**International Journal of Genetics and Genomics**
2023; 11(4): 133-138
http://www.sciencepublishinggroup.com/j/ijgg
doi: 10.11648/j.ijgg.20231104.14
ISSN: 2376-7340 (Print); ISSN: 2376-7359 (Online)

# Construction of a Cuproptosis-Related Gene Clinical Prediction Model for Juvenile Idiopathic Arthritis Using Machine Learning

**Yiwei Hong†, Xu Cai†, Xinpeng Chen, Xinmin Huang, Zhengbo Yan, Peihu Li, Jianwei Xiao***

Rheumatology, Shenzhen Futian Hospital for Rheumatic Diseases, Shenzhen, China

## Email address:

108812365@qq.com (Jianwei Xiao)

*Corresponding author

† Yiwei Hong and Xu Cai are co-first authors.

## To cite this article:

Yiwei Hong, Xu Cai, Xinpeng Chen, Xinmin Huang, Zhengbo Yan et al. (2023). Construction of a Cuproptosis-Related Gene Clinical Prediction Model for Juvenile Idiopathic Arthritis Using Machine Learning. *International Journal of Genetics and Genomics*, *11*(4), 133-138. https://doi.org/10.11648/j.ijgg.20231104.14

**Received**: November 23, 2023; **Accepted**: December 7, 2023; **Published**: December 22, 2023

**Abstract:** *Objective:* Juvenile Idiopathic Arthritis (JIA) is a chronic inflammatory joint disease affecting children and adolescents, where early diagnosis and treatment are crucial for improving prognosis. This study aimed to identify cuproptosis-related genes in JIA and develop a clinical predictive model. *Methods:* The GSE13849 dataset was retrieved from the GEO database to extract cuproptosis-related genes. Key JIA genes were selected using the Boruta and SVM-REF algorithms, followed by the construction of a clinical prediction model. The model's predictive capacity was validated using the concordance index (C-index), Receiver Operating Characteristic (ROC) curves, and calibration curves. Patient net benefit was evaluated through clinical decision curves, with internal validation conducted via Bootstrap. *Results:* The Boruta and SVM-REF algorithms identified four and five core cuproptosis-related genes, respectively, intersecting to yield three core genes (PDHA1, LIAS, DLAT). A clinical prediction model was established using multivariate logistic regression, exhibiting a C-index of 0.75 and an area under the ROC curve of 0.749. Clinical decision curve analysis demonstrated the highest net clinical benefit at a threshold probability range of 0.15 to 0.9, ensuring no harm to other patients. Internal validation reported a C-index of 0.755 and an area under the ROC curve of 0.736. *Conclusion:* The JIA clinical prediction model, based on three cuproptosis-related genes, demonstrates substantial predictive diagnostic capability, contributing to the early diagnosis of JIA patients.

**Keywords:** Juvenile Idiopathic Arthritis, Cuproptosis, Machine Learning, Clinical Prediction Model, Boruta, SVM-REF

## 1. Introduction

Juvenile Idiopathic Arthritis (JIA) is a chronic inflammatory joint disease affecting children and adolescents, characterized by arthritis persisting for more than six weeks. JIA represents the most common form of chronic arthritis in children and significantly impacts their quality of life and long-term health [1]. The exact etiology of JIA remains unclear, but studies indicate that genetics, environmental factors, and abnormal immune responses play key roles in its pathogenesis [2, 3]. Globally, the prevalence of JIA is estimated to be between 16-150 cases per 100,000 children, with an annual incidence rate of approximately 2-20 cases per 100,000 children per year [4]. JIA presents with diverse clinical manifestations and can involve multiple systems, including the eyes, skin, and internal organs. Early diagnosis and treatment are crucial for improving prognosis.

Cuproptosis is a novel cellular death mechanism, distinct from known cell death regulatory mechanisms such as apoptosis, ferroptosis, pyroptosis, and necroptosis. It is induced by copper ions and primarily depends on the accumulation of copper within the cell. The mechanism involves the direct binding of copper ions to lipoylated components in the tricarboxylic acid (TCA) cycle, forming aggregates that lead to cytotoxicity and ultimately cell death [5]. The function of cuproptosis in JIA remains unclear. A deeper understanding of cuproptosis's role in JIA could aid in the development of new biomarkers and therapeutic

approaches.

In this study, we employed bioinformatics and machine learning methods to identify cuproptosis-related genes in JIA and developed a predictive model for its diagnosis. Our findings may contribute to the early diagnosis of JIA.

# 2. Materials and Methods

### *2.1. Data Retrieval*

Utilizing the keyword "Juvenile Idiopathic Arthritis," a comprehensive search of the GEO database was conducted. This search identified the GSE13849 dataset, comprising 59 healthy controls and 61 JIA patients, with all samples tested on the GPL570 platform. For data processing, the Perl language was utilized to effectively merge sample data and re-annotate genes. This included mapping probes to genes, discarding non-contributing empty probes, and accurately selecting the median expression level for genes represented by multiple probes.

### *2.2. Acquisition of Cuproptosis-Related Gene Expression Profiles*

A detailed selection of cuproptosis-related genes (NFE2L2, NLRP3, ATP7B, ATP7A, SLC31A1, FDX1, LIAS, LIPT1, LIPT2, DLD, DLAT, PDHA1, PDHB, MTF1, GLS, CDKN2A, DBT, GCSH, DLST) was derived from relevant literature. To precisely extract expression profiles for these 19 cuproptosis-related genes, the R software package 'limma' was skillfully applied.

### *2.3. Identification of HUB Genes*

In an innovative approach, the cuproptosis-related gene expression profiles obtained in step 2.2 were meticulously screened for potential key JIA genes using the Boruta algorithm and SVM-REF. To ensure robust analysis, the Boruta algorithm was configured with a doTrace parameter of 2 and a maximum iteration limit of 60. Similarly, the SVM-REF algorithm was optimally set with a radial basis function kernel "CV" and cross-validated 5 times. The integration of results from both methods facilitated the identification of core cuproptosis-related genes in JIA.

### *2.4. Model Construction and Validation*

Leveraging the HUB Genes obtained in step 2.3, a predictive model was meticulously developed using multivariate logistic regression, coupled with the construction of nomograms. To ascertain the clinical utility of the predictive model, decision curve analysis methodically assessed the net benefit at different threshold probabilities. Furthermore, the performance of the predictive model was rigorously evaluated by measuring the C-index, ROC curves, and calibration curves.

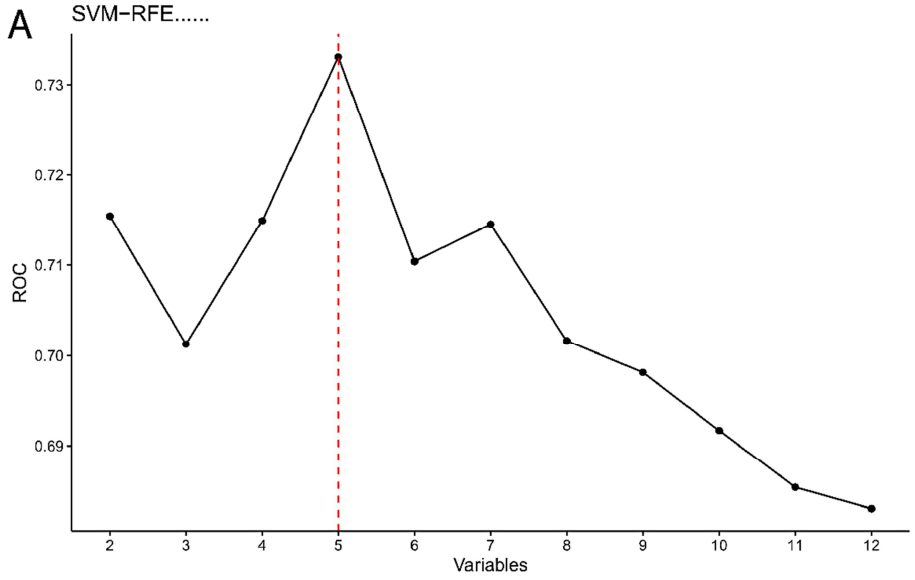
### *2.5. Internal Validation*

For rigorous testing, 70% of the data was strategically selected for internal validation. Undergoing internal validation with 1000 Bootstrap resamples, the model's robustness was confirmed by calculating the adjusted C-index. Additionally, the model's performance on the internal validation dataset was comprehensively assessed using the ROC curve.

# 3. Results

### *3.1. HUB Genes*

Four HUB genes (PDHA1, DBT, LIAS, DLAT) were conclusively identified using the Boruta algorithm. Simultaneously, the SVM-REF algorithm pinpointed five HUB Genes (PDHA1, LIAS, DLAT, LIPT1, PDHB). An intersection of these results highlighted three critical HUB Genes (PDHA1, LIAS, DLAT) as pivotal in JIA (Figure 1).



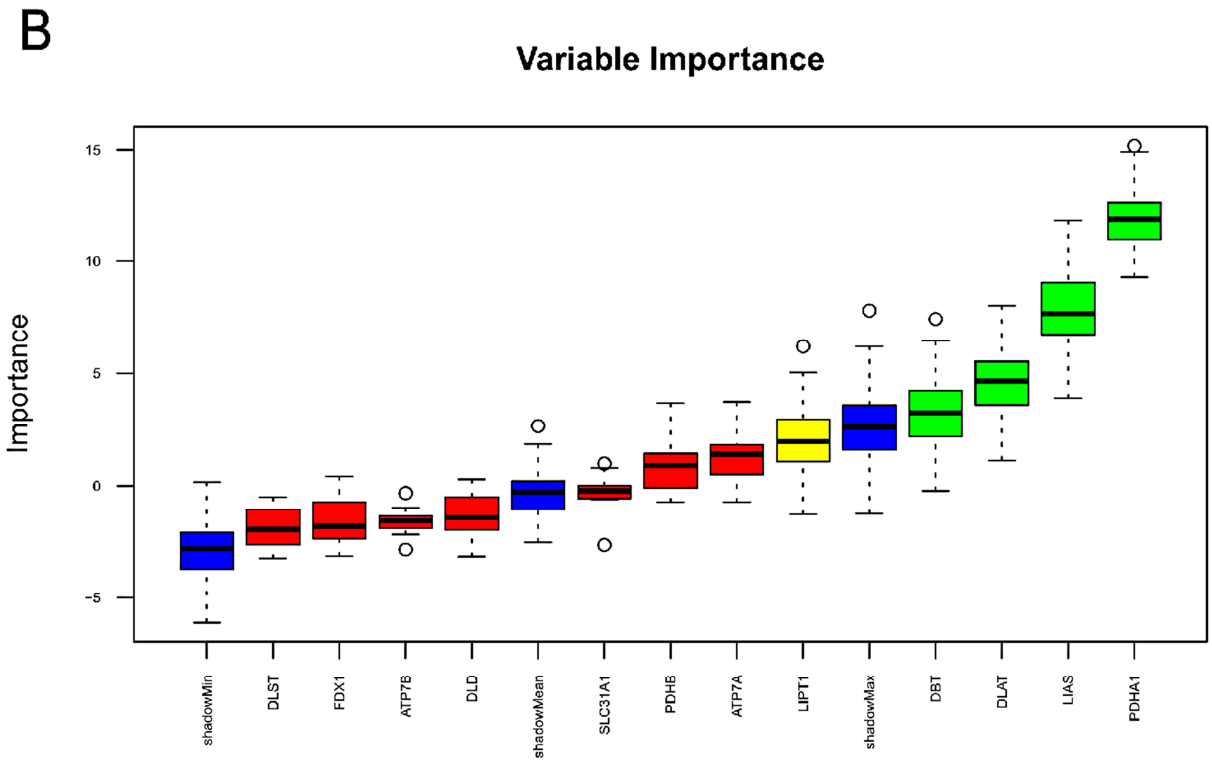



***Figure 1.*** *Key Gene Selection (A. Boruta B. SVM-REF).*

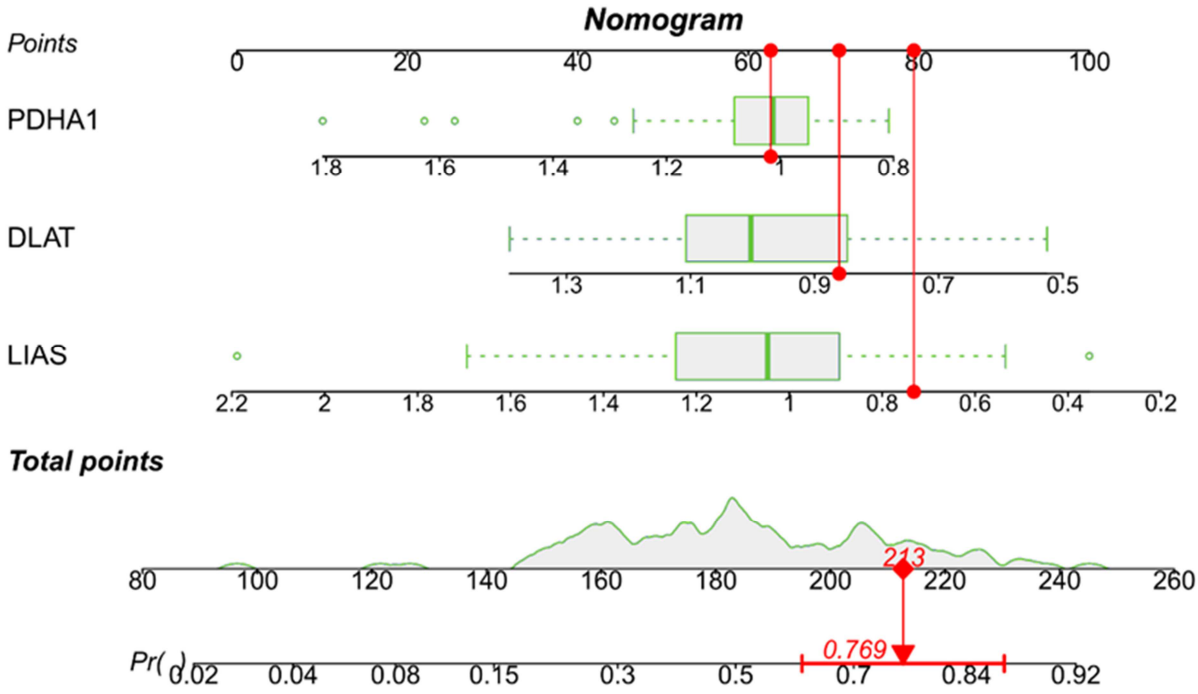



***Figure 2.*** *Clinical Prediction Model Nomogram.*

### *3.2. Clinical Prediction Model Results*

Utilizing the three identified HUB genes, a clinical prediction model was expertly constructed using multivariate logistic regression and effectively visualized with nomograms (Figure 2). The model's C-index stood at 0.75. The ROC curve demonstrated an area under the curve of 0.749 (Figure 3A), while the calibration curve exhibited good consistency (Figure 3B). Decision curve analysis revealed the highest net clinical benefit within a threshold probability range of 0.15 to 0.9, ensuring minimal risk to other patients (Figure 4).

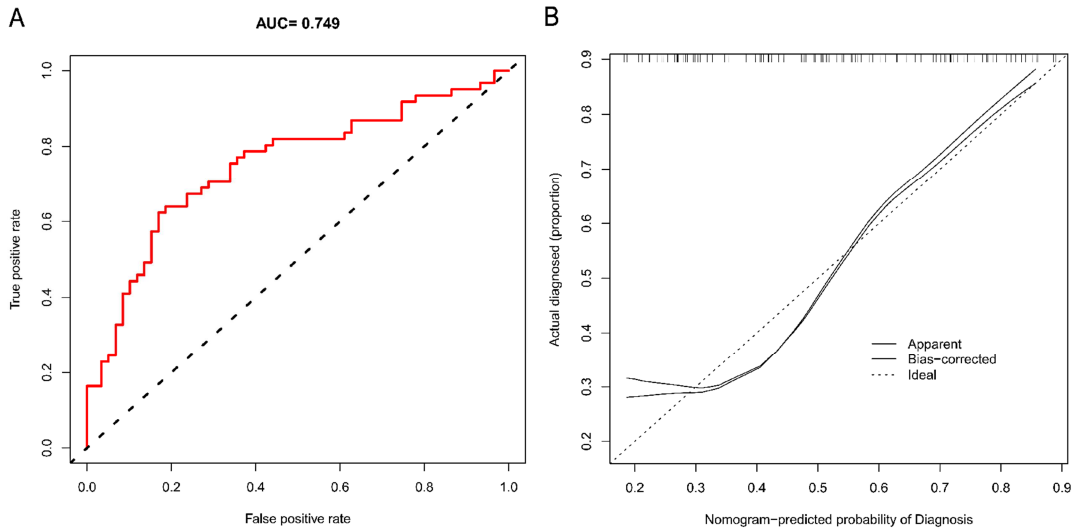



***Figure 3.*** *Predictive Model ROC Curve and Calibration curve (A: Area under the ROC curve is 0.749. B: The dashed line represents the ideal model, and the solid line represents the actual performance of the nomogram. The closer the proximity of the solid line to the dashed line, the more robust the predictive power of the model.)*

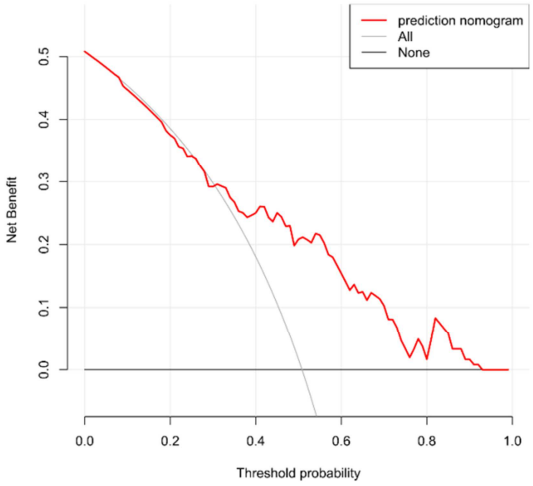



***Figure 4.*** *Clinical Decision Curve.*

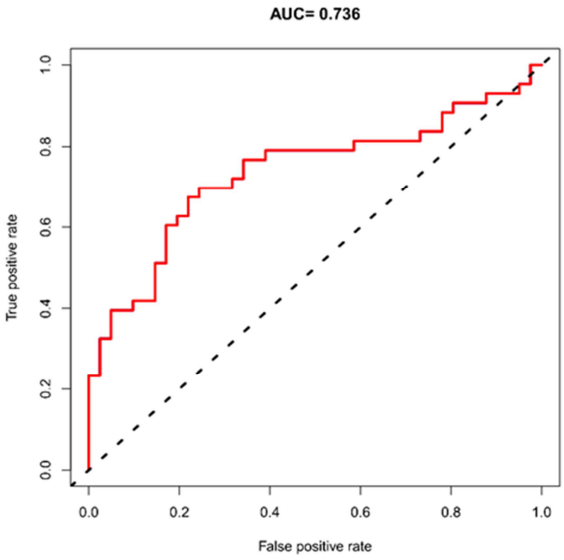



***Figure 5.*** *ROC Curve of Internal Validation Data Set.*

### *3.3. Internal Validation*

Internal validation conducted with 70% of the data revealed an internal validation set C-index of 0.755. The ROC curve further displayed an area under the curve of 0.736, corroborating the model's efficacy (Figure 5).

## 4. Discussion

The challenge of early diagnosis in Juvenile Idiopathic Arthritis (JIA) often stems from the need to differentiate it from various conditions such as infections, malignancies, and other auto-inflammatory or autoimmune diseases [6-8]. Therefore, the identification of specific biomarkers that can distinctly differentiate JIA and its differential diagnoses, potentially connected to the disease's pathogenesis, is vital for its early diagnosis. In recent years, significant advancements have been made in the study of biomarkers. For instance, research by Rosina et al. has indicated that certain biomarkers can predict disease activity, clinical remission, relapse, and drug response in JIA [9]. Additionally, Nziza et al. identified specific miRNA biomarkers, such as synovial fluid miRNA, which accurately differentiate between JIA and septic arthritis [10]. Souma et al.'s research suggests that serum levels of interleukin-18 may serve as a useful biomarker for the early diagnosis of JIA [11]. Moreover, advancements in ultrasound and magnetic resonance imaging techniques have provided more precise tools for the early diagnosis and assessment of disease activity in JIA [12].

Recent advances in machine learning within the medical field have been significant, with numerous studies highlighting its critical role in disease biomarker identification and prognosis evaluation [13, 14]. In JIA research, machine learning has been instrumental in developing predictive models for distinguishing JIA subtypes and predicting disease progression [15]. These models have also been effectively used in evaluating patient responses to JIA treatments. By

analyzing the variations in clinical data and biomarker levels before and after treatment, these models can anticipate patient reactions to specific medications, thereby enhancing clinical decision-making and optimizing treatment strategies [16].

Copper, an essential trace element, plays a crucial role in various biological processes, such as energy metabolism, antioxidant defense, and cellular signaling. In the TCA cycle, copper ions are vital cofactors for several enzymes. For instance, enzymes like cytochrome c oxidase require copper for their activity. Copper ions also participate in redox reactions within the electron transport chain, forming key components of electron transfer [17]. Recent studies have shown that intermediates and derivatives of the TCA cycle are extremely important in activating both innate and adaptive immune cells and in synthesizing pro-inflammatory mediators [18]. TCA cycle metabolites like succinate regulate inflammation and tissue repair by inhibiting inflammatory genes and promoting antioxidant gene expression [19]. Furthermore, succinate modulates inflammation by directly modifying cysteine sites on proteins involved in inflammasomes, signaling, transcription, and cell death [20]. Another TCA cycle intermediate, enolase, acts as a cell death inhibitor, influencing inflammation by converting GSDMD cysteine into S-(2-succinyl)-cysteine [21]. The pathology of JIA, characterized by synovial inflammation, involves an imbalance between pro-inflammatory effector cells and anti-inflammatory regulatory cells. Thus, regulating inflammation is a critical aspect of JIA treatment. In JIA, copper ions may regulate inflammation by inducing cuproptosis through the TCA cycle, thereby affecting the disease's progression.

This study developed a clinical predictive model for JIA diagnosis, identifying three cuproptosis-related genes as potential key contributors to JIA pathogenesis. The model's effectiveness in differentiating between normal individuals and JIA patients is evidenced by several indicators, such as ROC and calibration curves. Currently, these three genes are relatively underexplored in JIA, with most research focusing on oncology. For example, the Pyruvate Dehydrogenase E1 Component Subunit Alpha (PDHA1) is a crucial gene in cuproptosis, primarily involved in altering glucose metabolism in cancer cells. As a part of the pyruvate dehydrogenase complex, PDHA1 is essential for cellular metabolic processes. Studies indicate abnormal PDHA1 expression in most cancer types, correlating its expression levels with prognosis in cancers like lung adenocarcinoma, renal clear cell carcinoma, and gastric adenocarcinoma. PDHA1 also regulates mitochondrial signaling pathways, including oxidative phosphorylation, cellular respiration, and electron transport activities [22]. The Dihydrolipoamide S-Acetyltransferase (DLAT) gene, associated with cuproptosis, is notably upregulated in liver cancer tissues. Elevated DLAT expression is an independent prognostic factor for reduced overall survival in Hepatocellular Carcinoma (HCC). DLAT's roles encompass cellular metabolism, tumor progression, and immune regulation, with its expression levels linked to immune cell infiltration and checkpoint levels in HCC. Moreover, HCCs with high DLAT expression are predicted to be more responsive to sorafenib treatment [23]. LIAS, a component of the lipoic acid pathway that includes genes like FDX1, LIPT1, LIAS, and DLD, plays a significant role in cellular processes associated with cuproptosis. Suppressing LIAS expression results in reduced liver alpha-lipoic acid levels and increased tissue oxidative stress [24]. In summary, based on existing research, it is postulated that these three cuproptosis-related genes influence inflammatory pathways through metabolic pathways, playing a role in JIA pathogenesis. However, further in-depth research is required to understand the mechanisms of cuproptosis in JIA.

# 5. Conclusion

This study, utilizing machine learning and bioinformatics methods, identified three core genes associated with cuproptosis in JIA and developed a predictive model for these cuproptosis-related genes. The model exhibits high accuracy, providing a new approach for the early diagnosis of JIA. This holds significant importance for the early diagnosis and treatment of JIA, indicating that cuproptosis may play a crucial role in its pathogenesis. However, the study has limitations, as it did not confirm the expression levels of PDHA1, LIAS, and DLAT in JIA patients, nor did it delve into their specific mechanisms of action. Further experimental validation is required..

# Funding

The present study was funded by the Chinese Medicine Research Project of Traditional Chinese Medicine Bureau Of Guangdong Province (20221342), and the Shenzhen Futian District Health and Public Welfare Research Project (grant no. FTWS2022021, FTWS2022043, FTWS2023027, FTWS2023046).

# Author Contributions

Conceptualisation: Jianwei XIAO. Data curation: Yiwei HONG. Data analysis: Yiwei HONG, Xinpeng CHEN. Writing-original draft: Yiwei HONG, Xu Cai, Xinmin Huang, YAN Zhengbo. Approval of final manuscript: all authors.

# ORCID

Yiwei Hong: https://orcid.org/0009-0006-6399-0277
Xu Cai: https://orcid.org/0000-0002-8553-1585
Xinpeng Chen: https://orcid.org/0009-0008-7455-3173
Xinmin Huang: https://orcid.org/0009-0003-3183-8869
Zhengbo Yan: https://orcid.org/0000-0002-7954-5358,
Jianwei Xiao: https://orcid.org/0000-0002-1178-538X

# Acknowledgments

We sincerely thank the contributors of the GEO database

for making their data publicly available, which has facilitated this research.

## Conflicts of Interest

The authors declare no conflicts of interest.

## References

[1] Petty RE, Southwood TR, Manners P, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol. 2004; 31(2): 390-392.

[2] Ravelli A, Martini A. Juvenile idiopathic arthritis. Lancet. 2007; 369(9563): 767-778. doi: 10.1016/S0140-6736(07)60363-8.

[3] Prakken B, Albani S, Martini A. Juvenile idiopathic arthritis. Lancet. 2011; 377(9783): 2138-2149. doi: 10.1016/S0140-6736(11)60244-4.

[4] Thierry S, Fautrel B, Lemelle I, Guillemin F. Prevalence and incidence of juvenile idiopathic arthritis: a systematic review. Joint Bone Spine. 2014; 81(2): 112-117. doi: 10.1016/j.jbspin.2013.09.003.

[5] Tsvetkov P, Coy S, Petrova B, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. Science. 2022; 375(6586): 1254-1261. doi: 10.1126/science.abf0529.

[6] Wittkowski H, Frosch M, Wulffraat N, et al. S100A12 is a novel molecular marker differentiating systemic-onset juvenile idiopathic arthritis from other causes of fever of unknown origin. Arthritis Rheum. 2008; 58(12): 3924-3931. doi: 10.1002/art.24137.

[7] Shimizu M, Nakagishi Y, Yachie A. Distinct subsets of patients with systemic juvenile idiopathic arthritis based on their cytokine profiles. Cytokine. 2013; 61(2): 345-348. doi: 10.1016/j.cyto.2012.11.025.

[8] Park C, Miranda-Garcia M, Berendes R, et al. MRP8/14 serum levels as diagnostic markers for systemic juvenile idiopathic arthritis in children with prolonged fever. Rheumatology (Oxford). 2022; 61(7): 3082-3092. doi: 10.1093/rheumatology/keab729.

[9] Rosina S, Natoli V, Santaniello S, et al. Novel biomarkers for prediction of outcome and therapeutic response in juvenile idiopathic arthritis. Expert Rev Clin Immunol. 2021; 17(8): 853-870. doi: 10.1080/1744666X.2021.1945441.

[10] Nziza N, Jeziorski E, Delpont M, et al. Synovial-Fluid miRNA Signature for Diagnosis of Juvenile Idiopathic Arthritis. Cells. 2019; 8(12): 1521. doi: 10.3390/cells8121521.

[11] Souma K, Fujimura J, Nishiyama A, et al. Serum interleukin-18 level as a possible early diagnostic marker of systemic juvenile idiopathic arthritis. Int J Rheum Dis. 2023; 26(3): 551-553. doi: 10.1111/1756-185X.14534.

[12] Sudoł-Szopińska I, Herregods N, Doria AS, et al. Advances in Musculoskeletal Imaging in Juvenile Idiopathic Arthritis. Biomedicines. 2022; 10(10): 2417. doi: 10.3390/biomedicines10102417.

[13] Mello, A. C., Freitas, M., Coutinho, L., et al. Machine Learning Supports Long Noncoding RNAs as Expression Markers for Endometrial Carcinoma. BioMed research international, 2020, 3968279. https://doi.org/10.1155/2020/3968279.

[14] Hu Y, Guo G, Li J, et al. Screening key lncRNAs with diagnostic and prognostic value for head and neck squamous cell carcinoma based on machine learning and mRNA-lncRNA co-expression network analysis. Cancer Biomark. 2020; 27(2): 195-206. doi: 10.3233/CBM-19069.

[15] Van Nieuwenhove E, Lagou V, Van Eyck L, et al. Machine learning identifies an immunological pattern associated with multiple juvenile idiopathic arthritis subtypes. Ann Rheum Dis. 2019 May; 78(5): 617-628. doi: 10.1136/annrheumdis-2018-214354.

[16] Mo X, Chen X, Ieong C, et al Early Prediction of Clinical Response to Etanercept Treatment in Juvenile Idiopathic Arthritis Using Machine Learning. Front Pharmacol. 2020 Jul 31; 11: 1164. doi: 10.3389/fphar.2020.01164.

[17] Baker ZN, Cobine PA, Leary SC. The mitochondrion: a central architect of copper homeostasis. Metallomics. 2017; 9(11): 1501-1512. doi: 10.1039/c7mt00221a.

[18] Shi J, Cai C. Research Progress on the Mechanism of Itaconate Regulating Macrophage Immunometabolism. Front Immunol. 2022; 13: 937247. Published 2022 Jun 23. doi: 10.3389/fimmu.2022.937247.

[19] Nakkala JR, Yao Y, Zhai Z, et al. Dimethyl Itaconate-Loaded Nanofibers Rewrite Macrophage Polarization, Reduce Inflammation, and Enhance Repair of Myocardic Infarction. Small. 2021; 17(17): e2006992. doi: 10.1002/smll.202006992.

[20] Lin J, Ren J, Gao DS, et al. The Emerging Application of Itaconate: Promising Molecular Targets and Therapeutic Opportunities. Front Chem. 2021; 9: 669308. doi: 10.3389/fchem.2021.669308.

[21] Ji DY, Park SH, Park SJ, et al. Comparative assessment of Graves' disease and main extrathyroidal manifestation, Graves' ophthalmopathy, by non-targeted metabolite profiling of blood and orbital tissue. Sci Rep. 2018; 8(1): 9262. doi: 10.1038/s41598-018-27600-0.

[22] Ran XM, Xiao H, Tang YX, et al. The effect of cuproptosis-relevant genes on the immune infiltration and metabolism of gynecological oncology by multiply analysis and experiments validation. Sci Rep. 2023; 13(1): 19474. doi: 10.1038/s41598-023-45076-5.

[23] Zhang P, Zhao JH, Yuan LX, et al. DLAT is a promising prognostic marker and therapeutic target for hepatocellular carcinoma: a comprehensive study based on public databases. Sci Rep. 2023; 13(1): 17295. doi: 10.1038/s41598-023-43835-y.

[24] Yi X, Kim K, Yuan W, et al. Mice with heterozygous deficiency of lipoic acid synthase have an increased sensitivity to lipopolysaccharide-induced tissue injury. J Leukoc Biol. 2009; 85(1): 146-153. doi: 10.1189/jlb.0308161.