	0.00298
Scar	0.003191
minssO	0.003312
BCUTc-1h	0.003494
MLFER_A	0.003711
TopoPSA	0.003768
MDEO-12	0.003868
minHBa	0.004054
Smin33	0.004253
SHsOH	0.004402
GATSe8	0.004485
PEOEVSA6	0.00461
Mnc	0.004826
MATSe1	0.004978

Table 1: Continued.

Descriptor	Importance
LDI	0.005205
MDEC-33	0.005471
GATSe1	0.005843
bcute2	0.005866
VC-5	0.006197
nC	0.0064
nHBAcc	0.006408
LogP2	0.006781
SHBint10	0.006873
Hy	0.007517
kappam3	0.007695
VSAEstate1	0.007799
QNss	0.009891
minsssN	0.01014
LogP	0.011371
maxssO	0.011647
ATSp4.1	0.013044
QHmax	0.014923
C1SP2	0.01631
minHBint5	0.017225
ATSv5	0.017404
Smax35	0.018515
minHsOH	0.025883
maxHsOH	0.028849
LipoaffinityIndex	0.031403
QOmin	0.041317
Qmin	0.043254
MDEC-23	0.049635
ATSp5.1	0.143226

Finally, considering the nature of the Adam optimizer itself, we adopt the cosine annealing strategy [32] to update the learning rate to optimize the performance of the model as much as possible. In selecting the most suitable upper limit of the hyperparameter learning rate and the number of epochs, we also performed experimental training on the actual training set. Ultimately, we came to the conclusion that the upper limit of the learning rate is 0.0001 and the number of epochs is 100.

3.2. Qualitative Prediction Model for ADMET Properties. To simplify the problem, we trained the models separately for the five properties in ADMET. To further simplify the problem, we believe that each property has only two possibilities of "yes" or "no," which can be expressed by the values 0 and 1. In this way, the problem can be classified as a binary classification problem; then, we can reuse the divided dataset given in Section 3.1.2 and merely change the data label to pharmacokinetic properties.

Since the input data share a certain level of similarity, we still employ the MobileNetV3 structure as the backbone; therefore the desired model structure is essentially identical to the structure described in Figure 2. The only needed minor change is to adjust the output neuron of the bottom fully connected layer and add an extra sigmoid activation function; other designs shall not be repeated here.

To adapt to the binary classification problem, the model training method also needs to be adjusted. We changed the loss function to BCELoss (binary cross entropy loss), while the optimizer, learning rate adjustment strategy, and hyperparameter setup remain unchanged. The prediction accuracy is obtained based on the comparison between the predicted outputs and the real labels.

During the training process based on experiments on real datasets, the issue of data imbalance has been found. To prevent the trained model from being biased due to data imbalance, we redundantly expand the data, that is, expanding the samples of a relatively small number to be roughly equivalent to the other categories.

3.3. Optimization Model for Clinical Properties Based on Specific Features

3.3.1. Definition of Optimization Problem. We now define an optimization problem using these six prediction models that were trained in earlier sections, looking for the ideal circumstances for the 224 variable values that were chosen. Note that we assume that any "acceptable" compound must perform "well" at least three of the given ADMET properties; then, the problem could be defined as follows:

Definition 1. Given the selected molecular descriptor, what value of the molecular descriptor satisfied can make the compound have better biological activity for inhibiting ERα, meanwhile having better ADMET properties (at least three or better).

3.3.2. Optimization Problem Modeling. First, determine the decision variables; we follow the selected results in the previous section; consider the selected 224 molecular descriptors as decision variables, denoted as follows:

$$X = \{x_1, x_2, x_3, \cdots, x_{224}\}. \tag{1}$$

Now, determine the objective function. After analyzing the problem, we can find that the problem essentially is as follows: based on the given prediction models, under the premise that at least three properties of the given five ADMET properties are "good," by changing the value of selected features, the clinical properties (both inhibition potency and pharmacokinetic performance) are optimized to guide the production process. The biological activity of the compound is altered by the chosen feature's value; it is worth noting that this process will also alter the compounds' ADMET properties. Therefore, the relations between each model should not be ignored. By applying the prediction models to the input samples, the predicted value of ADMET properties of each sample can be obtained. Following the idea in Section 3.2, we take all the values representing good properties as 1 and the values of bad properties as 0; then, we get an optimization limit that the pharmacokinetic point

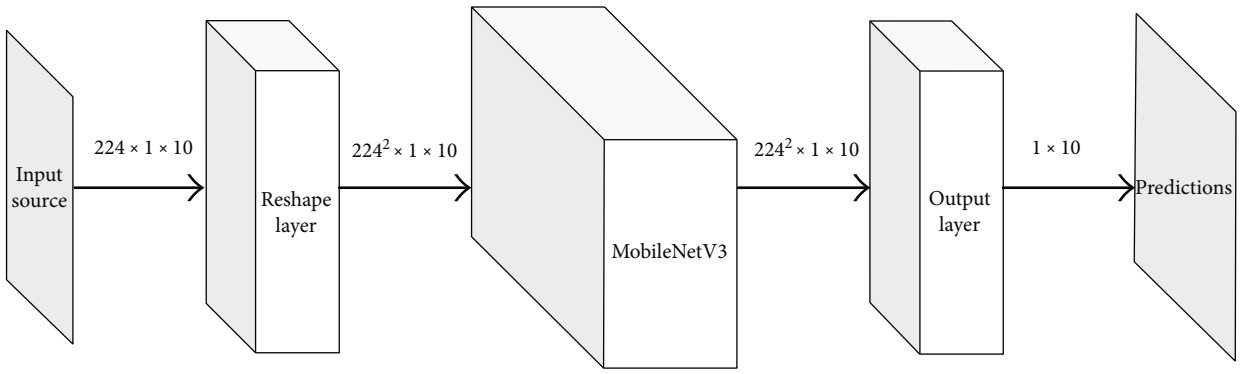


Figure 2: The schematic diagram for the model built.

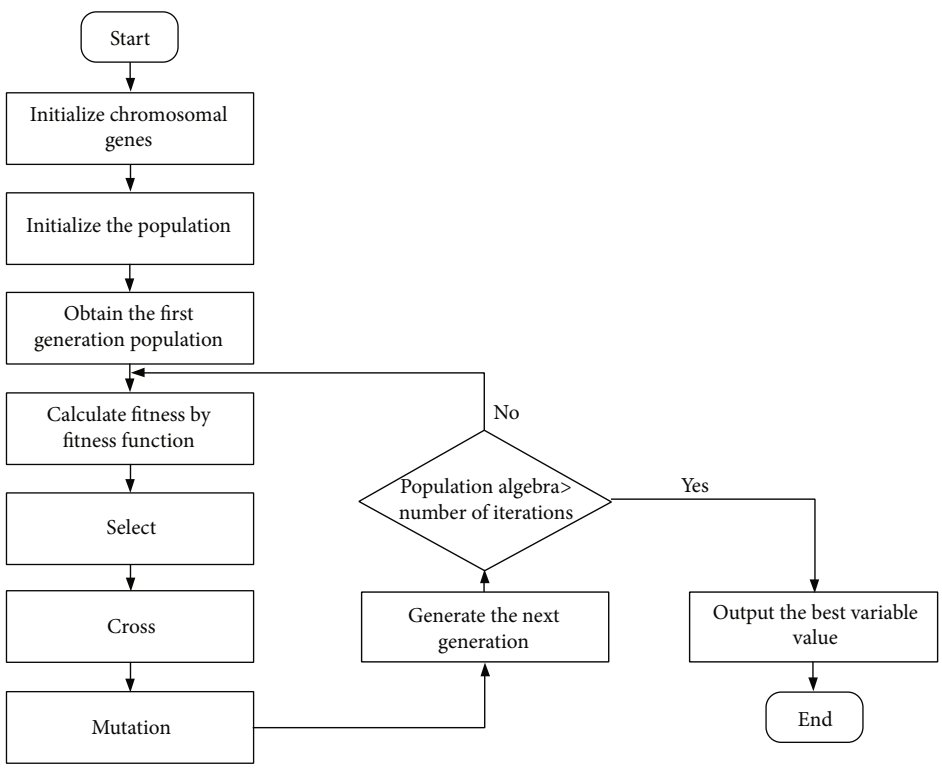


Figure 3: Flowchart of genetic algorithm.

(the sum of ADMET property marks) has a maximum value of 5. According to Definition 1, the ADMET marks of an "acceptable" compound should not be less than 3; then, the modified output function of the qualitative prediction model is derived, denoted original output as $\phi(X)$; then, denote our desired function as $\Phi(X)$, defined as follows:

$$\Phi(X) = \begin{cases} 0, & 0 \leq \sum_{\text{ADMETproperties}} \phi(X) < 3, \\ \sum_{\text{ADMETproperties}} \phi(X), & 3 \leq \sum_{\text{ADMETproperties}} \phi(X) \leq 5. \end{cases} \tag{2}$$

In addition to the ADMET properties, we need to consider the pIC50 value of the compound as well. The goal of this output function is to obtain the highest activity value under the premise of satisfying "acceptable" ADMET properties; combined with the definition of pIC50, the modified quantitative prediction model output function is given. We denoted it as $\Psi(X)$ and the original one as $\psi(X)$, define as follows:

$$\Psi(X) = \begin{cases} 0, & 10 < \psi(X), \\ 0, & \psi(X) < 0, \\ \psi(X), & 0 \leq \psi(X) \leq 10. \end{cases} \tag{3}$$

Table 2: Quantitative prediction accuracy of biological activity.

Biological activity	MSE	MAPE	MAE
pIC50	1.5720	0.1624	0.9880

Intuitively, the optimization problem should be a multi-objective nonlinear programming problem. To simplify the solution process, we transform it into a single-objective nonlinear programming problem to solve. Considering the optimization problem, it is desired that the compound's ADMET properties are as good as possible and the biological activity is as high as possible, that is, to find a set of selected feature values X_0 to maximize the sum of $\Phi(X_0)$ and $\Psi(X_0)$. In this way, we are able to extract the objective function of the optimization problem, which is defined as follows:

$$F(X) = \Psi(X) + \Phi(X). \tag{4}$$

Finally, determine the constraints: for this optimization problem, since the proposed 224 decision variables (dimensionless normalized molecular descriptors) have a certain range of actual values, there is constraint 1 as follows:

$$0 \le x_i \le 1, \quad i = 1, 2, 3, \cdots, 224. \tag{5}$$

Now, take into account the biological activity limitation. Considering the predicted pIC50 value, according to the definition of pIC50, it can be seen that there is a constraint on the value of $\psi(X)$ and we hope that the optimized biological activity value is not lower than the maximum value in the dataset for building the prediction model, so there is a constraint 2 as follows:

$$10 \ge \mathrm{pIC50}_{\mathrm{t\ arg\ et}} \ge \mathrm{pIC50}_{\mathrm{source}} \ge 0. \tag{6}$$

Combine constraints with the target function; in summary, the problem can be defined as follows:

$$\begin{aligned} &\mathrm{Max}\ F(X) \\ &\mathrm{s.t.} \begin{cases} 0 \le x_i \le 1, \quad i = 1, 2, 3, \cdots, 224, \\ 10 \ge \mathrm{pIC50}_{\mathrm{target}} \ge \mathrm{pIC50}_{\mathrm{source}} \ge 0. \end{cases} \end{aligned} \tag{7}$$

3.3.3. Optimization Problem Solving. We now address the optimization problem raised in Section 3.3.2. It can be said that the optimization problem is a single-objective nonlinear optimization problem given the complicated link between molecular descriptors and biological activity. Intelligent optimization algorithms, such the genetic algorithm [33], ant colony algorithm [34], and particle swarm optimization [35], can be used to solve this type of problem's model to acquire the optimal set of variables. We employ the genetic algorithm to address the optimization problem since it can frequently produce better optimization results more quickly than some traditional optimization methods when solving complex combinatorial optimization problems. Figure 3 depicts a typical genetic algorithm optimization procedure.

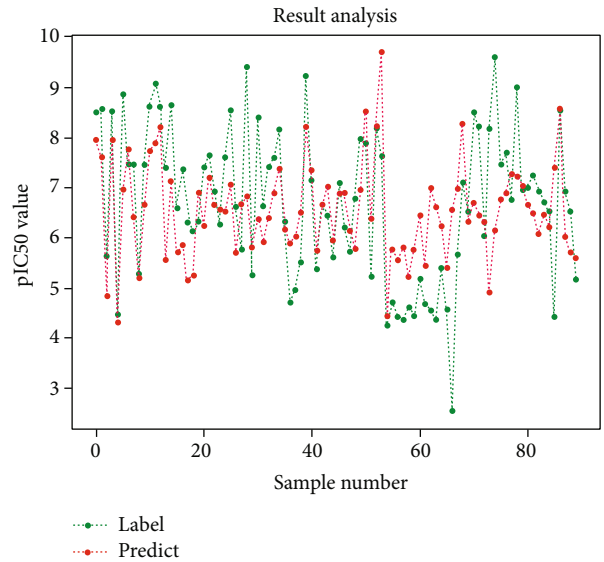


Figure 4: Performance of the quantitative prediction model.

Table 3: Qualitative prediction accuracy of biological activity.

ADMET	Accuracy	Precision	Recall	F1 score
Caco-2	0.8830	0.8158	0.3735	0.5124
CYP3A4	0.8230	0.8286	0.7436	0.7838
hERG	0.7660	0.7091	0.5147	0.6142
HOB	0.7979	0.5263	0.1333	0.2127
MN	0.9149	0.9444	0.7907	0.8608

The primary chromosomes of some members of the population are first constructed by performing binary coding on the sample operating variable's initial value, and the chromosomes of the remaining individuals are randomly generated within the value range of the operating variable. To determine the fitness of each chromosome in the population and to calculate the corresponding selection probability matrix, the binary-coded chromosomes are first decoded to the actual values of the altered variables before being input into the ADMET property prediction model and the biological activity prediction model. Chromosomes with higher fitness are more likely to be selected during evolution. Roulette selection is used in the selection strategy. Chromosomes interact with one another and mutate to create new chromosomes. Finally, if a combination of operational variables satisfies all requirements for biological activity and pharmacokinetic features, record the combination and optimize the following sample; if not, keep iterating until the ideal operating circumstances are discovered.

We built the solver in Python language to lessen the implementation's complexity. When using a genetic algorithm, simulating more complex "populations" takes longer and takes more effort. After simulation training and testing, we find the proper parameters for the solver. We randomly created the initial population and fixed the number to 2000, taking into account the difficulty of solving and the accuracy requirements. The number of iterations is limited

Table 4: The predicted values for best-performance candidate's operating variables.

Descriptor	Normalized	Real
ALogP	0.796372	22.52675
ALogp2	0.565526	301.9008
nC	0.911078	80.17483
ATSc1	0.49451	2.244443
ATSc2	0.939863	2.221841
ATSc3	0.15546	0.141572
ATSc4	0.452555	1.233483
ATSc5	0.454323	1.587603
BCUTc-1l	0.086871	0.020243
BCUTc-1h	0.197627	0.09002
BCUTp-1l	0.348526	1.416559
BCUTp-1h	0.30613	2.687486
C1SP2	0.842156	16.84312
C3SP2	0.44475	5.337003
SCH-7	0.570198	1.226224
VCH-5	0.136673	0.066583
VCH-7	0.284164	0.471843
SC-5	0.90026	2.358201
VC-4	0.678221	0.33911
VC-5	0.476908	0.706329
SPC-4	0.223485	4.435693
SPC-6	0.734395	25.70635
VPC-5	0.126582	1.515056
VPC-6	0.017549	0.319702
CrippenLogP	0.908239	23.01173
ECCEN	0.868879	13130.5
ndssC	0.095972	2.687217
SHBd	0.814323	14.78969
SHBint5	0.536753	51.89493
SHBint6	0.493409	88.24471
SHBint10	0.622465	71.61779
SHsOH	0.680758	1.868451
SHaaCH	0.943556	9.03814
SHCsats	0.361675	16.73021
SHCsatu	0.005772	0.112306
SaaCH	0.522805	20.98883
SdssC	0.131028	4.567775
SaasC	0.472381	10.71859
SsOH	0.945776	62.06689
minHBd	0.004693	0.004157
minHBa	0.381193	6.144505
minHBint5	0.414706	5.288915
minHBint6	0.647892	5.650089
minHBint7	0.123079	1.481101
minHBint10	0.785248	9.487377
minHsOH	0.825143	0.730969
minHCsats	0.59855	0.663469
minHCsatu	0.105526	0.118007

Table 4: Continued.

Descriptor	Normalized	Real
minssCH2	0.889895	2.358533
mindssC	0.268325	1.054303
minsssN	0.735504	2.011435
minsOH	0.967622	11.35247
mindO	0.231127	3.32619
minssO	0.222965	1.499747
maxHBd	0.573001	0.488639
maxHBint5	0.502949	5.811736
maxHBint7	0.117421	1.2678
maxHBint8	0.185541	1.72333
maxHsOH	0.609498	0.519763
maxHCsats	0.973025	1.252399
maxsssCH	0.255134	0.249398
maxdssC	0.322333	0.754448
maxdsN	0.148611	0.763495
maxsOH	0.982069	12.24723
maxssO	0.754777	5.077492
maxsF	0.375781	5.816332
hmin	0.202651	0.19802
gmin	0.196719	1.518661
LipoaffinityIndex	0.351299	9.693234
MAXDN	0.916639	5.952869
MAXDP	0.688238	4.738658
MAXDP2	0.41876	2.882345
ETA_Epsilon_1	0.744645	0.275913
ETA_dEpsilon_D	0.389612	0.062396
ETA_Shape_Y	0.912192	0.399887
ETA_BetaP_s	0.60827	0.118229
ETA_EtaP_F	0.156509	0.243636
nHBAcc	0.763162	50.36871
nAtomLAC	0.886048	15.94887
MDEC-14	0.861766	4.024826
MDEC-22	0.385049	13.60019
MDEC-23	0.253417	13.6896
MDEC-24	0.187219	2.285725
MDEC-33	0.722608	35.96141
MDEO-11	0.339665	1.592459
MDEO-12	0.690197	2.438236
MLFER_A	0.884812	7.618233
MLFER_BH	0.682306	15.75445
MLFER_S	0.983677	20.57753
TopoPSA	0.810281	968.5693
WTPT-3	0.939527	169.2643
WTPT-4	0.836226	42.47587
WTPT-5	0.580833	73.03263
XLogP	0.921274	16.46778
kappam3	0.591715	38.51945
phi	0.069879	4.911113
LDI	0.839175	0.266858

Table 4: Continued.

Descriptor	Normalized	Real
Mnc	0.401081	0.122731
QNss	0.717554	3.800167
QCss	0.401335	0.675045
QHss	0.816212	1.412047
Qmin	0.458487	0.197608
QOmin	0.902885	0.557983
QNmin	0.399317	0.186082
QOmax	0.2189	0.111201
QNmax	0.733215	0.462658
QCmax	0.219663	0.114005
QHmax	0.309301	0.08877
mChi1	0.70787	0.059461
knotp	0.463493	3.682452
Chiv3	0.504351	12.68139
dchi0	0.531787	14.54385
Chiv5ch	0.373625	0.159538
Chiv6ch	0.813352	0.252139
knotpv	0.332321	1.601785
naccr	0.252895	8.092637
PC6	0.68057	207.574
S17	0.617016	13.29916
S25	0.41233	3.873014
S34	0.471534	30.94445
Smax11	0.012002	0.028529
Smax16	0.270107	0.819234
Smax24	0.166683	0.618393
Smax35	0.118755	0.798862
Smin8	0.95032	2.519298
Smin12	0.345704	1.386271
Smin15	0.094406	0.351002
Smin24	0.795928	2.952893
Smin33	0.561019	6.58187
Smin34	0.555043	7.947103
Smin35	0.727286	4.891723
Scar	0.881448	89.12223
Smax	0.089706	1.067416
Smin	0.953752	6.591381
GATSm1	0.812402	1.001692
GATSm2	0.634299	0.793508
GATSm3	0.120694	0.223285
GATSm4	0.510951	0.973362
GATSm5	0.049626	0.135925
GATSm8	0.100571	0.559475
GATSv1	0.961747	0.951168
GATSv3	0.282006	0.472642
GATSe1	0.927808	0.905541
GATSe5	0.599223	1.181668
GATSe7	0.23859	1.199867
GATSe8	0.078374	0.43999

Table 4: Continued.

Descriptor	Normalized	Real
GATSp3	0.437899	0.708521
GATSp4	0.924518	1.591095
GATSp7	0.980757	4.932225
TPSA1	0.606654	694.4674
slogPVSA0	0.062831	13.0085
slogPVSA1	0.216952	76.48269
slogPVSA2	0.700064	67.12981
MRVSA5	0.64033	57.65721
MRVSA6	0.374169	51.50251
PEOEVSA1	0.302775	42.13574
PEOEVSA5	0.643223	105.9755
PEOEVSA6	0.965518	159.7043
PEOEVSA7	0.010657	0.809047
PEOEVSA8	0.929568	51.26194
EstateVSA1	0.420173	90.7544
EstateVSA4	0.150524	17.979
EstateVSA7	0.617028	85.38316
VSAEstate1	0.916056	260.2258
VSAEstate2	0.469295	61.3763
VSAEstate4	0.182115	5.816037
VSAEstate7	0.434015	9.820449
MATSm1	0.259205	0.199588
MATSm3	0.55142	0.600496
MATSm5	0.41914	0.43381
MATSm6	0.11384	0.224834
MATSm7	0.309463	0.536608
MATSm8	0.835615	8.049475
MATSv1	0.691362	0.486719
MATSv3	0.421519	0.464936
MATSv5	0.803173	1.224839
MATSv8	0.974739	10.77086
MATSe1	0.044383	0.029914
MATSe3	0.858027	0.967855
MATSe5	0.820808	0.897964
MATSe8	0.84285	8.119175
MATSp1	0.618676	0.638474
MATSp3	0.031658	0.035141
MATSp4	0.995234	1.393328
MATSp5	0.310054	0.464771
MATSp6	0.950285	1.76563
MATSp7	0.401141	0.920218
ATSv5	0.422807	1.369895
ATSe5	0.526873	1.836152
ATSp4.1	0.970895	2.954435
ATSp5.1	0.441556	1.408123
J	0.517201	2.905116
BertzCT	0.331505	0.429631
IDE	0.017944	0.054297
LogP	0.809566	20.63664

Table 4: Continued.

Descriptor	Normalized	Real
LogP2	0.935064	170.3864
Hy	0.258499	0.832883
CIC1	0.531694	2.274054
CIC3	0.195279	0.597359
CIC4	0.947302	2.624027
CIC6	0.315254	0.833533
SIC1	0.534191	0.238783
SIC2	0.548336	0.223173
SIC3	0.93592	0.379048
IC1	0.62343	1.283018
IC2	0.961285	2.260943
IC3	0.605757	1.788194
bcutm13	0.686797	0.80836
bcutm12	0.096662	0.106618
bcutm9	0.662882	0.380495
bcutm3	0.045307	0.169084
bcutm2	0.449717	3.285183
bcutv4	0.772798	0.863988
bcutv1	0.57612	0.338759
bcute12	0.330634	0.366674
bcute10	0.721806	0.510317
bcute9	0.529904	0.349736
bcute4	0.370874	0.411299
bcute2	0.741989	0.470421
bcute1	0.128208	0.061027
bcutp12	0.873497	0.828075
bcutp9	0.541204	0.357195
bcutp5	0.669777	0.953092
bcutp1	0.884249	0.579183

to 500. As for the chromosomes, the number is set to 224, the length of each chromosome is set to 20 bits, the crossover rate is set to 0.6, and the mutation rate is set to 0.1. Record the value of the operand variable that maximizes the objective function, and return it.

4. Experimental Results and Discussions

4.1. Analysis of the Biological Activity Prediction Model. Once the network has been trained, the prediction can be made by simply feeding the network the values of the main variables. The validation set was imported into the model after establishing the quantitative neural network-based prediction model of biological activity. The predicted outcomes were compared with their real labels, which are displayed in Table 2.

The change curve of the predicted value and its corresponding actual value are similar in Figure 4, which shows that the model has a decent prediction result and is able to accurately reflect the biological activity in theory. The mean square error of the model is 1.572, which is within the acceptable range when choosing MSE to measure the prediction accuracy.

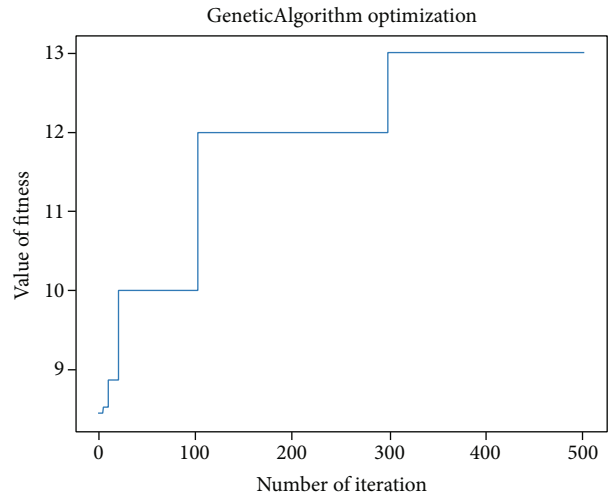


Figure 5: The figure of best fitness trending.

4.2. Analysis of the ADMET Property Prediction Model. Compare the predicted results with the actual results by importing the validation set into the trained ADMET property prediction model. The following table display the findings (Table 3).

The verification results for all five attributes are acceptable as considering the aforementioned tables, while the results for HOB are a little inferior. Nevertheless, the accuracy rate is still quite good. The prediction accuracy of HOB is the lowest result in terms of these five attributes, and this fact might be caused by data imbalance, since the neural network-based models tend to develop a preference on biased data. However, our model's average prediction accuracy is close to 85%, which is quite a satisfactory performance.

4.3. Analysis of the Clinical Property Optimizing Model. By resolving the optimization problem, the optimal fitness value of 13 is discovered and the relevant actual values for the molecular descriptors are resolved. Table 4 shows the values of the first 224 molecular descriptors in detail.

It can be found that due to the inconsistency of the definitions among the descriptors, the value difference is relatively large but it seems to not affect the results at last. Consider Figure 5, since the dataset that we created has a maximum fitness value of 12.86 while an optimal fitness value of 13 that could be attained by solving the problem; this fact proves that, by applying our method to existing chemical data, it might be possible to find a candidate which has better properties.

5. Conclusion

Breast cancer, as a common and influential disease, requires the development of new drugs to continuously improve the treatment methods. How to efficiently select possible drug

candidates to reduce the cost of drug development has a certain research value. In our work, we consider the use of machine learning methods to assist in the selection of compounds. We first used the known knowledge and computer compound molecular descriptor calculation software to construct a dataset and then used the random forest algorithm to screen the features and simplify the dataset; then, based on the MobileNetV3 structural deep convolutional network, the biological activity and pharmacokinetics were constructed.

The invention of novel medications is necessary for the ongoing development of effective treatments for breast cancer, a common and notable disease. There may be some research value in how to effectively choose potential medication candidates to lower the cost of drug development. In our study, we take into account the application of machine learning techniques to aid in compound selection. First, a dataset was created by combining already known inhibition potency knowledge and the compound molecular descriptors generated by modern calculation software. Next, features were screened out and the dataset was made simpler using the random forest algorithm. Finally, the MobileNetV3 structural deep convolutional network was introduced to construct the biological activity and pharmacokinetics. A genetic algorithm solver is utilized to solve an optimization problem based on the obtained prediction model to predict the best value for the chosen molecular descriptors. The analysis from the perspective of the entire drug design process, rather than constructing or modifying each child model, reflects the proposed model's innovation and applicability the most. The internal connection and progressive relationship of each model are emphasized in many places throughout this paper. We believe that our four-step method of influencing variable screening, biological activity prediction modeling, pharmacokinetic properties modeling, and clinical property optimization can successfully model and optimize the properties of drugs through various machine learning technologies and serve as a useful guide for drug manufacturers. According to the aforementioned analysis, our method not only offers a significant practical industrial application value but also some academic innovation and research value.

In the future, our research direction will mainly focus on giving weight to the properties of ADMET. For example, for the properties of hERG, we do not want the drug to be highly toxic to the human body, so for toxic compounds, we will appropriately reduce its evaluation, that is, the calculated adaptation value. Other properties are the same, and we expect to obtain antibreast cancer drugs with better efficacy and less harm to the human body through this method.

Data Availability

The datasets used during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. G. Waks and E. P. Winer, "Breast cancer treatment," *JAMA*, vol. 321, no. 3, pp. 288–300, 2019.

[2] M. J. Duffy, "Estrogen receptors: role in breast cancer," *Critical Reviews in Clinical Laboratory Sciences*, vol. 43, no. 4, pp. 325–347, 2006.

[3] S. Ali and R. C. Coombes, "Estrogen receptor alpha in human breast cancer: occurrence and significance," *Journal of Mammary Gland Biology and Neoplasia*, vol. 5, no. 3, pp. 271–281, 2000.

[4] M. Sarhadi, L. Aryan, and M. Zarei, "The estrogen receptor and breast cancer: a complete review," *CRPASE Transactions of Applied Sciences*, vol. 6, pp. 309–314, 2020.

[5] W. P. Bocchinfuso and K. S. Korach, "Mammary gland development and tumorigenesis in estrogen receptor knockout mice," *Journal of Mammary Gland Biology and Neoplasia*, vol. 2, no. 4, pp. 323–334, 1997.

[6] V. C. Jordan, "Tamoxifen: a most unlikely pioneering medicine," *Nature Reviews Drug Discovery*, vol. 2, no. 3, pp. 205–213, 2003.

[7] C. K. Osborne, "Tamoxifen in the treatment of breast cancer," *New England Journal of Medicine*, vol. 339, no. 22, pp. 1609–1618, 1998.

[8] B. Furr and V. Jordan, "The pharmacology and clinical uses of tamoxifen," *Pharmacology & Therapeutics*, vol. 25, no. 2, pp. 127–205, 1984.

[9] X. Fang, J. Cao, and A. Shen, "Advances in anti-breast cancer drugs and the application of nano-drug delivery systems in breast cancer therapy," *Journal of Drug Delivery Science and Technology*, vol. 57, article 101662, 2020.

[10] A. Emadi, R. J. Jones, and R. A. Brodsky, "Cyclophosphamide and cancer: golden anniversary," *Nature Reviews Clinical Oncology*, vol. 6, no. 11, pp. 638–647, 2009.

[11] L. Gianni, T. Pienkowski, Y. H. Im et al., "Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial," *The Lancet Oncology*, vol. 13, no. 1, pp. 25–32, 2012.

[12] M. Capelan, L. Pugliano, E. de Azambuja et al., "Pertuzumab: new hope for patients with HER2-positive breast cancer," *Annals of Oncology*, vol. 24, no. 2, pp. 273–282, 2013.

[13] S. M. Swain, J. Baselga, S. B. Kim et al., "Pertuzumab, trastuzumab, and docetaxel in HER2-positive metastatic breast cancer," *New England Journal of Medicine*, vol. 372, no. 8, pp. 724–734, 2015.

[14] R. Gozalbes, J. Doucet, and F. Derouin, "Application of topological descriptors in QSAR and drug design: history and new trends," *Current Drug Targets-Infectious Disorders*, vol. 2, no. 1, pp. 93–102, 2002.

[15] J. Verma, V. M. Khedkar, and E. C. Coutinho, "3D-QSAR in drug design - a review," *Current Topics in Medicinal Chemistry*, vol. 10, no. 1, pp. 95–115, 2010.

[16] Y. Wang, J. Xing, Y. Xu et al., "In silicoADME/T modelling for rational drug design," *Quarterly Reviews of Biophysics*, vol. 48, no. 4, pp. 488–515, 2015.

[17] F. Ghasemi, A. Mehridehnavi, A. Perez-Garrido, and H. Perez-Sanchez, "Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks," *Drug Discovery Today*, vol. 23, no. 10, pp. 1784–1790, 2018.

[18] S. P. Niculescu, "Artificial neural networks and genetic algorithms in QSAR," *Journal of Molecular Structure: THEOCHEM*, vol. 622, no. 1–2, pp. 71–83, 2003.

[19] W. Cui, A. Aouidate, S. Wang, Q. Yu, Y. Li, and S. Yuan, "Discovering anti-cancer drugs via computational methods," *Frontiers in Pharmacology*, vol. 11, p. 733, 2020.

[20] G. R. Marshall, "Computer-aided drug design," *Annual Review of Pharmacology and Toxicology*, vol. 27, no. 1, pp. 193–213, 1987.

[21] J. Xiong, Z. Xiong, K. Chen, H. Jiang, and M. Zheng, "Graph neural networks for automated _de novo_ drug design," *Drug Discovery Today*, vol. 26, no. 6, pp. 1382–1393, 2021.

[22] C. H. Liu, M. Korablyov, S. Jastrzębski, P. Włodarczyk-Pruszyński, Y. Bengio, and M. H. Segler, "RetroGNN: approximating retrosynthesis by graph neural networks for de novo drug design," 2020, http://arxiv.org/abs/2011.13042.

[23] A. Howard, M. Sandler, G. Chu et al., "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1314–1324, Seoul, Korea, 2019.

[24] X. Lu, H. Wu, and Y. Zeng, "Classification of Alzheimer's disease in MobileNet," *Journal of Physics: Conference Series*, vol. 1345, no. 4, article 042012, 2019.

[25] W. Sae-Lim, W. Wettayaprasit, and P. Aiyarak, "Convolutional neural networks using MobileNet for skin lesion classification," in *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, pp. 242–247, Chonburi, Thailand, 2019.

[26] D.-S. Cao, Q.-S. Xu, Q.-N. Hu, and Y.-Z. Liang, "ChemoPy: freely available python package for computational biology and chemoinformatics," *Bioinformatics*, vol. 29, no. 8, pp. 1092–1094, 2013.

[27] S. Patro and K. K. Sahu, "Normalization: A preprocessing stage," 2015, http://arxiv.org/abs/1503.06462.

[28] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[29] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, http://arxiv.org/abs/1409.1556.

[30] L. Bottou, "Large-scale machine learning with stochastic gradient descent," *Proceedings of COMPSTAT'2010*, , pp. 177–186, Springer, 2010.

[31] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," 2014, http://arxiv.org/abs/1412.6980.

[32] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," 2017, http://arxiv.org/abs/1711.05101.

[33] D. Whitley, "A genetic algorithm tutorial," *Statistics and Computing*, vol. 4, no. 2, pp. 65–85, 1994.

[34] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Computational Intelligence Magazine*, vol. 1, no. 4, pp. 28–39, 2006.

[35] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, pp. 1942–1948, Perth, WA, Australia, 1995.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8332278, 13 pages
https://doi.org/10.1155/2022/8332278

Research Article

An Untargeted Lipidomics Study of Acute Ischemic Stroke with Hyperglycemia Based on Ultrahigh-Performance Liquid Chromatography-Mass Spectrometry

Jia Guo, Hailan Wang, Xin Jiang, Yan Wang, Zhihao Zhang, Qingbin Liao, and Jia Xu

Department of Geriatrics, Dadukou District People's Hospital, Chongqing 400084, China

Correspondence should be addressed to Jia Xu; 879737168@qq.com

Received 9 July 2022; Revised 1 August 2022; Accepted 4 August 2022; Published 26 August 2022

Academic Editor: Xiaohan Ren

Patients with type 2 diabetes have twice as much of the risk of acute ischemic stroke (AIS) occurrence as healthy individuals, and the AIS patients with type 2 diabetes have a higher risk of death and a poorer prognosis. This study was to investigate the interrelationship between hyperglycemia and AIS and provided a reference for blood glucose management of AIS patients. The blood glucose level of AIS patients of the present study was controlled by insulin below 180 mg/dL (standard group) and between 80 and 130 mg/dL (management group). And the fasting venous blood samples were collected for determination of blood glucose level, homeostasis model assessment of insulin resistance (HOMA-IR), peptide C, and basal insulin level. Furthermore, lipids of the blood samples were detected using metabolomics, so as to clarify the similarities and differences in metabolic patterns in AIS patients with diabetes after the intervention of different glycemic strategies. The results revealed that compared to the standard group, the blood glucose level and HOMA-IR in the management group were significantly decreased, and levels of peptide C and basal insulin level were greatly increased. Through lipidomics detection, 83, 50, and 44 types of significantly upregulated differential lipids were detected in the standard vs. normal groups, the standard vs. management groups, and the management vs. normal groups, respectively, with triacylglycerol dominated. This study preliminarily revealed metabolic differences among AIS patients with hyperglycemia after different blood glucose intervention methods, hoping to provide a theoretical basis for clinical prevention and treatment of this disease.

1. Introduction

Acute ischemic stroke (AIS) is recognized as a serious brain vessel disease with extremely high fatality and disability rates [1], and the risk of AIS in patients with type 2 diabetes is about twice that of healthy individuals [2]. Diabetes treatment can effectively reduce the severity of AIS and improve the prognosis of patients [3]. Glucose control is a challenge during the acute care of AIS patients currently, which contributes to neuroprotection according to several studies [4]. Despite the fact that hyperglycemia is regarded as a risk factor for AIS patients, doubts on the effectiveness of blood sugar control has been raised by some scholars [5]. Research on Chinese AIS patients has indicated that blood sugar control below 200 mg/dL is necessary [6]. Some researchers have proposed that neuroprotective measures for AIS patients with hypoglycemia should be based on vascular thrombolysis and imaging-assisted classification [7]. Most previous studies on blood glucose in AIS patients have focused on the detection of blood glucose, high- and low-density lipoproteins, and the causal relationship between AIS and hyperglycemia is still unclear. In recent years, significant progress has been made in the metabolomics of small molecules. Typically, an untargeted metabolomics approach is aimed at characterizing global metabolism and quantifying hundreds or thousands of small- and medium-sized molecules, mainly in urine, serum, and tissue extracts [8]. Untargeted lipidomics can fully collect information in metabolic networks, thereby depicting a more comprehensive profile of metabolic mechanisms of disease. As the continuous improvement of databases, untargeted metabolomics research based on mass spectrometry is more convenient

Table 1: Criteria for inclusion and exclusion of AIS subjects.

Inclusion criteria	Exclusion criteria
(I) AIS diagnosed by brain CT or MRI (II) The time from onset to hospital admission was within 24 hours (III) No history of stroke (IV) Clear consciousness (V) The patients and their family members agreed and signed the informed consent, which was approved by the ethics committee of this hospital	(I) Patients in a critical condition or combined with severe trauma and uncooperative with examinations (II) Patients with severe cardiopulmonary disease requiring complex and comprehensive testing and treatment (III) Patients with respiratory diseases or liver and kidney failures (IV) Patients with history of mental illness (V) Patients under hormone or immune stimulant administration in recent time or currently (VI) Patients who discontinued the study

Abbreviations: CT: contrast-enhanced cerebral computed tomography; MRI: magnetic resonance imaging.

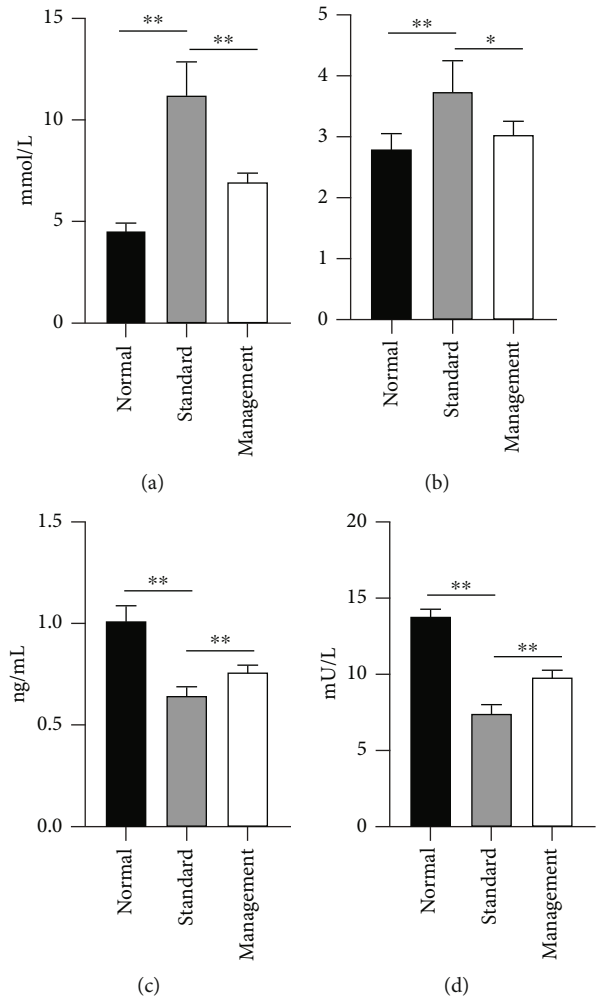


Figure 1: Results of blood biochemical indicators: (a) fasting blood glucose level; (b) HOMA-IR; (c) peptide C level; (d) basal insulin level. $^{*}p < 0.05$ and $^{**}p < 0.01$.

with a wider application than the method based on nuclear magnetic resonance [9].

Unluckily, previous research and the current clinical treatment protocols are not complete; it is of great practical significance to clarify the causal relationship between AIS and hyperglycemia and to investigate blood glucose control in the acute phase of AIS patients. In light of previous research findings and treatment guidelines, the present project classified the AIS patients into three groups: AIS normal group with normal glucose, AIS diabetes-standard group, and AIS diabetes-management group. In terms of hypoglycemic measures, the management group was treated using

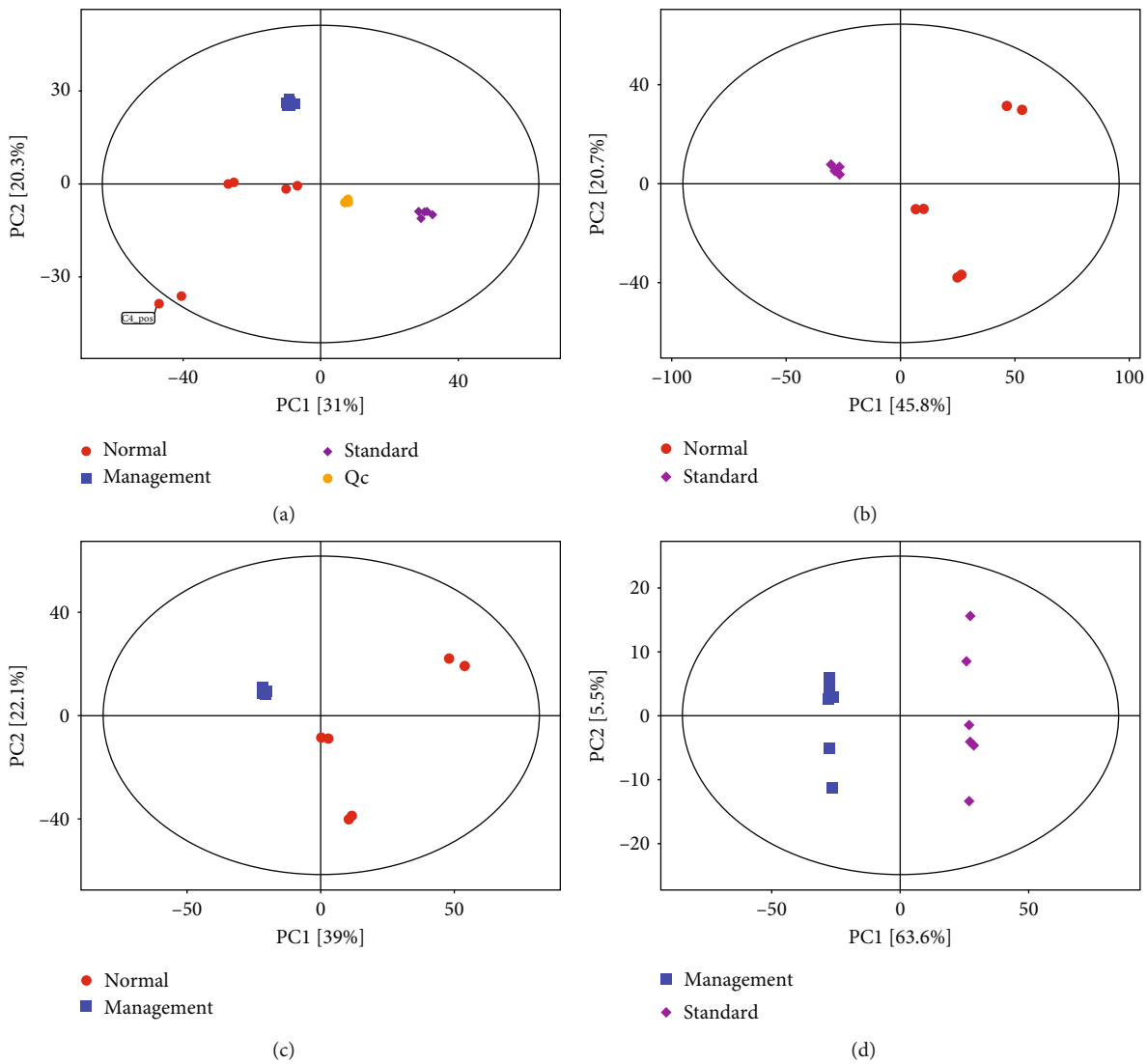


Figure 2: Scatter point diagrams based on PCA: (a) standard, management, and normal groups and QC samples; (b) normal group vs. standard group; (c) normal group vs. management group; (d) management group vs. standard group. Red dots represent the normal group, purple dots represent the standard group, blue dots represent the management group, and yellow dots represent QC.

a strict blood sugar control scheme and the standard group was treated using a standard blood sugar control scheme. Simultaneously, based on general blood biochemistry, metabolomics was applied to detect small molecular substances in patients and to explore the differences in lipid metabolism of AIS diabetic patients in each group at the molecular level.

2. Materials and Methods

2.1. Instruments, Reagents, and Materials. The instruments included the 1290 ultrahigh-performance liquid chromatography instrument (UHPLC) from Agilent, USA; Q Exactive Focus High-Resolution Mass Spectrometer, Heraeus Fresco 17 Centrifuge, and NanoDrop One Microplate Reader from Thermo Fisher, USA; and SYNCHRON LX-20 Biochemistry Analyzer from Beckman Coulter, USA. The reagents and materials included insulin kit (RX104930H) and serum peptide C kit (RX106372H). Both were purchased from Quanzhou Ruixin Biotechnology Co., Ltd. LC-MS, methanol, acetonitrile, ammonium formate, and dichloromethane were purchased from CNW, Germany.

2.2. Test Grouping. The patients with fasting blood glucose ≥ 6.1 mmol/L were divided into the hyperglycemia group, and those with fasting blood glucose < 6.1 mmol/L were divided into the normal group. The hyperglycemia group was subdivided into standard and management groups according to their history of diabetes. There were a total of 54 patients included, with 18 cases in each of the described three groups: AIS normal group with normal glucose, AIS diabetes-standard group, and AIS diabetes-management group, respectively. The inclusion criteria are shown in Table 1. Blood glucose of the patients was measured and recorded every 8 hours. Patients in the AIS normal group with normal glucose were given no treatment; patients in the standard group were injected with insulin once the blood

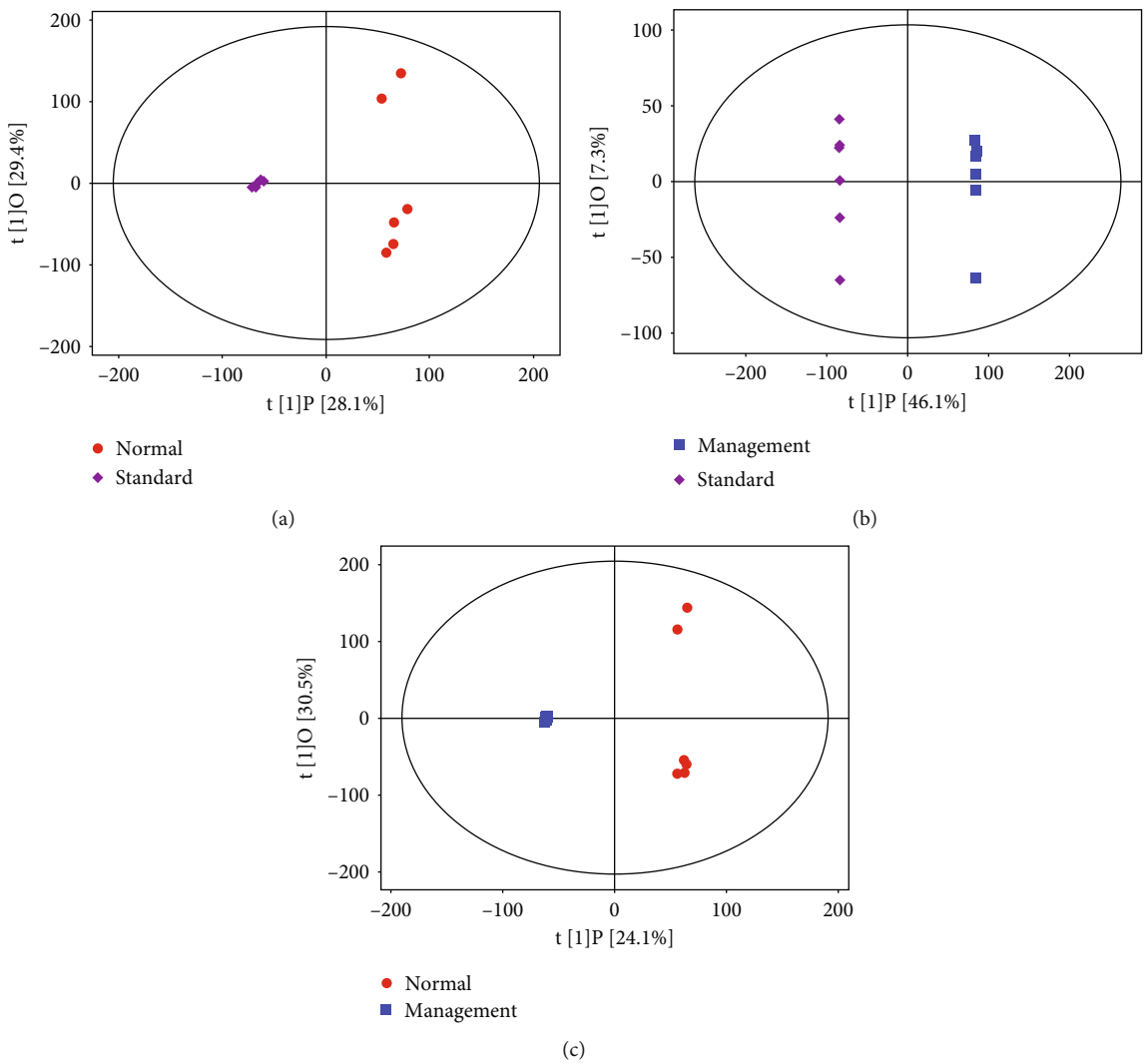


Figure 3: Scatter point diagrams based on OPLS-DA: (a) normal group vs. standard group; (b) standard group vs. management group; (c) management group vs. normal group. Red dots represent the normal group, purple dots represent the standard group, and blue dots represent the management group.

sugar level increased above 180 mg/dL, and the dose was increased if the target concentration was not reached within 12 h; the blood sugar concentration of the management group was strictly controlled within the 80-130 mg/dL range using insulin.

2.3. Determination of Blood Biochemical Indicators. The patients were treated with different management measures for blood sugar control, and fasting venous blood was collected 24 h later for the detection of blood biochemical indicators. Fasting blood glucose levels were detected using the SYNCHRON Biochemical Analyzer. The basal level of insulin and the concentration of C-peptide were detected using the test kit, performed strictly in accordance with the instructions of use. Insulin resistance index (HOMA-IR) = (fasting glucose level × basal insulin level)/22.5.

2.4. Sample Pretreatment. The complete blood samples were performed centrifugation at 4000 g 4°C for 15 min, and 30 μL of the supernatant was aspirated to a new tube for further extraction. Following the addition of 90 μL of precooled methanol-acetonitrile ($v/v = 1:1$) to the serum, the solution was vortexed for 1 min and kept in an ice-water bath for 15 min. The mixed solution was centrifuged at 12,000 g, 4°C for 15 min, and 100 μL of the supernatant was collected and kept in a sample vial for subsequent analysis.

2.5. LC-MS Sample Loading Detection. The samples were analyzed by ultrahigh-performance liquid chromatography-mass spectrometry (UHPLC-MS) technique, and the vacuum-dried solids were redissolved using 100 μL of 50% methanol. After centrifugation at 12,000 g 4°C, the supernatant was transferred to a sample vial. Chromatographic separation was performed using an ultrahigh-performance liquid chromatography system packed with a BEHC18 column (2.1 mm × 100 mm, 1.7 μm). The prepared sample volume was 5 μL, mobile phase A was 0.1% formic acid-acetonitrile, and mobile phase B was 0.1% formic acid-

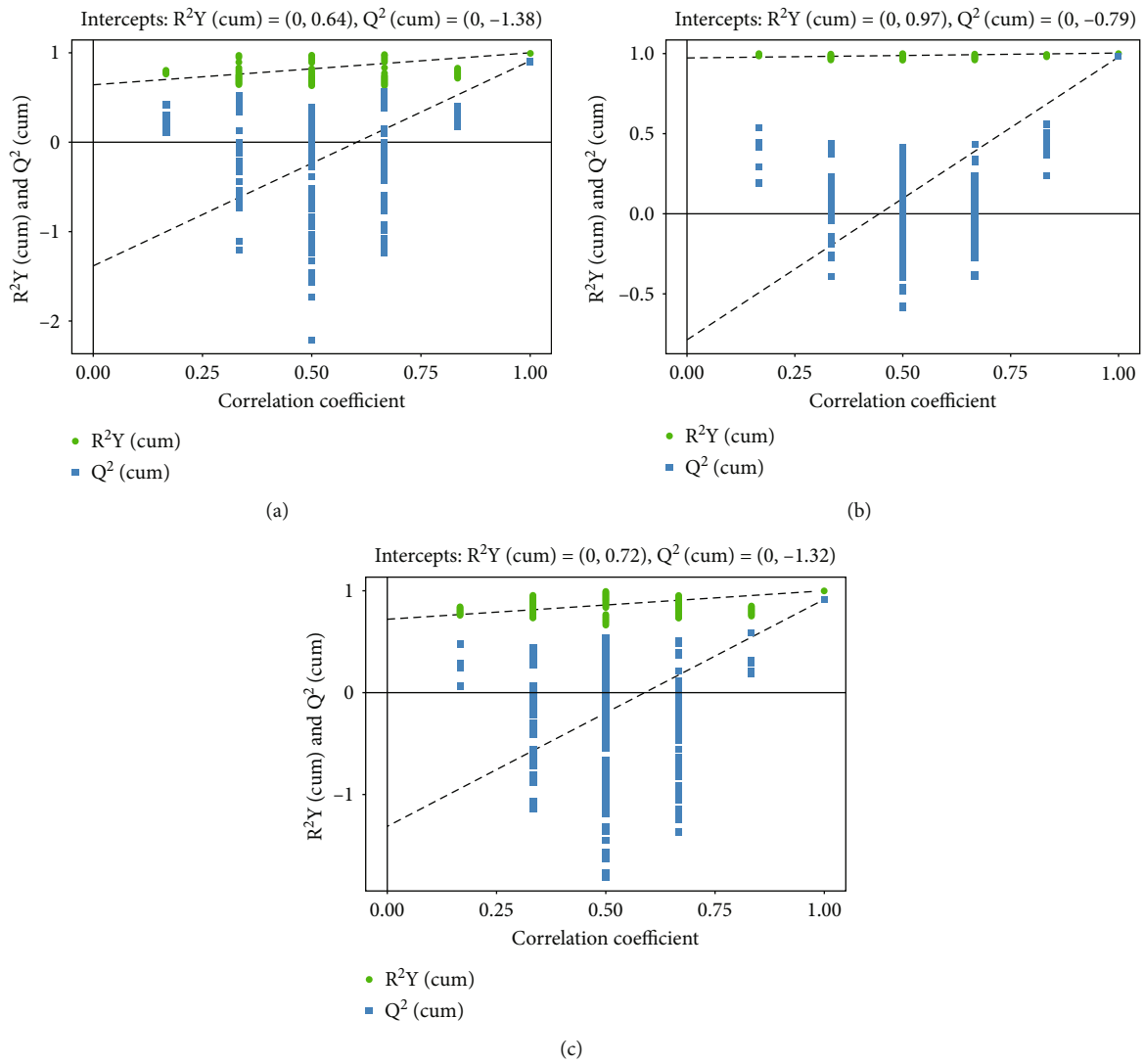


Figure 4: Permutation test plots based on OPLS-DA: (a) normal group vs. standard group; (b) standard group vs. management group; (c) management group vs. normal group.

water. The flow rate was 0.35 mL/min, and the column temperature was 40°C. The analysis was conducted with chromatographic gradient as follows: 0-0.5 min, 1% A; 0.5-3.5 min, 1-53% A; 3.5-7.5 min, 53-70% A; 7.5-9 min, 70-90% A; 9-13 min, 90% A; 13.1-15 min, and back to 1% A. To ensure a steady system, the quality control samples were used at the beginning and end of each analysis. Mass spectrometry data were collected using ESI source in positive and negative ion modes, and the specific scanning parameters were capillary voltage 4 kV in positive ion mode, 3.5 kV in negative ion mode, atomizing gas temperature at 330°C, atomizing gas flow rate 10 L/min, splitting voltage at 100 V, screening voltage at 65 V, scanning range m/z 70-1100, and scanning speed 1.5 scan/sec.

2.6. Bioinformatics Analysis. The LC-MS data preprocessing included peak detection, extraction, alignment, and integration which were mainly implemented using the xcms toolkit of R platform. The bandwidth was set to 15 sec, the peak broadening 5-30 sec, and other parameters default. The preprocessed data were imported into EZinfo software in matrix for principal component analysis (PCA) and orthogonal partial least squares-discriminant analysis (OPLS-DA); and the data were imported into MetaboAnalyst for hierarchical clustering analysis (HCA). Kruskal-Wallis tests, Student's t-tests, and logistic regression analysis were all implemented using the R software to assess the influence of confounding factors.

2.7. Statistical Analysis. The blood biochemical index data were statistically analyzed using GraphPad Prism 9.0 software, and the results were expressed as mean ± standard deviation (SD). One-way ANOVA was applied for significant analysis; when $p < 0.05$, the difference was considered significant.

3. Results

3.1. Differences in Blood Biochemical Indexes among Groups. The fasting glucose levels and HOMA-IR levels of the

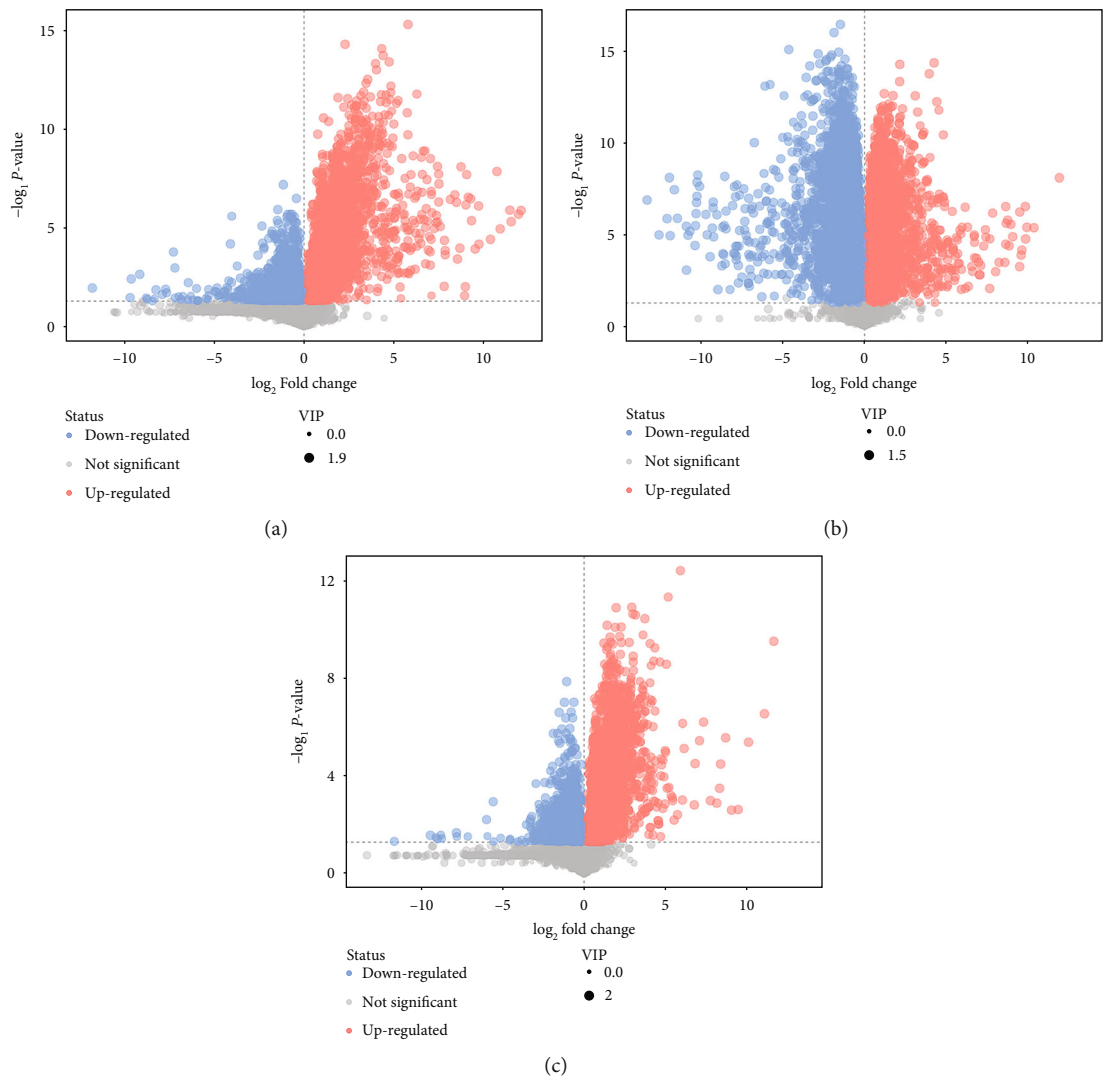
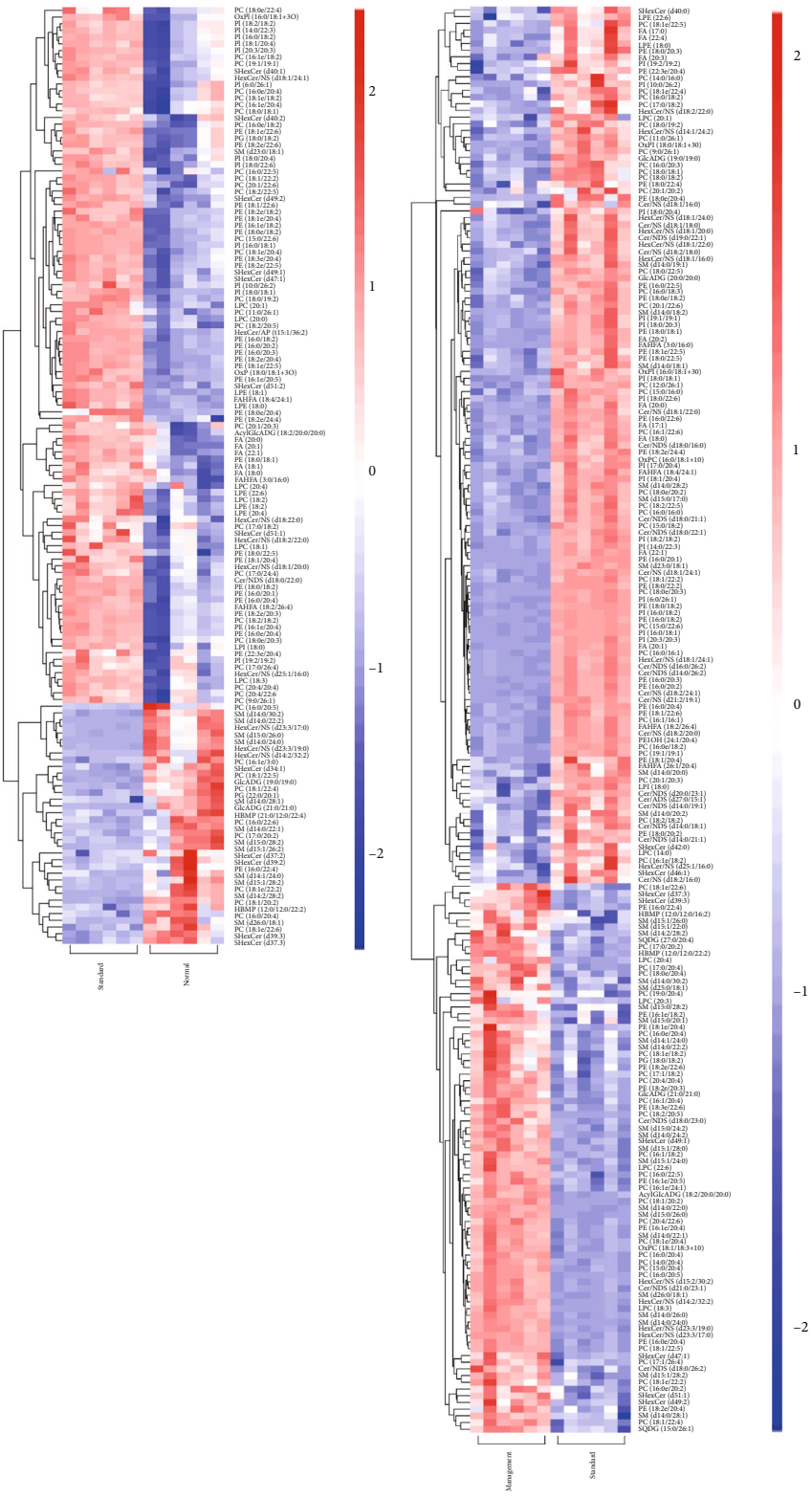


Figure 5: Volcano maps of differential lipid screening: (a) normal group vs. standard group; (b) standard group vs. management group; (c) management group vs. normal group.

standard group were significantly higher than those of the normal group ($p < 0.01$) and the management group ($p < 0.05$) (Figures 1(a) and 1(b)), while the serum peptide C and basal insulin levels were markedly lower than those of the control and management groups ($p < 0.01$) (Figures 1(c) and 1(d)).

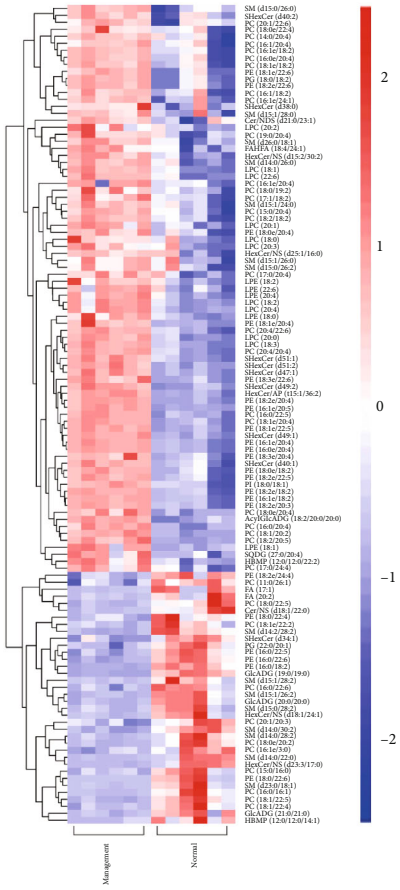
3.2. Multivariate Statistical Analysis. Sample PCA was conducted using multivariate statistics. Figure 2(a) presents $R^2X = 0.647 > 0.4$, and QC was clustered together, indicating that the instrument was stable during the sample collection and the data reproducibility was satisfactory which were applicable for subsequent analysis. PCA revealed that normal vs. standard, normal vs. management, and management vs. standard groups had the tendency to separate and were clearly clustered into two groups to be distinguished effectively, indicating that there was a significant difference (Figures 2(b)–2(d)). The scatter point diagram of the OPLS-DA model indicated evident separation of the normal group and the AIS diabetic group. In Figures 3(a)–3(c), R^2X were 0.575, 0.534, and 0.546; R^2Y were 0.991, 1.000, and 0.998; Q^2 were 0.904, 0.981, and 0.913, indicating that the predicted probabilities of the model were 90.4%, 98.1%, and 91.3%, respectively. Permutation test was performed on the model also with the results presented in Figures 4(a)–4(c). The y-axis intercept of Q^2 in each group was all below 0, implying that the model had been fitted and differential metabolites could be screened accordingly.

3.3. Identification of Differential Lipid Metabolism. Through the above analysis, combined with the statistical results of univariate and multivariate analysis, differential metabolites were screened out. To present a better visualization of the differences among groups, a volcano plot (Figures 5(a)–5(c)), cluster analysis (Figures 6(a)–6(c)), and a lipid group bubble plot (Figures 7(a)–7(c)) were constructed based on the identified lipid metabolites, and there were substantial differences in lipid products and evident hierarchical clustering among each of the groups. A total of 83 differentially upregulated lipid metabolites were screened from the



(a) (b)

Figure 6: Continued.



(c)

Figure 6: Heat maps of hierarchical cluster analysis: (a) normal group vs. standard group; (b) standard group vs. management group; (c) management group vs. normal group.

standard group compared with the normal group based on variable importance in projection (VIP) > 1, $p < 0.05$ of t-tests, and fold change (FC) > 2 of the maximum difference fold. Eighty-three significantly upregulated differential lipid metabolites were screened from the standard group compared with the normal group and were classified into nine categories. They were divided into nine categories: acylcarnitine (Acar), diacylglycerol (DAG), diacylglyceryl trimethylhomoserine (DGTS), hexosylceramide nonhydroxy fatty acid-sphingosine (HexCer/NS), lysophosphatidylcholine (LPC), lysophosphatidylethanolamine (LPE), phosphatidylcholine (PC), sulfur hexosylceramide hydroxy fatty acid (SHexCer), and triacylglycerol (TAG). A total of 50 differentially upregulated lipid metabolites were screened from the standard group compared with the management group, which were classified into eight categories: cholesteryl ester (CE), ceramide nonhydroxy fatty acid-dihydrosphingosine (Cer/NDS), ceramide nonhydroxyl fatty acid-sphingosine (Cer/NS), HerCer/NS, LPC, PC, sphingomyelin (SM), and TAG. A total of 44 differentially upregulated lipid metabolites were sorted out from the management group compared with the normal group, which were grouped into eight categories: acyl glucuronic acid glycolipid (AcylGlcADG), CE, DGTS, HerCer/NS, LPC, PC, SHexCer, and TAG. TAG dominated the differential metabolites screened from the three groups, and the specific differential metabolites are shown in Table 2.

4. Discussion

AIS has a very high morbidity and mortality, and it is often accompanied by hyperglycemia in the acute phase [10]. Many studies have indicated the interrelationship between hyperglycemia and poor clinical outcomes in AIS patients [11]. At present, there is no unified standard for the definition of hyperglycemia after AIS onset at home and abroad, and the majority of scholars set the blood glucose threshold as 6-8 mmol/L [12, 13]. Increasing evidence has revealed that hyperglycemia increases infarct size, risk of hemorrhagic transformation after reperfusion, and mortality in AIS patients, and there is a linear relationship between post-AIS hyperglycemia and poor clinical outcomes [14]. Some researchers have pointed out that there is a high mortality rate within 72 hours of onset for AIS patients with the admission blood glucose greater than 7.9 mmol/L [15]. Hyperglycemia adversely affects the prognosis of AIS patients, and the majority of cases develop cerebral infarction with large vessel occlusion.

Fasting glucose levels are commonly used clinical indicators for monitoring the degree of glucose metabolism and

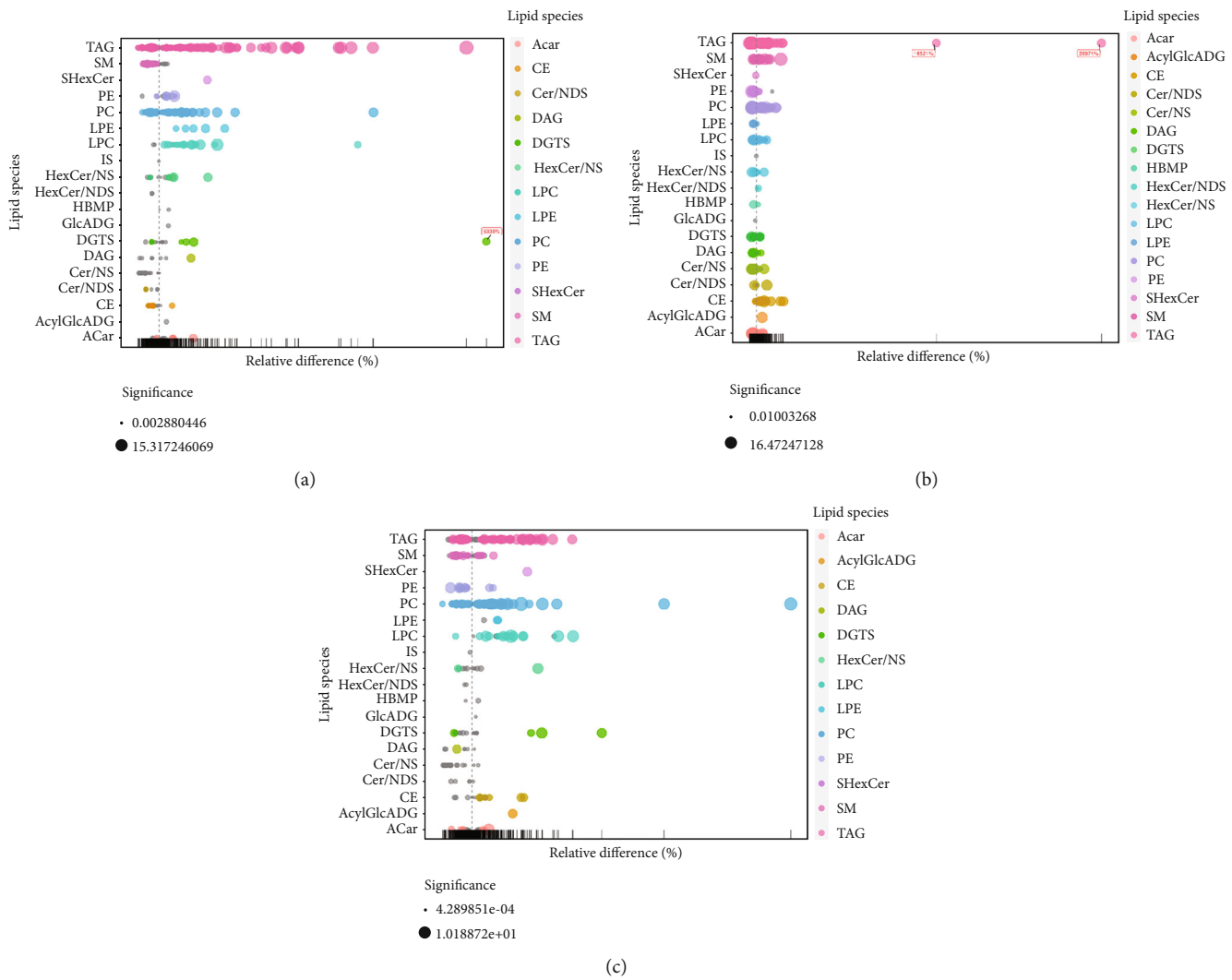


Figure 7: Bubble plots of lipid analysis: (a) normal group vs. standard group; (b) standard group vs. management group; (c) management group vs. normal group.

blood glucose control, which is an essential part for the prevention and treatment of diabetes complications [16]. Peptide C levels indicate if the secretory function of β cells is working well or not in the body, and it can be used as a guiding indicator for the differentiation of diabetic patients and the identification of hypoglycemia [17]. Some present studies have shown that C-peptide has a good protective effect on diabetes and its complications. It protects the functions of β cells and stimulates the direct secretion of insulin, thereby reducing high glucose load and relevant hyperglycemic effect on β cells and other cells of the entire body. Peptide C can also inhibit reactive oxygen species produced by excess glucose and fat and protect the functions of β cells through antioxidant pathways [18]. Insulin is an indispensable factor in the control of blood glucose levels, which regulates the expression and transport of glucose transporters in different types of tissues. Hence, insulin insensitivity results in increased intestinal glucose absorption, increased renal glucose reabsorption, and impaired peripheral tissue glucose uptake causing hyperglycemia. Conversely, high blood glucose level further reduces insulin sensitivity, creating a vicious circle [19]. Insulin resistance and β cell function failure are the two main causes of type 2 diabetes. As the disease progresses, more serious basal insulin secretion disorder occurs, so correcting insulin resistance is the key to controlling blood glucose levels [20].

The blood biochemical test findings indicated that the levels of fasting blood glucose and HOMA-IR in the standard group were markedly higher than those in the normal group and the management group, whereas the serum peptide C concentration and basal insulin level were substantially lower than those in the normal group and the management group. The results revealed that decreased basal insulin level impaired the absorption and processing ability of glucose. Fat cells and muscle cells cannot effectively utilize glucose substances, and liver cells fail to regulate glycogen decomposition and gluconeogenesis, resulting in an evident increase in blood glucose concentration [21, 22]. Conversely, after strict blood glucose management, the indicators in the management group differed largely from those

Table 2: Differences of lipid metabolites between groups.

Number	Standard vs. normal	Standard vs. management	Management vs. normal
1	Acar(20:1)	CE(18:1)	AcylGlcADG(22:1/22:1/14:0)
2	DAG(16:0/18:0)	CE(18:2)	CE (18:2)
3	DGTS(19:0/26:2)	CE(18:3)	CE (18:3)
4	DGTS(2:0/17:1)	CE(20:5)	DGTS(2:0/17:1)
5	DGTS(2:0/21:2)	CE(22:4)	DGTS(2:0/21:2)
6	DGTS(27:0/18:1)	Cer/NDS(d18:0/24:0)	DGTS(27:0/18:1)
7	HexCer/NS(d18:1/16:1)	Cer/NS(d18:1/24:2)	HexCer/NS(d18:1/16:1)
8	LPC(14:0)	HexCer/NS(d20:3/36:1)	LPC(14:0)
9	LPC(15:1)	LPC(20:4)	LPC(15:1)
10	LPC(18:1)	PC(14:0e/20:4)	LPC(18:1)
11	LPC(18:3)	PC(14:0e/21:0)	LPC(18:3)
12	LPC(19:0)	PC(14:0e/23:0)	LPC(19:0)
13	LPC(19:1)	PC(14:0e/24:0)	LPC(19:1)
14	LPC(20:0)	PC(16:1e/22:6)	LPC(20:0)
15	LPC(20:1)	PC(18:0e/22:3)	LPC(20:1)
16	LPC(20:5)	PC(19:1/20:4)	LPC(20:2)
17	LPC(24:1)	PC(20:3/20:3)	LPC(20:3)
18	LPE(16:0)	SM(d14:0/24:0)	LPC(20:4)
19	LPE(18:2)	SM(d14:0/26:0)	PC(14:0e/20:4)
20	LPE(20:4)	SM(d14:0/27:0)	PC(14:0e/4:0)
21	LPE(22:5)	SM(d14:0/30:1)	PC(14:1e/4:0)
22	PC(11:0/22:2)	TAG(12:1/19:1/19:1)	PC(14:1e/6:0)
23	PC(11:0/26:2)	TAG(13:0/13:0/21:4)	PC(16:1e/22:6)
24	PC(12:0/26:1)	TAG(13:0/13:0/21:5)	PC(18:2/18:2)
25	PC(14:0/22:5)	TAG(13:0/18:2/18:2)	PC(18:3/18:3)
26	PC(14:0e/20:4)	TAG(13:0/21:2/21:2)	PC(20:4/20:4)
27	PC(14:0e/22:3)	TAG(13:1/18:4/18:4)	PC(24:4/18:5)
28	PC(14:0e/4:0)	TAG(13:1/21:1/21:1)	PC(3:0/18:4)
29	PC(14:1e/4:0)	TAG(14:0/18:2/20:5)	SHexCer(d18:1/16:0)
30	PC(14:1e/6:0)	TAG(14:1/14:1/22:3)	TAG(12:0/22:0/22:0)
31	PC(16:2/22:6)	TAG(15:1/19:0/21:2)	TAG(12:1/12:1/19:5)
32	PC(18:2/18:2)	TAG(16:0/16:2/22:6)	TAG(12:1/22:0/22:0)
33	PC(18:3/18:3)	TAG(16:0/20:4/20:4)	TAG(13:0/13:0/16:0)
34	PC(20:1/20:1)	TAG(16:1/16:1/22:5)	TAG(13:1/21:1/21:1)
35	PC(21:2/21:2)	TAG(16:1/18:3/20:4)	TAG(16:0/18:1/22:0)
36	PC(22:6/22:6)	TAG(16:2/18:2/18:2)	TAG(16:0/20:0/20:0)
37	PC(24:4/18:5)	TAG(17:0/18:5/22:0)	TAG(16:0/22:1/22:1)
38	SHexCer(d18:1/16:0)	TAG(17:1/19:1/19:1)	TAG(18:0/18:0/20:0)
39	TAG(12:0/12:0/22:7)	TAG(18:0/20:4/22:6)	TAG(18:0/18:0/20:1)
40	TAG(12:0/14:0/20:2)	TAG(18:1/18:2/20:4)	TAG(18:0/18:0/22:0)
41	TAG(12:0/16:0/18:1)	TAG(18:1/18:2/22:4)	TAG(18:0/18:0/22:1)
42	TAG(12:0/22:0/22:0)	TAG(18:1/20:4/22:5)	TAG(18:0/18:1/22:1)
43	TAG(12:1/22:0/22:0)	TAG(18:1/20:4/22:6)	TAG(18:1/18:1/22:0)
44	TAG(12:2/17:0/17:0)	TAG(18:2/18:2/20:2)	TAG(18:1/18:1/22:1)
45	TAG(13:0/13:0/16:0)	TAG(18:2/18:2/20:4)	TAG(18:1/20:0/20:0)
46	TAG(13:1/20:0/20:0)	TAG(18:2/18:2/20:5)	TAG(18:1/21:0/21:0)
47	TAG(14:0/14:1/22:5)	TAG(18:2/18:2/22:6)	TAG(18:1/22:1/22:1)
48	TAG(14:0/16:0/16:0)	TAG(18:2/18:3/20:5)	TAG(18:2/18:2/22:0)

Table 2: Continued.

Number	Standard vs. normal	Standard vs. management	Management vs. normal
49	TAG(14:0/16:0/18:1)	TAG(18:2/18:3/22:6)	TAG(20:1/20:1/22:0)
50	TAG(16:0/16:0/16:0)	TAG(18:3/18:3/22:5)	
51	TAG(16:0/16:0/17:0)		
52	TAG(16:0/16:0/18:0)		
53	TAG(16:0/17:0/18:1)		
54	TAG(16:0/18:0/18:0)		
55	TAG(16:0/18:0/18:1)		
56	TAG(16:0/18:0/22:1)		
57	TAG(16:0/18:1/22:0)		
58	TAG(16:0/20:0/20:0)		
59	TAG(16:0/22:1/22:1)		
60	TAG(16:1/16:2/20:4)		
61	TAG(16:2/16:2/16:2)		
62	TAG(17:0/17:0/19:0)		
63	TAG(17:0/17:2/19:0)		
64	TAG(17:0/18:0/18:0)		
65	TAG(17:2/17:2/18:3)		
66	TAG(18:0/18:0/18:0)		
67	TAG(18:0/18:0/20:0)		
68	TAG(18:0/18:0/20:1)		
69	TAG(18:0/18:0/22:0)		
70	TAG(18:0/18:0/22:1)		
71	TAG(18:0/18:1/20:3)		
72	TAG(18:0/18:1/21:0)		
73	TAG(18:0/18:1/22:1)		
74	TAG(18:1/18:1/20:0)		
75	TAG(18:1/18:1/22:0)		
76	TAG(18:1/18:1/22:1)		
77	TAG(18:1/20:0/20:0)		
78	TAG(18:1/21:0/21:0)		
79	TAG(18:1/22:1/22:1)		
80	TAG(18:2/18:2/22:0)		
81	TAG(19:0/19:0/19:4)		
82	TAG(20:1/20:1/20:1)		
83	TAG(20:1/20:1/22:0)		

Notes: the number of carbon chains that make up the lipid and the structure of each carbon chain is indicated in parentheses. DGTS (19:0/26:2) indicates that the diacylglyceryl trimethylhomoserine contains two carbon chains, one of which consists of 19 carbon atoms without double bonds, and the other chain consists of 26 carbon atoms containing two double bonds. PC(14:0e/20:4), e represents an ether bond in the chain structure. HexCer/NS(d18:1/16:1), d represents 2 hydroxyl groups in the chain structure. Abbreviations: Acar: acylcarnitine; AcylGlcADG: acylglucuronosyldiacylglycerol; CE: cholesteryl ester; Cer/NDS: ceramide nonhydroxy fatty acid-dihydrosphingosine; Cer/NS: ceramide nonhydroxy fatty acid-sphingosine; DAG: diacylglycerol; DGTS: diacylglyceryl trimethylhomoserine; HexCer/NS: nonhydroxy fatty acid-sphingosine; LPC: lysophosphatidylcholine; LPE: lysophosphatidylethanolamine; PC: phosphatidylcholine; SM: sphingomyelin; SHexCer: sulfur hexosylceramide hydroxy fatty acid; TAG: triacylglycerol.

in the standard group, suggesting that the treatment option was effective. The UHPLC-MS detection results revealed that triacylglycerol, the differential lipid metabolites in the standard group compared with the normal group, the management group compared with the normal group, and the standard group compared with the normal group were all markedly upregulated. Normally, insulin inhibits the lipolysis of triacylglycerol in adipose tissue. In the presence of insulin resistance, increased lipolysis produces more free fatty acids, thereby inhibiting the antilipolytic effect of insulin. As more free fatty acids enter the liver, hepatic triacylglycerol synthesis increased, resulting in hypertriglyceridemia [23], decreased cerebral hemodynamic indicators, and increased risk of stroke. It is an independent risk factor for stroke, suggesting that elevated triacylglycerol is closely related to the occurrence of stroke [24].

Taken together, the experimental findings revealed that insulin injection could reduce blood glucose levels and

correct the insulin resistance levels in patients with AIS diabetes mellitus. However, as diabetic patients are not sensitive to insulin, strict blood sugar management regimens are more effective in lowering the glucose levels in AIS diabetic patients. Additionally, triacylglycerol is also an independent factor affecting the occurrence of stroke and diabetes. Lower blood triacylglycerol levels can reduce the antilipolysis effect of decomposed fatty acids on insulin and better exert hypoglycemic effect of insulin, allowing to resume normal glucose levels. Meanwhile, it can reduce the effect of triacylglycerol on hemodynamics and minimize the risk of stroke.

5. Conclusion

The use of strictly managed glycemic control strategy can reduce blood glucose levels, insulin resistance levels in patients with AIS diabetes mellitus. Triacylglycerol is an independent risk factor for stroke occurrence, and lowering blood triacylglycerol levels can prevent elevated blood glucose and reduce the probability of stroke. Our study provides a theoretical reference for the prevention and treatment of hyperglycemic AIS.

Data Availability

All data, models, and code generated or used during the study appear in the submitted article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This study was supported by the Chongqing Science and Health Joint Medical Research Project (No. 2020FYYX220).

References

[1] N. Townsend, M. Nichols, P. Scarborough, and M. Rayner, "Cardiovascular disease in Europe 2015: epidemiological update," *European Heart Journal*, vol. 36, no. 40, pp. 2673-2674, 2015.

[2] N. Sarwar, P. Gao, S. R. Seshasai et al., "Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies," *Lancet*, vol. 375, no. 9733, pp. 2215–2222, 2010.

[3] D. Magkou and K. Tziomalos, "Antidiabetic treatment, stroke severity and outcome," *World Journal of Diabetes*, vol. 5, no. 2, pp. 84–88, 2014.

[4] J. S. Lim, C. Kim, M. S. Oh et al., "Effects of glycemic variability and hyperglycemia in acute ischemic stroke on post-stroke cognitive impairments," *Journal of Diabetes and its Complications*, vol. 32, no. 7, pp. 682–687, 2018.

[5] M. Luitse, T. V. Seeters, A. D. Horsch et al., "Admission hyperglycaemia and cerebral perfusion deficits in acute ischaemic stroke," *Cerebrovascular Diseases*, vol. 35, no. 2, pp. 163–167, 2013.

[6] S. F. Lin, A. C. Chao, H. H. Hu et al., "Hyperglycemia predicts unfavorable outcomes in acute ischemic stroke patients treated with intravenous thrombolysis among a Chinese population: a prospective cohort study," *Journal of the Neurological Sciences*, vol. 388, pp. 195–202, 2018.

[7] L. Suissa, E. Panicucci, C. Perot et al., "Effect of hyperglycemia on stroke outcome is not homogeneous to all patients treated with mechanical thrombectomy," *Clinical Neurology and Neurosurgery*, vol. 194, article 105750, 2020.

[8] C. Brunius, L. Shi, and R. Landberg, "Large-scale untargeted LC-MS metabolomics data correction using between-batch feature alignment and cluster-based within-batch signal intensity drift correction," *Metabolomics*, vol. 12, no. 11, p. 173, 2016.

[9] K. Cho, N. G. Mahieu, S. L. Johnson, and G. J. Patti, "After the feature presentation: technologies bridging untargeted metabolomics and biology," *Curr Opin Biotech*, vol. 28, pp. 143–148, 2014.

[10] J. T. Kim, S. Y. Lee, D. S. Yoo et al., "Clinical implications of serial glucose measurements in acute ischemic stroke patients treated with intravenous thrombolysis," *Scientific Reports*, vol. 8, no. 1, p. 11761, 2018.

[11] J. Zhou, J. Wu, J. Zhang et al., "Association of stroke clinical outcomes with coexistence of hyperglycemia and biomarkers of inflammation," *Journal of Stroke and Cerebrovascular Diseases*, vol. 24, no. 6, pp. 1250–1255, 2015.

[12] S. Mahalingaiah and E. Diamanti-Kandarakis, "Targets to treat metabolic syndrome in polycystic ovary syndrome," *Expert Opinion on Therapeutic Targets*, vol. 19, no. 11, pp. 1561–1574, 2015.

[13] L. Wang, Y. Mu, J. Zhao, X. Wang, and H. Che, "IGRNet: a deep learning model for non-invasive, real-time diagnosis of prediabetes through electrocardiograms," *Sensors (Basel)*, vol. 20, no. 9, p. 2556, 2020.

[14] A. Bruno, S. R. Levine, M. R. Frankel et al., "Admission glucose level and clinical outcomes in the NINDS rt-PA Stroke Trial," *Neurology*, vol. 59, no. 5, pp. 669–674, 2002.

[15] K. Nardi, P. Milia, P. Eusebi, M. Paciaroni, V. Caso, and G. Agnelli, "Predictive value of admission blood glucose level on short-term mortality in acute cerebral ischemia," *Journal of Diabetes & Its Complications*, vol. 26, no. 2, pp. 70–76, 2012.

[16] D. Su, J. Zhou, M. S. Kelley et al., "Does telemedicine improve treatment outcomes for diabetes? A meta-analysis of results from 55 randomized controlled trials," *Diabetes Research and Clinical Practice*, vol. 116, pp. 136–148, 2016.

[17] H. J. Choe and Y. M. Cho, "Invincible β-cells in type 1 diabetes," *J Diabetes Investig*, vol. 12, no. 2, pp. 137–139, 2021.

[18] P. Luppi and P. Drain, "C-peptide antioxidant adaptive pathways in β cells and diabetes," *Journal of Internal Medicine*, vol. 281, no. 1, pp. 7–24, 2017.

[19] E. Mccracken, M. Monaghan, and S. Sreenivasan, "Pathophysiology of the metabolic syndrome," *Clinics in Dermatology*, vol. 36, no. 1, pp. 14–20, 2018.

[20] R. Haselgrübler, F. Stübl, K. Essl, M. Iken, K. Schröder, and J. Weghuber, "Gluc-HET, a complementary chick embryo model for the characterization of antidiabetic compounds," *PLoS One*, vol. 12, no. 8, article e182788, 2017.

[21] C. R. M. Ruth and J. W. Matthew, "Hepatokines: linking nonalcoholic fatty liver disease and insulin resistance," *Nature Reviews. Endocrinology*, vol. 13, no. 9, pp. 509–520, 2017.

[22] B. C. Bergman, J. T. Brozinick, A. Strauss et al., "Muscle sphingolipids during rest and exercise: a C18:0 signature for insulin resistance in humans," *Diabetologia*, vol. 59, no. 4, pp. 785–798, 2016.

[23] K. Janaka and G. Luigi, "Diabetes mellitus, a complex and heterogeneous disease, and the role of insulin resistance as a determinant of diabetic kidney disease," *Nephrology Dialysis Transplantation Official Publication of the European Dialysis & Transplant Association European Renal Association*, vol. 31, no. 2, pp. 206–213, 2016.

[24] G. Tripepi, F. Mattace-Raso, F. Rapisarda et al., "Traditional and nontraditional risk factors as predictors of cerebrovascular events in patients with end stage renal disease," *Journal of Hypertension*, vol. 28, no. 12, pp. 2468–2474, 2010.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 1507125, 12 pages
https://doi.org/10.1155/2022/1507125

Research Article

Radiomic Signature Based on Dynamic Contrast-Enhanced MRI for Evaluation of Axillary Lymph Node Metastasis in Breast Cancer

Yanqiu Tang,[1] Lin Chen,[2] Yating Qiao,[3] Weifeng Li,[4] Rong Deng,[2] and Mengdi Liang[1]

[1]*Department of Breast Surgery, The First Affiliated Hospital with Nanjing Medical University, Nanjing 210029, China*
[2]*Department of General Surgery, The Affiliated Cancer Hospital of Nanjing Medical University & Jiangsu Cancer Hospital & Jiangsu Institute of Cancer Research, Nanjing 210009, China*
[3]*Department of Gastrointestinal Surgery, Affiliated Hospital of Hebei University, Baoding, China*
[4]*School of Electronic Science and Engineering, Nanjing University, Nanjing 210046, China*

Correspondence should be addressed to Mengdi Liang; sjoy1100@126.com

Received 30 June 2022; Revised 17 July 2022; Accepted 21 July 2022; Published 17 August 2022

Academic Editor: Hongye Wang

Background. To construct and validate a radiomic-based model for estimating axillary lymph node (ALN) metastasis in patients with breast cancer by dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI). *Methods*. In this retrospective study, a radiomic-based model was established in a training cohort of 236 patients with breast cancer. Radiomic features were extracted from breast DCE-MRI scans. A method named the least absolute shrinkage and selection operator (LASSO) was applied to select radiomic features based on highly reproducible features. A radiomic signature was built by a support vector machine (SVM). Multivariate logistic regression analysis was adopted to establish a clinical characteristic-based model. The performance of models was analysed through discrimination ability and clinical benefits. *Results*. The radiomic signature comprised 6 features related to ALN metastasis and showed significant differences between the patients with ALN metastasis and without ALN metastasis ($P < 0.001$). The area under the curve (AUC) of the radiomic model was 0.990 and 0.858, respectively, in the training and validation sets. The clinical feature-based model, including MRI-reported status and palpability, performed slightly worse, with an AUC of 0.784 in the training cohort and 0.789 in the validation cohort. The radiomic signature was confirmed to provide more clinical benefits by decision curve analysis. *Conclusions*. The radiomic-based model developed in this study can successfully diagnose the status of lymph nodes in patients with breast cancer, which may reduce unnecessary invasive clinical operations.

1. Introduction

Breast cancer has severely threatened women's physical health and quality of life. It accounts for the highest incidence of malignancy and the second leading cause of tumour-related deaths in females [1]. The 5-year survival rate is up to 99% for patients with localized breast cancer, while the rate of patients with lymph node metastasis descends to 89% [2].

The identification of axillary lymph node (ALN) is crucial for therapy of patients with breast cancer, which determines whether postoperative chemical or radiation therapy is needed [3]. Sentinel lymph node dissection (SLND) is recommended to predict the status of nonpalpable axillary lymph nodes [4]. Axillary lymph node dissection (ALND) or ultrasound-guided fine-needle aspiration (US-FNA) is clinically operated in patients with palpable axillary lymph nodes [5]. However, both ALND and SLND are invasive, potentially resulting in impaired shoulder range of motion, limber edema, numbness, and pain syndromes [6, 7]. Therefore, it would be beneficial if there is a reliable noninvasive evaluation of ALN metastasis.

Radiomics has attracted great attention as a potential way to preoperatively detect breast cancer and assess ALN

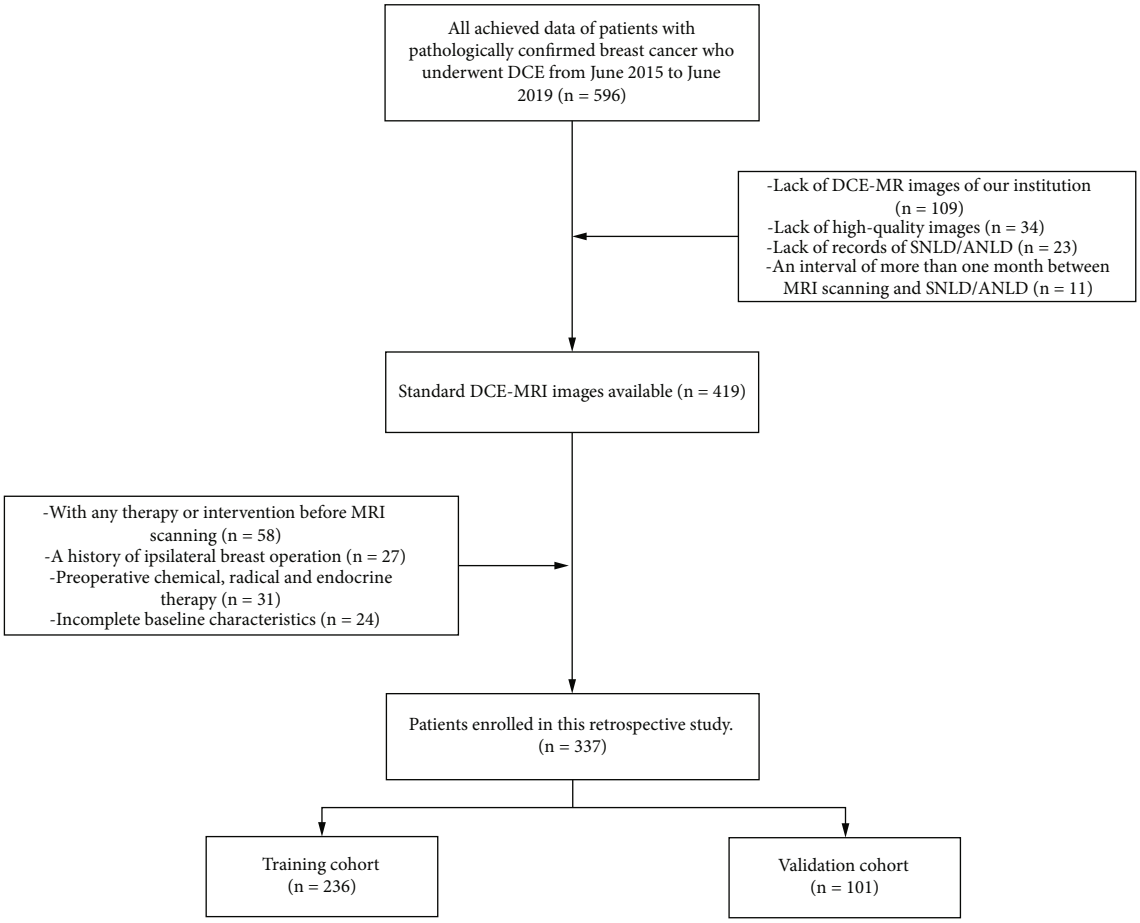


Figure 1: Patient selection flow chart. In all, 337 breast cancer patients with high-quality DCE-MRI scans were included in this study.

status. Quantitative imaging features extracted from magnetic resonance imaging (MRI), combined with other clinical information, can be used to support clinical decisions [8]. In the study by Chai et al. [9], dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) showed the best performance to predict ALN metastasis among multiparametric MR sequences.

There have been several studies evaluating SLN status by radiomic analysis of DCE-MRI combined with clinicopathologic features. However, the use of ALND shows no superior survival compared with SLND alone for patients with T1 or T2 breast cancer if there are only one or two positive SLNs [10, 11]. Thus, patients with positive SLNs may undergo unnecessary ALND due to the lack of evidence for ALN metastasis [12]. The purpose of this study is to predict ALN metastasis by developing and validating a radiomics-based model.

2. Methods

2.1. Patients. 573 patients with histologically confirmed breast cancer who received DCE-MRI and SLND/ALND from June 2015 to June 2019 were retrospectively reviewed. The exclusion criteria were (1) lack of DCE-MR images at The Affiliated Cancer Hospital of Nanjing Medical University ($n = 109$). (2) lack of high-quality images ($n = 34$), (3) an interval of more than one month between MRI scanning and SNLD/ANLD ($n = 11$), (4) with any therapy or intervention before MRI scanning (e.g., a history of ipsilateral breast operation and chemical, radical, and endocrine therapy) ($n = 58$), and (5) incomplete baseline characteristics ($n = 24$). A total of 337 patients were enrolled in this study (Figure 1). Based on the date of surgery, 236 patients treated from June 2015 to December 2017 were allocated to the training cohort and 101 patients from January 2018 to June 2019 were for the validation cohort.

Clinical characteristics and the data of baseline MR imaging were collected from medical records. The clinical data included age, immunophenotype (according to progesterone receptor (PR) status, estrogen receptor (ER) status, human epidermal growth factor receptor 2 (HER-2) status, and Ki67 proliferation index), histological type (invasive ductal carcinoma, invasive lobular carcinoma, and other types), and ALN palpability. The serum tumour markers (cancer antigen 15-3 (CA15-3), cancer antigen 125 (CA125), and carcinoembryonic antigen (CEA)) were considered due to the diagnostic value in metastatic breast cancer [13].

2.2. SLND/ALND and Pathological Assessment. SLND was performed for all participants by the dye method within one week after MRI examination. Negative indications of

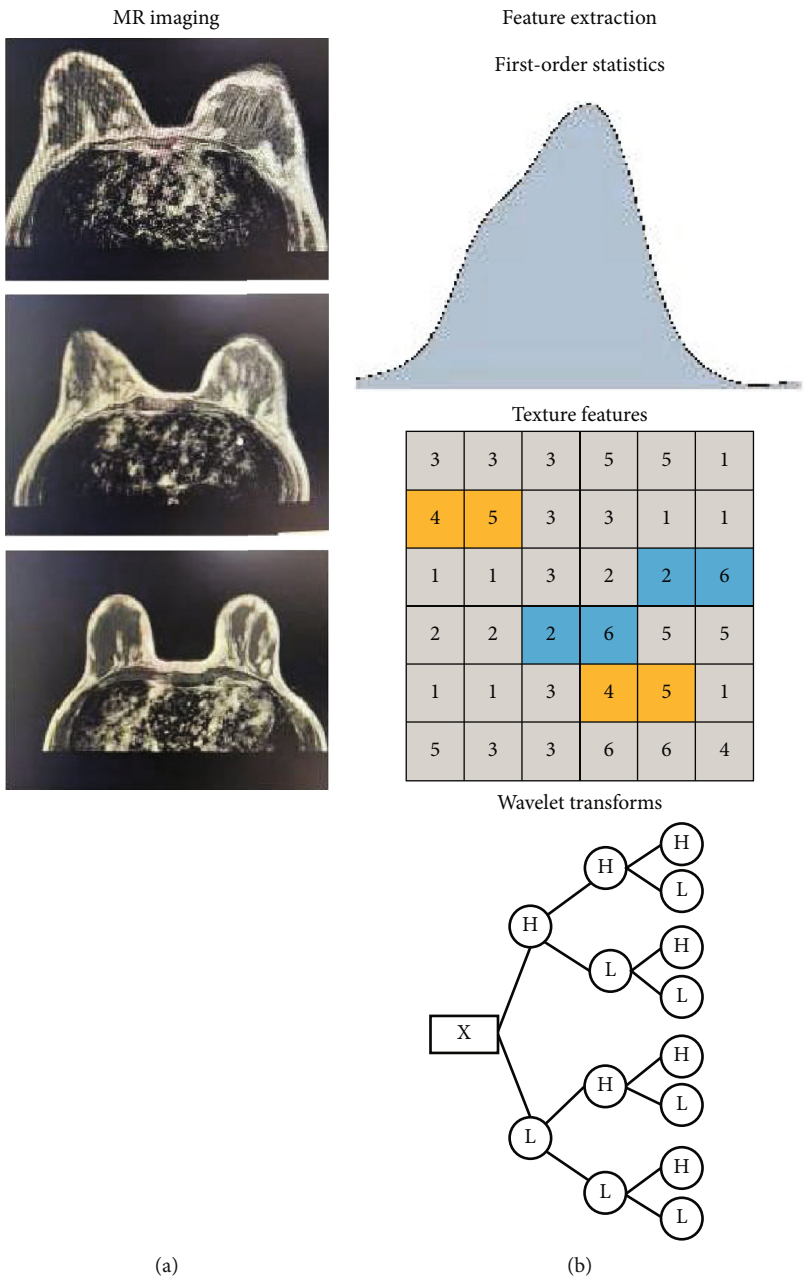


Figure 2: Continued.

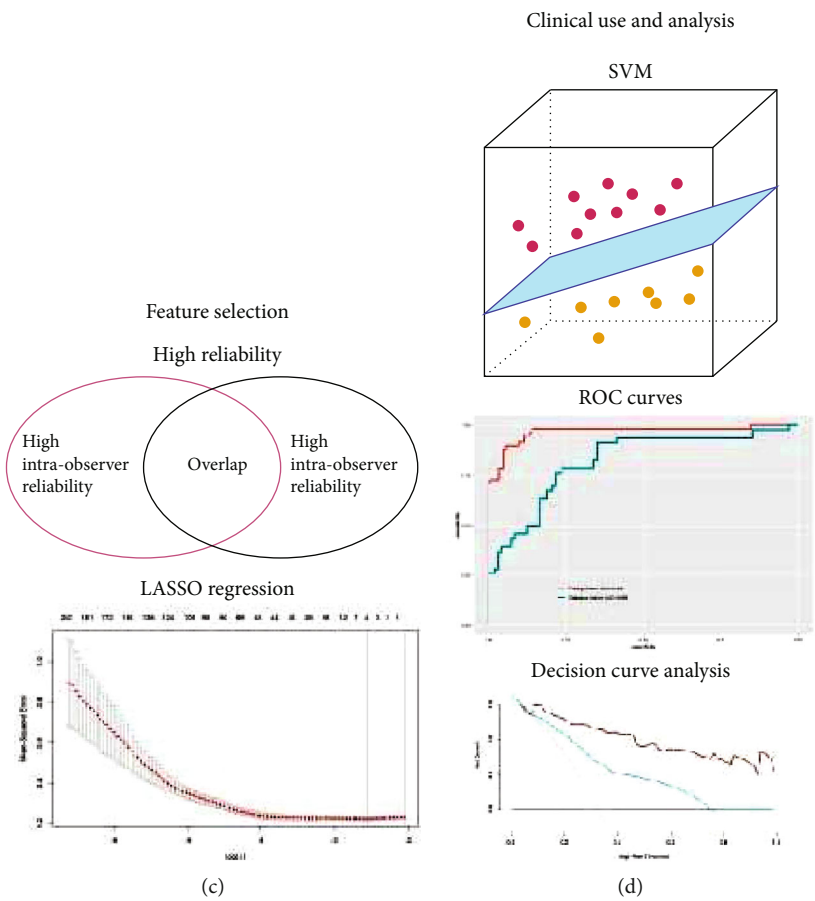


Figure 2: The workflow of necessary steps in this study. (a) ROI was manually delineated on the DCE-MR images. (b) Radiomic features were extracted, including the first-order statistics, textural features, and wavelet transforms. (c) Radiomic features highly related to LN metastasis were selected to construct a radiomic signature. (d) The radiomic model was constructed by SVM, and the performance of models was evaluated by ROC and DCA curves.

ALND were defined as isolated tumour cells and micrometastasis in the SN (isolated tumour cells: <0.2 mm or <200 tumour cells; micrometastasis: tumour diameter > 0.2 mm, ≤2 mm, or <200 tumour cells). Macrometastasis (tumour diameter > 2 mm) was considered a positive indication of axillary clearance [14].

Histological types of breast cancer were analysed by two pathologists with over 5-year experience in consensus. Each pathologist was blinded to the clinical situation. Immunophenotypes were based on expression levels of ER, PR, HER2, and Ki67: luminal A (ER+ and/or PR+, HER2-, and Ki67-low), luminal B (ER+ and/or PR+, HER2+; ER+ and/or PR+, HER2-, and Ki67-high), HER2 positive (ER-, PR-, and HER2+), and triple-negative (ER-, PR-, and HER2-) [15].

2.3. MR Image Acquisition and Radiologic Evaluation. DCE-MRI scanning was performed with a 3 T MRI system (SIEMENS) and 8-channel breast coils in a prone position. The contrast enhancement agent, gadopentetate dimeglumine penta-acetic acid (Gd-DPTA), was intravenously injected at a dose of 0.1 mmol/kg and a rate of 3.0 ml/s. A total of five phases, one precontrast and four postcontrast phases, were obtained with a sagittal VIBRANT multiphase sequence: repetition time (TR) = 4.46-7.80 msec, echo time (TE) = 1.54 -4.20 msec, flip angle = 10°, field of view (FOV) = 36 ∗ 36 c m^2, matrix = 256 ∗ 256, and slice thickness without a gap = 2 mm.

The workflow is depicted in Figure 2. Two radiologists, one with 8 years of breast cancer MRI experience and the other with 10 years, evaluated the following traits of all images with blindness to the clinical and pathological details: (1) tumour size: defined by the maximum transverse diameter of the largest lesion; (2) multifocality: defined by more than one lesion; and (3) LN status: defined by shape, fatty hilum, cortical thickness, and thickening pattern [16]. Consensus would be reached through discussion if the two radiologists disagreed with the LN status.

2.4. Region-of-Interest Segmentation and Radiomic Feature Extraction. The signal intensity of tumour lesions in different phases was calculated with GE Advanced Workstation ADW4.4. In the images with the strongest enhanced phase, region of interests were manually segmented along the 4 mm dilated tumour contour by a 3D slicer (version 4.4.0), which contains both intra- and peritumoural data.

Table 1: Main characteristics of patients in the training and validation cohorts.

Characteristics	Training ($n = 236$)	Validation ($n = 101$)	P
Age*	53.21 ± 9.72	52.7 ± 10.13	0.67
Immunophenotype			0.415
Luminal A	76 (32.2)	38 (37.6)	
Luminal B	126 (53.4)	46 (45.5)	
HER2	11 (4.7)	8 (7.9)	
Triple-negative	23 (9.7)	9 (8.9)	
Histological type			0.547
Invasive carcinoma	189 (80.1)	87 (86.1)	
Precursor lesions	43 (18.2)	14 (13.9)	
Tumour markers*			
CEA	0.83 ± 0.31	0.76 ± 0.42	0.09
CA15-3	12.32 ± 5.41	11.71 ± 5.53	0.34
CA125	20.61 ± 8.57	21.13 ± 7.49	0.59
Tumour size	18.44 ± 8.34	19.52 ± 9.12	0.29
Multifocality			0.56
Yes	21 (8.9)	11 (10.9)	
No	215 (91.1)	90 (89.1)	
LN palpability			0.88
Yes	29 (12.3)	13 (12.9)	
No	207 (87.1)	88 (87.1)	
MRI-reported LN status			0.48
Positive	50 (21.2)	18 (17.8)	
Negative	186 (78.8)	83 (82.2)	
LN metastasis			0.68
Yes	78 (33.1)	36 (35.6)	
No	158 (66.9)	65 (64.4)	

Note. Data are numbers of patients, with percentages in parentheses. CEA: carcinoembryonic antigen; CA15-3: cancer antigen 15-3; CA125: cancer antigen 125. *Data are presented as means ± standard deviations.

Table 2: Main characteristics of patients with and without LN metastasis.

Characteristics	LN metastasis ($n = 114$)	Non-LN metastasis ($n = 223$)	P
Age*	52.13 ± 9.18	52.75 ± 8.93	0.54
Immunophenotype			0.29
Luminal A	41 (40.0)	73 (32.7)	
Luminal B	51 (44.7)	121 (54.2)	
HER2	9 (7.9)	10 (4.5)	
Triple-negative	13 (11.4)	19 (8.5)	
Histological type			<0.001
Invasive carcinoma	111 (97.4)	169 (75.8)	
Precursor lesions	3 (2.6)	54 (24.2)	
Tumour markers*			
CEA	0.77 ± 0.45	0.82 ± 0.31	0.26
CA15-3	11.69 ± 5.61	12.37 ± 5.32	0.27
CA125	20.53 ± 7.86	20.88 ± 8.41	0.71
Tumour size	19.82 ± 7.58	18.22 ± 9.42	0.12
Multifocality			<0.001
Yes	22 (19.3)	10 (4.5)	
No	92 (80.7)	213 (95.5)	
LN palpability			<0.001
Yes	34 (29.8)	8 (3.6)	
No	80 (80.2)	215 (96.4)	
MRI-reported LN status			<0.001
Positive	46 (40.4)	22 (9.9)	
Negative	68 (59.4)	201 (90.1)	
LN metastasis			
Yes	—	—	
No	—	—	

Note. Data are numbers of patients, with percentages in parentheses. CEA: carcinoembryonic antigen; CA15-3: cancer antigen 15-3; CA125: cancer antigen 125. *Data are presented as means ± standard deviations.

Firstly, to assess inter- and intraobserver reliability, the ROIs were performed by two experienced radiologists (one with 10-year experience and the other with 16-year experience). Reader 1 repeated the segmentation twice a week, and reader 2 independently extracted ROIs to, respectively, calculate intra- and interobserver reproductivity with intraclass correlation coefficient [17] (ICC). Radiomic features with both intra- and interobserver ICC greater than 0.8 were subsequently analysed in this study. Secondly, the least absolute shrinkage and selection operator [18] (LASSO), with penalty parameter tuning conducted by 10-fold cross-validation, was applied to select features of ALN status with nonzero coefficients in the training cohort.

2.5. Establishment, Performance, and Validation of a Radiomic Model. The support vector machine (SVM) is a kind of supervised model for regression analysis with robust prediction ability. Based on selected radiomic features, SVM was applied to generate a radiomic signature, using "e1071" package (https://CRAN.R-project.org/package=e1071) on R software (version 3.6.1, http://www.r-project.org).

To assess the association between ALN metastasis and clinical features, the features with significant differences ($P < 0.05$) between the training and validation cohorts were selected for further analysis. Next, multivariable logistic regression was applied to build the clinical model in the training cohort. The cutoff value of each independent risk factor was evaluated by receiver operating characteristic (ROC) analysis with the maximum Youden index.

This study established three models to predict ALN metastasis using logistic regression, including the radiomic signature alone, clinical factors alone, and the model combining the radiomic signature and clinical risk factors. The discrimination performance of each model was determined by ROC analysis and area under the curve (AUC). The Delong test was used to compare each model with the

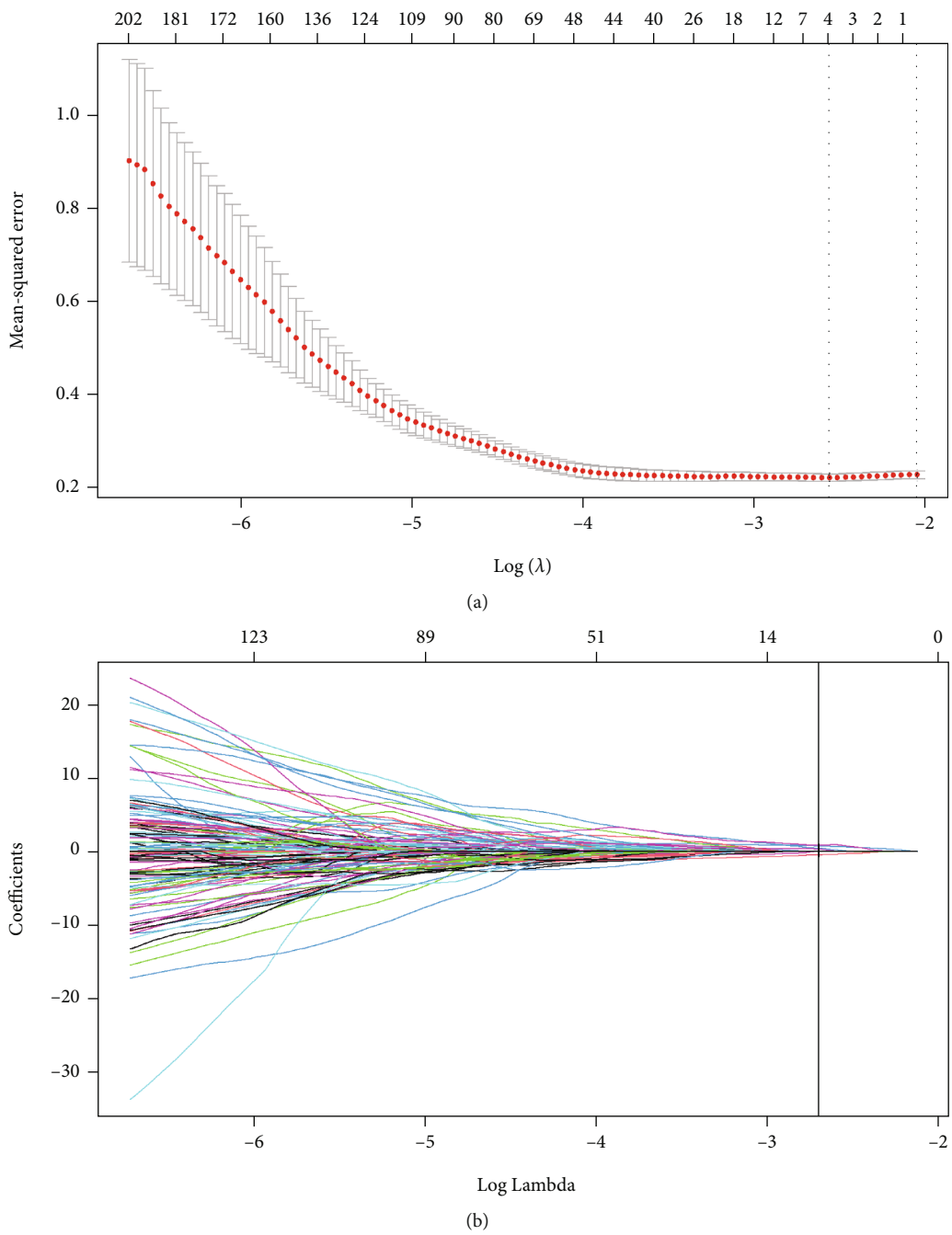


Figure 3: Radiomic feature selection by the least absolute shrinkage and selection operator (LASSO) regression. (a) LASSO coefficient profiles of the 841 selected features. (b) Optimal λ value was determined by the LASSO model using 10-fold cross-validation via minimum criteria. The mean-squared error was plotted versus log(λ). Dotted vertical lines were drawn at the optimal values by using the minimum criteria and the 1 standard error of the minimum criteria. The optimal λ value of 0.067 was chosen.

Table 3: LN status-related radiomic features.

Types	Features
Shape ($n = 2$)	Maximum 3D diameter, minor axis
Texture features ($n = 1$)	Original GLDM gray level variance
Wavelet transforms ($n = 2$)	LHL GLSZM gray level nonuniformity, HHL glcm difference entropy

Note. GLDM: gray level difference matrix; GLSZM: gray level size zone matrix; H: high; L: low.

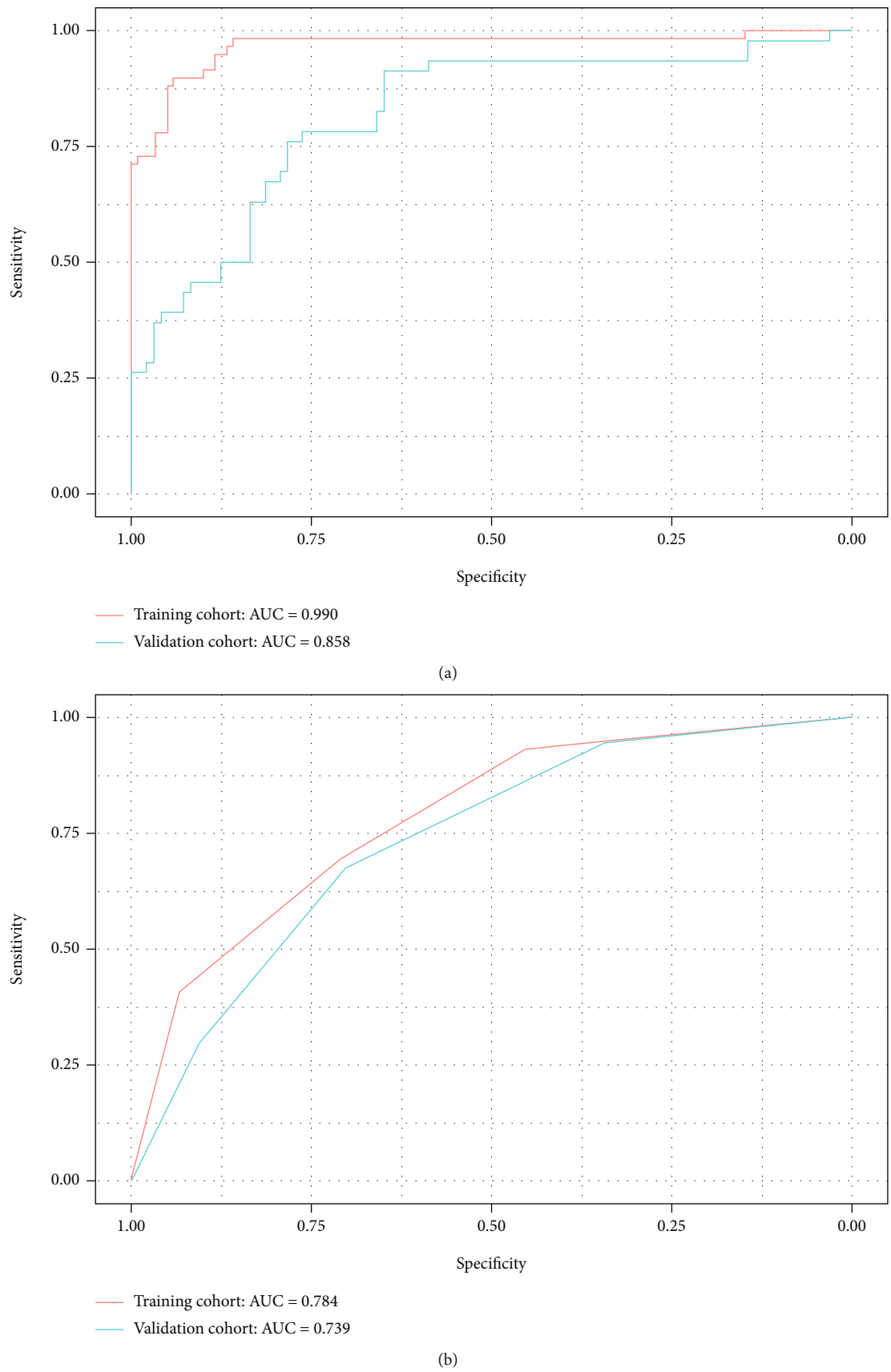


Figure 4: ROC curves using radiomic features (a) and clinical features (b) for the training and validation sets.

Table 4: Clinical risk factors for axillary lymph node metastasis.

	b coefficient	Odds ratio	*P* value
Histological type	0.26	1.3 (0.23-11.45)	0.7864
Multifocality	0.12	1.13 (0.33-3.67)	0.8444
MRI-reported status	1.66	5.28 (2.52-11.79)	<0.001
LN palpability	2	7.35 (3.48-16.57)	<0.001

Note. *b* coefficient was from multivariable logistic regression. Clinical factors found to be significantly related to the LN metastasis entered into the clinical model.

AUC value. The performance of models was then tested in the independent validation cohort with the formula from the training cohort.

2.6. Statistical Analysis. For categorical variables, the chi-squared test or Fisher exact test was performed to analyse the equality of variances between the training and validation cohorts, and the Student *t*-test or Mann–Whitney *U* test was used to compare continuous variables. Decision curve analysis (DCA) was applied in the validation cohort to assess the benefit of each model at different threshold probabilities. A two-sided *P* value less than 0.05 was considered statistical significance.

3. Results

3.1. Patient Characteristics. As summarized in Table 1, there were no statistical differences in clinical and radiological characteristics between the training and validation cohorts. The rates of LN metastasis were, respectively, 33.1% (78 of 236) and 35.6% (36 of 101) in the training and validation cohorts, whereas no difference was found between the two cohorts ($\chi^2, P = 0.64$). The overall discrimination accuracy of MRI report of LN status was 63%, with a sensitivity of 48.7% (38 of 78), a specificity of 66.4% (105 of 158), a positive value of 76.0% (38 of 50), and a negative predictive value of 56.4% (105 of 186). Statistical differences were found between non-LN metastasis and LN metastasis in multifocality, LN palpability, and MRI-reported LN status (Table 2).

3.2. Radiomic Signature. In total, 841 features (13 shape features, 18 first-order features, 74 textural features, and 736 wavelet-based features) were automatically extracted from each ROI with the opensource Pyradiomics package (http://www.radiomics.io/pyradiomics.html). Details about the extracted features are shown in the supplement material (Table S1).

Of 841 extracted features, 332 LN-related features were selected for the following analysis, including 11 shape features, 9 first-order features, 37 textural features, and 275 wavelet-based features. Five LN status-related features with nonzero coefficients in the LASSO regression model were selected based on the training cohort, including two shape features, one textural feature, and two wavelet-transformed features (Figure 3). The five features are shown in Table 3.

The SVM algorithm was applied to construct a radiomic signature. A difference in the decision values was observed between patients with and those without LN metastasis in the training cohort (mean, -0.581 vs. 0.737, $P < 0.001$) and also obtained in the validation cohort (mean, 0.129 vs. 0.470, $P < 0.001$). As is shown in Figure 4, the radiomic signature displayed a favourable discriminatory ability with an AUC of 0.990 (95% confidence interval (CI): 0.990-1) in the training cohort and 0.858 (95% CI: 0.834-0.950) in the validation cohort. The optimal cutoff value of 0.8567 for the radiomic signature was calculated at the point of the maximum Youden index from the entire cohort. The radiomic model performs well in the training cohort, whose sensitivity and specificity were 100% and 94%, respectively. In the validation cohort, the sensitivity was as high as 87% and the specificity was 78%. The accuracies were 96% and 81% in the training and validation cohorts, respectively. The calibration curve of the radiomic signature yielded great agreement between the predicted and actual metastases in the training cohort.

3.3. LN Status-Related Clinical Factors. As shown in Table 1, histological type, multifocality, MRI-reported LN status, and LN palpability were significantly related to LN metastasis ($P < 0.001$, chi-squared test). Table 4 displays the odds ratios of the above clinical factors. The odds ratios of MRI-reported status and LN palpability were statistically significant, respectively, 5.28 (95% CI: 2.52-11.79) and 7.35 (95% CI: 3.48-16.57). Then, the clinical prediction model was built by multivariable logistic regression based on MRI-reported status and LN palpability. The model displayed an AUC of 0.784 (95% CI: 0.716-0.851) in the training cohort, and the sensitivity, specificity, and accuracy were, respectively, 69%, 71%, and 70%. The performance in the validation cohort was similar, with an AUC of 0.739 (95% CI: 0.644-0.833), a sensitivity of 68%, a specificity of 70%, and an accuracy of 69%. The optimal cutoff value of the clinical model was -0.671, determined from the whole cohort. Compared with the radiomic signature, the clinical model yielded poorer results in the training and validation sets.

3.4. Combining Radiomic Signature and Clinical Factors. The discriminatory ability of the combined model was poorer than that of the radiomic signature, with an AUC of 0.987 (95% CI: 0.9743-1; $P = 0.877$) in the training cohort and of 0.826 (95% CI: 0.742-0.909; $P = 0.11$). The sensitivity, specificity, and accuracy of the third model were, respectively, 96%, 93%, and 95% in the training cohort and 77%, 81%, and 78% in the validation cohort.

3.5. Clinical Use. As summarized in Table 5, the radiomic signature had the best discriminatory ability in the training and validation cohorts. In the training cohort, the AUC value of the radiomic signature was significantly higher than that of the clinical model (AUC: 0.784; CI: 0.716-0.851), MRI-reported metastasis alone (AUC: 0.661; CI: 0.586-0.735; $P < 0.001$), and palpability (AUC: 0.703; CI: 0.631-0.775; $P < 0.001$). In the validation cohort, the radiomic signature displayed the best results, compared with the clinical model (AUC: 0.739; CI: 0.644-0.833; $P = 0.020$), the MRI-reported metastasis (AUC: 0.557; CI: 0.456-0.659; $P <$

Table 5: Performances of all methods for predicting LN metastasis.

	Training	Validation	Training vs. validation
	AUROC	AUROC	Delong test
Radiomic signature	0.990 (0.990, 1)	0.858 (0.834, 0.950)	0.004
Clinical model	0.784 (0.716, 0.851)	0.739 (0.644, 0.833)	0.4439
Combined model	0.987 (0.9743-1)	0.826 (0.742-0.909)	<0.001
MRI-reported status	0.661 (0.586-0.735)	0.557 (0.456-0.659)	0.1083
LN palpability	0.703 (0.631-0.775)	0.689 (0.594-0.784)	0.8253
Comparison of AUCROC			
Radiomic signature vs. clinical model	<0.001	0.02	
Radiomic signature vs. MRI-reported status	<0.001	<0.001	
Radiomic signature vs. LN palpability	<0.001	0.002	
Radiomic signature vs. combined model	0.877	0.11	
Clinical model vs. MRI-reported status	0.005	0.006	
Clinical model vs. LN palpability	<0.001	0.028	
MRI-reported status vs. palpability	0.462	0.103	

Note. 95% confidence intervals were shown in parentheses.

0.001), and palpability (AUC: 0.689; CI: 0.594-0.784; $P = 0.002$). Though the AUC of the radiomic signature was slightly higher than the combined model, the difference showed no statistical significance. Figure 5 presents the decision curve analysis for the clinical prediction model and the radiomic signature. The radiomic signature indicates more benefit to predict LN metastasis, with the threshold probabilities of more than 10%.

4. Discussion

This study constructed and validated a radiomics-based model and a clinical model to predict LN metastasis in patients with breast cancer. Six stable radiomic features effectively identify patients as LN metastasis or non-LN metastasis. Compared with the clinical model consisting of the MRI-reported LN status and LN palpability, the radiomic model performed much better with an AUC of 0.858 in the validation cohort.

Radiomics is termed as extracting quantitative features that convert images into mineable data, and analyse these data to improve diagnosis, prediction power, and much other decision support. MRI-based radiomic analysis can provide an efficient method to estimate the existence of ALN metastasis and probably change the clinical routine in the future. In this study, we constructed a radiomic model based on images extracted from DCE-MRI to find the LN metastasis, and the results of the model were satisfactory. The AUC, sensitivity, and specificity were 0.858, 87%, and 78%, respectively, in the validation cohort. While MRI has been the main noninvasive method to assess LNs, its sensitivity and specificity were merely 68% and 70%.

Like other research, this study found that the clinical characteristics of patients with breast cancer were related to LN status, such as multifocality, palpability of LNs, and MRI-reported LN status. Tan et al. selected four factors, including age, HER2 status, size of tumour, and vascular thrombus accompanied or not, into the clinical model to predict SLN metastasis [19]. The accuracy of the model relying on clinicopathological features was merely 70.26%, much lower than the radiomic signatures, and the damage must happen if histological information is needed for the model.

A radiomic signature, based on ROI extracted from DCE-MR images, was constructed in this study to evaluate the ALN metastasis. Dong et al. predicted SLN metastasis based on radiomics of diffusion-weighted (DWI) and T2-weighted fat-suppression MRI (T2-FS). The AUC of radiomic model-based DWI and T2-FS were, respectively, 0.77 and 0.79 in the validation set, slightly lower than ours. To date, there have been some articles published combining the radiomic signature and clinicopathological features. For instance, Han et al. evaluated the ALN with the radiomic signature and clinical characteristics including palpability of LN and MRI-reported LN status, achieving an AUC of 0.78 [20]. Similar clinical features associated with LN metastasis were found to predict the LN status in our study, but the radiomic-based model combining the clinical features was slightly worse than the model using the radiomic signature alone. One of the reasons was that the accuracy of the clinical information is much dependent on the experience of doctors.

There were several limitations in this study. First, the retrospective analysis had inherent bias influencing the outcomes. More independent samples from different centres will be needed to validate the results. Second, the images to extract radiomic features are the primary tumour instead of LNs. The MR images with LNs, however, are only a small part of all samples. Third, the ROI is circled manually. Although the intra- and interobserver ICC were more than 0.8, some studies demonstrated that automated or semiautomated methods show higher accuracy and stability [21, 22]. Fourth, many scholars have applied radiogenomics to cancer research [23–25]. This study ignored the genomic data due to economic limitations.

In conclusion, the radiomic model is a promising noninvasive method to predict LN metastasis for breast cancer.

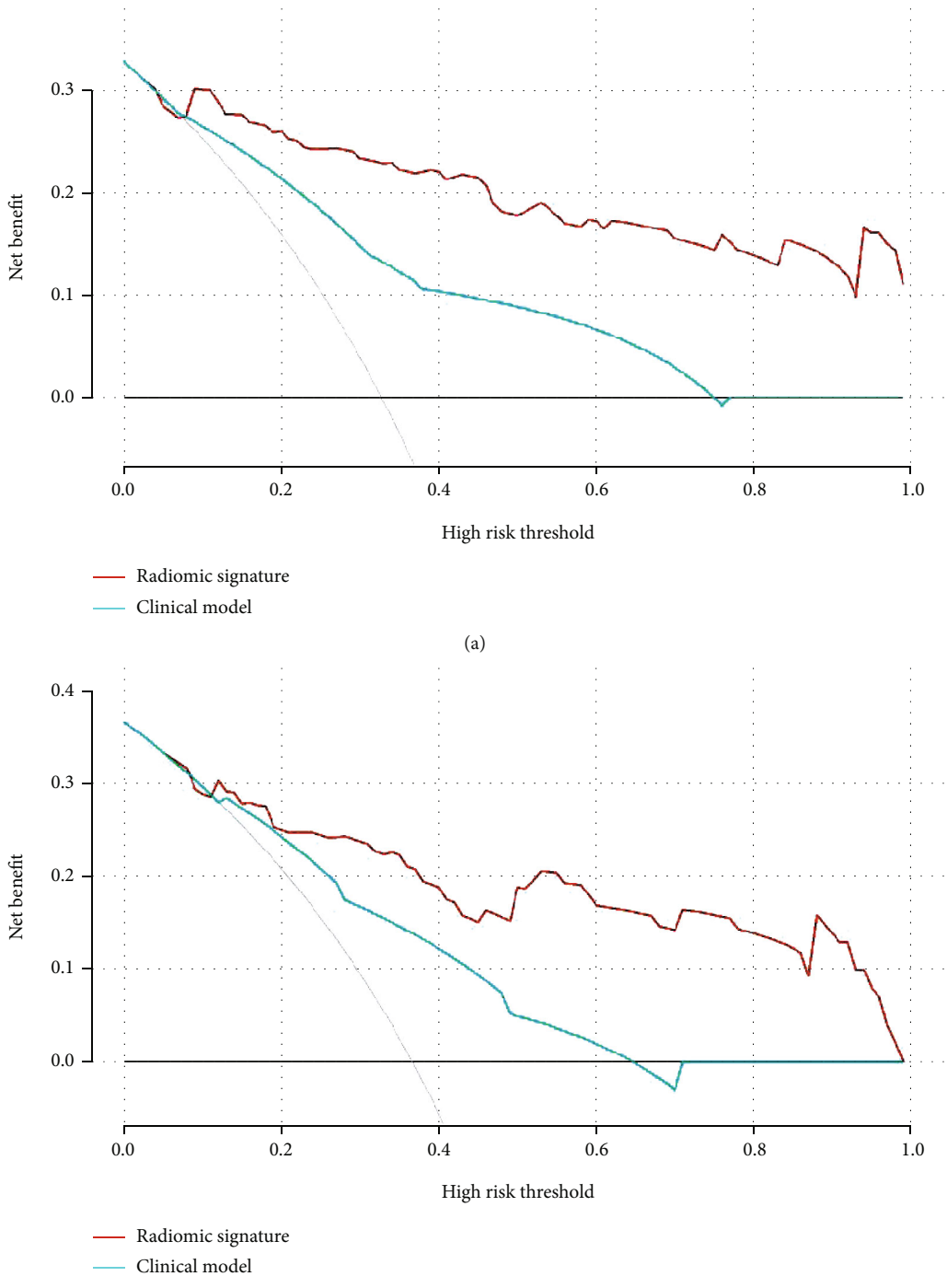


Figure 5: Decision curve analysis for each model in the training (a) and validation (b) cohorts. The Y-axis measures the net benefit, calculated by the true-positive findings and false-positive findings. Across the threshold probability, the application of radiomic signature to predict LN status provides more benefits than clinical features.

Further study with a larger sample size is needed to achieve the application.

4.1. Clinical Practice Points. Breast cancer patients with positive SLNs are advised to undergo ALND to confirm the ALN status which is important for treatment strategy, while some results of ALND are negative. We hope to find a convenient and noninvasive method for patients with breast cancer to assess ALN status. MRI has been extensively applied to the diagnosis of breast cancer, and it is easy for

patients to obtain MR images. The MRI-based radiomic model performs well in evaluating the ALN metastasis and may reduce unnecessary lesions.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Yanqiu Tang, Lin Chen, and Yating Qiao share co-first authorship.

Acknowledgments

We would like to express our great appreciation to the R Development Core Team and contributors for R packages used in this study. This study was supported by the Foundation of State Key Laboratory of Ultrasound in Medicine and Engineering (Grant No. 2021KFKT002 and Grant No. 2021KFKT014) and the First Affiliated Hospital of Nanjing Medical University.

Supplementary Materials

The supplementary material contains a table showing all 841 features extracted from ROI. Table S1: list of all extracted features. *(Supplementary Materials)*

References

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.

[2] K. D. Miller, R. L. Siegel, C. C. Lin et al., "Cancer treatment and survivorship statistics, 2016," *CA: a Cancer Journal for Clinicians*, vol. 66, no. 4, pp. 271–289, 2016.

[3] I. Jatoi, S. G. Hilsenbeck, G. M. Clark, and C. K. Osborne, "Significance of axillary lymph node metastasis in primary breast cancer," *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, vol. 17, no. 8, pp. 2334–2340, 1999.

[4] C. Coutant, C. Olivier, E. Lambaudie et al., "Comparison of models to predict nonsentinel lymph node status in breast cancer patients with metastatic sentinel lymph nodes: a prospective multicenter study," *Journal of Clinical Oncology: Official Journal of the American Society of Clinical Oncology*, vol. 27, no. 17, pp. 2800–2808, 2009.

[5] S. Usmani, N. Ahmed, N. Al Saleh et al., "The clinical utility of combining pre-operative axillary ultrasonography and fine needle aspiration cytology with radionuclide guided sentinel lymph node biopsy in breast cancer patients with palpable axillary lymph nodes," *European Journal of Radiology*, vol. 84, no. 12, pp. 2515–2520, 2015.

[6] I. Langer, U. Guller, G. Berclaz et al., "Morbidity of sentinel lymph node biopsy (SLN) alone versus SLN and completion axillary lymph node dissection after breast cancer surgery: a prospective Swiss multicenter study on 659 patients," *Annals of Surgery*, vol. 245, no. 3, pp. 452–461, 2007.

[7] J. J. Kootstra, P. U. Dijkstra, H. Rietman et al., "A longitudinal study of shoulder and arm morbidity in breast cancer survivors 7 years after sentinel lymph node biopsy or axillary lymph node dissection," *Breast Cancer Research and Treatment*, vol. 139, no. 1, pp. 125–134, 2013.

[8] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, 2016.

[9] R. Chai, H. Ma, M. Xu et al., "Differentiating axillary lymph node metastasis in invasive breast cancer patients: a comparison of radiomic signatures from multiparametric breast MR sequences," *Journal of Magnetic Resonance Imaging: JMRI*, vol. 50, no. 4, pp. 1125–1132, 2019.

[10] A. E. Giuliano, K. V. Ballman, L. McCall et al., "Effect of axillary dissection vs no axillary dissection on 10-year overall survival among women with invasive breast cancer and sentinel node metastasis: the ACOSOG Z0011 (Alliance) randomized clinical trial," *Journal of the American Medical Association*, vol. 318, no. 10, pp. 918–926, 2017.

[11] A. E. Giuliano, K. K. Hunt, K. V. Ballman et al., "Axillary dissection vs no axillary dissection in women with invasive breast cancer and sentinel node metastasis: a randomized clinical trial," *Journal of the American Medical Association*, vol. 305, no. 6, pp. 569–575, 2011.

[12] K. U. Chu, R. R. Turner, N. M. Hansen, M. B. Brennan, and A. E. Giuliano, "Sentinel node metastasis in patients with breast carcinoma accurately predicts immunohistochemically detectable nonsentinel node metastasis," *Annals of Surgical Oncology*, vol. 6, no. 8, pp. 756–761, 1999.

[13] W. Wang, X. Xu, B. Tian et al., "The diagnostic value of serum tumour markers CEA, CA19-9, CA125, CA15-3, and TPS in metastatic breast cancer," *Clinica Chimica Acta*, vol. 470, pp. 51–55, 2017.

[14] K. Takada, S. Kashiwagi, W. Goto et al., "Possibility of avoiding axillary lymph node dissection by immune microenvironment monitoring in preoperative chemotherapy for breast cancer," *Journal of Translational Medicine*, vol. 16, no. 1, p. 318, 2018.

[15] A. Goldhirsch, W. C. Wood, A. S. Coates et al., "Strategies for subtypes–dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011," *Annals of Oncology*, vol. 22, no. 8, pp. 1736–1747, 2011.

[16] E. J. Kim, S. H. Kim, B. J. Kang, B. G. Choi, B. J. Song, and J. J. Choi, "Diagnostic value of breast MRI for predicting metastatic axillary lymph nodes in breast cancer patients: diffusion-weighted MRI and conventional MRI," *Magnetic Resonance Imaging*, vol. 32, no. 10, pp. 1230–1236, 2014.

[17] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," *Journal of Chiropractic Medicine*, vol. 15, no. 2, pp. 155–163, 2016.

[18] H. Wang, G. Li, and C.-L. Tsai, "Regression coefficient and autoregressive order shrinkage and selection via the lasso," *Journal of the Royal Statistical Society: Series B*, vol. 69, no. 1, pp. 63–78, 2007.

[19] H. Tan, F. Gan, Y. Wu et al., "Preoperative prediction of axillary lymph node metastasis in breast carcinoma using

radiomics features based on the fat-suppressed T2 sequence," *Academic Radiology*, vol. 27, no. 9, pp. 1217–1225, 2020.

[20] L. Han, Y. Zhu, Z. Liu et al., "Radiomic nomogram for prediction of axillary lymph node metastasis in breast cancer," *European Radiology*, vol. 29, no. 7, pp. 3820–3829, 2019.

[21] N. S. M. Haniff, M. K. Abdul Karim, N. H. Osman, M. I. Saripan, I. N. Che Isa, and M. J. Ibahim, "Stability and reproducibility of radiomic features based various segmentation technique on MR images of hepatocellular carcinoma (HCC)," *Diagnostics*, vol. 11, no. 9, p. 1573, 2021.

[22] L. Wang, J. Tan, Y. Ge et al., "Assessment of liver metastases radiomic feature reproducibility with deep-learning-based semi-automatic segmentation software," *Acta Radiologica*, vol. 62, no. 3, pp. 291–301, 2021.

[23] K. Pinker, J. Chin, A. N. Melsaether, E. A. Morris, and L. Moy, "Precision medicine and radiogenomics in breast cancer: new approaches toward diagnosis and treatment," *Radiology*, vol. 287, no. 3, pp. 732–747, 2018.

[24] Z. Bodalal, S. Trebeschi, T. D. L. Nguyen-Kim, W. Schats, and R. Beets-Tan, "Radiogenomics: bridging imaging and genomics," *Abdominal Radiology*, vol. 44, no. 6, pp. 1960–1984, 2019.

[25] G. Singh, S. Manjila, N. Sakla et al., "Radiomics and radiogenomics in gliomas: a contemporary update," *British Journal of Cancer*, vol. 125, no. 5, pp. 641–657, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 8954606, 8 pages
https://doi.org/10.1155/2022/8954606

Research Article

Causal Association of Thyroid Signaling with C-Reactive Protein: A Bidirectional Mendelian Randomization

Tingting Li,[1] Haigang Geng,[2] Yuquan Wang,[1] Zhaorong Wu,[3] Siqian Yang,[1] and Yue-Qing Hu[1,4]

[1]*State Key Laboratory of Genetic Engineering, Human Phenome Institute, Institute of Biostatistics, School of Life Sciences, Fudan University, Shanghai, China*
[2]*Department of Gastrointestinal Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China*
[3]*Department of Interventional Oncology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China*
[4]*Shanghai Center for Mathematical Sciences, Fudan University, Shanghai, China*

Correspondence should be addressed to Yue-Qing Hu; yuehu@fudan.edu.cn

Received 8 July 2022; Revised 23 July 2022; Accepted 27 July 2022; Published 13 August 2022

Academic Editor: Jincheng Wang

Context. Existing literature demonstrated that thyroid-stimulating hormone (TSH) and free thyroid hormone (fT4) were associated with the C-reactive protein (CRP), an inflammatory risk factor of cardiovascular diseases (CVD), but the causal relationship between them remained unclear. *Methods*. Based on the latest genome-wide association study summary data, bidirectional two-sample Mendelian randomization (MR) was employed to detect the causal relationship and effect direction between TSH, fT4, and CRP. Furthermore, in view of obesity being an important risk factor of CVD, obesity trait waist-hip ratio (WHR) and body mass index (BMI) were treated as the research objects in MR analyses for exploring the causal effects of TSH and fT4 on them, respectively. *Results*. Genetically increased CRP was associated with increased TSH ($\beta = -0.02$, $P = 0.011$) and with increased fT4 ($\beta = 0.043$, $P = 0.001$), respectively, but there was no evidence that TSH or fT4 could affect CRP. In further analyses, genetically increased TSH was associated with decreased WHR ($\beta = -0.02$, $P = 3.99e-4$). Genetically increased WHR was associated with decreased fT4 ($\beta = -0.081$, $P = 0.002$). Genetically increased BMI was associated with increased TSH ($\beta = 0.03$, $P = 0.028$) and with decreased fT4 ($\beta = -0.078$, $P = 1.05e-4$). Causal associations of WHR and BMI with thyroid signaling were not supported by weighted median analysis in sensitivity analyses. *Conclusion*. TSH and fT4 were increased due to the higher genetically predicted CRP. WHR was decreased due to the higher genetically predicted TSH. These findings will provide reference for the prevention and treatment of inflammation and metabolic syndrome.

1. Introduction

Subclinical thyroid disease is a common public health issue. In hypothyroidism due to thyroid dysfunction, serum thyroid-stimulating hormone (TSH) levels are appropriately elevated while serum free thyroxine (fT4) levels are within normal range [1, 2]. Subclinical hypothyroidism affects up to 10% of the adult population [3]. A lot of previous studies showed variation in TSH or fT4 may increase the risk of future cardiovascular diseases (CVD) [4–6]. Recently, many investigations indicated that even in normal thyroid function individuals, variation in TSH and fT4 was associated with an increased risk of CVD and metabolic diseases [7–9], including obesity [1]. Therefore, it is important and urgent to pay more attention to TSH and fT4.

C-reactive protein (CRP) is an acute-phase inflammatory protein, which has been traditionally utilized as a clinical marker of inflammation, infection, and tissue damage [10]. Generally, CRP exhibits elevated expression during inflammatory disorders, such as CVD, rheumatoid arthritis, and some acute or chronic infection [11, 12]. Recently, research outputs showed that minor CRP elevation could contribute to an increased future risk of major cardiovascular events [13, 14]. In addition, there was growing evidence that elevated CRP levels are associated with cancer disease risk [15, 16]. Hence, CRP measurements have potential

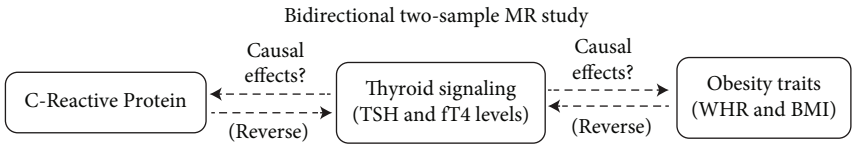


Figure 1: Schematic diagram. Bidirectional two-sample MR approach based on the summary level data from large scale meta-analyses of the GWASs was used to investigate the causal relationships between thyroid signaling and CRP. Further bidirectional two-sample MR approach was used to investigate the causal relationships between thyroid signaling and obesity traits. All data sets used in this study are publicly available at the GWAS Catalog, dbGaP, and the GIANT websites. TSH: thyroid-stimulating hormone; fT4: free thyroxine; GWAS: genome-wide association study; CRP: C-reactive protein; MR: Mendelian randomization; WHR: waist-hip ratio; BMI: body mass index.

utility as a clinical tool in assessing disease status and progression, including CVD, some infections, and cancer. Therefore, with the important role of CRP, more studies are needed to understand the complex mechanism of CRP production.

Existing literature demonstrated that subclinical hypothyroidism may be associated with elevated high-sensitive CRP, although the clinical implications were uncertain [17–19]. Some researchers found that there was a significant positive correlation between TSH and CRP [17, 20]. Meanwhile, a Brazilian longitudinal study of adult health also investigated the association between TSH and CRP, but this study showed that TSH was not associated with CRP because of the existence of confounders [21]. So, it was controversial about the relationship between thyroid signaling and CRP. Similarly, there were researches investigating the association of fT4 with CRP [18, 22, 23], and their conclusions were also controversial. Therefore, the association between thyroid signaling and CRP is hard to uncover.

Furthermore, as an important public health problem, obesity is also an important risk factor for CVD [11]. Previous investigations revealed a significant relationship between CRP and obesity [24–26]. In obese and overweight adults, CRP levels are significantly increased [27]. So it is meaningful to study obesity traits, such as waist-hip ratio (WHR) and body mass index (BMI).

Correlation describes whether two variables "go together." However, the fact that two variables change together does not necessarily mean that we know whether one variable causes the other to change or vice versa [28]. Therefore, it is necessary to study causal association. To this end, one powerful method is Mendelian randomization (MR) [29], which uses genetic variants as instrumental variants (IVs) and has been widely used [30]. MR can minimize the influence of confounding factors on the causal association between two variables, exposure and outcome. Note both individual data and publicly available genome-wide association study (GWAS) summary statistics are applicable in MR analyses. Moreover, bidirectional two-sample MR can explore the nature and direction of the links between them.

To date, as far as we know, no studies investigated the causal associations of TSH and fT4 levels with CRP levels. In this paper, we studied the causal association between thyroid signaling and CRP level. To further detect the possible causes of CVD, we also studied two obesity traits, WHR and BMI. For this, we utilized summary data from the latest and largest GWASs [31–33] and inferred causality in bidirectional two-sample MR analyses.

2. Materials and Methods

2.1. Data Sources. The first is the source of summary data related to thyroid signaling. Summary data for TSH within reference range were obtained from a GWAS meta-analysis that is the largest GWAS on thyroid function to date, including 120000 subjects, with more than 22 million single nucleotide polymorphisms (SNPs) [31]. These data are accessed through the GWAS Catalog (https://www.ebi.ac.uk/gwas). Summary data for fT4 within reference range were obtained from a GWAS meta-analysis in up to 72167 individuals with 8 million SNPs [32], which can be downloaded on dbGaP website under the accession number phs000930 (https://www.ncbi.nlm.nih.gov/gap).

Second is the source of summary data on inflammatory factor CRP. Summary data for CRP were obtained from a GWAS meta-analysis which is the largest data set on inflammatory factors lately, including 49839 subjects (CRP: mean $= 4.114$ (SD $= 4.836$)) [33]. These summary data can be available through the GWAS Catalog (https://www.ebi.ac.uk/gwas). Summary data for WHR were obtained from a GWAS meta-analysis in 694649 individuals of European ancestry with 2.7 million SNPs combining UK Biobank and GIANT [34]. Summary data for BMI were obtained from a GWAS meta-analysis which included about 700000 participants of European ancestry with 2.3 million SNPs from GIANT [35].

2.2. Two-Sample MR. We conducted bidirectional two-sample MR analyses using data published by GWAS (Figure 1). Because the data is public, there is no need of ethical review.

2.3. Selection of SNPs. Based on the GWAS results [31–33] on TSH, fT4, and CRP, we used independent SNPs which are strongly associated at a genome-wide significant level ($P < 5 \times 10^{-8}$) with TSH, fT4, and CRP, respectively. The selected SNPs were used as IVs in using MR method.

2.4. Statistical Analysis. In order to avoid the estimator bias caused by weak IVs as much as possible, we calculated the F statistic ($F = \beta^2_{\text{exposure}}/\text{SE}^2_{\text{exposure}}$) as a measure of strength for each SNP. According to the existing literature, criterion of $F \geq 10$ was adopted for screening strong IVs (F statistic was in 30.01–1231.188 for TSH, 30.25–455.33 for fT4, and 27.94–528.51 for CRP) [36]. The primary analysis used to examine the causality between exposure and outcome was inverse-variance weighted (IVW) method [37]. IVW

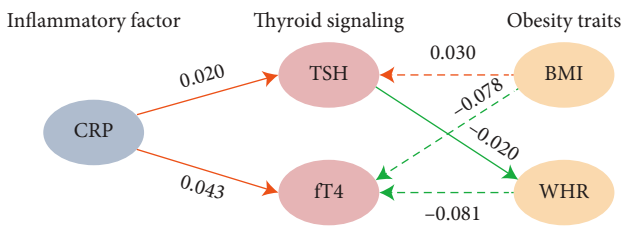


Figure 2: Causal effects between thyroid signaling, CRP, and obesity traits. TSH and fT4 are shown in pink ovals, CRP is shown in gray ovals, and BMI and WHR are shown in yellow ovals. The arrows' direction denotes causal direction. The solid line and the dotted line, respectively, indicate whether the causal relationship is robust or not. The red and green arrows denote positive and negative causal relationships, respectively, and the number beside each arrow is the causal effects. All causal relationships are significant at $P \leq 0.05$.

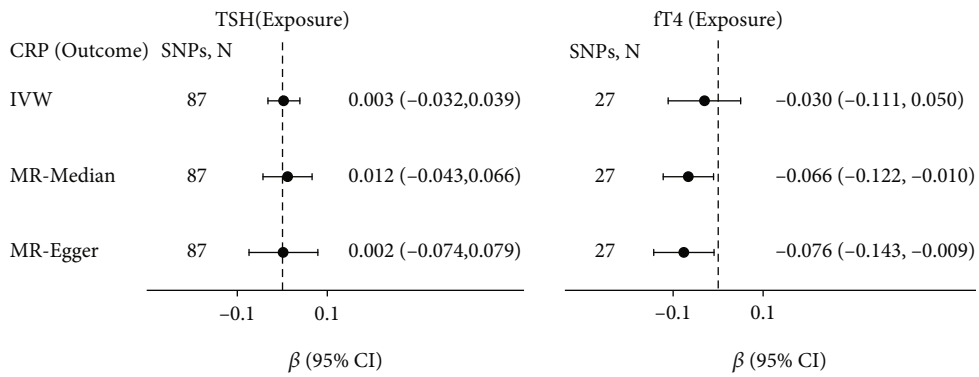


Figure 3: Causal effects of variation in TSH and fT4 levels on CRP. Presented βs and CIs (horizontal lines and their corresponding numerical interval form on the right side) correspond to the effects of a one SD change in TSH or fT4 levels on the outcome CRP levels. The results of MR analyses using various analysis methods (IVW, MR-Median, MR-Egger) are presented for comparison. The number of SNPs indicates the number of genetic variants used as instrument variables for MR analysis. MR: Mendelian randomization; SNPs: single nucleotide polymorphisms; CI: confidence interval; IVW: inverse-variance weighted; MR-Median: weighted median method; MR-Egger: Egger regression.

method aggregated two or more IVs to minimize the variance of the weighted average, and the weight given to each IV was the inverse of the variance of the effect estimate [38]. Note that the estimated effect obtained by IVW may be biased, which may be due to the violation of one assumption of IV. Specifically, IVs and outcome are not only related through exposure but also directly related, which is termed as pleiotropy. We addressed the problem of pleiotropy in sensitivity analyses.

In sensitivity analyses, we assessed the robustness of IVW in two complementary sensitivity analyses with different assumptions about horizontal pleiotropy: weighted median (MR-Median) [38] and MR-Egger regression [39]. MR-Median yielded consistent causal effect estimates compared with IVW method. Egger intercept in MR-Egger represented the average horizontal pleiotropic effect across the IVs. We used I^2 statistic and Cochran's Q test to quantify heterogeneity across all SNPs. If the results indicated the presence of horizontal pleiotropy or significant heterogeneity suggesting pleiotropy [40], we calculated individual Q statistic for each SNP, and SNPs were identified as potential pleiotropic variants if their individual Q statistics exceeded the 95th percentile of the chi-square distribution with one degree of freedom [41–43]. After excluding these potential pleiotropic IVs, the IVW, MR-Median, and MR-Egger methods were performed on the remaining IVs.

For the estimated causal effect of the exposure on the outcome, a P value of less than 0.05 was considered as statistically significant. Statistical analysis was performed with R package "MendelianRandomization" version 0.5.1 in R version 4.1.0.

3. Results

Use the MR analysis method in previous sections to explore the causal relationships between thyroid signaling, CRP, and obesity traits. The MR-Egger intercepts were insignificant ($P > 0.05$) in all analyses. The result diagram is shown in Figure 2. The diagram showed whether there was a causal relationship between two subjects and showed the magnitudes and directions of the causal relationships. All causal relationships are significant at $P \leq 0.05$.

3.1. Causal Relationships between Thyroid Signaling and CRP. The results of MR analyses between genetically predicted TSH and fT4 levels (exposure) and CRP levels (outcome) are presented in Figure 3. 87 SNPs were in consideration when we investigated the causal association between TSH and CRP and 30 SNPs for fT4 and CRP. Based on this analysis, we found neither serum TSH nor fT4 levels could cause changes in CRP (TSH: $\beta = 0.003$, 95% CI = −0.032–0.039, $P = 0.856$; fT4: $\beta = 0.003$, 95% CI = −0.084–0.089, $P = 0.953$) (also, see details in Supplementary Table S1).

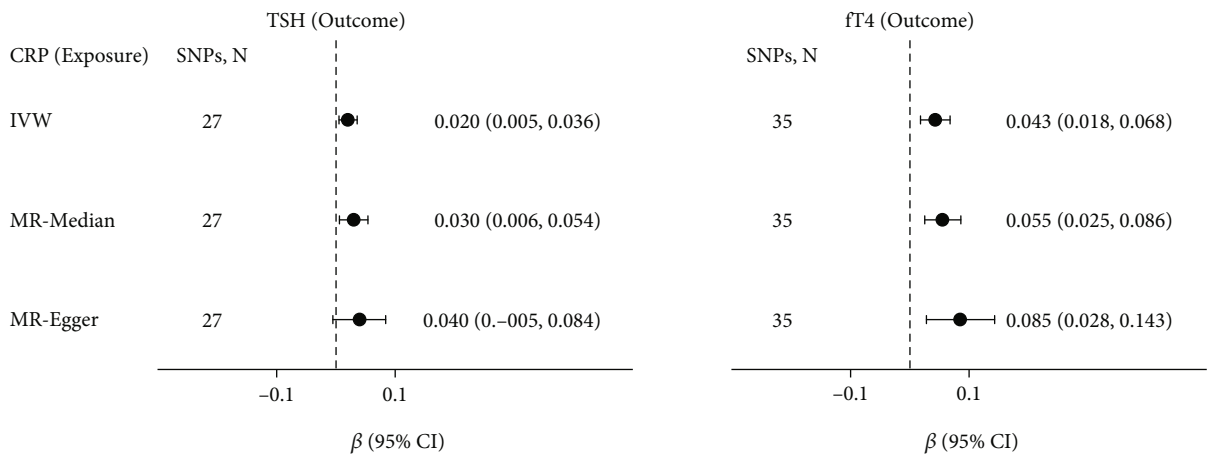


Figure 4: Causal effects of variation in CRP levels on TSH and fT4. Presented βs and CIs (horizontal lines and their corresponding numerical interval form on the right side) correspond to the effects of a one SD change in CRP levels on the outcome TSH or fT4 levels. The results of MR analyses using various analysis methods (IVW, MR-Median, MR-Egger) are presented for comparison. The number of SNPs indicates the number of genetic variants used as instruments for MR analysis.

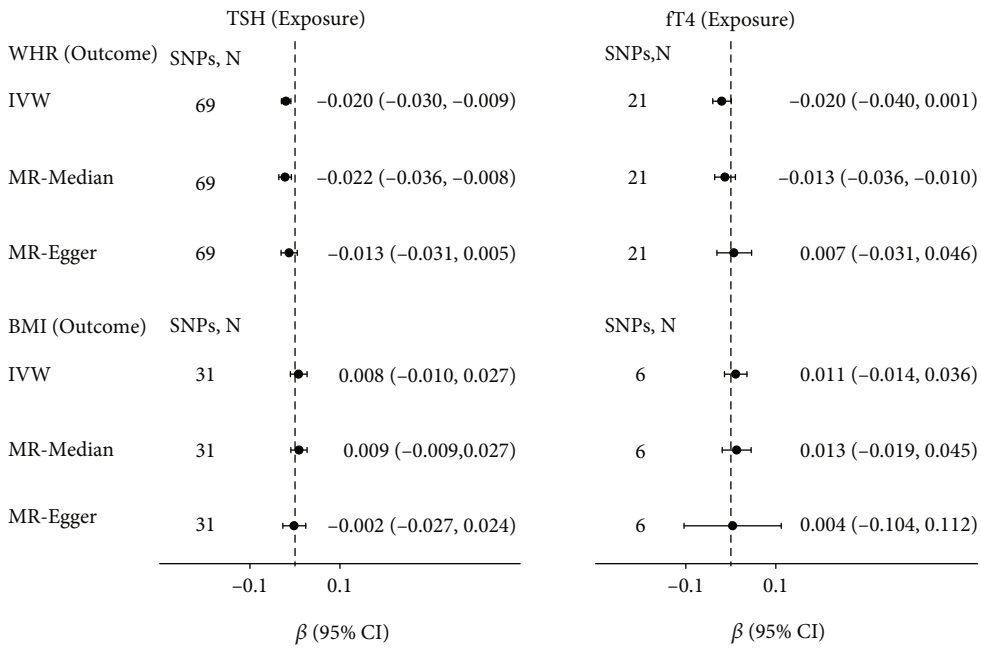


Figure 5: Causal effects of variation in TSH and fT4 levels on WHR and BMI. Presented βs and CIs (horizontal lines and their corresponding numerical interval form on the right side) correspond to the effects of a one SD change in TSH or fT4 levels on the outcome WHR or BMI levels. The results of MR analyses using various analysis methods (IVW, MR-Median, MR-Egger) are presented for comparison. The number of SNPs indicates the number of genetic variants used as instruments for MR analysis.

However, exchanging the exposure and outcome of interest in MR yielded different results (Figure 4). 27 SNPs were considered as IVs when we investigated the causal association between CRP and TSH. There was some evidence that higher CRP levels might cause higher TSH levels ($\beta = 0.02$, 95% CI = 0.005–0.036, $P = 0.011$), which was confirmed in sensitivity analyses using MR-Median method (see Supplementary Table S2).

For the association between CRP and fT4, 35 SNPs were taken as IVs as shown in Figure 4. There was some strong evidence that higher genetically predicted FT4 might cause higher CRP levels ($\beta = 0.013$, 95% CI = 0.018–0.068, $P = 0.001$), which was also in line with the results of sensitivity analyses using the MR-Median and MR-Egger method (see Supplementary Table S2).

3.2. Causal Relationships between Thyroid Signaling and Obesity Traits. Due to elevated CRP levels in overweight and obese adults [27], we also wanted to know whether there exist causal associations between thyroid signaling and obesity traits. Therefore, we conducted MR analyses of TSH and fT4 on obesity traits, respectively (Figure 5). MR analyses

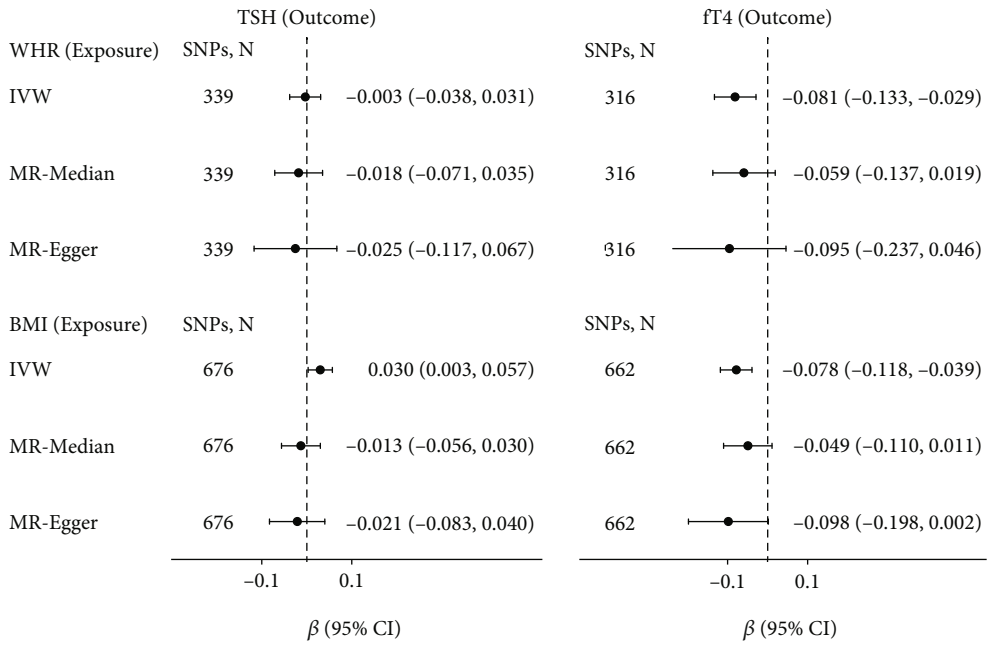


Figure 6: Causal effects of variation in WHR and BMI levels on TSH and fT4. Presented βs and CIs (horizontal lines and their corresponding numerical interval form on the right side) correspond to the effects of a one SD change in WHR or BMI levels on TSH or fT4 levels. The results of MR analyses using various analysis methods (IVW, MR-Median, MR-Egger) are presented for comparison. The number of SNPs indicates the number of genetic variants used as instruments for MR analysis.

showed higher genetically predicted TSH could cause decreased WHR levels ($\beta = -0.02$, 95% CI = −0.030–0.009, $P = 3.99e-4$), and MR-Median also led to similar results (see Supplementary Table S3 for more information).

For the association between obesity traits and thyroid signals (Figure 6), MR analyses showed higher genetically predicted WHR could cause decreased fT4 levels ($\beta = -0.081$, 95% CI = −0.133–-0.029, $P = 0.002$), and higher genetically predicted BMI could cause higher TSH levels ($\beta = 0.030$, 95% CI = 0.003–0.057, $P = 0.028$) and lower fT4 ($\beta = 0.02$, 95% CI = −0.118–0.068, $P = 1.05e-4$). This causal relationship was not robust because it is not supported by MR-Median (see Supplementary Table S4).

4. Discussion

In this study, the bidirectional two-sample MR analyses between thyroid signaling (TSH and fT4) and CRP levels were accomplished based on the current largest GWAS summary statistics. We studied the causal relationships between TSH and fT4 levels and CRP levels and found TSH and fT4 levels could be affected by CRP, whereas TSH and fT4 levels could not affect CRP levels. Furthermore, we found some evidence that there were associations between obesity traits (BMI and WHR) and fT4 levels. We also found that TSH could be significantly affected by BMI.

CRP responds quickly to inflammatory processes and is utilized as one of the best inflammatory markers. Various research results showed that there was a significant positive correlation between TSH and CRP [17, 44, 45]. However, their underlying causality was still unclear. A prospective study indicated that patients with subclinical hypothyroidism had increased levels of signs of low-grade inflammation (CRP levels) [20, 46]. On the other hand, the conclusions of many studies were not consistent with this prospective study. For example, some authors believed that serum CRP was not significantly affected by the thyroid dysfunction's degree [47]. An observational study found that CRP was not correlated with fT4 and TSH [48]. These studies indicated that further evidence was needed to determine the causal link between TSH and fT4 levels and CRP levels. In this study, based on MR analysis results, we found that there was a causal association between CRP and thyroid signaling (TSH and fT4). TSH and fT4 levels could be positively affected by CRP levels, but not vice versa. The underlying cause of CRP affecting thyroid signaling is still unclear, possibly because severe inflammation may significantly affect the thyroid gland, leading to changes in thyroid signaling. Besides, we thought there were some potential effects of inflammation on deiodinase activity. Inflammation (elevated in CRP levels) which was related to infection or injury led to a reduction in deiodinase activity. This results in decreased conversion of fT4 to fT3, leading to high fT4 [17]. In the future, the causal relationship between CRP levels and TSH and fT4 levels may be confirmed with larger populations and more precise statistical methods.

Interestingly, in a Brazilian longitudinal study of adult health, obesity was considered as one of the most important confounders in the association study between TSH and CRP [21]. Some researches showed that CRP was correlated with obesity and the role of obesity in inflammation can not be ignored [49]. This promoted us to study further the obesity traits.

In the subsequent MR analyses, we found that increased TSH could cause decreased WHR. In reverse MR analyses, increased WHR and BMI could cause decrease in fT4, and increased BMI could cause increase in TSH. Previous literature showed that lower fT4 was consistently associated with obesity in healthy euthyroid people [50, 51]. One research indicated that serum fT4 levels were negatively correlated with BMI and serum TSH levels were positively correlated with WHR and BMI [52]. It was suggested that the increase in fT3 levels in obese people may be a compensatory mechanism for the fat accumulation increase [53]. In obese people, thyroxine 5-deiodinase increased activity, inducing the increased peripheral conversion of fT4 to fT3 [49, 54]. The lower fT4 in obese and overweight people might partially result from this cause. These were consistent with the results of our MR study. Conclusion of a recent MR analysis was also consistent with our study; i.e., genetically predicted BMI was inversely associated with fT4 levels [55].

Another MR analysis pointed out that TSH could be significantly elevated by the genetically driven BMI, while fT4 could not be affected by BMI [56]. Notice in our study, we used the latest and largest GWAS summary data, where fT4 cohorts included nearly 70000 participants. Moreover, we performed sensitivity analyses to exclude pleiotropic and heterogeneous IVs, because these heterogenous SNPs could partially result in bias in MR analysis. In our analysis, TSH and fT4 both could be affected by BMI.

Advantages of this study design were that (1) GWAS data were freely available obtained from the largest recent GWAS on TSH, fT4, and CRP, respectively; (2) sensitivity analyses were performed in order to reduce potential bias resulting from potential pleiotropic and heterogeneous IVs. It is the first time to reach a conclusion based on MR analysis that higher genetically predicted CRP may induce an increase in TSH and fT4. However, this study has certain limitations. (1) Due to the accession of the public databases, we used people of diverse ancestry for CRP and people of European ancestry for thyroid signaling and obesity traits. MR analysis for population-stratification and other populations should be considered if related data can be available; (2) generally speaking, thyroid function is sex-specific; due to the limitation of TSH and fT4 summary data, we did not perform the sex-specific MR analyses.

Taken together, the bidirectional MR study demonstrated that higher TSH and fT4 levels were causally affected by higher CRP levels, but not vice versa. Further MR analyses provided evidence that higher obesity traits could cause lower fT4 and higher BMI could cause higher TSH.

Data Availability

All data generated or analyzed during this study are included in this published article or in the data repositories listed in References. Summary statistic data for genetic associations with thyroid signaling have been contributed by the thyroid GWAS meta-analysis of Hunt and the ThyroidOmics consortium. Summary statistic data for genetic association with body mass index and waist-hip ratio have been contributed by the GIANT consortium and MEGASTROKE consortium. Summary statistic data for genetic association with C-reactive protein have been contributed by the Population Architecture Using Genomics and Epidemiology study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

T.L., H.G., and Y.H. conceptualized the study. T. L., Y. W. and Y. H. were responsible for the methodology. Y.H. was responsible for the funding acquisition. T.L. was responsible for the investigation. Y.H. was responsible for the project administration. T.L. was responsible for the resources. T.L. was responsible for the software. Y.W. and Y.H. supervised the study. T.L. and S.Y. validated the study. T.L. was responsible for the writing—original draft. H. G., Z. W., S. Y., and Y. H. were responsible for the writing, review and editing. All authors have read and agreed to the published version of the manuscript. Tingting Li, Haigang Geng, and Yuquan Wang contributed equally to this work.

Acknowledgments

This study was supported by grants to YH from the National Natural Science Foundation of China (grants nos. 11971117 and 11571082).

Supplementary Materials

Supplementary Table S1: results of Mendelian randomization analyses between genetically predicted thyroid-stimulating hormone (TSH) and free thyroid hormone (fT4) levels (exposure) and C-reactive protein (CRP) levels (outcome). Supplementary Table S2: results of Mendelian randomization analyses between genetically predicted CRP levels (exposure) and TSH and fT4 levels (outcome). Supplementary Table S3: results of Mendelian randomization analyses between genetically predicted TSH and fT4 levels (exposure) and obesity traits (outcome). Supplementary Table S4: results of Mendelian randomization analyses between genetically predicted obesity traits (exposure) and TSH and fT4 levels (outcome). *(Supplementary Materials)*

References

[1] M. A. Michalaki, A. G. Vagenakis, A. S. Leonardou et al., "Thyroid function in humans with morbid obesity," *Thyroid*, vol. 16, no. 1, pp. 73–78, 2006.

[2] B. Biondi, A. R. Cappola, and D. S. Cooper, "Subclinical hypothyroidism," *Journal of the American Medical Association*, vol. 322, no. 2, pp. 153–160, 2019.

[3] J. G. Hollowell, N. W. Staehling, W. D. Flanders et al., "Serum TSH, T4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III)," *The Journal of Clinical Endocrinology & Metabolism.*, vol. 87, no. 2, pp. 489–499, 2002.

[4] N. Rodondi, W. P. J. den Elzen, D. C. Bauer et al., "Subclinical hypothyroidism and the risk of coronary heart disease and mortality," *Journal of the American Medical Association*, vol. 304, no. 12, pp. 1365–1374, 2010.

[5] C. Floriani, B. Gencer, T.-H. Collet, and N. Rodondi, "Subclinical thyroid dysfunction and cardiovascular diseases: 2016 update," *European Heart Journal*, vol. 39, no. 7, pp. 503–507, 2018.

[6] A. E. Hak, H. A. Pols, T. J. Visser, H. A. Drexhage, A. Hofman, and J. C. Witteman, "Subclinical hypothyroidism is an independent risk factor for atherosclerosis and myocardial infarction in elderly women: the Rotterdam study," *Annals of Internal Medicine*, vol. 132, no. 4, pp. 270–278, 2000.

[7] L. J. van Tienhoven-Wind and R. P. Dullaart, "Low–normal thyroid function and the pathogenesis of common cardiometabolic disorders," *European Journal of Clinical Investigation*, vol. 45, no. 5, pp. 494–503, 2015.

[8] B. O. Åsvold, T. Bjøro, C. Platou, and L. J. Vatten, "Thyroid function and the risk of coronary heart disease: 12-year follow-up of the HUNT study in Norway," *Clinical Endocrinology.*, vol. 77, no. 6, pp. 911–917, 2012.

[9] N. Takamura, A. Akilzhanova, N. Hayashida et al., "Thyroid function is associated with carotid intima-media thickness in euthyroid subjects," *Atherosclerosis*, vol. 204, no. 2, pp. e77–e81, 2009.

[10] M. A. Albert, E. Danielson, N. Rifai, P. M. Ridker, and PRINCE Investigators, "Effect of statin therapy on C-reactive protein levels: the pravastatin inflammation/CRP evaluation (PRINCE): a randomized trial and cohort study," *Journal of the American Medical Association*, vol. 286, no. 1, pp. 64–70, 2001.

[11] N. R. Sproston and J. J. Ashworth, "Role of C-reactive protein at sites of inflammation and infection," *Frontiers in Immunology.*, vol. 9, no. 754, 2018.

[12] T. P. Erlinger, E. A. Platz, N. Rifai, and K. J. Helzlsouer, "C-reactive protein and the risk of incident colorectal cancer," *Journal of the American Medical Association*, vol. 291, no. 5, pp. 585–590, 2004.

[13] P. M. Ridker, M. Cushman, M. J. Stampfer, R. P. Tracy, and C. H. Hennekens, "Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men," *New England Journal of Medicine.*, vol. 336, no. 14, pp. 973–979, 1997.

[14] P. M. Ridker, C. H. Hennekens, J. E. Buring, and N. Rifai, "C-reactive protein and other markers of inflammation in the prediction of cardiovascular disease in women," *New England Journal of Medicine.*, vol. 342, no. 12, pp. 836–843, 2000.

[15] P. C. Hart, I. M. Rajab, M. Alebraheem, and L. A. Potempa, "C-reactive protein and cancer—diagnostic and therapeutic insights," *Frontiers in Immunology.*, vol. 11, article 595835, 2020.

[16] K. H. Allin and B. G. Nordestgaard, "Elevated C-reactive protein in the diagnosis, prognosis, and cause of cancer," *Critical Reviews in Clinical Laboratory Sciences.*, vol. 48, no. 4, pp. 155–170, 2011.

[17] R. Sharma, T. K. Sharma, G. Kaushik, S. Sharma, S. Vardey, and M. Sinha, "Subclinical hypothyroidism and its association with cardiovascular risk factors," *Clinical Laboratory*, vol. 57, no. 9-10, pp. 719–724, 2011.

[18] M. Christ-Crain, C. Meier, M. Guglielmetti et al., "Elevated C-reactive protein and homocysteine values: cardiovascular risk factors in hypothyroidism? A cross-sectional and a double-blind, placebo- controlled trial," *Atherosclerosis*, vol. 166, no. 2, pp. 379–386, 2003.

[19] N. Boulman, Y. Levy, R. Leiba et al., "Increased C-reactive protein levels in the polycystic ovary syndrome: a marker of cardiovascular disease," *The Journal of Clinical Endocrinology & Metabolism.*, vol. 89, no. 5, pp. 2160–2165, 2004.

[20] A. Tuzcu, M. Bahceci, D. Gokalp, Y. Tuzun, and K. Gunes, "Subclinical hypothyroidism may be associated with elevated high-sensitive C-reactive protein (low grade inflammation) and fasting hyperinsulinemia," *Endocrine Journal.*, vol. 52, no. 1, pp. 89–94, 2005.

[21] É. J. F. P. de Miranda, M. S. Bittencourt, I. S. Santos, P. A. Lotufo, and I. M. Benseñor, "Thyroid function and high-sensitivity C-reactive protein in cross-sectional results from the Brazilian longitudinal study of adult health (Elsa-Brasil): effect of adiposity and insulin resistance," *European Thyroid Journal*, vol. 5, no. 4, pp. 240–246, 2016.

[22] C. Jublanc, E. Bruckert, P. Giral et al., "Relationship of circulating C-reactive protein levels to thyroid status and cardiovascular risk in hyperlipidemic euthyroid subjects: low free thyroxine is associated with elevated hsCRP," *Atherosclerosis*, vol. 172, no. 1, pp. 7–11, 2004.

[23] S. Singh, "Serum lipids, tHcy, hs-CRP, MDA and PON-1 levels in SCH and overt hypothyroidism: effect of treatment," *Acta Biomedica Atenei Parmensis*, vol. 85, no. 2, pp. 127–134, 2014.

[24] J. R. Greenfield, K. Samaras, A. B. Jenkins et al., "Obesity is an important determinant of baseline serum C-reactive protein concentration in monozygotic twins, independent of genetic influences," *Circulation*, vol. 109, no. 24, pp. 3022–3028, 2004.

[25] M. Hiura, T. Kikuchi, K. Nagasaki, and M. Uchiyama, "Elevation of serum C-reactive protein levels is associated with obesity in boys," *Hypertension Research.*, vol. 26, no. 7, pp. 541–546, 2003.

[26] M. Ebrahimi, A. R. Heidari-Bakavoli, S. Shoeibi et al., "Association of serum hs-CRP levels with the presence of obesity, diabetes mellitus, and other cardiovascular risk factors," *Journal of Clinical Laboratory Analysis.*, vol. 30, no. 5, pp. 672–676, 2016.

[27] M. Visser, L. M. Bouter, G. M. McQuillan, M. H. Wener, and T. B. Harris, "Elevated C-reactive protein levels in overweight and obese adults," *Journal of the American Medical Association*, vol. 282, no. 22, pp. 2131–2135, 1999.

[28] P. W. Holland, "Statistics and causal inference," *Journal of the American Statistical Association.*, vol. 81, no. 396, pp. 945–960, 1986.

[29] I. Postmus, J. Deelen, S. Sedaghat et al., "LDL cholesterol still a problem in old age? A Mendelian randomization study," *International Journal of Epidemiology.*, vol. 44, no. 2, pp. 604–612, 2015.

[30] J. R. Thompson, C. Minelli, K. R. Abrams, M. D. Tobin, and R. D. Riley, "Meta-analysis of genetic studies using Mendelian randomization—a multivariate approach," *Statistics in Medicine.*, vol. 24, no. 14, pp. 2241–2254, 2005.

[31] W. Zhou, B. Brumpton, O. Kabil et al., "GWAS of thyroid stimulating hormone highlights pleiotropic effects and inverse association with thyroid cancer," *Nature Communications*, vol. 11, no. 1, pp. 1–13, 2020.

[32] A. Teumer, L. Chaker, S. Groeneweg et al., "Genome-wide analyses identify a role for SLC17A4 and AADAT in thyroid hormone regulation," *Nature Communications*, vol. 9, no. 1, pp. 1–14, 2018.

[33] G. L. Wojcik, M. Graff, K. K. Nishimura et al., "Genetic analyses of diverse populations improves discovery for complex traits," *Nature Communications*, vol. 570, no. 7762, pp. 514–518, 2019.

[34] S. L. Pulit, C. Stoneman, A. P. Morris et al., "Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry," *Human Molecular Genetics.*, vol. 28, no. 1, pp. 166–174, 2019.

[35] L. Yengo, J. Sidorenko, K. E. Kemper et al., "Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry," *Human Molecular Genetics.*, vol. 27, no. 20, pp. 3641–3649, 2018.

[36] B. L. Pierce, H. Ahsan, and T. J. Vander Weele, "Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants," *International Journal of Epidemiology*, vol. 40, no. 3, pp. 740–752, 2011.

[37] S. Burgess, A. Butterworth, and S. G. Thompson, "Mendelian randomization analysis with multiple genetic variants using summarized data," *Genetic Epidemiology*, vol. 37, no. 7, pp. 658–665, 2013.

[38] J. Bowden, G. Davey Smith, P. C. Haycock, and S. Burgess, "Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator," *Genetic Epidemiology*, vol. 40, no. 4, pp. 304–314, 2016.

[39] J. Bowden, G. Davey Smith, and S. Burgess, "Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression," *International Journal of Epidemiology.*, vol. 44, no. 2, pp. 512–525, 2015.

[40] M. F. D. Greco, C. Minelli, N. A. Sheehan, and J. R. Thompson, "Detecting pleiotropy in Mendelian randomisation studies with summary data and a continuous outcome," *Statistics in Medicine*, vol. 34, no. 21, pp. 2926–2940, 2015.

[41] S. Burgess, J. Bowden, T. Fall, E. Ingelsson, and S. G. Thompson, "Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants," *Epidemiology*, vol. 28, no. 1, pp. 30–42, 2017.

[42] W. G. Cochran, "The comparison of percentages in matched samples," *Biometrika*, vol. 37, no. 3-4, pp. 256–266, 1950.

[43] B. Baujat, C. Mahé, J. P. Pignon, and C. Hill, "A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials," *Statistics in Medicine.*, vol. 21, no. 18, pp. 2641–2652, 2002.

[44] G. Gupta, P. Sharma, P. Kumar, and M. Itagappa, "Study on subclinical hypothyroidism and its association with various inflammatory markers," *Journal of Clinical and Diagnostic Research: JCDR.*, vol. 9, no. 11, pp. BC04–BC06, 2015.

[45] J. Wu, Y. Tao, H. Gu, and J. Sui, "Association between cardiovascular risk factors and serum thyrotropin concentration among healthy Chinese subjects and subjects with unsuspected subclinical hypothyroidism," *Clinical Laboratory*, vol. 62, no. 5, pp. 807–814, 2016.

[46] J. Kvetny, P. Heldgaard, E. M. Bladbjerg, and J. Gram, "Subclinical hypothyroidism is associated with a low-grade inflammation, increased triglyceride levels and predicts cardiovascular disease in males below 50 years," *Clinical Endocrinology*, vol. 61, no. 2, pp. 232–238, 2004.

[47] W.-Y. Lee, J.-Y. Suh, E.-J. Rhee, J.-S. Park, K.-C. Sung, and S.-W. Kim, "Plasma CRP, apolipoprotein A-1, apolipoprotein B and Lp(a) levels according to thyroid function status," *Archives of Medical Research*, vol. 35, no. 6, pp. 540–545, 2004.

[48] V. Lubrano, A. Pingitore, A. Carpi, and G. Iervasi, "Relationship between triiodothyronine and proinflammatory cytokines in chronic heart failure," *Biomedicine & Pharmacotherapy.*, vol. 64, no. 3, pp. 165–169, 2010.

[49] D. Y. Aksoy, N. Cinar, A. Harmanci et al., "Serum resistin and high sensitive CRP levels in patients with subclinical hypothyroidism before and after L-thyroxine therapy," *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research.*, vol. 19, pp. 210–215, 2013.

[50] G. L. Roef, E. R. Rietzschel, C. M. Van Daele et al., "Triiodothyronine and free thyroxine levels are differentially associated with metabolic profile and adiposity-related cardiovascular risk markers in euthyroid middle-aged subjects," *Thyroid*, vol. 24, no. 2, pp. 223–231, 2014.

[51] C. M. Kitahara, E. A. Platz, P. W. Ladenson, A. M. Mondul, A. Menke, and A. B. de González, "Body fatness and markers of thyroid function among U.S. men and women," *PLoS One*, vol. 7, no. 4, article e34979, 2012.

[52] F.-M. Du, H.-Y. Kuang, B.-H. Duan, D.-N. Liu, and X.-Y. Yu, "Effects of thyroid hormone and depression on common components of central obesity," *Journal of International Medical Research.*, vol. 47, no. 7, pp. 3040–3049, 2019.

[53] G. De Pergola, A. Ciampolillo, S. Paolotti, P. Trerotoli, and R. Giorgino, "Free triiodothyronine and thyroid stimulating hormone are directly associated with waist circumference, independently of insulin resistance, metabolic parameters and blood pressure in overweight and obese women," *Clinical Endocrinology.*, vol. 67, no. 2, pp. 265–269, 2007.

[54] R. Mullur, Y.-Y. Liu, and G. A. Brent, "Thyroid hormone regulation of metabolism," *Physiological Reviews*, vol. 94, no. 2, pp. 355–382, 2014.

[55] A. Kuś, E. Marouli, M. F. Del Greco et al., "Variation in normal range thyroid function affects serum cholesterol levels, blood pressure, and type 2 diabetes risk: a Mendelian randomization study," *Thyroid*, vol. 31, no. 5, pp. 721–731, 2021.

[56] X. Wang, X. Gao, Y. Han et al., "Causal association between serum thyrotropin and obesity: a bidirectional, Mendelian randomization study," *The Journal of Clinical Endocrinology & Metabolism*, vol. 106, no. 10, pp. e4251–e4259, 2021.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 1750834, 12 pages
https://doi.org/10.1155/2022/1750834

Research Article

Diagnostic and Prognostic Nomograms for Lung Metastasis in Triple-Negative Breast Cancer

Jianguo Wang, Hongjun Zhao, Lifen Ye, Jingyong Li, Huaixiao Zhang, Chao Zhang, Qishuo Rao, Yurong Cai, Yiping Xu, and Youyuan Deng

Department of General Surgery, Xiangtan Central Hospital, Xiangtan, China

Correspondence should be addressed to Youyuan Deng; dengyouyuan@smail.hunnu.edu.cn

Received 15 June 2022; Accepted 21 July 2022; Published 11 August 2022

Academic Editor: Xiaohan Ren

Background. The lungs are one of the common sites of metastasis of triple-negative breast cancer (TNBC). Patients with lung metastases (LM) have a shorter duration of survival. This study is aimed at determining the prognostic factors of patients with TNBC with LM and constructing two nomograms to assess the risk of LM and the prognosis of patients with TNBC with LM. *Methods*. Clinicopathological and follow-up data of patients with TNBC between 2010 and 2015 were retrieved from the Surveillance, Epidemiology, and End Results (SEER) database. Univariate and multivariate Cox regression analyses were used to screen for independent predictors of LM in patients with TNBC and identify the independent prognostic factors of patients with TNBC with LM. The two nomograms were appraised using calibration curves, receiver operating characteristic (ROC) curves, and decision curve analysis (DCA). *Results*. A total of 27,048 patients with TNBC were included in this study. Age, tumour size, T stage, and N stage were identified as independent risk factors for LM in patients with TNBC. Histological type, marital status, prior surgery, chemotherapy, bone metastases, brain metastases, and LM were confirmed as independent prognostic factors for patients with TNBC with LM. The area under the ROC curve (AUC) of the diagnostic nomogram was 0.838 (95% confidence interval 0.817-0.860) in the training cohort and 0.894 (95% confidence interval 0.875-0.917) in the verification cohort. The AUC values of the 6-, 12-, and 18-month prognostic nomograms in the training cohort were 0.809 (95% confidence interval 0.771-0.868), 0.779 (95% confidence interval 0.737-0.834), and 0.735 (95% confidence interval 0.699-0.811), respectively, and the corresponding AUC values in the validation cohort were 0.735(95% confidence interval 0.642-0.820), 0.672 (95% confidence interval 0.575-0.758), and 0.705 (95% confidence interval 0.598-0.782), respectively. According to the calibration curves and data analysis, both nomograms exhibited good performance. *Conclusion*. We successfully constructed and verified two valuable nomograms for predicting the incidence of LM and prognosis of patients TNBC with LM.

1. Introduction

Breast cancer is the most common malignant tumour in women, and breast cancer-specific deaths accounted for approximately 15% of cancer-related deaths in women in 2018 [1]. Triple-negative breast cancer (TNBC) accounts for approximately 10–20% of all breast cancer cases [2, 3]. Chemotherapy is the mainstay of treatment for TNBC because of the lack of expression of oestrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) in patients [4]. Patients with TNBC have a worse prognosis than those with other types of breast cancer and have a mortality rate of 40% within the first 5 years of diagnosis [5].

Moreover, approximately 50% of patients with TNBC develop distant metastasis [6]. The mortality rate of patients with distant metastases is higher than that of patients with carcinoma in situ [7]. The lungs are one of the most common sites of distant metastasis, accounting for 40% of the cases of metastasis. The median survival time of patients with metastatic TNBC is 1–1.5 years [8]. Therefore, determining a new method to predict the risk of lung metastasis (LM) and the prognosis of patients with TNBC is extremely important. ENY2, KCNK9, TNFRSF11B, KXNMB2, race,

Table 1: Clinical and pathological features of patients with TNBC.

	Training cohort (18936)	Validation cohort (8112)		*P* value
Age, years			4.953	0.084
≤40	2004 (10.6%)	933 (11.5%)		
41-60	9033 (47.7%)	3833 (47.3%)		
>60	7899 (41.7%)	3346 (41.2%)		
Tumour size, cm			0.033	0.983
≤5	16573 (87.5%)	7099 (87.5%)		
5.1-10	2005 (10.6%)	862 (10.6%)		
>10	358 (1.9%)	151 (1.9%)		
Race			0.293	0.864
White	13570 (71.7%)	5832 (71.9%)		
Black	3913 (20.7%)	1653 (20.4%)		
Other	1453 (7.7%)	627 (7.7%)		
Grade			8.162	0.043
I	358 (1.9%)	179 (2.2%)		
II	3249 (17.2%)	1321 (16.3%)		
III	15166 (80.1%)	6557 (80.8%)		
IV	163 (0.9%)	55 (0.7%)		
Histological type			<0.001	0.999
8500	16268 (85.9%)	6969 (85.9%)		
Other	2668 (14.1%)	1143 (14.1%)		
T stage			0.610	0.894
T1	7976 (42.1%)	3394 (41.8%)		
T2	8133 (42.9%)	3485 (43.0%)		
T3	1735 (9.2%)	747 (9.2%)		
T4	1092 (5.8%)	486 (6.0%)		
N stage			0.132	0.716
N0	12153 (64.2%)	5225 (64.4%)		
Nx	6783 (35.8%)	2887 (35.6%)		
Lung metastasis			1.635	0.201
No	18590 (98.2%)	7945 (97.9%)		
Yes	346 (1.8%)	167 (2.1%)		
Insurance			0.513	0.474
No	435 (2.3%)	198 (2.4%)		
Yes	18501 (97.7%)	7914 (97.6%)		
Marital status			0.274	0.601
No	8289 (43.8%)	3523 (43.4%)		
Yes	10647 (56.2%)	4589 (56.6%)		

and marital status have been identified as risk factors and prognostic variables of LM [9, 10]. To the best of our knowledge, no in-depth studies performed thus far have used predictive models to determine the incidence and prognosis of TNBC with LM; therefore, risk factors cannot be combined to effectively assess individual outcomes, and implementation of precision medicine is thus hampered.

A nomogram is a convenient tool that can accurately predict individual outcomes and exhibit good accuracy in assessing the prognosis of various cancers [11]. In this study, we aimed to construct two nomograms to predict the risk of LM in patients with TNBC and the prognosis of these patients based on the data from the Surveillance, Epidemiology, and End Results (SEER) database.

2. Methods

2.1. Study Population Selection. The SEER∗Stat software (version 8.3.6) was used to download patient data from the SEER database. Patients diagnosed with TNBC from 2010

Table 2: Univariate and multivariate logistic analyses of lung metastasis in patients with TNBC.

	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*P* value	OR (95% CI)	*P* value
Age, years				
≤40				
41-60	1.061 (0.699-1.611)	0.782	1.295 (0.846-1.983)	0.233
>60	1.805 (1.202-2.709)	0.004	2.372 (1.561-3.602)	<0.001
Tumour size, cm				
≤5				
5.1-10	6.460 (5.091-8.195)	<0.001	1.220 (0.804-1.852)	0.349
>10	18.717 (13.567-25.821)	<0.001	2.411 (1.548-3.756)	<0.001
Race				
White				
Black	1.109 (0.722-1.704)	0.637		
Other	0.959 (0.737-1.249)	0.757		
Grade				
I				
II	1.411 (0.507-3.928)	0.509		
III	1.713 (0.635-4.621)	0.288		
IV	1.659 (0.367-7.501)	0.511		
Histological type				
8500				
Other	1.366 (1.036-1.801)	0.027		
T stage				
T1				
T2	3.806 (2.470-5.867)	<0.001	3.318 (2.138-5.149)	<0.001
T3	15.752 (10.120-24.518)	<0.001	9.164 (5.005-16.780)	<0.001
T4	43.134 (28.197-65.983)	<0.001	18.977 (11.179-32.217)	<0.001
N stage				
N0				
Nx	5.178 (4.068-6.590)	<0.001	2.182 (1.671-2.848)	<0.001
Lung metastasis				
No				
Yes	0.235 (0.073-0.760)	0.016		
Insurance				
No				
Yes	0.418 (0.258-0.679)	<0.001		
Marital status				
No				
Yes	0.562 (0.453-0.697)	<0.001		

to 2015 were included in this study. The exclusion criteria were as follows: (1) patients in whom TNBC was not the primary tumour; (2) death of patients with an unknown cause; and (3) patients with unknown information, including age, tumour size, race, grade, histological type, T stage, N stage, LM, insurance status, and marital status. Eventually, 27,048 patients in the cohort were enrolled to examine the risk factors of TNBC with LM and establish a predictive nomogram. Subsequently, patients with TNBC with LM who survived ≥1 month; underwent surgery, radiotherapy, and chemotherapy; and had specific metastasis data, including bone, brain, and liver metastases, were included to form a new cohort to identify the prognostic factors of TNBC with LM and establish a prognostic nomogram. Eventually, we included 480 patients to investigate the prognostic factors of TNBC with LM; these patients were randomly divided into the training and validation cohort in a ratio of 7 : 3 (caret package (version: 6.0.88) of the R studio). The training

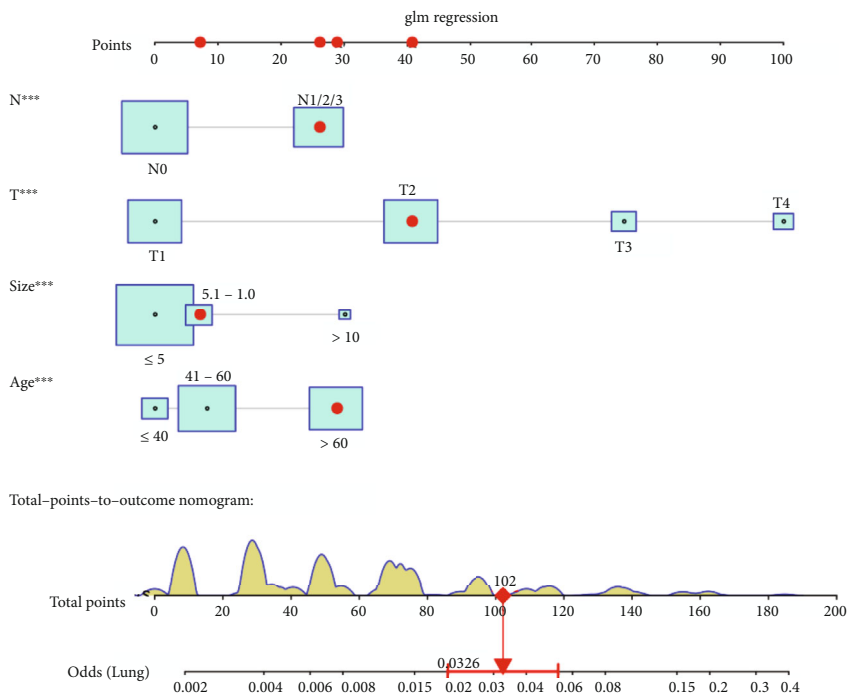


Figure 1: Diagnostic nomogram of LM for patients with TNBC. LM: lung metastasis; TNBC: triple-negative breast cancer.

cohort was used to develop a nomogram, which was externally verified in the verification cohort.

2.2. Data Collection. We used demographic variables, including age, race, insurance status, marital status, tumour characteristics, tumour size, grade, histological type, T stage, and N stage, to identify the risk factors of TNBC with LM. Additionally, we used data pertaining to metastasis to the bone, brain, and liver and treatment modalities, including surgery, chemotherapy, and radiotherapy, to determine the prognostic factors associated with TNBC with LM.

2.3. Statistical Analysis. All statistical analyses were performed using SPSS 25.0 and R software (version 3.6.1). A P value < 0.05 (bilateral) was considered statistically significant. Univariate logistic analysis and multivariate binary logistic regression analysis were performed to determine the independent risk factors of TNBC with LM. Univariate and multivariate Cox regression analyses were performed to identify the independent prognostic factors.

The receiver operating characteristic (ROC) curve and time-dependent ROC curve of the predicted nomogram were generated. The area under the ROC curve (AUC) signified the distinctiveness of the nomogram and was further compared with the AUC of all independent prognostic factors. In addition, calibration curves were established to compare the consistency between the actual results and those predicted by the line graph. The range of threshold probability and the size of benefits were determined using decision curve analysis (DCA). In addition, patients were divided into the high- and low-risk groups based on the median risk score. Kaplan–Meier (KM) curves were generated, and the logarithmic rank test was performed.

3. Results

3.1. Baseline Characteristics of Patients. The baseline characteristics of 27,048 patients with TNBC are shown in Table 1. The tumour size of most patients with TNBC was <5 cm, and of all patients, 19,402 (71.7%), were Caucasian. Most patients (80.3%) had stage III disease.

3.2. Risk Factors of TNBC with LM. To determine the LM-related variables of TNBC, we used single-factor logistic analysis to screen for risk factors and found that the age, tumour size, histological type, T stage, N stage, insurance status, and marital status of patients with TNBC were related to LM (Table 2). Furthermore, multivariate logistic analysis showed that age, tumour size, T stage, and N stage were independent predictors of LM in patients with TNBC (Table 2).

3.3. Construction and Validation of a Diagnostic Nomogram for TNBC with LM. The diagnostic nomogram of LM for patients with TNBC was constructed by including the corresponding independent risk factors (Figure 1). The AUC values of the training and verification cohort were 0.838 (95% confidence interval 0.817-0.860) and 0.894 (95% confidence interval 0.875-0.917), respectively (Figures 2(a) and 2(d)). Additionally, we generated ROC curves for each independent predictor (Figure 3) and found that the AUC of the nomogram was higher than that of all individual

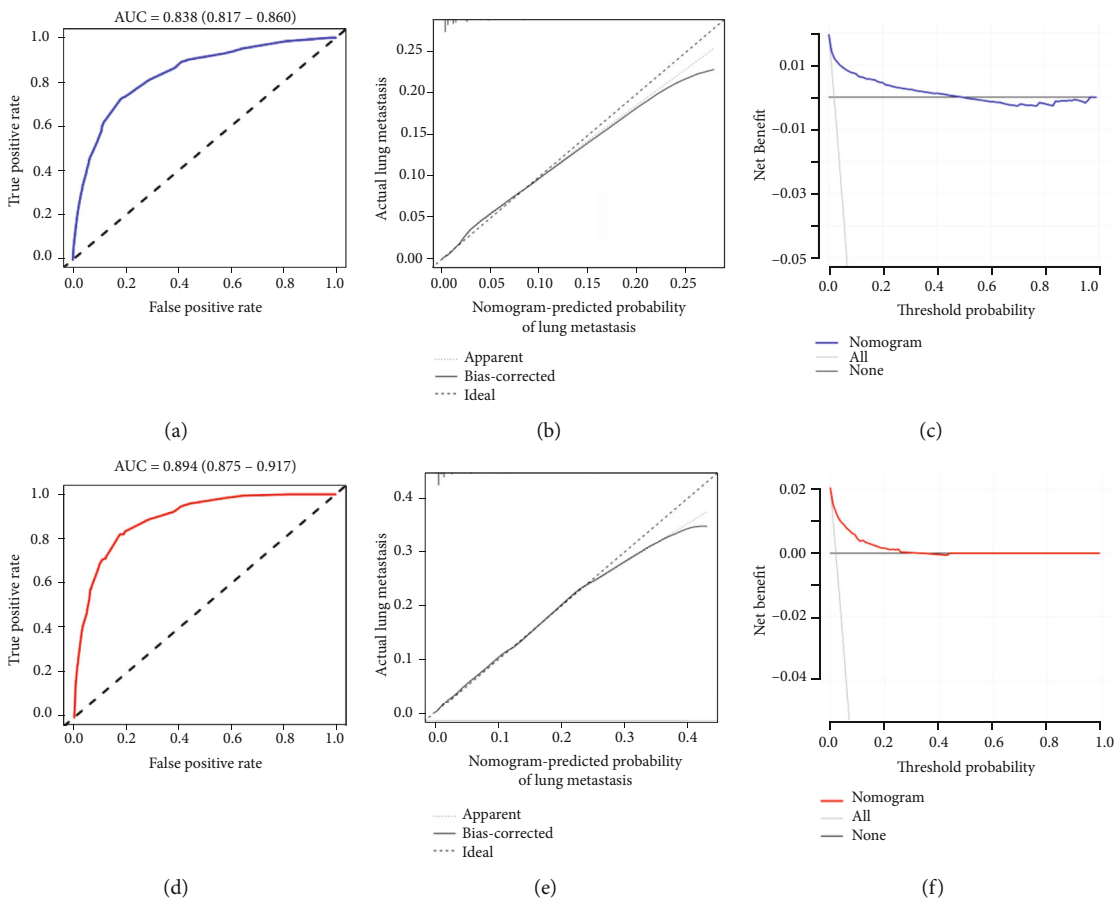


Figure 2: ROC curves (a), calibration curve (b), and DCA curve (c) of the diagnostic nomogram for patients with TNBC with LM in the training cohort. ROC curves (d), calibration curve (e), and DCA curve (f) of the diagnostic nomogram for patients with TNBC with LM in the validation cohort. ROC: receiver operating characteristic; DCA: decision curve analysis; LM: lung metastasis; TNBC: triple-negative breast cancer.

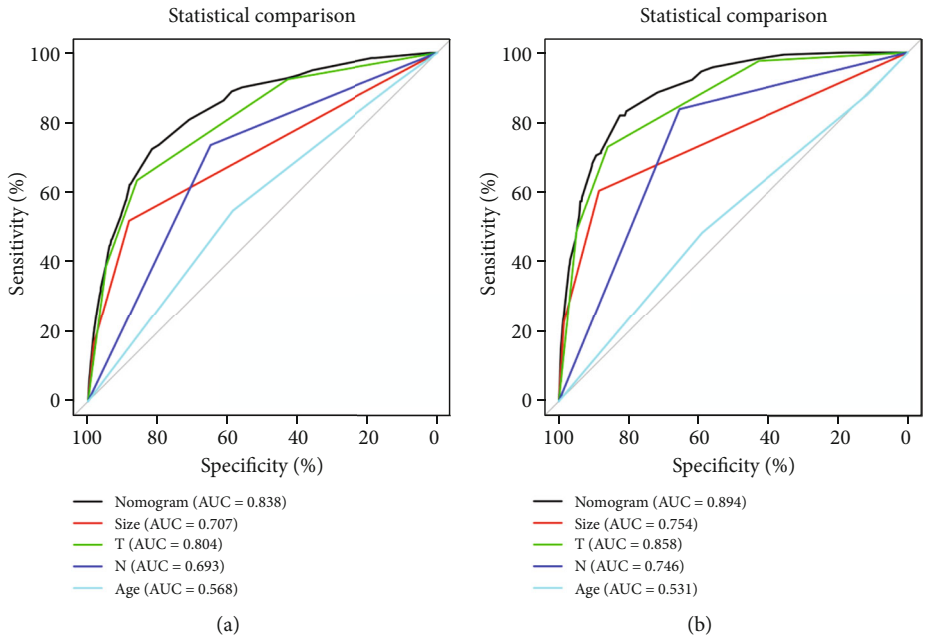


Figure 3: Independent predictor of ROC curves in the training cohort (a) and validation cohort (b). ROC: receiver operating characteristic.

Table 3: Clinical and pathological features of patients with TNBC with lung metastasis.

	Training cohort (336)	Validation cohort (144)
Age, years		
≤40	25	19
41-60	118	59
>60	193	66
Tumour size, cm		
≤5	149	69
5.1-10	126	49
>10	61	26
Race		
White	230	94
Black	84	36
Other	22	14
Grade		
I	2	3
II	38	19
III	291	118
IV	5	4
Histological type		
8500	271	120
Other	65	24
T stage		
T1	18	10
T2	91	39
T3	79	35
T4	148	60
N stage		
N0	71	36
N1	164	70
N2	28	15
N3	73	23
Surgery		
No	190	76
Yes	146	68
Radiotherapy		
No	4	2
Yes	332	142
Chemotherapy		
No	91	30
Yes	245	114
Bone metastasis		
No	242	89
Yes	94	55
Brain metastasis		
No	291	126
Yes	45	18
Liver metastasis		
No	240	110
Yes	96	34
Insurance		
No	17	9
Yes	319	135
Marital status		
No	198	84
Yes	138	60

predictors. In addition, the calibration curve showed consistent results in the training and validation cohorts (Figures 2(b) and 2(e)). The DCA curve showed that the nomogram had high accuracy for the diagnosis of TNBC with LM (Figures 2(c) and 2(f)).

3.4. Prognostic Factors of TNBC with LM. To determine the prognostic factors, we examined the data from 480 patients with TNBC with LM (Table 3). Of these patients, 324 (67.5%) were Caucasian, 120 (25.0%) were Black, and 36 (7.5%) belonged to other races. Most patients received radiotherapy. Univariate and multivariate Cox proportional hazards regression were used to identify histological type, prior surgery, chemotherapy, bone metastasis, brain metastasis, liver metastasis, and marital status as independent prognostic factors of TNBC with LM (Table 4).

3.5. Construction and Validation of a Prognostic Nomogram for TNBC with LM. By integrating the identified independent prognostic factors, a prognostic nomogram was established for TNBC with LM (Figure 4). The AUC values for predicting prognosis at 6, 12, and 18 months were 0.809 (95% confidence interval 0.771-0.868), 0.779 (95% confidence interval 0.737-0.834), and 0.735 (95% confidence interval 0.699-0.811), respectively, in the training cohort and 0.735 (95% confidence interval 0.642-0.820), 0.672 (95% confidence interval 0.575-0.758), and 0.705 (95% confidence interval 0.598-0.782), respectively, in the verification cohort (Figures 5(a) and 5(c)). In addition, in the training and validation cohorts, the probability calibration curves for 6, 12, and 18 months showed good agreement (Figures 6(a) and 6(c)). The DCA curve showed that the predictive performance of the nomogram was relatively accurate (Figures 6(b) and 6(d)).

3.6. Comparison of Discrimination between Prognostic Nomogram and Independent Prognostic Factors. To assess the advantages of the prognostic nomogram, we generated ROC curves inclusive of independent prognostic factors and found that the AUC value of part prognostic factor was >0.650, which signified that part individual factors can

Table 4: Univariate and multivariate Cox analyses of lung metastasis in patients with TNBC.

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Age, years				
≤40				
41-60	1.382 (0.864-2.211)	0.177		
>60	1.532 (0.972-2.414)	0.066		
Tumour size, cm				
≤5				
5.1-10	1.191 (0.922-1.539)	0.180		
>10	1.061 (0.768-1.466)	0.719		
Race				
White				
Black	0.888 (0.529-1.491)	0.654		
Other	0.907 (0.690-1.191)	0.481		
Grade				
I				
II	1.146 (0.275-4.771)	0.851		
III	0.673 (0.167-2.711)	0.578		
IV	0.943 (0.172-5.158)	0.946		
Histological type				
8500				
Other	1.407 (1.054-1.878)	0.021	1.424 (1.061-1.910)	0.019
T stage				
T1				
T2	1.359 (0.783-2.361)	0.276		
T3	1.412 (0.808-2.470)	0.226		
T4	1.335 (0.781-2.284)	0.291		
N stage				
N0				
Nx	1.153 (0.864-1.539)	0.333		
Surgery				
No				
Yes	0.553 (0.436-0.701)	<0.001	0.619 (0.485-0.790)	<0.001
Radiotherapy				
No				
Yes	0.830 (0.309-2.231)	0.712		
Chemotherapy				
No				
Yes	0.399 (0.308-0.517)	<0.001	0.360 (0.273-0.474)	<0.001
Bone metastasis				
No				
Yes	1.487 (1.152-1.919)	0.002	1.386 (1.060-1.812)	0.017
Brain metastasis				
No				
Yes	1.970 (1.414-2.747)	<0.001	1.810 (1.281-2.558)	<0.001
Liver metastasis				
No				
Yes	1.833 (1.424-2.358)	<0.001	2.050 (1.576-2.667)	<0.001

Table 4: Continued.

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P* value	HR (95% CI)	*P* value
Insurance				
No				
Yes	0.739 (0.452-1.209)	0.228		
Marital status				
No				
Yes	0.676 (0.532-0.857)	0.001	0.665 (0.522-0.849)	0.001

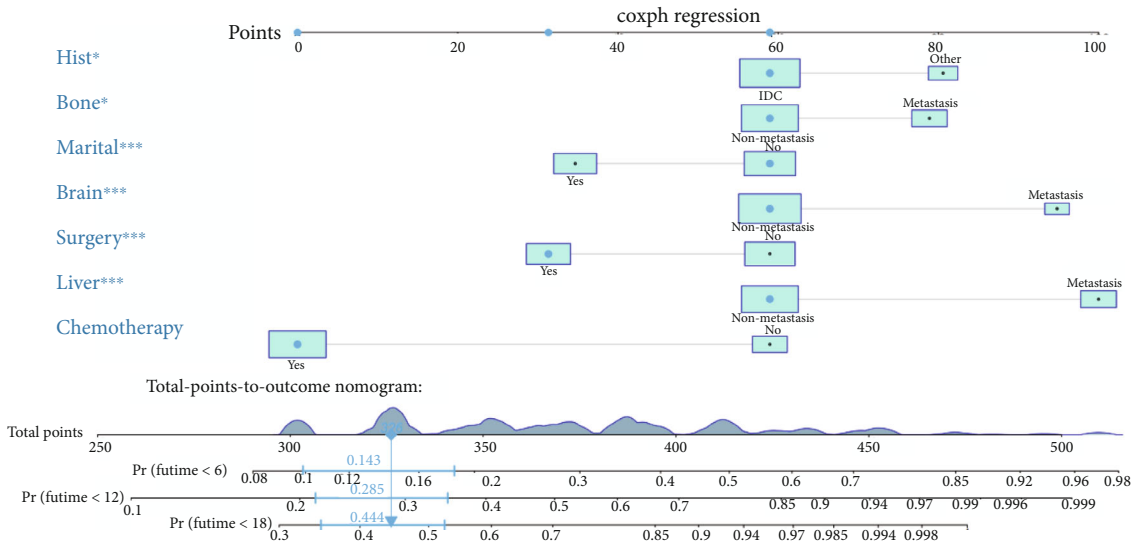


Figure 4: Prognostic nomogram for patients with TNBC with LM. LM: lung metastasis; TNBC: triple-negative breast cancer.

be used as reliable prognostic factors. However, the AUCs of all prognostic factors were lower than those of the prognostic nomogram (Figure 7). Overall, we confirmed that the function of a rosette combining different information from individual patients is superior to the predictive power of evaluating individual risk factors.

3.7. Role of Prognostic Nomograms in Risk Stratification of Patients with TNBC with LM. The overall prognostic score of all patients TNBC with LM was calculated based on the nomogram. The KM curve showed that patients in the low-risk group survived longer than those in the high-risk group (Figures 5(b) and 5(d)). The threshold values identified in the training cohort were used for the validation cohort. The prognosis in both risk groups was significantly different ($P < 0.0001$). Overall, our system of risk stratification worked very well.

4. Discussion

TNBC is an aggressive tumour, which is prone to distant metastasis [12]. The lungs are a common site for distant metastasis. In this study, we constructed a diagnostic and a prognostic nomogram to predict LM in patients with TNBC. The risk of LM can be easily identified using these nomograms. The prognostic nomogram was used to assess the prognosis of TNBC patients with LM and provide guidance for further clinical management. The two nomograms accurately assessed the risk of LM and predicted survival and may help clinicians in decision-making and disease monitoring.

Although the prognosis of patients with TNBC with LM is extremely poor, early detection of LM is essential for patients to receive appropriate treatment [13]. Therefore, identifying risk factors of TNBC with LM is very important to guide clinical treatment. Several biomarkers and prognostic factors have been identified, including linc-ZNF469-3, miR-629-3p, age, T stage, and N stage [14–17]. However, to the best of our knowledge, no risk-prediction nomograms have been constructed to date; therefore, individual risk of LM cannot be quantified. Our results showed that age, tumour size, T stage, and N stage are independent predictors of TNBC with LM, which is consistent with the results of previous studies.

In addition, our results showed that patients with invasive ductal carcinoma (IDC), who were married, with no

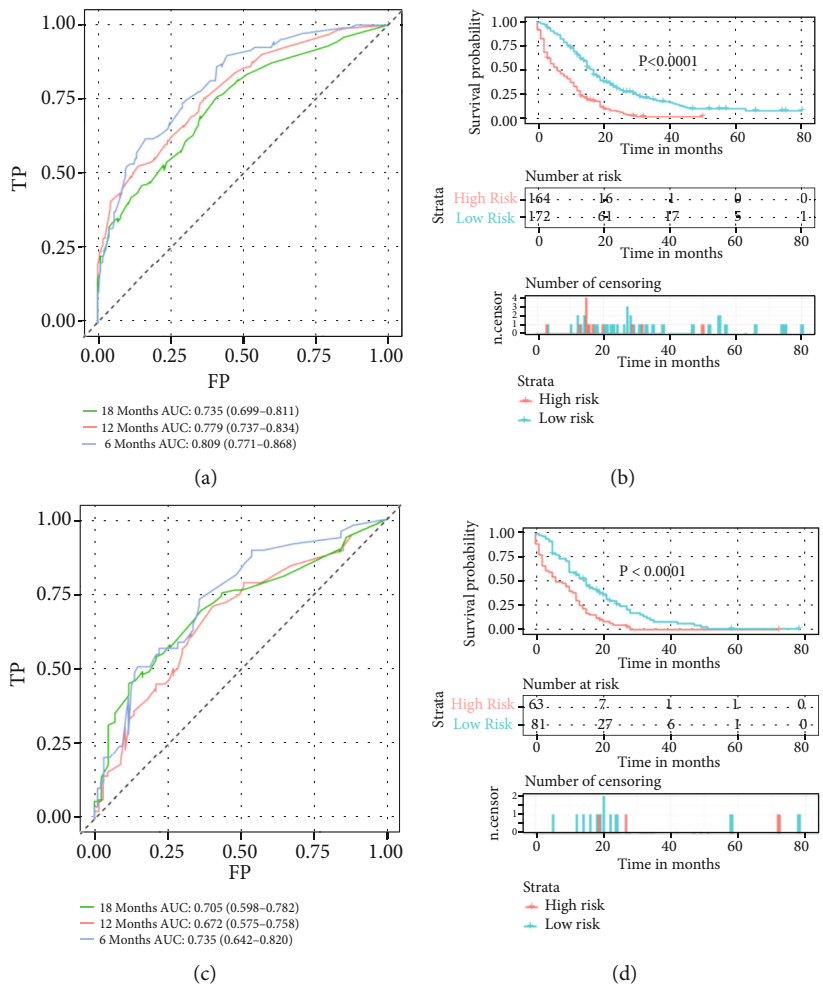


Figure 5: ROC curve of the prognostic survival model for patients with TNBC with LM in the training cohort (a) and validation cohort (c). Survival curves of the high- and low-group were generated using the prognostic total score calculated from the nomogram in the training cohort (b) and validation cohort (d). ROC: receiver operating characteristic; LM: lung metastasis; TNBC: triple-negative breast cancer.

brain metastases, no liver metastases, and no bone metastases had a better prognosis after undergoing surgery and chemotherapy. Based on seven independent prognostic factors, a prognostic nomogram was constructed, which can be used as an effective tool to identify high-risk patients. IDC is the most common histological type of patients with TNBC [18]. Zhao et al. [19] and Li et al. [20] reported that the survival rate of patients with IDC is higher than that of patients with other histological types. These findings are consistent with those reported in our study. Patients who were married had a better prognosis. Previous studies have shown that the risk of cancer metastasis and cancer-related deaths is lower in married patients than in unmarried patients [21]. This relationship may be attributed to the important role of marital status in regulating the functions of the endocrine and immune systems [22]. Moreover, TNBC has a high recurrence rate [12]. We found that patients with distant metastases had a lower survival rate, which was consistent with the results of a study by Wang et al. [23]. In addition, different sites of metastasis affect the survival of patients with TNBC. Studies have reported that the prognosis of patients with visceral metastasis is worse than that of patients with bone metastasis [24]. Typically, the treatment of patients with advanced disease should focus on improving the survival rate of these patients. Chemotherapy and surgery are both favourable prognostic factors for patients with TNBC with LM. In addition, previous studies have shown that chemotherapy and surgery can significantly improve the prognosis of patients with LM [25]. Currently, chemotherapy remains the standard treatment for patients with TNBC [26]. The NCCN guidelines recommend a combination treatment plan based on taxanes, anthracycline, cyclophosphamide, cisplatin, and fluorouracil [6]. Previous studies have shown that despite metastasis to distant organs, patients can benefit from surgery [27, 28]. Our prognostic nomogram showed that surgery and chemotherapy were beneficial for the survival of patients with TNBC with LM. Therefore, identification of independent prognostic factors may help to identify high-risk patients and establish a specific monitoring plan.

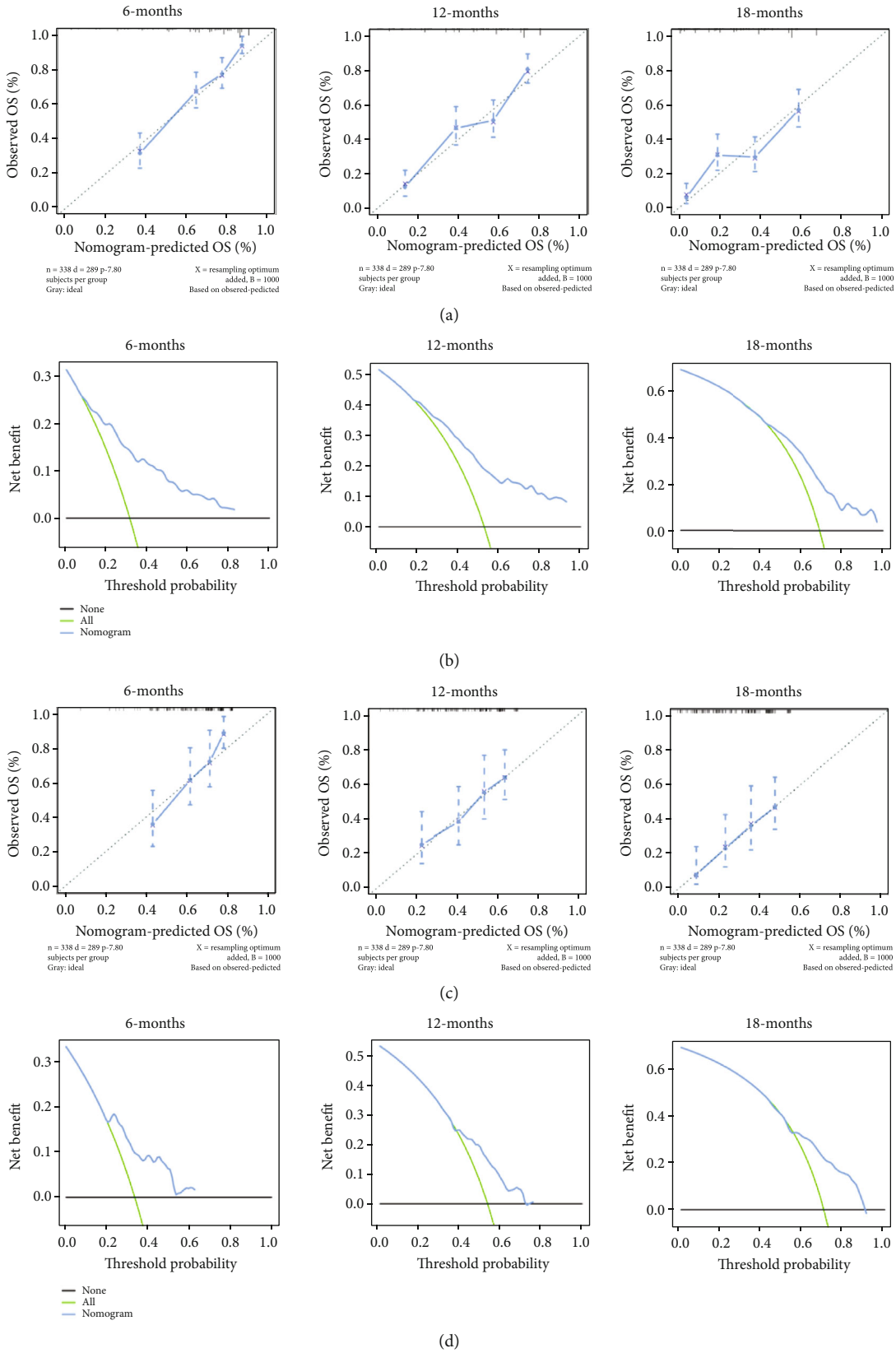


Figure 6: Calibration curves of the prognostic nomogram model for patients with TNBC with LM in the training cohort (a) and validation cohort (c) at 6, 12, and 18 months; the DCA curve of the prognostic nomogram model for patients with TNBC with LM in the training cohort (b) and validation cohort (d) at 6, 12, and 18 months. ROC: receiver operating characteristic; DCA: decision curve analysis; LM: lung metastasis; TNBC: triple-negative breast cancer.

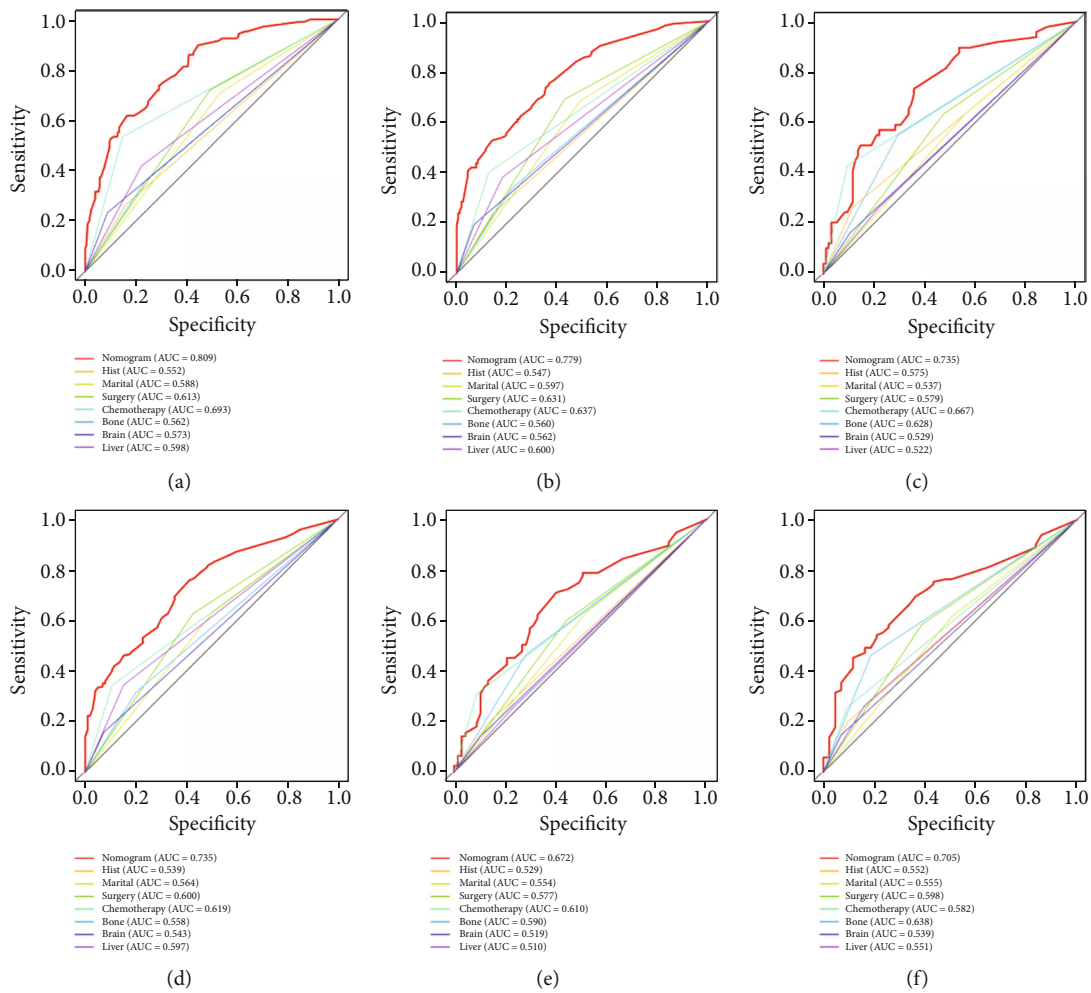


Figure 7: ROC curves of independent risk factors for the prognosis of patients with TNBC with LM in the training cohort at 6 months (a), 12 months (b), and 18 months (c). ROC curves of independent risk factors for the prognosis of patients with TNBC with LM in the validation cohort at 6 months (a), 12 months (b), and 18 months (c). ROC: receiver operating characteristic; LM: lung metastasis; TNBC: triple-negative breast cancer.

However, our study has several limitations. The SEER database contains data regarding the first diagnosis of the disease; therefore, LM that occurs in advanced stages cannot be recorded. Furthermore, this is a retrospective study with large sample size; therefore, selection bias is inevitable. The prognostic nomogram was constructed and verified at a single institution, which may affect its clinical applicability to a certain extent. Therefore, further calibration of the nomogram is required in future studies.

5. Conclusion

Age, tumour size, T stage, and N stage are the risk factors of TNBC with LM. Histological type, marital status, brain metastasis, liver metastasis, bone metastasis, surgery, and chemotherapy are independent prognostic factors. The results of this study are very useful for determining individualised treatment and ensuring appropriate management of patients with TNBC with LM.

Data Availability

The dataset from SEER database generated and/or analyzed during the current study are available in the SEER dataset repository (https://seer.cancer.gov/).

Ethical Approval

The research did not involve animal experiments and human specimens and no ethical-related issues.

Conflicts of Interest

The authors declare that they have no competing interests.

Authors' Contributions

YD conceived and directed the project. JW designed the study and analyzed the data. JW, HZ, and LY wrote the manuscript. JL, HZ, CZ, QR, YC, and YX reviewed the data.

All authors have read and approved the final manuscript for publication.

Acknowledgments

The authors thank the contributors of SEER databases for the availability of the data. This work was supported by the Xiangtan Medical Research Project (no. 2020xtyx-3).

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] R. Ismail-Khan and M. M. Bui, "A review of triple-negative breast cancer," *Cancer Control*, vol. 17, no. 3, pp. 173–176, 2010.

[3] S. J. Dawson, E. Provenzano, and C. Caldas, "Triple negative breast cancers: clinical and prognostic implications," *European Journal of Cancer*, vol. 45, Supplement 1, pp. 27–40, 2009.

[4] W. D. Foulkes, I. E. Smith, and J. S. Reis-Filho, "Triple-negative breast cancer," *The New England Journal of Medicine*, vol. 363, no. 20, pp. 1938–1948, 2010.

[5] R. Dent, M. Trudeau, K. I. Pritchard et al., "Triple-negative breast cancer: clinical features and patterns of recurrence," *Clinical Cancer Research*, vol. 13, no. 15, pp. 4429–4434, 2007.

[6] L. Yin, J.-J. Duan, X.-W. Bian, and S.-C. Yu, "Triple-negative breast cancer molecular subtyping and treatment progress," *Breast Cancer Research*, vol. 22, no. 1, p. 61, 2020.

[7] O. Gluz, C. Liedtke, N. Gottschalk, L. Pusztai, U. Nitz, and N. Harbeck, "Triple-negative breast cancer–current status and future directions," *Annals of Oncology*, vol. 20, no. 12, pp. 1913–1927, 2009.

[8] J. Zhang, Y. Lin, X. J. Sun et al., "Biomarker assessment of the CBCSG006 trial: a randomized phase III trial of cisplatin plus gemcitabine compared with paclitaxel plus gemcitabine as first-line therapy for patients with metastatic triple-negative breast cancer," *Annals of Oncology*, vol. 29, no. 8, pp. 1741–1747, 2018.

[9] G. Xie, H. Yang, D. Ma et al., "Integration of whole-genome sequencing and functional screening identifies a prognostic signature for lung metastasis in triple-negative breast cancer," *International Journal of Cancer*, vol. 145, no. 10, pp. 2850–2860, 2019.

[10] W. Xiao, S. Zheng, P. Liu et al., "Risk factors and survival outcomes in patients with breast cancer and lung metastasis: a population-based study," *Cancer Medicine*, vol. 7, no. 3, pp. 922–930, 2018.

[11] V. P. Balachandran, M. Gonen, J. J. Smith, and R. P. DeMatteo, "Nomograms in oncology: more than meets the eye," *The Lancet Oncology*, vol. 16, no. 4, pp. e173–e180, 2015.

[12] P. Boyle, "Triple-negative breast cancer: epidemiological considerations and recommendations," *Annals of Oncology*, vol. 23, Supplement 6, pp. vi7–vi12, 2012.

[13] L. M. Tseng, N. C. Hsu, S. C. Chen et al., "Distant metastasis in triple-negative breast cancer," *Neoplasma*, vol. 60, no. 3, pp. 290–294, 2013.

[14] P.-S. Wang, C.-H. Chou, C.-H. Lin et al., "A novel long noncoding RNA linc-ZNF469-3 promotes lung metastasis through miR-574-5p-ZEB1 axis in triple negative breast cancer," *Oncogene*, vol. 37, no. 34, pp. 4662–4678, 2018.

[15] J. Wang, C. Song, H. Tang et al., "miR-629-3p may serve as a novel biomarker and potential therapeutic target for lung metastases of triple-negative breast cancer," *Breast Cancer Research*, vol. 19, no. 1, p. 72, 2017.

[16] W. Xiao, S. Zheng, A. Yang et al., "Breast cancer subtypes and the risk of distant metastasis at initial diagnosis: a population-based study," *Cancer Management and Research*, vol. Volume 10, pp. 5329–5338, 2018.

[17] Y. Yao, Y. Chu, B. Xu, Q. Hu, and Q. Song, "Risk factors for distant metastasis of patients with primary triple-negative breast cancer," *Bioscience Reports*, vol. 39, no. 6, 2019.

[18] R. Konigsberg, G. Pfeiler, T. Klement et al., "Tumor characteristics and recurrence patterns in triple negative breast cancer: a comparison between younger (<65) and elderly (>/=65) patients," *European Journal of Cancer*, vol. 48, no. 16, pp. 2962–2968, 2012.

[19] S. Zhao, D. Ma, Y. Xiao, Y.-Z. Jiang, and Z.-M. Shao, "Clinicopathologic features and prognoses of different histologic types of triple-negative breast cancer: a large population-based analysis," *European Journal of Surgical Oncology*, vol. 44, no. 4, pp. 420–428, 2018.

[20] Y. Li, P. Su, Y. Wang et al., "Impact of histotypes on preferential organ-specific metastasis in triple-negative breast cancer," *Cancer Medicine*, vol. 9, no. 3, pp. 872–881, 2020.

[21] Y. Li, M.-X. Zhu, and S.-H. Qi, "Marital status and survival in patients with renal cell carcinoma," *Medicine (Baltimore)*, vol. 97, no. 16, p. e0385, 2018.

[22] R. B. Herberman and J. R. Ortaldo, "Natural killer cells: their roles in defenses against disease," *Science*, vol. 214, no. 4516, pp. 24–30, 1981.

[23] Z. Wang, H. Wang, X. Sun et al., "A risk stratification model for predicting overall survival and surgical benefit in triple-negative breast cancer patients with de novo distant metastasis," *Frontiers in Oncology*, vol. 10, p. 14, 2020.

[24] S. G. Ahn, H. M. Lee, S.-H. Cho et al., "Prognostic factors for patients with bone-only metastasis in breast cancer," *Yonsei Medical Journal*, vol. 54, no. 5, pp. 1168–1177, 2013.

[25] Y. Gu, G. Wu, X. Zou, P. Huang, and L. Yi, "Prognostic value of site-specific metastases and surgery in de novo stage IV triple-negative breast cancer: a population-based analysis," *Medical Science Monitor*, vol. 26, p. e920432, 2020.

[26] G. Bianchini, J. M. Balko, I. A. Mayer, M. E. Sanders, and L. Gianni, "Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease," *Nature Reviews. Clinical Oncology*, vol. 13, no. 11, pp. 674–690, 2016.

[27] M.-T. Chen, H.-F. Sun, Y. Zhao et al., "Comparison of patterns and prognosis among distant metastatic breast cancer patients by age groups: a SEER population-based analysis," *Scientific Reports*, vol. 7, no. 1, p. 9254, 2017.

[28] K. Wang, Y. Shi, Z.-Y. Li et al., "Metastatic pattern discriminates survival benefit of primary surgery for de novo stage IV breast cancer: a real-world observational study," *European Journal of Surgical Oncology*, vol. 45, no. 8, pp. 1364–1372, 2019.

Hindawi
Computational and Mathematical Methods in Medicine
Volume 2022, Article ID 5062798, 8 pages
https://doi.org/10.1155/2022/5062798

Research Article

Association between HBV Infection and the Prevalence of Coronary Artery Disease in the US Population

Zun-Ping Ke,[1] Miao Gong,[1] Gang Zhao,[2] Yue Geng,[1] and Kuan Cheng[2]

[1]*Department of Geriatrics, Shanghai Fifth People's Hospital, Fudan University, Shanghai, China*
[2]*Department of Cardiology, Zhongshan Hospital, Fudan University, Shanghai Institute of Cardiovascular Diseases, Shanghai, China*

Correspondence should be addressed to Kuan Cheng; chengkuanck@126.com

Received 15 June 2022; Revised 18 July 2022; Accepted 20 July 2022; Published 8 August 2022

Academic Editor: Neelam Yadav

Aims. This study aims to investigate the association between HBV infection and coronary artery disease (CAD) prevalence in the US population. A nomogram was proposed to predict CAD based on HBV infection. *Methods*. 25,749 individuals were collected from the 2001-2014 National Health and Nutrition Examination Survey. Participants with hepatitis B core antibody seropositivity were identified with HBV infection, including current and previous HBV infection status. We used adjusted logistic regression and performed sensitivity analysis to investigate the association between HBV infection and the prevalence of CAD. The effect size was evaluated by odds ratio (OR) with a 95% confidence interval (CI). Then, we created a nomogram to predict coronary artery disease. Additionally, we applied the Cox regression model to assess the association between HBV infection and all-cause mortality in those with baseline CAD. *Results*. 1790 (6.95%) individuals were with HBV infection. In the adjusted model, individuals with HBV showed a decreased CAD risk than those without (OR, 0.81; 95% CI, 0.67-0.98). Consistently, reduced risk in self-reported angina (OR, 0.72; 95% CI, 0.52-0.98) and coronary heart disease (OR, 0.76; 95% CI, 0.58-0.98) was observed in the hepatitis B core antibody seropositivity group. The subgroup analysis showed a consistent trend in the subgroups of age (<45 or ≥45), gender (male or female), hypertension (no or yes), and diabetes (no or yes). In the testing set, the proposed predictive model showed good performance with an area under the curve of 0.85 (95% CI, 0.83-0.86). There was no significant association between HBV infection and all-cause mortality in CAD patients (adjusted $P = 0.202$). *Conclusion*. Our study suggests that HBV infection was associated with lower CAD risk. The proposed nomogram showed good performance in predicting CAD. However, no significant association was observed between HBV and all-cause mortality in CAD patients.

1. Introduction

Despite the availability of prophylactic vaccines, hepatitis B virus (HBV) infection remains one of the most common infections worldwide, which constitutes a significant worldwide socioeconomic burden [1]. Serological evidence suggested that about 30% population worldwide were currently or previously infected with HBV in their lifetime [2, 3]. HBV infection could result in many hepatic or extrahepatic diseases, including acute/chronic hepatitis, liver cirrhosis, hepatocellular carcinoma, metabolic syndrome, and kidney injury[4–6].

Due to the lack of direct cytotoxicity, HBV-induced liver injury highly depends on the immune response against infected hepatocytes [7]. Lobular disarray and hepatocyte swelling were important features of acute HBV infection, whereas chronic infection was characteristic of varying lymphocytic portal inflammation [8]. In chronic HBV infection, the hepatitis B surface antigen accumulates in the endoplasmic reticulum, thus causing the so-called ground-glass hepatocyte [9].

Considering the essential role of the immune system in cardiovascular diseases [10], a few studies explored the association between HBV infection and cardiovascular diseases, and conflicting results were reported. Some studies observed that HBV infection was positively associated with coronary artery disease (CAD) [11–13], while some others reported a neural or negative association [14, 15]. Most previous

studies were based on patients with positive hepatitis B surface antigen, which suggested acute/chronic clinical hepatitis. Differently, hepatitis B core antibody remains positive in both current and previous HBV infection status. There lacks sufficient evidence on the association between hepatitis B core antibody and CAD. Also, it is unclear whether HBV infection increases all-cause death in individuals with established CAD. Further analysis of patients with hepatitis B core antibody seropositivity should provide a more in-depth insight into the role of HBV infection in the development and progression of CAD.

Therefore, this study aims to investigate the association between HBV infection and the prevalence of coronary artery disease in the US population. A nomogram was created to predict coronary artery disease based on HBV infection. Additionally, we investigated whether HBV infection could increase all-cause mortality in patients with established CAD.

2. Methods

2.1. Data Source and Study Population. The National Health and Nutrition Examination Survey (NHANES) is an open-access cross-sectional database, which collects demographic, socioeconomic, dietary, and health-related information from the US population in a 2-year cycle (https://www.cdc.gov/nchs/nhanes/index.htm). The National Death Index (NDI) is a linked centralized mortality database that collects follow-up survival information and death certificate records for participants from the NHANES survey. In this study, we downloaded seven consecutive NHANES cycles from 2001 to 2014, and the linked survival information was acquired from the NDI database.

Individuals with records on demographic information, blood pressure, cigarette/alcohol consumption, diabetes, medical conditions, standard biochemistry profiles, and hepatitis B core antibody test were enrolled in this study. Participants would be excluded if they met the following criteria (1) aged <18 or>80 years, (2) pregnant individuals, (3) missing hepatitis B core antibody test, or (4) missing survival information in the linked NDI database. Finally, a total of 25,749 participants were enrolled in this study. The National Center for Health Statistics Research Ethics Review Board approved the NHANES survey (https://www.cdc.gov/nchs/nhanes/irba98.htm). Informed consent was acquired from all the individuals.

2.2. Definition of HBV Infection. After collection, serum samples were processed, stored, and shipped to the Centers for Disease Control and Prevention. The certificated examiner would test hepatitis B core antibody by the VITROS Anti-HBc Reagent Pack and VITROS Immunodiagnostic Products Anti-HBc Calibrator on the VITROS Immunodiagnostic System. Detailed protocols were presented in an online Laboratory Procedure Manual of NHANES survey [16]. Participants with hepatitis B core antibody seropositivity were identified with HBV infection, which included both current and previous infection status.

2.3. Definition of Coronary Artery Disease. In the NHANES questionnaires, participants were asked the following questions: (1) Did a doctor or other health professional ever told that you had angina, also called angina pectoris? (2) Did a doctor or other health professional ever told that you had coronary heart disease? (3)Did a doctor or other health professional ever told that you had heart attack, also called myocardial infarction? Participants were identified with coronary artery disease if they were with self-reported angina, coronary heart disease, or myocardial infarction.

2.4. Covariates. Demographic records, including age, gender, income, and education (below high school, high school, and above high school), were collected by questionnaires. The family income-to-poverty ratio (PIR) was used to evaluate the household poverty level, calculated as family income divided by the federal poverty level. We categorized PIR into three groups (<1.33, 1.33-3.50, and ≥3.50) according to the recommendation by the Supplemental Nutrition Assistance Program [17]. Anthropometric index, cardiometabolic profiles, and health risk behaviors were collected, including body weight, height, triglycerides, total cholesterol, high-density lipoprotein cholesterol (HDL), fasting plasma glucose, hemoglobin A1c, creatinine, and cigarette/alcohol consumption. Body mass index (BMI) was calculated as weight divided by the square of height (kg/m^2). Diabetes was defined as (1) fasting plasma glucose ≥ 126 mg/dL, (2) hemoglobin A1c ≥ 6.5%, or (3) self-reported diabetes. The estimated glomerular filtration rate was calculated by the Chronic Kidney Disease-Epidemiology Collaboration equation [18]. Self-reported cardiovascular history (angina, coronary heart disease, myocardial infarction, heart failure, and stroke) was collected by questionnaires. The detailed methodology and protocols for all the examinations were presented on the NHANES website.

2.5. Statistical Analysis. We first applied multiple multivariate imputation strategies to fill the missing covariates and control the bias due to missing covariables [19, 20]. Kolmogorov-Smirnov test was used to evaluate the data distribution. If normally distributed, continuous variables were presented as mean ± standard deviation and compared by the one-way ANOVA test. Otherwise, continuous variables were presented as median with Q1-Q3 and compared by the Kruskal-Wallis test. Categorical variables were presented as percentages and compared by the Chi-square test.

We used logistic regression to investigate the association between HBV infection and the prevalence of coronary artery disease. The association was evaluated by odds ratio (OR) with a 95% confidence interval (CI). We adjusted for no covariate in the crude model, whereas the following variables were adjusted for in the adjusted model: age, gender, PIR level, BMI, total-to-HDL cholesterol, diabetes, hypertension, smoking, and alcohol consumption. All these covariates were well-established risk factors for CVDs [17]. Then, we performed sensitivity analyses to evaluate the consistency of the association in different subgroups, including age (<45 or ≥45), gender (male or female), hypertension (no or yes), and diabetes (no or yes). Additionally, we created a

Table 1: Characteristics of the enrolled participants by hepatitis B core antibody status.

	Seropositivity	Seronegative	*P*
N	1790	23959	
Age (years)	55.00 (44.00; 65.00)	45.00 (31.00; 60.00)	<0.001
Male, *n* (%)	1000 (55.87%)	11441 (47.75%)	<0.001
PIR class, *n* (%)			<0.001
<1.33	628 (35.08%)	7417 (30.96%)	
1.33 - 3.5	607 (33.91%)	7747 (32.33%)	
≥ 3.5	555 (31.01%)	8795 (36.71%)	
Education, *n* (%)			<0.001
Below high school	587 (32.79%)	6150 (25.67%)	
High school	422 (23.58%)	5467 (22.82%)	
Above high school	781 (43.63%)	12342 (51.51%)	
Hypertension, *n* (%)	720 (40.22%)	7546 (31.50%)	<0.001
Diabetes, *n* (%)	384 (21.45%)	3586 (14.97%)	<0.001
Triglycerides (mg/dL)	114.00 (77.00; 176.00)	116.00 (78.00; 179.00)	0.541
Total-to-HDL cholesterol	3.70 (2.95; 4.70)	3.69 (2.97; 4.63)	0.786
eGFR (mL/min/1.73m^2)	94.78 (74.81; 118.85)	111.80 (87.80; 141.44)	<0.001
Coronary artery disease, *n* (%)	139 (7.77%)	1473 (6.15%)	0.007
Angina, *n* (%)	45 (2.51%)	574 (2.40%)	0.814
Coronary heart disease, *n* (%)	71 (3.97%)	793 (3.31%)	0.156
Myocardial infarction, *n* (%)	89 (4.97%)	837 (3.49%)	0.001
Stroke, *n* (%)	77 (4.30%)	689 (2.88%)	0.001
HF, *n* (%)	62 (3.46%)	604 (2.52%)	0.019

PIR: income-to-poverty ratio; eGFR: estimated glomerular filtration rate; HDL: high-density lipoprotein.

nomogram to predict coronary artery disease based on the logistic regression model. Each factor (HBV infection and other covariates) was assigned a preliminary score ranging from 0 to 100, and all variables' scores were added to generate a total score. The total score was then converted to estimate the probability of the CAD. All participants were randomly divided into a training set or testing set at a ratio of 7 : 3. The training set was used to create the model, whereas the testing set was used to evaluate its performance. We applied area under the curve (AUC), sensitivity, and specificity to access the model performance in the testing set.

Moreover, the crude and adjusted cox regression analysis was used to evaluate the association between HBV infection and all-cause mortality in those with baseline coronary artery disease. Hazard ratio (HR) with 95% CI was accordingly calculated. In the adjusted model, we adjusted for age, gender, PIR level, BMI, total-to-HDL cholesterol, diabetes, hypertension, smoking, and alcohol consumption. All statistical analyses were performed by R software (version 4.1). *P* value < 0.05 was considered statistically significant.

3. Results

3.1. Characteristics of the Enrolled Participants. We show the demographic characteristics, cardiometabolic risk factors, and the prevalence of cardiovascular diseases by hepatitis B core antibody status in Table 1. The median age was 46 years, and 48.32% were males. 1790 (6.95%) individuals were currently or previously infected with HBV, and 864 (3.36%) individuals were diagnosed with CAD. After a median follow-up of 5.29 years, 264 death events were observed in patients with established CAD. In the seropositivity group, individuals were significantly older, and more were males. Also, hypertension, diabetes, CAD, myocardial infarction, stroke, and heart failure were more observed in the seropositivity group (all $P < 0.05$). There was no significant difference in total-to-HDL cholesterol and angina.

3.2. The Association between HBV Infection and the Prevalence of Ischemic Heart Disease. We summarized the association between HBV infection and CAD prevalence in Table 2. A significant association was observed in the crude model with an OR of 1.29 (95% CI: 1.07-1.53; $P < 0.01$). When we adjusted for age, gender, PIR level, BMI, total-to-HDL cholesterol, diabetes, hypertension, smoking, and alcohol consumption, individuals with HBV infection showed a lower risk of CAD than those without (OR, 0.81; 95% CI, 0.67-0.98). Consistently, reduced risk in self-reported angina (OR, 0.72; 95% CI, 0.52-0.98) and coronary heart disease (OR, 0.76; 95% CI, 0.58-0.98) was also observed. However, no significant association was observed in myocardial infarction (OR, 0.91; 95% CI, 0.72-1.15). In the sensitivity analysis (Figure 1), the negative trend remained in the subgroups of age (<45 or ≥45), gender (male or female), hypertension (no or yes), and diabetes (no or yes).

Table 2: Association between HBV infection and the prevalence of coronary artery disease.

	Crude model		Adjusted model	
	Odds ratio	*P* value	Odds ratio	*P* value
Coronary artery disease	1.29 (1.07-1.53)	<0.001	0.81 (0.67-0.98)	<0.001
Angina	1.05 (0.76-1.41)	0.753	0.72 (0.52-0.98)	0.041
Coronary heart disease	1.21 (0.93-1.53)	0.137	0.76 (0.58-0.98)	0.036
Myocardial infarction	1.45 (1.15-1.80)	0.001	0.91 (0.72-1.15)	0.450

Crude model: non-adjusted model. Adjusted model: age, gender, poverty-income ratio level, body mass index, total-to-high-density lipoprotein cholesterol, diabetes, hypertension, smoking, and alcohol consumption.

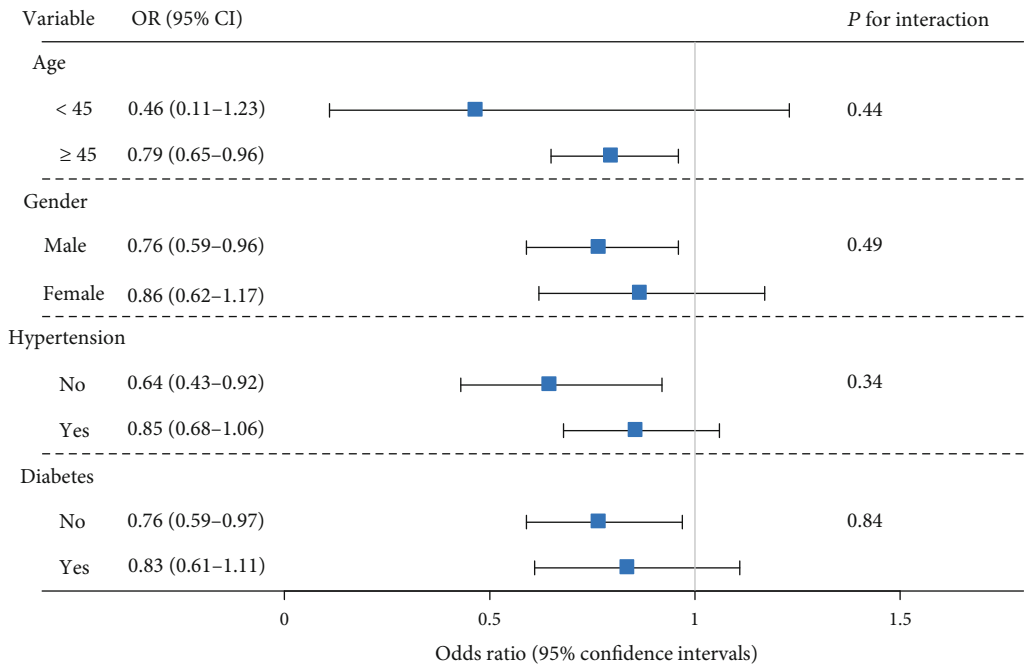


Figure 1: Subgroup analysis for the association between HBV infection and the prevalence of CAD by logistic regression. The association was adjusted for age, gender, PIR level, total-to-HDL cholesterol, diabetes, hypertension, smoking, and alcohol consumption. When we performed the analysis in the subgroup, the specific variable would be removed from the adjusted covariables. The seronegative group was set as a reference.

3.3. The Construction and Validation of the Nomogram. Figure 2 shows the nomogram for CAD based on HBV infection, age, BMI, gender, diabetes, hypertension, smoking, alcohol consumption, PIR level, and total-to-HDL cholesterol. In the testing set, the proposed predictive model showed good performance with an AUC of 0.85 (95% CI, 0.83-0.86), sensitivity of 0.85, and specificity of 0.71. Figure 3 shows the receiver operating curve of the nomogram.

3.4. The Association between HBV Infection and All-Cause Mortality in Patients with Established CAD. We created a Kaplan-Meier curve of all-cause mortality for patients with established CAD. Figure 4 shows the survival outcomes in patients with or without HBV infection. No significant difference was observed between the two groups ($P = 0.42$). Still, we performed Cox regression to investigate the association between HBV infection and all-cause mortality in patients with established CAD (Table 3). The crude Cox regression model showed no significant association between HBV infection and all-cause mortality ($P = 0.41$). When we adjusted for multiple covariates, a similar trend remained with an OR of 0.73 (95% CI, 0.45-1.18).

4. Discussion

Complex interactions are observed between HBV infection and the immune system, and the altered immune status might play a potential role in the development of CAD [10]. However, the association between HBV infection and coronary artery disease remains controversial. Based on 25749 US individuals, our study showed that hepatitis B core antibody seropositivity was significantly associated with a reduced prevalence of CAD. This trend remained in the age, gender, hypertension, and diabetes subgroups. However, there was no significant association between hepatitis B core antibody and all-cause mortality in those with established CAD.

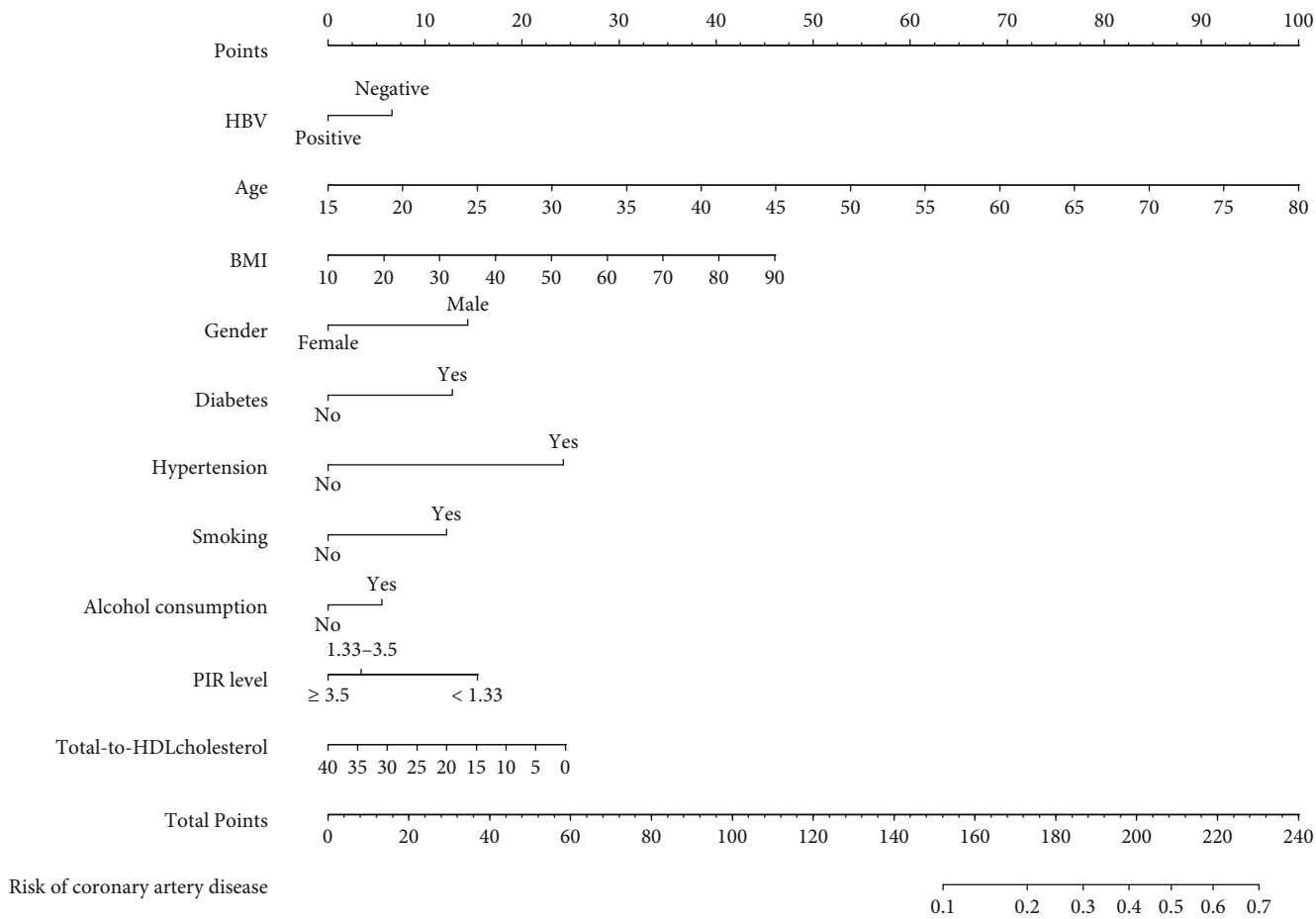


Figure 2: The nomogram to predict the prevalence of coronary artery disease. The proposed nomogram was designed to predict coronary artery disease based on the logistic regression model. Each factor (HBV infection and other covariates) was assigned a preliminary score ranging from 0 to 100, and all variables' scores were added to generate a total score. The total score was then converted to estimate the probability of coronary artery disease.

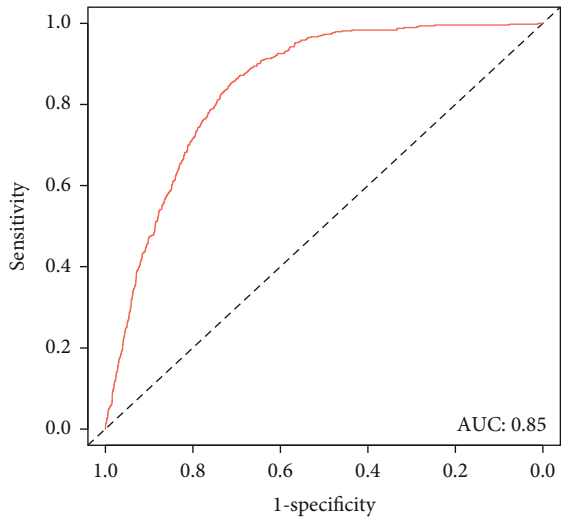


Figure 3: The receiver operating curve of the nomogram. In the testing set, the proposed predictive model showed good performance with an AUC of 0.85 (95% CI, 0.83-0.86), sensitivity of 0.85, and specificity of 0.71. AUC: area under the curve.

Interestingly, the prevalence of CAD was higher in patients with HBV infection, and the seropositivity group showed an increased risk for CAD in the non-adjusted model (OR, 1.29; 95% CI, 1.07-1.53). This observation contradicted the results from the adjusted model when considering multiple covariates to reduce bias. It should be noted that the age of the seropositivity group was about ten years higher than the negative group. The paradox may be caused by the distinct age distribution between groups since older patients are at a higher risk of multiple cardiovascular diseases. Therefore, when we adjusted for age and many other cardiovascular risk factors, HBV infection was associated with a reduced prevalence of CAD (adjusted OR, 0.81; 95% CI, 0.67-0.98).

Although our study suggested a potential association, the mechanisms underlying this observation remain vague. Our results should be interpreted cautiously since the association's immune mechanism remains uncertain. No sufficient evidence currently supports the direct influence of hepatitis B core antibody on coronary artery disease. Still, there are some possible explanations. First, acute/chronic hepatitis due to HBV impairs normal liver function, contributes to progressive fibrosis, and even causes liver cirrhosis. In

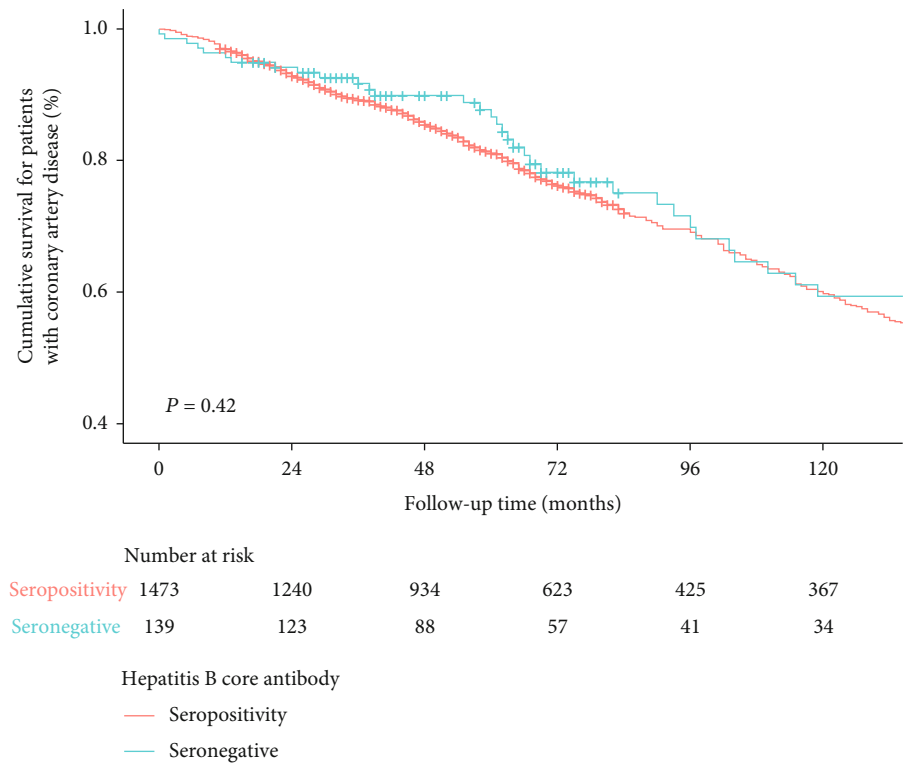


Figure 4: Kaplan-Meier curve of all-cause mortality for patients with established CAD. The survival outcomes are shown in patients with or without HBV infection.

Table 3: Association between HBV infection and all-cause mortality in patients with established coronary artery disease.

	Crude model		Adjusted model	
	Hazard ratio	*P* value	Hazard ratio	*P* value
Coronary artery disease	0.82 (0.51, 1.32)	0.412	0.73 (0.45, 1.18)	0.202

Crude model: non-adjusted model. Adjusted model: age, gender, poverty-income ratio level, body mass index, total-to-high-density lipoprotein cholesterol, diabetes, hypertension, smoking, and alcohol consumption.

contrast, HBV infection also disturbs lipid metabolism, thus decreasing the risk of non-alcoholic fatty liver disease [21, 22]. Sung et al. [12] proposed that the association between HBV infection and reduced risk of cardiovascular diseases might be attributed to liver dysfunction [12]. The impaired liver metabolic function would reduce multiple atherogenic cardiometabolic risk factors (e.g., triglyceride, cholesterol, and lipoprotein A), thus reducing the risk of atherosclerosis [23, 24]. Second, systemic inflammation, which facilitates endothelial dysfunction and arterial atherosclerosis, may play an essential role in the association between hepatitis B core antibody and CAD. HBC infection was negatively correlated with systemic inflammation evaluated by C-reactive protein, an independent risk factor for atherosclerosis [11]. A previous study indicates the possible protective effect on atherogenesis might be attributed to lower inflammation levels [11]. Additionally, the upregulated anti-atherosclerosis cytokines (e.g., hepatocyte growth factor) were also observed in patients with HBV infection [25]. The observation suggested a low systemic inflammation burden in patients with HBV infection.

Apart from CAD, the reduced lipid synthesis and lower inflammation level due to HBV infection could decrease the risk of other cardiovascular diseases. For example, patients with HBV infection were less likely to develop an acute ischemic stroke than those without. Based on the Taiwan National Health Insurance program, Tseng et al. reported that HBV infection significantly reduced the risk of acute ischemic stroke compared with the control group in a 7-year follow-up (adjusted HR, 0.77; 95% CI 0.66-0.89) [26].

Some limitations of this study should be noticed. First, the study design is cross-sectional, which is insufficient to establish causality. More prospective studies are necessary to demonstrate the causality between HBV infection and the development of CAD. Second, the diagnosis of cardiovascular disease was self-reported based on questionnaires. We are unaware of which diagnostic criteria were adopted for angina, coronary heart disease, or myocardial infarction. Third, although we adjusted for multiple covariates in the adjusted model, other factors (such as physical activity, liver function, co-infections, and antiviral treatment) may potentially induce bias in the association. Forth, the number of participants with seropositivity or seronegative was unbalanced, which might cause additional bias. Fifth, the

NHANES were based on the US population, and further studies are required to demonstrate whether these findings can be extended to different ethnicities. Also, no significant association was observed in some subgroups. More studies should be performed to understand the association better.

5. Conclusion

Our study suggested that HBV infection might be associated with lower CAD risk. The proposed nomogram showed good performance in predicting CAD. However, no significant association was observed between HBV infection and all-cause mortality in patients with established CAD. Further external validation should be performed on the nomogram in the following study.

Abbreviations

BMI: Body mass index
CAD: Coronary artery disease
CI: Confidence interval
HBV: Hepatitis B virus
HDL: High-density lipoprotein cholesterol
HR: Hazard ratio
NDI: National Death Index
OR: Odds ratio
PIR: Income-to-poverty ratio.

Data Availability

All the data were acquired from the National Health and Nutrition Examination Survey database (https://www.cdc.gov/nchs/nhanes/index.htm).

Ethical Approval

National Center for Health Statistics Research Ethics Review Board approved the NHANES survey (https://www.cdc.gov/nchs/nhanes/irba98.htm).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Zun-Ping Ke and Kuan Cheng conceived and designed the study; Zun-Ping Ke, Miao Gong, Gang Zhao, and Yue Geng analyzed the data. Zun-Ping Ke and Miao Gong wrote the paper. All authors provided critical revisions of the manuscript and approved the final manuscript. Zun-Ping Ke and Miao Gong are co-first authors.

References

[1] N. Schmit, S. Nayagam, M. R. Thursz, and T. B. Hallett, "The global burden of chronic hepatitis B virus infection: comparison of country-level prevalence estimates from four research groups," *International Journal of Epidemiology*, vol. 50, no. 2, pp. 560–569, 2021.

[2] P. A. Revill, F. V. Chisari, J. M. Block et al., "A global scientific strategy to cure hepatitis B," *The Lancet Gastroenterology & Hepatology*, vol. 4, no. 7, pp. 545–558, 2019.

[3] C. Trepo, H. L. Chan, and A. Lok, "Hepatitis B virus infection," *Lancet*, vol. 384, no. 9959, pp. 2053–2063, 2014.

[4] Y. Du, S. Zhang, M. Hu et al., "Association between hepatitis B virus infection and chronic kidney disease: a cross-sectional study from 3 million population aged 20 to 49 years in rural China," *Medicine (Baltimore)*, vol. 98, no. 5, article e14262, 2019.

[5] L. B. Yan, J. Liao, N. Han et al., "Association between hepatitis B virus infection and metabolic syndrome in Southwest China: a cross-sectional study," *Scientific Reports*, vol. 10, no. 1, p. 6738, 2020.

[6] C. Cai, J. Zeng, H. Wu et al., "Association between hepatitis B virus infection and diabetes mellitus: a meta-analysis," *Experimental and Therapeutic Medicine*, vol. 10, no. 2, pp. 693–698, 2015.

[7] K. Morikawa, T. Shimazaki, R. Takeda, T. Izumi, M. Umumura, and N. Sakamoto, "Hepatitis B: progress in understanding chronicity, the innate immune response, and cccDNA protection," *Annals of translational medicine*, vol. 4, no. 18, p. 337, 2016.

[8] H. Mani and D. E. Kleiner, "Liver biopsy findings in chronic hepatitis B," *Hepatology*, vol. 49, no. S5, pp. S61–S71, 2009.

[9] B. Rehermann and M. Nascimbeni, "Immunology of hepatitis B virus and hepatitis C virus infection," *Nature Reviews. Immunology*, vol. 5, no. 3, pp. 215–229, 2005.

[10] I. Fernandez-Ruiz, "Immune system and cardiovascular disease," *Nature Reviews. Cardiology*, vol. 13, no. 9, p. 503, 2016.

[11] D. Y. Tong, X. H. Wang, C. F. Xu, Y. Z. Yang, and S. D. Xiong, "Hepatitis B virus infection and coronary atherosclerosis: results from a population with relatively high prevalence of hepatitis B virus," *World Journal of Gastroenterology*, vol. 11, no. 9, pp. 1292–1296, 2005.

[12] J. Sung, Y. M. Song, Y. H. Choi, S. Ebrahim, and S. G. Davey, "Hepatitis B virus seropositivity and the risk of stroke and myocardial infarction," *Stroke*, vol. 38, no. 5, pp. 1436–1441, 2007.

[13] A. Kumar, M. Shariff, and R. Doshi, "Association between past hepatitis B infection and ischemic heart disease: an analysis from the 2007-2016 NHANES data," *The American Journal of the Medical Sciences*, vol. 360, no. 4, pp. 372–377, 2020.

[14] Y. Momiyama, R. Ohmori, R. Kato, H. Taniguchi, H. Nakamura, and F. Ohsuzu, "Lack of any association between persistent hepatitis B or C virus infection and coronary artery disease," *Atherosclerosis*, vol. 181, no. 1, pp. 211–213, 2005.

[15] A. Amirzadegan, G. Davoodi, M. A. Boroumand, S. Darabyan, M. R. Dehkordi, and H. Goodarzynejad, "Association between hepatitis B surface antibody seropositivity and coronary artery disease," *Indian Journal of Medical Sciences*, vol. 61, no. 12, pp. 648–655, 2007.

[16] The National Health and Nutrition Examination Survey, "*Laboratory Procedure Manual for Hepatitis B Core Antibody*," 2013, https://wwwn.cdc.gov/nchs/data/nhanes/2011-2012/labmethods/hepbd_g_met_hepatitis-b-core-antibody.pdf.

[17] S. H. Saydah, K. R. Siegel, G. Imperatore, C. Mercado, and E. W. Gregg, "The cardiometabolic risk profile of young adults with diabetes in the U.S," *Diabetes Care*, vol. 42, no. 10, pp. 1895–1902, 2019.

[18] A. S. Levey, L. A. Stevens, C. H. Schmid et al., "A new equation to estimate glomerular filtration rate," *Annals of Internal Medicine*, vol. 150, no. 9, pp. 604–612, 2009.

[19] J. A. Sterne, I. R. White, J. B. Carlin et al., "Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls," *BMJ*, vol. 338, no. jun29 1, article b2393, 2009.

[20] J. C. Jakobsen, C. Gluud, J. Wetterslev, and P. Winkel, "When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts," *BMC Medical Research Methodology*, vol. 17, no. 1, p. 162, 2017.

[21] T. C. Su, Y. T. Lee, T. J. Cheng, H. P. Chien, and J. D. Wang, "Chronic hepatitis B virus infection and dyslipidemia," *Journal of the Formosan Medical Association*, vol. 103, no. 4, pp. 286–291, 2004.

[22] H. Li, Q. Y. Xu, Y. Xie, J. J. Luo, H. X. Cao, and Q. Pan, "Effects of chronic HBV infection on lipid metabolism in non-alcoholic fatty liver disease: a lipidomic analysis," *Annals of Hepatology*, vol. 24, article 100316, 2021.

[23] V. W. Wong, G. L. Wong, W. C. Chu et al., "Hepatitis B virus infection and fatty liver in the general population," *Journal of Hepatology*, vol. 56, no. 3, pp. 533–540, 2012.

[24] E. J. Joo, Y. Chang, J. S. Yeom, and S. Ryu, "Hepatitis B virus infection and decreased risk of nonalcoholic fatty liver disease: a cohort study," *Hepatology*, vol. 65, no. 3, pp. 828–835, 2017.

[25] Y. Wang, J. Xiong, M. Niu, W. Xu, K. Xu, and H. Zhong, "Hepatitis B virus and the risk of coronary heart disease: a comprehensive systematic review and meta-analyses of observational studies," *International Journal of Cardiology*, vol. 265, pp. 204–209, 2018.

[26] C. H. Tseng, C. H. Muo, C. Y. Hsu, and C. H. Kao, "Association of hepatitis B virus infection with decreased ischemic stroke," *Acta Neurologica Scandinavica*, vol. 134, no. 5, pp. 339–345, 2016.