Volume 7, Issue 3, 2025

Print ISSN: 2663-1954
Online ISSN: 2663-1962

JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

Journal of Pharmaceutical and Medical Research

Volume 7, Issue 3, 2025

Published by Upubscience Publisher

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954 Online ISSN: 2663-1962
Email: info@upubscience.com
Website: http://www.upubscience.com/

Table of Content

NEOADJUVANT CHEMOTHERAPY IN HER2-POSITIVE BREAST CANCER

FeiFei Shan, JunQi Sun*, Wei Zhou, ChengYuan Zheng

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3050

EXPRESSION AND PROGNOSTIC VALUE OF HUMAN EPIDERMAL GROWTH FACTOR RECEPTOR 2 (HER2) AND VASCULAR ENDOTHELIAL GROWTH FACTOR C (VEGFC) IN NON-SMALL CELL LUNG CANCER

Lu Xu
Department of Respiratory Medicine, The Affiliated Taizhou People's Hospital of Nanjing Medical University, Taizhou School of Clinical Medicine, Nanjing Medical University, Taizhou 225300, Jiangsu, China.

Abstract: Non-small cell lung cancer (NSCLC) represents the most prevalent subtype of lung cancer, exhibiting high mortality rates globally. Identification of robust prognostic biomarkers is critical for early diagnosis, risk stratification, and the development of targeted therapies. Human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor C (VEGFC) have been implicated in tumor proliferation, angiogenesis, and metastasis. However, their expression patterns and prognostic significance in NSCLC remain incompletely characterized. In this study, we performed a comprehensive bioinformatics analysis using publicly available datasets, including The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), Oncomine, and Human Protein Atlas (HPA). Differential expression, clinical correlation, survival, immune infiltration, functional enrichment, and protein–protein interaction (PPI) analyses were conducted. Our results demonstrated that HER2 and VEGFC were significantly upregulated in NSCLC tissues compared to normal controls. Elevated expression of either gene was associated with advanced tumor stage, lymph node metastasis, and poor overall survival. Co-expression analysis indicated a synergistic negative impact on prognosis. Functional enrichment suggested involvement of HER2 in PI3K-Akt signaling and cell proliferation, while VEGFC was associated with lymphangiogenesis and vascular development. Immune infiltration analysis revealed significant correlations with CD8+ T cells, regulatory T cells, and macrophage populations. These findings suggest that HER2 and VEGFC serve as prognostic biomarkers and potential therapeutic targets in NSCLC, warranting further validation in clinical studies.

Keywords: Non-small cell lung cancer; Human epidermal growth factor receptor 2; Vascular endothelial growth factor C

1 INTRODUCTION

Lung cancer remains the leading cause of cancer-related mortality worldwide, accounting for over 1.8 million deaths annually. Non-small cell lung cancer (NSCLC) constitutes approximately 85% of all lung cancer cases and is characterized by considerable histological heterogeneity and variable clinical outcomes. Despite advances in surgery, chemotherapy, targeted therapy, and immunotherapy, the 5-year survival rate for NSCLC remains below 20%, largely due to late-stage diagnosis and tumor metastasis. Identification of reliable biomarkers for early detection, prognosis, and therapeutic guidance is therefore crucial[1-3]. Human epidermal growth factor receptor 2 (HER2), encoded by the ERBB2 gene, is a member of the ErbB family of receptor tyrosine kinases. HER2 plays a pivotal role in regulating cell proliferation, survival, and differentiation. In breast and gastric cancers, HER2 amplification and overexpression have been established as critical prognostic markers and therapeutic targets, guiding the use of HER2-targeted agents. In NSCLC, HER2 alterations, including gene amplification, mutation, and protein overexpression, have been reported at variable frequencies, ranging from 2% to 30%, depending on detection methods and patient cohorts. Accumulating evidence indicates that HER2 may contribute to tumor progression, metastasis, and resistance to conventional therapies in NSCLC, but its prognostic value requires further elucidation[4-5]. Vascular endothelial growth factor C (VEGFC) is a key mediator of lymphangiogenesis and angiogenesis. By binding to VEGFR-3 on lymphatic endothelial cells, VEGFC promotes the formation of new lymphatic vessels, facilitating tumor metastasis via the lymphatic system. Elevated VEGFC expression has been associated with tumor aggressiveness and poor prognosis in multiple malignancies, including breast, gastric, colorectal, and lung cancers. In NSCLC, VEGFC overexpression has been linked to lymph node metastasis and disease progression, highlighting its potential utility as a prognostic biomarker[6-7]. Bioinformatics analyses using high-throughput transcriptomic data provide a powerful approach to identify and validate cancer biomarkers. Databases such as TCGA and GEO offer extensive gene expression profiles along with clinical annotations, enabling systematic evaluation of gene expression patterns, prognostic significance, and potential molecular mechanisms. In addition, tools like TIMER and CIBERSORT facilitate exploration of tumor immune microenvironment interactions, while STRING and Cytoscape allow the construction of protein–protein interaction (PPI) networks to elucidate molecular crosstalk[8-11]. This study aimed to comprehensively analyze the expression and prognostic value of HER2 and VEGFC in NSCLC using integrated bioinformatics approaches. Specifically, we sought to (1) evaluate differential expression in tumor versus normal tissues, (2) assess associations with clinical parameters and survival outcomes, (3) explore immune infiltration patterns related to gene expression, and (4) investigate potential functional roles through gene enrichment and PPI network analyses. By systematically characterizing HER2 and VEGFC, we aim to provide insights into their clinical relevance and potential as therapeutic targets in NSCLC[12-13].

2 MATERIALS AND METHODS

2.1 Data Sources

RNA sequencing data along with corresponding clinical annotations for two major subtypes of non-small cell lung cancer (NSCLC), namely lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), were systematically obtained from The Cancer Genome Atlas (TCGA) database. These datasets provided comprehensive transcriptomic profiles and detailed patient clinical information, including survival outcomes, tumor stage, and demographic variables, which were essential for downstream bioinformatics analyses. To further validate the findings derived from TCGA, independent gene expression datasets were retrieved from the Gene Expression Omnibus (GEO), specifically GSE31210 and GSE50081, which contain well-characterized cohorts of NSCLC patients with long-term follow-up information. In addition to transcriptomic validation, the protein-level expression of key biomarkers, including human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor C (VEGFC), was assessed using the Human Protein Atlas (HPA), allowing verification of their spatial distribution and expression intensity in normal versus cancerous lung tissues. Moreover, Oncomine, a robust cancer microarray database, was employed to perform cross-dataset comparisons of HER2 and VEGFC expression across multiple independent studies, providing further confidence in the reproducibility and consistency of expression patterns observed in NSCLC.

2.2 Differential Expression Analysis

Differential gene expression between non-small cell lung cancer (NSCLC) tissues and corresponding normal lung tissues was systematically analyzed using the widely adopted “DESeq2” package in R (version 4.3.0). This approach allowed for rigorous normalization and statistical evaluation of RNA sequencing data, ensuring that observed differences in gene expression were biologically meaningful and not due to technical variability. Genes exhibiting a |log2 fold change| greater than 1, along with an adjusted p-value less than 0.05, were considered to be significantly differentially expressed, reflecting substantial upregulation or downregulation in tumor tissues compared with normal controls. These criteria were applied to comprehensively identify candidate genes that may contribute to NSCLC pathogenesis.Specifically, the expression levels of human epidermal growth factor receptor 2 (HER2) and vascular endothelial growth factor C (VEGFC), two genes of particular interest due to their known roles in tumor growth, angiogenesis, and metastasis, were extracted from the dataset. Their expression patterns across tumor and normal samples were visually represented using boxplots, providing an intuitive overview of the differential expression and enabling easy comparison between NSCLC subtypes and normal lung tissues.To further validate the reproducibility of these findings, independent Gene Expression Omnibus (GEO) datasets were analyzed using the “limma” package in R, which is optimized for the analysis of microarray and RNA expression data. This additional validation step confirmed the observed expression trends of HER2 and VEGFC, ensuring that the differential expression patterns identified in TCGA were consistent across multiple patient cohorts and experimental platforms. The combination of DESeq2 analysis for discovery and limma-based validation strengthened the reliability of the results and highlighted the potential clinical relevance of these two biomarkers in NSCLC.

2.3 Clinical Correlation Analysis

Clinical parameters, including patient age, sex, tumor stage, presence of lymph node metastasis, and histological subtype, were systematically analyzed to explore their potential correlations with the expression levels of HER2 and VEGFC in NSCLC patients. These clinical variables are widely recognized as important prognostic and diagnostic factors in lung cancer, and examining their association with key molecular markers can provide insights into tumor biology and patient stratification. To facilitate this analysis, patients were stratified into high-expression and low-expression groups for each gene based on the median expression level, allowing for clear comparison between groups with relatively elevated versus reduced gene expression.Statistical analyses were performed to determine the significance of these associations. For categorical variables, such as sex, tumor stage, lymph node metastasis status, and histological subtype, Chi-square tests were primarily employed to assess the independence between gene expression levels and clinical characteristics. When the expected frequency in any subgroup was less than five, Fisher’s exact test was used as a more accurate alternative to account for small sample sizes and ensure valid statistical inference. Continuous variables, such as patient age, were compared between high- and low-expression groups using independent samples t-tests, allowing for assessment of mean differences and potential trends related to gene expression. By integrating both categorical and continuous clinical parameters with HER2 and VEGFC expression data, this analysis provided a comprehensive understanding of how these molecular markers may relate to clinicopathological features and potentially influence prognosis in NSCLC patients.

2.4 Survival Analysis

Overall survival (OS) and progression-free survival (PFS) of NSCLC patients were systematically evaluated to investigate the prognostic significance of HER2 and VEGFC expression. Kaplan–Meier survival analysis was employed to estimate survival probabilities over time, providing a visual and quantitative depiction of differences between patient subgroups. Patients were stratified into high- and low-expression groups based on median gene expression levels, and survival curves were generated accordingly. The log-rank test was subsequently applied to compare survival distributions between groups, allowing for assessment of whether observed differences in OS and PFS were statistically significant. This approach is widely used in clinical research to identify potential molecular markers that may impact patient outcomes. In addition to survival curve analysis, both univariate and multivariate Cox proportional hazards regression models were constructed to identify independent prognostic factors among the clinical and molecular variables. The univariate analysis first evaluated each variable individually, determining its association with survival outcomes and highlighting candidates that might influence prognosis. Variables found to be significant in univariate analysis were then incorporated into multivariate models, which account for the simultaneous effects of multiple factors and thereby allow determination of whether HER2 and VEGFC expression independently predict patient survival outcomes. Hazard ratios (HRs) along with corresponding 95% confidence intervals (CIs) were reported to quantify the magnitude and precision of the associations, providing clear insight into the risk of death or disease progression associated with elevated or reduced expression of these genes. Collectively, this analytical framework enabled a comprehensive evaluation of both the individual and combined prognostic value of HER2 and VEGFC in NSCLC.

2.5 Immune Infiltration Analysis

To investigate the potential relationship between key molecular markers and the tumor immune microenvironment in NSCLC, immune cell infiltration levels were systematically estimated using two widely recognized computational algorithms: TIMER 2.0 and CIBERSORT. TIMER 2.0 provides a robust platform for the quantification of six major immune cell types across various cancer types using RNA sequencing data, while CIBERSORT employs a deconvolution approach to estimate the relative proportions of a more comprehensive set of 22 immune cell subtypes within bulk tumor transcriptomic profiles. By applying these complementary methods, we were able to obtain a detailed and reliable characterization of immune cell composition in NSCLC tissues. Following the estimation of immune cell abundance, the potential correlations between HER2 and VEGFC expression levels and the infiltration of specific immune cell populations were examined. This included key effector and regulatory cells such as CD8+ T cells, which are critical for anti-tumor immunity; regulatory T cells, which can suppress immune responses and facilitate tumor progression; and macrophages, which may exhibit pro- or anti-tumor functions depending on their polarization state. Spearman correlation analysis was employed to evaluate these associations due to its ability to capture monotonic relationships between continuous variables without assuming linearity. This approach provided quantitative insight into how HER2 and VEGFC expression might influence, or be influenced by, the immune landscape of NSCLC, offering potential mechanistic clues regarding their roles in tumor progression, immune evasion, and therapeutic response.

2.6 Functional Enrichment Analysis

Genes that were co-expressed with HER2 and VEGFC in NSCLC samples were systematically identified using Pearson correlation analysis. Only genes exhibiting a correlation coefficient (r) greater than 0.4 and a p-value less than 0.05 were considered significantly co-expressed, ensuring that the selected genes demonstrated a moderately strong and statistically meaningful relationship with the expression levels of HER2 and VEGFC. This approach allowed for the identification of gene networks and potential functional partners that may be involved in similar biological processes or regulatory pathways, providing a foundation for understanding the molecular mechanisms underlying NSCLC progression.To explore the biological significance of these co-expressed genes, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were performed using the “clusterProfiler” package in R. GO analysis enabled classification of genes into three main categories—biological process, cellular component, and molecular function—highlighting the specific roles that these genes may play in cellular activities, structural localization, and molecular interactions. KEGG pathway analysis, on the other hand, provided insight into the signaling pathways and metabolic processes in which the co-expressed genes might be involved, shedding light on the potential mechanisms through which HER2 and VEGFC contribute to tumor development, angiogenesis, and immune regulation.Significantly enriched GO terms and KEGG pathways were visually represented using bubble plots, which not only display the magnitude of enrichment but also the statistical significance of each term or pathway. This visualization method allows for intuitive interpretation of complex enrichment results, facilitating identification of key biological processes and pathways that may be closely associated with HER2 and VEGFC function in NSCLC. Collectively, these analyses provided a comprehensive overview of the functional landscape of genes co-expressed with HER2 and VEGFC, offering valuable clues for future experimental validation and potential therapeutic targeting.

2.7 Protein–Protein Interaction Network

Protein–protein interaction (PPI) networks for genes related to HER2 and VEGFC were systematically constructed using the STRING database (version 11.5), which integrates known and predicted interactions based on experimental data, computational prediction, and public literature. This approach enabled the identification of functional connections and interaction patterns among HER2- and VEGFC-associated genes, providing insights into the molecular networks that may contribute to NSCLC development and progression. The resulting PPI networks were then imported into Cytoscape, a widely used bioinformatics platform for network visualization and analysis, allowing for intuitive graphical representation of complex interactions and facilitating the identification of highly interconnected regions within the network. To further pinpoint the most biologically relevant subnetworks, the Molecular Complex Detection (MCODE) plugin in Cytoscape was employed to identify core gene modules. These modules represent clusters of tightly interconnected genes that may act synergistically in critical biological processes, such as cell proliferation, angiogenesis, and immune modulation. By focusing on these modules, we were able to prioritize genes that are likely to play central roles in NSCLC pathogenesis. Hub genes within the PPI network and MCODE-identified modules were subsequently subjected to functional analysis to evaluate their biological significance. This included assessing their involvement in key signaling pathways, cellular processes, and potential regulatory mechanisms. Moreover, the therapeutic relevance of these hub genes was explored, providing a foundation for identifying candidate targets for precision medicine approaches in NSCLC. Collectively, this integrative PPI network analysis allowed for a comprehensive understanding of the molecular interactions surrounding HER2 and VEGFC and highlighted critical genes that may serve as biomarkers or therapeutic targets in lung cancer.

2.8 Statistical Analysis

All statistical analyses in this study were conducted using the R software environment (version 4.3.0), which provides a comprehensive platform for advanced statistical computing and graphical visualization, as well as GraphPad Prism (version 9.5), a widely used software for biostatistical analysis and scientific graphing. These tools allowed for rigorous data processing, analysis, and visualization across different types of datasets, including RNA sequencing, microarray, and clinical parameters.Throughout the analyses, a p-value threshold of less than 0.05 was used to determine statistical significance, in line with widely accepted standards in biomedical research. This criterion ensured that the observed associations, differences, or correlations were unlikely to have occurred by chance, providing confidence in the reliability and reproducibility of the results. By combining the computational power of R with the user-friendly interface and visualization capabilities of GraphPad Prism, all statistical tests—including differential expression analysis, correlation analysis, survival analysis, and comparison of clinical parameters—were performed in a systematic and reproducible manner, supporting the robustness of the study findings.

3 RESULTS

3.1 Differential Expression of HER2 and VEGFC in NSCLC

Analysis of TCGA-LUAD and TCGA-LUSC datasets demonstrated that both HER2 and VEGFC were significantly upregulated in tumor tissues compared to adjacent normal lung tissues (HER2: log2FC = 1.52, $p < 0.001$; VEGFC: log2FC = 1.87, $p < 0.001$) (Figure 1A–B). GEO validation datasets (GSE31210, GSE50081) confirmed the upregulation of HER2 and VEGFC in NSCLC (Figure 1C–D). Oncomine cross-dataset comparison supported consistent overexpression across multiple cohorts.

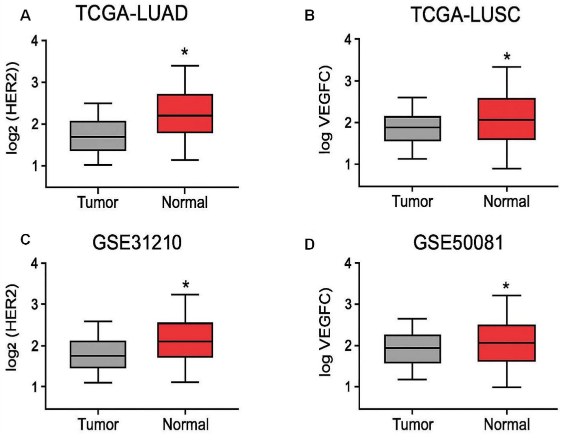
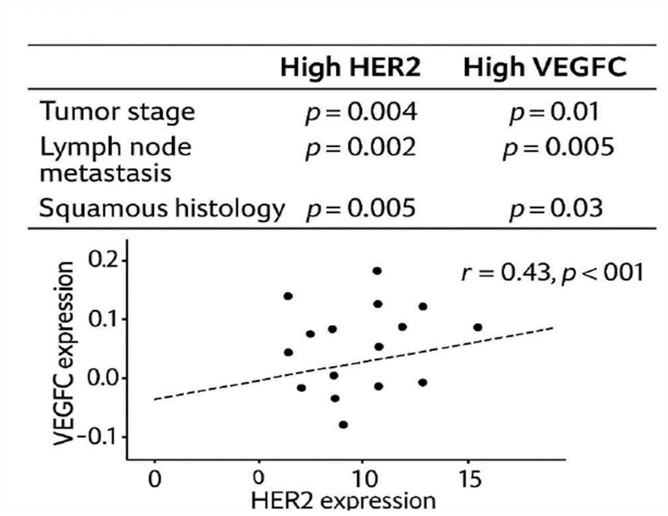


Figure 1 Differential Expression of HER2 and VEGFC in NSCLC

3.2 Clinical Correlation Analysis

High HER2 expression was significantly associated with advanced tumor stage (III–IV, $p = 0.004$) and lymph node metastasis ($p = 0.002$). VEGFC high expression correlated with tumor stage ($p = 0.01$), lymph node involvement ($p = 0.005$), and squamous histology ($p = 0.03$) (Table 1). Co-expression analysis revealed a moderate positive correlation between HER2 and VEGFC ($r = 0.43$, $p < 0.001$).

Table 1 Clinical Correlation Analysis



	High HER2	High VEGFC
Tumor stage	$p = 0.004$	$p = 0.01$
Lymph node metastasis	$p = 0.002$	$p = 0.005$
Squamous histology	$p = 0.005$	$p = 0.03$

3.3 Survival Analysis

Kaplan–Meier curves indicated that patients with high HER2 expression exhibited significantly poorer OS (median OS: 34 vs. 60 months, p < 0.001) and PFS (median PFS: 22 vs. 48 months, p < 0.001). Similarly, high VEGFC expression predicted shorter OS (median OS: 32 vs. 61 months, p < 0.001) and PFS (median PFS: 21 vs. 50 months, p < 0.001) (Figure 2A–B). Combined high expression of HER2 and VEGFC was associated with the worst prognosis (median OS: 28 months, p < 0.001) (Figure 2C). Multivariate Cox regression confirmed HER2 (HR = 1.85, 95% CI: 1.32–2.60, p = 0.001) and VEGFC (HR = 1.70, 95% CI: 1.21–2.40, p = 0.003) as independent prognostic factors (Figure 3).

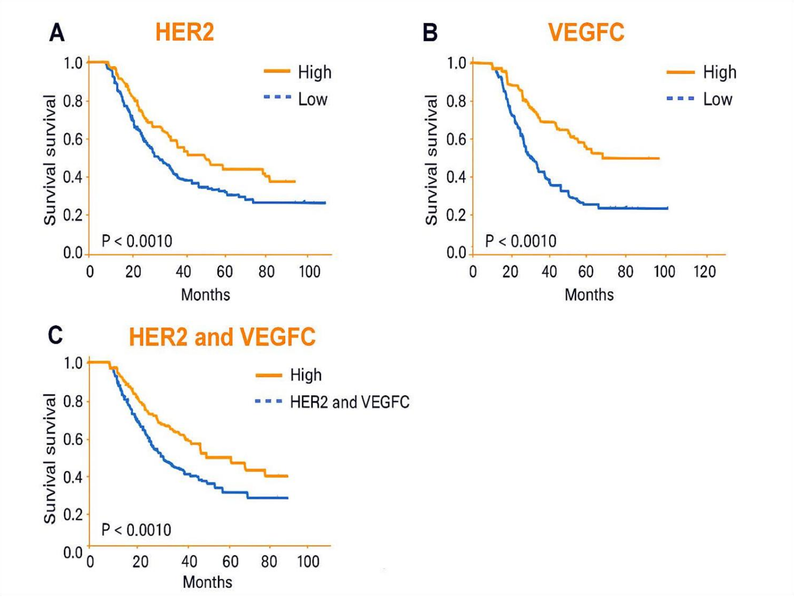


Figure 3 Survival Analysis

3.4 Immune Infiltration Analysis

TIMER and CIBERSORT analyses revealed significant correlations between HER2 and VEGFC expression and immune cell infiltration. HER2 was positively correlated with CD8+ T cells (r = 0.22, p = 0.01) and M2 macrophages. VEGFC expression showed positive correlations with regulatory T cells (r = 0.31, p < 0.001) and dendritic cells (Figure 4). These results suggest potential involvement of HER2 and VEGFC in modulating the tumor immune microenvironment.

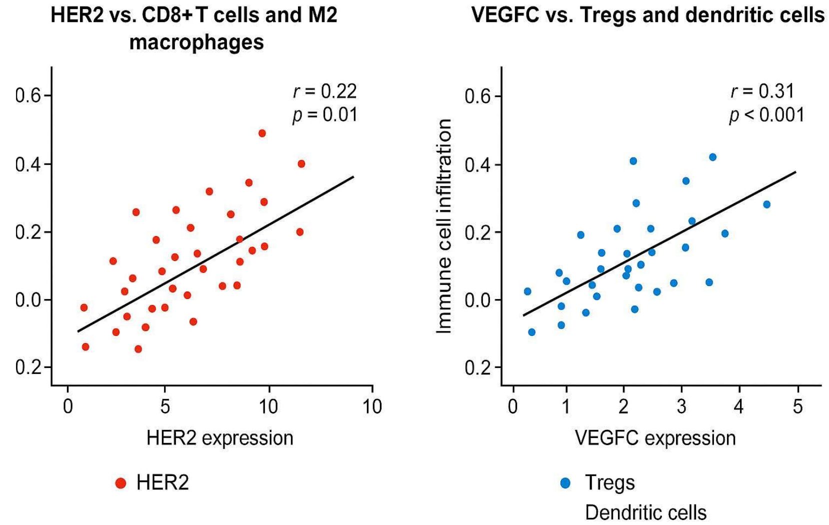


Figure 4 Immune Infiltration Analysis

3.5 Functional Enrichment Analysis

Co-expressed genes with HER2 were enriched in GO terms including "cell proliferation," "epidermal growth factor receptor signaling pathway," and KEGG pathways such as "PI3K-Akt signaling" and "MAPK signaling." VEGFC co-expressed genes were enriched in "lymphangiogenesis," "angiogenesis," and "VEGF signaling pathway" (Figure 5A–B). These findings indicate complementary roles of HER2 and VEGFC in tumor growth and metastasis.

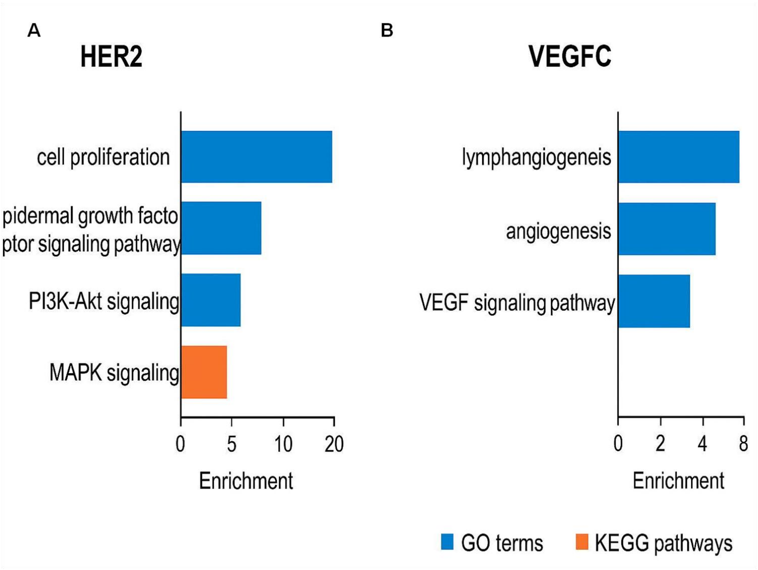


Figure 5 Functional Enrichment Analysis

3.6 Protein–Protein Interaction Network

STRING-based PPI network analysis identified key hub genes interacting with HER2 (EGFR, JUN, LARC) and VEGFC (VEGFR3/FLT4, NRP2, PROX) (Figure 6). MCODE module analysis highlighted densely connected clusters potentially regulating tumor proliferation and lymphatic metastasis, suggesting candidate targets for combinatorial therapy.

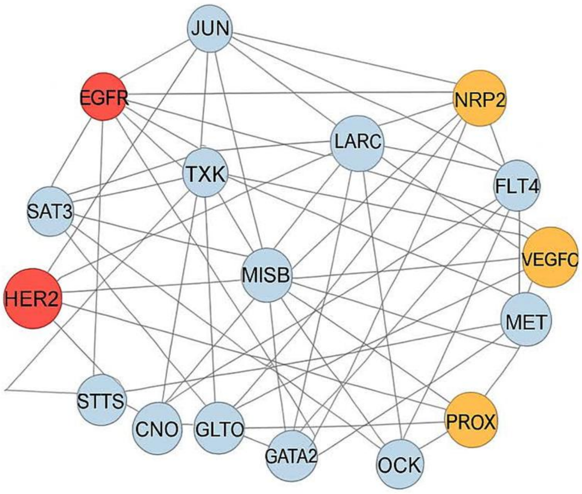


Figure 6 Protein–Protein Interaction Network

4 DISCUSSION

This study provides a comprehensive bioinformatics-based evaluation of HER2 and VEGFC in NSCLC. Our results demonstrated that both genes are significantly upregulated in tumor tissues and are associated with advanced stage, lymph node metastasis, and poor survival. Importantly, combined overexpression further exacerbates adverse prognosis, emphasizing the potential additive effect of these biomarkers[14-16]. HER2's role in activating PI3K-Akt and MAPK signaling pathways likely contributes to uncontrolled tumor proliferation. VEGFC's involvement in lymphangiogenesis and angiogenesis supports its role in metastatic dissemination. Immune infiltration analyses suggest that HER2 and VEGFC may modulate the tumor microenvironment, particularly affecting CD8+ T cells, regulatory T cells, and macrophage populations, potentially impacting response to immunotherapy[17]. Compared with previous studies primarily relying on small patient cohorts or experimental assays, our bioinformatics approach integrates large-scale multi-omic datasets, increasing statistical power and generalizability. However, limitations include the retrospective nature of publicly available datasets and lack of experimental validation. Future studies should combine bioinformatics predictions with functional assays and clinical trials to confirm therapeutic implications[18]. Our findings highlight HER2 and VEGFC as promising prognostic biomarkers and potential therapeutic targets in NSCLC. Targeted inhibition of HER2 signaling combined with VEGFC/lymphangiogenesis blockade may offer synergistic benefits, warranting further preclinical and clinical exploration[19-21].

5 CONCLUSION

HER2 and VEGFC are consistently upregulated in NSCLC and correlate with adverse clinicopathological features and poor survival. Co-expression analysis suggests a synergistic negative impact on prognosis. Functional enrichment and immune infiltration analyses reveal potential mechanistic insights into tumor progression and immune evasion. These findings underscore the potential of HER2 and VEGFC as prognostic biomarkers and therapeutic targets, providing a foundation for future translational and clinical studies.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

The project was supported by Taizhou School of Clinical Medicine, Nanjing Medical University (TZKY20230110).

REFERENCES

[1] Yufei Sheng, Lulu Yang, Boyang Wang, et al. Plasma-derived circALG8 and circCAMTA1 as a panel for early diagnosis of non-small cell lung cancer. Biomarkers in medicine, 2025, 19(16): 725-736.

[2] Unchalee P, Sumitra T, Buntitabhon S. Cost-Utility Analysis of First-Line Pemetrexed Plus Cisplatin in Non-Small Cell Lung Cancer in Thailand. Value in health regional issues, 2020, 21: 9-16.

[3] Giaccone G, Bazhenova L A, Nemunaitis J, et al. A phase III study of belagenpumatucel-L, an allogeneic tumour cell vaccine, as maintenance therapy for non-small cell lung cancer. European journal of cancer (Oxford, England: 1990), 2015, 51(16): 2321-2329.

[4] Harun M, Patel I, Ahmad R, et al. <i>In silico</i> search of triple mutant T790M/C797S allosteric inhibitors to conquer acquired resistance problem in non-small cell lung cancer (NSCLC): a combined approach of structure-

based virtual screening and molecular dynamics simulation. Journal of biomolecular structure & dynamics, 2021, 39(4): 1491-1505.

[5] Patricia L L, Lizet S, Danay S, et al. Identifying predictive biomarkers of CIMAvaxEGF success in non-small cell lung cancer patients. BMC cancer, 2020, 20(1): 772.

[6] Mark M, Awad R, Govindan K N, et al. Personalized neoantigen vaccine NEO-PV-01 with chemotherapy and anti -PD-1 as first-line treatment for non-squamous non-small cell lung cancer. Cancer cell, 2022, 40(9): 1010-1026.

[7] Marco T, Erika R, Giulia B, et al. CIMAvax-EGF, a therapeutic non-small cell lung cancer vaccine. Expert opinion on biological therapy, 2018, 18(7): 829-835.

[8] Juan C, Trujillo J B, Soriano M, et al. Cost-effectiveness of a machine learning risk prediction model (LungFlag) in the selection of high-risk individuals for non-small cell lung cancer screening in Spain. Journal of medical economics, 2025, 28(1): 147-156.

[9] Xiaomu Wang, Yunping Niu, Fang Bian. The progress of tumor vaccines clinical trials in non-small cell lung cancer. Clinical & translational oncology: official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico, 2025, 27(3): 1062-1074.

[10] Hamada A A, Noreldeen, Lijie D, et al. Serum lipidomic biomarkers for non-small cell lung cancer in nonsmoking female patients. Journal of pharmaceutical and biomedical analysis, 2020, 185: 113220.

[11] Tao Jiang, Changyun Zhai, Chunxia Su, et al. The diagnostic value of circulating cell free DNA quantification in non-small cell lung cancer: A systematic review with meta-analysis. Lung cancer (Amsterdam, Netherlands), 2016, 100: 63-70.

[12] Filippo L, Massimiliano P, Cristian R, et al. Preliminary Evidence on the Diagnostic and Molecular Role of Circulating Soluble EGFR in Non-Small Cell Lung Cancer. International journal of molecular sciences, 2015, 16(8): 19612-19630.

[13] Qi Cai, Shuhui You, Jinglong Huang, et al. Cost-effectiveness of trastuzumab deruxtecan as a second-line treatment for HER2-mutant advanced non-small cell lung cancer. Human vaccines & immunotherapeutics, 2025, 21(1): 2468070.

[14] Fang Fang, Mei Zhao, Jinming Meng, et al. Upregulation of TTYH3 by lncRNA LUCAT1 through interacting with ALYREF facilitates the metastasis in non-small cell lung cancer. Cancer biology & therapy, 2025, 26(1): 2464966.

[15] Yingying Xu, Jinping Li, Xiang Ji, et al. Lymphocyte-to-C-reactive protein ratio predicts prognosis in unresectable locally advanced non-small cell lung cancer patients. Annals of medicine, 2025, 57(1): 2487629.

[16] Guomeng Sha, Zhengwen Wu, Biao Wang, et al. Intratumorally specific microbial-derived lipopolysaccharide contributes to non-small cell lung cancer progression. Virulence, 2025, 16(1): 2548626.

[17] Yuan Xu, Dongjie Ma, Yingzhi Qin, et al. Prognostic significance of pathological response and lymph node status in neoadjuvant immunotherapy for potentially resectable non-small cell lung cancer. Annals of medicine, 2025, 57(1): 2453825.

[18] Fuze Zhu, Xudong Yang, Yanlong Yang, et al. The role of histone methyltransferases in therapeutic resistance of NSCLC. Epigenetics, 2025, 20(1): 2536786.

[19] Wenyi Liu, Jinming Tang, Xu Li, et al. Reduction in surgical scope after neoadjuvant chemotherapy and immunotherapy for non-small cell lung cancer. Oncology letters, 2025, 30(5): 501.

[20] Obada Alhalabi, Lukas Klein, David Wasilewski, et al. Managing hydrocephalus in patients with leptomeningeal disease: A multicenter retrospective analysis. International journal of cancer, 2025, 157(8): 1613-1624.

[21] Gee-Chen Chang, Akhil Kapoor, Chee Khoon Lee, et al. Optimizing management of stage IV EGFR mutant non-small cell lung cancer in Asia: An expert opinion. International journal of cancer, 2025, 157(8): 1648-1661.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3049

ADVANCES IN THE APPLICATION OF DEEP LEARNING IN CERVICAL OSSIFICATION OF THE POSTERIOR LONGITUDINAL LIGAMENT

ZhongLiang Wang[1], Xiang Guo[2*]
[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.
[2]Department of Spine Surgery, Shanghai Changzheng Hospital, Shanghai 200093, China.
[*]Corresponding Author: Xiang Guo

Abstract: Cervical ossification of the posterior longitudinal ligament (OPLL) is relatively common among Asian populations. Its progression can cause spinal canal stenosis and compression of the spinal cord and nerve roots, leading to neurological dysfunction and increasing surgical complexity and the risk of complications. In recent years, early identification, precise evaluation, and appropriate intervention for OPLL have become major focuses in radiology and spine surgery. Artificial intelligence, particularly deep learning, has shown new potential in the detection, lesion segmentation, and prognostic evaluation of this disease. This article integrates existing studies to summarize the advances of deep learning in multimodal imaging and discusses its value in clinical decision support, aiming to provide methodological references and clinical insights for related disciplines.

Keywords: Deep learning; Cervical ossification of the posterior longitudinal ligament; Multimodal medical imaging; Risk prediction

1 INTRODUCTION

Cervical ossification of the posterior longitudinal ligament (OPLL) is a distinctive subtype of degenerative spinal disease. Its defining feature is ectopic ossification of the posterior longitudinal ligament, with ossified lesions protruding into the spinal canal, resulting in canal stenosis, spinal cord compression, and subsequent neurological dysfunction[1–4]. The prevalence of OPLL is markedly higher in Asian populations than in Western countries. Epidemiological data indicate that middle-aged and elderly men in Korea and other East Asian regions are more susceptible to OPLL, which is closely related to genetic predisposition, metabolic abnormalities, and long-term mechanical stress[5]. As the disease progresses, the ossified mass gradually thickens and compresses the spinal cord, leading to symptoms such as upper-limb numbness, gait disturbance, and limb weakness. Even minor trauma can precipitate central cord injury in advanced cases[6-9]. Previous surgical follow-up studies have demonstrated that OPLL is one of the major causes of cervical myelopathy and a key factor influencing surgical strategy and prognosis[10]. Therefore, early detection and timely intervention are essential for preventing disease progression and improving neurological outcomes.

At present, radiographic screening primarily relies on anteroposterior and lateral cervical X-rays, which can reveal characteristic ossified lesions or high-density shadows within the spinal canal. However, this approach is insensitive to early or segmental lesions, with a reported miss rate of approximately 48%[11]. In comparison, computed tomography (CT) provides detailed visualization of bony structures and is regarded as the preferred modality for diagnosing OPLL. CT enables accurate assessment of ossification morphology, thickness, and segmental distribution, thereby assisting in canal stenosis evaluation and preoperative planning. Systematic studies consistently report that CT surpasses plain radiography in both lesion detection and anatomical delineation; however, its higher radiation dose and cost limit its use in long-term follow-up, especially for postoperative patients who require periodic monitoring of ossification progression[12–14]. Magnetic resonance imaging (MRI), on the other hand, offers superior evaluation of soft tissue and spinal cord pathology, allowing direct visualization of spinal cord compression and intramedullary signal changes. Nevertheless, MRI has limited capability for depicting calcified or ossified components, and OPLL may be confounded with ligamentous hypertrophy or intervertebral disc calcification, leading to reduced diagnostic accuracy on conventional sequences[16–17].

With the advancement of artificial intelligence (AI) and deep learning technologies, novel strategies have emerged for the early diagnosis and quantitative assessment of OPLL. Current studies mainly focus on three domains: (1) automatic detection of ossified regions based on X-ray or CT images; (2) assessment of spinal cord involvement and direct identification of OPLL using MRI combined with deep learning algorithms; and (3) clinical prediction models for surgical risk and postoperative prognosis, including complication rates and functional recovery. This review systematically summarizes recent progress in the application of deep learning to multimodal imaging—X-ray, CT, and MRI—for the detection, segmentation, and prognostic evaluation of OPLL, aiming to provide methodological references and clinical perspectives for future research.

2 MATERIALS AND METHODS

A systematic literature search was conducted on September 1, 2025, across the PubMed, MEDLINE, and Web of Science databases. The search strategy followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The following medical subject headings (MeSH) and keywords were used: *"artificial intelligence"* OR *"AI"* OR *"machine learning"* OR *"ML"* OR *"deep learning"* OR *"DL"*, combined with *"ossification of the posterior longitudinal ligament"* OR *"OPLL"*. The search was primarily limited to English-language studies published within the past eight years (2017–2025).

To ensure the comprehensiveness of this review, several earlier seminal studies with substantial academic or clinical influence were also cited in the Discussion section as supplementary references.

Inclusion criteria:

(1) Studies applying artificial intelligence (AI) or machine learning (ML) techniques to cervical ossification of the posterior longitudinal ligament (OPLL);
(2) Human subject studies;
(3) English-language publications;
(4) Peer-reviewed original research articles.

Exclusion criteria:

(1) Non-original research (e.g., review articles, commentaries, conference abstracts, or case reports) and unpublished studies;
(2) Studies focusing on non-OPLL topics or using non-clinical data.

The initial database search identified 163 records. After sequentially excluding studies published outside the specified time frame, papers with inappropriate article types, and non-English publications, a total of 73 studies remained for full-text screening. Following rigorous evaluation according to predefined inclusion criteria, 12 original research articles were finally included in this review.

3 APPLICATION OF DEEP LEARNING IN THE DIAGNOSIS AND QUANTIFICATION OF OPLL

Deep learning technology has been applied to the automatic recognition and segmentation of OPLL across various imaging modalities. This article first reviews deep learning–based methods for the diagnosis and quantification of OPLL using X-ray, CT, and MRI images, and compares the similarities and differences in model performance among these different modalities.

3.1 OPLL Auxiliary Diagnosis Based on X-Ray

Traditional anteroposterior and lateral cervical X-rays are prone to missed diagnoses and misjudgments due to overlapping tissues and subtle calcified lesions. In recent years, several studies have attempted to apply deep learning to the identification of OPLL on X-rays to enhance screening capability.

Murata et al. trained a residual network model on 2,318 radiographs (672 cases) and achieved an overall accuracy of 98.9%, sensitivity of 97.0% [18], specificity of 99.4%, and an AUC of 0.99, demonstrating the potential of OPLL binary screening. Ogawa et al. conducted a small-sample study (50 OPLL cases and 50 controls) and showed that their convolutional neural network performed well in detecting OPLL on X-rays (AUC = 0.924, sensitivity 80%, specificity 100%) [19]. Miura et al. first demonstrated that using the EfficientNet-B4 model for three-category classification of lateral cervical radiographs could reach expert-level performance (accuracy 0.86) [20]; however, its performance was limited in cases with segmental or localized OPLL and in low-quality images, indicating a need for improvement. The study also emphasized the clinical potential of AI for early OPLL screening.

Notably, Chae et al. conducted an observer study showing that incorporating AI model results into film reading assistance could improve diagnostic performance [11]. When the AI model analyzed patients independently, its AUC was 0.851 (sensitivity 91%, specificity 69%); when radiologists performed a second review with AI assistance after their initial interpretation, the average patient-based AUC increased from 0.841 to 0.911, suggesting that AI is better suited as an auxiliary reading tool to enhance overall diagnostic performance.

Overall, deep learning can screen for OPLL on cervical X-rays with high sensitivity, reducing missed and incorrect diagnoses caused by limited CT examinations, and has potential value in health check-ups or primary care. AI-based automatic screening using plain radiographs can not only reduce unnecessary CT radiation exposure but also assist physicians in early diagnosis. However, differences remain among studies in sample size, external validation, and annotation standards, and further multicenter and standardized studies are needed to verify clinical generalizability.

3.2 Quantitative Segmentation of OPLL Based on CT

Cervical CT is considered the gold standard for diagnosing OPLL because ossified lesions can be clearly visualized, and manual reading maintains high sensitivity and specificity. Therefore, deep learning studies directly targeting OPLL detection on plain CT scans are relatively limited. Clinically, the advantage of CT lies more in the precise quantification of the range and volume of ossified lesions, providing an important basis for surgical planning. In recent years, as deep learning technology has matured, it has become feasible to quantify the extent and volume of OPLL ossification from CT images.

Jiang et al. conducted a dual-center study based on CT data from 307 patients [21], constructing a 3D U-Net segmentation model that simultaneously quantified the maximum ossification thickness, residual spinal canal diameter at the corresponding level, and the "spinal cord compression coefficient." The model achieved a Dice coefficient of 0.71 and an average surface distance (ASD) of 2.63 mm on the external test set, showing little difference from manual measurement. External validation demonstrated that automatic segmentation and compression quantification on the CT platform have clinical feasibility. Overall, CT imaging—with its rich anatomical detail—is highly suitable for the segmentation and quantification of OPLL ossified lesions using deep learning. The automatically segmented results can also be applied to preoperative 3D visualization and simulated surgical planning. However, CT segmentation models require large amounts of slice-by-slice annotated data for training, which entails a heavy labeling workload. Currently, publicly reported studies of this type remain limited, but their performance demonstrates that deep learning has the potential to significantly reduce physicians' measurement workload while maintaining high accuracy.

3.3 OPLL Auxiliary Diagnosis Based on MRI

MRI is highly sensitive for displaying soft tissue and spinal cord lesions, but its ability to visualize ossified lesions is inferior to that of CT. Nevertheless, some researchers have attempted to develop deep learning models based on MRI, hoping to identify OPLL patients directly from MRI without relying on CT.

Shemesh et al. developed an MRI-AI tool based on a single-center cohort of consecutive patients (900 cases within 36 months, including 65 OPLL cases) [22]. The tool can automatically layer and segment the vertebral bodies, posterior longitudinal ligament, and disc-ligament complex, and identify OPLL at the patient level. In a reader comparison study, the model achieved a sensitivity of 85%, specificity of 98%, and overall accuracy of 98%, and additionally detected five OPLL cases that were initially missed by manual interpretation, suggesting that MRI combined with AI also has high discriminative ability even without CT.

In addition, Qu et al. constructed MRI classification models based on ResNet34/50/101 using surgical cases (272 OPLL cases and 412 cases of degenerative stenosis) [23]. In the test set, the model accuracies reached 92.98%, 95.32%, and 97.66%, respectively. Among them, the accuracy and specificity of ResNet50/101 were significantly higher than those of three spine surgeons, and ResNet101 also achieved higher sensitivity than two of them, further demonstrating that deep neural networks can enhance the detection performance of OPLL on MRI.

Overall, MRI-AI has shown considerable potential in the automatic stratification and detection of OPLL. However, multicenter external validation and finer-grained lesion-level segmentation studies are still needed to achieve CT-equivalent quantitative assessment of ossified lesions and promote clinical application.

4 APPLICATION OF DEEP LEARNING IN PROGNOSIS

In addition to assisting diagnosis, deep learning has begun to play a role in preoperative decision support and prognosis prediction for OPLL patients. By integrating imaging and clinical data, AI models can be used to predict surgical approach selection, postoperative complication risks, and functional recovery, thereby helping clinicians develop more optimized treatment strategies. The following section summarizes several representative studies on predictive models.

4.1 Preoperative Decision Support for OPLL

The choice of surgical approach for cervical OPLL (anterior, posterior, or combined) has long been controversial. The decision requires comprehensive consideration of the ossification range, spinal canal occupancy ratio, and the K-line (cervical lordotic line). Traditionally, this judgment is qualitatively made by experienced spine surgeons based on imaging, which inevitably introduces subjectivity. Deep learning can transform such empirical knowledge into reusable and standardized models.

Li et al. proposed a two-stage process of "Dilated TransUNet + discriminative post-processing. [24]" First, a U-Net variant combining dilated convolution and Transformer was used to detect the four boundary points (C2–C7) on lateral cervical CT images and reconstruct the K-line. Then, within the ROI formed by these four points, a bimodal histogram–based dynamic threshold and left–right "white zone" continuity difference were used to distinguish the artifacts of "ossification crossing the line" and "vertebral crossing the line," outputting the K-line determination. The test results showed an average keypoint detection error of 1.463 ± 3.007 pixels, SDR-5 of 98.49%, and SDR-9 of 98.90%; 97.8% of images had all four points located within 5 pixels. Compared with baseline models such as ASM, U-Net, and TransUNet, the new method achieved better overall accuracy and significantly reduced misidentification at the easily confused C7 level, effectively suppressing interference from "ossification crossing" or "vertebral crossing." Converting K-line determination from manual drawing to standardized, reproducible automatic evaluation facilitates standardized surgical approach selection, surgeon–patient communication, and structured preoperative stratification.

4.2 Postoperative Risk and Prognostic Prediction for OPLL

Surgery for OPLL is technically demanding, and postoperative complications such as neurological injury, C5 nerve root palsy, dural tears, and infection are not uncommon. Identifying high-risk patients preoperatively and implementing targeted preventive measures could improve surgical safety.

Ito et al. conducted a multicenter prospective study (28 institutions, 478 cases) and trained a deep learning model using preoperative clinical and imaging features [25]. The model achieved an overall prediction accuracy for complications of 74.6%, comparable to logistic regression (74.1%); for neurological complications specifically, accuracy reached 91.7%, indicating feasibility for use in risk communication.

Kim et al. developed various machine learning models using a single-center retrospective cohort of 901 cases. The best-performing model achieved an AUC of 0.88, significantly higher than logistic regression (AUC = 0.69) [26]. The model identified age, surgical approach, involvement of the C1–C3 segments, and immediate postoperative shoulder pain as factors associated with C5 palsy, suggesting enhanced intraoperative protection and postoperative monitoring. This finding indicates that machine learning can be used for early warning of specific complications.

Regarding postoperative functional improvement, Maki et al. developed a preoperative prediction model using data from 478 cases to determine whether cervical OPLL patients would achieve “minimal clinically important difference (MCID)” at 1 year and 2 years after surgery [27]. The XGBoost model performed best at 1 year (AUC = 0.72, accuracy 67.8%), while the random forest model performed best at 2 years (AUC = 0.75, accuracy 69.6%). These results demonstrate the feasibility of constructing postoperative prognostic ML models in OPLL populations, with current discriminative power at a moderate level (evidence level 4), providing a reference for expectation management and rehabilitation planning.

For ossification progression prediction, Qin et al. developed an interpretable model to predict postoperative re-ossification and progression risks [28]. They extracted radiomics features from preoperative CT images of 473 followed-up postoperative patients and combined them with clinical variables. The combined model achieved an AUC of 0.751 on the test set, outperforming radiomics-only (AUC = 0.693) and clinical-only models (AUC = 0.620). The results support that “preoperative CT radiomics + clinical variables” can enable individualized prediction of postoperative progression risk of OPLL, facilitating clinical interpretation and personalized follow-up.

5 CURRENT RESEARCH STATUS AND LIMITATIONS

Based on existing findings, the application of deep learning to multimodal imaging in OPLL has become increasingly well defined. X-ray–based intelligent models are mainly used for rapid screening and reducing missed diagnoses. CT-based algorithms focus on lesion segmentation and quantitative feature extraction, such as ossified volume and thickness, providing objective references for preoperative planning. MRI-related models can avoid ionizing radiation and have achieved high accuracy in individualized diagnosis. Notably, integrated models combining imaging and clinical variables have demonstrated diagnostic performance comparable to or even superior to traditional methods in predicting postoperative complications and functional recovery. Overall, AI tools in the OPLL field are evolving from simple recognition toward quantitative analysis and clinical decision-support systems, showing practical value in early screening, surgical pathway selection, and risk communication.

Despite the remarkable progress, several critical bottlenecks remain at the current stage. Most models are built on retrospective, single-center data, lacking external validation and prospective study design, which limits their applicability to different populations. Meanwhile, inconsistency in the delineation of lesion boundaries and classification standards across studies and the absence of universally accepted annotation protocols hinder cross-study comparisons and affect model stability and generalization performance. In addition, model interpretability and uncertainty assessment have not received sufficient attention, reducing clinicians’ trust in algorithmic outputs and restricting their integration into real clinical workflows. Therefore, the translation of AI models into clinical practice for OPLL still requires further breakthroughs.

Future research can proceed along several directions. Strengthening multicenter collaboration and prospective study design will be essential to build diverse, high-quality datasets and improve model generalizability and robustness through cross-institutional validation. Meanwhile, establishing unified imaging annotation and grading systems to clearly define ossification boundaries and lesion types will facilitate model training and result comparison. Model transparency is equally important; techniques such as saliency visualization, feature-importance ranking, and confidence-interval estimation can help clinicians understand model reasoning and assess its reliability. Moreover, the stability of models under different scanner vendors, acquisition parameters, and image quality conditions also needs to be verified. With systematic improvements in data sharing, interpretability, and cross-domain validation, deep learning is expected to achieve a key transition—from algorithmic performance optimization to tangible patient benefit.

6 SUMMARY AND OUTLOOK

Cervical ossification of the posterior longitudinal ligament (OPLL) is a spinal disorder associated with a high risk of disability. Imaging examinations play a central role in establishing the diagnosis and formulating treatment strategies. In recent years, the rapid development of deep learning has created new opportunities for imaging-based recognition and preoperative decision-making in OPLL. Artificial intelligence–assisted analysis of multimodal imaging has demonstrated high accuracy and feasibility in multiple studies, while preoperative predictive models have shown clinical value in surgical strategy selection. Notably, deep learning–based approaches for complication risk assessment and postoperative functional prediction have achieved performance comparable to or even exceeding that of traditional models in some studies. Although these algorithms remain in the validation and optimization stage, the research direction of intelligent decision support for OPLL is becoming increasingly clear—by quantitatively evaluating preoperative risks and potential benefits, they can provide clinicians with more informative decision-making references.

Overall, deep learning is driving a paradigm shift in the diagnosis and treatment of OPLL. Its applications are expanding from image recognition to quantitative analysis and clinical decision support, leading to continuous optimization of diagnostic and therapeutic workflows. Further prospective validation and inter-institutional collaboration are required to ensure model robustness and generalizability. With the deepening integration of medicine and computational technology, deep learning is expected to become a routine tool for multidisciplinary teams in spinal disease management, enhancing the clarity of imaging interpretation, the precision of quantitative evaluation, and the reliability of therapeutic decision-making. Close collaboration among radiologists, spine surgeons, and algorithm researchers will accelerate the translation of technological advances into clinical benefits. Looking ahead, the role of deep learning in precision medicine for OPLL will continue to strengthen, ultimately aiming to provide patients with more accurate, safer, and personalized treatment strategies.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This study was supported by the National Natural Science Foundation of China (82272454) and the Shanghai Municipal Health Commission Program for Outstanding Young Medical Talents (20234Z0016).

REFERENCES

[1] Xi Y, Xu J, Zhang X, et al. Risk factor analysis and risk prediction model construction of ossification progression after postoperative cervical ossification of the posterior longitudinal ligament. Spine (Phila Pa 1976), 2025, 50(15): 1025-1034.
[2] Liu Y, Zhang Y, Wang J, et al. Surgical options for ossification of the posterior longitudinal ligament of the cervical spine: a narrative review. Journal of Orthopaedic Surgery and Research, 2024, 19(1): 707.
[3] Ahn DK, Park HS, Choi DJ, et al. Prevalence and characteristics of cervical ossification of the posterior longitudinal ligament in the Korean population: a population-based study using computed tomography. Spine (Phila Pa 1976), 2020, 45(23): E1529-E1536.
[4] Kobayashi K, Imagama S, Ando K, et al. Diagnosis and assessment of the ossification of the posterior longitudinal ligament: problems, current issues, and future perspectives. Spine Surgery and Related Research, 2020, 4(4): 252-261.
[5] Sohn S, Chung CK, Yun TJ, et al. Epidemiological survey of ossification of the posterior longitudinal ligament in an adult Korean population: three-dimensional computed tomographic observation of 3,240 cases. Calcified Tissue International, 2014, 94(6): 613-620.
[6] Liu WKT, Siu KHY, Cheung JPY, et al. Radiographic characterization of OPLL progression in patients receiving laminoplasty with a minimum of two-years follow-up. Neurosurgical Review, 2024, 47(1): 505.
[7] Saetia K, Cho D, Lee S, et al. Ossification of the posterior longitudinal ligament: a review. Neurosurgical Focus, 2011, 30(3): E1.
[8] Nouri A, Cheng JS, Davies B, et al. Degenerative cervical myelopathy: a brief review of past perspectives, present developments, and future directions. Journal of Clinical Medicine, 2020, 9(2): 535.
[9] Matsunaga S. Trauma-induced myelopathy in patients with ossification of the posterior longitudinal ligament. Journal of Neurosurgery-Spine, 2002, 97(2): 172–175.
[10] Smith ZA, Buchanan CC, Raphael D, et al. Ossification of the posterior longitudinal ligament: pathogenesis, management, and current surgical approaches. Neurosurgical Focus, 2011, 30(3): E10.
[11] Chae H D, Jeon T S, Park J H, et al. Improved diagnostic performance of plain radiography for ossification of the posterior longitudinal ligament of the cervical spine using deep learning. Plos One, 2022, 17(4): e0267643.
[12] Liawrungrueang W, Tseng C, Corr F, et al. A systematic review of computed tomography imaging for cervical ossification of the posterior longitudinal ligament: diagnostic strengths and limitations in radiation exposure and cost. Quantitative Imaging in Medicine and Surgery, 2025, 15(3): 1719-1740.
[13] Tseng C, Wang H, Lin Y, et al. Computed tomography is required for definitive diagnosis of cervical ossification of the posterior longitudinal ligament but is associated with increased radiation exposure and higher costs. Spine (Phila Pa 1976), 2025, 50(12): 845-853.
[14] Jeong H S, Park C, Kim K S, et al. Clinical feasibility of MR-generated synthetic CT images of the cervical spine: diagnostic performance for detection of ossification of the posterior longitudinal ligament and comparison of CT number. Medicine (Baltimore), 2021, 100(18): e25800.
[15] Kang MS, Lee JW, Zhang HY, et al. Diagnosis of cervical ossification of the posterior longitudinal ligament in lateral radiograph and MRI: is it reliable? Korean Journal of Spine, 2012, 9(3): 205-208.
[16] Hsiung W, Lin H Y, Lin H H, et al. MRI-based lesion quality score assessing ossification of the posterior longitudinal ligament of the cervical spine. The Spine Journal, 2024, 24(7): 1162-1169.
[17] Caffard T, Blum A, Lecocq S, et al. Diagnostic accuracy of zero-echo time MRI for detecting cervical ossification of the posterior longitudinal ligament compared to CT. International Orthopaedics, 2024, 48(5): 993-1001.

[18] Murata K, Endo K, Aihara T, et al. Use of residual neural network for the detection of ossification of the posterior longitudinal ligament on plain cervical radiography. European Spine Journal, 2021, 30(8): 2185-2190.

[19] Ogawa T, Yoshii T, Oyama J, et al. Detecting ossification of the posterior longitudinal ligament on plain radiographs using a deep convolutional neural network: a pilot study. The Spine Journal, 2022, 22(6): 934-940.

[20] Miura M, Maki S, Miura K, et al. Automated detection of cervical ossification of the posterior longitudinal ligament in plain lateral radiographs of the cervical spine using a convolutional neural network. Scientific Reports, 2021, 11(1): 12702.

[21] Jiang B, Yan J, Xu S, et al. Deep learning-based segmentation of cervical posterior longitudinal ligament ossification in computed tomography images and assessment of spinal cord compression: a two-center study. World Neurosurgery, 2025, 194: 123567.

[22] Shemesh S, Kimchi G, Yaniv G, et al. MRI-based detection of cervical ossification of the posterior longitudinal ligament using a novel automated machine-learning diagnostic tool. Neurosurgical Focus, 2023, 54(6): E11.

[23] Qu Z, Deng B, Sun W, et al. A convolutional neural network for automated detection of cervical ossification of the posterior longitudinal ligament using magnetic resonance imaging. Clinical Spine Surgery, 2024, 37(3): E106-E112.

[24] Li X, Ma D, Zhang H, et al. An end-to-end deep learning framework for the automated diagnosis of OPLL in CT images. Biomedical Signal Processing and Control, 2025, 110(Pt B): 108150.

[25] Ito S, Nakashima H, Yoshii T, et al. Deep learning-based prediction model for postoperative complications of cervical posterior longitudinal ligament ossification. European Spine Journal, 2023, 32(11): 3797-3806.

[26] Kim S H, Lee S H, Shin D A. Could machine learning better predict postoperative C5 palsy of cervical ossification of the posterior longitudinal ligament? Clinical Spine Surgery, 2022, 35(5): E419-E425.

[27] Maki S, Furuya T, Yoshii T, et al. Machine-learning approach in predicting clinically significant improvements after surgery in patients with cervical ossification of the posterior longitudinal ligament. Spine (Phila Pa 1976), 2021, 46(24): 1683-1689.

[28] Qin S, Qu R, Liu K, et al. Predicting postoperative progression of ossification of the posterior longitudinal ligament in the cervical spine using interpretable radiomics models. Neurospine, 2025, 22(1): 144-156.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3048

MECHANISM STUDY OF SHENFU INJECTION REGULATING MAPK PATHWAY TO ALLEVIATE MYOCARDIAL ISCHEMIA-REPERFUSION INJURY IN RATS BASED ON BIOINFORMATICS AND EXPERIMENTAL VERIFICATION

Hong Wang[1], Liu Yang[1], HuaiGang Chen[2], GaoMing Hong[3], ShengQiang Zeng[1], WenYan Xu[1,2], Yao Hu[1*]
[1]Department of Cardiology, Jiangxi Provincial People's Hospital, The First Affiliated Hospital of Nanchang Medical College, Nanchang 330006, Jiangxi, China.
[2]Medical College of Nanchang University, Nanchang 330006, Jiangxi, China.
[3]The People's Hospital of Yujiang District, Yingtan 335000, Jiangxi, China.
[*]Corresponding Author: Yao Hu

Abstract: To elucidate SFI's protective mechanism against MIRI, this study identified the MAPK pathway via bioinformatics. Experimental validation in MIRI rats confirmed that SFI improves myocardial injury by suppressing MAPK activation and inflammatory factors (IL-1β, IL-6, TNF-α), supporting its clinical potential.

Keywords: Shenfu Injection (SFI); Myocardial Ischemia-Reperfusion Injury (MIRI); Mitogen-Activated Protein Kinase (MAPK) signaling pathway; Bioinformatics; Inflammatory factors

1 INTRODUCTION

Myocardial is chemia-reperfusion injury (MIRI), a serious post-revascularization complication, is characterized by oxidative stress, inflammation, and cell death[1-2], The outbreak of reactive oxygen species (ROS) and calcium ion (Ca^2) overload in the early stage of reperfusion trigger the opening of mitochondrial permeability transition pore (mPTP) and energy imbalance[3], subsequently, damage-associated molecular patterns (DAMPs) mediate the activation of pathways including Toll-like receptor-nuclear factor-kappa B (TLR-NF-kB), nucleotide-binding oligomerization domain-like receptor protein 3 (NLRP3), and mitogen-activated protein kinase(MAPK), leading to the increase of pro-inflammatory factors (IL-1β, IL-6, TNF-α) and further aggravation of tissue damage[4]. Shenfu Injection's documented multi-faceted pharmacological profile aligns with the network pathology of MIRI, suggesting its potential as a comprehensive therapeutic agent [5]. However, its upstream molecular targets and signaling networks have not been systematically clarified, which restricts the establishment of mechanistic and translational evidence. Among the multiple parallel pathways in MIRI, the MAPK pathway (p38/JNK/ERK) plays a core role [6]: p38 and c-Jun N-terminal kinase (JNK) dominate the stress and inflammatory transcriptional programs, while extracellular signal-regulated kinase (ERK) is associated with cell survival/ repair [7]. Since re-perfusion often over-activates pro-inflammatory/ apoptosis p38/JNK, ideal cardio protection requires their inhibition while maintaining ERK activity. SFI's ability to remodel these pathways at this hub provides a unified mechanism for its therapeutic effect.

Integrating bioinformatics prediction with in vivo validation, this study demonstrated that SFI alleviates MIRI by inhibiting the MAPK pathway and subsequent inflammation.

2 MATERIALS AND METHODS

2.1 Bioinformatics Analysis

2.1.1 GEO data mining and biofunctional analysis

This study obtained the GSE6381 dataset from the GEO database, applying log 2transformation and normalization. Probes were mapped to gene symbols, redundant records were removed, and mean values were calculated. Batch effects were addressed using the limma package, while differential expression analysis was performed with GEO2R based on the criteria $|logFC|>1.0$ and adjusted $p <0.05$. Target genes were integrated to construct miRNA-mRNA networks, and GO and KEGG analyses were conducted using DAVID with adjusted $p <0.05$ as the significance threshold.

2.1.2 Animal modeling and grouping

Male SD rats (SPF grade,8-10 weeks old,250-300 g) were acclimatized for 7 days under standard SPF conditions (22±2°C,50±5% humidity,12-h light/ dark cycle) with free access to food and water. All procedures were approved by the Institutional Animal Ethics Committee and designed as terminal experiments. Rats were randomly divided into five groups (n=5): Sham, I/R, SFI-L, SFI-M, and SFI-H.

2.1.3 Drug administration

Following a 7-day pretreatment with SFI or saline, I/R was induced by LAD clamping. Investigators were blinded to the treatment groups throughout the experiment. Myocardial I/R injury was induced by 30- minute occlusion of the left anterior descending coronary artery followed by60- minute reperfusion, with sham rats undergoing identical surgery without occlusion.

2.2 Detection of Molecular Mechanisms and Inflammatory Indicators

Sampling and Time Points: According to the animal experimental protocol, the left ventricular anterior wall ischemic zone and plasma were collected immediately after 60 minutes of reperfusion for the following assays.

2.3 qRT-PCR (p38, JNK, ERK mRNA)

The mRNA expression of MAPK pathway genes (p38/JNK/ERK) was analyzed by qPCR in triplicate and calculated relative to Sham controls using the 2^- Δ△Ct method.

2.4 Western Blot (p-p38、p-JNK、p-ERK)

Phosphorylation levels of MAPK proteins (p38, JNK, ERK) were determined by Western blot analysis. Protein extracts were separated by SDS-PAGE, transferred to PVDF membranes, and probed with specific antibodies. The p-MAPK/ total-MAPK ratio was quantified using ImageJ, with technical replicates for validation.

2.5 ELISA (IL-1β、IL-6、TNF-α)

Plasma analyte levels were measured in duplicate using validated kits (CV ≤ 10%) and are expressed in pg/mL.

3 STATISTICAL ANALYSIS

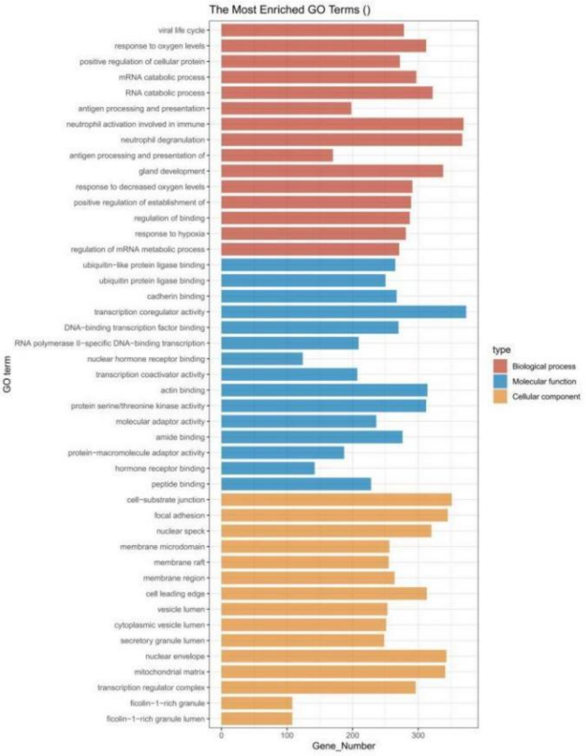
Data are presented as the mean±standard deviation. Intergroup comparisons were performed using one-way analysis of variance (ANOVA), with a P-value of less than 0.05 considered statistically significant.

4 RESULTS

4.1 Bio-informatics Prediction Results

Enrichment analysis highlighted key processes: in BP, hypoxia response and immune activation; in MF, kinase/ ubiquity-related activities; in CC, secretary and nuclear structures. This collectively suggests crucial roles in immune secretion, energy metabolism, and transcriptional regulation during MIRI.

Significantly enriched KEGG pathways included the MAPK signaling pathway, Rap1 signaling, Apoptosis, and pathways related to cardiovascular-metabolic injury, such as ' Fluid shear stress and atherosclerosis', 'AGE-RAGE signaling in diabetic complications',' Lipid and atherosclerosis', and' Diabetic cardiomyopathy'. Neurodegenerative disease pathways (e.g., Alzheimer's, Parkinson's, Huntington's) and several pathogen infection pathways (e.g., Epstein-Barr virus infection, Hepatitis B, Human cytomegalovirus infection) were also significantly enriched.



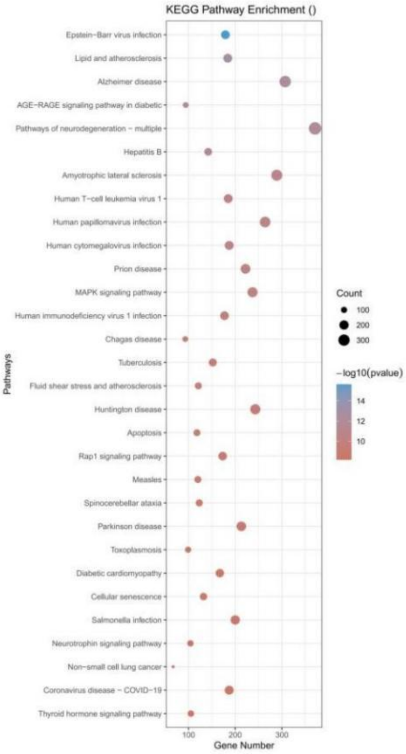


Figure 1 Functional Enrichment Analysis of Differential Expressed Genes

Collectively, the enrichment profile points to stress-inflammatory cascades and metabolic/ mitochondrial imbalance as key pathological hubs. This bio-information prediction aligns with the in vivo observation of MAPK axis activation and its suppression by Shenfu Injection (SFI), as illustrated in Figure 1.

GO/KEGG enrichment was analyzed (cluster Profiler v3.18.0) and visualized as bar/ bubble plots showing significance (-log10(P/FDR)) and gene count. Significant terms/ pathways met P/FDR <0.05.

4.2 Experimental Verification Results

4.2.1 MAPK pathway regulation

The MAPK pathway was activated: phosphorylation levels of p-p38, p-JNK, and p-ERK were significantly higher in the I/R group compared to the Sham group. Shenfu Injection (SFI) treatment resulted in a dose-dependent reduction (High> Medium> Low) in phosphorylation, consistent with the enriched MAPK signaling pathway. At the mRNA level, expressions of p38, JNK, and ERK were up-regulated in the I/R group, while SFI treatment induced dose-dependent suppression, with the high-dose group approaching baseline levels (See Table 1).

Table 1 Shenfu Injection (SFI) Attenuates Ischemia-Reperfusion (I/R)-Induced MAPK Pathway Activation in a Dose-Dependent Manner: Results from qRT-PCR and Phosphoprotein Quantification

Group	p38mRNA	JNKmRNA	ERKmRNA	p-p38	p-JNK	p-ERK
I/R	1.98±0.1	2.39±0.18	1.84±0.31	5198±323	4719±500	5228±260
SFI-H	1.12±0.10	1.15±0.08	1.18±0.21	1869±40	1804±233	1633±193
SFI-L	1.63±0.11	1.80±0.19	1.53±0.16	3967±207	3527±229	3766±544
SFI-M	1.41±0.12	1.38±0.10	1.27±0.11	2915±322	2571±216	2664±286
Sham	1.01±0.06	0.91±0.07	1.00±0.11	1227±96	1010±138	1029±96

4.2.2 Changes in Inflammatory Cytokines

Table 2 Plasma Inflammatory Factor Levels (ELISA, pg/mL, mean ± SD)

Group	IL-1β	IL-6	TNF-α
Sham	11.8±1.2	22.2±2.2	14.7±1.5
I/R	83.2±8.1	197.3±19.6	128.3±12.6
SFI-L	57.7±5.5	160.5±16.0	92.6±9.3
SFI-M	42.7±4.3	93.5±9.4	62.5±6.3
SFI-H	25.4±2.5	55.4±5.5	40.1±4.0

Compared to the Sham group, the I/R group exhibited significant increases in IL-1β, IL-6, andTNF-α levels. All doses of Shenfu Injection (SFI) significantly reduced the levels of these three inflammatory cytokines in a dose-dependent manner (High > Medium > Low) compared to the I/Rgroup. One-way ANOVA revealed significant intergroup differences(IL-1β:F=387.41, $p<10^{-15}$;IL-6:F=259.57, $p<10^{-14}$; TNF-α: F=143.35, $p<10^{-12}$). Post-hoc Bonferroni tests indicated that the differences between each SFI dose group (Low, Medium, High) and the I/R group were statistically significant (most p<0.001; e. g., for IL-6: SFI-L p=0.028, SFI-M $p=2.2\times10^{-4}$, SFI-H,$p=1.6\times10^{-4}$. Specifically, compared to the I/R group, SFI Low, Medium, and High doses reducedIL-1β levels by 30.7%,48.7%, and 69.4%; The levels of IL-6 were reduced by 18.6%,52.6%,ahd 71.9%; and TNF-αby 27.8%,51.3%, and 68.8%, respectively (Bonferroni correction following one-way ANOVA, all P<0.05). These results suggest that SFI effectively attenuates"I/R-induced amplification of the inflammatory response (See Table 2).

5 DISCUSSION

Myocardial ischemia-reperfusion injury (MIRI) is driven by early MAPK axis activation,which amplifies inflammatory damage. This study demonstrates that Shenfu Injection (SFI)addresses this mechanism. Employing a bioinformatics-to-experimental approach, we found that SFI significantly suppressed I/R-induced activation of the p38/JNK/ERK pathway and dose-dependently reduced key inflammatory cytokines(IL-1β,IL-6, TNF-α) by up to 69.4%, 71.9%,and 68.8%, respectively.

Bioinformatic enrichment analysis highlighted hypoxic stress and innate immunity activation in MIRI, as evidenced by key GO-BP terms including" response to hypoxia" and" granulocyte activation. "[8]; GO-MF terms including" protein serine/ threonine kinase activity" and" transcription factor binding" directly reflected core molecular functions of the MAPK pathway, while GO-CCtérms like" secretory granule lumen" and" mitochondrial matrix" suggested organelle dysfunction and enhanced immune secretion. KEGG analysis not only confirmed the central role of the MAPK signaling pathway but also revealed enriched pathways such as" Apoptosis" and" Lipid and atherosclerosis, " collectively depicting MIRI's pathological network involving energy metabolism disruption, inflammation, and cell death. Notably, the enrichment of neurodegenerative disease pathways may imply share stress-inflammatory mechanisms between MIRI and neurological disorders [9]. Overall, the enrichment profile consistently highlights stress-inflammatory cascades and metabolic/ mitochondrial imbalance as key pathological hubs, which aligns with in vivo observations of MAPK axis activation and its suppression by SFI, providing multi-omics evidence for SFI's cardio-protective role via MAPK pathway modulation.

The observed synchronous inhibition of p38/JNK and reduction in inflammation align with established understanding that early stress pathways drive inflammatory amplification in MIRI. The dose-dependent suppression of pro-inflammatory cytokines by SFI is consistent with its reported anti-inflammatory properties [10]. While some studies suggest that ERK may mediate survival/ repair signals during early re-perfusion [11], while some studies suggest that ERK may mediate survival/ repair signals during early re-perfusion [12]. High-dose Shenfu Injection (SFI)significantly and simultaneously reduced phosphorylation levels of p38(p-p38) and JNK (p-JNK) in myocardial tissue and decreased plasma concentrations of key inflammatory cytokines includingIL-1β, IL-6, and TNF-α in the MIRI model. This indicates that one of the primary mechanisms underlying SFI's cardio-protective effect is through effectively suppressing the excessive activation of the p38 and JNK branches within the stress-responsive MAPK pathway [13-14]; Thereby, SFI interrupts the key upstream circuit of inflammatory signal amplification, preventing the vicious cycle of inflammatory response and subsequent tissue damage [15].

In summary, high-dose SFI treats MIRI by coordinately modulating the MAPK pathway: suppressing the pro-inflammatory p38/JNK branches while preserving pro-repair ERK activity, thereby integrating anti-inflammatory and reparative effects.

6 CONCLUSION

This study demonstrates that SFI alleviates myocardial injury by inhibiting the MAPK pathway and its associated inflammation, providing mechanistic support for its clinical use.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDINFG

The project was supported by Science and Technology Project of Jiangxi Provincial Administration of Traditional Chinese Medicine: Study on MAPK Signaling Pathway of Shenfu Injection in Myocardial Ischemia-Re-perfusion Injury in Rats Based on Data Mining NO: 2022B969.

ACKNOWLEDGMENTS

The texts in this paper are polished and significantly improved by Stork's Writing Assistant.

REFERENCES

[1] Zhao C, Liu T, Wei H, et al. Serum Oxidative Stress Factors Predict Myocardial Ischemia Reperfusion Injury after Percutaneous Coronary Intervention in Patients with Acute Myocardial Infarction and Type 2 Diabetes Mellitus. Postepy Kardiol Interwencyjnej, 2023, 19(04): 333-342.

[2] Zhou Y, Liang Q, Wu X, et al. siRNA Delivery against Myocardial Ischemia Reperfusion Injury Mediated by Reversibly Camouflaged Biomimetic Nanocomplexes. Advanced Materials, 2023, 35(23): e2210691.

[3] Huang P, Wu SP, Wang N, et al. Hydroxysafflor Yellow A Alleviates Cerebral Ischemia Reperfusion Injury by Suppressing Apoptosis via Mitochondrial Permeability Transition Pore. Phytomedicine, 2021, 85(01): 153532.

[4] Lin H, Xiong W, Fu L, et al. Damage-Associated Molecular Patterns (DAMPs) in Diseases: Implications for Therapy. Molecular Biomedicine, 2025, 6(01): 60.

[5] Shi Y, Han B, Yu X, et al. Ginsenoside Rb3 Ameliorates Myocardial Ischemia-Reperfusion Injury in Rats. Pharmaceutical Biology, 2011, 49(09): 900-906.

[6] Yang W, Bai X, Tang Y, et al. miR-379-3p Activates the MAPK/JNK/p38 Pathway by Negatively Targeting DUSP1 Expression to Promote Myocardial Ischemia-Reperfusion Injury. Clinical Hemorheology and Microcirculation, 2025, 90(02): 73-83.

[7] Shi Y, Xu S, Ngoi NYL, et al. PRL-3 Dephosphorylates p38 MAPK to Promote Cell Survival under Stress. Free Radical Biology and Medicine, 2021, 177(01): 72-87.

[8] Zhang L, Jiang Y, Jia W, et al. Modelling Myocardial Ischemia/Reperfusion Injury with Inflammatory Response in Human Ventricular Cardiac Organoids. Cell Proliferation, 2025, 58(03): e13762.

[9] Prem PN, Sivakumar B, Boovarahan SR, et al. Recent Advances in Potential of Fisetin in the Management of Myocardial Ischemia-Reperfusion Injury-A Systematic Review. Phytomedicine, 2022, 101(01): 154123.

[10] Liu X, Ai F, Li H, et al. Anti-Inflammatory Effects of Shenfu Injection against Acute Lung Injury through Inhibiting HMGB1-NF-κB Pathway in a Rat Model of Endotoxin Shock. Evidence-Based Complementary and Alternative Medicine, 2019(01), 9857683.

[11] Qin X, Liu B, Gao F, et al. Gluconolactone Alleviates Myocardial Ischemia/Reperfusion Injury and Arrhythmias via Activating PKCε/Extracellular Signal-Regulated Kinase Signaling. Front Physiol, 2022, 13(01): 856699.

[12] Negm S, Wolf M, Craveiro RB, et al. Resveratrol Modulates Phosphorylation of ERK and AKT in Murine Cementoblasts during in Vitro Orthodontic Compression. BMC Oral Health, 2025, 25(01): 226.

[13] Zhao L, Wu X, Li Q, et al. Inhibition of CCN5 Protects Against Apoptosis and Endoplasmic Reticulum Stress in Bisphenol A-Induced Sertoli Cells via p38/JNK MAPK Signaling Pathway. DNA and Cell Biology, 2025, 44(04): 174-185.

[14] Song D, Liu Y, Yao Y, et al. Melatonin improves bisphenol A-induced cell apoptosis, oxidative stress and autophagy impairment via inhibition of the p38 MAPK signaling pathway in FLK-BLV cells. Environmental Toxicology, 2022, 37(7): 1551-1562.

[15] Doganyigit Z, Eroglu E, Akyuz E. Inflammatory Mediators of Cytokines and Chemokines in Sepsis: From Bench to Bedside. Human & Experimental Toxicology, 2022, 41(01): 9603271221078871.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3051

ANALYSIS OF THE STATUS OF NEWLY REPORTED OCCUPATIONAL DISEASES IN EMPLOYING UNITS IN A CERTAIN CITY FROM 2015 TO 2024

Wei Gao[1*], Yong Zhang[2], Xin Ge[1]
[1]*Huaibei Occupational Disease Prevention and Control Institute, Huaibei 235000, Anhui, China.*
[2]*Huaibei Center for Disease Control and Prevention, Huaibei 235000, Anhui, China.*
**Corresponding Author: Wei Gao*

Abstract: Objective: To analyze the epidemiological characteristics and trends of newly reported occupational diseases among employing units in a certain City from 2015 to 2024, and to provide a scientific basis for developing targeted occupational disease prevention and control strategies. Methods: A retrospective epidemiological investigation was conducted. Data from 469 newly reported occupational disease cases in a certain City between 2015 and 2024 were collected from the occupational disease reporting system. Descriptive epidemiological methods and statistical analyses (chi-square test, analysis of variance) were employed to analyze the distribution characteristics and trends of occupational diseases. Results: Pneumoconiosis was the predominant occupational disease in a certain City, accounting for 95.10% of all cases. Among these, coal workers' pneumoconiosis (316 cases, 67.38%) and silicosis (130 cases, 27.72%) were the most common. The year 2019 marked the peak incidence (89 cases), with a fluctuating downward trend observed in recent years. A significant difference was found in the average length of exposure before onset among different occupational diseases ($F = 11.41$, $P = 0.001$), with coal workers' pneumoconiosis having the longest average exposure period (25.66 ± 6.63 years). Conclusions: While certain achievements have been made in occupational disease prevention and control in a certain City, pneumoconiosis remains the primary concern. Targeted prevention and control measures should be strengthened in high-risk areas and industries, and for specific exposure length groups. Additionally, emerging occupational health risks require ongoing attention.

Keyword: Energy city; Occupational diseases; Incidence trend

1 INTRODUCTION

Occupational diseases refer to diseases caused by exposure to dust, radioactive substances, and other toxic or harmful factors in occupational activities among workers of employers such as enterprises, public institutions, and individual economic organizations [1]. Occupational diseases not only severely harm the health and life safety of workers but also impose a heavy economic burden on families and society, affecting social harmony and stability. In 2023, a total of 10,307 new cases of various occupational diseases were reported in China, including 8,186 cases of pneumoconiosis, accounting for 79.41% of the total new occupational disease cases reported[2, 3].

As an important resource-based city in Anhui Province, a certain city has long taken industries such as coal, electric power, and chemical engineering as its economic pillars[4]. While these industries have promoted the economic development of the city, they have also brought relatively serious occupational disease hazards, and the city bears a heavy burden of occupational diseases[5]. To understand the epidemiological characteristics and changing trends of new occupational diseases among employers in the city from 2015 to 2024, this article systematically sorted out and analyzed the reported data on new occupational diseases among employers in the city over the past decade (2015—2024). The report is as follows:

2 RESEARCH OBJECTS AND METHODS

2.1 Research Objects

All new occupational disease report card information from employers under the jurisdiction of a certain city from 2015 to 2024 was collected as the research objects of this survey.

2.2 Survey Contents

Mainly collect information such as the types of occupational diseases of new occupational disease cases, gender, basic information about employers (location, name, category, enterprise scale, economic type), the age of occupational disease patients at the time of diagnosis, length of exposure to hazards, and year of diagnosis.

2.3 Methods

Occupational diseases were classified according to the national Classification and Catalog of Occupational Diseases (covering 10 categories and 132 types of occupational diseases). The basic situation, trend changes, industry distribution, length of exposure to hazards, age distribution of onset, and other characteristics and trend changes of new occupational diseases in employers in the city over the past 10 years were statistically analyzed.

2.4 Statistical Analysis

Data were input, organized, and processed using EXCEL, and statistically analyzed using SPSS 24.0 software. For data that passed the normality test and conformed to a normal distribution, they were described using mean and standard deviation. One-way ANOVA was used for comparison between groups; chi-square test was used for comparison of rates between groups; the test level α=0.05 (two-sided).

3 RESULTS

3.1 General Situation of New Occupational Diseases

In 2024, the number of workers exposed to different occupational hazard factors in a certain city was as follows: 12,640 for coal dust, 9,768 for silica dust, 382 for cement dust, 1,276 for welding fumes, 168 for benzene, 1,295 for lead, and 12,251 for noise.

The distribution and characteristics of different occupational disease cases in a certain City over the past decade are presented in Table 1.

Over the past decade, employers in the city reported a total of 469 occupational disease cases, including 526 males (99.6%) and 2 females (0.4%). Occupational diseases were mainly dominated by pneumoconiosis. Among them, coal worker's pneumoconiosis had the largest number of cases, with 316 cases accounting for 67.38% of the total. The average age of patients was (51.87±8.45) years, and the average length of exposure to hazards was (25.20±6.64) years. Silicosis ranked second, with 130 cases accounting for 27.72%. The average age of patients was (51.32±10.17) years, and the average length of exposure to hazards was (21.95±8.73) years. Other occupational diseases such as noise-induced deafness (9 cases, 1.92%) and welder's pneumoconiosis (5 cases, 1.07%) had relatively fewer cases. In addition, it also included heat stroke (3 cases) and 6 other sporadic cases (1 case each of other pneumoconioses, dimethylformamide poisoning, cement pneumoconiosis, contact dermatitis, brucellosis, and bursitis (limited to underground workers)). The length of exposure to hazards for the patient with dimethylformamide poisoning was 0 years, suggesting acute chemical poisoning (Table 1).

Table 1 Distribution and Characteristics of Different Occupational Disease Cases

Name of Occupational Disease	Number of Cases n (%)	Age at Onset (years)			Length of Exposure to Hazards (years)		
		Mean ± Standard Deviation	Minimum	Maximum	Mean ± Standard Deviation	Minimum	Maximum
Coal Worker's Pneumoconiosis	316（67.38）	51.87±8.45	34	81	25.20±6.64	8	40
Silicosis	130（27.72）	51.32±10.17	36	82	21.95±8.73	4	37
Noise-induced Deafness	9（1.92）	53.67±1.50	51	56	18.22±9.05	8	36
Welder's Pneumoconiosis	5（1.70）	47.20±4.44	43	52	13.60±10.26	5	31
Heat Stroke	3（0.64）	52.00±13.75	40	67	15.67±14.01	0	27
Other Pneumoconioses	1	49.70			30.08		
Cement Pneumoconiosis	1	73.43			26.92		
Dimethylformamide Poisoning	1	50.25			0.00		
Brucellosis	1	29.58			1.17		
Contact Dermatitis	1	53.54			1.17		
Bursitis (limited to underground workers)	1	55.08			28.50		

3.2 Trends in New Occupational Disease Incidence

This study systematically analyzed the incidence trends of occupational diseases in the city from 2015 to 2024. Data showed that pneumoconioses (including coal worker's pneumoconiosis, silicosis, and other pneumoconioses) were the main type of occupational diseases. The distribution of newly reported occupational diseases across different years is summarized in Table 2.Among them, coal worker's pneumoconiosis peaked in 2019 with 64 cases, accounting for 71.9% of the total occupational diseases that year. Notably, the total number of occupational disease cases reached 89 in 2019, the highest in the decade, mainly due to a significant increase in coal worker's pneumoconiosis and silicosis cases (Table 2).

In the long term, coal worker's pneumoconiosis showed a fluctuating downward trend, decreasing from 46 cases in 2015 to 12 cases in 2024, but with a temporary rebound in 2023 (26 cases). Silicosis cases declined after peaking at 24 cases in 2019, with 5 cases reported in 2024. Only 1 case of other pneumoconioses was reported in 2021.

In recent years, new occupational diseases have shown a diversification trend: noise-induced deafness has occurred continuously since 2021 (2 cases in 2021, 2 in 2022, 2 in 2023, and 3 in 2024), indicating that the risk of noise exposure deserves attention. Heat stroke occurred in 3 cases from 2021 to 2022, 1 case of dimethylformamide poisoning was reported in 2022, and brucellosis and contact dermatitis were first reported in 2024 (1 case each). In addition, 1 case of bursitis (limited to underground workers) was reported in 2023.

Figure 1 presents the trend in occupational disease incidence in a certain city over the past decade.The total number of new occupational disease cases fluctuated downward from 63 in 2015 to 22 in 2024, but rebounded to 32 cases in 2023. 2017 and 2020 were the lowest (38 cases) and second lowest (54 cases) points in the decade, respectively. It is worth noting that from 2021 to 2024, although the dominant position of pneumoconioses weakened, the types of occupational diseases increased significantly—from 4 types in 2021 to 5 types in 2024 .

Table 2 Incidence of New Occupational Diseases in Different Years

Name of Occupational Disease	Incidence Year (Number of Cases)										Total	Proportion (%)
	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024		
Coal Worker's Pneumoconiosis	46	39	25	29	64	41	19	19	26	12	320	68.2
Silicosis	17	18	12	20	24	13	16	2	3	5	134	28.6
Noise-induced Deafness	0	0	0	0	0	0	0	2	2	3	7	1.5
Welder's Pneumoconiosis	0	1	0	1	1	0	2	0	0	0	5	1.1
Heat Stroke	0	0	0	0	0	0	0	2	1	0	3	0.6
Other Occupational Diseases*	0	0	1	0	0	0	0	1	1	2	5	1.1
Total	63	58	38	50	89	54	42	21	32	22	469	100.0

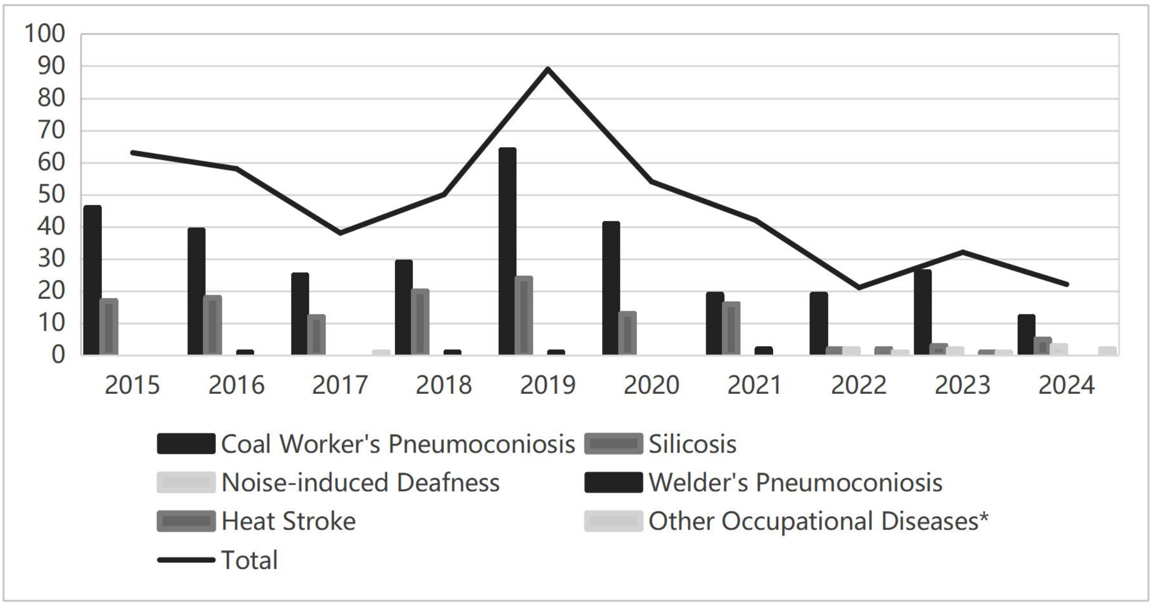


Figure 1 Incidence Trend of Occupational Diseases in a Certain City Over the Past Decade

3.3 Age and Length of Exposure to Hazards at Onset of Key Occupational Diseases

Table 3 presents the characteristics of exposure time and age at onset across different key new occupational diseases.Significant differences were found in the lengths of exposure to hazards at onset among different key occupational diseases in the city (F=11.41, P=0.001<0.05). Among them, coal worker's pneumoconiosis had the longest average length of exposure to hazards at onset (25.66±6.63 years), followed by silicosis (22.41±8.70 years), noise-induced deafness (18.55±8.91 years), and welder's pneumoconiosis (14.08±9.99 years). The length of exposure to hazards at onset ranged from 4.08 to 40.92 years. Welder's pneumoconiosis had the shortest average length (14.08±9.99 years); however, due to the small number of cases and a large standard deviation, significant individual differences were observed. The average diagnostic age for all occupational diseases was 52.15±8.85 years. No statistically significant difference was found in diagnostic age among different types of occupational diseases (F=0.689, P=0.559) (Table 3)

Table 3 Statistics on Length of Exposure to Hazards and Age at Onset of Different Key New Occupational Diseases

Type of Occupational Disease	Number of Cases	Length of Exposure to Hazards (years/Mean ± SD)	Minimum	Maximum	Diagnostic Age (years/Mean ± SD)	Minimum	Maximum
Coal Worker's Pneumoconiosis	316	25.66±6.63	8.75	40.92	52.32±8.43	34.95	81.85
Silicosis	130	22.41±8.70	4.08	37.00	51.77±10.15	36.36	82.69
Welder's Pneumoconiosis	5	14.08±9.99	5.92	31.17	47.66±4.39	43.21	52.68
Noise-induced Deafness	9	18.55±8.91	8.33	36.25	54.10±1.43	51.26	56.28
F		11.41			0.689		
P		0.002<0.001			0.559>0.05		
Total	460	24.48±7.60	4.08	40.92	52.15±8.85	34.95	82.69

4 DISCUSSION

The results of this study showed that pneumoconioses were the dominant occupational diseases in the city, accounting for 95.10% of the total cases. This is highly consistent with the industrial structure of the region—where the coal industry serves as the economic pillar—and aligns with relevant reports on Anhui Province as a whole. As an important resource-based city in Anhui, the city has long relied on industries such as coal, electric power, and chemical engineering. The production processes of these industries involve significant dust hazards, which explains why pneumoconioses have become the most prevalent type of occupational disease.

An analysis of the incidence trend from 2015 to 2024 revealed that 2019 was the peak year for occupational diseases (89 cases), followed by a fluctuating downward trend, with the number dropping to 22 cases in 2024. This trend is generally consistent with national reports[3]. Multiple factors may have contributed to this pattern: First, the 2019 peak may be linked to the active production activities in the coal industry at that time, as well as changes in diagnostic criteria and requirements. With industrial structure adjustment and tighter environmental policies, coal production capacity has been controlled—an important reason for the decline in incidence after 2020. Second, strengthened occupational disease prevention and control efforts (e.g., controlling workplace dust concentrations, improving workers' protection awareness, and popularizing occupational health examinations) have also played a positive role. An increasing number of enterprises now prioritize post protection and employee health record management, which has driven the overall downward trend in incidence.

Notably, although the number of pneumoconiosis cases has generally decreased, the types of occupational diseases have diversified—with non-traditional conditions such as noise-induced deafness, heat stroke, and dimethylformamide poisoning beginning to emerge. This change reflects the effect of the city's industrial structure adjustment and suggests that occupational disease prevention and control must keep pace with the times, focusing on emerging occupational hazards[6].

An analysis of the length of exposure to hazards at onset of key occupational diseases showed significant differences among different disease types (F=11.41, *P*=0.001). Coal worker's pneumoconiosis had the longest average length of exposure at onset (25.66±6.63 years), followed by silicosis (22.41±8.70 years), while welder's pneumoconiosis had a relatively shorter length (14.08±9.99 years)—consistent with relevant reports[7]. This difference reflects the distinct pathogenesis and latency characteristics of different occupational diseases. As a chronic cumulative disease, pneumoconioses require long-term exposure to develop; the 21–30-year exposure segment was the peak period for onset (225 cases, 48.91%), which aligns with the long latency of pneumoconioses and matches findings from related surveys[8]. From a preventive perspective, this analysis of exposure length provides a basis for identifying key monitoring populations: Workers with more than 10 years of exposure—especially those with 21–30 years of exposure—should be the focus of occupational health monitoring.

5 CONCLUSION

The epidemiological analysis of occupational diseases in the city from 2015 to 2024 indicates that pneumoconioses are the most prevalent occupational diseases, closely linked to the city's resource-based economic structure. The results of this study have important guiding significance for occupational disease prevention and control. It is recommended to:Adopt targeted prevention and control strategies and strengthen pneumoconiosis prevention;Pay attention to emerging occupational health risks;Optimize resource allocation; Establish a key monitoring mechanism based on the length of exposure to comprehensively improve the effectiveness of occupational disease prevention and control.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCE

[1] Wang B, Wu C, Kang L, et al. What are the new challenges, goals, and tasks of occupational health in China's Thirteenth Five-Year Plan (13th FYP) period?. Journal of Occupational Healt, 2018, 60(3): 208-228.

[2] Zhong X, Zeng Y, Peng L, et al. Levels and related factors of occupational stress among nurses: hospital-based evidence from China, 2023. Frontiers in Psychology, 2024, 15: 1471640.

[3] Jia N, Wang Z, Zhang M, et al. Prevalence and Risk Factors of Lower Extremity Musculoskeletal Disorders Among Occupational Groups in Key Industries - China, 2018-2023. China CDC Weekly, 2024, 6(52): 1388-1395.

[4] Li H, Long R, Chen H. Economic transition policies in Chinese resource-based cities: An overview of government efforts. Energy Policy, 2013, 55: 251-260.

[5] Zheng J N, Yu Y, Zhang S Y, et al. Analysis of the status of occupational disease prevention and control of enterprises in mining and manufacturing industries in China in 2019. Chinese Journal of Industrial Hygiene and Occupational Diseases, 2023, 41(5): 338-344.

[6] Zhou S, Huang Y, Wen X, et al. Epidemiological Characteristics and Diagnostic Outcomes of Suspected Occupational Noise-Induced Deafness - Guangdong Province, China, 2014-2023. China CDC Weekly, 2024, 6(52): 1381-1387.

[7] Wyman A E, Hines S E. Update on metal-induced occupational lung disease. Current Opinion in Allergy and Clinical Immunology, 2018, 18(2): 73-79.

[8] Lee J, Bambrick M, Lau A, et al. Computed Tomography of Contemporary Occupational Lung Disease: A Pictorial Review. Diagnostics (Basel), 2024, 14(16).

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3052

THE ANIMAL MODEL ESTABLISHMENT METHODS AND MODEL EVALUATION FOR LIVER STAGNATION AND SPLEEN DEFICIENCY SYNDROME

Feng Wang[1], HuiYong Zhang[2*], GuanLin Yang[2]
[1]Liaoning University of Traditional Chinese Medicine, Shenyang 110847, Liaoning, China.
[2]Affiliated Hospital of Liaoning University of Traditional Chinese Medicine, Shenyang 110032, Liaoning, China.
Corresponding Author: HuiYong Zhang, Email: a997_0815@qq.com

Abstract: The fast pace of modern life has increased stress, leading to a faster incidence of liver qi stagnation and spleen deficiency syndrome in diseases. Therefore, liver qi stagnation and spleen deficiency syndrome has become a common clinical syndrome. Animal models of liver qi stagnation and spleen deficiency syndrome are used in various disease areas, and numerous methods exist for evaluating these models, resulting in problems such as confusion regarding counter-evidence formulas. This article mainly evaluates the modeling methods, behavioral testing methods, and counter-evidence formulas for liver qi stagnation and spleen deficiency syndrome in various diseases, providing theoretical support for experimental research on this syndrome.

Keywords: Liver qi stagnation and spleen deficiency syndrome; Animal model; Counter-evidence formula; Behavioral science; Tail clamping; Restraint

1 INTRODUCTION

With the increasingly rapid pace of modern life and rising levels of stress in both work and study, the incidence of liver stagnation and spleen deficiency syndrome has shown a marked upward trend. As clinical research must be grounded in solid basic research, it is essential to establish reliable animal models of liver stagnation and spleen deficiency syndrome and translate the findings into clinical practice. Since their initial application in digestive system disorders such as gastritis and liver cancer, these animal models have gradually been extended to a broader range of diseases, including tic disorder and depression [1,2]. Over more than 30 years of continuous refinement and innovation, the modeling system has become increasingly sophisticated; however, challenges remain—particularly regarding the accuracy of behavioral indicators and the lack of standardized medicinal prescriptions following model establishment. To gain a clearer understanding of current research progress, this study retrieved literature from CNKI published before February 1, 2020, using the topic search terms “liver stagnation and spleen deficiency + animal model.” A total of 267 relevant articles were identified, from which 26 representative studies were selected for detailed review and discussion.

2 APPLICATION OF ANIMAL MODELS OF LIVER STAGNATION AND SPLEEN DEFICIENCY SYNDROME IN VARIOUS DISEASES

Animal models of liver stagnation and spleen deficiency syndrome have been widely used across multiple disease areas. The selection of experimental animals varies considerably in terms of species, weight, and age, while the primary modeling methods include tail-clamping stimulation, restraint stress, and combined modeling approaches. Modeling duration, detection indicators, and observable macroscopic features differ according to disease type. These key characteristics and their applications are summarized in Table 1.

Table 1 Application of Liver Qi Stagnation and Spleen Deficiency Syndrome in Various Systems

Disease	animal models	Weight and age	Modeling methods	Molding time	detection indicators	Macroscopic representation
Irritable bowel syndrome [3]	SD pregnant mice Pregnancy duration: 18.2 ± days	-	Mother-infant separation	9 weeks	D-xylose, serotonin (5-HT) , brain-derived neurotrophic factor, immunoglobulin A, immunoglobulin G	Reduced activity, sluggish response, and clustering
Fatty liver [4]	Wistar male rats	160-200 g	20% ethanol solution administered by gavage	2 weeks	Alanine aminotransferase (ALT), aspartate aminotransferase (AST), gamma-glutamyl transferase, total cholesterol (GGT), triglycerides, low-density lipoprotein cholesterol, and high-density lipoprotein cholesterol.	Irritability, decreased activity, loss of appetite, yellowish coat, loose stools, and a few bruises on the tail.
Liver injury [5]	Wistar male rats	180-220 g	Gavage with ethanol	1 week	ALT, AST, prealbumin, total bilirubin, total bile acids	Laziness, weight gain
Duck hepatitis B virus infection [6]	Guangzhou Ma Duck	40-50 g 1 day old	Restraint, dietary disorder	2 weeks	duck hepatitis B virus	listless, with sparse downy hair, irritable, and lacking appetite.
Liver cancer [7]	Wistar rats	100-200 g	Restraint, dietary disorder	9 weeks	T lymphocytes, inflammatory cytokine-2 (IL-2) , and tumor necrosis factor -γ (TNF-γ)	Irritability, anger, weakness, lack of exercise, loose stools, dry and yellow hair
Depression [8]	SD rats, male	180-200 g 6-8 months	Restraint, tilted cage, swimming, tail clamping, shaking the cage, water restriction, and odor stimulation.	7 weeks	IL-6, TNF-α	Reduced appetite and decreased physical condition
Nervous system [9]	Wistar male rats	240 ±20g	Electrical stimulation	9 days	Hypothalamic CRF content, locus coeruleus CRF content	Avoidance, reduced activity, irritability
Tourette syndrome [10]	SD rats, male	40-60g 3 weeks old	Restraint, swimming, and dietary indiscretion	3 weeks	Gastrin, D-xylosin	Decreased appetite, low mood, increased bowel movements
Tinnitus [11]	Wistar male rats	222-238 g 6-8 weeks old	Restraint, swimming, and dietary indiscretion	20 days	False positive reaction	Disheveled hair, slow movement, huddling together, and closed eyes
Chronic pharyngitis [12]	Wistar male rats	180-220 g	Restraint, dietary disorder	3 weeks	5-HT1B, 5-HT2C	Loose hair, occasional loose stools, low activity level, and dirty anus.

3 BEHAVIORAL TESTING METHODS

3.1 Open Field Experiment

Yue Lifeng et al. constructed a custom open-field apparatus measuring 1 m × 1 m × 40 cm, with a camera positioned at the center of the base. All experiments were conducted in a dark, enclosed environment. Each rat was placed in the central area of the box, and its behavior was recorded and timed simultaneously. The animal’s activity was observed for 5 minutes. After each trial, the rat was removed and the apparatus was thoroughly cleaned to eliminate any residual odor that might influence subsequent observations. Behavioral indicators included the total number of grid crossings, the frequency of rearing, and the frequency of grooming behaviors (such as scratching, washing the face, and licking the paws) [13].

Similarly, Zhao Bo et al. designed an open-field box measuring 122 cm in length and 45 cm in height. The bottom was divided into 16 equal-area squares, and each rat was placed at the center of the apparatus. The number of squares crossed served as the activity score, and the behavioral indicators were consistent with those used by Yue Lifeng et al.

[14].

3.2 Sucrose Preference Test

Sucrose preference percentage is widely recognized as an objective and reliable indicator for evaluating anhedonia and reduced reward sensitivity. Liu Yuan et al. fasted rats for 24 hours and then provided each animal with 200 mL of 1% sucrose solution and an equal volume of sterile water in a quiet environment. After 24 hours of free drinking, the consumption of each liquid was recorded, and the sucrose preference percentage was calculated [15]. Liu Yueyun et al. conducted a similar assessment but first trained the rats to develop a preference for 1% sucrose solution for two consecutive days, beginning three days before the completion of modeling. On the third day, following a fasting period, the rats were allowed access to both sucrose water and pure water for 1 hour, after which the intake of each was measured to determine sucrose preference [16].

3.3 Forced Swimming Experiment

Zhou Guoer et al. conducted the forced swimming test by placing mice in a container filled with 25-cm-deep water maintained at 25 °C. Each mouse was allowed to swim for 6 minutes. Immobility was defined as the state in which the mouse floated with minimal limb movement and ceased active struggling for at least 3 seconds. The cumulative immobility time during the final 4 minutes of the test was recorded and used as the primary behavioral indicator [17].

3.4 Elevated Plus Maze Test

The elevated plus maze was positioned 50 cm above the ground and consisted of two open arms, two closed arms, and a central platform, with dimensions of 45 cm × 15 cm for the open arms, 45 cm × 15 cm × 30 cm for the closed arms, and 15 cm × 15 cm for the central platform. The experiment was conducted in a quiet environment, and the maze was kept dark prior to behavioral recording. Each rat was placed on the central platform facing one of the closed arms, and its activity was monitored for 5 minutes. Behavioral indicators included the number of entries into the open arms and the total time spent in these arms. After each session, the apparatus was thoroughly cleaned to remove odor cues [18].

3.5 Tail Suspension Test

Zhang Beihua et al. performed the tail suspension test in a quiet, disturbance-free environment. Each rat was suspended by its tail and allowed to hang freely, and the duration of complete immobility within a 5-minute observation period was recorded as the primary behavioral indicator [19].

4 COUNTER-EVIDENCE PRESCRIPTIONS

4.1 Xiaoyao Pills

After artificial modeling, the solution was adjusted to a concentration of mL^{-1}, and each rat received an oral dose of $3.24\ mL \cdot kg^{-1} \cdot d^{-1}$. Following administration, the liver volume decreased and the tissue became softer compared with untreated rats exhibiting liver stagnation and spleen deficiency. In addition, Xiaoyao Pills were shown to improve blood lipid profiles and alleviate disturbances in glucose and amino acid metabolism in rats [20].

4.2 Tongxie Yaofang

After establishing the model, Wang Min et al. administered Tongxie Yaofang to rats by gavage at doses equivalent to 4, 8, and 16 times the adult dosage. Following 21 days of treatment, rats in the Tongxie Yaofang groups showed a marked attenuation of body-weight loss and a significant reduction in gastric mucosal injury associated with liver stagnation and spleen deficiency syndrome [21].

4.3 Chaihu Shugan San Combined with Sijunzi Tang (Chai Shu Si Jun Zi Tang)

After establishing the model, Li Cong et al. administered the drug to rats at a dose of 3.57 g/kg for two weeks. The results indicated that Chai Shu Si Jun Zi Tang exerted a noticeable ameliorative effect on the abnormal hemorheology observed in rats with liver stagnation and spleen deficiency syndrome [22].

5 CONCLUSION

Animal models of liver stagnation and spleen deficiency syndrome are primarily established using rats and mice, with few studies employing large animals. As a result, key diagnostic features in traditional Chinese medicine (TCM)—such as tongue appearance, pulse, and facial expression—cannot be observed. In addition, inconsistencies in the sex, age, and body weight of experimental animals may influence modeling outcomes [23,24]. Therefore, when replicating this model, the selection of animals should align with the experimental objectives, considering anatomical and physiological characteristics, and their age and weight should be standardized.

A second challenge is the lack of standardized criteria for evaluating model-inducing stimuli. Techniques such as restraint stress and tail clamping not only lack quantifiable stimulation intensity but may also cause physical injury, thereby influencing subsequent behavioral assessments. Moreover, after modeling, distinguishing between liver stagnation alone and liver stagnation combined with spleen deficiency is often difficult.

From a modern biomedical perspective, animal models of liver stagnation and spleen deficiency involve functional disturbances across multiple systems—including the nervous, digestive, and immune systems—which complicates the identification of specific biomarkers. Beyond commonly measured indicators such as AST, ALT, gastrin, and 5-HT, physicochemical markers reported in the literature remain inconsistent. During model evaluation, syndrome classification relies heavily on macroscopic signs; however, liver qi stagnation and spleen qi deficiency represent two major TCM syndrome groups that often overlap. Furthermore, differentiating spleen qi deficiency from spleen yang deficiency is challenging, leading to confusion and variability in macroscopic assessments after modeling [25].

Counter-evidence prescriptions play a crucial role in validating whether a model has been successfully established. Yet, current practice remains problematic: the selection of formulas is often arbitrary, the variety is overly broad, and in many cases, counter-evidence prescriptions are not used at all after modeling. This approach deviates from the TCM principle of integrating theory, method, formula, and medicine as a unified system [26]. Additionally, behavioral testing methods vary widely in terms of environmental requirements and apparatus specifications, raising concerns about the reliability and comparability of experimental results. Therefore, it is essential to standardize both counter-evidence prescriptions and behavioral testing methods to ensure the accurate and scientifically sound verification of model success.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Lü Aiping, Chen Xiaoye, Zou Shijie, et al. Experimental study on spleen deficiency and liver stagnation type gastritis in rats. Chinese Journal of Traditional Chinese Medicine Science and Technology, 1994, 11-14+4.

[2] Peng Guiying, Gu Ligang, Wang Qingguo, et al. Effects of Liver Qi Stagnation and Spleen Deficiency on Immune Function in DEN-Induced Liver Cancer Rats and the Preventive and Treatment Effects of Liver-Soothing and Spleen-Strengthening Formulas. Dissertation, Editorial Department of Journals of China Association of Traditional Chinese Medicine, 2002.

[3] Zhang Beihua. Establishment and evaluation of a rat model of IBS-D with liver stagnation and spleen deficiency syndrome. Conference Paper, China Academy of Chinese Medical Sciences, 2013.

[4] Sun Xiaoqian, Sun Rong. A model of fatty liver with liver stagnation and spleen deficiency in rats. Chinese Journal of Experimental Traditional Medical Formulae, 2017, 23, 100-106.

[5] Huang Nana, Sun Xiaoqian, Yang Qian, et al. Study on the combined pathogenesis and syndrome model of acute liver injury (liver stagnation and spleen deficiency syndrome) in rats. Chinese Journal of Traditional Chinese Medicine, 2016, 41, 124-131.

[6] Luo Huanhuan. Study on the intervention of liver-soothing and spleen-strengthening method in a duck animal model of hepatitis B virus infection with liver stagnation and spleen deficiency syndrome. New Chinese Medicine, 2010, 42(432): 107-109.

[7] Peng Guiying, Gu Ligang, Wang Qingguo, et al. Effects of liver stagnation and spleen deficiency on immune function in rats with DEN-induced liver cancer. Journal of Beijing University of Chinese Medicine, 2002, 37-40.

[8] Wan Xiaomin, Zhou Rong, Huang Xintong, et al. Effects of Xiaoyao San on IL-6 and TNF-α in the hippocampus of rats with chronic unpredictable mild stress and liver stagnation and spleen deficiency type depression. Journal of Liaoning University of Traditional Chinese Medicine, 2020, 22(190): 53-57.

[9] Cai Gan, Zhang Yuxi, Liu Qun, et al. Study on changes in CRF content in hypothalamus and locus coeruleus in patients with liver stagnation syndrome, spleen deficiency syndrome, and liver stagnation and spleen deficiency syndrome. Shanghai Journal of Traditional Chinese Medicine, 2006, 5-7.

[10] Liu Xiaofang, Xue Xiaona, Wang Sumei, et al. Establishment and evaluation of rat model of multiple tic disorder with liver stagnation and spleen deficiency syndrome. World Journal of Traditional Chinese Medicine, 2018, 13: 166-169.

[11] Li Jinfei, Jia Zhijiao, Ding Lei, et al. A rat model of tinnitus with liver stagnation and spleen deficiency syndrome was established by a combined etiology method. Chinese Journal of Integrated Traditional and Western Medicine on Otorhinolaryngology, 2019, 27(144): 63-65+67.

[12] Wang Yi, Zhou Jiaxuan, Huang Chunjiang, et al. Study on the establishment of a rat model of chronic pharyngitis with liver stagnation and spleen deficiency based on the theory of “emotional pathogenesis”. Journal of Yunnan University of Traditional Chinese Medicine, 2018, 41(185): 15-19.

[13] Yue Lifeng, Wang Zonghua, Chen Jiaxu, et al. Behavioral changes and regulatory effects of Xiaoyao San on bilateral amygdala injection in rats with liver stagnation and spleen deficiency syndrome. Chinese Journal of Traditional Chinese Medicine, 2011, 26, 90-93.

[14] Zhao Bo, Chai Li, Wu Damei, et al. Experimental study on replicating animal model of liver qi stagnation and spleen deficiency syndrome using a compound multifactorial method. Journal of Chengdu University of Traditional Chinese Medicine, 2013, 36, 12-16+25.

[15] Liu Yuan, Tang Hongmei, Zhong Rufan, et al. Study on the effects and mechanisms of Jianpi Tongfu Granules on rat model of constipation-predominant irritable bowel syndrome with liver stagnation and spleen deficiency. Chinese Journal of Traditional Chinese Medicine, 2018, 36, 56-60.

[16] Liu Yueyun, Guo Xiaoling, Zhao Hongbo, et al. Behavioral evaluation of the efficacy of Xiaoyao San extract on rats with liver stagnation and spleen deficiency syndrome. Chinese Journal of Traditional Chinese Medicine, 2013, 28, 222-225.

[17] Zhou Guoer, Wu Jing, Huang Yunjuan, et al. Preliminary study on the establishment of animal models of depression with “liver qi stagnation” and “liver qi stagnation and spleen deficiency”. Chinese Journal of Traditional Chinese Medicine, 2014, 32, 77-80.

[18] Jin Zhongye, Li Na, Zhao Hongbo, et al. Effects of Xiaoyao San on serum IL-1β, IL-6 and TNF-α in rats with chronic restraint stress and liver stagnation and spleen deficiency anxiety model. Chinese Journal of Traditional Medicine, 2016, 31(217): 59-63.

[19] Zhang Beihua. Establishment and evaluation of a rat model of IBS-D liver stagnation and spleen deficiency syndrome. China Academy of Chinese Medical Sciences, 2013.

[20] Guan Wei, Li Ruoyu, Guo Jilong, et al. Analysis of the mechanism of action of Xiaoyao Wan in rats with non-alcoholic fatty liver disease and liver stagnation and spleen deficiency based on ~1H-NMR metabolomics technology. Chinese Journal of Experimental Traditional Medical Formulae, 2018, 24, 107-113.

[21] Wang Min, Liu Jiemin, Chen Ling, et al. Study on the therapeutic effect and mechanism of Tongxie Yaofang on gastric ulcer of liver stagnation and spleen deficiency in rats. Shizhen Guoyi Guoyao, 2016, 27(254): 63-67.

[22] Li Cong, Xie Ming, Zhao Ronghua, et al. Hemorheological changes and effects of Shugan Jianpi Decoction on rats with different syndromes of liver stagnation-spleen deficiency. Journal of Guangzhou University of Chinese Medicine, 2014, 31(140): 66-70.

[23] Xiao Chengyu, Zhang Huiyong, Yang Guanlin, et al. Application of animal behavior in medical experimental research. Journal of Liaoning University of Traditional Chinese Medicine, 2017, 19(155): 52-56.

[24] Zhai Guoyuan, Qin Jian, Liu Yang, et al. Laboratory animals and their application in scientific research. Gansu Animal Husbandry and Veterinary Medicine, 2018, 48(306): 16-18.

[25] Li Qin, Zhang Huiyong, Wu Tianshi, et al. Research progress on animal model establishment methods and model evaluation of spleen yang deficiency syndrome. World Science and Technology - Modernization of Traditional Chinese Medicine, 2015, 17, 141-148.

[26] Wu Tianshi, Zhang Huiyong, Zhang Zhe, et al. A review of animal modeling methods for spleen deficiency syndrome. Journal of Traditional Chinese Medicine, 2015, 56, 88-93.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3053

REGRESSION ANALYSIS-BASED INVESTIGATION OF FACTORS INFLUENCING MALE FETAL Y CHROMOSOME CONCENTRATION AND STRATIFIED OPTIMIZATION OF OPTIMAL TIMING FOR NON-INVASIVE PRENATAL TESTING

Ting Li
Mathematics and Computer Science, Yan'an University, Yan'an 716000, Shaanxi, China.

Abstract: This study systematically analyzed the quantitative relationship between male fetal Y chromosome concentration and maternal gestational age and body mass index, aiming to optimize the timing for non-invasive prenatal testing. To investigate these relationships, gestational age data formats were standardized, and missing BMI values were calculated and imputed. Pearson and Spearman correlation analyses revealed a weak positive correlation between Weeks and, and a weak negative correlation with BMI. Further OLS regression modeling revealed that Weeks exerted a significant positive effect on Y and BMI a significant negative effect , both statistically significant. Although the linear model exhibited low explanatory power and heteroscedasticity, incorporating LOWESS nonlinear smoothing clearly revealed a “rapid-then-slow” increase pattern of Y with Weeks. This confirmed the clinical observation that Y significantly declines when BMI exceeds 30. Subsequently, to determine the optimal testing window, pregnant women were stratified into four BMI groups. Intra-group OLS regression and LOWESS fitting were performed, supplemented by sensitivity analysis using 3.5%, 4.0%, and 4.5% thresholds. Results indicated that detection was feasible around 11–12 weeks for low-BMI women; while high-BMI pregnant women required delaying until approximately 12.5 weeks to achieve stable detection rates. This study provides quantitative evidence for optimizing NIPT testing strategies in clinical practice.

Keywords: Non-invasive prenatal testing; Regression analysis; Optimal testing timing; Y Chromosome Concentration

1 INTRODUCTION

With continuous advances in medical technology, non-invasive prenatal testing (NIPT) has become a vital tool for assessing fetal health. By analyzing fetal cell-free DNA fragments from maternal blood samples, NIPT enables early detection of fetal chromosomal abnormalities, thereby promoting the healthcare system's shift toward a “prevention-first” approach. The accuracy of NIPT technology primarily relies on analyzing fetal Y chromosome concentration, which is significantly influenced by multiple factors including gestational age and body mass index (BMI). Clinical experience generally suggests that Y increases with gestational age, but high BMI reduces detection sensitivity, leading to lower Y[1-2]. However, whether this relationship is strictly linear and how to determine the optimal testing time for different pregnant women based on the nonlinear interactions of these factors remain critical unresolved issues. This study aims to establish a mathematical model that accurately quantifies the relationship between Y and factors like gestational age and BMI. Through BMI grouping and timing sensitivity analysis, it seeks to determine optimal non-invasive prenatal testing timepoints for pregnant women of different body types, thereby enhancing test accuracy and clinical feasibility. The primary innovations of this section are: 1. Employing a combined strategy of correlation analysis, OLS regression, and LOWESS nonparametric smoothing to balance interpretability and flexibility, comprehensively revealing the linear and nonlinear influence patterns of gestational age and BMI on Y. 2. Addressing the Y threshold issue, we introduced BMI grouping and threshold sensitivity analysis modeling strategies to evaluate BMI's impact on optimal testing timing and quantify measurement error uncertainty[3-4]. 3. It is explicitly demonstrated that LOWESS nonlinear fitting better reflects the true trend of Y as a function of Weeks and BMI, avoiding the failure of linear regression in specific BMI groups. This section's research plan comprises two phases: First, OLS regression and LOWESS smoothing were used to explore the quantitative relationship between Y and gestational weeks/BMI. Second, pregnant women were grouped into four BMI categories, and grouped OLS regression, LOWESS fitting, and threshold sensitivity analysis were employed to determine and optimize the optimal NIPT detection timing for each BMI group[5].

2 MODELING THE ASSOCIATION BETWEEN Y CHROMOSOME CONCENTRATION AND GESTATIONAL AGE/BMI AND ANALYSIS OF NONLINEAR TRENDS

In Non-invasive Prenatal Testing (NIPT), fetal Y-chromosome concentration (Y_conc) is a crucial indicator for determining test effectiveness. Clinical experience generally holds that: as gestational age (Weeks) increases, Y-chromosome concentration gradually rises; while in pregnant women with a high Body Mass Index (BMI), test sensitivity decreases and Y-chromosome concentration is relatively low.
However, it is questionable whether a simple linear model can capture this relationship:

(1) The relationship between gestational age and concentration may not be strictly linear: Y-concentration rises rapidly in early gestational weeks, then slows down in the second trimester, possibly exhibiting diminishing marginal effects.
(2) The effect of BMI may have a threshold effect: it has little impact within the normal range (BMI < 28); but becomes significantly enhanced in obese pregnant women (BMI ≥ 30).
(3) Sequencing errors and individual differences: cfDNA fragmentation, sequencing depth, and individual differences among pregnant women may introduce random noise, leading to low linear correlation coefficients.
Based on the above considerations, we cannot rely solely on Pearson correlation or simple OLS regression, but should adopt a combined strategy of correlation analysis + OLS regression + LOWESS smoothing:
Correlation analysis is used to test the overall direction and significance[6-7];
(1) OLS regression is used to quantify the marginal effects of Weeks and BMI and conduct statistical tests;
(2) LOWESS smoothing is used to reveal potential non-linear patterns as a supplementary validation of the model.

2.1 Data Preprocessing

After reading the original table, we conducted a preliminary inspection of the data and identified several issues that need to be addressed:
(1) Inconsistent formats of gestational age field: Forms such as "12w+3" and "12w" exist, which will cause numerical calculation errors if not unified.
Processing: Write a parsing function to unify these representations into floating-point weeks (e.g., 11w+6 → 11 + 6/7 weeks) and check for records that fail to parse;
(2) BMI may be missing or need to be calculated: Some records lack the pregnant woman's BMI field but have height/weight fields.
Processing: If missing, calculate BMI using the formula:

$$\mathrm{BMI} = \frac{\mathrm{Weight}}{(\mathrm{Height}/100)^2} \tag{1}$$

(3) Missing value handling and sample screening: Exclude samples with complete missing key variables; for a small number of missing values, median/mean imputation can be used (or the impact of imputation on results can be tested in sensitivity experiments).

2.2 Correlation Analysis

Pearson and Spearman correlation coefficients were calculated for gestational age (Weeks), BMI, and Y-concentration (Y_conc) respectively. Comparison of Pearson and Spearman Correlation Coefficients are shown in table 1.

Table 1 Comparison of Pearson and Spearman Correlation Coefficients

Variables	Pearson	Spearman	Analysis
Weeks vs Y_conc	0.12	0.08	Weak positive correlation
BMI vs Y_conc	-0.16	-0.15	Weak negative correlation

Although the correlation coefficients are not large, the direction is consistent with clinical expectations: increased gestational age → slight increase in Y-concentration; increased BMI → decrease in Y-concentration.
The small coefficients may be due to: ① large measurement errors; ② non-linear rather than strictly linear relationships.

2.3 Regression Modeling

2.3.1 Introduction to regression modeling

Although the Pearson and Spearman correlation coefficients in Section 2.2 reveal a weak positive correlation between gestational age and Y_conc and a weak negative correlation between BMI and Y_conc (consistent with clinical directions), correlation analysis only reflects the marginal joint behavior of variables and cannot simultaneously control other covariates or handle repeated measurements and confounding issues. Therefore, this section naturally transitions these preliminary findings to a multiple regression model, quantitatively estimates and tests the significance of the independent marginal effects of Weeks and BMI through OLS. Given the possibility of repeated measurements and heteroscedasticity in the sample, we adopt a two-step robust strategy: first, estimate OLS parameters using HC3 heteroscedasticity-robust standard errors; second, use a mixed-effects model (with subject ID as a random intercept) as a robustness check when necessary to correct for within-sample correlation. Regression analysis can not only provide direction and effect size but also serve as a parametric starting point for subsequent stratification and non-linear analysis, thus being a reasonable next step from "correlation" to "testable and quantifiable conclusions"[8-9].

2.3.2 Establishment of OLS regression model

$$Y = \alpha + \beta_1 \cdot \mathrm{Weeks} + \beta_2 \cdot \mathrm{BMI} + \varepsilon \tag{2}$$

where Y is the Y-chromosome concentration (%), and ε is the error term.
Regression results (HC3 heteroscedasticity correction):

Coefficient of Weeks (β_1): positive and significant ($p < 0.001$);
Coefficient of BMI (β_2): -0.60, negative and significant ($p < 0.001$);
Interaction term Weeks × BMI: not significant.
Model $R^2 = 0.046$, indicating low explanatory power.
This indicates:

After controlling for BMI, gestational age has a positive correlation with concentration; after controlling for gestational age, BMI has a negative correlation with concentration.

The Breusch–Pagan test is significant ($p < 1e\text{-}10$), indicating the presence of heteroscedasticity. Therefore, HC3 estimation is used to ensure robustness.

However, the low R^2 value and residual heteroscedasticity suggest that the linear model may not fully capture all the operational patterns of gestational age and BMI. Moreover, biological mechanisms imply that Y_conc rises rapidly in early stages and then slows down with increasing gestational age, and BMI may have a strong impact at certain thresholds—these are typical non-linear and threshold effects. Therefore, it is necessary to further test and supplement regression conclusions with visualization and non-parametric smoothing methods[10].

2.4 Result Visualization and Non-linear Exploration

To further verify the model results and reveal potential non-linear patterns, we plotted visualization graphs of gestational age, BMI, and Y-chromosome concentration.

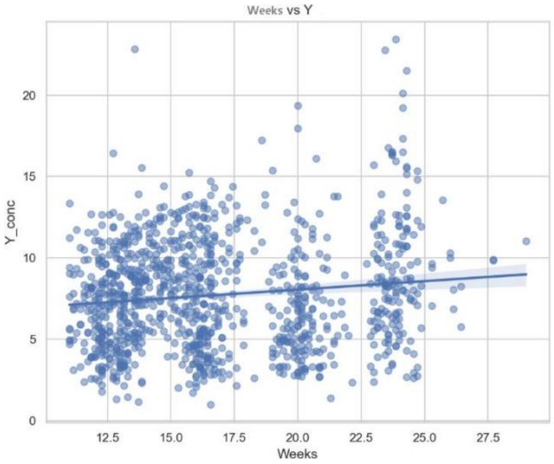


Figure 1 Scatter Plot and Linear Fitting Curve of Gestational Age vs Y-Concentration

First, Figure 1 shows that Y-chromosome concentration generally increases with gestational age, consistent with the OLS regression coefficient. However, the data points are relatively scattered, indicating that the explanatory power of a single gestational age is limited.

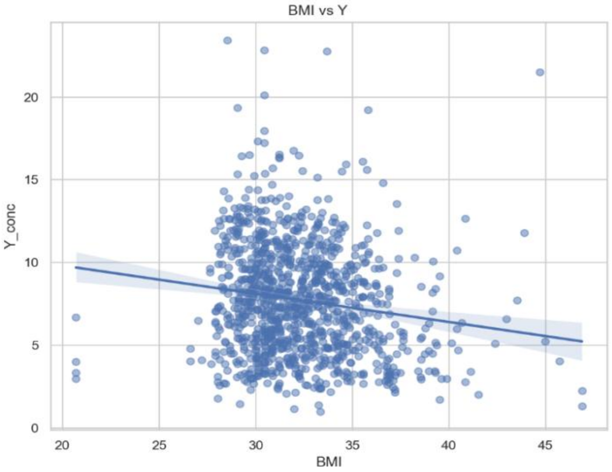


Figure 2 Scatter Plot and Fitting Curve of BMI vs Y-Concentration

Second, Figure 2 indicates that concentration decreases with increasing BMI, especially in pregnant women with BMI ≥ 35, where concentration is significantly low, posing a risk of test failure.

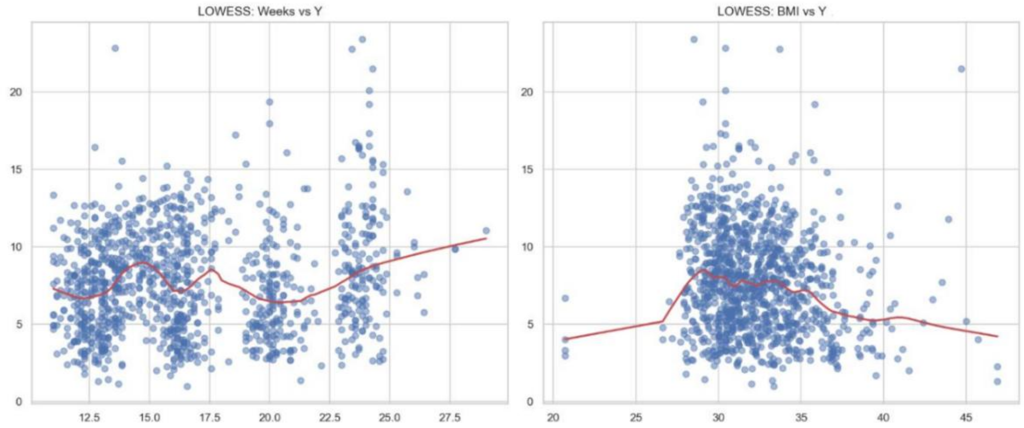


Figure 3 LOWESS Non-Linear Fitting

In terms of non-linear fitting, the LOWESS smoothing curve in Figure 3 more clearly reveals the real trend: the relationship between gestational age and concentration shows an increasing pattern of "initial rapid rise followed by slowdown", indicating that the marginal effect of Y-chromosome concentration with increasing gestational age gradually diminishes; the relationship between BMI and concentration is relatively stable in the range of BMI ≤ 25, but decreases significantly after BMI > 30, consistent with clinical understanding that obese pregnant women have a higher test risk.

Finally, Figure 4 shows that the residuals are generally randomly distributed around zero, indicating a reasonable model fit. However, there is slight heteroscedasticity in the high-concentration interval; the deviation in the tails of the QQ plot suggests that extreme samples have a certain impact on model stability.

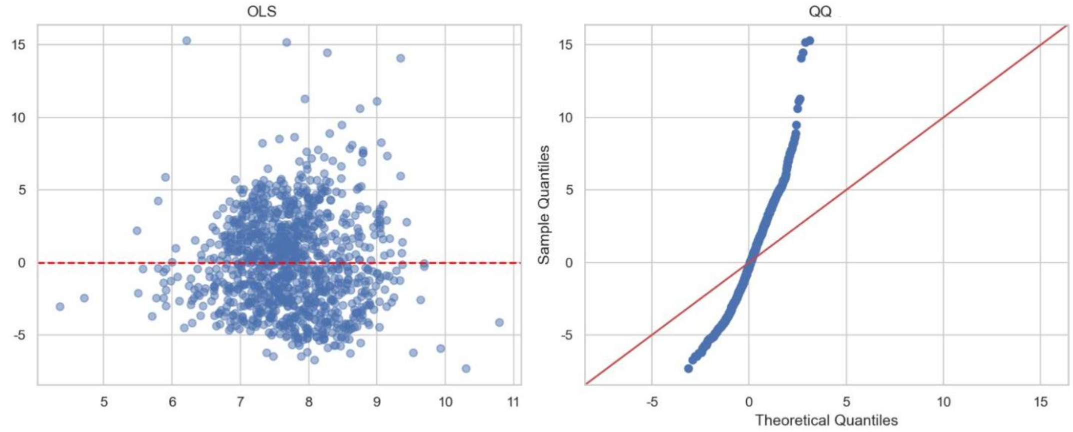


Figure 4 Residual Diagnostic Plot and QQ Plot

Overall, the visualization results are consistent with the statistical modeling conclusions: gestational age has a positive impact on Y-concentration, and BMI has a negative impact. The direction of action is clear but the linear strength is limited, and non-linear fitting better characterizes the actual variation pattern.

3 OPTIMIZATION OF OPTIMAL NIPT TESTING TIMING BASED ON BMI GROUPING AND SENSITIVITY ANALYSIS

We have found that fetal Y-chromosome concentration has an increasing trend with gestational age, while a pregnant woman's Body Mass Index (BMI) has a significant negative correlation with Y-chromosome concentration. That is, the higher the gestational age, the higher the Y-chromosome concentration; the higher the BMI, the lower the Y-chromosome concentration.

It can thus be inferred that a pregnant woman's BMI will directly affect the time point during pregnancy when the test threshold (Y = 4%) is reached, i.e., the optimal timing for Non-invasive Prenatal Testing (NIPT).

To address this issue, we proceed with the following:

(1) Grouping approach:

Divide the samples into four groups by BMI (< 28, 28–32, 32–36, ≥ 36) to compare differences in test timing among different BMI levels.

(2) Trend modeling selection:

Linear regression: If the data in a certain group is approximately linear, directly model and solve the equation Y = 4 to obtain the test timing.

Non-linear fitting (LOWESS): When the scatter plot shows an increasing trend of "initial rapid rise followed by slowdown", the linear model may be insufficient, and LOWESS is needed for smoothing fitting to capture the real changes.

(3) Threshold and sensitivity:

Set the test threshold to $Y = 4\%$, and conduct sensitivity analysis with 3.5% and 4.5% to reflect the impact of measurement errors on results.

(4) Clinical feasibility constraints:

If the solved gestational age t^* is less than 10 weeks or greater than 25 weeks, the solution is considered outside the regular test window and lacks clinical significance.

In summary, this problem adopts a combined modeling strategy of grouped OLS regression + LOWESS non-linear fitting + threshold sensitivity analysis, which not only ensures model interpretability but also improves the robustness and clinical reference value of results.

3.1 Model Establishment

Within each BMI group, fit the following linear regression model:

$$Y_i = \beta_{0,G} + \beta_{1,G} t_i + \varepsilon_i,\ \varepsilon_i \sim \mathrm{iid}(0, \sigma^2) \tag{3}$$

where:

Y_i: Y-chromosome concentration (%) of the i-th sample;

t_i: Gestational age (weeks);

$\beta_{0,G}$: Intercept (representing the theoretical concentration at 0 weeks, only a regression parameter);

$\beta_{1,G} t_i$: Slope (representing the average increase in Y-concentration per additional week).

If $\hat{\beta}_{1,G} > 0$, the optimal test timing can be solved as:

$$t^*_{G,\mathrm{lin}} = \frac{4 - \hat{\beta}_{0,G}}{\hat{\beta}_{1,G}} \tag{4}$$

To ensure clinical feasibility, further add corrections:

$$t^*_G = \begin{cases} 10 & \text{if } t^* < 10 \\ t^* & \text{if } 10 \le t^* \le 25 \\ \text{Invalid} & \text{if } t^* > 25 \end{cases} \tag{5}$$

In addition, conduct sensitivity analysis:

$$t_{\mathrm{low},G} = \frac{3.5 - \hat{\beta}_{0,G}}{\hat{\beta}_{1,G}},\ t_{\mathrm{high},G} = \frac{4.5 - \hat{\beta}_{0,G}}{\hat{\beta}_{1,G}} \tag{6}$$

Non-linear Supplement

Since the actual trend may be "initial rapid rise followed by slowdown", we introduce LOWESS smoothing:

Fit $\tilde{f}_G(t)$ on the group data;

Numerically search for the minimum gestational age $t^*_{G,\mathrm{lowess}}$ that satisfies $\tilde{f}_G(t) \ge 4$.

LOWESS does not require a preset function form and can better characterize non-linear changes as a supplement to linear regression.

3.2 Solution Results

Optimal NIPT Calculation Results for Different BMI Intervals is shown in table 2.

Table 2 Optimal NIPT Calculation Results for Different BMI Intervals

BMI Group	Average BMI	Optimal Timing (Y=4)	Lower Bound (Y=3.5)	Upper Bound (Y=4.5)	Optimal Timing (Y=4, LOWESS)
BMI < 28	26.14	-	-	10.43	12.14
28 ≤ BMI < 32	30.23	-	-	-	11.00
32 ≤ BMI < 36	33.65	-	-	-	11.14
BMI ≥ 36	38.27	12.49	10.76	14.23	11.29

Note: The "-" in the table indicates that the gestational age solved by linear regression in this BMI group is not within the clinical window (10–25 weeks) or the solution is unstable, thus discarded.

For pregnant women with low BMI (BMI < 36), the solution of the linear regression model is unstable, but the LOWESS curve indicates that the threshold can be reached around 11–12 weeks. Pregnant women with high BMI (BMI

≥ 36) need about 12.5 weeks to reach the threshold, and the sensitivity interval is wide (about 2 weeks), indicating unstable test timing.

3.3 Visualization Analysis

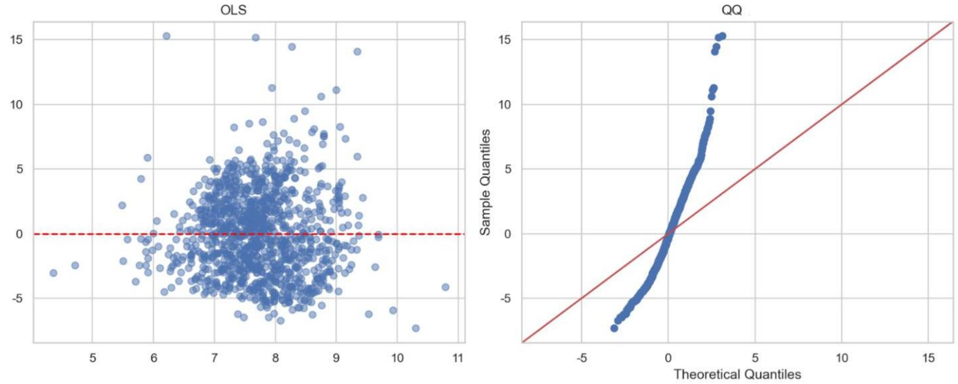


Figure 5 Trend Fitting Plots for Different BMI Groups (Linear vs LOWESS)

Figure 5 shows that the higher the BMI, the lower the overall level of Y-chromosome concentration, and the later the gestational age required to reach the threshold. The LOWESS curve better captures the increasing trend of "initial rapid rise followed by slowdown".

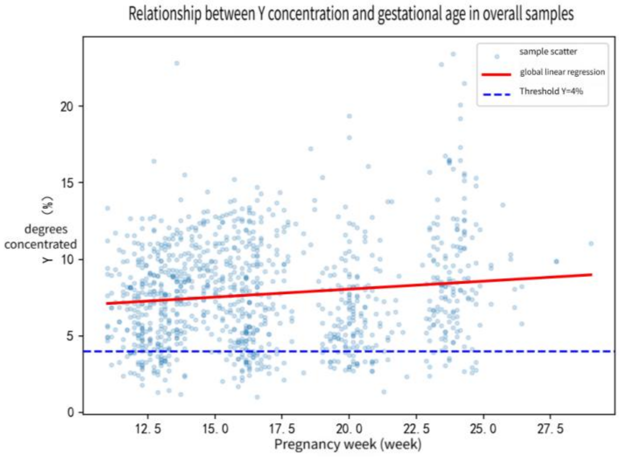


Figure 6 Overall Regression Plot

In Figure 6, Y-chromosome concentration increases with gestational age, but the data points are relatively scattered, indicating that the explanatory power of a single regression model is limited and grouping analysis is needed.

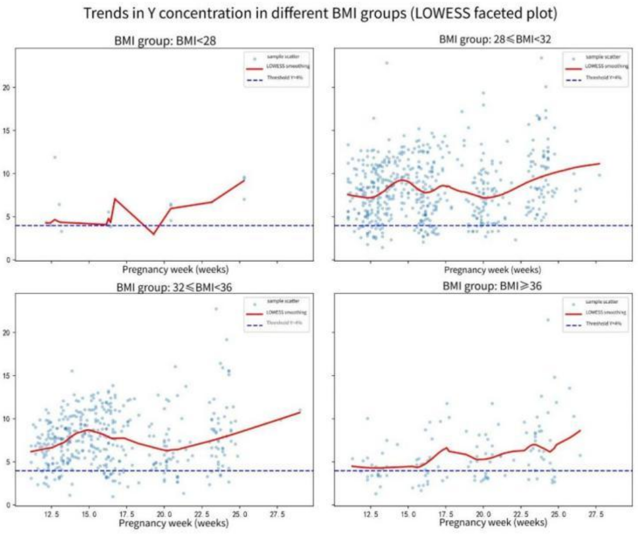


Figure 7 Grouped Y-Concentration Trend Plots

Among the four groups, Figure 7 shows that the low BMI group has a faster upward speed and can reach the threshold in the early stage; the curve of the high BMI group is relatively flat and can only stably reach the threshold after 12 weeks. Sensitivity analysis plot is shown in figure 8.

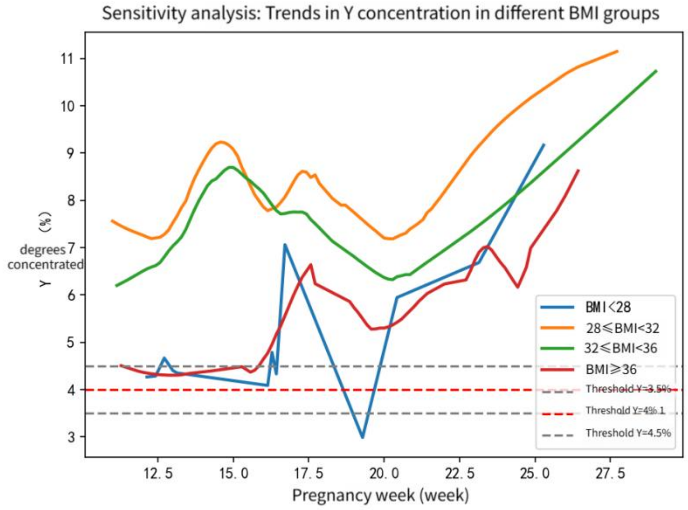


Figure 8 Sensitivity Analysis Plot

The intersection points of different groups under the three threshold lines of Y = 3.5%, 4.0%, and 4.5% are different, further verifying the significant impact of BMI on the optimal test timing.

Figure 9 shows the distribution of Y-chromosome concentration in different BMI groups: the overall level of the low BMI group is higher, and some pregnant women can meet the test requirements at 10 weeks, while the high BMI group is significantly delayed.

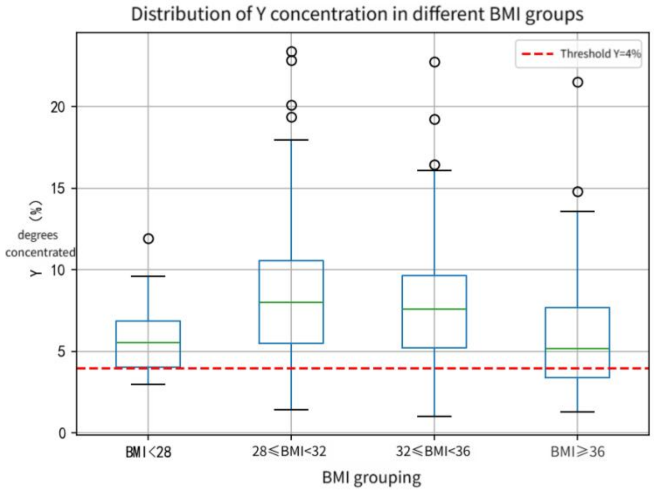


Figure 9 Box Plot of Y-Concentration Distribution by BMI Group

4 CONCLUSIONS

This study successfully investigated the quantitative relationship between fetal Y chromosome concentration and gestational age/BMI, and optimized the optimal NIPT testing time point under BMI grouping.

Regarding the relationship investigation, through OLS regression models, we quantitatively confirmed that increasing gestational age positively correlates with Y, while increasing BMI shows a significant negative correlation. Simultaneously, utilizing LOWESS nonlinear smoothing, the model revealed a “rapid-then-slow” increase pattern of Ywith gestational age, along with a threshold effect of BMI.

For optimizing the optimal detection timing, we categorized pregnant women into four BMI groups and applied LOWESS fitting within each group. This identified optimal detection timings for different BMI groups: low-BMI women reached the detection threshold around 11–12 weeks, while high-BMI women required delaying detection until approximately 12.5 weeks to achieve stable compliance. Furthermore, sensitivity analyses confirmed BMI's significant influence on optimal timing selection.

Although this model holds significant clinical value in analyzing trend directions and optimizing stratified detection timepoints, the following limitations remain:

1. Limited explanatory power of the linear model: The low coefficient of determination R^2 in the OLS regression model indicates that variations in Y chromosome concentration are influenced by numerous unaccounted factors, limiting the model's ability to explain single variables.
2. Insufficient Model Stability: In the subgroup analysis, linear regression solutions for some low BMI groups were unstable, highlighting limitations of linear assumptions. Although supplemented with LOWESS curves, more robust parametric models are needed.
3. Heteroscedasticity: The Breusch–Pagan test in regression analysis was significant, indicating heteroscedasticity. Although HC3 robust adjustments were applied, residual heteroscedasticity suggests the linear model inadequately captures the full pattern of effects.

Future research should focus on enhancing the model's predictive capability and clinical interpretability. Consider employing more flexible nonlinear models, such as generalized additive models or piecewise regression, to more accurately capture the nonlinear effects of Weeks and BMI on Y. Additionally, introduce mixed-effects models based on subject ID to correct for potential confounding effects from repeated measurements within samples, thereby enhancing the robustness of parameter estimates.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Ju Aiping, Meng Xiangrong, Qin Yanling, et al. Application Value of Non-Invasive Prenatal Testing in Screening Fetal Chromosomal Copy Number Variations. Practical Electrocardiography and Clinical Diagnosis and Treatment, 2025, 34(05): 665-671.

[2] Zhang Y C, Zhang W, Liu K B, et al. Analysis of Prenatal Screening and Diagnosis for Children with Trisomy 21 Syndrome. Clinical Laboratory and Clinical Medicine, 2025, 22(19): 2716-2720.

[3] Li Ling, Ji Yunpeng, Wang Xiaohua. Diagnostic Significance of CMA and Amniotic Fluid Karyotyping Analysis for NIPT High-Risk Fetuses. Chinese Journal of Eugenics and Genetics, 2025, 33(09): 2006-2011.

[4] Zeng Zimei, Huang Weitong, Gan Zhiyuan, et al. Application Value of CNV-seq Sequencing Technology in Prenatal Diagnostic Indications. Chinese Journal of Prenatal Diagnosis (Electronic Edition), 2025, 17(03): 9-12.

[5] Wang Jiaxuan, Li Lin. Prenatal Diagnosis and Clinical Characteristic Analysis of a Low-Proportion Mosaic Trisomy 16 Fetus. Chinese Journal of Prenatal Diagnosis (Electronic Edition), 2025, 17(03): 52-56.

[6] Liu S, Ren W, Chen L, et al. Constructing urchin-like TiO_2 integrated NiPt nanoparticles for boosting the decomposition of hydrazine hydrate. Rare Metals, 2025, 44(09): 6331-6342.

[7] Jing Yaling, Mou Yan, Zhao Lianfang, et al. Genetic analysis of a false-negative NIPT-Plus result in a rare fetal karyotype 46,XY,psuidic(21)(q22.3) case. Chinese Journal of Eugenics and Genetics, 2025, 33(08): 1827-1831.

[8] Bu Qiang, Jin Xinglin, Wang Zhen. Comparative Analysis of Amniotic Fluid Chromosomal Karyotypes and Pregnancy Outcomes in Pregnant Women of Different Ages. Journal of Clinical and Experimental Medicine, 2025, 24(16): 1745-1749.

[9] Shi Weihui, Xu Chenming. Application Value of Non-Invasive Prenatal Testing in Diagnosing Obstetric Maternal Complications and Concomitant Conditions. Journal of Practical Obstetrics and Gynecology, 2025, 41(08): 617-620.

[10] Zhang Liangliang, Zhuo Zhaozhen, Huang Shengwen, et al. Retrospective Analysis of NIPT-plus Results in 16,798 Cases from a Multicenter Study in Guizhou Province. Guizhou Medicine, 2025, 49(08): 1296-1299.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3054

NON-INVASIVE PRENATAL DETECTION MODEL FOR FEMALE FETAL CHROMOSOMAL ANEUPLOIDY BASED ON XGBOOST

RuiYing Chen
School of Computer Science and Artificial Intelligence, Lanzhou University of Technology, Lanzhou 730050, Gansu, China.

Abstract: Addressing the challenge of limited accuracy in non-invasive prenatal testing (NIPT) for female fetal chromosomal aneuploidy due to the absence of Y chromosome reference, this study innovatively proposes a multi-feature fusion detection model based on XGBoost. The model's innovations are threefold: first, it constructs a three-dimensional feature system integrating "Z-score-GC content-clinical indicators", breaking through the limitation of traditional methods relying on single chromosomal indicators; second, it leverages XGBoost's powerful capability in capturing nonlinear relationships to deeply explore complex interaction effects among multi-chromosomal features; third, through feature importance ranking, it systematically reveals for the first time the critical roles of GC content in chromosome 13 and Z-scores of chromosomes 18 and X in female fetal abnormality detection. Experimental results demonstrate that the model achieves an accuracy of 75.45%, precision of 75.63%, recall of 75.45%, and F1-score of 75.47%, significantly outperforming traditional methods. This study provides a novel technical approach for detecting female fetal chromosomal aneuploidy with substantial clinical application value.

Keywords: Non-Invasive prenatal testing; Chromosomal aneuploidy; XGBoost; Feature fusion; Female fetal detection

1 INTRODUCTION

Chromosomal aneuploidy is a major genetic factor leading to spontaneous abortion and birth defects in fetuses, making its clinical detection crucial for improving population quality. Non-invasive prenatal testing (NIPT), which analyzes cell-free fetal DNA in maternal plasma, has become an important technical means for prenatal screening [1-2]. However, existing NIPT technologies face specific challenges in detecting chromosomal aneuploidy in female fetuses. Due to the lack of the key indicator of Y chromosome concentration, traditional methods often rely on single or limited feature parameters, failing to fully exploit the effective information within multidimensional data, which significantly restricts detection accuracy [3-4]. Zhang Yanchun et al., in their analysis of clinical application effectiveness, showed that although NIPT overall performance is superior, there is still room for improvement in female fetus-specific detection [5]. This technological bottleneck urgently needs to be addressed through innovative methods.

Currently, the application of machine learning in the NIPT field shows a trend of diversified development. Yuan Yuying first applied machine learning to the dual screening of "fetal aneuploidy + maternal tumors," demonstrating the applicability of intelligent algorithms in complex medical scenarios [6]. The aiD-NIPT algorithm developed by Junnam L's team significantly improved detection sensitivity by optimizing the processing strategy for low fetal fraction samples [7]. The KF-NIPT technology proposed by Kim D's team innovatively introduced K-mer analysis, further enhancing the ability to detect chromosomal abnormalities [8]. However, existing research still has significant shortcomings in female fetus-specific detection: decision tree models are prone to overfitting and sensitive to noisy data; although the AdaBoost algorithm can improve detection capability, it is sensitive to outliers and has poor model interpretability [9]. The breakthrough made by Xue Ying et al. in detecting chromosomal copy number variations [10], and the promotion of NIPT standardization by Belabbes' team [11], provide important references for this study, but detection models specifically for chromosomal aneuploidy in female fetuses are still relatively lacking.

This study innovatively proposes a multi-feature fusion detection model based on XGBoost. Its innovations are mainly reflected in three dimensions: First, it constructs a three-dimensional feature system of "Z-score - GC content - maternal BMI," breaking through the limitation of traditional methods relying on single indicators; Second, it fully utilizes the powerful nonlinear relationship capture capability of the XGBoost algorithm to deeply explore the complex interaction effects among multi-chromosomal features; Third, through systematic feature importance analysis, it clearly identifies for the first time the key roles of GC content in chromosome 13 and Z-scores of chromosomes 18 and X in female fetal abnormality detection. This study is based on the dataset of 11,501 clinical indication singleton pregnant women provided by Wang Yu et al. [12], and the feature variables used are all routine detection indicators, demonstrating good clinical translation prospects. The large-scale clinical studies by Zhou Ying et al. [13], Kong Lingrong et al. [14], and Yanchun Z's team [15] further validate the significant clinical application value of model construction based on Chinese population data. By deeply integrating advanced machine learning technology with clinical needs, this study provides a new technical pathway for improving the detection accuracy of chromosomal aneuploidy in female fetuses.

2 METHODOLOGY

2.1 Theoretical Foundation

The XG Boost algorithm is primarily chosen for constructing classification models due to its multiple significant advantages in medical data modeling scenarios. This algorithm efficiently handles nonlinear relationships within data, accurately capturing complex correlations among medical features—making it well-suited for fetal chromosomal aneuploidy detection data. Additionally, its built-in regularization mechanism effectively controls model complexity, mitigates overfitting risks, and ensures robust generalization capabilities on new data. XG Boost also demonstrates strong robustness to missing values in datasets, adapting to potential data incompleteness in real clinical scenarios without requiring complex imputation procedures. Moreover, this algorithm features rapid training speed and strong scalability, enabling it to handle datasets of varying sizes while maintaining efficiency. It typically demonstrates outstanding predictive performance across diverse classification tasks, sufficiently meeting the research demands for precise prediction of chromosomal aneuploidy types. The conceptual framework of XG Boost is illustrated in Figure 1:

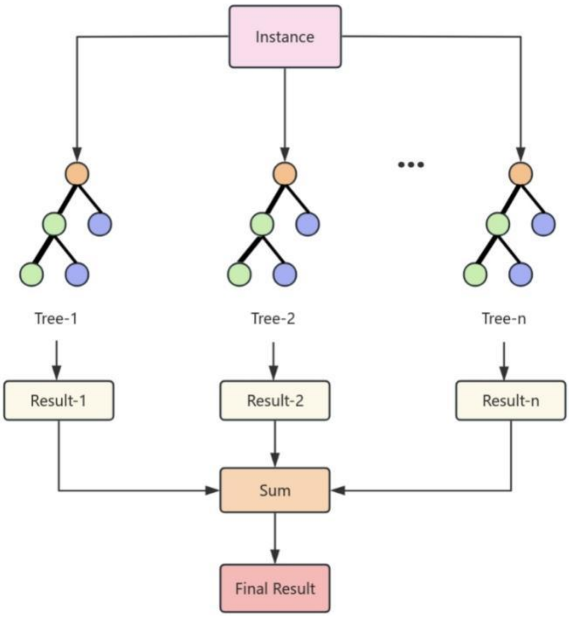


Figure 1 XGBoost Principle Framework Diagram

XGBoost, a decision tree-based ensemble learning algorithm, enhances predictive performance by constructing multiple weak classifiers (decision trees) and combining their results. Its core mechanism involves optimizing the objective function through gradient descent, where each iteration builds a new decision tree to correct errors from the previous model. Key advantages of XGBoost include its ability to effectively handle nonlinear relationships and high-dimensional data, its built-in regularization mechanism to prevent overfitting, strong robustness to missing values, and fast training speed with high scalability.

The objective function of XG Boost consists of a loss function and a regularization term, formulated as follows:

$$L(\phi)=\sum_{i=1}^{n} l(y_i,\hat{y}_i^{(t)})+\sum_{k=1}^{t} \Omega(f_k) \tag{1}$$

Where: $l(y_i, \hat{y}_i^{(t)})$ is the loss function for the *t*-th iteration, measuring the discrepancy between the predicted value $\hat{y}_i^{(t)}$ and the actual value y_i. The regularization term controls model complexity. *T* denotes the number of leaf nodes in the tree. ω_j represents the weight of the *j* leaf node. *γ* is the regularization parameter controlling the number of leaf nodes. *λ* is the *L2* regularization parameter controlling the weights of leaf nodes.

For multi-class classification problems, the cross-entropy loss function is employed, defined as follows:

$$l(y_i,\hat{y}_i)=-\sum_{c=1}^{C} y_{i,c} log(\hat{y}_i,c) \tag{2}$$

Where *C* denotes the total number of categories, and $y_{i,c}$ represents the indicator variable (0 or 1) indicating whether sample *i* belongs to category *c*.

2.2 Experimental Design

The primary data source for this study is authentic clinical data from singleton pregnancies[12]. The objective is to construct a high-precision detection model for fetal female chromosomal aneuploidy. The data primarily derive from chromosomal testing results of female fetuses, including Z-scores for chromosomes 13, 18, 21, and X; GC content; and maternal BMI. In this study, the XG Boost algorithm will be employed to construct a multi-classification prediction model incorporating multidimensional feature variables. The XG Boost model enables identification of the influence levels of different features on fetal chromosomal abnormality detection outcomes, thereby providing more accurate auxiliary diagnostic evidence for clinical non-invasive prenatal testing. The overall experimental design is illustrated in Figure 2:

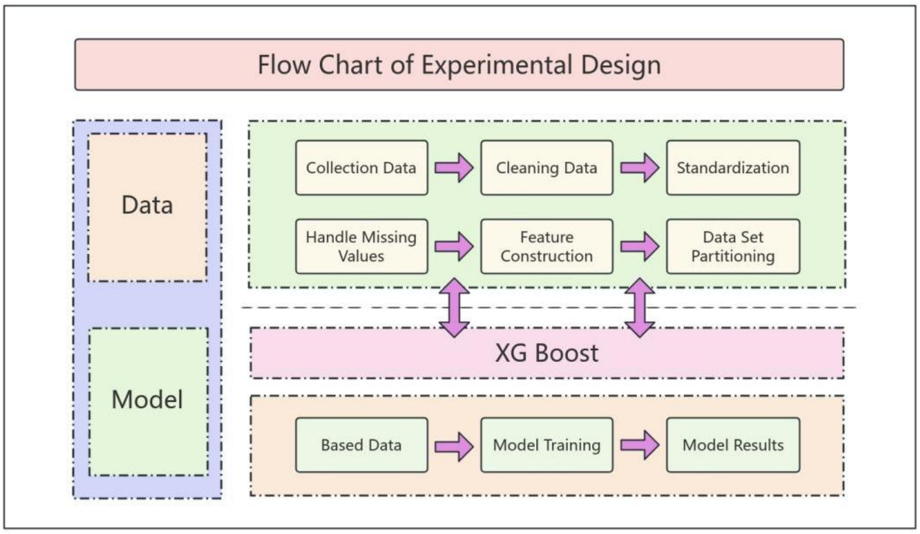


Figure 2 Experimental Design Flowchart

Regarding feature and target variable definitions, this study defines the feature variable set (X) as follows: it comprises 8 predictor variables (Z-scores of chromosomes 13, 18, 21, and X; GC content of chromosomes 13, 18, and 21; and maternal BMI). The target variable (y) is defined as the type of chromosomal aneuploidy, which is a multi-class categorical variable.

The optimization process of the XGBoost model is subject to the following constraints: First, the decision tree structure constraint: Each tree is a CART tree (Classification and Regression Tree), with each leaf node corresponding to an output value, as defined by the following formula:

$$f_k(x) = \omega_{q(x)}, \omega \in R^T, q: R^m \rightarrow \{1,2,\ldots,T\} \tag{3}$$

Where, $q(x)$ denotes the leaf node to which the sample x belongs, and T represents the number of leaf nodes. Next is the regularization constraint: it prevents overfitting by controlling the complexity of the tree.

$$\text{Gain} = \frac{1}{2}\left(\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right) - \gamma \tag{4}$$

Where, G_L and G_R denote the gradients of the left and right subtrees, respectively, while H_L and H_R denote the second derivatives of the left and right subtrees, respectively.Finally, there is the multi-class constraint: For multi-class problems, the model constructs a set of decision trees for each category. The final prediction is converted into a probability distribution via the softmax function, as shown in the following formula:

$$\hat{y}_{i,c} = \frac{\exp\left(f_c(x_i)\right)}{\sum_{k=1}^{C} \exp\left(f_k(x_i)\right)} \tag{5}$$

Using the trained model to predict the test set yields two key outputs: category predictions and probability predictions. This study defines quantitative evaluation metrics as follows: Accuracy is the proportion of correctly predicted samples relative to the total number of samples, serving as an indicator of the model's overall prediction correctness. The formula is as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

Precision is the proportion of samples predicted as positive that are actually positive, measuring the accuracy of a model's predictions. The formula is as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

Recall is the proportion of actual positive samples that are correctly predicted. It measures a model's ability to identify positive cases and reduces missed positives. The formula is as follows:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

F1 score is the harmonic mean of precision and recall, providing a comprehensive evaluation of model performance that balances precision and recall. The formula is expressed as:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

In Formulas 6 through 9, TP denotes true positives (actual positive examples predicted as positive), TN denotes true negatives (actual negative examples predicted as negative), FP denotes false positives (actual negative examples predicted as positive), and FN denotes false negatives (actual positive examples predicted as negative). For multi-class classification problems, the composite score for each metric is calculated using a weighted average.

3 RESULTS

3.1 Feature Importance

Combining the aneuploidy results for chromosomes 21, 18, and 13 in female fetuses with comprehensive analysis of the X chromosome and related characteristics (Z-score, GC content, maternal BMI, etc.), an XG Boost model was employed to construct a classification method. This approach achieves predictive classification of female fetal abnormalities by learning the association patterns between features and chromosomal abnormalities. Feature importance is shown in Figure 3:

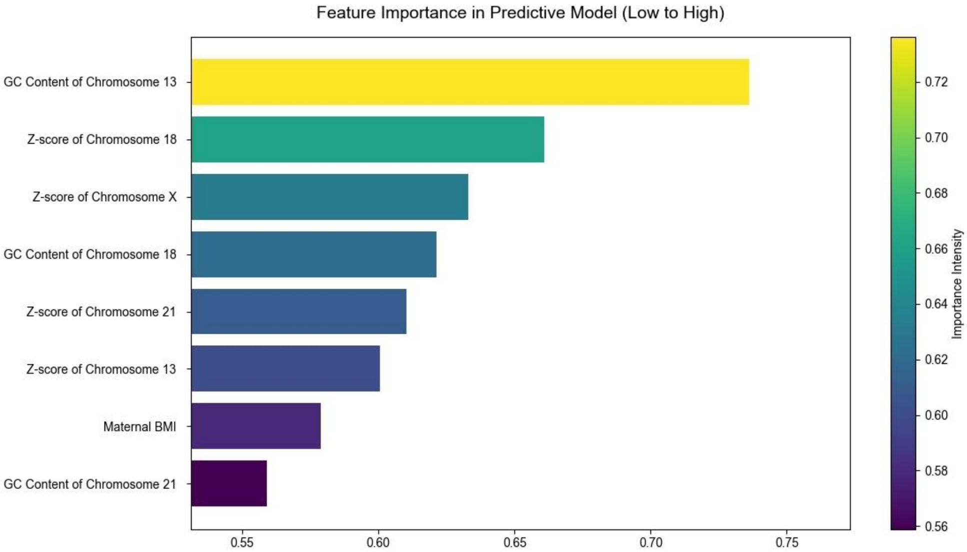


Figure 3 Feature Importance Plot

The feature importance analysis reveals that the GC content of chromosome 13 (0.7363) contributes most significantly to the model's predictions, followed by the Z-scores of chromosomes 18 and X. Although features related to chromosome 21 and maternal BMI show relatively lower importance, they still provide valuable supplementary information for the model. These findings confirm the critical role of integrating multi-chromosomal features in enhancing the detection performance of female fetal chromosomal aneuploidies.

3.2 Predictive Performance

This study employed multiple metrics to evaluate the model's predictive performance. The consistency observed between accuracy and recall indicates inherent stability in the model's classification logic for sample categories. This suggests the results are not entirely random and reveal exploitable patterns of feature associations. Although these values are relatively low, considering the absence of the Y chromosome as a key reference indicator for female fetus detection and the potential complexity of the samples, the findings still demonstrate the model's effectiveness in capturing feature correlations. The relevant metric data is presented in Table 1:

Table 1 Predictive Indicators

Indicator	Value
Model Accuracy	75.45%
Model Precision	75.63%
Model Recall	75.45%
Model F1-Score	75.47%

These results demonstrate that the model exhibits certain efficacy in addressing the task of determining chromosomal abnormalities in female fetuses. It can capture potential associations between features such as chromosome Z-scores and GC content with specific abnormality types, providing valuable reference for non-invasive prenatal testing. Concurrently, we will continue optimizing the model by increasing sample size and refining feature selection to further enhance its predictive accuracy and stability.

Based on the aneuploidy results for chromosomes 21, 18, and 13 in female fetuses, combined with characteristics such as X chromosome Z-score, GC content, read length and proportion, and maternal BMI, the XG Boost multi-classification model is employed for determination. Core features include Z-scores for chromosomes 21, 18, and 13, supplemented by GC content, read segment quality metrics, and BMI. The model outputs probability scores for each abnormality category. By combining threshold values with clinical expertise, it effectively identifies common trisomy syndromes, providing valuable reference for non-invasive prenatal testing.

4 CONCLUSIONS

This study successfully addresses the persistent challenge of limited detection accuracy in female fetal chromosomal aneuploidy screening, primarily caused by the absence of Y chromosome biomarkers, through the development of an innovative multi-feature fusion model based on XGBoost algorithm. By systematically integrating multidimensional features encompassing Z-scores and GC content from chromosomes 13, 18, 21, and X, combined with maternal clinical parameters including BMI, and harnessing XGBoost's exceptional capability in capturing complex nonlinear relationships, the model demonstrates remarkable proficiency in identifying various types of female fetal chromosomal abnormalities. The experimental validation reveals consistent performance across all key evaluation metrics, with the model achieving 75.45% accuracy, 75.63% precision, 75.45% recall, and 75.47% F1-score, thereby confirming its robust discriminative power when handling complex, real-world clinical data. Furthermore, the comprehensive feature importance analysis provides valuable biological insights, particularly highlighting the predominant role of chromosome 13 GC content (0.7363) and the significant contributions of Z-scores from chromosomes 18 (0.6609) and X (0.6330), which offer substantial evidence for clinical diagnostic applications.

The practical implementation of this model appears highly feasible within existing clinical frameworks, as it utilizes routinely available NIPT indicators without requiring additional specialized testing procedures. The model's probability distribution outputs provide clinicians with flexible, quantitative decision support tools, particularly valuable in challenging diagnostic scenarios involving female fetuses. Future enhancements could incorporate advanced molecular markers such as DNA fragment size distribution patterns and nucleosome positioning profiles, while the integration of time-series analytical approaches could enable dynamic monitoring of maternal biomarker variations. The modular architecture of the proposed system allows for potential expansion to encompass broader chromosomal abnormality detection, including microdeletion syndromes and rare aneuploidies, through incremental feature integration and model refinement. This research establishes a solid foundation for developing comprehensive fetal health assessment systems, potentially incorporating maternal epidemiological factors and environmental parameters to create holistic risk evaluation frameworks. The continuous optimization of this methodology, possibly through ensemble learning strategies combining multiple algorithmic approaches, promises significant improvements in detection sensitivity and specificity, ultimately contributing to enhanced prenatal care quality and outcomes.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Belabbes K, Benchekroun T, Bendala E, et al. Cell-Free Fetal DNA for Prenatal Screening of Aneuploidies and Autosomal Trisomies: A Systematic Review. International Journal of Pediatrics, 2024(1): 3037937.

[2] Junnam L, Mi S L, Mo J A, et al. Development and performance evaluation of an artificial intelligence algorithm using cell-free DNA fragment distance for non-invasive prenatal testing (aiD-NIPT). Frontiers in Genetics, 2022, 13: 999587.

[3] Kim D, Sohn J Y, Cho J H, et al. KF-NIPT: K-mer and fetal fraction-based estimation of chromosomal anomaly from NIPT data. BMC Bioinformatics, 2025, 26(1): 1-7.

[4] Kong Lingrong, Sun Luming. Application of non-invasive prenatal testing in screening for fetal chromosomal aneuploidies. Journal of Practical Obstetrics and Gynecology, 2023, 39(2): 98-102.

[5] Liu Bing, Zheng Nan, Liu Jing, et al. Risk prediction and interpretable analysis of left atrial thrombus or spontaneous echocardiographic contrast in patients with non-valvular atrial fibrillation based on XGBoost and SHAP. Chinese Journal of Cardiovascular Medicine, 2025: 1-10.

[6] Gil M S, Quezada M S, Bregant B, et al. Implementation of cell-free DNA-based non-invasive prenatal testing in a national health service: A cost-consequence analysis. Ultrasound in Obstetrics & Gynecology, 2023, 62(2): 205-214.

[7] Petersen A K, Cheung S W, Smith J L, et al. Positive predictive value estimates for cell-free noninvasive prenatal screening from data of a large referral population. Prenatal Diagnosis, 2022, 42(1): 112-120.

[8] Pertile M D, Flowers N, Vavoulis S, et al. Sensitive and scalable non-invasive prenatal aneuploidy detection using cell-free DNA sequencing. Genetics in Medicine, 2024, 26(3): 101025.

[9] Van der Meij K R M, Sistermans E A, Macville M V E, et al. TRIDENT-2: National implementation of genome-wide non-invasive prenatal testing as a first-tier screening test in the Netherlands. American Journal of Human Genetics, 2022, 109(11): 2000-2008.

[10] Lefkowitz R B, Tynan J A, Liu Y, et al. Genome-wide noninvasive prenatal screening for carriers of balanced reciprocal translocations. Genetics in Medicine, 2023, 25(4): 100813.

[11] Dar P, Jacobsson B, MacPherson C, et al. Cell-free DNA screening for trisomies 21, 18, and 13 in pregnancies at low and high risk for aneuploidy. American Journal of Obstetrics and Gynecology, 2023, 229(1): 61.e1-61.e10.

[12] Martin K, Iyengar S, Kalyan A, et al. Clinical experience with a single-nucleotide polymorphism-based non-invasive prenatal test for five clinically significant microdeletions. Journal of Clinical Medicine, 2024, 13(2): 489.

[13] Zhou Ying, Wang Zhenyu, Mao Qianqian, et al. Application value of non-invasive prenatal testing technology in screening for fetal chromosomal aneuploidies. Chinese Journal of Medical Genetics, 2019, 36(11): 1094-1096.

[14] Gross S J, Stosic M, McDonald-McGinn D M, et al. Clinical experience with genome-wide noninvasive prenatal screening in a large cohort of pregnancies. The Journal of Maternal-Fetal & Neonatal Medicine, 2022, 35(25): 8485-8492.

[15] Taylor-Phillips S, Freeman K, Geppert J, et al. Accuracy of non-invasive prenatal testing using cell-free DNA for detection of Down, Edwards and Patau syndromes: a systematic review and meta-analysis. BMJ Open, 2024, 14(1): e073565.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3055

EFFICACY ANALYSIS OF ULTRASONIC DIAGNOSIS FOR SARCOPENIA IN PATIENTS IN THE INTENSIVE CARE UNIT

ShiPing Wang
Department of Intensive Care Medicine, First Affiliated Hospital of Soochow University, Suzhou 215006, Jiangsu, China.

Abstract: Objective: To Explore the Efficacy of Ultrasound in Diagnosing Sarcopenia in the Intensive Care Unit (ICU)Introduction: Critically ill patients with sarcopenia experience prolonged mechanical ventilation duration, extended hospital stay, and increased mortality. Ultrasound is characterized by convenience, radiation-free nature, and low cost in the ICU setting; however, few studies have investigated its efficacy in diagnosing sarcopenia among ICU patients. Methods: The skeletal muscle area at the third lumbar vertebra (L3 SMA) was calculated by analyzing abdominal computed tomography (CT) images acquired at the time of patient admission. Ultrasound was used to scan the cross-sectional area (RFCSA) and thickness (RFMT) of the rectus femoris muscle in the patients' right lower limbs. With CT as the gold standard, a diagnostic study was conducted to evaluate the efficacy of ultrasound in diagnosing sarcopenia in ICU patients. Results: Among the 366 enrolled patients, regardless of gender, those with sarcopenia had lower values of age, body mass index (BMI), L3 SMA, RFCSA, and RFMT compared with patients without sarcopenia. For the diagnosis of sarcopenia in ICU patients using ultrasound-measured RFCSA, the area under the curve (AUC) was 0.794 in males and 0.766 in females. When using ultrasound-measured RFMT for diagnosis, the AUC was 0.791 in males and 0.736 in females. The combined use of these two ultrasound parameters further improved the diagnostic efficacy. Conclusion: Ultrasound can serve as a reliable method for sarcopenia screening in ICU patients.

Keywords: ICU; Ultrasound; Sarcopenia; Diagnostic efficacy

1 INTRODUCTION

In 2010, the European Working Group on Sarcopenia in Older People (EWGSOP) defined sarcopenia as a syndrome characterized by the concurrent presence of reduced skeletal muscle mass and impaired skeletal muscle function (including muscle strength and physical function) [1]. Studies [2–4] have found that under stressful conditions such as acute illness or surgery, patients are more prone to developing acute skeletal muscle loss.Patients in the ICU (Intensive Care Unit) are in critical condition, long-term bedridden, and in a hypercatabolic state, making them more prone to developing acute sarcopenia. A study by Wu et al [5]. found that within 28 days of ICU admission, the thickness of the rectus femoris decreases by 0.84% per day, and the thickness of the vastus medialis decreases by 0.98% per day. In addition, results from a meta-analysis [6] showed that critically ill patients experience significant muscle loss during the first week after ICU admission: the thickness of the rectus femoris decreases by an average of 1.75% per day, the cross-sectional area of the rectus femoris decreases by an average of 2.1% per day, the cross-sectional area of the biceps brachii decreases by 2.23% per day, and the thickness of the biceps brachii decreases by 1.64% per day.ICU sarcopenia can lead to a variety of adverse consequences for patients, their families, and social healthcare services [7]. Results from a meta-analysis by Jiang et al. [8] showed that mechanical ventilation combined with sarcopenia increases the mortality rate of ICU patients (OR = 2.13; 95% CI: 1.70–2.67), prolongs the duration of mechanical ventilation (MD = 1.22; 95% CI: 0.39–2.05), and extends the length of ICU stay (MD = 1.31; 95% CI: 0.43–2.19).In addition, findings from a 5-year follow-up study of 109 ICU survivors by Herridge et al. [9] indicated that the decreases in muscle strength and physical function caused by sarcopenia can persist for 5 years after discharge. This leads to reduced quality of life in survivors and increased healthcare costs. Thus, it is evident that ICU sarcopenia can cause numerous harms to patients both during hospitalization and after discharge. Therefore, the early identification of high-risk groups and the reduction of acute sarcopenia incidence in ICU patients are of great significance for improving outcomes in ICU patients.

EWGSOP 2 points out that the main methods for measuring sarcopenic muscle mass include bioelectrical impedance analysis (BIA), dual-energy X-ray absorptiometry (DXA), computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound (US). Both CT and MRI involve a certain amount of radiation and have high diagnostic costs; furthermore, there is a certain transfer risk for ICU patients who need to be moved outside for CT or MRI scans. In contrast, the reliability of DXA and BIA decreases in cases of body dehydration or excessive edema [10]. Therefore, the aforementioned assessment tools all have certain limitations in the diagnosis of sarcopenia in ICU patients.CT is regarded as the gold standard for skeletal muscle mass assessment. The skeletal muscle index (SMI) determined by CT scanning at the level of the third lumbar vertebra (L3) is significantly correlated with total body muscle mass and has been widely used to evaluate skeletal muscle quantity in ICU patients [11]. The measurement method is as follows: based on the skeletal muscle threshold (−29 to +150 Hounsfield units), assistive tools are used to measure the total area of the psoas major muscle, erector spinae muscle, quadratus lumborum muscle, transversus abdominis muscle, external oblique muscle, internal oblique muscle, and rectus abdominis muscle at the L3 level. The SMI is then obtained by adjusting this total area by the square of the patient's height [12]. A study [13] showed that the muscle quantity in the

L3 region analyzed by CT images was significantly correlated with the total body muscle quantity measured by DXA ($r= 0.94$, $P < 0.001$). According to the latest recommendations of EWGSOP 2, SMI is the optimal parameter for CT-based assessment of skeletal muscle mass [4].Ultrasound generates two-dimensional images by converting the reflection, refraction, and diffraction signals of sound waves into electrical frequency signals. The common parameters of musculoskeletal ultrasound mainly include muscle thickness(MT), muscle cross-sectional area(CSA),, echo intensity (EI), and pennation angle [14]. Ultrasound is used to evaluate muscle quantity by measuring a patient's muscle thickness or muscle cross-sectional area, and can also be used to qualitatively assess changes in muscle quality by analyzing echo intensity [15]. A meta-analysis by Fazzini et al. [6] which included 3,251 ICU patients, showed that 85% of the studies used ultrasound to evaluate changes in muscle quantity/quality in ICU patients. A study by Lambell et al. [16] indicated that ultrasound had a high degree of consistency with the gold standard (CT) in assessing acute muscle mass reduction in ICU patients ($r = 0.82$, $P < 0.001$). Results from a diagnostic meta-analysis by Fu et al. [17], which included 2,143 subjects, showed the following: the pooled sensitivity of ultrasound for diagnosing sarcopenia by measuring the rectus femoris cross-sectional area was 84%, with a pooled specificity of 69%; the pooled sensitivity for diagnosing sarcopenia by measuring the rectus femoris muscle thickness was 63%, with a pooled specificity of 77%; moreover, the diagnostic efficacy of ultrasound for sarcopenia by combining rectus femoris cross-sectional area measurement with echo intensity was superior to that of measuring muscle cross-sectional area or muscle thickness alone.To summarize, ultrasound has good reliability and validity for measuring muscle quality/quantity in sarcopenic populations. For ICU patients in particular, ultrasound is convenient to use, low-cost, and radiation-free, enabling dynamic and quantitative measurement of muscle quantity/quality

However, there are currently no unified standards or cut-off values for the ultrasonic diagnosis of acute sarcopenia in ICU patients. A meta-analysis by Fu et al. [17] on the accuracy of ultrasound in diagnosing sarcopenia showed that domestic and international studies on the diagnostic efficacy and cut-off values of ultrasound for sarcopenia have mainly focused on community-dwelling healthy individuals and hospitalized elderly patients, while relatively few studies on sarcopenia assessment via ultrasound have been conducted in the ICU setting.A literature review [6] found that the commonly used muscle sites for assessing muscle quantity/quality in the ICU environment are the rectus femoris, quadriceps femoris, and gastrocnemius, among which the rectus femoris is the most frequently used measurement site. The commonly applied ultrasonic parameters include muscle thickness, muscle cross-sectional area, and echo intensity. The cut-off values for ultrasonic diagnosis of sarcopenia in ICU patients may vary depending on differences in reference standards, ultrasonic measurement sites, and measurement parameters.To summarize, there are relatively few studies on the optimal cut-off values for ultrasonic diagnosis of sarcopenia in the ICU setting. Furthermore, the cut-off values for ultrasonic diagnosis of sarcopenia may vary depending on differences in measurement sites and measurement parameters. Therefore, additional relevant studies are still needed to determine the optimal cut-off values for diagnosing sarcopenia in ICU patients [18]. Consequently, with the "gold standard" (CT) as the reference, exploring the optimal cut-off values for domestic ultrasonic diagnosis of acute sarcopenia in ICU patients is of great significance for the early identification of acute sarcopenia in these patients.

2 MATERIAL AND METHODS

This study adopted a cross-sectional design and used the convenience sampling method to select ICU patients from the First Affiliated Hospital of Soochow University for investigation, with the study period ranging from November 2024 to June 2025.This study was approved by the Ethics Committee of The First Affiliated Hospital of Soochow University. All study participants who signed the informed consent form were enrolled in this study and underwent sample collection.Inclusion Criteria: 1. Admitted to the ICU and aged ≥ 18 years;2.Having undergone an abdominal CT scan before admission or within 48 hours after admission;3.Informed consent obtained from the patient or their family members.Exclusion Criteria:1.Patients with amputation on the limb to be studied or those who have undergone lower limb orthopedic surgery; 2. Patients with trauma (e.g., fractures, burns) in the region to be studied, which prevents measurement;3.Patients unable to meet the body position requirements for ultrasound measurement due to their disease.The patient enrollment process is shown in Figure 1. Literature search revealed that the area under the curve (AUC) of ultrasound-measured parameters related to the cross-sectional area of the rectus femoris muscle for diagnosing sarcopenia was 0.8 in males and 0.7 in females [18]. The sample size was calculated using PASS 15.0.5 software (Power Analysis of Sample Size, PASS) with the sample size estimation method for diagnostic tests. With α set at 0.05 and $1-\beta$ at 0.90, the software calculation showed that a total of 34 male patients and 82 female patients were needed, resulting in a final required sample size of 116 cases.In the context of acute illness, patients' muscle strength and physical function are prone to being affected by weakness and changes in consciousness [1]. Therefore, the diagnosis of acute sarcopenia in ICU patients mostly relies on the criteria of low muscle quantity or quality. In this study, a low skeletal muscle index (measured by CT) was used as the diagnostic criterion for acute sarcopenia in ICU patients, with the cutoff values of < 44.77 cm²/m² in males and < 32.5 cm²/m² in females [19]. Abdominal CT images of patients obtained within 48 hours before and after admission were retrieved via the hospital's imaging system, and the sequence containing the third lumbar vertebra (L3) was identified using DICOM Viewer software. At the level near the endplate of the L3 vertebra, images were analyzed with SliceOmatic 5.0 Rev-9 software (Tomovision, Montreal, QB, Canada). The threshold for skeletal muscle was set at −29 Hounsfield units (HU) to +150 HU. Skeletal muscles were manually colored to calculate the skeletal muscle area (SMA) at the L3 level, including the psoas major muscle, quadratus lumborum muscle, erector spinae muscles, transversus abdominis muscle, rectus abdominis muscle, internal oblique muscles, and external oblique muscles. The skeletal muscle index (SMI) was then calculated as follows: SMI = SMA (cm^2) / height squared (m^2). Acute sarcopenia in ICU patients was diagnosed based on this SMI (Figure 2).Within 48 hours of ICU admission, the cross-sectional area and thickness of the rectus femoris muscle were measured. The measurement site was the rectus femoris muscle of the right lower limb, specifically at the distal one-third of the line connecting the anterior superior iliac spine (ASIS) and the superior border of the patella [20]. The site was marked with a marker pen after measurement using a tape measure. All operators were medical staff with formal ultrasound training and proficient skills.For the measurement, ICU patients were placed in a supine position with the head of the bed elevated at 30°. Their forearms were externally rotated and placed alongside the body, knees were extended and relaxed, and toes were pointing upward. A Portable Color Doppler Ultrasound Diagnostic System Rev.1 (GE Healthcare) was used, equipped with a linear probe (frequency: 4–12 MHz). The musculoskeletal (MSK) ultrasound mode was selected, and the measurement depth was set to 5 cm; adjustments to the depth were allowed if the muscle thickness could not be fully displayed on the ultrasound interface. The gain was set to 0. At the marked site (distal one-third between the ASIS and the superior border of the patella on the right lower limb), the ultrasound probe was placed perpendicular to the leg skin surface, with a sufficient amount of ultrasound gel applied to maintain minimal pressure between the probe and the skin, thereby obtaining images of the rectus femoris (RF)muscle.Two investigators independently measured the data on the original ultrasound screen, and the average value was taken. If the error between the two measurements exceeded 10%, a third experienced ultrasound physician performed the measurement again, and the average of the three measurements was calculated (Figure 3).

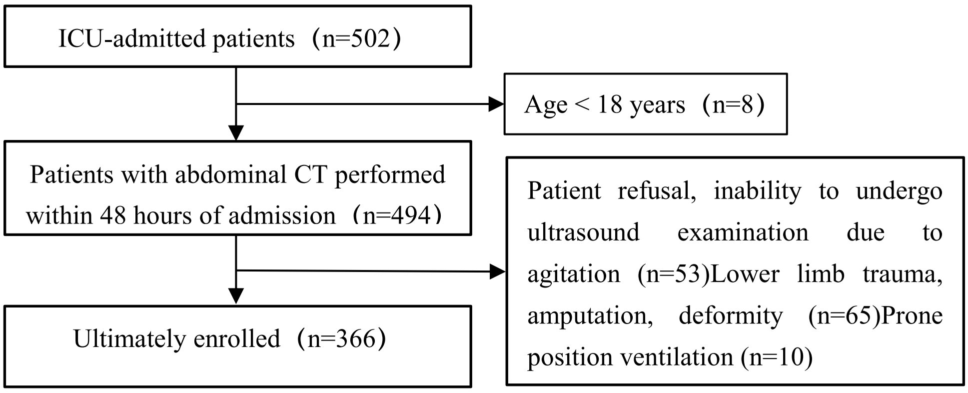


Figure1 Flow Chart

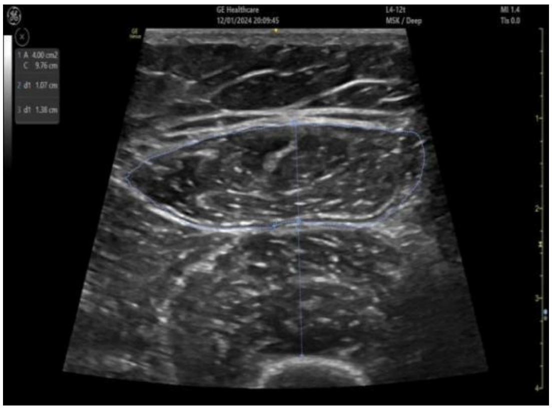


Figure 2 L3 CT Skeletal Muscle Area Image

Figure 3 Rectus Femoris Muscle image

2.1 Statistical Analysis

All data in this study were subjected to statistical analysis using SPSS 26.0 software.Descriptive Analysis of DataFor quantitative data: The Shapiro-Wilk test was used to determine whether the data conformed to a normal distribution.If the data conformed to a normal distribution, they were expressed as mean ± standard deviation ($\bar{x} \pm s$);If the data did not conform to a normal distribution, they were expressed as median M (P25, P75) .For categorical data: They were expressed as frequency (percentage).Inferential Statistical Analysis For comparisons of normally distributed quantitative data, the t-test was used.For comparisons of non-normally distributed quantitative data and ordinal data, the rank-sum test (Mann-Whitney Test) was used.The diagnostic efficacy of ultrasound for sarcopenia was analyzed using three indicators: sensitivity, specificity, and the receiver operating characteristic (ROC) curve.

Table1 CT-based Population Characteristics of ICU Sarcopenia Patients

	Population(n=366)		Z	*P*	Man(n=185)		Z	*P*	Female(n=181)		Z	*P*
	Sacopemia (n=122)	No sarcopenia (n=244)			Sacopemia (n=77)	No sarcopenia (n=108)			Sacopemia (n=45)	No sarcopenia (n=136)		
Age（year）	66.00(57.00,81.00)	58.00(45.00,71.00)	-4.557	＜0.001	68.00(55.00,77.00)	57.00(43.00,93.50)	-3.927	＜0.001	65.00(59.00,82.00)	60.00(46.00,74.00)	-2.932	0.003
BMI(kg/m^2）	21.15(18.37,23.10)	24.20 (21.45,26.65)	-7.084	＜0.001	20.94(18.40,22.90)	25.10(22.50,27.60)	-6.901	＜0.001	21.50 (18.25,25.39)	23.40 (20.80,25.40)	-3.267	0.001
SMA(cm^2)	100.69(77.42,120.80)	126.85(99.02,152.20)	-7.790	＜0.001	116.50 (102.1,126.70)	153.40 (140.60,169.55)	-10.978	＜0.001	75.34(70.56,80.11)	101.00 (91.68,114.35)	-9.868	＜0.001
SMI（cm^2/m^2）	35.71(31,53,41.73)	46.49(39.38,52.48)	-5.071	＜0.001	39.86(36.48,43.25)	48.95(48.72,56.67)	-11.581	＜0.001	31.53(27.37,32.22)	40.24(37.13,45.04)	-10.045	＜0.001
RFCSA（cm^2）	2.70(1.93,3.58)	3.13(2.49,4.15)	-10.862	＜0.001	3.11(2.65,371)	4.25(3.57,5.19)	-6.817	＜0.001	1.90(1.70,2.33)	2.66(2.21,3.23)	-4.734	＜0.001
RFMT（cm）	0.93(0.83,1.21)	1.21(0.98,1.42)	-5.993	＜0.001	1.00(0.93,1.26)	1.35(1.23,1.62)	-6.749	＜0.001	0.83(0.71,0.91)	1.05(0.90,1.26)	-5.351	＜0.001

3 RESULTS

Analysis of the characteristics of sarcopenia in ICU patients based on CT diagnosis is shown in Table 1.atients with sarcopenia are more likely to be older and have a lower BMI, and this trend is consistent in both males and females.Whether referring to indicators reflecting "whole-body muscle mass" (SMA, SMI) or those reflecting "local muscle mass of the lower limbs" (RFCSA, RFMT), the sarcopenia group showed significantly lower values than the non-sarcopenia group (all $P<0.001$ or $P=0.001$)—a finding that fully aligns with the core definition of sarcopenia as "reduced muscle mass".The absolute values of SMA, SMI, RFCSA, and RFMT in males were all higher than those in females (a physiological difference). However, the trend of differences between "sarcopenia vs. non-sarcopenia" was completely consistent in both male and female subgroups, indicating that the muscle mass characteristics of sarcopenia have no gender specificity.Analysis of the diagnostic efficacy of ultrasound for sarcopenia in ICU patients is shown in Table 2.Both RFCSA and RFMT exhibit moderate to moderately good diagnostic efficacy for sarcopenia in ICU patients (AUC: 0.736–0.794), with no "ineffective diagnostic" indicators observed.Females performed better in terms of sensitivity (84.4%, resulting in fewer missed diagnoses), while males showed superior specificity in RFCSA (72.2%, leading to fewer misdiagnoses).The combined use of RFCSA and RFMT can slightly improve diagnostic efficacy, with a more pronounced improvement observed in male patients (with the AUC exceeding 0.8).The ROC curve for ultrasound diagnosis of sarcopenia in ICU patients is shown in Figure 4 and Figure 5.

Table2 Efficacy Analysis of Ultrasonic Diagnosis for Sarcopenia in ICU

	RFCSA		RFMT		Joint Prediction	
	Man	Female	Man	Female	Man	Female
AUC	0.794	0.766	0.791	0.736	0.807	0.768
Cut off	3.730	2.340	1.250	0.920	/	/
Sensitivity	0.753	0.844	0.753	0.844	/	/
Specificity	0.722	0.706	0.657	0.721	/	/

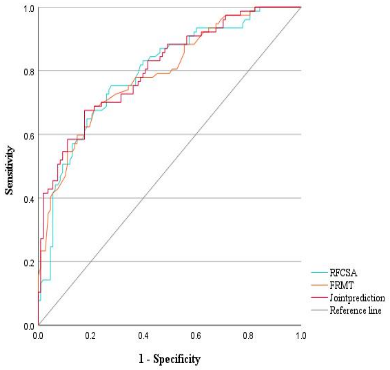


Figure 4 ROC Curve for Ultrasonic Diagnosis in Males

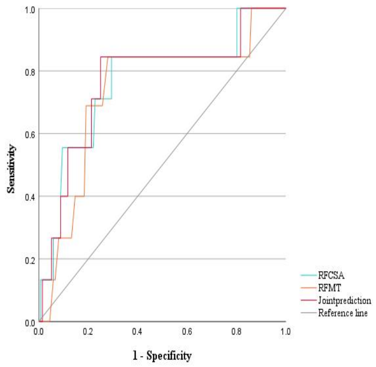


Figure 5 ROC Curve for Ultrasonic Diagnosis in Females

4 DISCUSSION

Ultrasound can achieve moderate diagnostic efficacy using either RFCSA or RFMT alone, meeting the need for "rapid screening of sarcopenia" in the ICU—given the high portability of bedside ultrasound, which is suitable for critically ill patients.While the combined prediction of RFCSA + RFMT can improve diagnostic accuracy, it is particularly suitable for scenarios where "precise identification of sarcopenia is required to guide nutritional support/rehabilitation interventions".It is essential to set cut-off values based on gender; otherwise, diagnostic accuracy will be severely compromised.If prioritizing the avoidance of missed diagnoses (e.g., screening for high-risk patients in the ICU), RFCSA/RFMT may be preferred (with 84.4% sensitivity in females). If prioritizing the avoidance of misdiagnoses (e.g., preventing unnecessary interventions), RFCSA is preferred for males (72.2% specificity) and RFMT is preferred for females (72.1% specificity).

Ultrasound has emerged as the preferred tool for diagnosing sarcopenia in ICU patients, primarily due to its bedside immediacy and radiation-free nature—attributes that align perfectly with the special needs of critically ill individuals who require frequent, non-harmful assessments. Notably, its efficacy conclusions are consistent with findings from multiple field studies, further validating its clinical value.In a multicenter study published by [18] et al., ultrasound was used to diagnose sarcopenia in ICU patients. The results showed that the Area Under the Curve (AUC) of the Rectus Femoris Cross-Sectional Area (RFCSA) was 0.80 for male patients and 0.70 for female patients. These figures are highly consistent with the data from the present study, where the AUC of RFCSA was 0.794 for males and 0.766 for females[17]. Both studies collectively confirm that a single ultrasound parameter (e.g., RFCSA) can meet the demand for rapid sarcopenia screening in ICU settings, streamlining the diagnostic process without compromising accuracy.Regarding diagnostic efficacy, the present study also verified that the diagnostic performance of rectus femoris cross-sectional area (RFCSA) for sarcopenia is superior to that of rectus femoris thickness—a conclusion that aligns with the consensus of most current researchers. This finding provides clear guidance for clinical practice: when selecting muscle sites for ultrasound assessment, researchers and clinicians should prioritize lower limb muscles. This preference is largely attributed to the relatively large volume and stable anatomical structure of lower limb muscles, which are less affected by patient positioning— a critical advantage when examining critically ill individuals in the ICU who often have limited mobility and require stable life support. When choosing measurement parameters, cross-sectional area should be the priority, as it more comprehensively reflects overall muscle mass compared to linear thickness, which may be influenced by local tissue edema or compression. Additionally, it is essential to establish different cutoff values based on gender, considering the inherent physiological differences in muscle distribution and baseline muscle mass between males and females—this personalized adjustment helps further improve the accuracy of ultrasound diagnosis for sarcopenia in ICU patients.

This study verified that the combined diagnostic value of skeletal muscle cross-sectional area (CSA) + muscle thickness (MT) based on ultrasound imaging for sarcopenia in ICU patients is higher than that of CSA or MT alone, providing evidence for the following conclusions:Ultrasound may offer higher diagnostic accuracy for sarcopenia in critically ill patients;When diagnosing sarcopenia using different ultrasound parameters, priority can be given to muscle cross-sectional area (CSA).However, this study also has certain limitations. It only used conventional two-dimensional ultrasound (transverse section) to construct a single-modal ultrasound radiomics model. In the future, researchers can further integrate conventional two-dimensional ultrasound (transverse and longitudinal sections), contrast-enhanced ultrasound, and elastography to construct multi-modal ultrasound radiomics models. This integration will enable multi-dimensional and more detailed feature extraction, which helps reduce subjective variability among clinicians, thereby further improving diagnostic accuracy.

5 CONCLUSION

In conclusion, this study reveals that ultrasound, as an assessment tool for diagnosing sarcopenia in ICU patients, exhibits favorable diagnostic efficacy. Specifically, the combination of the Rectus Femoris Cross-Sectional Area

(RFCSA) and Rectus Femoris Thickness (RFMT) yields higher diagnostic cutoff values. In the future, ultrasound—this non-invasive tool—can be used to perform early screening for sarcopenia in ICU patients.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Liu Dandan, Tan Lingling, Yang Ming, et al. Research Progress of Acute Sarcopenia. Chinese Journal of Multiple Organ Diseases in the Elderly, 2023, 22(1): 71-75.
[2] Huang D-D, Ji Y-B, Zhou D-L, et al. Effect of surgery-induced acute muscle wasting on postoperative outcomes and quality of life. Journal of Surgical Research, 2017, 218: 58–66.
[3] Mayer K P, Thompson Bastin M L, Montgomery-Yates A A, et al. Acute skeletal muscle wasting and dysfunction predict physical disability at hospital discharge in patients with critical illness. Critical Care, 2020, 24(1): 637.
[4] Cruz-Jentoft A J, Bahat G, Bauer J, et al. Sarcopenia: revised european consensus on definition and diagnosis. Age and Ageing, 2019, 48(1): 16–31.
[5] Wu R-Y, Sung W-H, Cheng H-C, et al. Investigating the rate of skeletal muscle atrophy in men and women in the intensive care unit: a prospective observational study. Scientific Reports, 2022, 12(1): 16629.
[6] Fazzini B, Märkl T, Costas C, et al. The rate and assessment of muscle wasting during critical illness: a systematic review and meta-analysis. Critical Care, 2023, 27(1): 2.
[7] Akan B, Gokcinar D, Damgaci L, et al. Sarcopenia detected in aged patients in intensive care units is associated with poor prognosis. Authorea Preprints, 2021.
[8] Jiang T, Lin T, Shu X, et al. Prevalence and prognostic value of preexisting sarcopenia in patients with mechanical ventilation: a systematic review and meta-analysis. Critical Care, 2022, 26(1): 140.
[9] Herridge M S, Diaz-Granados N, Cooper A, et al. Functional disability 5 years after acute respiratory distress syndrome. The New England Journal of Medicine, 2011.
[10] Mundi M S, Patel J J, Martindale R. Body composition technology: implications for the icu. Nutrition in Clinical Practice, 2019, 34(1): 48–58.
[11] Lopez-Ruiz A, Kashani K. Assessment of muscle mass in critically ill patients: role of the sarcopenia index and images studies. Current Opinion in Clinical Nutrition & Metabolic Care, 2020, 23(5): 302–311.
[12] Steele S, Lin F, Le T-L, et al. Segmentation and linear measurement for body composition analysis using slice-o-matic and horos. Journal of Visualized Experiments, 2021(169): 61674.
[13] Mourtzakis M, Prado C M M, Lieffers J R, et al. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. Applied Physiology, Nutrition, and Metabolism, 2008, 33(5): 997–1006.
[14] Tagliafico A S, Bignotti B, Torri L, et al. Sarcopenia: how to measure, when and why. La radiologia Medica, 2022, 127(3): 228–237.
[15] Giovannini S, Brau F, Forino R, et al. Sarcopenia: diagnosis and management, state of the art and contribution of ultrasound. Journal of Clinical Medicine, 2021, 10(23): 5552.
[16] Lambell K J, Tierney A C, Wang J C, et al. Comparison of ultrasound-derived muscle thickness with computed tomography muscle cross-sectional area on admission to the intensive care unit: a pilot cross-sectional study. Journal of Parenteral and Enteral Nutrition, 2021, 45(1): 136–145.
[17] Fu H, Wang L, Zhang W, et al. Diagnostic test accuracy of ultrasound for sarcopenia diagnosis: a systematic review and meta-analysis. Journal of Cachexia, Sarcopenia and Muscle, 2023, 14(1): 57–70.
[18] Nakanishi N, Miyajima I, Saijo T, et al. Cut-off values for an ultrasound-based muscle mass assessment at intensive care unit admission: a retrospective study. Clinical Nutrition Open Science, 2024, 55: 263–273.
[19] Zeng X, Shi Z, Yu J, et al. Sarcopenia as a prognostic predictor of liver cirrhosis: a multicentre study in china. Journal of Cachexia, Sarcopenia and Muscle, 2021, 12(6): 1948–1958.
[20] Van Ruijven I M, Stapel S N, Molinger J, et al. Monitoring muscle mass using ultrasound: a key role in critical care. Current Opinion in Critical Care, 2021, 27(4): 354–360.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3056

THE EFFICACY OF ELECTROTHERAPY WITH BIAN STONE IN TREATING FUNCTIONAL CONSTIPATION IN THE ELDERLY

Xia Li#, Fei Wang#, Yan Qin, Erdi Wang, Jie Zhang*
Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 200071, China.
#Xia Li and Fei Wang are both the first authors.
**Corresponding Author: Jie Zhang*

Abstract: Objective: The objective of the study is to assess the effect of a combined therapy of electric heating and Bian stone use with functional constipation in aged patients aiming to come up with a non-drug teaching method and to obtain a subjective reference of applying the intervention of traditional Chinese medicine non-drug to treat chronic diseases. **Methods:** Sixty-six elderly constipated patients were recruited notwithstanding the inclusion criteria. The patients were randomly split into a group of patients undergoing stone therapy and a control group (n=33 each) using Stata 15.1 software. Both conditions were treated to conventional monotherapy, using lactulose and glycerol suppositories, and the control group also had the pseudo clinical nursing care. Moreover, the Bian stone therapy group went through an extensive course of electrotherapy using Bian stones. This system involved the use of Chinese herbal medicine as a vehicle, the application of acupoints, and localized four energy methods, including rubbing, vibration, tapping, and warming. The treatment was done as 20 minutes per day during a 10-day course of the therapy. The scores of constipation symptoms were measured in both groups before the treatment and in day five and day ten. Symptomatic efficacy and quality of life were also observed by the time to the first bowel movement after treatment. **Results:** The overall efficacy rate of the stone treatment group, at the end of the treatment, was 93.94% which was much better than that of the control group, which was 54.84. Median time taken to pass bowel movement was 18hours in the stone therapy group and 34hours in the control group and the differences were statistically significant ($P<0.01$). Day 5 of treatment, both groups proved to have a significant improvement in the constipation symptom score levels with the stone therapy group proving better than the other in terms of the time to first bowel movement score, feeling of incomplete evacuation score, and the abdominal distension score ($P < 0.05$). Compared to the situation prior to the intervention, quality of life symptom scores improved in both groups following the intervention ($P < 0.01$) and stone therapy but not the control group had significantly lower scores across all domains ($P < 0.01$). **Conclusions:** To treat functional constipation of older patients, a combination of conventional clinical and electro-therapeutic stone therapy has better effects. The approach reduces the time to the first bowel movement by a significant margin, decreases the symptoms of constipation, and increases the quality of life of patients. Hence, electric-therapeutic stone method has unique paths to cure functional constipation among elderly people.

Keywords: Electro-heated stone therapy; Elderly; Functional constipation; Non-pharmacological traditional Chinese medicine therapy; Clinical efficacy

1 INTRODUCTION

Constipation is a condition, which is predominantly characterized by the inability to pass stools, longer periods between bowel movements, dry and hard stool, or an unresolved feeling of incomplete evacuation regardless of defecation urgency [1]. One of the gastrointestinal problems that is the most prevalent with the elderly population is functional constipation. As the number of elderly people is rapidly increasing in China because of the ageing population, it has been observed that the instances of constipation in older adults are also increasing by the year end and more than a quarter of them are reporting symptoms [2]. According to several surveys, local and foreign, the prevalence of constipation among older persons, in home care provision or nursing facilities, can even reach 50% [3,4].

Even though constipation per se is not a significant mortal threat, its presence raises the risk of the related diseases considerably, in particular, it can advance the risks of the onset of colorectal polyps and, possibly, lead to the development of colorectal cancer. Secondly, long-term straining when using the restroom may lead to cardiovascular and cerebral accidents, including congestive heart illness and hemorrhage of the brain [5,6]. A significant psychological load on the elderly is also constipation, which seriously impacts their everyday life and the overall quality of health, which should be given universal consideration [7].

It is widely agreed nowadays that the pharmacological therapies of functional constipation in the aging population depend basically on the use of chemical medications. Although quick symptom relief is achieved with these medications the drugs are always accompanied by side effects including high recurrence rate, nonoptimal efficacy at a long term and dependence on the drugs. The Chinese non-pharmacological therapies traditionally are aimed at preventing and curing the disease without medication using the application of traditional methods to regulation of the yin-yang balance of the body. One of the four major non-pharmacological therapies of traditional Chinese medicine (along with acupuncture, moxibustion and Daoyin massage) is called Bian stone therapy because the object in this

therapy is the Bian stones that have specific physical characteristics and which are used to treat acupoints and meridians. Its efficacy in the treatment of functional constipation has been verified in the previous works.
Nonetheless, keeping a constant temperature of the stone is a technical problem with Bian stone therapy. High temperature may lead to certain undesirable side effects like burns and blisters whereas low temperature does not bring the intended therapeutic effect [8]. Electrically heated stone that is employed in the current study is made of natural stone material that is supplemented with semiconductor parts and is governed by a microcomputer program. The innovation helps in overcoming the conventional limitations of the inconveniency in heating and inability to sustain the temperatures over long durations making sure that the stone is kept at the same temperature over exercising and tackling the technical challenges posed by stone therapy. Informed by the traditional Chinese medicine principles in the areas of the viscera and meridians, a complex conventionalized electrothermally heated stone integrated diagnosis and treatment regimen was further refined. In this research, they used the electrothermally heated stone protocol to intervene in elderly patients with constipation that was functional and the results of this intervention were favorable as outlined below.

2 SUBJECTS AND METHODS

2.1 Study Population

Sixty-six elderly patients who came to the tertiary hospital in Shanghai and got diagnosed with functional constipation between February 2023 and February 2024 were randomly assigned to one of the two conditions, i. e. an observation group and a control group (n=33 each). This study received approval from the hospital's ethics committee. During the study, two patients were excluded from the control group: one for poor compliance (self-administering other constipation medications) and the other for premature discharge before completing the required treatment duration. At the study's completion, there were 33 effective cases in the stone therapy group and 31 in the control group. A comparison of general characteristics (age, gender, height, weight) between the groups showed no statistically significant differences ($P > 0.05$), satisfying the study's requirements. Detailed results are presented in Table 1.

2.2 Inclusion Criteria

① Patients meeting the Rome IV diagnostic criteria for functional constipation; ②Aged between 60 and 89 years, regardless of gender; ③ Voluntarily signed informed consent and able to cooperate.

2.3 Exclusion Criteria

① Patients with concurrent cardiovascular disease or organic organ dysfunction (e.g., pulmonary, hepatic, renal); ②Patients with psychiatric disorders or cognitive impairment that impede communication; ③ Constipation due to intestinal strictures caused by colorectal organic lesions; ④Patients with advanced tumors; ⑤ Constipation resulting from medication factors; ⑥ Patients with abdominal skin damage or severe skin diseases that prevent massage and heat application; ⑦ Patients with bleeding or coagulation abnormalities due to severe primary hematopoietic system diseases.

Table 1 Comparison of General Characteristics Between Two Groups (n=64)

Group	Number of Cases	Age (years)	Gender (n, %)		Height (m)	Weight(kg)
		[M(P25, P75)]	Male	Female	(x±s)	(x±s)
Stone therapy group	33	67(63,70)	18(54.5)	15(45.4)	1.67±0.086	63.68±12.218
Control group	31	70(66,75)	11(35.4))	20(64.5)	1.64±0.073	62.31±13.469
Statistic		Z=-1.476	χ^2=2.344		t=-1.576	t=-0.428
P		0.14	0.126		0.12	0.67

2.4 Research Methods

2.4.1 Control group intervention

Patients in the control group received conventional therapeutic medication, specifically lactulose oral solution, taken orally after meals at a dose of 2 vials once daily for 10 consecutive days. Enema suppositories (20ml) were used as needed. Additionally, they were provided with routine clinical care, including personal care, dietary management, exercise guidance, abdominal massage instruction, and psychological support.

2.4.2 Intervention method for the Bian Stone Group

In addition to routine clinical care provided to the control group, patients in the Bian Stone Group underwent the following specific procedures: The practitioner stood to the patient's left side while the patient lay supine with a pillow under the knees. The formula was Tongfu Formula which included Raw rhubarb 2g, Zhi shi 1g, Aloe 2g,1g Areca nut that was used to enhance the spleen, improve digestion, and balance qi and blood. Electrically heated stone was applied on the stone as follows:

1. Abdominal Rubbing: The stone was applied round the belly button, and rubbed in a clockwise rotary motion of 10 minutes until warmth was experienced and a feeling of distension being felt along the lower left abdomen was obtained;
2. Vibration Massage: Viable, One side of the stone was laid against the left lower abdomen of the patient and the vibrations, with a high frequency of more than 200 vibrations per minute were done by flexing up and down in 1 minute;
3. Acupoint Stimulation: The bilateral Tianshu (ST25), Guanyuan (CV4) and Zhongwan (CV12) points were stimulated with the corner of the electrically heated stone with 1 minute pressure per point with the aim of a sore and distending feeling;
4. Warming Technique: A 45degC and warm rock was rolled over the umbilicus and lasted a point of 5 minutes. The entire treatment period was 20 minutes and the patients underwent this whole course of electro-heated stones therapy after one day of treatment of 10 days in a row.

2.5 Evaluation Methods

Four indicators were used to evaluate clinical efficacy in both groups namely: (1) Differences in syndrome efficacy (pre-treatment, day 10 post-treatment); (2) Time to first bowel movement post-treatment; (3) Constipation Symptom Score (pre-treatment, day 5 post-treatment, day 10 post-treatment); (4) Constipation Patient Quality of Life Scale (pre-treatment, day 10 post-treatment).

Details are as follows:

(1) Syndrome-Based Therapeutic Efficacy Assessment: The disease efficacy was assessed according to the guidelines on clinical research of new traditional Chinese medicines [9], when the nimodipine method was used. The syndrome score was calculated as follows: [(Pre-treatment score - Post-treatment score) / Pre-treatment score] × 100%. Complete Recovery: Normal bowel function was restored, and all related symptoms were resolved or substantially alleviated, with a syndrome score ≥ 95%. Marked Improvement: Primary symptoms and signs showed significant amelioration, with a syndrome score ranging from 70% to less than 95%. Effective: Primary symptoms and signs were significantly ameliorated, with a symptom score between 30% and less than 70%. Ineffective: There was no marked improvement or even a worsening of primary symptoms and signs, with a symptom score < 30%. Efficacy rate (%) = [(Cured + Markedly effective + Effective) ÷ Total number of cases] × 100%.

(2) Time to the First Bowel Movement after Treatment Initiation: The time to the first bowel movement measured after administration of treatment or the beginning of medication administration was taken on the two patient groups, starting at the beginning of medication or treatment.

(3) Constipation Symptom Score Sheet: This questionnaire was compiled in 2005 by the Surgery Division of the Chinese Medical Association that evaluated six aspects namely: ① Severity of defecation difficulty; ②_Faecal consistency; Time taken to defecate; ③Sensation of heaviness, ④incomplete evacuation, or distension; ⑤Weekly frequency of bowel movements; ⑥Abdominal bloating. The scoring system used was four points (0-3) in the sense that the higher score, the worse the symptoms of constipation. The symptom score constipation was the tally of scores in every of the six domains [10].(4) Chinese Version of the Patient-Specific Constipation Quality of Life Questionnaire (PAC-QOL): licensed, Mapi Research Trust (France) this special tool assessed the quality of life in patients with chronic constipation. It comprised 28 questions and four domains; physical status, psychological status, concerns, and social relationships/satisfaction with a 5-point Likert scale. The scales of 0-4 points were used to rate the level of discomfort as not at all, somewhat, somewhat not, somewhat, not, and extremely. Some items were reverse-scored. An increase in the scores marked a decrease in the quality of life and the total Cronbach a coefficient of the scale was 0.93 [11].

2.6 Statistical Methods

Statistical analysis was conducted using IBM SPSS 24 software. Quantitative data were expressed as mean ± standard deviation (x±s) and analyzed using t-tests; data not normally distributed were presented as medians and analyzed using the Mann-Whitney U test. Categorical data were presented as frequencies and percentages and analyzed using chi-square tests. Repeated measures analysis of variance was employed to compare pre- and post-treatment constipation symptom scores between patient groups. $P < 0.05$ was considered statistically significant.

3 RESULTS

3.1 Therapeutic Efficacy Evaluation in Both Groups Post-Treatment

After 10 days of treatment, the efficacy evaluation results for patients in the Bian stone group and the control group are presented in Table 2. A comparison indicated that the comprehensive therapy integrating electro-heated Bian stones significantly outperformed conventional clinical nursing paired with basic medication alone. Although the number of effective cases was similar between the two groups, the Bian stone group exhibited significantly more cured and markedly effective cases than the control group. Additionally, the control group had a greater proportion of ineffective cases, suggesting that conventional clinical care combined with basic medication yields less satisfactory results for elderly patients with functional constipation.

Table 2 Therapeutic Efficacy of Syndromes in Both Groups After Treatment (n=64)

Group	Number of Cases	Cured	Markedlyeffective	Effective	Ineffective	Overallresponserate %	Z	*P*
Stone therapy group	33	6	14	11	2	93.94%	-3.528	<0.001
Control group	31	3	4	10	14	54.84%		

3.2 Time to First Bowel Movement After Treatment in Both Groups

A Mann-Whitney U test was conducted to analyze the time to first anal defecation post-treatment in both groups. The results showed a statistically significant difference, with $P=0.015$ ($P<0.05$). The electro-thermotherapy group had a significantly shorter time to first bowel movement compared to the control group, indicating that the electro-thermotherapy treatment was notably more effective. See Table 3 for detailed information.

Table 3 Comparison of Time to First Anal Defecation Post-Treatment Between Two Patient Groups [M(P25, P75)]

Group	Time to First Anal DefecationPost-Treatment (h)	Z	*P*
Stone therapy group (n=33)	18(5.75,43)	-2.432	0.015
Control group (n=31)	34(12.5,70.5)		

3.3 Evaluation of Constipation Symptom Scores Before and After Treatment in Both Groups

When comparing the efficacy between the stone therapy group and the control group, it was found that the stone therapy group consistently had significantly lower total constipation symptom scores than the control group at each intervention time point ($p<0.05$). The control group did see statistically significant score reductions during treatment compared to pre-treatment levels ($P<0.05$). However, the differences in scores between the fifth and tenth days of treatment were not statistically significant ($P>0.05$), see Table 4.

Table 4 Comparison of Constipation Symptom Scores Before and After Treatment Between the Two Groups (x±s, points)

Group	Pre-treatment	Day 5 of Treatment	Day 10 of Treatment	F	*P*
Control group (n=31)	11.42±3.33	6.84±4.46△	6.84±4.10△◆	$F_{time}=158.91$	$P_{time}<0.001$
Stone therapy group (n=33)	11.88±2.90	3.12±2.85△	4.36±2.99△◇	$F_{group}=6.91$	$P_{group}=0.01$
t/Z	-5.950*	-2.725**	-3.396**	$F_{Interaction}=13.46$	$P_{Interaction}<0.001$
P	0.557	0.010	0.001		

Note: *Two-sample t-test, **Mann-Whitney U test: Compared with pre-intervention values in the same group, $^{\triangle}P < 0.05$, $^{\blacktriangle}P > 0.05$: Compared with values on day 5 of intervention in the same group, $^{\diamond}P < 0.05$, $^{\blacklozenge}P > 0.05$

3.4 Quality of Life Scores Before and After Treatment in Both Groups

Prior to treatment, the stone therapy group had a quality of life score of 49 ± 15.91, and the control group scored 46.39 ± 13.87, with no statistically significant difference between them ($P > 0.05$). After treatment, the stone therapy group's quality of life score improved to 16.91 ± 11.28, while the control group's was 31.16 ± 13.71. This post-treatment improvement was statistically significant ($P < 0.001$) (Table 5). These results highlight the added benefits of incorporating electrotherapy stone therapy.

Table 5 Comparison of Quality of Life Scores Before and After Treatment Between Groups (x±s, points)

Group	Pre-treatment	Aftertreatment	Difference	t	*P*
Stone therapy group (n=33)	49±15.91	16.9±11.28	32.09±16.9	10.905	＜0.001
Control group (n=31)	46.39±13.87	31.1±13.71	15.2±11.87	7.145	＜0.001
t/Z	-0.699^{*}	-3.965^{**}	-4.592^{*}		
P	0.487	＜0.001	＜0.005		

Note: *Two-sample t-test, **Mann-Whitney U test

4 DISCUSSION AND CONCLUSION

The findings of this paper indicate the patients of the stone therapy exhibited longer time to first anal defecation and symptomatic efficacy than the controls who underwent the intervention therapy after 10 days of therapy. This shows that the introduction of electrotherapy stone therapy into the mainstream clinical practice has better results that attest its high effectiveness in functional constipation treatment in the elderly. Such findings are in agreement with Tang Ying et al [12], who found that there was a lower rate of improvement in the symptom of constipation amongst patients undergoing dietary counseling and abdominal massage versus the group of patients receiving supplementary electro-thermal stone acupoint massage and acupoint plaster application.

The electrothermally heated group of stones had a much earlier first bowel movement after the treatment as compared to the control group and thus it would appear this treatment accelerates the process by comparison to the usual care. It is believed that this positive effect can be attributed to the resonance in which the energy waves within the human cells when they are touched by the stone are exposed to the internal electrons in the cells leading to the production of pulse frequencies ranging between 20-2000 kHz. With a combination of far-infrared, the application causes an increase of 0.5-2degC in body temperature and allows the formation of infrared thermal images following the meridian paths [13, 14]. As a result, aside from the anticonstipation properties of this practice, the electrically heated stone therapy used as an adjunct to treat functional constipation among the elderly population promotes the peripopular blood and microcirculation of their bowel and stool volumes. This is a general facilitation of visceral performance to aid the propulsion capacity of the large intestine waste.

The older patients who experience constipation frequently experience reduced organ functionality as a result of being old and weak. The impairment of the large intestine propulsion and fluid circulation is caused by kidney yin deficiency and qi deficit. A deficiency in the ability of the spleen-stomach, especially deficiency of the spleen, results in insufficient transformation and transportation which results in abdominal distension and poor appetite. This interferes with qi and blood production causing blood deprivation and consequent intestinal parchedness which materialized as dry and hard stool and constipated bowel movements. Consequently, functional constipation in the elderly is normally chronic and difficult to manage. The hours that are spent on it are often accompanied by depression, anxiety, and psychological discomfort significantly worsening the quality of life of patients [11]. The Expert Consensus on Constipation Grading and Clinical Management also indicates [15] that constipation can cause a number of negative emotional and psychological disorders. These negative effects, in turn, impact the gastrointestinal functioning via the autonomic route, increasing the rectal sensory threshold, decreasing the reflex defecation reflex, and increasing pelvic floor muscle tension, which leads to aggravation of the symptoms of constipation and formation of a vicious circle. This paper will prove that when integrated stone therapy intervention was administered, patients had better quality of life scores than in the control group which points to an improved quality of life. The action of electrically heated Bian stones has more force than the hand massage alone, penetrate to deeper areas and points of the body that are not available to the hand massage. It alters temporarily the abdominal wall by meaning that the lumen on the intestines, this acts to stimulate and enhance gastrointestinal reflexes, increase muscular activity to the digestive tract, and relaxation of the sphincters to hasten faecal expulsion [16]. Another method used in this study was the use of meridian acupoints and the four points that were most applicable (billetal Tianshu, Zhongwan and Qihai) were used. The Tianshu point is one of these, the point selected by adhering to the principle of choosing points near the affected area, and it is the Mu point of the large intestine in its area, mostly curing large intestine diseases. The Mu point of the stomach is Zhongwan on the Ren Meridian on the anterior midline, and the intersection of the viscera in all the Eight Convergence Points. It balances the stomach, strengthens the spleen, promotes qi, produces blood and eases constipation. According to the Universal Relief Formulas: in case of hard bowel movements, treat Zhongwan. One of the most important acupoints in the treatment of constipation is the Qihai point and it is the first point to be considered when restoring original qi. Together, all these four make the viscera feed and cause intestinal peristalsis as well as eliminate stool. This is also augmented with topical deposition of the herbal formula "Tongfu Fang" (Intestinal-Opening Formula) which gets into the skin and is able to internally loosen and control the intestines. The combination of these three complementary methods forms the electro-thermal stone integrated therapy to the treatment of functional constipation in elderly.

The combined results of the research prove that the use of the electro-thermal stone therapy actually enhances the first bowel movement time, symptom efficacy scores, and the quality of life of patients with constipation. This treatment mostly relieves the symptoms of constipation and minimizes the pain of patients by increasing the manual force of the practitioner to reach the depths and section that could not be achieved in the treatment solely using manual methods. Patient satisfaction was improved, compliance was also good, and there were no adverse reactions throughout the treatment process. Further studies can prove this combined therapy as a successful standardized TCM intervention of functional constipation in the elderly that can be applied clinically.

Treatment of constipation involves long-term strategies in treatment; this research project used short-term treatment and therefore a follow-up study should ensure a prolonged intervention period where the effect of the intervention is lasting and individual use of laxatives may have reduced.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This study was supported by the Shanghai Municipal Three-Year Action Plan for the Inheritance, Innovation, and Development of Traditional Chinese Medicine: Promotion System Construction Project for Shanghai-style Traditional Chinese Medicine Non-Pharmaceutical Therapies (Project No. ZY(2025-2027)-3-2-1).

REFERENCES

[1] Leng Yan, Wei Wei, Tang Xudong. Expert Consensus on Traditional Chinese Medicine Diagnosis and Treatment of Constipation. Journal of Traditional Chinese Medicine, 2025, 66(03): 321-328.

[2] SONG H J. Constipation in community-dwelling elders: prevalence and associated factors. Journal of Wound Ostomy and Continence Nursing, 2012, 39(6): 640-645.

[3] KOMIYA H, UMEGAKI H, ASAI A, et al. Prevalence and risk factors of constipation and pollakisuria among older home-care patients. International Journal of Geriatrics and Gerontology, 2019,19(4): 277-281.

[4] GUSTAFSSON M, LÄMÅS K, ISAKSSON U, et al. Constipation and laxative use among people living in nursing homes in 2007 and 2013. Maria Gustafsson; Kristina Lämås; Ulf Isaksson; Per-Olof Sandman; Hugo Lövheim, 2019, 19(1).

[5] SUNDBØLL J, SZÉPLIGETI S K, ADELBORG K, et al. Constipation and risk of cardiovascular diseases: a Danish population-based matched cohort study. BMJ open, 2020,10(9): e37080.

[6] ISHIYAMA Y, HOSHIDE S, MIZUNO H, et al. Constipation-induced pressor effects as triggers for cardiovascular events. J Clin Hypertens (Greenwich), 2019, 21(3): 421-425.

[7] Huang Caifeng, Liu Rui, Deng Yuqin, et al. Analysis of constitutional factors in constipation among the elderly. Nursing Research, 2023, 37(02): 347-350.

[8] Zheng Fang, Chen Changxiang, Cui Zhaoyi. Analysis of defecation abnormalities among elderly individuals across different age groups. Journal of North ChinaUniversity of Science and Technology (Medical Edition), 2021, 23(02): 143-147.

[9] Chinese Society of Coloproctology. Assessment of Constipation Symptoms and Treatment Efficacy. Chinese Journal of Gastrointestinal Surgery, 2005, 8(04): 355.

[10] Zhao Zhenzhen, Lin Zheng, Lin Lin, et al. Research on the Reliability and Validity of the Chinese Version of the Patient Constipation Status Assessment Scale in Application Evaluation. Chinese Journal of Nursing, 2010, 45(12): 1124-1126.

[11] Xu Yihui. Correlation study between related symptoms and quality of life in middle-aged and elderly patients with functional constipation. Electronic Journal of Modern Medicine and Health Research, 2020, 4(02): 190-192.

[12] Tang Ying, Xie Hongqiong, Zou Feiping. Application of Bian Stone Acupoint Massage Combined with Acupoint Plaster Application in Postoperative Constipation Following Lower Limb Fracture Surgery. Modern Distance Education in Chinese Medicine, 2021, 19(14): 159-161.

[13] Xie Xiande, Wang Fuyi, Xie Nanchu, et al. Mineralogical Study of Sibin Bian Stone I: Relationship between Rock Chemistry, Structural Characteristics and Infrared Emission Functionality. Bulletin of Mineralogy, Petrology and Geochemistry, 2008(01): 1-5.

[14] Xie Xiande, Sun Zhenya, Wang Fuyi, et al. Mineralogical Study of Sibin Bian Stone II: Relationship between Mineral Composition Characteristics and Infrared Emission Functionality. Bulletin of Mineralogy, Petrology and Geochemistry, 2008(01): 6-12.

[15] 2017 Expert Consensus on Classification and Clinical Management Strategies for Constipation. Chinese Journal of Gastrointestinal Surgery, 2018, 21(03): 345-346.

[16] SINCLAIR M L. The use of abdominal massage to treat chronic constipation. Journal of Bodywork and Movement Therapies, 2010, 15(4): 436-445.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3057

FACTORS INFLUENCING ENTERAL FEEDING INTOLERANCE IN NEUROCRITICAL PATIENTS BASED ON ADMISSION DATA

ZhongYi Wu
Department of Critical Care Medicine, The First Affiliated Hospital of Soochow University, Suzhou 215001, Jiangsu, China.

Abstract: Objective: To identify risk factors for feeding intolerance (FI) among neurocritical patients based on admission data.Methods: A retrospective study was conducted using convenience sampling to analyze neurocritical patients who were admitted to the neurology department of a tertiary hospital in Suzhou and received enteral nutrition between January 2018 and December 2022. Univariate and multivariate logistic regression analyses were performed to determine independent predictors of FI.Results: A total of 291 patients were included, of whom 119 (40.9%) developed FI. Univariate analysis showed that age, history of diabetes, history of stroke, Glasgow Coma Scale (GCS) score, and Nutritional Risk Screening 2002 (NRS-2002) score were significantly associated with FI ($P<0.05$). Multivariate logistic regression identified history of diabetes (OR=2.687), history of stroke (OR=2.352), lower GCS score (OR=0.750), and higher NRS-2002 score (OR=1.467) as independent risk factors for FI.Conclusion: History of diabetes, history of stroke, lower GCS score, and higher NRS-2002 score are major risk factors for FI in neurocritical patients. Early identification of these high-risk characteristics using admission data may support timely, targeted interventions to improve feeding tolerance and enhance the effectiveness of enteral nutrition.

Keywords: Neurocritical patients; Enteral nutrition; Feeding intolerance; Nursing

1 INTRODUCTION

Neurocritical patients often present with dysphagia and impaired consciousness, resulting in restricted or absent oral intake[1]. Early initiation of enteral nutrition is essential for maintaining the integrity of the gastrointestinal mucosal barrier and preventing gut-derived infections[2]. However, due to multiple pathological factors—such as elevated intracranial pressure, dysregulation of the brain–gut axis, and central nervous system suppression[3]—these patients are highly susceptible to feeding intolerance (FI), most commonly occurring within 1–3 days after the initiation of enteral nutrition, with an incidence reported as high as 50%[4]. FI not only compromises the safety and effectiveness of enteral nutrition but may also delay recovery and adversely affect patient outcomes. Therefore, early identification of patients at high risk for FI is crucial for optimizing enteral nutrition strategies.Admission-stage clinical data—including demographic characteristics, comorbidities, neurological status, and nutritional risk—are easy to obtain, stable, and suitable for early risk assessment and rapid clinical decision-making. However, previous studies have largely focused on gastrointestinal function and disease severity–related dynamic indicators, with insufficient attention to high-risk factors that can be identified at admission.Accordingly, this study aimed to identify accessible and clinically meaningful independent risk factors for FI in neurocritical patients based on admission information. Univariate and multivariate logistic regression analyses were performed to screen predictors of FI, with the goal of supporting early risk assessment, guiding feeding strategies, and promoting individualized management to improve the safety and effectiveness of enteral nutrition.

2 SUBJECTS AND METHODS

2.1 Study Population

This retrospective study used a convenience sampling method to include neurocritical patients who were hospitalized in the neurology department of a tertiary hospital in Suzhou and received enteral nutrition between January 2018 and December 2022. The inclusion criteria were as follows: (1) age ≥18 years; (2) confirmed diagnosis of neurological disease; (3) issuance of a "critical condition" medical order after admission; and (4) initiation of enteral nutrition via a nasogastric tube within 24 hours of admission. The exclusion criteria were: (1) contraindications to enteral nutrition, such as intestinal obstruction, gastrointestinal ischemia, severe diarrhea, or intractable vomiting; (2) receipt of enteral nutrition prior to admission; and (3) discontinuation of enteral nutrition for reasons other than feeding intolerance. A total of 291 patients met the criteria and were included in the final analysis. This study was approved by the Ethics Committee of the hospital (Approval No. 2024-618).

2.2 Enteral Nutrition Procedure

Commercial enteral nutrition formulas were administered according to physician orders and delivered continuously via an enteral feeding pump. The initial infusion rate was set at 20–50 mL/h, and the feeding solution was maintained at

38°C using a heater. During feeding, the head of the patient's bed was elevated to at least 30°. The responsible nurse assessed gastrointestinal tolerance every 4 hours, including gastric residual volume, vomiting, diarrhea, and aspiration.If the patient exhibited good tolerance, the infusion rate was increased to 50–80 mL/h on the following day and further adjusted to 80–120 mL/h on the third day, progressing toward the target feeding volume based on the patient's condition. If signs of gastrointestinal intolerance occurred, enteral feeding was immediately stopped. Tolerance was reassessed after 6 hours, and the infusion rate and total volume were readjusted accordingly.

2.3 Criteria for Feeding Intolerance

Currently, there is no universally accepted international standard for defining feeding intolerance (FI). Based on clinical practice and the European Society of Intensive Care Medicine (ESICM) abdominal compartment group's definition of FI[5], the criteria used in this study for determining FI in neurocritical patients receiving enteral nutrition were as follows:Vomiting: expulsion of gastric contents accompanied by visible gastric material;Diarrhea: ≥3 bowel movements per day with watery or loose stool consistency;Abdominal distension: decreased bowel sounds, abdominal bloating, or tympanic percussion note;Gastric retention: gastric residual volume >250 mL within 4 hours;Gastrointestinal bleeding: coffee-ground gastric contents, recurrent hematemesis or melena, or unexplained reductions in hemoglobin.

2.4 Instruments and Data Collection

Based on literature review, availability of clinical data, and expert consultation, a self-designed General Information Questionnaire was developed, consisting of three components:Demographic and medical history information: sex, age, body mass index (BMI), admission temperature, admission diagnosis, smoking history, alcohol use, history of hypertension, diabetes, heart disease, and stroke;Admission assessment indicators: Glasgow Coma Scale (GCS) score, Braden Pressure Ulcer Risk Score, Nutritional Risk Screening 2002 (NRS-2002) score, Barthel Index (BI), and Morse Fall Scale (MFS) score;Neurological assessments: results of the Kubota water-swallowing test and limb muscle strength evaluation.Data were obtained from the Haitai electronic medical record system and the inpatient nursing information system. Two trained researchers independently entered the data in duplicate and performed cross-checking to ensure accuracy and completeness.

2.5 Statistical Analysis

Data were analyzed using SPSS version 24.0. Quantitative variables with a normal distribution were expressed as mean ± standard deviation and compared between groups using the independent-samples t test. Quantitative variables with a non-normal distribution were expressed as median and percentiles, and between-group comparisons were performed using the Mann–Whitney U test. Categorical variables were described as frequencies and percentages, and compared using the chi-square test or Fisher's exact test.Univariate analysis and collinearity diagnostics were conducted in SPSS 24.0. A variance inflation factor (VIF) ≥10 was considered to indicate severe multicollinearity among independent variables. Variables with $P < 0.1$ in univariate analysis were included in the multivariate logistic regression model. A P value < 0.05 was regarded as statistically significant.

3 RESULTS

3.1 General Characteristics of the Study Population

A total of 291 patients were included in this study, among whom 119 (40.9%) developed feeding intolerance (FI). Patients were divided into an FI group (n = 119) and a non-FI group (n = 172) based on the occurrence of FI. The general characteristics of the two groups are presented in Table 1.

3.2 Univariate Analysis of Factors Associated With Feeding Intolerance in Neurocritical Patients

Univariate analyses were conducted to compare demographic characteristics, admission assessment indicators, and neurological evaluation variables between patients who developed FI and those who did not. Significant differences were observed between the two groups in age, history of diabetes, history of stroke, GCS score, and NRS-2002 score ($P < 0.05$). Detailed results are shown in Table 1.

Table 1 Univariate Analysis of Factors Associated With Feeding Intolerance in Neurocritical Patients

Variable	Category	FI group (n = 119)	Non-FI group (n = 172)	Statistic	*P*
Sex, n (%)	Male	72 (24.7)	93 (32.0)	1.186[1]	0.276
	Female	47 (16.2)	79 (27.1)		
Age, n (%)	<65 years	22 (7.6)	57 (19.6)	7.635[1]	0.006
	≥65 years	97 (33.3)	115 (39.5)		
BMI, n (%)	Underweight	18 (6.2)	15 (5.1)	4.945[1]	0.084
	Normal	81 (27.8)	123 (42.3)		
	Overweight	20 (6.9)	34 (11.7)		

Temperature, °C, M (P25, P75)	—	36.8 (36.5, 37.3)	36.9(36.6, 37.3)	−0.496[2]	0.620
Admission diagnosis, n (%)	Cerebral infarction	103 (35.4)	148 (50.8)	1.077[1]	0.584
	Cerebral hemorrhage	10 (3.4)	11 (3.8)		
	Others	6 (2.1)	13 (4.5)		
Smoking history, n (%)	Yes	17 (5.8)	32 (11.0)	0.973[1]	0.333
	No	102 (35.1)	140 (48.1)		
Alcohol history, n (%)	Yes	12 (4.1)	19 (6.5)	0.068[1]	0.794
	No	107 (36.8)	153 (52.6)		
Hypertension, n (%)	Yes	92 (31.6)	124 (42.6)	1.001[1]	0.317
	No	27 (9.3)	48 (16.5)		
Diabetes, n (%)	Yes	46 (15.8)	38 (13.0)	9.397[1]	0.002
	No	73 (25.1)	134 (46.1)		
Heart disease, n (%)	Yes	40 (13.8)	44 (15.1)	2.210[1]	0.137
	No	79 (27.1)	128 (44.0)		
History of stroke, n (%)	Yes	34 (11.7)	27 (9.3)	7.036[1]	0.008
	No	85 (29.2)	145 (49.8)		
Kubota water-swallowing test, n (%)	<grade 3	15 (5.2)	33 (11.3)	2.525[1]	0.112
	≥grade 3	106 (36.4)	137 (47.1)		
Muscle strength, n (%)	<grade 3	39 (13.4)	45 (15.5)	1.497[1]	0.221
	≥grade 3	80 (27.5)	127 (43.6)		
GCS score, M (P25, P75)	—	11 (9, 12)	12 (11, 14)	−5.509[2]	<0.001
Braden score, M (P25, P75)	—	12 (11, 14)	13 (11, 15)	−0.668[2]	0.504
NRS-2002 score, M (P25, P75)	—	3 (2, 4)	3 (2, 3)	−5.188[2]	<0.001
BI score, M (P25, P75)	—	0 (0, 20)	0 (0, 30)	−1.295[2]	0.195
MFS score, M (P25, P75)	—	50 (35, 50)	50 (35, 55)	−0.194[2]	0.881

NOTE:[1] Chi-square statistic,[2] Z statistic

3.3 Logistic Regression Analysis of Factors Associated With Feeding Intolerance in Neurocritical Patients

The five variables with $P < 0.05$ identified in the univariate analysis were examined for multicollinearity, and no significant multicollinearity was detected. Variables that showed statistically significant differences in the univariate analysis were then entered into a multivariate logistic regression model, with variable coding presented in Table 2. The results indicated that history of diabetes, history of stroke, Glasgow Coma Scale (GCS) score, and Nutritional Risk Screening 2002 (NRS-2002) score were independent predictors of feeding intolerance (FI) in neurocritical patients (Table 3).

Table 2 Variable Coding for Logistic Regression Analysis

Variable	Coding Method
Age	<65 years = 0; ≥65 years = 1
History of diabetes	No = 0; Yes = 1
History of stroke	No = 0; Yes = 1
GCS score	Entered as a continuous variable
NRS-2002 score	Entered as a continuous variable

Table 3 Logistic Regression Analysis of Factors Associated With Feeding Intolerance in Neurocritical Patients

Variable	β	SE	Wald χ^2	P value	OR	95% CI
Age	0.163	0.341	0.230	0.632	1.177	0.604–2.296
History of diabetes	0.988	0.297	11.057	0.001	2.687	1.501–4.812
History of stroke	0.855	0.326	6.862	0.009	2.352	1.240–4.459
GCS score	−0.288	0.065	19.952	<0.001	0.750	0.661–0.851
NRS-2002 score	0.383	0.129	8.876	0.003	1.467	1.140–1.888

4 CONCLUSION AND DISCUSSION

4.1 Current Status of Feeding Intolerance in Neurocritical Patients Receiving Enteral Nutrition

Neurocritical patients have a high incidence of feeding intolerance (FI) due to multiple influencing factors, including autonomic dysfunction, impairment of the gastrointestinal mucosal barrier, and medication effects. In this study, the incidence of FI was 40.9%, which is consistent with the 30%–50% incidence reported in relevant guidelines[6], indicating that FI poses a substantial clinical risk in this population. In addition, we found that the peak occurrence of FI was within 1–3 days after the initiation of enteral nutrition, whereas the incidence markedly decreased during days 4–7. This pattern is consistent with the finding of Li[7]. During the early phase of enteral nutrition, patients are often in an acute stress state accompanied by gut microbiota imbalance and gastrointestinal dysfunction, which may be key mechanisms underlying the high early incidence of FI[8]. Therefore, clinical staff should strengthen monitoring and intervention during the early stage of enteral nutrition, identify high-risk patients promptly, and ensure the safe and effective implementation of nutritional support.

4.2 Analysis of Factors Influencing Feeding Intolerance in Neurocritical Patients

4.2.1 History of diabetes

This study found that patients with a history of diabetes were more likely to develop FI, consistent with the findings of Bu[9]. Previous studies[10] have shown that hyperglycemia enhances antral contractility while disrupting the coordination of gastric–duodenal contractions, resulting in delayed gastric emptying. Elevated blood glucose also reduces smooth muscle tone in the stomach, further impairing gastrointestinal motility and increasing the risk of gastric retention. Moreover, diabetes is often accompanied by autonomic neuropathy, which may lead to gastroparesis, diarrhea, and fecal incontinence, further increasing the likelihood of FI. Therefore, in patients with diabetes, clinicians should closely monitor blood glucose levels during enteral nutrition and implement timely interventions to maintain glycemic stability, thereby reducing the risk of FI and enhancing feeding safety.

4.2.2 History of stroke

Our findings indicate that patients with a history of stroke are more susceptible to FI, which is consistent with the results reported by Hu [11]. Several mechanisms may underlie this association. Post-stroke dysphagia, impaired consciousness, prolonged immobilization, and reduced gastrointestinal motility can collectively lead to delayed gastric emptying and diminished feeding tolerance. In addition, stroke-related disruptions in brain–gut axis regulation, along with the frequent use of sedatives and opioids, may further exacerbate gastrointestinal dysfunction and increase the likelihood of FI. Therefore, comprehensive gastrointestinal assessment and early, individualized feeding strategies are particularly crucial for this population. Nonetheless, the precise mechanisms linking stroke to FI warrant further investigation.

4.2.3 Glasgow Coma Scale (GCS) score

This study identified lower GCS scores as a risk factor for FI. The GCS is a key tool for evaluating a patient’s level of consciousness and the severity of neurological impairment, with lower scores indicating more critical illness. Previous studies[12] have shown that when the GCS score is <8, patients are more likely to experience intestinal barrier dysfunction and bacterial translocation, which can lead to impaired gastrointestinal motility and substantially increase the risk of FI. Therefore, during the administration of enteral nutrition, clinicians should closely monitor patients with low GCS scores, promptly adjust feeding methods and infusion rates, and implement preventive strategies to reduce the incidence of FI and improve feeding tolerance.

4.2.4 NRS-2002 score

In this study, a higher NRS-2002 score was identified as a predictive factor for FI. The NRS-2002 incorporates age, disease severity, and nutritional status, and is widely used for screening malnutrition risk[13]. Neurocritical patients are often in a hypermetabolic state during the early stage of illness and commonly present with impaired consciousness and dysphagia upon admission, predisposing them to malnutrition[14]. Inadequate nutritional intake weakens various organ systems, including gastrointestinal digestive and absorptive functions, thereby increasing the risk of FI. Hence, strengthening nutritional assessment and implementing targeted interventions for high-risk patients with elevated NRS-2002 scores are essential measures to reduce the incidence of FI.

5 CONCLUSION

In summary, a history of diabetes or stroke, lower GCS scores, and higher NRS-2002 scores were identified as major factors associated with FI in neurocritical patients. These findings may assist clinicians in the early identification of high-risk individuals and in formulating individualized nutritional management strategies. As a single-center retrospective study based on admission data with only internal validation, the present study has certain limitations. The generalizability of the findings requires further confirmation through multicenter studies with larger sample sizes.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Robinson KA, Davis WE, Dinglas VD, et al. A systematic review finds limited data on measurement properties of instruments measuring outcomes in adult intensive care unit survivors. Journal of Clinical Epidemiology. 2017, 82: 37-46.

[2] Reintam Blaser A, Starkopf J, et al. Early enteral nutrition in critically ill patients: ESICM clinical practice guidelines. Intensive Care Medicine. 2017, 43(3): 380-398.

[3] Liu F, Gao L, Wang XY, et al. Expert consensus on enteral feeding care for neurocritical patients. Chinese Journal of Nursing. 2022, 57(3): 261–264.

[4] O'Mara KL, Islam S, Taylor JA, et al. Gabapentin improves oral feeding in neurologically intact infants with abdominal disorders. Journal of Pediatric Pharmacology and Therapeutics. 2018, 23(1): 59–63.

[5] Reintam Blaser A, Malbrain ML, Starkopf J, et al. Gastrointestinal function in intensive care patients: terminology, definitions and management. Recommendations of the ESICM Working Group on Abdominal Problems. Intensive Care Medicine. 2012, 38(3): 384-394.

[6] McClave SA, Martindale RG, Vanek VW, et al. Guidelines for the Provision and Assessment of Nutrition Support Therapy in the Adult Critically Ill Patient: Society of Critical Care Medicine (SCCM) and American Society for

Parenteral and Enteral Nutrition (A.S.P.E.N.). JPEN: Journal of Parenteral and Enteral Nutrition. 2009, 33(3): 277-316.

[7] Li H, Yang Z, Tian F. Risk factors associated with intolerance to enteral nutrition in moderately severe acute pancreatitis: A retrospective study of 568 patients. Saudi Journal of Gastroenterology. 2019, 25(6): 362-368.

[8] Du B, An Y, Kang Y, et al. Characteristics of critically ill patients in ICUs in mainland China. Critical Care Medicine. 2013, 41(1): 84-92.

[9] Bu LJ, Cheng FE, Zhang AQ, et al. Development and validation of a risk prediction model for enteral feeding intolerance in critically ill patients. Chinese Journal of Nursing. 2024, 59(15): 1877–1883.

[10] Halland M, Bharucha AE. Relationship Between Control of Glycemia and Gastric Emptying Disturbances in Diabetes Mellitus. Clinical Gastroenterology and Hepatology. 2016, 14(7): 929-936.

[11] Hu K, Deng XL, Han L, et al. Development and validation of a predictive model for feeding intolerance in intensive care unit patients with sepsis. Saudi Journal of Gastroenterology. 2022, 28(1): 32-38.

[12] Yuan R, Liu L, Mi J, et al. Development and validation of a risk prediction model for feeding intolerance in neurocritical patients with enteral nutrition. Frontiers in Nutrition. 2024, 11: 1481279. DOI: 10.3389/fnut.2024.1481279

[13] Cortés-Aguilar R, Malih N, Abbate M, et al. Validity of nutrition screening tools for risk of malnutrition among hospitalized adult patients: a systematic review and meta-analysis. Clinical Nutrition. 2024, 43(5): 1094–1116.

[14] Elke G, van Zanten AR, Lemieux M, et al. Enteral versus parenteral nutrition in critically ill patients: an updated systematic review and meta-analysis of randomized controlled trials. Critical Care. 2016, 20(1): 117.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3058

MACROPHAGES IN ACUTE MYOCARDIAL INFARCTION: HETEROGENEITY AND TARGETED THERAPIES

HaoDi Fu, YiHuan Chen*
Department of Cardiovascular Surgery, The First Affiliated Hospital of Soochow University, Suzhou 215006, Jiangsu, China.
**Corresponding Author: YiHuan Chen*

Abstract: Cardiac repair after acute myocardial infarction depends on the precise regulation of the immune microenvironment. Macrophages are now understood beyond the M1/M2 dichotomy, with dual origins and a continuous functional spectrum. This review summarizes the spatiotemporal roles of cardiac resident and monocyte-derived macrophages, as well as their coordination in the resolution of inflammation, tissue repair, and scar maturation. It also outlines new therapeutic strategies targeting the recruitment of macrophages, their phenotypic transition, and metabolic reprogramming to improve post-infarction cardiac remodeling.

Keywords: Myocardial infarction; Macrophages; Ventricular remodeling; Immunotherapy

1 INTRODUCTIONS

Acute myocardial infarction (AMI) is one of the leading causes of death and disability worldwide[1]. Its pathological process begins with acute coronary artery occlusion, which leads to sustained myocardial ischemia and cardiomyocyte necrosis. Large numbers of necrotic cells subsequently release damage-associated molecular patterns, rapidly activating the innate immune response and triggering a strong and complex inflammatory cascade. While this response is essential for the clearance of necrotic tissue, it also causes collateral injury to surviving cardiomyocytes[2]. At the same time, the endogenous renewal capacity of adult human cardiomyocytes is extremely limited. Using ^{14}C isotope dating, Bergmann et al. clearly demonstrated that the annual turnover rate of cardiomyocytes is less than 1%, rendering effective regeneration of injured myocardium unlikely[3]. This finding provides a strong scientific rationale for the development of therapeutic strategies that target inflammatory regulation.

Currently, the main clinical strategies for AMI focus on early reperfusion of ischemic myocardium and suppression of excessive neuroendocrine activation[4]. However, despite standardized in-hospital management, nearly one quarter of AMI patients without a prior history of heart failure develop new-onset heart failure within three months after discharge. Moreover, mortality is significantly higher in patients with new-onset heart failure than in those who do not develop heart failure[5]. A key reason for this unfavorable outcome is that the regulatory mechanisms governing post-infarction cardiac inflammation and repair remain incompletely understood, and existing interventions fail to precisely target the critical balance between inflammation and repair. Therefore, elucidating the key cellular and molecular mechanisms that control cardiac repair after AMI has become a major priority for improving patient prognosis.

Macrophages are the most abundant innate immune cells in the heart[6] and exert dual roles under both physiological and pathological conditions. Under homeostatic conditions, they participate in the clearance of metabolic waste from cardiomyocytes, maintenance of endothelial function, and immune surveillance. Under pathological conditions, macrophages regulate the initiation and resolution of inflammation through dynamic phenotypic transitions, remove necrotic tissue and cellular debris, and promote tissue repair, angiogenesis, and collagen deposition. Dysregulated macrophage responses and persistent inflammation have been clearly identified as key pathological contributors to heart failure following myocardial infarction[7]. Traditionally, macrophages have been classified into pro-inflammatory (M1) and anti-inflammatory or reparative (M2) phenotypes[8,9]. However, with the application of single-cell RNA sequencing (scRNA-seq) and cell lineage tracing, it has become evident that cardiac macrophages after AMI comprise multiple functional subsets that extend beyond the M1/M2 paradigm[10,11]. Importantly, macrophages of different origins and subsets show distinct temporal dynamics and spatial localization in the infarcted heart[11]. Together, these macrophage populations, differing in origin, phenotype, and spatiotemporal distribution, collectively shape the outcome of cardiac repair and remodeling after AMI[12]. Based on these insights, this review systematically summarizes the roles of macrophages in AMI from three perspectives: cellular origin, phenotypic heterogeneity and functional diversity, and emerging macrophage-targeted therapeutic strategies.

2 ORIGIN OF CARDIAC MACROPHAGES

Multiple studies using fate-mapping approaches, bone marrow chimera models, and single-cell RNA sequencing have demonstrated that cardiac macrophages after AMI exhibit a dual origin, comprising both embryonic and adult-derived populations. Embryonically derived resident macrophages maintain cardiac immune homeostasis through self-renewal and play a dominant role during the early phase after myocardial infarction. In contrast, monocytes derived from the adult bone marrow are extensively recruited from the peripheral circulation during the inflammatory and reparative phases, infiltrating the infarcted region and its border zone, where they undergo phenotypic differentiation and further participate in inflammatory responses and tissue remodeling[13,14].

Although these two macrophage populations differ in their origin, renewal capacity, and functional specialization, they act in a coordinated manner to maintain the dynamic balance of the inflammatory immune microenvironment following myocardial infarction.

2.1 Cardiac Resident Macrophages (CRMs)

Cardiac resident macrophages (CRMs) are a key component of the intrinsic immune system of the heart. Their origin, maintenance, and functional roles have been extensively characterized over the past decade. Lineage-tracing and bone marrow chimera studies first demonstrated that, under steady-state conditions, the majority of cardiac macrophages are not continuously replenished by circulating monocytes. Instead, they originate from embryonic yolk sac-derived primitive macrophages and fetal liver-derived monocytes, which seed the heart during embryonic development. After birth, these cells are mainly maintained through local self-renewal, exhibiting a renewal pattern that is largely independent of bone marrow hematopoiesis[13,15].

In human cardiac tissue, a similar dual-lineage organization has been confirmed by multi-omics analyses. $CCR2^-$ macrophages, which are predominantly embryonically derived, and $CCR2^+$ macrophages, which are largely bone marrow-derived, coexist long term and display marked differences in metabolic profiles, phagocytic capacity, and immunoregulatory functions[10]. Notably, under homeostatic conditions, CRMs perform essential roles in the clearance of metabolic waste and the maintenance of extracellular matrix and vascular homeostasis. After myocardial infarction, however, their numbers decline to varying degrees[13]. Importantly, studies have shown that delaying CRM depletion can attenuate adverse post-infarction remodeling, whereas premature loss of CRMs exacerbates cardiac injury[12].

2.2 Monocyte-Derived Macrophages (MDMs)

Although CRMs represent the predominant macrophage population in the heart under steady-state conditions, monocyte-derived macrophages (MDMs) become the major compensatory population when CRMs are depleted under stress conditions such as myocardial infarction or ischemia-reperfusion injury. This compensatory response marks a fundamental shift in the macrophage landscape from homeostatic maintenance toward injury-driven repair. Following myocardial infarction, large numbers of necrotic cardiomyocytes rapidly release damage-associated molecular patterns (DAMPs), including HMGB1 and S100A9. These molecules activate CRMs through receptors such as Toll-like receptors (TLRs) and the receptor for advanced glycation end products (RAGE), thereby initiating a local pro-inflammatory response[16,17].

Activated CRMs and cardiac stromal cells secrete a range of chemokines, including CCL2, CCL7, CXCL1, and CXCL2, which establish a strong inflammatory chemotactic gradient. This gradient drives the massive mobilization of monocytes from the bone marrow and spleen into the circulation[18-20]. In parallel, sympathetic nervous system activation and angiotensin II signaling further promote the release of monocytes from the splenic marginal zone[21,22]. Once in the peripheral blood, $Ly6C^{high}$ monocytes roll and adhere to the vascular endothelium via adhesion molecules such as VCAM-1 and ICAM-1, and subsequently infiltrate the infarcted myocardium and border zones in a CCR2- and CX3CR1-dependent manner[18,23].

Within the local inflammatory microenvironment, these recruited monocytes differentiate into pro-inflammatory macrophages, amplify the early inflammatory response, participate in the clearance of necrotic tissue, and regulate subsequent repair processes[18,24]. Compared with CRMs, the role of MDMs after myocardial infarction is more stage-dependent. During the early phase, they facilitate debris clearance and modulate inflammation, whereas in later stages, they contribute to fibrosis, angiogenesis, and scar formation. However, excessive recruitment or dysfunctional activation of MDMs can lead to uncontrolled inflammation and exaggerated fibrosis, ultimately aggravating ventricular remodeling and impairing functional recovery[25]. As repair progresses, the proportion of MDMs gradually declines due to differentiation and cellular turnover, and the heart eventually re-establishes an immune steady state dominated by CRMs[26]. Therefore, clarifying the functional division of labor and dynamic balance between CRMs and MDMs is essential for achieving precise immune modulation after AMI.

3 PHENOTYPIC HETEROGENEITY AND FUNCTIONAL DIVERSITY OF CARDIAC MACROPHAGES

Phenotypic plasticity is a key mechanism that enables macrophages to adapt to the changing demands of different pathological stages. Although the traditional M1/M2 dichotomous model provides a basic conceptual framework, it fails to capture the high degree of heterogeneity and dynamic changes of macrophages after AMI. With the application of single-cell transcriptomic sequencing (scRNA-seq) and cytometry by time-of-flight (CyTOF), multiple macrophage subsets with distinct functional properties have been identified in the infarcted heart. The emergence and interconversion of these subsets are tightly regulated by diverse signaling pathways and metabolic programs.

3.1 Polarization Characteristics and Regulatory Mechanisms in the Classical M1/M2 Paradigm

The M1/M2 classification of macrophages was first proposed by Mills and colleagues, by analogy to Th1/Th2 immune responses, to describe two opposing functional states in inflammatory regulation[27]. This dichotomous framework has provided an important basis for understanding macrophage roles in tissue injury and repair.

Classical M1 macrophage polarization is driven by the coordinated activation of the IFN-γ/JAK-STAT1 and TLR4/NF-κB signaling pathways. IFN-γ induces STAT1 phosphorylation and initiates the transcription of pro-inflammatory genes such as inducible nitric oxide synthase (iNOS), tumor necrosis factor-α (TNF-α), and interleukin-1β (IL-1β)[28,29]. In parallel, lipopolysaccharide (LPS) activates NF-κB[30-32] and hypoxia-inducible factor-1α (HIF-1α), thereby enhancing the production of pro-inflammatory cytokines and reactive oxygen species (ROS)[33]. The transcription factor interferon regulatory factor 5 (IRF5) also plays a critical role in promoting the M1 phenotype[34,35]. Metabolically, M1 macrophages display enhanced glycolysis, accompanied by a partial “break” in the tricarboxylic acid cycle and accumulation of succinate. Succinate stabilizes HIF-1α and further promotes IL-1β expression. Typical markers of the M1 phenotype include iNOS, CD86, and high expression of major histocompatibility complex class II (MHC II$^{\text{high}}$)[36,37]. Functionally, M1 macrophages rapidly initiate inflammatory responses and clear pathogens and cellular debris; however, the excessive release of pro-inflammatory mediators and ROS may cause secondary injury to surrounding viable cardiomyocytes[18,38]. Therefore, during the early phase of AMI, M1 polarization must be tightly regulated in a spatiotemporal manner, as sustained or excessive M1 activation can exacerbate tissue damage and adverse remodeling.

In contrast, M2 macrophage polarization is primarily mediated by the IL-4/IL-13-JAK-STAT6 signaling pathway. Upon activation, STAT6 cooperates with transcription factors such as IRF4, peroxisome proliferator-activated receptor-γ (PPARγ), and Krüppel-like factor 4 (KLF4) to induce the expression of repair-associated genes, including arginase-1 (Arg-1), CD206, and CD163[39-41]. Interleukin-10 (IL-10) further promotes anti-inflammatory programs through STAT3 activation and upregulation of negative feedback regulators such as suppressor of cytokine signaling 3 (SOCS3), which inhibit NF-κB-mediated pro-inflammatory signaling[42,43]. Metabolically, M2 macrophages preferentially rely on oxidative phosphorylation and fatty acid oxidation. Arg-1 converts arginine into ornithine, which can be further metabolized into polyamines that support cell proliferation or proline that contributes to collagen synthesis. Through these pathways, M2 macrophages facilitate fibroblast proliferation, extracellular matrix deposition, and scar formation[44,45]. In addition, M2 macrophages secrete pro-angiogenic factors such as vascular endothelial growth factor (VEGF) and platelet-derived growth factor (PDGF), thereby promoting angiogenesis and tissue repair[46-48].

Macrophages play a central role in the inflammatory and reparative processes following acute myocardial infarction. Traditionally, their functional diversity has been described using the polarization paradigm of classical activation (M1) and alternative activation (M2). With deeper investigation, both phenotypes have been further subdivided into functionally distinct subtypes, the detailed characteristics of which are summarized in the table below. In recent years, under conditions of intense stress such as ischemia-reperfusion injury, a mixed M3 phenotype displaying features of both M1 and M2 macrophages has also been identified. Together, these subsets constitute a complex and dynamic functional spectrum, and their precise temporal regulation is a key determinant of the reparative outcome after myocardial injury.

Table1 Classification and Characteristics of Classical Macrophages[49-54]

Phenotype	Inducers	Markers	Main Functions
M1(Classically Activated)	IFN-γ, LPS	iNOS, CD86, MHC II	Pro-inflammatory, phagocytosis, oxidative stress
M1a	IFN-γ	STAT1, IRF5	Immune response
M1b	LPS, DAMPs	NF-κB, HIF-1α	Pro-inflammatory, pathogen and necrotic tissue clearance
M1c	TNF	TNF-α, ROS	Pro-inflammatory, oxidative stress
M2(Alternatively Activated)	IL-4,IL-13,IL-10	Arg-1, CD206, CD163	Anti-inflammatory, tissue repair, fibrosis, angiogenesis
M2a	IL-4, IL-13	Arg-1, CD206	Matrix remodeling, pro-fibrotic
M2b	Immune complexes	IL-10, TNF	Immune regulation
M2c	IL-10, TGF-β	MerTK	Anti-inflammatory, matrix degradation
M2d	TLR antagonists, IL-6	VEGF	Pro-angiogenic
M3(Mixed Phenotype)	Strong stress (e.g., I/R)	iNOS, Arg-1	Dual effects, highly metabolically active

It should be noted that the traditional M1/M2 classification is merely a simplified model to describe macrophage functional polarization. Multi-omics studies have revealed the existence of numerous novel macrophage subsets after myocardial infarction, often exhibiting overlapping markers and functions. These findings indicate that macrophages form a highly complex and dynamic network in terms of spatiotemporal distribution, metabolic characteristics, and epigenetic regulation.

3.2 Novel Macrophage Phenotypes and Functional Subsets

After clarifying the differences between CRMs and MDMs, it is particularly important to further understand macrophage heterogeneity based on cellular origin, as macrophages of different origins exhibit markedly distinct responses in inflammation, repair, and metabolic states after myocardial infarction. This perspective has prompted researchers to identify novel macrophage subsets with specific lineage features and functional orientations, beyond the classical M1/M2 functional framework.

Single-cell transcriptomic analyses have revealed substantial heterogeneity among CRMs. Although CRMs share the common feature of being CCR2$^-$, they can be further subdivided based on embryonic lineage markers and functional characteristics. The TIMD4$^+$/LYVE1$^+$/FOLR2$^+$ subset, derived from yolk sac and fetal liver, represents the classical embryonic phenotype, exhibiting high efficiency in homeostatic maintenance and phagocytic clearance; these cells are often referred to as "barrier macrophages"[12,55,56]. A smaller MHC II^{high} subset possesses strong antigen-presenting capabilities, contributing to immune surveillance and certain pro-inflammatory functions[11].

scRNA-seq has also identified a CCR2^{low}/ISG$^+$ interferon-responsive CRM population, characterized by high expression of antiviral genes such as ISG15, MX1, and IFIT1. These macrophages are primarily responsible for cardiac antiviral surveillance and systemic inflammatory stress responses, and they are rapidly activated in conditions such as myocardial infarction and viral myocarditis[57]. Additionally, a small population of conduction system-specific CRMs has been identified, mainly localized to the atrioventricular node, expressing molecules such as CX3CR1 and performing a unique role in maintaining electrical conduction stability. Ablation of this population leads to atrioventricular conduction block, highlighting their critical importance in cardiac electrophysiological homeostasis[58].

In the context of myocardial injury, these CRM subsets exhibit pronounced spatiotemporal dynamics. Within hours after AMI, CRMs recognize DAMPs, ATP, and mitochondrial DNA released by necrotic cells and are rapidly activated to participate in the early inflammatory response. However, as tissue damage progresses and inflammation spreads, the majority of CRMs are depleted, and the remaining cells shift toward reparative programs. During the repair phase, a subset of CRMs expresses high levels of MerTK, markedly enhancing their phagocytic capacity for apoptotic cells[12].

At the same time, CCR2$^-$ CRMs secrete insulin-like growth factor 1 (IGF-1) to promote tissue repair[59,60] and vascular endothelial growth factor (VEGF) to facilitate angiogenesis and reperfusion in the ischemic region[61]. Together, these factors drive the transition from inflammation to repair, limiting adverse ventricular remodeling. Importantly, studies have shown that preserving or enhancing the self-renewal capacity of CRMs significantly improves post-infarction cardiac function, whereas CRM depletion or dysfunction leads to delayed clearance of necrotic debris and apoptotic cells, sustained inflammation, and exacerbated remodeling[12].

After myocardial infarction, MDMs play a central regulatory role in cardiac injury repair. Single-cell sequencing studies consistently show that post-infarction MDMs undergo a multi-stage differentiation process—"pro-inflammatory response, anti-inflammatory transition, fibrotic repair"—forming multiple functional subsets within defined temporal windows.

In the early phase of infarction, circulating CCR2$^+$Ly6C^{high} monocytes are robustly recruited to the infarct core by DAMPs and CCL2[62], where they differentiate into pro-inflammatory CCR2$^+$ MDMs[63]. Driven strongly by IRF5 and NLRP3 inflammasome signaling, these cells express high levels of IL-1β, TNF-α, S100A8/A9, and other pro-inflammatory genes[64-66], localizing to the infarct core to clear necrotic tissue and amplify inflammation. Concurrently, a subset of macrophages at the border zone differentiates into antigen-presenting MHC II$^+$ MDMs under the regulation of the IFN-γ-IRF1-CIITA axis, expressing H2-Aa, CD74, and other factors to coordinate local T cell immunity, effectively compensating for the depletion of MHC II$^+$ CRMs[13,63].

Subsequently, pro-inflammatory MDMs begin to transition toward an anti-inflammatory phenotype, co-expressing markers such as SPP1 and ARG1. This transition is mediated through IL-10/STAT3, TGF-β/SMAD3, and transcription factors including KLF4 and PPARγ, which collectively suppress excessive inflammation while initiating early reparative functions[39,40,67].

Finally, as inflammatory cytokines decline and necrotic debris is cleared, reparative TREM2$^+$ MDMs become the dominant subset[11,56,68]. These cells accumulate in maturing fibrotic regions and, under the regulation of metabolic and phagocytosis-related pathways such as PPARγ, LXR, and TFEB, express high levels of TREM2, CD9, and LGALS3. They perform lipid clearance, apoptotic cell removal, and extracellular matrix remodeling, thereby promoting scar stabilization[56,67,69,70].

In summary, macrophage functions during the immune response following AMI exhibit precise spatiotemporal dynamics and high heterogeneity, with phenotypic lineages far exceeding the traditional M1/M2 dichotomy. In the early inflammatory phase, CRMs are rapidly activated by DAMPs and, together with newly recruited CCR2$^+$ MDMs, coordinate the response. The CCR2$^+$ MDMs highly express IL-1β and TNF-α, performing necrotic tissue clearance and inflammation amplification functions similar to classical M1 macrophages, yet single-cell analyses reveal distinct internal states within this population.

As necrotic debris decreases, macrophages enter a transitional phase, exemplified by SPP1$^+$ARG1$^+$ MDMs, in which their function shifts from pro-inflammatory to reparative. This transition is regulated by pathways such as IL-10/STAT3, initiating early repair programs. During the repair and remodeling phase, TREM2$^+$ MDMs become the dominant subset, expressing high levels of genes associated with lipid metabolism and phagocytosis. They execute functions akin to M2 macrophages in fibrosis regulation and scar stabilization, but with a more precise molecular definition.

Meanwhile, specific CRM subsets, such as MerTK$^+$ cells, continue to clear apoptotic cells and secrete reparative factors, collectively promoting inflammation resolution, angiogenesis, and organized extracellular matrix remodeling. Therefore, cardiac repair after AMI is orchestrated by multiple macrophage subsets of distinct origin, clearly defined molecular identity, and continuous functional spectrum, precisely coordinated over time. Systematic analysis of this continuous lineage provides a new perspective for targeted therapeutic interventions.

Table2 Spatiotemporal Heterogeneity and Functional Specialization of CRMs and MDMs Post-AMI

Subtype	Spatiotemporal Distribution	Markers	Key Pathways	Core Functions
CRMs				
Barrier type	Throughout the heart at steady state	TIMD4$^+$, LYVE1$^+$, FOLR2$^+$, CCR2$^-$	—	Phagocytosis of senescent cell debris; maintenance of cardiac homeostasis
Antigen-presenting type	Throughout the heart at steady state	MHC II^{high} (H2-Aa, CD74)	—	Antigen presentation and immune surveillance
Interferon-responsive type	Activated under stress	ISG15$^+$,MX1$^+$, IFIT1$^+$, CCR2^{low}	—	Antiviral defense; early response to systemic inflammatory signals
Conduction-specific type	Atrioventricular node	CX3CR1$^+$	—	Maintenance of normal AV node electrical conduction
Reparative/activated type	Infarct and border zones	MerTK	Apoptosis-related pathways	Clearance of apoptotic cells
MDMs				
Pro-inflammatory	Days 1-3 post-MI, infarct core	CCR2$^+$,IL-1β, TNF-α, S100A8/A9	CCL2/CCR2, IRF5, NLRP3	Clearance of necrotic tissue; amplification of inflammatory response
Antigen-presenting	Days 2-5 post-MI, infarct border	MHC II ^{high} (H2-Aa, CD74)	IFN-γ/IRF1/CIITA	Antigen presentation; modulation of local immune response
Transitional	Days 3-6 post-MI, infarct border	SPP1$^+$, ARG1$^+$	IL-10/STAT3, TGF-β/SMAD3, KLF4, PPARγ	Suppression of excessive inflammation; promotion of tissue repair
Reparative	Days 5-14 post-MI, fibrotic regions	TREM2$^+$,CD9$^+$, LGALS3$^+$,FABP5$^+$	PPARγ, LXR, TFEB, APOE/TREM2	Phagocytosis of residual debris; extracellular matrix remodeling; scar formation

4 NOVEL THERAPEUTIC STRATEGIES TARGETING MACROPHAGES

The inflammatory response following myocardial infarction is double-edged: early inflammation is essential for clearing necrotic tissue, but excessive or persistent inflammation can lead to adverse remodeling and heart failure. Traditional systemic broad-spectrum anti-inflammatory approaches, such as glucocorticoids, can significantly suppress inflammation but may also delay the clearance of necrotic tissue, interfere with scar formation, and ultimately exacerbate adverse remodeling[71,72]. Therefore, broad anti-inflammatory interventions are not beneficial across all temporal windows, making the timing and specificity of intervention critical.

In recent years, accumulating preclinical evidence has highlighted the central role of macrophages in bridging inflammation and repair[73]. As a result, research focus has shifted from non-specific, systemic suppression of inflammation toward precision immunomodulation centered on monocyte/macrophage lineages. Some of these strategies have already progressed to preclinical or early-phase clinical trials, offering promising new avenues to improve outcomes in patients with AMI.

4.1 Targeting Monocyte Recruitment and Infiltration

The key to regulating monocyte recruitment is targeting chemokine-receptor axes. After myocardial infarction, CCR2$^+$ monocytes from the bone marrow and spleen are extensively mobilized and migrate into the heart, exacerbating early inflammation. Experimental studies have shown that angiotensin-converting enzyme inhibitors (ACEIs) can suppress the expression of chemokines and adhesion molecules induced by Ang II, thereby reducing the release of peripheral monocytes and their recruitment to the infarcted region[22]. In mouse MI models, nanoparticles delivering CCR2-targeted siRNA significantly reduced Ly6C^{high} monocyte infiltration into the infarct zone and improved ejection fraction[23].

However, other studies indicate that in CCR2-deficient mice, collagen deposition and tissue stiffness in the infarct scar are markedly reduced[74], suggesting that insufficient or mistimed monocyte/macrophage recruitment may delay the formation of collagen and elastin in the scar, ultimately impairing long-term functional recovery.

4.2 Targeting Macrophage Phenotypic Polarization

In the early phase of AMI, pro-inflammatory macrophages are essential for clearing necrotic tissue, but their prolonged presence can exacerbate tissue damage. In contrast, timely differentiation into reparative macrophages promotes angiogenesis, collagen deposition, and tissue remodeling through the secretion of factors such as IGF-1, TGF-β, and VEGF. Therefore, modulating the local inflammatory microenvironment to facilitate the temporal transition of macrophages from a pro-inflammatory to a reparative phenotype has become one of the most active areas in immunomodulatory research.

Several pharmacological agents have been shown to directly regulate macrophage polarization. For example, the SGLT2 inhibitor dapagliflozin can activate JAK-STAT3 signaling to promote M2 polarization, potentially improving ventricular remodeling by reducing post-MI fibroblast infiltration[75]. Resveratrol similarly induces M2 polarization via STAT3 phosphorylation, thereby improving cardiac function after MI [76]. Moreover, PPARγ agonists such as rosiglitazone, originally used for diabetes treatment, have been shown to induce M2 polarization through STAT6 phosphorylation[76]. Moreover, PPARγ agonists such as rosiglitazone, originally used for diabetes treatment, have been shown to induce M2 polarization through STAT6 phosphorylation[77,78], and in ischemia-reperfusion models, they reduce pro-inflammatory cytokine expression and decrease myocardial necrosis[79,80].

In the field of cell therapy, stem cells and their derivatives—particularly mesenchymal stem cell-derived exosomes (MSC-Exo)—are recognized as key modulators of macrophage fate and function, which is considered one of the central mechanisms underlying their cardioprotective effects[81,82]. Studies indicate that mesenchymal stem cells (MSCs) do not repair the heart by directly differentiating into cardiomyocytes; rather, they actively remodel the immune microenvironment of the infarcted region through continuous secretion of exosomes, microvesicles, and soluble factors[83,84].

Firstly, MSC-Exo can deliver critical microRNAs such as miR-21, miR-146a, miR-182, and miR-223, targeting upstream regulators of inflammation. This suppresses classical pro-inflammatory signaling pathways, including TLR4/NF-κB and NLRP3, and significantly reduces the secretion of pro-inflammatory cytokines such as TNF-α, IL-1β, and IL-6[85-87]. Secondly, MSC-Exo activate reparative signaling axes, including PI3K/AKT, STAT6, and S1P/SK1/S1PR1, upregulating genes such as Arg1, CD206, IGF-1, and VEGF. This drives macrophages toward a reparative M2 phenotype, promoting inflammation resolution and tissue remodeling[88-90].

Additionally, some studies suggest that MSCs or their exosomes may indirectly exert anti-inflammatory effects by enhancing macrophage phagocytosis of apoptotic cells or infiltrating neutrophils, thereby further supporting cardiac repair after injury[91,92].

4.3 Targeting Macrophage Metabolic Reprogramming

In recent years, post-MI immunomodulatory strategies have advanced from simply regulating macrophage phenotypic polarization to precise interventions targeting their underlying metabolic programs. Studies have shown that macrophage functional polarization is closely coupled with metabolic state: pro-inflammatory macrophages primarily rely on glycolysis for energy[93,94], and the accumulation of metabolic intermediates such as succinate stabilizes HIF-1α, driving the expression of pro-inflammatory cytokines like IL-1β[36]. In contrast, following the phagocytosis of apoptotic cells, exogenous fatty acids enter macrophage mitochondria for β-oxidation, providing the main energy source for reparative and biosynthetic functions[95].

Based on these metabolic insights, several interventions targeting specific pathways have demonstrated effective immunomodulation. For example, inhibition of pyruvate kinase isozyme M2 (PKM2), a key glycolytic enzyme, alleviates post-ischemic inflammation[96]. Activation of the IRG1-itaconate pathway or exogenous administration of its derivative 4-octyl itaconate reduces macrophage glycolysis through NADPH-dependent mechanisms, exerting anti-inflammatory effects[97-99]. Conversely, upregulation of the PPARγ-PGC1β axis drives fatty acid β-oxidation, promoting macrophage transition toward a reparative phenotype and suppressing excessive inflammation[100].

Collectively, targeting macrophage metabolic reprogramming represents a novel frontier in post-MI immune microenvironment modulation and offers promising strategies to improve cardiac repair outcomes.

5 CONCLUSIONS AND PERSPECTIVES

Macrophages serve as central regulators throughout the course of AMI, with research paradigms evolving from static phenotypic classification to dynamic functional continua and spatiotemporally specific interaction networks. However, current intervention strategies still face major challenges, primarily the difficulty of precisely distinguishing and targeting macrophage subsets with opposing functions. Achieving selective inhibition of harmful subpopulations while preserving protective ones is key to successful translational applications.

To address these challenges, future research could focus on three frontier directions. First, leveraging spatial multi-omics and in situ imaging technologies to map the spatial distribution and intercellular interactions of distinct macrophage subsets at the microscale, laying the foundation for precise targeting. Second, developing spatiotemporally intelligent delivery systems that enable selective release of therapeutics within defined time windows and infarct regions, thereby coordinating inflammation resolution and tissue repair. Third, exploring personalized therapeutic strategies that integrate patient-specific multi-omics data to understand individual variations in immune responses, ultimately enabling a paradigm shift from broad-spectrum anti-inflammation to precision immune remodeling.

In-depth exploration of these mechanisms holds promise for bridging the gap between scientific understanding and clinical translation of post-MI heart failure, providing a solid basis for the development of targeted prevention and treatment strategies.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Vaduganathan M, Mensah G A, Turco J V, et al. The Global Burden of Cardiovascular Diseases and Risk: A Compass for Future Health. Journal of the American College of Cardiology, 2022, 80(25): 2361-2371.

[2] Zhang W, Lavine K J, Epelman S, et al. Necrotic myocardial cells release damage-associated molecular patterns that provoke fibroblast activation in vitro and trigger myocardial inflammation and fibrosis in vivo. Journal of the American Heart Association, 2015, 4(6): e001993.

[3] Bergmann O, Bhardwaj R D, Bernard S, et al. Evidence for cardiomyocyte renewal in humans. Science, 2009, 324(5923): 98-102.

[4] Sim H W, Zheng H, Richards A M, et al. Beta-blockers and renin-angiotensin system inhibitors in acute myocardial infarction managed with inhospital coronary revascularization. Scientific Reports, 2020, 10(1): 15184.

[5] Butler J, Hammonds K, Talha K M, et al. Incident heart failure and recurrent coronary events following acute myocardial infarction. European Heart Journal, 2025, 46(16): 1540-1550.

[6] Pinto A R, Ilinykh A, Ivey M J, et al. Revisiting Cardiac Cellular Composition. Circulation Research, 2016, 118(3): 400-409.

[7] Lavine K J, Pinto A R, Epelman S, et al. The Macrophage in Cardiac Homeostasis and Disease: JACC Macrophage in CVD Series (Part 4). Journal of the American College of Cardiology, 2018, 72(18): 2213-2230.

[8] Martinez F O, Helming L, Milde R, et al. Genetic programs expressed in resting and IL-4 alternatively activated mouse and human macrophages: similarities and differences. Blood, 2013, 121(9): e57-e69.

[9] Murray P J, Allen J E, Biswas S K, et al. Macrophage activation and polarization: nomenclature and experimental guidelines. Immunity, 2014, 41(1): 14-20.

[10] Bajpai G, Schneider C, Wong N, et al. The human heart contains distinct macrophage subsets with divergent origins and functions. Nature Medicine, 2018, 24(8): 1234-1245.

[11] Jung S H, Hwang B H, Shin S, et al. Spatiotemporal dynamics of macrophage heterogeneity and a potential function of Trem2(hi) macrophages in infarcted hearts. Nature Communications, 2022, 13(1): 4580.

[12] Dick S A, Macklin J A, Nejat S, et al. Self-renewing resident cardiac macrophages limit adverse remodeling following myocardial infarction. Nature Immunology, 2019, 20(1): 29-39.

[13] Epelman S, Lavine K J, Beaudin A E, et al. Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. Immunity, 2014, 40(1): 91-104.

[14] Koenig A L, Kadyrov F F, Amrute J M, et al. Genetic Mapping of Monocyte Fate Decisions Following Myocardial Infarction. bioRxiv, 2025.

[15] Heidt T, Courties G, Dutta P, et al. Differential contribution of monocytes to heart macrophages in steady-state and after myocardial infarction. Circulation Research, 2014, 115(2): 284-295.

[16] Kitahara T, Takeishi Y, Harada M, et al. High-mobility group box 1 restores cardiac function after myocardial infarction in transgenic mice. Cardiovascular Research, 2008, 80(1): 40-46.

[17] Chalise U, Becirovic-Agic M, Daseke M J, 2nd, et al. S100A9 is a functional effector of infarct wall thinning after myocardial infarction. American Journal of Physiology-Heart and Circulatory Physiology, 2022, 322(2): H145-H155.

[18] Nahrendorf M, Swirski F K, Aikawa E, et al. The healing myocardium sequentially mobilizes two monocyte subsets with divergent and complementary functions. Journal of Experimental Medicine, 2007, 204(12): 3037-3047.

[19] Anzai A, Choi J L, He S, et al. The infarcted myocardium solicits GM-CSF for the detrimental oversupply of inflammatory leukocytes. Journal of Experimental Medicine, 2017, 214(11): 3293-3310.

[20] Heusch G, Kleinbongard P. The spleen in ischaemic heart disease. Nature Reviews Cardiology, 2025, 22(7): 497-509.

[21] Swirski F K, Nahrendorf M, Etzrodt M, et al. Identification of splenic reservoir monocytes and their deployment to inflammatory sites. Science, 2009, 325(5940): 612-616.

[22] Leuschner F, Panizzi P, Chico-Calero I, et al. Angiotensin-converting enzyme inhibition prevents the release of monocytes from their splenic reservoir in mice with myocardial infarction. Circulation Research, 2010, 107(11): 1364-1373.

[23] Majmudar M D, Keliher E J, Heidt T, et al. Monocyte-directed RNAi targeting CCR2 improves infarct healing in atherosclerosis-prone mice. Circulation, 2013, 127(20): 2038-2046.

[24] Hilgendorf I, Gerhardt L M, Tan T C, et al. Ly-6C^high monocytes depend on Nr4a1 to balance both inflammatory and reparative phases in the infarcted myocardium. Circulation Research, 2014, 114(10): 1611-1622.

[25] Panizzi P, Swirski F K, Figueiredo J L, et al. Impaired infarct healing in atherosclerotic mice with Ly-6C(hi) monocytosis. Journal of the American College of Cardiology, 2010, 55(15): 1629-1638.

[26] Wu Y, Wei P, Li B, et al. Temporal dynamics of the multi-omic response reveals the modulation of macrophage subsets post-myocardial infarction. Journal of Translational Medicine, 2025, 23(1): 777.

[27] Mills C D, Kincaid K, Alt J M, et al. M-1/M-2 macrophages and the Th1/Th2 paradigm. Journal of Immunology, 2000, 164(12): 6166-6173.

[28] Varinou L, Ramsauer K, Karaghiosoff M, et al. Phosphorylation of the Stat1 transactivation domain is required for full-fledged IFN-gamma-dependent innate immunity. Immunity, 2003, 19(6): 793-802.

[29] Yang S H, Li L, Xie Y Q, et al. IFN-γ-STAT1-iNOS Induces Myeloid Progenitors to Acquire Immunosuppressive Activity. Frontiers in Immunology, 2017, 8: 1192.

[30] Sakai J, Cammarota E, Wright J A, et al. Lipopolysaccharide-induced NF-κB nuclear translocation is primarily dependent on MyD88, but TNFα expression requires TRIF and MyD88. Scientific Reports, 2017, 7(1): 1428.

[31] Lu Y C, Yeh W C, Ohashi P S. LPS/TLR4 signal transduction pathway. Cytokine, 2008, 42(2): 145-151.

[32] Guijarro-Muñoz I, Compte M, Álvarez-Cienfuegos A, et al. Lipopolysaccharide activates Toll-like receptor 4 (TLR4)-mediated NF-κB signaling pathway and proinflammatory response in human pericytes. Journal of Biological Chemistry, 2014, 289(4): 2457-2468.

[33] Nishi K, Oda T, Takabuchi S, et al. LPS induces hypoxia-inducible factor 1 activation in macrophage-differentiated cells in a reactive oxygen species-dependent manner. Antioxidants & Redox Signaling, 2008, 10(5): 983-995.

[34] Krausgruber T, Blazek K, Smallie T, et al. IRF5 promotes inflammatory macrophage polarization and TH1-TH17 responses. Nature Immunology, 2011, 12(3): 231-238.

[35] Weiss M, Blazek K, Byrne A J, et al. IRF5 is a specific marker of inflammatory macrophages in vivo. Mediators of Inflammation, 2013, 2013: 245804.

[36] Tannahill G M, Curtis A M, Adamik J, et al. Succinate is an inflammatory signal that induces IL-1β through HIF-1α. Nature, 2013, 496(7444): 238-242.

[37] Jha A K, Huang S C, Sergushichev A, et al. Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. Immunity, 2015, 42(3): 419-430.

[38] Mouton A J, Deleon-Pennell K Y, Rivera Gonzalez O J, et al. Mapping macrophage polarization over the myocardial infarction time continuum. Basic Research in Cardiology, 2018, 113(4): 26.

[39] Liao X, Sharma N, Kapadia F, et al. Krüppel-like factor 4 regulates macrophage polarization. Journal of Clinical Investigation, 2011, 121(7): 2736-2749.

[40] Odegaard J I, Ricardo-Gonzalez R R, Goforth M H, et al. Macrophage-specific PPARgamma controls alternative activation and improves insulin resistance. Nature, 2007, 447(7148): 1116-1120.

[41] Satoh T, Takeuchi O, Vandenbon A, et al. The Jmjd3-Irf4 axis regulates M2 macrophage polarization and host responses against helminth infection. Nature Immunology, 2010, 11(10): 936-944.

[42] Cassatella M A, Gasperini S, Bovolenta C, et al. Interleukin-10 (IL-10) selectively enhances CIS3/SOCS3 mRNA expression in human neutrophils: evidence for an IL-10-induced pathway that is independent of STAT protein activation. Blood, 1999, 94(8): 2880-2889.

[43] Hovsepian E, Penas F, Siffo S, et al. IL-10 inhibits the NF-κB and ERK/MAPK-mediated production of pro-inflammatory mediators by up-regulation of SOCS-3 in Trypanosoma cruzi-infected cardiomyocytes. PLoS One, 2013, 8(11): e79445.

[44] Huang S C, Smith A M, Everts B, et al. Metabolic Reprogramming Mediated by the mTORC2-IRF4 Signaling Axis Is Essential for Macrophage Alternative Activation. Immunity, 2016, 45(4): 817-830.

[45] Rath M, Müller I, Kropf P, et al. Metabolism via Arginase or Nitric Oxide Synthase: Two Competing Arginine Pathways in Macrophages. Frontiers in Immunology, 2014, 5: 532.

[46] Bingle L, Lewis C E, Corke K P, et al. Macrophages promote angiogenesis in human breast tumour spheroids in vivo. British Journal of Cancer, 2006, 94(1): 101-107.

[47] Corliss B A, Azimi M S, Munson J M, et al. Macrophages: An Inflammatory Link Between Angiogenesis and Lymphangiogenesis. Microcirculation, 2016, 23(2): 95-121.

[48] Han X, Hu J, Zhao W, et al. Hexapeptide induces M2 macrophage polarization via the JAK1/STAT6 pathway to promote angiogenesis in bone repair. Experimental Cell Research, 2022, 413(1): 113064.

[49] Fuchs A L, Costello S M, Schiller S M, et al. Primary Human M2 Macrophage Subtypes Are Distinguishable by Aqueous Metabolite Profiles. International Journal of Molecular Sciences, 2024, 25(4).

[50] Liu B, Nguyen P L, Li P, et al. Identification of novel protein biomarkers of macrophage polarization using comparative proteomic analyses of murine primary macrophages. Journal of Immunology, 2025, 214(10): 2730-2742.

[51] Wang J, Lu S, Yang F, et al. The role of macrophage polarization and associated mechanisms in regulating the anti-inflammatory action of acupuncture: a literature review and perspectives. Chinese Medicine, 2021, 16(1): 56.

[52] Sun J X, Xu X H, Jin L. Effects of Metabolism on Macrophage Polarization Under Different Disease Backgrounds. Frontiers in Immunology, 2022, 13: 880286.

[53] Dousdampanis P, Aggeletopoulou I, Mouzaki A. The role of M1/M2 macrophage polarization in the pathogenesis of obesity-related kidney disease and related pathologies. Frontiers in Immunology, 2024, 15: 1534823.

[54] Sezginer O, Unver N. Dissection of pro-tumoral macrophage subtypes and immunosuppressive cells participating in M2 polarization. Inflammation Research, 2024, 73(9): 1411-1423.

[55] Dick S A, Wong A, Hamidzada H, et al. Three tissue resident macrophage subsets coexist across organs with conserved origins and life cycles. Science Immunology, 2022, 7(67): eabf7777.

[56] Rizzo G, Gropper J, Piollet M, et al. Dynamics of monocyte-derived macrophage diversity in experimental myocardial infarction. Cardiovascular Research, 2023, 119(3): 772-785.

[57] Calcagno D M, Ng R P, Jr., Toomu A, et al. The myeloid type I interferon response to myocardial infarction begins in bone marrow and is regulated by Nrf2-activated macrophages. Science Immunology, 2020, 5(51).

[58] Hulsmans M, Clauss S, Xiao L, et al. Macrophages Facilitate Electrical Conduction in the Heart. Cell, 2017, 169(3): 510-522.e20.

[59] Yang S, Penna V, Lavine K J. Functional diversity of cardiac macrophages in health and disease. Nature Reviews Cardiology, 2025, 22(6): 431-442.

[60] Zaman R, Hamidzada H, Kantores C, et al. Selective loss of resident macrophage-derived insulin-like growth factor-1 abolishes adaptive cardiac growth to stress. Immunity, 2021, 54(9): 2057-2071.e6.

[61] Ferraro B, Leoni G, Hinkel R, et al. Pro-Angiogenic Macrophage Phenotype to Promote Myocardial Repair. Journal of the American College of Cardiology, 2019, 73(23): 2990-3002.

[62] Dewald O, Zymek P, Winkelmann K, et al. CCL2/Monocyte Chemoattractant Protein-1 regulates inflammatory responses critical to healing myocardial infarcts. Circulation Research, 2005, 96(8): 881-889.

[63] Bajpai G, Bredemeyer A, Li W, et al. Tissue Resident CCR2- and CCR2+ Cardiac Macrophages Differentially Orchestrate Monocyte Recruitment and Fate Specification Following Myocardial Injury. Circulation Research, 2019, 124(2): 263-278.

[64] Courties G, Heidt T, Sebas M, et al. In vivo silencing of the transcription factor IRF5 reprograms the macrophage phenotype and improves infarct healing. Journal of the American College of Cardiology, 2014, 63(15): 1556-1566.

[65] Toldo S, Marchetti C, Mauro A G, et al. Inhibition of the NLRP3 inflammasome limits the inflammatory injury following myocardial ischemia-reperfusion in the mouse. International Journal of Cardiology, 2016, 209: 215-220.

[66] Sreejit G, Abdel-Latif A, Athmanathan B, et al. Neutrophil-Derived S100A8/A9 Amplify Granulopoiesis After Myocardial Infarction. Circulation, 2020, 141(13): 1080-1094.

[67] Shirakawa K, Endo J, Kataoka M, et al. IL (Interleukin)-10-STAT3-Galectin-3 Axis Is Essential for Osteopontin-Producing Reparative Macrophage Polarization After Myocardial Infarction. Circulation, 2018, 138(18): 2021-2035.

[68] Gong S, Zhai M, Shi J, et al. TREM2 macrophage promotes cardiac repair in myocardial infarction by reprogramming metabolism via SLC25A53. Cell Death & Differentiation, 2024, 31(2): 239-253.

[69] Javaheri A, Bajpai G, Picataggi A, et al. TFEB activation in macrophages attenuates postmyocardial infarction ventricular dysfunction independently of ATG5-mediated autophagy. JCI Insight, 2019, 4(21).

[70] Che X, Xiao Q, Song W, et al. Protective Functions of Liver X Receptor α in Established Vulnerable Plaques: Involvement of Regulating Endoplasmic Reticulum-Mediated Macrophage Apoptosis and Efferocytosis. Journal of the American Heart Association, 2021, 10(10): e018455.

[71] Kloner R A, Fishbein M C, Lew H, et al. Mummification of the infarcted myocardium by high dose corticosteroids. Circulation, 1978, 57(1): 56-63.

[72] Seropian I M, Toldo S, Van Tassell B W, et al. Anti-inflammatory strategies for ventricular remodeling following ST-segment elevation acute myocardial infarction. Journal of the American College of Cardiology, 2014, 63(16): 1593-1603.

[73] Prabhu S D, Frangogiannis N G. The Biological Basis for Cardiac Repair After Myocardial Infarction: From Inflammation to Fibrosis. Circulation Research, 2016, 119(1): 91-112.

[74] Schumacher D, Curaj A, Staudt M, et al. Endogenous Modulation of Extracellular Matrix Collagen during Scar Formation after Myocardial Infarction. International Journal of Molecular Sciences, 2022, 23(23).

[75] Lee T M, Chang N C, Lin S Z. Dapagliflozin, a selective SGLT2 Inhibitor, attenuated cardiac fibrosis by regulating the macrophage polarization via STAT3 signaling in infarcted rat hearts. Free Radical Biology and Medicine, 2017, 104: 298-310.

[76] Liu S, Du Y, Shi K, et al. Resveratrol improves cardiac function by promoting M2-like polarization of macrophages in mice with myocardial infarction. American Journal of Translational Research, 2019, 11(8): 5212-5226.

[77] Lee Y J, Kim B M, Ahn Y H, et al. STAT6 Signaling Mediates PPARγ Activation and Resolution of Acute Sterile Inflammation in Mice. Cells, 2021, 10(3).

[78] Luo W, Xu Q, Wang Q, et al. Effect of modulation of PPAR-γ activity on Kupffer cells M1/M2 polarization in the development of non-alcoholic fatty liver disease. Scientific Reports, 2017, 7: 44612.

[79] Yue T L, Chen J, Bao W, et al. In vivo myocardial protection from ischemia/reperfusion injury by the peroxisome proliferator-activated receptor-gamma agonist rosiglitazone. Circulation, 2001, 104(21): 2588-2594.

[80] Morrison A, Yan X, Tong C, et al. Acute rosiglitazone treatment is cardioprotective against ischemia-reperfusion injury by modulating AMPK, Akt, and JNK signaling in nondiabetic mice. American Journal of Physiology-Heart and Circulatory Physiology, 2011, 301(3): H895-H902.

[81] Comità S, Penna C, Pagliaro P. The dual role of mesenchymal stem cell-derived exosomes modulates the phagocytic activity of macrophages and promotes cardiac recovery after heart attack. ExRNA, 2021, 4.

[82] Wang J, Xia J, Huang R, et al. Mesenchymal stem cell-derived extracellular vesicles alter disease outcomes via endorsement of macrophage polarization. Stem Cell Research & Therapy, 2020, 11(1): 424.

[83] Pan Y, Wu W, Jiang X, et al. Mesenchymal stem cell-derived exosomes in cardiovascular and cerebrovascular diseases: From mechanisms to therapy. Biomedicine & Pharmacotherapy, 2023, 163: 114817.

[84] Rayat Pisheh H, Sani M. Mesenchymal stem cells derived exosomes: a new era in cardiac regeneration. Stem Cell Research & Therapy, 2025, 16(1): 16.

[85] Huo W, Li Y, Zhang Y, et al. Mesenchymal stem cells-derived exosomal microRNA-21-5p downregulates PDCD4 and ameliorates erectile dysfunction in a rat model of diabetes mellitus. The FASEB Journal, 2020, 34(10): 13345-13360.

[86] Yao M, Cui B, Zhang W, et al. Exosomal miR-21 secreted by IL-1β-primed-mesenchymal stem cells induces macrophage M2 polarization and ameliorates sepsis. Life Sciences, 2021, 264: 118658.

[87] Zhao J, Li X, Hu J, et al. Mesenchymal stromal cell-derived exosomes attenuate myocardial ischaemia-reperfusion injury through miR-182-regulated macrophage polarization. Cardiovascular Research, 2019, 115(7): 1205-1216.

[88] Deng S, Zhou X, Ge Z, et al. Exosomes from adipose-derived mesenchymal stem cells ameliorate cardiac damage after myocardial infarction by activating S1P/SK1/S1PR1 signaling and promoting macrophage M2 polarization. International Journal of Biochemistry & Cell Biology, 2019, 114: 105564.

[89] Wang X, Chen S, Lu R, et al. Adipose-derived stem cell-secreted exosomes enhance angiogenesis by promoting macrophage M2 polarization in type 2 diabetic mice with limb ischemia via the JAK/STAT6 pathway. Heliyon, 2022, 8(11): e11495.

[90] Wang J, Wang N, Zheng Z, et al. Exosomal lncRNA HOTAIR induce macrophages to M2 polarization via PI3K/p-AKT/AKT pathway and promote EMT and metastasis in laryngeal squamous cell carcinoma. BMC Cancer, 2022, 22(1): 1208.

[91] Zhang Z, Tian H, Yang C, et al. Mesenchymal Stem Cells Promote the Resolution of Cardiac Inflammation After Ischemia Reperfusion Via Enhancing Efferocytosis of Neutrophils. Journal of the American Heart Association, 2020, 9(5): e014397.

[92] Patil M, Saheera S, Dubey P K, et al. Novel Mechanisms of Exosome-Mediated Phagocytosis of Dead Cells in Injured Heart. Circulation Research, 2021, 129(11): 1006-1020.

[93] Kelly B, O'Neill L A. Metabolic reprogramming in macrophages and dendritic cells in innate immunity. Cell Research, 2015, 25(7): 771-784.

[94] Haschemi A, Kosma P, Gille L, et al. The sedoheptulose kinase CARKL directs macrophage polarization through control of glucose metabolism. Cell Metabolism, 2012, 15(6): 813-826.

[95] Zhang S, Weinberg S, Deberge M, et al. Efferocytosis Fuels Requirements of Fatty Acid Oxidation and the Electron Transport Chain to Polarize Macrophages for Tissue Repair. Cell Metabolism, 2019, 29(2): 443-456.e5.

[96] Palsson-McDermott E M, Curtis A M, Goel G, et al. Pyruvate kinase M2 regulates Hif-1α activity and IL-1β induction and is a critical determinant of the Warburg effect in LPS-activated macrophages. Cell Metabolism, 2015, 21(1): 65-80.

[97] Michelucci A, Cordes T, Ghelfi J, et al. Immune-responsive gene 1 protein links metabolism to immunity by catalyzing itaconic acid production. Proceedings of the National Academy of Sciences of the United States of America, 2013, 110(19): 7820-7825.

[98] Mills E L, Ryan D G, Prag H A, et al. Itaconate is an anti-inflammatory metabolite that activates Nrf2 via alkylation of KEAP1. Nature, 2018, 556(7699): 113-117.

[99] Liao S T, Han C, Xu D Q, et al. 4-Octyl itaconate inhibits aerobic glycolysis by targeting GAPDH to exert anti-inflammatory effects. Nature Communications, 2019, 10(1): 5091.

[100]Vats D, Mukundan L, Odegaard J I, et al. Oxidative metabolism and PGC-1beta attenuate macrophage-mediated inflammation. Cell Metabolism, 2006, 4(1): 13-24.

Journal of Pharmaceutical and Medical Research
Print ISSN: 2663-1954
Online ISSN: 2663-1962
DOI: https://doi.org/10.61784/jpmr3059

PRETREATMENT ULTRASOUND-BASED RADIOMICS NOMOGRAM FOR PREDICTING PATHOLOGICAL COMPLETE RESPONSE TO NEOADJUVANT CHEMOTHERAPY IN HER2-POSITIVE BREAST CANCER

FeiFei Shan[1], JunQi Sun[2*], Wei Zhou[2], ChengYuan Zheng[2]
[1]Department of Ultrasound, Affiliated Yuebei People's Hospital of Shan University, Shaoguan 512026, Guangdong, China.
[2]Department of Radiology, Affiliated Yuebei People's Hospital of Shan University, Shaoguan 512026, Guangdong, China.
[*]Corresponding Author: JunQi Sun

Abstract: Background: Accurately predicting pathological complete response (PCR) to neoadjuvant chemotherapy (NAC) in patients with human epidermal growth factor receptor 2 (HER2)-positive breast cancer is crucial for tailoring individualized treatment strategies. This study seeks to develop and validate a radiomics-based nomogram utilizing pretreatment ultrasound (US) data to predict PCR on an individual basis. Methods: In this retrospective analysis, patients diagnosed with HER2-positive breast cancer who underwent NAC were included and randomly assigned to either a training cohort (70%) or a validation cohort (30%) between January 2020 and December 2024. Radiomics features were extracted from the entire tumor volume as delineated on pretreatment US images. Feature selection was conducted using a combination of stability analysis, the Least Absolute Shrinkage and Selection Operator (LASSO)to establish a comprehensive radiomics signature (Rad-score). A predictive nomogram was then constructed by integrating the Rad-score with independent clinical predictors through multivariate logistic regression. The model's performance was assessed using the area under the receiver operating characteristic curve (AUC), calibration curves, and decision curve analysis (DCA). Results:The radiomics signature, consisting of 10 features derived from first-order statistics and texture matrices, exhibited significant predictive capability, achieving an area under the curve (AUC) of 0.705 in the validation cohort. The final nomogram, integrating the Rad-score, progesterone receptor (PR), demonstrated enhanced performance, with AUCs of 0.749 and 0.744 in the training and validation cohorts, respectively. Calibration curves and decision curve analyses confirmed the model's robust calibration and clinical applicability. Conclusion: The pretreatment ultrasonography-based radiomics nomogram offers a non-invasive and effective tool for the individualized prediction of pathological complete response (PCR) in HER2-positive breast cancer patients undergoing neoadjuvant chemotherapy (NAC), potentially aiding in the customization of therapeutic strategies.

Keywords: Breast cancer; HER2-Positive; Neoadjuvant chemotherapy; Ultrasonography; Radiomics; Machine learning

1 INTRODUCTION

Breast cancer remains a leading cause of cancer-related mortality among women globally [1].The HER2-positive subtype, accounting for 15-20% of breast cancer cases, is distinguished by its aggressive tumor biology and historically poor prognosis. The incorporation of HER2-targeted therapies, such as trastuzumab and pertuzumab, into neoadjuvant chemotherapy (NAC) regimens has markedly enhanced pathological complete response (PCR) rates and survival outcomes[2-3]. Achieving PCR, defined as the absence of invasive cancer in both the breast and lymph nodes (ypT0/is ypN0), serves as a robust surrogate endpoint, strongly associated with excellent long-term event-free and overall survival[4].

Despite these advancements, the response to NAC remains variable, with PCR rates ranging from 40% to 65%, indicating that a significant proportion of patients experience the toxicity, cost, and delay of ineffective systemic therapy[5]. Presently, clinical assessment predominantly relies on postoperative histopathological evaluation, which is inherently retrospective. Conventional imaging techniques, such as ultrasound (US) and magnetic resonance imaging (MRI), primarily evaluate response based on changes in tumor size, which often lag behind molecular alterations and lack sufficient accuracy for early, individualized prediction[3-10].Consequently, there is an urgent demand for reliable, non-invasive biomarkers that can predict treatment outcomes at the pretreatment stage.

Radiomics, a high-throughput computational approach, addresses this need by extracting and analyzing a comprehensive array of quantitative features from standard medical images. These features, which characterize tumor intensity, shape, and texture heterogeneity, can elucidate the underlying tumor phenotype and microenvironment—information that often remains imperceptible to the human eye [11-13]. Although much of the foundational research in breast cancer radiomics has focused on MRI, breast ultrasound (US) offers distinct practical advantages. It is widely accessible, cost-effective, non-ionizing, and serves as a cornerstone for breast cancer diagnosis and staging worldwide. Emerging evidence robustly supports the predictive value of US-based radiomics for assessing the response to neoadjuvant chemotherapy (NAC). Many research demonstrated that machine learning models leveraging US radiomics features alongside clinical data could effectively predict NAC efficacy [14-18]. Moreover, dynamic models that incorporate features from multiple time points exhibit enhanced predictive power by capturing treatment dynamics.

A considerable portion radiomics research emphasizes feature extraction from subdivided regions of interest (ROI), such as intra- and peritumoral areas[17]. Although this method offers valuable insights, it introduces challenges in segmentation standardization and may not be easily generalizable across diverse clinical environments. There is an acknowledged necessity for models that utilize robust features extracted from the entire tumor volume, as this represents a more straightforward and clinically replicable segmentation task [19-21]. Our study proposes that employing an advanced multi-category feature selection strategy on the whole tumor ROI in pretreatment ultrasound can produce a highly predictive radiomics signature. We hypothesize that integrating this signature with essential clinicopathological factors into a nomogram will yield an enhanced, clinically applicable tool for individualized pathologic complete response (PCR) prediction in HER2-positive breast cancer, thereby eliminating the need for complex regional subdivision.

2 MATERIALS AND METHODS

2.1 Study Population and Design

This retrospective study received approval from the Institutional Review Boards of the participating academic medical centers, with the requirement for informed consent being waived.We conducted a screening of consecutive female patients diagnosed with primary, non-metastatic, HER2-positive invasive breast cancer who underwent neoadjuvant chemotherapy (NAC) followed by definitive surgery between January 2020 and December 2024.

The inclusion criteria encompassed: histologically confirmed HER2-positive invasive breast carcinoma, characterized by immunohistochemistry (IHC) scores of 3+ or 2+ with positive fluorescence in situ hybridization (FISH); completion of a standard NAC regimen incorporating trastuzumab, with or without pertuzumab; availability of high-quality pretreatment grayscale ultrasound (US) images in Digital Imaging and JPG format, conducted within four weeks prior to the initiation of NAC; and availability of a comprehensive pathological report from the surgical specimen for the assessment of pathological complete response (PCR).

The exclusion criteria included incomplete NAC cycles or surgeries performed at external institutions, as well as poor-quality US images with significant artifacts that impede accurate segmentation.

Using a computer-generated random number sequence, the cohort was stratified by institution and PCR statusand subsequently divided into a training set (70%) for model development and an internal validation set (30%) for performance testing.

2.2 Treatment Evaluation and Pathological Evaluation

All patients were administered chemotherapy regimens based on anthracyclines and/or taxanes, in conjunction with anti-HER2 targeted therapy, in alignment with NCCN guidelines. The choice of specific regimen was determined by the treating oncologist. Upon completion of neoadjuvant chemotherapy (NAC), patients underwent either mastectomy or breast-conserving surgery, accompanied by sentinel lymph node biopsy or axillary lymph node dissection. The pathological evaluation of the surgical specimen was employed as the definitive standard for assessing treatment response. Pathological complete response (PCR) was rigorously defined as ypT0/Tis, indicating no residual invasive carcinoma in the breast, with the presence of ductal carcinoma in situ permitted, and ypN0, indicating no tumor deposits in any examined axillary lymph nodes. All pathology slides were independently reviewed by two specialized breast pathologists who were blinded to the imaging analyses and clinical data. Any discrepancies were resolved through a consensus review process.

2.3 Ultrasound Image Acquisition and Tumor Segmentation

Pre-treatment breast ultrasound examinations were conducted by experienced sonographers utilizing high-frequency linear array transducers (frequency range: 9-15 MHz) from major manufacturers (GE LOGIQ E9). Imaging parameters, including gain, depth, and focal zone, were optimized for each patient to ensure maximal lesion visibility.The imaging parameters, including gain, depth, and focal zone, were meticulously optimized for each patient to ensure maximal lesion conspicuity. The standard protocol entailed saving representative grayscale images in both longitudinal and transverse planes, capturing the largest tumor cross-section.

For image segmentation, all JPG images were anonymized and imported into ITK-SNAP 4.4. Careful attention was given to include all hypoechoic tumor tissue while excluding adjacent normal breast parenchyma, fat, and calcifications. To evaluate inter-observer reproducibility, a second radiologist (Reader 2) independently conducted the segmentation on 30 randomly selected cases from the entire cohort.

2.4 Radiomics Feature Extraction and Signature Construction

Radiomics feature extraction were performed using the Onekey AI, adhering to the Image Biomarker Standardization Initiative (IBSI) guidelines to ensure reproducibility.

Feature Extraction: From each segmented tumor region of interest (ROI), a comprehensive set of 1562 radiomic features was extracted. These features were systematically categorized into distinct groups: First-order Statistics: These are descriptors of the distribution of voxel intensities within the ROI, including metrics such as mean, median, 10th percentile, 90th percentile, entropy, kurtosis, and skewness.

Shape-based Features (2D): These features provide geometric descriptors of the ROI, encompassing measurements such as area, perimeter, major axis length, minor axis length, eccentricity, and solidity.

Texture Features: These features quantify intra-tumoral heterogeneity and are derived from various matrices, including: Gray-Level Co-occurrence Matrix (GLCM): Metrics such as contrast, correlation, energy, and homogeneity. Gray-Level Run Length Matrix (GLRLM): Metrics such as short-run emphasis, long-run emphasis, and run-length non-uniformity. Gray-Level Size Zone Matrix (GLSZM): Metrics such as zone entropy and size-zone non-uniformity. Gray-Level Dependence Matrix (GLDM): Metrics such as dependence entropy and dependence non-uniformity.Gray-Level Size Zone Matrix (GLSZM) metrics, such as zone entropy and size-zone non-uniformity, and Gray-Level Dependence Matrix (GLDM) metrics, including dependence entropy and dependence non-uniformity, were analyzed.

Feature preprocessing and selection: Feature preprocessing and selection involved a meticulous multi-step procedure applied exclusively to the training cohort to mitigate risks of data leakage and overfitting.

Stability filtering:features with an inter-observer intraclass correlation coefficient (ICC) below 0.90, based on 30 double-segmented cases were excluded to ensure robustness.

Variance and correlation filtering: features exhibiting near-zero variance (variance < 0.01) were removed. Subsequently, features with high correlation (absolute Pearson correlation coefficient > 0.95) were identified, and from each cluster of correlated features, the feature with the highest mean absolute correlation to others was retained.

Predictive feature selection:the remaining features underwent standardization through z-score normalization, followed by a two-stage selection process.

The Least Absolute Shrinkage and Selection Operator (LASSO) regression, utilizing 10-fold cross-validation, was employed to reduce coefficients and select a subset of non-redundant features predictive of pathological complete response (PCR). Subsequently, Recursive Feature Elimination (RFE) based on a Support Vector Machine (SVM) classifier was conducted on the features selected by LASSO to determine the most optimal and parsimonious feature set.

Radiomics Signature (Rad-score) Building: the final selected features were assigned weights according to their coefficients obtained from a logistic regression model fitted to the training dataset. The Rad-score for each patient was computed as a linear combination of these selected feature values and their corresponding weights.

2.5 Clinical Model and Combined Nomogram

Clinicopathological data were extracted from electronic medical records, encompassing variables such as age, menopausal status, clinical T and N stage (according to the AJCC 8th edition), histological grade, and biomarker status (estrogen receptor [ER], progesterone receptor [PR], and Ki-67 index). Univariate logistic regression analysis was performed to identify clinical factors significantly associated with PCR within the training cohort.In the multivariate logistic regression analysis, variables with a P-value less than 0.05 were identified as independent predictors and subsequently utilized to construct a clinical-only prediction model. The final predictive nomogram was developed by integrating the Rad-score with the independent clinical predictors into a multivariable logistic regression model within the training cohort.

2.6 Model Evaluation and Statistical Analysis

The efficacy of the radiomics signature, the clinical model, and the combined nomogram was assessed in both the training and independent validation cohorts.

Discrimination: The primary metric was the area under the receiver operating characteristic curve (AUC). Differences in AUC between models were compared using DeLong's test.

Calibration: The agreement between predicted probabilities and observed outcomes was assessed using calibration plots and the Hosmer-Lemeshow goodness-of-fit test, where a non-significant P-value greater than 0.05 indicates good calibration.

Clinical Utility: Decision curve analysis (DCA) was conducted to determine the net benefit of each model across a spectrum of threshold probabilities, thereby quantifying their clinical value in comparison to PCR or NO PCR strategies.

All statistical analyses were performed using Onekey AI software. A two-sided P value < 0.05 was considered statistically significant.

3 RESULTS

3.1 Patient Characteristics

The baseline demographic and clinicopathological characteristics of the158 patients are summarized (Table 1). The overall PCR rate was 56.33% (89 out of 158). There were no significant differences in baseline characteristics between the training and validation sets (all P>0.05).

Table 1 Baseline Characteristics of Patients in the Training and Validation Cohorts

Characteristic	Total Cohort (n=158)	Validation Cohort (n=48)	Train Cohort (n=110)	P-value
Age (years), mean ± SD	52.418±8.507	53.667±9.217	51.873±8.162	0.224
Clinical T Stage, n (%)				0.727
cT1-2	124(78.481)	39(81.250)	85(77.273)	
cT3-4	34(21.519)	9(18.750)	25(22.727）	
Clinical N Stage, n (%)				0.218
cN0	38(24.051)	8(16.667)	30(27.273）	
cN+	120(75.949)	40(83.333)	80(72.727）	
ER Status, n (%)				0.572
Positive	72(45.570)	24(50.000)	48(43.636）	
Negative	86(54.430)	24(50.000)	62(56.364)	
PR Status, n (%)				0.391
Positive	89(56.329)	30(62.500)	59(53.636）	
Negative	69(43.671)	18(37.500)	51(46.364	
Clinical ALL Stage				0.69
I1-2	104(65.823)	30(62.500)	74(67.273）	
3-4	54(34.177)	18(37.500)	36(32.727）	
PCR Status, n (%)				0.604
PCR	89(56.329)	30(62.500)	59(53.636）	
Non-PCR	69(43.671)	18(37.500)	51(46.364)	

3.2 Radiomics Feature Selection and model Performance

Following variance and correlation filtering, 15 features were retained for predictive selection.The LASSO-RFE pipeline selected 10 optimal features for the final signature, originating from first-order statistics and texture matrices (Figure 1), highlighting the importance of both intensity distribution and complex spatial heterogeneity patterns within the tumor. The radiomics signature showed good and stable predictive performance, with max AUC of 0.717 (95% CI: 0.66-0.81) in the training cohort and 0.705 (95% CI: 0.56-0.85) in the validation cohort in MLP model (Table 2).

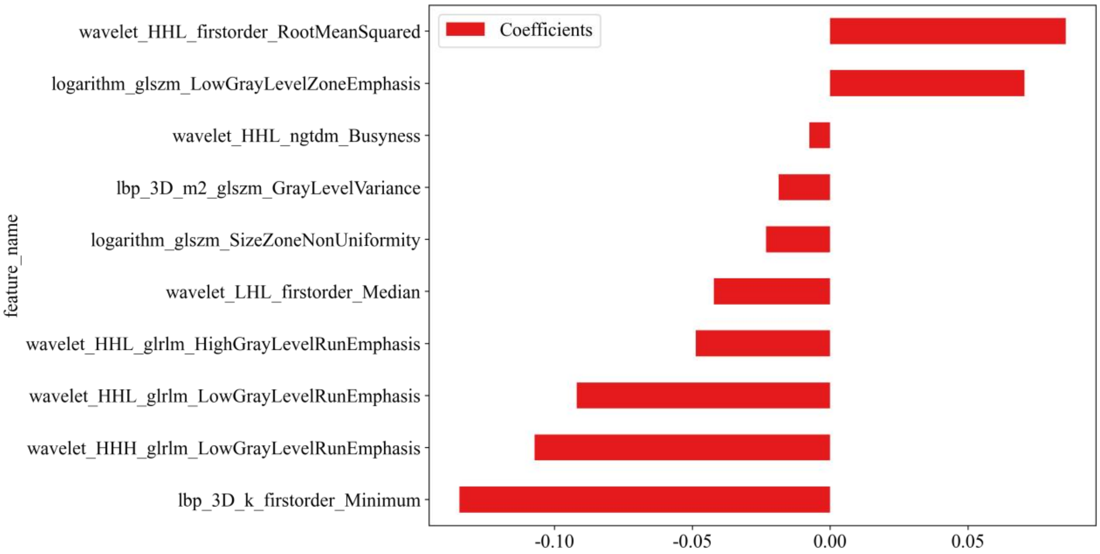


Figure 1 Features Coefficients of Patients in the Training and Validation Cohorts

Table 2 Model predictions of Radiomics in the Training and Validation Cohorts

Model	Cohort	AUC（95% CI）	Sensitivity	Specificity	Accuracy
LR	train	0.800（0.7189-0.8812）	0.717	0.754	0.736
	Validation	0.526（0.3557 - 0.6954）	0.407	0.714	0.542
NaiveBayes	train	0.693（0.5952 - 0.7904）	0.925	0.404	0.655
	Validation	0.651（0.4937 - 0.8079）	0.852	0.429	0.667
SVM	train	0.762（0.6745 - 0.8501）	0.623	0.772	0.700
	Validation	0.543（0.3751 - 0.7113）	0.370	0.857	0.583
KNN	train	0.770 0.6857 - 0.8545	0.717	0.684	0.700
	Validation	0.547 0.3820 - 0.7115	0.593	0.571	0.583
RandomForest	train	0.881 0.8180 - 0.9440	0.925	0.702	0.809
	Validation	0.436（0.2669 - 0.6044）	0.296	0.905	0.562
ExtraTrees	train	0.964（0.9355 - 0.9930）	0.981	0.825	0.900
	Validation	0.529（0.3612 - 0.6970）	0.593	0.571	0.583
XGBoost	train	0.959（0.9252 - 0.9927）	0.943	0.860	0.900
	Validation	0.441（0.2723 - 0.6095）	0.185	0.952	0.521
LightGBM	train	0.661（0.5799 - 0.7429）	0.849	0.474	0.655
	Validation	0.397（0.2633 - 0.5304）	1.000	0.000	0.562
GradientBoosting	train	1.000（1.0000 - 1.0000）	1.000	1.000	1.000
	Validation	0.594（0.4312 - 0.7575）	0.333	0.952	0.604
AdaBoost	train	0.934（0.8923 - 0.9763）	0.755	0.965	0.864
	Validation	0.587（0.4239 - 0.7507）	0.667	0.524	0.604

Model	Cohort	AUC（95% CI）	Sensitivity	Specificity	Accuracy
MLP	train	0.717（0.6220 - 0.8113）	0.660	0.702	0.682
	Validation	0.705（0.5584 - 0.8526）	0.556	0.810	0.667

3.3 Development and Performance of the Clinical Model and Final Nomogram

Multivariate logistic regression analysis identified PR negativity (Odds Ratio [OR]: -0.606, 95% CI: -1.088-[-10.124], P=0.039) as independent clinical predictors of PCR. This combined model achieved the highest discriminatory ability, with an AUC of 0.749(0.659 - 0.839) in the training cohort and 0.744(0.605 - 0.884) in the validation cohort (Table 3, Figure 2). The final nomogram integrated the Rad-score with PR status (Figure 3).The nomogram's performance was significantly superior to that of the clinical-only model (DeLong's test, P<0.001 in validation) and showed a strong trend towards improvement over the radiomics-only model.

Table 3 Predictive Performance of Different Models in the Training and Validation Cohorts

Model	Cohort	AUC (95% CI)	Sensitivity	Specificity	Accuracy
Clinic	Train	0.620(0.5284-0.7109)	0.660	0.579	0.618
Radiomics	Train	0.717(0.6220-0.8113)	0.660	0.702	0.682
Combined	Train	0.749(0.6589-0.8386)	0.849	0.526	0.682
Clinic	Validation	0.675(0.5399 - 0.8093)	0.778	0.571	0.688
Radiomics	Validation	0.705(0.5584 - 0.8526)	0.556	0.810	0.667
Combined	Validation	0.744(0.6046 - 0.8840)	0.593	0.810	0.688

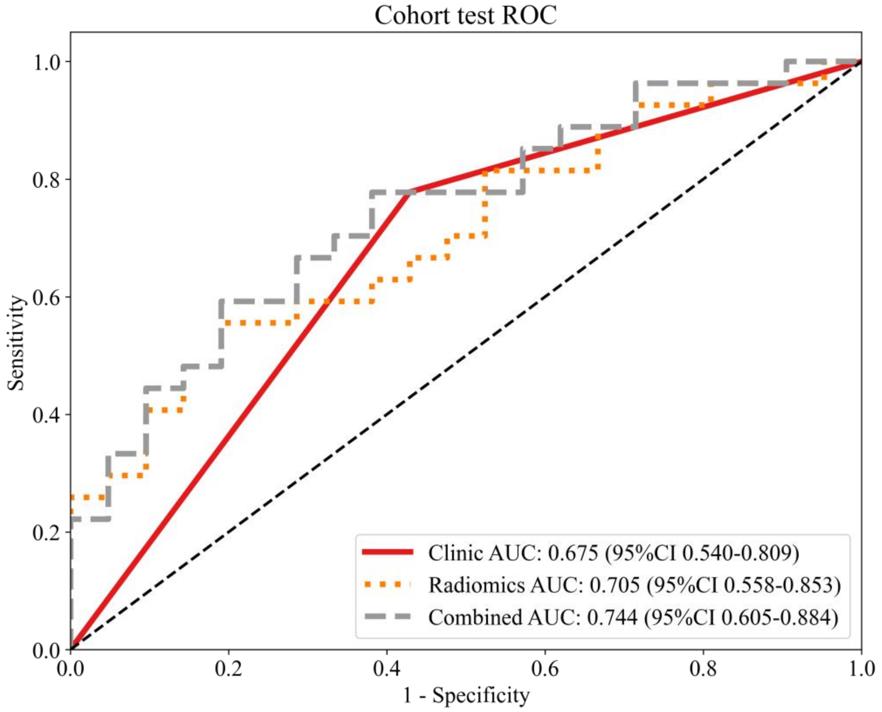


Figure 2 ROC of Patients in the Training and Validation Cohorts

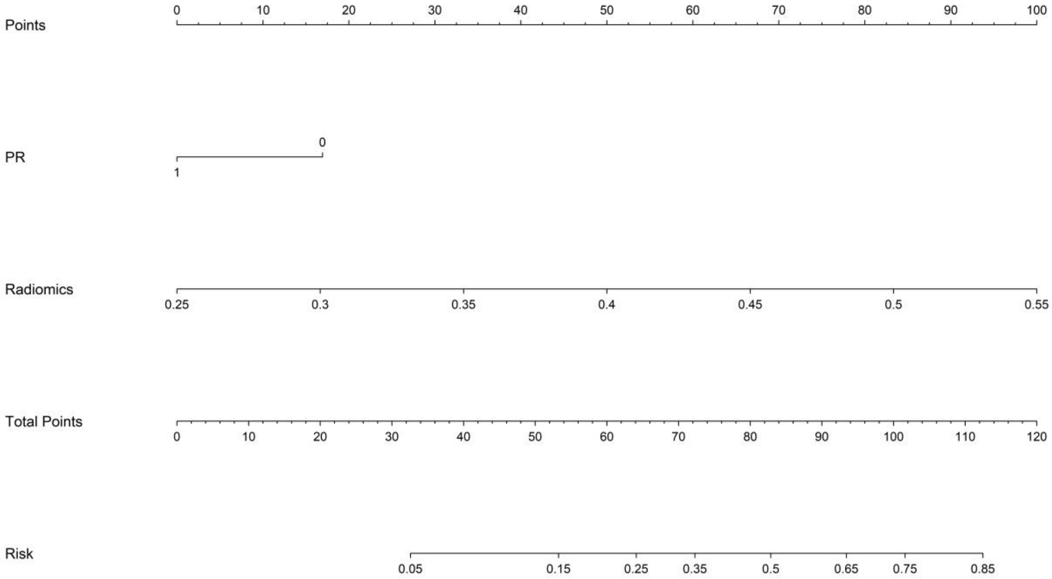


Figure 3 Nomogram of Patients in the Training and Validation Cohorts

Decision Curve Analysis (DCA) indicated that utilizing the nomogram for clinical decision-making offered a higher net benefit compared to the "treat-all" or "treat-none" strategies over a broad spectrum of clinically reasonable threshold probabilities (Figure 4).The calibration curve for the nomogram exhibited a strong concordance with the ideal 45-degree line in both the training and validation cohorts (Figure 5).

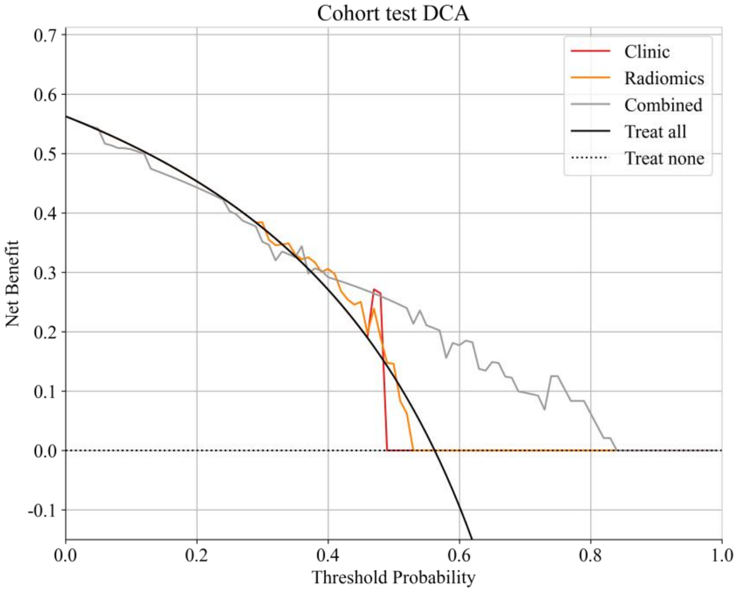


Figure 4 DCA of Patients in the Training and Validation Cohorts

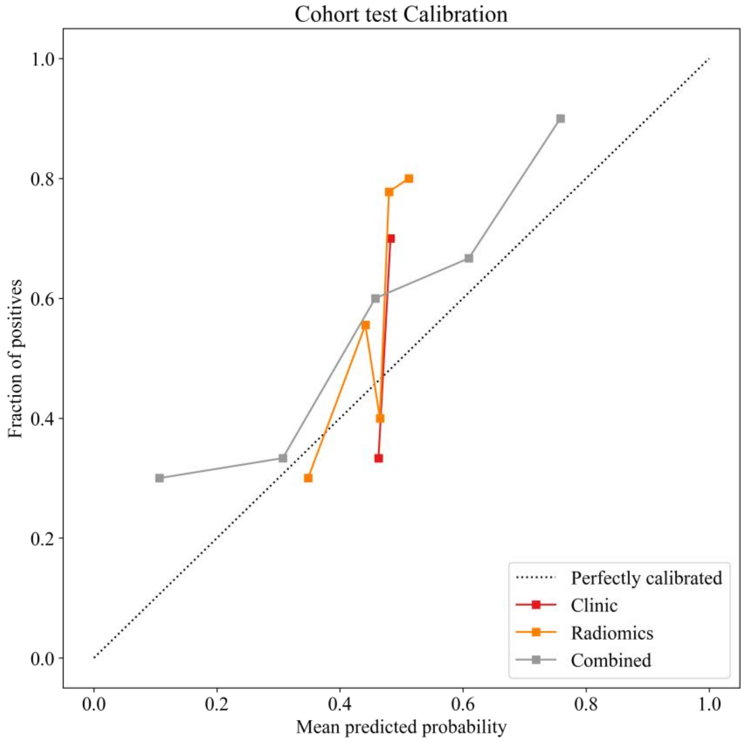


Figure 5 Calibration Curve in the Training and Validation Cohorts

4 DISCUSSION

In this retrospective study, we developed and validated a pretreatment ultrasound-based radiomics nomogram for the individualized prediction of pathological complete response (PCR) to neoadjuvant chemotherapy (NAC) in patients with HER2-positive breast cancer. Our model is based on a sophisticated selection of radiomics features extracted from the entire tumor volume, integrated with key clinical biomarkers. The final nomogram demonstrated predictive accuracy, good calibration, and favorable clinical utility, indicating its potential as a practical tool for personalizing NAC strategies.

The success of our model is rooted in the multi-step, rigorous feature engineering process applied to the whole-tumor region of interest (ROI). The selected 10-feature signature comprises a combination of first-order intensity statistics and higher-order texture features. This is consistent with the biological understanding that tumors with high internal heterogeneity, often indicative of genomic instability and clonal diversity, may exhibit differential responses to therapy [4-13]. Our approach illustrates that meaningful predictive information can be effectively captured from the global tumor phenotype without necessitating complex subdivision of the ROI, potentially enhancing the method's clinical translatability and reproducibility across different institutions and operators[14-17].

The integration of the radiomics signature with progesterone receptor (PR) has resulted in the development of a robust multi-parametric predictive tool. Progesterone receptor (PR) is well-recognized indicators of a more aggressive and proliferative tumor biology, which is typically more responsive to cytotoxic chemotherapy. This aligns with their roles as independent predictors within our model [18-19]. The nomogram effectively combines these established clinical biomarkers with the novel, quantitative imaging phenotype provided by radiomics. This integration likely captures a more comprehensive view of the tumor, encompassing both its molecular characteristics (via PR) and its macroscopic phenotypic expression (via radiomics), thereby leading to enhanced predictive performance compared to the use of either component in isolation[20].

The clinical implications of our findings are significant. An accurate pretreatment prediction model could profoundly influence patient management in the neoadjuvant setting. For patients predicted to have a very high likelihood of achieving a pathologic complete response (PCR), clinicians might consider exploring strategies within clinical trial frameworks, such as evaluating de-escalated chemotherapy regimens or shorter treatment durations， The objective is to minimize both acute and long-term toxicity without compromising oncologic outcomes [2-13]. In contrast, for patients predicted to be non-responders, early treatment intensification may be considered. This could involve the initial incorporation of novel antibody-drug conjugates, such as trastuzumab deruxtecan, dual HER2 blockade with pertuzumab if not previously utilized, or participation in clinical trials exploring new targeted agents or immunotherapy combinations [12-16]. Such a risk-adapted strategy exemplifies the principles of precision oncology.

Our study has several limitations that merit consideration. Firstly, the retrospective design may introduce inherent selection bias. Secondly, although we conducted internal validation, external validation in a larger, prospective, and multinational cohort is crucial to confirm the generalizability and robustness of our nomogram. Thirdly, our model is based on two-dimensional ultrasound images. The utilization of three-dimensional automated breast ultrasound (ABUS) could offer more comprehensive volumetric feature extraction.Fourth, we concentrated on predicting outcomes before treatment. By incorporating serial ultrasound imaging during treatment to analyze delta-radiomics, which involves changes in features over time, we could capture dynamic response patterns and improve prediction accuracy, as suggested by recent research. Future studies should also consider integrating ultrasound radiomics with other data types, such as MRI radiomics or circulating tumor DNA, to create a more comprehensive multi-omics predictive profile.

5 CONCLUSION

We have developed and validated a pretreatment ultrasound-based radiomics nomogram that effectively predicts the likelihood of achieving a complete response following neoadjuvant chemotherapy in HER2-positive breast cancer patients. By using a thorough whole-tumor feature selection strategy and combining the resulting radiomics signature with key clinical biomarkers, this non-invasive tool offers individualized risk assessment with high accuracy. It represents a practical step towards personalized neoadjuvant therapy, potentially guiding clinical decisions by identifying patients who may benefit from treatment de-escalation or early intensification, ultimately aiming to improve therapeutic efficacy and patient quality of life.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

FUNDING

This work was supported by the Science and Technology Program of Shaoguan City (No. 210922204531943, No. 230330098033759), Medical Science and Technology Research Fund Project of Guangdong Province (No. B2023468), and Health Research Project of Shaoguan City (No. Y22103, No. Y2023080).

REFERENCES

[1] Wagle NS, Nogueira L, Devasia TP, et al. Cancer treatment and survivorship statistics. CA: A Cancer Journal for Clinicians, 2025, 75(4): 308-340.

[2] Bardia A, Jhaveri K, Im SA, et al. Datopotamab Deruxtecan Versus Chemotherapy in Previously Treated Inoperable/Metastatic Hormone Receptor-Positive Human Epidermal Growth Factor Receptor 2-Negative Breast Cancer: Primary Results From TROPION-Breast01. Journal of Clinical Oncology, 2025, 43(3): 285-296.

[3] Gillies R J, Kinahan P E, Hricak H, et al. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016, 278(2): 563-577.

[4] Lin Z, Zheng M, Li Z, et al. Development and validation of a delta ultrasomics model for predicting treatment response to neoadjuvant chemotherapy in breast cancer. Translational Cancer Research, 2025, 14(11): 7967-7979.

[5] Moore-Palhares D, Alberico D, Chan AW, et al. Quantitative ultrasound imaging for predicting response and guiding personalized neoadjuvant chemotherapy in breast cancer: randomized phase 2 clinical trial results. NPJ Precision Oncology, 2025, 9(1): 390.

[6] Liu J, Leng X, Yuan Z, et al. Predicting breast cancer response to neoadjuvant chemotherapy with ultrasound-based deep learning radiomics models -- dual-center study. BMC Cancer, 2025, 25(1): 1737.

[7] Wang M, Huang Z, Tian H, et al. Longitudinal Ultrasound Delta Radiomics for Early Stratified Prediction of Tumor Response to Neoadjuvant Chemotherapy in Breast Cancer. Academic Radiology, 2025, 32(12): 7119-7133.

[8] Peng Q, Ji Z, Xu N, et al. Prediction of neoadjuvant chemotherapy efficacy in patients with HER2-low breast cancer based on ultrasound radiomics. Cancer Imaging, 2025, 25(1): 112.

[9] Wei C, Jia Y, Gu Y, et al. Predictive Analysis of Neoadjuvant Chemotherapy Efficacy in Breast Cancer Using Multi-Region Ultrasound Imaging Features Combined With Pathological Parameters. Ultrasound in Medicine and Biology, 2025, 51(12): 2205-2216.

[10] Moore-Palhares D, Sannachi L, Chan AW, et al. Validation of a Quantitative Ultrasound Texture Analysis Model for Early Prediction of Neoadjuvant Chemotherapy Response in Breast Cancer: A Prospective Serial Imaging Study. Cancers (Basel), 2025, 17(15): 2594.

[11] Liu J, Xue X, Yan Y, et al. Prediction of breast cancer HER2 status changes based on ultrasound radiomics attention network. Computer Methods and Programs in Biomedicine, 2025, 271: 108987.

[12] Valizadeh P, Jannatdoust P, Moradi N, et al. Ultrasound-based machine learning models for predicting response to neoadjuvant chemotherapy in breast cancer: A meta-analysis. Clinical Imaging, 2025, 125: 110574.

[13] Chen M, Hong T, Wang Y, et al. Effect of a machine learning prediction model on the false-negative rate of sentinel lymph node biopsy for clinically node-positive breast cancer after neoadjuvant chemotherapy. Breast, 2025, 83: 104543.

[14] Tenghui W, Xinyi L, Ziyi S Y, et al. Combination of ultrasound-based radiomics and deep learning with clinical data to predict response in breast cancer patients treated with neoadjuvant chemotherapy. Frontiers in Oncology, 2025, 15: 1525285.

[15] Zhang H, Lang M, Shen H, et al. Machine learning-based fusion model for predicting HER2 expression in breast cancer by Sonazoid-enhanced ultrasound: a multicenter study. Frontiers of Medicine (Lausanne), 2025, 12: 1585823.

[16] Fan Y, Sun K, Xiao Y, et al. Deep learning predicts HER2 status in invasive breast cancer from multimodal ultrasound and MRI. Biomolecules and Biomedicine, 2025, 25(10): 2243-2251.

[17] Yao J, Zhou W, Jia X, et al. Machine learning prediction of pathological complete response to neoadjuvant chemotherapy with peritumoral breast tumor ultrasound radiomics: compare with intratumoral radiomics and clinicopathologic predictors. Breast Cancer Research and Treatment, 2025, 212(2): 325-336.

[18] Chan AW, Sannachi L, Moore-Palhares D, et al. Validation of Quantitative Ultrasound and Texture Derivative Analyses-Based Model for Upfront Prediction of Neoadjuvant Chemotherapy Response in Breast Cancer. Journal of Imaging, 2025, 11(4): 109.

[19] Feng X, Shi Y, Wu M, et al. Predicting the efficacy of neoadjuvant chemotherapy in breast cancer patients based on ultrasound longitudinal temporal depth network fusion model. Breast Cancer Research, 2025, 27(1): 30.

[20] Zhou P, Qian H, Zhu P, et al. Machine learning for predicting neoadjuvant chemotherapy effectiveness using ultrasound radiomics features and routine clinical data of patients with breast cancer. Frontiers in Oncology, 2025, 14: 1485681.

[21] Xie J, Wei J, Shi H, et al. A deep learning approach for early prediction of breast cancer neoadjuvant chemotherapy response on multistage bimodal ultrasound images. BMC Med Imaging, 2025, 25(1): 26.