

# Advances

## in Clinical and Experimental Medicine

MONTHLY ISSN 1899-5276 (PRINT) ISSN 2451-2680 (ONLINE)

[advances.umw.edu.pl](http://advances.umw.edu.pl)

2023, Vol. 32, No. 2 (February)

Impact Factor (IF) – 1.736  
Ministry of Science and Higher Education – 70 pts  
Index Copernicus (ICV) – 168.52 pts



WROCLAW  
MEDICAL UNIVERSITY

Advances  
in Clinical and Experimental  
Medicine



# Advances in Clinical and Experimental Medicine

ISSN 1899-5276 (PRINT)

ISSN 2451-2680 (ONLINE)

advances.umw.edu.pl

**MONTHLY 2023**  
**Vol. 32, No. 2**  
**(February)**

Advances in Clinical and Experimental Medicine (*Adv Clin Exp Med*) publishes high-quality original articles, research-in-progress, research letters and systematic reviews and meta-analyses of recognized scientists that deal with all clinical and experimental medicine.

## Editorial Office

ul. Marcinkowskiego 2–6  
50-368 Wrocław, Poland  
Tel.: +48 71 784 12 05  
E-mail: redakcja@umw.edu.pl

## Publisher

Wrocław Medical University  
Wybrzeże L. Pasteura 1  
50-367 Wrocław, Poland

Online edition is the original version  
of the journal

## Editor-in-Chief

Prof. Donata Kurpas

## Deputy Editor

Prof. Wojciech Kosmala

## Managing Editor

Marek Misiak, MA

## Statistical Editors

Wojciech Bombała, MSc  
Anna Kopszak, MSc  
Dr. Krzysztof Kujawa

## Manuscript editing

Marek Misiak, MA, Jolanta Krzyżak, MA

## Scientific Committee

Prof. Sabine Bährer-Kohler  
Prof. Antonio Cano  
Prof. Breno Diniz  
Prof. Erwan Donal  
Prof. Chris Fox  
Prof. Naomi Hachiya  
Prof. Carol Holland  
Prof. Markku Kurkinen  
Prof. Christos Lionis

Prof. Raimundo Mateos  
Prof. Zbigniew W. Raś  
Prof. Jerzy W. Rozenblit  
Prof. Silvana Santana  
Prof. James Sharman  
Prof. Jamil Shibli  
Prof. Michał Toborek  
Prof. László Vécsei  
Prof. Cristiana Vitale

## Section Editors

### Anesthesiology

Prof. Marzena Zielińska

### Basic Sciences

Prof. Iwona Bil-Lula  
Prof. Bartosz Kempisty  
Dr. Wiesława Kranc  
Dr. Anna Lebedeva  
Dr. Maciej Sobczyński

### Clinical Anatomy, Legal Medicine, Innovative Technologies

Prof. Rafael Boscolo-Berto

### Dentistry

Prof. Marzena Dominiak  
Prof. Tomasz Gedrange  
Prof. Jamil Shibli

### Dermatology

Prof. Jacek Szepietowski

### Emergency Medicine, Innovative Technologies

Prof. Jacek Smereka

### Gynecology and Obstetrics

Prof. Olimpia Sipak-Szmigiel

### Histology and Embryology

Dr. Mateusz Olbromski

### Internal Medicine

#### Angiology

Dr. Angelika Chachaj

#### Cardiology

Prof. Wojciech Kosmala  
Dr. Daniel Morris

#### Endocrinology

Prof. Marek Bolanowski

### Gastroenterology

Assoc. Prof. Katarzyna Neubauer

### Hematology

Prof. Andrzej Deptała

Prof. Dariusz Wołowicz

### Nephrology and Transplantology

Assoc. Prof. Dorota Kamińska

Assoc. Prof. Krzysztof Letachowicz

### Pulmonology

Prof. Anna Brzecka

### Microbiology

Prof. Marzenna Bartoszewicz

Assoc. Prof. Adam Junka

### Molecular Biology

Dr. Monika Bielecka

Prof. Jolanta Saczko

### Neurology

Assoc. Prof. Magdalena Koszewicz

Assoc. Prof. Anna Pokryszko-Dragan

Dr. Masaru Tanaka

### Neuroscience

Dr. Simone Battaglia

### Oncology

Prof. Andrzej Deptała

Dr. Marcin Jędryka

### Gynecological Oncology

Dr. Marcin Jędryka

### Orthopedics

Prof. Paweł Reichert

### Otolaryngology

Assoc. Prof. Tomasz Zatoński

### Pediatrics

Pediatrics, Metabolic Pediatrics, Clinical Genetics, Neonatology, Rare Disorders

Prof. Robert Śmigiel

### Pediatric Nephrology

Prof. Katarzyna Kiliś-Pstrusińska

### Pediatric Oncology and Hematology

Assoc. Prof. Marek Ussowicz

### Pharmaceutical Sciences

Assoc. Prof. Marta Kepinska

Prof. Adam Matkowski

### Pharmacoeconomics, Rheumatology

Dr. Sylwia Szafraniec-Buryło

### Psychiatry

Prof. Jerzy Leszek

### Public Health

Prof. Monika Sawhney

Prof. Izabella Uchmanowicz

### Qualitative Studies, Quality of Care

Prof. Ludmiła Marcinowicz

### Radiology

Prof. Marek Szaśniadek

### Rehabilitation

Dr. Elżbieta Rajkowska-Labon

### Surgery

Assoc. Prof. Mariusz Chabowski

Prof. Renata Tabała

### Telemedicine, Geriatrics, Multimorbidity

Assoc. Prof. Maria Magdalena

Bujnowska-Fedak

## Editorial Policy

Advances in Clinical and Experimental Medicine (Adv Clin Exp Med) is an independent multidisciplinary forum for exchange of scientific and clinical information, publishing original research and news encompassing all aspects of medicine, including molecular biology, biochemistry, genetics, biotechnology and other areas. During the review process, the Editorial Board conforms to the "Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication" approved by the International Committee of Medical Journal Editors ([www.ICMJE.org](http://www.ICMJE.org)). The journal publishes (in English only) original papers and reviews. Short works considered original, novel and significant are given priority. Experimental studies must include a statement that the experimental protocol and informed consent procedure were in compliance with the Helsinki Convention and were approved by an ethics committee.

For all subscription-related queries please contact our Editorial Office: [redakcja@umw.edu.pl](mailto:redakcja@umw.edu.pl)

For more information visit the journal's website: [advances.umw.edu.pl](http://advances.umw.edu.pl)

Pursuant to the ordinance of the Rector of Wrocław Medical University No. 12/XVI R/2023, from February 1, 2023, authors are required to pay a fee for each manuscript accepted for publication in the journal Advances in Clinical and Experimental Medicine. The fee amounts to 990 EUR for original papers and meta-analyses, 700 EUR for reviews, and 350 EUR for research-in-progress (RIP) papers and research letters.

Advances in Clinical and Experimental Medicine has received financial support from the resources of Ministry of Science and Higher Education within the "Social Responsibility of Science – Support for Academic Publishing" project based on agreement No. RCN/SP/0584/2021.



Ministry of Education and Science  
Republic of Poland

Czasopismo Advances in Clinical and Experimental Medicine korzysta ze wsparcia finansowego ze środków Ministerstwa Edukacji i Nauki w ramach programu „Społeczna Odpowiedzialność Nauki – Rozwój Czasopism Naukowych” na podstawie umowy nr RCN/SP/0584/2021.



Ministerstwo  
Edukacji i Nauki

Indexed in: MEDLINE, Science Citation Index Expanded, Journal Citation Reports/Science Edition, Scopus, EMBASE/Excerpta Medica, Ulrich's™ International Periodicals Directory, Index Copernicus

Typographic design: Piotr Gil, Monika Kołęda

DTP: Wydawnictwo UMW

Cover: Monika Kołęda

Printing and binding: Drukarnia I-BiS Bierońscy Sp.k.

## Contents

### Editorials

- 141 Marcin Jędryka  
**Translational molecular research has established a novel clinical approach in endometrial cancer patients**
- 147 Shlomo Vinker  
**Innovations in family medicine and the implication to rural and remote primary care**

### Meta-analyses

- 151 Xin Zhao, Hong Zhang, Yanyan Wu, Chun Yu  
**The efficacy and safety of St. John's wort extract in depression therapy compared to SSRIs in adults: A meta-analysis of randomized clinical trials**
- 163 Xiangxiang Cui, Dan Zhong, Jinou Zheng  
**A meta-analysis to investigate the role of magnetic resonance spectroscopy in the detection of temporal lobe epilepsy**
- 175 Wanying Guo  
**Evaluation of the impact of kangaroo mother care on neonatal mortality and hospitalization: A meta-analysis**

### Original papers

- 185 Yang Zhang, Lin Li, Ye Li, Zhihe Zeng  
**Machine learning model-based risk prediction of severe complications after off-pump coronary artery bypass grafting**
- 195 Huijuan Zhou, Shuangdi Li  
**Evaluation of pretreatment and early treatment changes in serum  $\beta$ -hCG with methotrexate**
- 203 Aleksandra Królikowska, Agnieszka Maj, Maciej Dejneka, Robert Prill, Anna Skotowska-Machaj, Anna Kołcz  
**Wrist motion assessment using Microsoft Azure Kinect DK: A reliability study in healthy individuals**
- 211 Olga Kisiel, Agnieszka Ewa Siennicka, Krystian Josiak, Robert Zymliński, Waldemar Banasiak, Kinga Węgrzynowska-Teodorczyk  
**Impact of assisted exercises on skeletal muscle oxygenation levels in men with acutely decompensated heart failure**
- 219 Feng Yang, Wei Wang, Yiling Zhang, Jifei Nong, Longdan Zhang  
**Effects of ferroptosis in myocardial ischemia/reperfusion model of rat and its association with Sestrin 1**
- 233 Lanlan Zhang, Xiaomin Chen, Xiangchai Guo, Huajuan Shen, Danying Qiu, Weiqun Jin, Dongjie Hou  
**Comprehensive analysis of cell death genes in hepatocellular carcinoma based on multi-omics data**
- 245 Konrad Kisielowski, Bogna Drozdowska, Mariusz Szuta, Tomasz Kaczmarzyk  
**Prognostic relevance of clinicopathological factors in sporadic and syndromic odontogenic keratocysts: A comparative study**



# Translational molecular research has established a novel clinical approach in endometrial cancer patients

Marcin Jędryka<sup>1,2,A–F</sup>

<sup>1</sup> Chair of Oncology, Department of Oncological Gynaecology, Wrocław Medical University, Poland

<sup>2</sup> Department of Oncological Gynaecology, Lower Silesian Oncology Center, Wrocław, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):141–145

## Address for correspondence

Marcin Jędryka

E-mail: marcin.jedryka@umw.edu.pl

## Funding sources

None declared

## Conflict of interest

None declared

Received on October 9, 2022

Reviewed on December 23, 2022

Accepted on December 26, 2022

Published online on January 13, 2023

## Abstract

Endometrial cancer (EC) is the most common gynecological tumor in developed countries. Nowadays, molecular biomarkers have become increasingly important in the management of EC patients, helping in early detection, risk stratification, prognosis and response to the treatment, and qualification for novel immunotherapies, especially in EC patients with metastatic disease or recurrence. When EC tumor molecular profiling is combined with the standard histopathological features such as clinical stage, histologic grading and evaluation of lymphovascular space invasion, the final therapeutic outcome may bring benefits in terms of personalized and efficient management.

**Key words:** endometrial cancer, molecular markers, translational researches

## Cite as

Jędryka M. Translational molecular research has established a novel clinical approach in endometrial cancer patients.

*Adv Clin Exp Med.* 2023;32(2):141–145.

doi:10.17219/acem/158556

## DOI

10.17219/acem/158556

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0)

(<https://creativecommons.org/licenses/by/3.0/>)

## Background

Endometrial cancer (EC) is the most common gynecological tumor in developed countries, and the 4<sup>th</sup> most frequent malignancy in women, with 61,800 estimated new cases (7% incidence) and 12,160 deaths (4%) in the USA in 2019.<sup>1,2</sup> Undoubtedly, the EC incidence has been rising together with the average age and obesity in these populations. This fact determines recent translational research effort for defining EC risk groups more precisely, including molecular characterization of tumor. Combining the standard approach, based mainly on histopathological features, with new molecular technologies of EC profiling translates this growing knowledge into novel diagnostic and therapeutic approach, resulting in well-tailored personalized management of this cancer.<sup>3,4</sup>

## Endometrial cancer traditional and novel classification systems

For many years, the Bokhman classification of 2 types endometrial carcinomas based on clinical, endocrine and epidemiological features has been used.<sup>5</sup> Type I tumors comprise majority of the cases (60–70%), with mostly endometrioid histology, estrogen-dependent – with high sex hormones receptors expression and endometrial hyperplasia in the background, associated with obesity, diabetes and reproductive dysfunctions, but on the other hand with predominant low-grade tumor, superficial myometrial invasion, low risk for lymph nodes metastases, and better prognosis in general. Type II tumors, usually serous carcinomas, are less common, arise in a non-estrogenic environment, and are associated with an atrophic endometrium of older, postmenopausal and non-obese women without reproductive and metabolic disturbances. Such tumors mostly show high-grade differentiation with deep myometrial invasion, increased risk for lymphogenic spread and poor final outcome.<sup>6</sup>

Another traditional World Health Organization (WHO) classification of endometrial neoplasms, based on histological features, distinguishes among epithelial carcinomas endometrioid, serous, clear cell, mucinous, squamous cell, transitional cell, small cell, and undifferentiated tumors.<sup>7</sup> Most common endometrioid adenocarcinomas can be differentiated into low-, intermediate- or high-grade tumors, whilst serous and clear cell cancers are aggressive and high-grade by definition.

The Bokhman and WHO classifications created the template for an easy-to-use, dichotomous model, practical for diagnostic, therapeutic and research approach. However, the wide implementation of this traditional classification is imperfect and sometimes misleading as biological, genetic and pathological features of endometrial carcinoma are complex and heterogeneous. There are not only overlaps between type I and II tumors, but heterogeneity is also

present within each of these types.<sup>6</sup> Furthermore, this dualistic model has been established almost 40 years ago and epidemiological characteristics of women suffering from EC has changed since that time due to widespread obesity followed by diabetes, increased use of hormone-replacement therapy, and many other multiple risk factors shared in both types.<sup>8</sup> Most of high-grade endometrioid carcinomas exhibit clinical, histopathological and molecular patterns placing them close to type II cancers, including no association with endometrial hyperplasia and poor outcome.<sup>8,9</sup> By contrast, some serous carcinomas arise in the background of endometrial hyperplasia, and about 20% of them lack deep myometrial infiltration.<sup>10</sup> Besides, the Bokhman model does not encompass such high-grade tumors as carcinosarcomas and clear cell or undifferentiated carcinomas, which are not so rare.<sup>6,11</sup> Likewise, endometrial cancer cases with predominant genetic predisposition, like Lynch syndrome, not associated with obesity and endometrial hyperplasia, have not been taken into account in this traditional classification.<sup>12</sup>

In recent years, next-generation sequencing (NGS) technology has changed the research approach from studying single genes or pathways into the insight of genome-wide genetic alterations in endometrial carcinomas.<sup>13</sup> The Cancer Genome Atlas Research Network (TCGA) assessed a large group of common endometrial cancers (endometrioid, serous and mixed) with NGS tests of mutation load and copy-number aberrations combined with microsatellite instability, reverse-phase protein array and DNA methylation analysis.<sup>14</sup> The results allowed to categorize endometrial tumors into 4 genomic clusters: 1) *POLE* tumors with very high mutation rates (ultramutated) and hotspot mutations in the exonuclease domain of polymerase  $\epsilon$  (*POLE*), few copy-number aberrations, mutations in *PTEN*, *PIK3R1*, *PIK3CA*, *FBXW7*, *KRAS*, and good prognosis; 2) a microsatellite-unstable group (MSI; hypermutated) characterized by microsatellite instability due to *MLH1* promoter methylation, high mutations rates, few copy-number aberrations, and *KRAS* and *PTEN* mutations; 3) copy-number low cluster compromising endometrioid microsatellite-stable tumors grade 1 and 2, with low mutation rates but frequent *CTNNB1* mutations; 4) copy-number high tumors (serous and some grade 3 endometrioid adenocarcinomas), characterized by extensive copy-number aberrations and low mutation rates, frequent and recurrent *TP53*, *FBXW7*, and *PPP2R1A* mutations, seldom *PTEN* and *KRAS* mutations, and poor outcome.<sup>6,14</sup> Only *POLE* hotspot mutations were unique for cluster 1, while other alterations and aberrations in studied genes and their pathways were expressed more or less in all 4 classes, overlapping themselves. The TCGA characterization of endometrial tumors 4 genomic classes in comparison with the traditional classification system of EC has been presented in Table 1.

This novel, molecular classification system of endometrial cancer has a great potential to develop better prognostic accuracy and provide predictive information.<sup>6</sup>



**Table 1.** Characteristics of the endometrial cancer genomic classification created using The Cancer Genome Atlas (TCGA) Network<sup>14</sup> in comparison with the dualistic Bokhman classification<sup>5</sup>

Feature	TCGA genomic classification				Bokhman's classification	
	POLE (ultramutated)	MSI (hypermutated)	copy-number low (endometrioid)	copy-number high (serous-like)	type I	type II
Mutation rate	very high	high	low	low	–	–
Microsatellite instability	mixed MSI high, low, stable	MSI high	MSI stable	MSI stable	28–40%	0–2%
Copy-number aberrations	low	low	low	high	–	–
Prevalent gene mutations	POLE (100%) PTEN (94%) PIK3CA (71%) PIK3R1 (65%) FBXW7 (82%) ARID1A (76%) KRAS (53%) ARID5B (47%)	PTEN (88%) RPL22 (37%) KRAS (35%) PIK3CA (54%) PIK3R1 (40%) ARID1A (37%)	PTEN (77%) CTNNB1 (52%) PIK3CA (53%) PIK3R1 (33%) ARID1A (42%)	TP53 (92%) PPP2R1A (22%) PIK3CA (47%)	PTEN (52–78%) PIK3CA (36–52%) PIK3R1 (21–43%) KRAS (15–43%) ARID1A (25–48%) CTNNB1 (23–24%)	TP53 (60–91%) HER2 (27–44%) PIK3CA (24–42%) PPP2R1A (15–43%)
Histological type	endometrioid	endometrioid	endometrioid	serous, endometrioid, mixed serous and endometrioid	endometrioid	serous
Grading	mixed (G1–3)	mixed (G1–3)	low (G1–2)	high (G3)	low (G1–2)	high (G3)
Prognosis	good	intermediate	intermediate	poor	good	poor

MSI – microsatellite instability; POLE – polymerase  $\epsilon$ ; G – grading.

The integration of genomic markers of an individual, endometrial tumor with its clinicopathological features has become an important tool for clinical decision-making, especially in the context of escalation or de-escalation of clinical management and introducing new targeted immunotherapies.

## Clinical management recommendations for EC patients based on integrated traditional pathologic and novel molecular features

The results reported by TCGA were retrospectively confirmed by analyzing data from PORTEC-1 and PORTEC-2 trials.<sup>15</sup> Molecular subgroups were combined with other significant EC biomarkers such as *LICAM*, *PTEN*,  $\beta$ -*catenin*, *ARID1a*, estrogen and progesterone expression, and lymphovascular space invasion. Integration of prognostic molecular alterations with established clinicopathological factors resulted in a stronger model with improved risk prognostication in high–intermediate risk cohort stratified into 3 subgroups of 50% (favorable), 35% (intermediate) and 15% (unfavorable) features. Another retrospective analysis (PORTEC-3 trial) of a cohort of high-risk EC patients randomized for adjuvant chemoradiotherapy compared to radiotherapy alone investigated the prognostic value of molecular EC classification, concluding that incorporation of this

classification into risk stratification is essential. Patients with p53abnormal (p53abn) mutation should be referred for adjuvant chemotherapy whilst high-risk *POLE*-mutated patients had much better outcome; therefore, de-escalation of adjuvant treatment should be considered.<sup>16</sup> While most EC cases can be classified based on a single TCGA classifier (*POLE* exonuclease domain mutations – *POL-Emut*, MMR deficiency – *MMRd*, p53 abnormal – p53abn), a small but clinically relevant group of tumors harbor more than 1 molecular classifying feature and are referred to as ‘multiple-classifier’ ECs.<sup>17</sup> This small population (3% of all ECs) consists of *MMRd*-p53abn, *POLEmut*-p53abn and *MMRd-POLEmut*-p53abn subgroups, and shows much better 5-year recurrence-free survival than single-classifier p53abn (94.1% compared to 70.8% for stage I), and thus could be de-escalated from the aggressive treatment, similarly to the single-classifier *POLEmut* EC.

Based on these recent translational studies, the newest, 5<sup>th</sup> edition of WHO classification of tumors of female genital tract has included molecular markers.<sup>18</sup> Furthermore, the National Comprehensive Cancer Network (NCCN) in its newest edition of *Clinical Practice Guidelines in Oncology* (2022) recommends molecular analysis of the endometrial tumor that could have an impact on the management of metastatic and recurrent cases presenting mismatch repair deficiency (dMMR/MSI-H tumors) regarding the therapeutic use of efficient immunotherapy (checkpoint inhibitors such as pembrolizumab, dostarlimab or nivolumab).<sup>19</sup> However, introducing molecular biomarkers into the clinical practice may be difficult and depends on availability of resources and the multidisciplinary team of each hospital

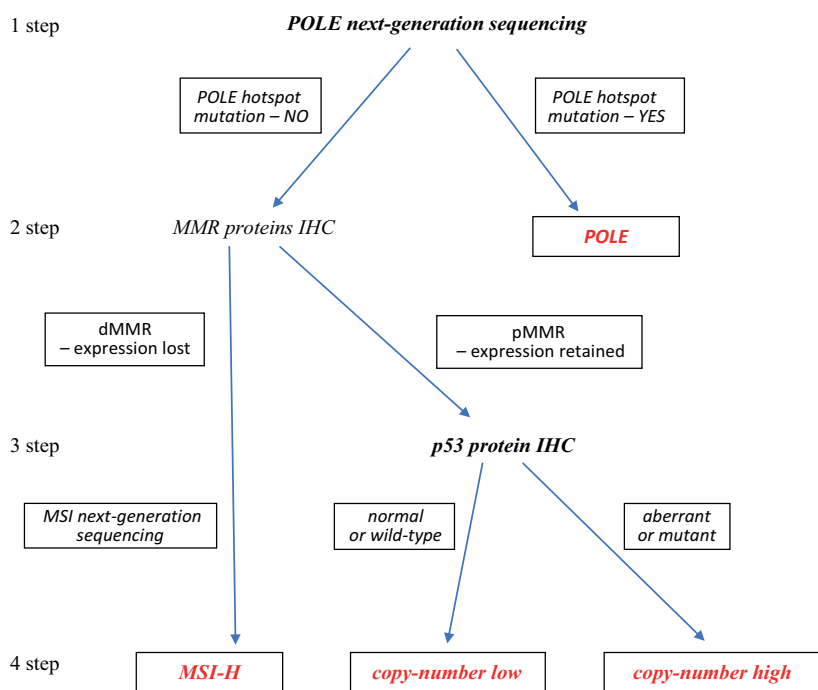


Fig. 1. Combine, molecular and histological diagnostic algorithm of endometrial cancer patients stratifying them into 4 groups of the new molecular classification (based on National Comprehensive Cancer Network (NCCN) 2022 guidelines)<sup>19</sup>

MSI – microsatellite instability; POLE – polymerase  $\epsilon$ ; IHC – immunohistochemistry; dMMR – mismatch repair deficient; pMMR – mismatch repair proficient.

dealing with the endometrial cancer management. Therefore, some immunohistochemical surrogates have been proposed to be combined with genomics and thus enable classifying each case into 1 from 4 clusters of the molecular classification.<sup>19,20</sup> Such molecular and histologic evaluation algorithm of each endometrial cancer case has been introduced in author's institution since the beginning of 2022 and is presented in Fig. 1.

The newest European Society of Gynecological Society/European Society for Radiotherapy and Oncology/European Society of Pathology (ESGO/ESTRO/ESP) 2021 guidelines for the management of patients with EC recommend using molecular classification in EC patients to define a prognostic risk group and to choose the best possible therapeutic strategy.<sup>21</sup> These comprehensive guidelines are a very good example of decision-making for escalation or de-escalation management of endometrial cancer patients depending on the presence or absence of molecular biomarkers. For instance, stage I–II patients with pathogenic *POLE* mutation may omit the adjuvant treatment. On the other hand, EC patients staged I–IVA with myometrial invasion and no residual disease but with p53abn mutation expressed should undergo external beam radiotherapy (EBRT) with concurrent and adjuvant chemotherapy.

To date, other sophisticated molecular studies in EC patients have been conducted in order to obtain better translational impact on clinical practice. Under the auspices of the National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC), multi-omic characterization of EC samples and appropriate normal tissues from a prospective cohort of 95 EC patients were extensively studied. Integrated measurements of DNA, RNA, proteins, and post-translational modifications (phosphorylation and acetylation) were used to identify novel regulatory relationships

and potential pathways for identifying therapeutic targets. This analysis revealed possible new consequences of perturbations to the p53 and Wnt/ $\beta$ -catenin pathways, identified a potential role for circRNAs in the epithelial–mesenchymal transition, and provided new information about proteomic markers of clinical and genomic tumor subgroups, including relationships to known druggable pathways.<sup>22</sup>

Another recent analysis used the Immuno-oncology panel and the Target 96 Oncology III panel to detect 92 specific cancer-related serum protein in early-stage EC cancer patients. The authors generated an algorithm based on multivariate logistic regression model that was able to discriminate early EC type I patients from controls with high specificity and sensitivity (83.7% and 97.6%, respectively) thanks to the analysis of *Gal-1*, *Gal-9*, *MMP7*, *COL9A1*, and *FASLG* serum levels.<sup>23</sup>

## Conclusions

The diagnostic and clinical landscape in EC has changed a lot in recent years. The translational studies on genomics and proteomics have a practical impact on management of this cancer. Novel molecular biomarkers are essential for stratifying the risk of each EC patient and creating a novel diagnostic tool for optimal therapeutic scenario and cancer surveillance regimens after the treatment completion. The use of new immunotherapies has been based on molecular-integrated risk profile to determine the treatment efficacy in patients with advanced and metastatic disease. However, further prospective studies are necessary to develop complete and commonly used molecular profiling that should be useful for tailoring the best individual EC management.

## ORCID iDs

Marcin Jędryka  <https://orcid.org/0000-0001-8935-0311>

## References

- Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J Clin.* 2021;71(3):209–249. doi:10.3322/caac.21660
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA A Cancer J Clin.* 2019;69(1):7–34. doi:10.3322/caac.21551
- Králíčková M, Větvicka V, Laganà AS. Endometrial cancer: Is our knowledge changing? *Transl Cancer Res.* 2020;9(12):7734–7745. doi:10.21037/tcr-20-1720
- Li L, Chen F, Liu J, et al. Molecular classification grade 3 endometrial endometrioid carcinoma using a next-generation sequencing-based gene panel. *Front Oncol.* 2022;12:935694. doi:10.3389/fonc.2022.935694
- Bokhman JV. Two pathogenetic types of endometrial carcinoma. *Gynecol Oncol.* 1983;15(1):10–17. doi:10.1016/0090-8258(83)90111-7
- Murali R, Soslow RA, Weigelt B. Classification of endometrial carcinoma: More than two types. *Lancet Oncol.* 2014;15(7):e268–e278. doi:10.1016/S1470-2045(13)70591-6
- Kurman RJ, ed. *WHO Classification of Tumours of Female Reproductive Organs.* 4<sup>th</sup> ed. Lyon, France: International Agency for Research on Cancer; 2014. ISBN:978-92-832-2435-8.
- Brinton LA, Felix AS, McMeekin DS, et al. Etiologic heterogeneity in endometrial cancer: Evidence from a Gynecologic Oncology Group trial. *Gynecol Oncol.* 2013;129(2):277–284. doi:10.1016/j.ygyno.2013.02.023
- Setiawan VW, Yang HP, Pike MC, et al. Type I and II endometrial cancers: Have they different risk factors? *J Clin Oncol.* 2013;31(20):2607–2618. doi:10.1200/JCO.2012.48.2596
- Soslow RA, Bissonnette JP, Wilton A, et al. Clinicopathologic analysis of 187 high-grade endometrial carcinomas of different histologic subtypes: Similar outcomes belie distinctive biologic differences. *Am J Surg Pathol.* 2007;31(7):979–987. doi:10.1097/PAS.0b013e31802ee494
- Zannoni GF, Scambia G, Gallo D. The dualistic model of endometrial cancer: The challenge of classifying grade 3 endometrioid carcinoma. *Gynecol Oncol.* 2012;127(1):262–263. doi:10.1016/j.ygyno.2011.09.036
- Garg K, Soslow RA. Lynch syndrome (hereditary non-polyposis colorectal cancer) and endometrial carcinoma. *J Clin Pathol.* 2009;62(8):679–684. doi:10.1136/jcp.2009.064949
- Talhok A, McConechy MK, Leung S, et al. A clinically applicable molecular-based classification for endometrial cancers. *Br J Cancer.* 2015;113(2):299–310. doi:10.1038/bjc.2015.190
- The Cancer Genome Atlas Research Network; Kandoth C, Schultz N, Cherniack AD, et al. Integrated genomic characterization of endometrial carcinoma. *Nature.* 2013;497(7447):67–73. doi:10.1038/nature12113
- Stelloo E, Nout RA, Osse EM, et al. Improved risk assessment by integrating molecular and clinicopathological factors in early-stage endometrial cancer: Combined analysis of the PORTEC cohorts. *Clin Cancer Res.* 2016;22(16):4215–4224. doi:10.1158/1078-0432.CCR-15-2878
- León-Castillo A, de Boer SM, Powell ME, et al. Molecular classification of the PORTEC-3 trial for high-risk endometrial cancer: Impact on prognosis and benefit from adjuvant therapy. *J Clin Oncol.* 2020;38(29):3388–3397. doi:10.1200/JCO.20.00549
- León-Castillo A, Gilvazquez E, Nout R, et al. Clinicopathological and molecular characterisation of ‘multiple-classifier’ endometrial carcinomas. *J Pathol.* 2020;250(3):312–322. doi:10.1002/path.5373
- McCluggage WG, Singh N, Gilks CB. Key changes to the World Health Organization (WHO) classification of female genital tumours introduced in the 5<sup>th</sup> edition (2020). *Histopathology.* 2022;80(5):762–778. doi:10.1111/his.14609
- National Comprehensive Cancer Network. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines®): Uterine Neoplasms. Version 1.2022.* Plymouth Meeting, USA: National Comprehensive Cancer Network; 2022. <https://www.nccn.org/guidelines/guidelines-detail?category=1&id=1473>. Accessed October 7, 2022.
- Vermij L, Smit V, Nout R, Bosse T. Incorporation of molecular characteristics into endometrial cancer management. *Histopathology.* 2020;76(1):52–63. doi:10.1111/his.14015
- Concin N, Matias-Guiu X, Vergote I, et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. *Int J Gynecol Cancer.* 2021;31(1):12–39. doi:10.1136/ijgc-2020-002230
- Dou Y, Kawaler EA, Cui Zhou D, et al. Proteogenomic characterization of endometrial carcinoma. *Cell.* 2020;180(4):729–748.e26. doi:10.1016/j.cell.2020.01.026
- Ura B, Capaci V, Aloisio M, et al. A targeted proteomics approach for screening serum biomarkers observed in the early stage of type I endometrial cancer. *Biomedicines.* 2022;10(8):1857. doi:10.3390/biomedicines10081857



# Innovations in family medicine and the implication to rural and remote primary care\*

Shlomo Vinker<sup>1,2,A–F</sup>

<sup>1</sup> Department of Family Medicine, Sackler School of Medicine, Tel Aviv University, Israel

<sup>2</sup> World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA) Europe President

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):147–150

## Address for correspondence

Shlomo Vinker

E-mail: [vinker01@gmail.com](mailto:vinker01@gmail.com)

## Funding sources

None declared

## Conflict of interest

None declared

\* A lecture provided during the 11<sup>th</sup> European Rural and Isolated Practitioners Association (EURIPA) Rural Health Forum (Catania, Sicily, Italy, October 6–8, 2022), <https://advances.umw.edu.pl/en/abstract-book-2022-euripa>

Received on October 14, 2022

Reviewed on November 16, 2022

Accepted on December 22, 2022

Published online on January 5, 2023

## Abstract

Modern medicine is characterized by introducing new innovative medications specially designed after understanding the mechanisms of diseases. Such process revolutionized the management of many diseases, but is usually more relevant to secondary and tertiary care. However, in parallel, we are observing an emerging wave of new technologies, devices and applications with particular relevance to primary care and significant implications for rural and remote areas. The approval processes by regulatory authorities all around the world are more flexible and less demanding in comparison to the approval of new medications. This process may lead to innovative treatments where the balance between benefit and harm is not clear and well documented in the conventional way of prospective randomized clinical trials. On the other hand, these technologies are adopted by the patients, in the case of free-of-charge applications or “on the shelf” devices, or by authorities, for example in the case of remote and telemedicine consultations.

In such a way, disciplines that ignore these technologies or are too cautious and slow in adaptation can find themselves trailing behind. Family medicine and especially remote and rural medicine can benefit a lot. We should be on the frontline in the considerate adoption of relevant technologies, and in this way improve our patients' health and keep family medicine attractive to the young generation of physicians. This editorial is aimed to serve as a window to this new era with the presentation of some of these technologies.

**Key words:** C-reactive protein, point-of-care ultrasound, point-of-care laboratory tests, medical applications, rural medicine

## Cite as

Vinker S. Innovations in family medicine and the implication to rural and remote primary care.

*Adv Clin Exp Med.* 2023;32(2):147–150.

doi:10.17219/acem/158171

## DOI

10.17219/acem/158171

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0)

(<https://creativecommons.org/licenses/by/3.0/>)

## Point of care ultrasound

Point of care ultrasound (POCUS) is an example of a technology that had been adopted by other disciplines as complementary to or replacing the physical examination. In a meta-analysis evaluating 34 studies, the overall pooled sensitivity for lung auscultation was 37% and specificity was 89%. Likelihood ratios (LRs), and area under the curve (AUC) of auscultation for congestive heart failure, pneumonia and obstructive lung diseases were low.<sup>1</sup> The authors summarize “Lung auscultation has a low sensitivity in different clinical settings and patient populations, thereby hampering its clinical utility. When better diagnostic modalities are available, they should replace lung auscultation. Only in resource-limited settings, with a high prevalence of disease and in experienced hands, lung auscultation has still a role”. The editor of *Chest* had an opposite point of view and in an editorial, he answers the question (Should point-of-care ultrasound examination be routine practice in the evaluation of the acutely breathless patient?) by a clear “No”.<sup>2</sup> These 2 different points of view may represent different clinical scenarios, case-mix and interpretations of the scientific evidence, but clearly represent the need to discuss the place of the stethoscope compared to POCUS in evaluating patients with respiratory symptoms in primary care.

Point-of-care ultrasound will not replace the specialist in imaging. It is complementary to the traditional physical examination and maybe even superior in specific cases. For example, in the case of hepatomegaly, we can better estimate its size and cause using POCUS. On suspicion of pneumonia, we can make the diagnosis faster and without using radiation. Using POCUS will improve the diagnosis and may even save time during the visit.

In another systematic review, the authors evaluated the use of POCUS by family physicians.<sup>3</sup> They included a total of 51 full-text articles. The POCUS was applied for a variety of purposes, with the majority of scans focused on abdominal and obstetric indications. Focused POCUS scans were reported to have higher diagnostic accuracy and be associated with less harm than more comprehensive scans or screening scans. In the cited systematic review, the length of a focused POCUS procedure and its accuracy had been evaluated – and for example, it would take 5–12 min to have an abdominal POCUS with a specificity of 99% and sensitivity of 82% while evaluating the kidneys.

In conclusion, POCUS has the potential to be an important tool for the family physician and may possibly reduce healthcare costs, especially in remote and rural areas. It is time for the wide implementation of POCUS in family medicine. The plan should be multidimensional, focusing on training, reimbursement, and continuous evaluation and research.

## Point-of-care C-reactive protein test

Several systematic reviews evaluated the place of point-of-care laboratory tests in primary care.<sup>4–6</sup> There are many such tests relevant in primary care. Their utilization is variable according to availability, costs, reimbursement plans, alternative traditional laboratory tests availability, the complexity of performance, and accuracy, among many other considerations. They may be especially useful in remote and rural areas where the availability of laboratories is lower, and on-the-spot results may narrow the differential diagnosis and spare the need for referral to the emergency room or other specialists that may request costly and complicated evacuation.

One such attractive test is the point-of-care C-reactive protein (CRP) test. C-reactive protein point-of-care testing (POCT) may reduce diagnostic uncertainty and enhance antibiotic stewardship. In primary care, respiratory tract infections (RTIs) are the most common reason for inappropriate antibiotic prescribing, which is a major driver for antibiotic resistance.

The impact of CRP-POCT on antibiotic prescribing for RTIs in primary care has been evaluated in a systematic review.<sup>7</sup> Thirteen studies comprising 9844 participants met the inclusion criteria. Meta-analyses showed that CRP-POCT significantly reduced immediate antibiotic prescribing at the index consultation compared with usual care (risk ratio (RR): 0.79, 95% confidence interval (95% CI): 0.70–0.90,  $p = 0.0003$ ,  $I^2 = 76\%$ ) but not during 28-day follow-up. The immediate effect was sustained at 12 months. In children, CRP-POCT reduced antibiotic prescribing when CRP (cutoff) guidance was provided. Meta-analyses showed significantly higher rates of re-consultation within 30 days. Clinical recovery, resolution of symptoms and number of hospital admissions were not significantly different between CRP-POCT and usual care, which points to the safety of the test. The authors concluded that CRP-POCT can reduce immediate antibiotic prescribing for RTIs in primary care (number needed to (NNT) for benefit = 8) at the expense of increased re-consultations (NNT for harm = 27).

In conclusion, CRP-POCT is highly available at a reasonable cost and should be adopted in primary care. Other useful tests that should be considered for widespread use are troponin, D-dimer and specific tests for various infectious diseases (viral, bacterial and fungal).

## Telemedicine and teleconsultations

A quick PubMed search will yield more than 50,000 publications about telemedicine, and narrowing the search to primary care and family medicine will yield around 10,000 papers. Further narrowing the search to rural and

remote settings will give us around 1000 publications, most of them from recent years, especially published during the COVID-19 pandemic.

Petrazzuoli et al. concluded that telemedicine can be used as an effective tool among disadvantaged populations and may reduce inequities, but care must be taken to avoid it as a tool to cut or replace services in rural and remote areas.<sup>8</sup> This had been stated at the World Health Organization (WHO) Europe 70<sup>th</sup> Regional Meeting by World Organization of National Colleges, Academies and Academic Associations of General Practitioners/Family Physicians (WONCA) Europe and European Rural and Isolated Practitioners Association (EURIPA). It seems that phone consultations gained more popularity in Europe than video consultations.<sup>9</sup> It is beyond the scope of this editorial to discuss the benefits, challenges and obstacles of telemedicine in remote and rural settings, and it is time to perform a systematic review and evaluation to give us insights into the post-pandemic era.

## Software solutions for remote measurement of physiological parameters

In the past, we used laboratories to evaluate physiological parameters. But for decades we have been moving to portable devices that are easy to wear and use, and the evaluations are being performed in ambulatory and primary care settings. A few examples are portable 24-hour blood pressure monitoring, loop recorders to detect rhythm abnormalities, and sleep tests for the diagnosis of sleep apnea, among many others.

New wearable devices can detect rhythm abnormalities, and their accuracy had been evaluated in a systematic review.<sup>10</sup> Nine observational studies (n = 1581) assessing the sensitivity and specificity of wrist-worn wearables in detecting AF in patients with and without a history of AF were included. In patients with a history of AF, the overall sensitivity was high. Specificity significantly differed between the devices. Wrist-worn wearable devices demonstrate promising results in detecting AF in patients with paroxysmal AF. However, more rigorous prospective data are needed to understand the limitations of these devices in regard to varying specificities that may lead to unintended overdiagnosis and overtreatment.

Other wearable devices are used to measure steps, energy expenditure and heart rate, and their accuracy had been evaluated in 2 systematic reviews.<sup>11,12</sup> They concluded that commercial wearable devices are accurate for measuring steps and heart rate in laboratory-based settings, but this varies by manufacturer and device type.<sup>11</sup> However, none of the tested devices proved to be accurate in measuring energy expenditure.<sup>12</sup>

There are new and promising technologies that use artificial intelligence (AI)-powered, video-based, 100% software

solution that easily integrates into any app or workflow to allow measurement of a wide range of physiologic parameters, using only the end user's smartphone, tablet or laptop.<sup>13</sup> With this technology, blood pressure, O<sub>2</sub> saturation, respiratory rate, and heart rate can be assessed. After regulatory approval, using these technologies can be easily applicable in remote and rural areas for distance monitoring of patients and in aiding the clinical decision-making about treatment, management and referrals for emergency room or specialist consultation.

## Conclusion

In this short editorial, I gave a few examples of new and innovative technologies that should be considered for widespread use in primary care and especially in remote and rural settings. Rural primary care should not take it as a threat, but rather as an opportunity to introduce advanced modalities of care without the need for secondary and tertiary expensive and hard-to-attend healthcare facilities. It can be used to empower family physicians and primary care providers and to reduce the inequity between the peripheral and central areas all around Europe. Each of these new technologies deserves deep discussions and research about their utility and effectiveness in primary care, as well as the ways to introduce them, teach the patients how to use them, and develop mechanisms of the reimbursement of the time and costs spent on the specific technology. However, these considerations should not delay the implementation – otherwise, the outlined solutions will be adopted, for example, as home-based tests with remote or application-derived interpretation, and such implementation will bypass the family physician.

### ORCID iDs

Shlomo Vinker  <https://orcid.org/0000-0001-9804-7103>

### References

1. Arts L, Lim EHT, van de Ven PM, Heunks L, Tuinman PR. The diagnostic accuracy of lung auscultation in adult patients with acute pulmonary pathologies: A meta-analysis. *Sci Rep*. 2020;10(1):7347. doi:10.1038/s41598-020-64405-6
2. Corcoran JP, Laursen CB. Counterpoint: Should point-of-care ultrasound examination be routine practice in the evaluation of the acutely breathless patient? No. *Chest*. 2019;156(3):426–428. doi:10.1016/j.chest.2019.04.119
3. Andersen CA, Holden S, Vela J, Rathleff MS, Jensen MB. Point-of-care ultrasound in general practice: A systematic review. *Ann Fam Med*. 2019;17(1):61–69. doi:10.1370/afm.2330
4. Goyder C, Tan PS, Verbakel J, et al. Impact of point-of-care panel tests in ambulatory care: A systematic review and meta-analysis. *BMJ Open*. 2020;10(2):e032132. doi:10.1136/bmjopen-2019-032132
5. Lingervelder D, Koffijberg H, Kusters R, IJzerman MJ. Point-of-care testing in primary care: A systematic review on implementation aspects addressed in test evaluations. *Int J Clin Pract*. 2019;73(10):e13392. doi:10.1111/ijcp.13392
6. Lingervelder D, Koffijberg H, Kusters R, IJzerman MJ. Health economic evidence of point-of-care testing: A systematic review. *Pharmacoeconomics Open*. 2021;5(2):157–173. doi:10.1007/s41669-020-00248-1

7. Martínez-González NA, Keizer E, Plate A, et al. Point-of-care C-reactive protein testing to reduce antibiotic prescribing for respiratory tract infections in primary care: Systematic review and meta-analysis of randomised controlled trials. *Antibiotics (Basel)*. 2020;9(9):610. doi:10.3390/antibiotics9090610
8. Petrazzuoli F, Kurpas D, Vinker S, et al. COVID-19 pandemic and the great impulse to telemedicine: The basis of the WONCA Europe Statement on Telemedicine at the WHO Europe 70<sup>th</sup> Regional Meeting September 2020. *Prim Health Care Res Dev*. 2021;22:e80. doi:10.1017/S1463423621000633
9. Petrazzuoli F, Gokdemir O, Antonopoulou M, et al. Patient consultations during SARS-CoV-2 pandemic: A mixed-method cross-sectional study in 16 European countries. *Remote Rural Health*. 2022;22(4):7196. doi:10.22605/RRH7196
10. Belani S, Wahood W, Hardigan P, Placzek AN, Ely S. Accuracy of detecting atrial fibrillation: A systematic review and meta-analysis of wrist-worn wearable technology. *Cureus*. 2021;13(12):e20362. doi:10.7759/cureus.20362
11. Fuller D, Colwell E, Low J, et al. Reliability and validity of commercially available wearable devices for measuring steps, energy expenditure, and heart rate: Systematic review. *JMIR Mhealth Uhealth*. 2020;8(9):e18694. doi:10.2196/18694
12. Germini F, Noronha N, Borg Debono V, et al. Accuracy and acceptability of wrist-wearable activity-tracking devices: Systematic review of the literature. *J Med Internet Res*. 2022;24(1):e30791. doi:10.2196/30791
13. Binah.ai Health Data Platform. <https://www.binah.ai/>. Accessed October 14, 2022.



# The efficacy and safety of St. John's wort extract in depression therapy compared to SSRIs in adults: A meta-analysis of randomized clinical trials

\*Xin Zhao<sup>1,A</sup>, \*Hong Zhang<sup>2,B</sup>, Yanyan Wu<sup>3,C</sup>, Chun Yu<sup>4,D–F</sup>

<sup>1</sup> Department of Rehabilitation Medicine, No. 215 Hospital of Shaanxi Nuclear Industry, Xianyang, China

<sup>2</sup> Department of Geriatrics, Meishan Hospital of Traditional Chinese Medicine, China

<sup>3</sup> Department of Pharmacy, Women and Children's Hospital of Chongqing Medical University, Chongqing Health Center for Women and Children, China

<sup>4</sup> Department of Internal Medicine, Xi'an Traditional Chinese Medicine Hospital of Encephalopathy, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;

D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):151–161

## Address for correspondence

Chun Yu

E-mail: yuruyi661858@gmail.com

## Funding sources

None declared

## Conflict of interest

None declared

\*Xin Zhao and Hong Zhang contributed equally to this work.

Received on January 6, 2022

Reviewed on July 12, 2022

Accepted on August 23, 2022

Published online on October 11, 2022

## Cite as

Zhao X, Zhang H, Wu Y, Yu C. The efficacy and safety of St. John's wort extract in depression therapy compared to SSRIs in adults: A meta-analysis of randomized clinical trials. *Adv Clin Exp Med.* 2023;32(2):151–161. doi:10.17219/acem/152942

## DOI

10.17219/acem/152942

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Depression is the most common mental disorder, affecting about 3.8% of the population worldwide. Clinical symptoms of depression include sadness, anxiety and frequent mood swings, among others. Selective serotonin reuptake inhibitors (SSRIs), psychotherapy and behavioral therapy are commonly used for the treatment of this condition. Since SSRIs are associated with various side effects, extract of St. John's wort (SJW) has been suggested as an effective alternative. However, there are conflicting studies regarding its efficacy. Many studies have reported positive outcomes with low adverse effects, while others did not find it to be a suitable alternative.

**Objectives.** To analyze the available studies using SJW for depression therapy and to thoroughly evaluate its effectiveness compared to SSRIs and placebo.

**Materials and methods.** Relevant articles for our meta-analysis were found using Medline (via PubMed), Cinahl (via EBSCO), Scopus, and Web of Sciences databases. Studies were included as per the predefined Population, Intervention, Comparison, Outcomes and Study (PICOS) criteria. A demographic summary of the patients treated with either SJW, placebo or SSRIs was collected and Hamilton Depression Rating Scale (HAMD) scores were extracted. Risks of bias analysis, diagnostic odds ratio (OR), risk ratio (RR), and sensitivity calculation were evaluated using Revman software, and the publication bias was assessed using MedCalc software.

**Results.** Fourteen clinical trials with a total of 2270 depression patients were included in accordance with the inclusion criteria. All analyzed papers were published between 2000 and 2022. For patients treated with either SSRIs or SJW, a pooled OR of 2.44 with a 95% confidence interval (95% CI) of 1.33–4.45 was obtained. The data were heterogeneous, with a  $\tau^2$  value of 0.54,  $\chi^2$  value of 31.05, degrees of freedom (df) value of 7,  $I^2$  value of 77%, and an overall Z-value of 2.90 with  $p = 0.004$ .

**Conclusions.** Our research supports the use of SJW as it reduced the number of depressive patients and their HAMD scores while having fewer risks and side effects than conventional medications.

**Key words:** mental disorder, SSRIs, depression, randomized controlled trials, St. John's wort

## Introduction

Technology is developing day by day as is the lifestyle. To keep pace with the ongoing developments, people are involved in a lot of activities, which in turn causes anxiety, stress and various mental disorders. Currently, depression is the most common mental illness worldwide in people of all age groups. The clinical symptoms associated with depression are frequent mood swings, persistent sadness, anxiety, low appetite, irritation, sleeplessness, and loss of interest and pleasure in all activities, among others.<sup>1,2</sup> Major depressive disorders arise due to frontal lobe dysfunction out of threat and fear,<sup>3,4</sup> alterations in neuronal networks, and other cognitive impairments.<sup>5–9</sup> Medications, psychotherapy and behavioral therapy are the main treatment strategies for treating depression. The primary medicines used to treat depression are selective serotonin reuptake inhibitors (SSRIs) such as fluoxetine, paroxetine, sertraline, etc., as shown in Fig. 1, which outlines depression, its causes and medications used. Along with these, other drugs such as serotonin-noradrenaline reuptake inhibitor – noradrenaline and specific serotonergic antidepressants (SNRI–NaSSA) – can be used. However, these medications are associated with various adverse side effects like dizziness, indigestion, diarrhea, blurred vision, dry mouth, and others.<sup>10,11</sup> Therefore, for the effective treatment of depression with minimum side effects, various studies have reported the use of St. John's wort (SJW) extract.

Perforate SJW, also known as *Hypericum perforatum*, is a flowering plant from the Hypericaceae family; its extract is used as a herbal remedy for various mental illnesses like

mild to moderate depression, obsessive-compulsive disorder, insomnia, etc. It is also an antioxidant with marked antiviral and antibacterial properties. St. John's wort is reported to reduce neuralgia, anxiety and stress by regulating neurotransmitters like serotonin, gamma-aminobutyric acid (GABA), dopamine, and others in the brain. Therefore, it has been used in the treatment of many neurological issues.

Various research articles and randomized controlled trials have reported SJW to effectively reduce the Hamilton Depression Rating Scale (HAMD) score and clinical symptoms in depressed patients.<sup>12–25</sup> For example, in a meta-analysis conducted by Linde and Mulrow, a total of 29 trials including 5489 patients were analyzed and the conclusion was that SJW is a safe medication for depression, superior when compared to existing SSRIs.<sup>26</sup> Similarly, Apaydin et al.,<sup>27</sup> Hübner and Kirste<sup>28</sup> and Behnke et al.<sup>29</sup> concluded that SJW used for depression is safe and more effective than placebo, with mild side effects.

However, the evidence regarding St. John's wort efficacy for the treatment of depression with fewer adverse effects is still insufficient since many studies, like the case study by Ferrara et al.<sup>30</sup> and Arold et al.,<sup>31</sup> did not find it perfectly safe and reported risk associated with its use in the treatment of depression. The contradictions regarding St. John's wort's safety and efficacy arise due to the limited amount of available case reports, use of unblinded assessments, and lack of refined analysis and evidence-level grade assessment. Therefore, the present study analyzes existing studies related to the use of SJW for depression, and summarizes the available literature in terms of efficacy and safety of SJW use to treat depression.

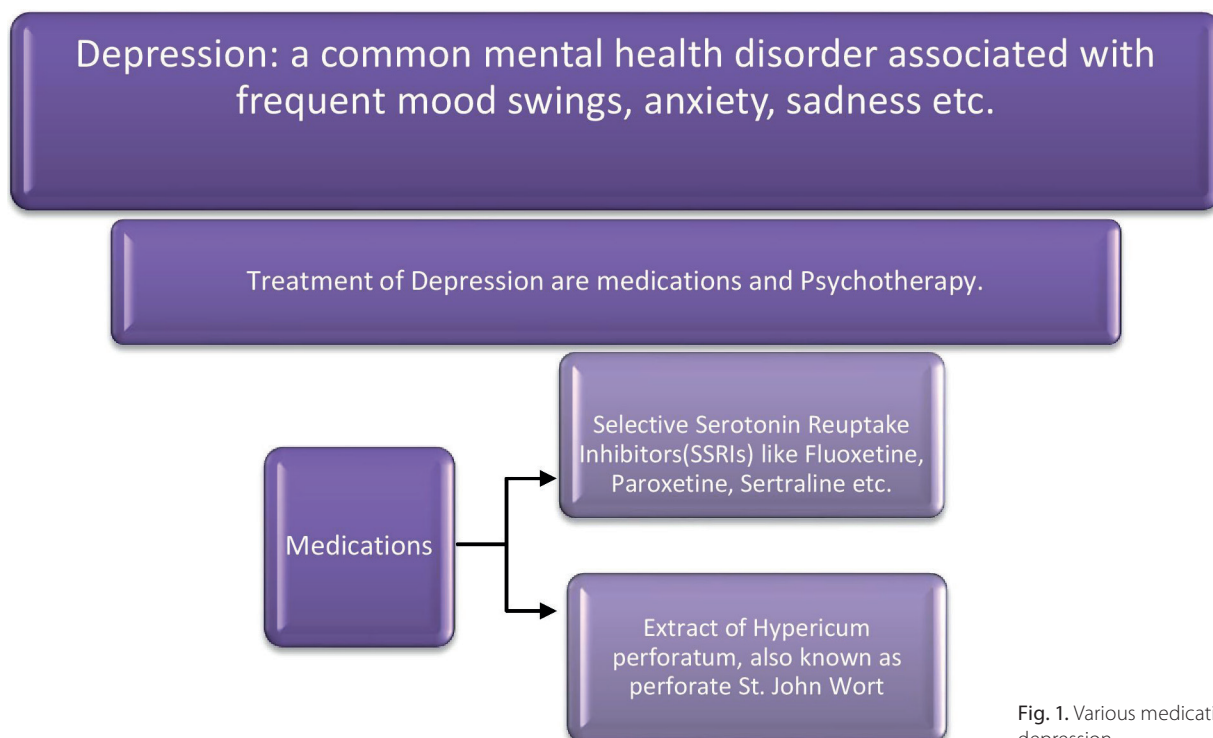


Fig. 1. Various medications used for depression

## Objectives

The goal of this study was to analyze the efficacy of SJW extract in depression compared to SSRIs in adults.

## Materials and methods

We followed the normative recommendations in the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) with a registration No. XU#/IRB/2021/1125.

### Search strategy

This meta-analysis is based on an extensive search in Medline (via PubMed), Cinahl (via EBSCO), Scopus, and Web of Science databases for relevant studies published between 2000 and 2022. The following keywords were used to search for relevant studies: depression, neurological disorder, neurotransmitters, SSRIs, randomized controlled trials, anxiety, and HAMD score. All the included articles were selected as required by the PRISMA guidelines and studies were selected randomly irrespective of language, publication status or type of study (prospective, retrospective, clinical trial). A demographic summary of the patients and event data was extracted from the included studies.

Two authors (XZ and HZ) independently scanned the relevant sources for related studies. Mainly full-text articles were collected, and abstracts were used only if they included sufficient information for the meta-analysis. Obsolete references were excluded and valuable studies were included according to the inclusion criteria. In addition, 2 researchers (YW and CY) independently collected event data on useful variables.

### Inclusion and exclusion criteria

We included studies that reported on the safety and efficacy of SJW in the treatment of adult patients with mild to moderate depression. Studies were selected from the year 2000 to 2022. Studies with insufficient data, reporting on the use of medications other than SSRIs and SJW, and published before 2000 were excluded.

### Evaluation of the analytical standard

Two reviewers (XZ and HZ) independently evaluated the methodological validity of each included study and calculated the heterogeneity of the studies, while author CY was responsible for resolving any type of disagreement between XZ and HZ. To investigate the heterogeneity, a Deek's funnel plot, Cochran's Q statistic and I<sup>2</sup> index in random bivariate mode were calculated using RevMan software v. 5.0 (<https://training.cochrane.org/online-learning/core-software/revman/>).

## Sources of heterogeneity and risk of bias assessment

The heterogeneity sources investigated included the use of full-text publications compared to abstracts, varied age groups, different numbers of patients, variable durations of treatment, different scales of analysis, and comparison of SJW to placebo and different SSRI medicines.

The risk of bias assessment for the included studies was performed, and the corresponding risk of bias summary and graph were created using RevMan v. 5.0.<sup>32</sup>

### Statistical analyses

The meta-analysis was performed using extracted data, and statistical parameters such as diagnostic odds ratios (ORs), relative risks with a 95% confidence interval (95% CI), and sensitivities were calculated using the Mantel–Haenszel method with random bivariate effects using RevMan v. 5.0, along with their respective forest plots.<sup>33</sup> Meta-analyses were performed using a random-effects model (Mantel–Haenszel method), and the heterogeneity of the included studies was evaluated using a tau<sup>2</sup> value,  $\chi^2$  value, I<sup>2</sup> value, and Z-value. A value of  $p < 0.05$  was considered statistically significant. Diagnostic ORs were calculated using the DerSimonian–Laird technique. For this, a 2×2 table was made and a meta-analysis was performed using RevMan software v. 5.0. A pooled diagnostic OR with a 95% CI was calculated and respective forest plots were created. The publication bias of the included studies was assessed using Begg's and Egger's tests,<sup>34</sup> and a funnel plot was prepared by plotting the log risk ratio (RR) of each study against its standard error using MedCalc software v. 20.115 (MedCalc Software Ltd., Ostend, Belgium).

## Results

### Literature search results

We found a total of 1135 studies through electronic searches of different databases. We excluded 148 studies after reading their titles and abstracts, leaving 987 records to be screened. Furthermore, due to invalid references and duplications, we excluded 881 studies, and included only 106 studies for the final screening. Out of these, 75 studies were excluded based on our inclusion criteria, and the eligibility of the remaining 31 was further assessed. The key reasons for exclusion included inadequate evidence and comparison criteria inappropriate to create 2×2 tables for review. The 14 studies that fulfilled the inclusion criteria (i.e., the use of SJW compared to SSRI) and were chosen for this meta-analysis are shown in Fig. 2.

The demographic details of the included studies are shown in Table 1. They include the author of the study, year of publication, duration of the study, total sample size,

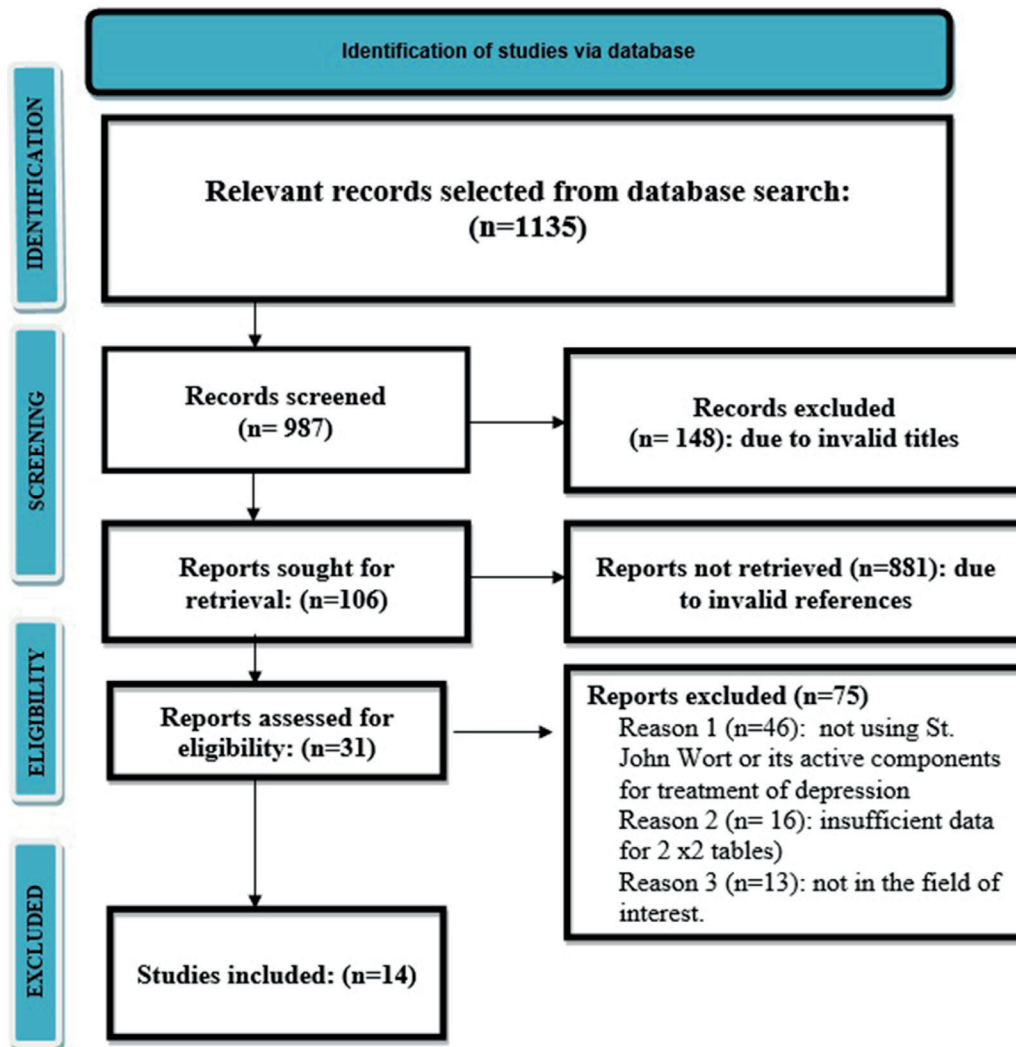


Fig. 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of the study group

type of SSRI drug used as the control, dose of the drugs used, age and gender of the patients, number of patients with positive outcomes, the initial HAMD score of patients, and the mean value and standard deviation ( $M \pm SD$ ) of improved HAMD scores, with their respective p-values for statistical significance.

Fourteen clinical trials with a total of 2270 patients suffering from depression were included according to the inclusion criteria. The included studies evaluated adult patients from different age groups who were randomly chosen and treated with either a conventional SSRI like fluoxetine and sertraline, or placebo, and SJW. In both cases, the number of patients with a reduction in their HAMD scores and a marked improvement in clinical symptoms were extracted as event data and analyzed statistically.

### Risk of bias assessment

The risk of bias assessment for the included studies is reported in Table 2. The risk of bias was low as evident from the corresponding risk of bias summary shown in Fig. 3 and the risk of bias graph presented as Fig. 4.

## Meta-analysis results

### Results for improvement in HAMD score of patients using either SJW or placebo

The risk of publication bias in the included studies related to the treatment with either SJW or placebo was low, as shown in the funnel plot in Fig. 5. We obtained the value of  $p = 0.85$  for Begg's test and  $p = 0.68$  for Egger's test, which confirms a low risk of publication bias.

We calculated a pooled OR of 0.72 with a 95% CI of 0.64–0.81, as shown in the forest plot in Fig. 6. The data were heterogeneous with a  $\tau^2$  value of 0.36, a  $\chi^2$  value of 17.93, a degrees of freedom (df) value of 5, an  $I^2$  value of 72%, and an overall Z-value of 2.60 with  $p = 0.009$ . We obtained a pooled RR of 0.46 with a 95% CI of 0.26–0.83, as shown in the forest plot in Fig. 7. The data were heterogeneous, with a  $\chi^2$  value of 5.34, df value of 5, and an overall Z-value of 5.36 with  $p = 0.00001$ . All of these results were statistically significant with  $p < 0.05$ .

Table 1. Demographic summary of the included studies

Study and year	Sample size, n	Age [years]	Gender (F/M)	Duration [weeks]	Control drug		Test drug (SJW)			
					dose	positive response in patients with test drug	SJW extract dose	positive response in patients with control drugs		
Hypericum depression trial study group 2002 [12]	340 mean HAMID ≥ 20	above 18 years	147/77	8	sertraline 50–100 mg/day	n = 79/111 M ±SD = 10.53 ±0.72	900–1500 mg	n = 82/113 M ±SD = 8.68 ±6.8	p = 0.21	p = 0.26
Lecrubier et al. 2002 [17]	375 mean HAMID ≥ 25	18–65	287/88	6	placebo	n = 80/189 M ±SD = 8.5 ±7.7	300 mg/day	n = 98/189 M ±SD = 10.5 ±7	p = 0.04	p = 0.237
Van Gorp et al. 2002 [14]	87 mean HAMID ≥ 16	18–65	52/33	12	sertraline 50–100 mg/day	n = 28/43 M ±SD = 11.5 ±8.4	900–1800 mg/day	n = 30/43 M ±SD = 9.4 ±8.3	p = 0.237	p = 0.237
Moreno et al. 2005 [18]	53 mean HAMID ≥ 20	19–64	44/9	8	fluoxetine 20 mg/day	n = 3/16 mean HAMID ≤ 10	900 mg/day	n = 6/18 mean HAMID ≤ 10	p = 0.3794	p = 0.0167
Fava et al. 2005 [13]	135 mean HAMID = 19.7 ±3.2	18–65	77/58	12	placebo 20 mg/day	n = 9/43 M ±SD = 12.6 ±6.4	900 mg/day	n = 17/43 M ±SD = 10.2 ±6.6	p = 0.096	p < 0.03
Bjerkenstedt et al. 2005 [12]	163 mean HAMID = 30	18–70	129/34	6	placebo 20 mg/day	n = 21/55 M ±SD = 15 ±8.4	900 mg/day	n = 22/55 M ±SD = 15.5 ±6.7	p = 0.7	p = 0.7
Szegedi et al. 2005 [24]	251 mean HAMID = 22	20–60	168/83	6	paroxetine 20 mg	n = 73/125 M ±SD = 11.4 ±8.6	900 mg/day	n = 86/125 M ±SD = 14.4 ±8.8	p = 0.01	p = 0.01
Simeon et al. 2005 [22]	26 mean HAMID = 18	12–17	14/12	8	sertraline 50 mg/day	n = 5/26 M ±SD = 10.8 ±5.3	900 mg/day	n = 21/26 M ±SD = 10.8 ±5.3	p = 0.016	p = 0.016
Kasper et al. 2006 [16]	332 mean HAMID = 18	18–65	205/127	6	placebo	n = 74/119 M ±SD = 17.6 ±8.8	600 mg/day	n = 111/119 M ±SD = 11.8 ±8.3	p < 0.001	p < 0.001
Ur-Rahman et al. 2008 [19]	112 mean HAMID = 18	18–65	87/25	6	placebo 300 mg thrice daily	n = 23/56 M ±SD = 10.04 ±0.8	300 mg thrice daily	n = 35/56 M ±SD = 8.04 ±1.0	p = 0.4	p = 0.4
Weber et al. 2008 [25]	54 mean HAMID = 18	6–17	20/34	8	placebo	n = 12/27 mean HAMID ≤ 10	300 mg/day	n = 14/27 mean HAMID ≤ 10	p = 0.89	p = 0.68
Singer et al. 2011 [23]	154 mean HAMID ≥ 20	18–74	73/35	6	20 mg citalopram	n = 24/54 mean HAMID ≤ 10	900 mg/day	n = 26/54 mean HAMID ≤ 10	p = 0.0288	p = 0.0288
Sarris et al. 2012 [20]	124 mean HAMID = 18	18–60	81/43	26	sertraline 50–100 mg/day	n = 64/124 M ±SD = 7.1 ±5.4	900–1500 mg/day	n = 105/124 M ±SD = 6.6 ±4.5	p = 0.036	p = 0.03
Seifritz et al. 2016 [21]	64 mean HAMID = 25	18–70	31/32	6	paroxetine 20 mg/day	n = 8/33 M ±SD = 22.9 ±0.8	3 × 300 mg/day	n = 20/31 M ±SD = 23.1 ±0.9	p = 0.248	p = 0.248

SJW – St. John's wort; HAMID – Hamilton Depression Rating Scale; M ±SD – mean value ± standard deviation.

Table 2. Risk assessment for included studies

Question	Weber et al. 2008 [25]	Szegedi et al. 2005 [24]	Singer et al. 2011 [23]	Simeon et al. 2005 [22]	Seifritz et al. 2016 [21]	Sarris et al. 2012 [20]	Ur-Rahman et al. 2008 [19]	Moreno et al. 2005 [18]	Lecrubier et al. 2002 [17]	Kasper et al. 2006 [16]	Hypericum Depression Trial (HTD) group 2002 [15]	Van Gulp et al. 2002 [14]	Fava et al. 2005 [13]	Bjerkenstedt et al. 2005 [12]
Was a consecutive or random sample of patients enrolled?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Did the study avoid inappropriate exclusions?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Did all patients receive the same reference standard?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were all patients included in the analysis?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the sample frame appropriate to address the target population?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were study participants sampled in an appropriate way?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the study subjects and the setting described in detail?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were valid methods used for the identification of the condition?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the condition measured in a standard, reliable way for all participants?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was there appropriate statistical analysis?	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

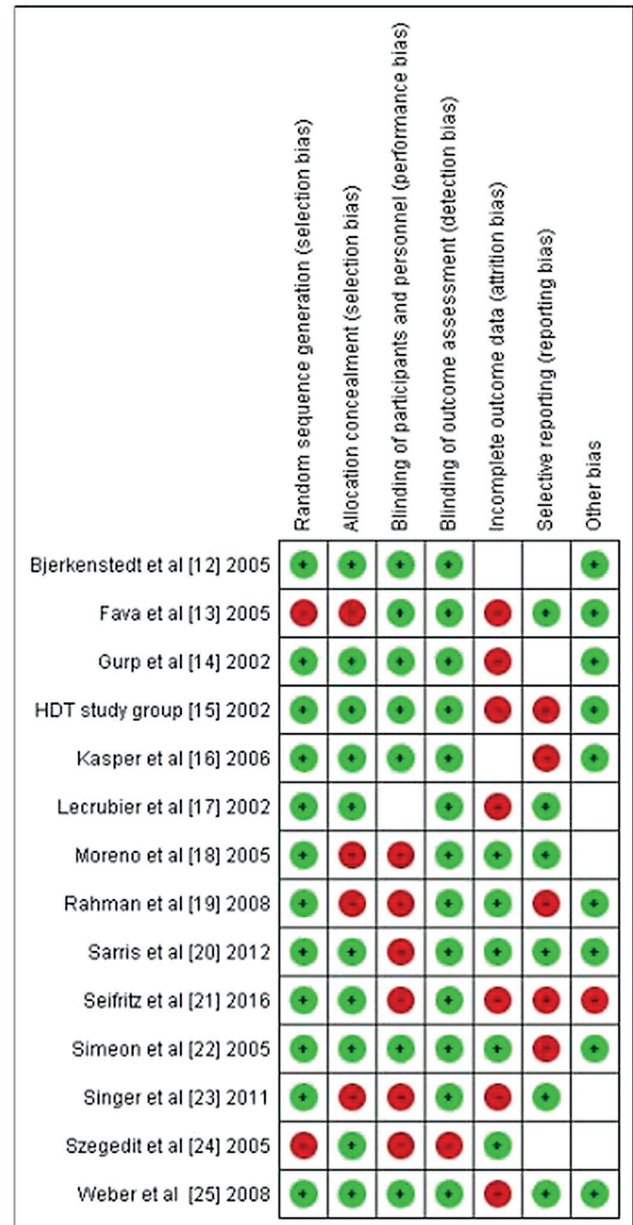


Fig. 3. Risk of bias summary

**Results for improvements in HAMD scores of patients using either SJW or SSRIs**

The risk of publication bias in the included studies related to the treatment with either SJW or SSRIs was low, as shown in the funnel plot in Fig. 8. We obtained the value of  $p = 0.80$  for Begg's test and  $p = 0.76$  for Egger's test, which confirms a low risk of publication bias.

We calculated a pooled OR of 2.44 with a 95% CI of 1.33–4.45, as shown in the forest plot in Fig. 9. The data were heterogeneous, with a  $\tau^2$  value of 0.54, a  $\chi^2$  value of 31.05, a df value of 7, an  $I^2$  value of 77%, and an overall Z-value of 2.90 with  $p = 0.004$ . We obtained a pooled RR of 1.39 with a 95% CI of 1.09–1.76, as shown in the forest plot in Fig. 10. The data were heterogeneous with a  $\tau^2$  value of 0.07, a  $\chi^2$  value of 31.43, a df value of 7, an  $I^2$  value of 78%,

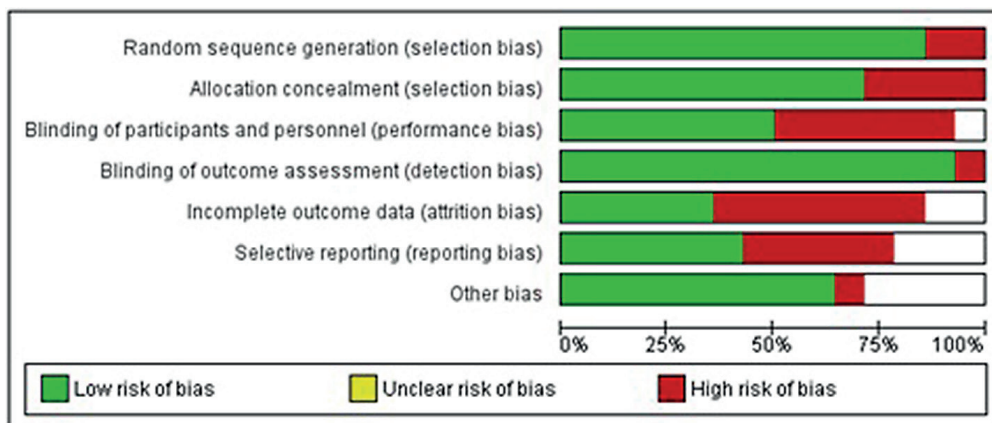
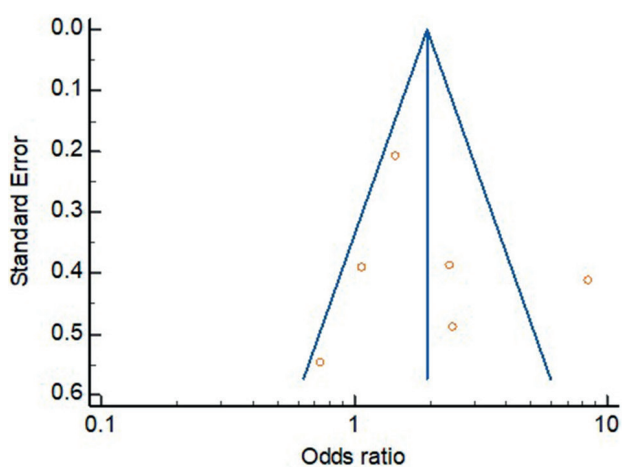


Fig. 4. Risk of bias graph



Publication bias	
Egger's test	
Intercept	1.1607
95% CI	-6.1467 to 8.4681
Significance level	P = 0.6820
Begg's test	
Kendall's Tau	-0.06667
Significance level	P = 0.8510

Fig. 5. Funnel plot of St. John's wort (SJW) compared to placebo

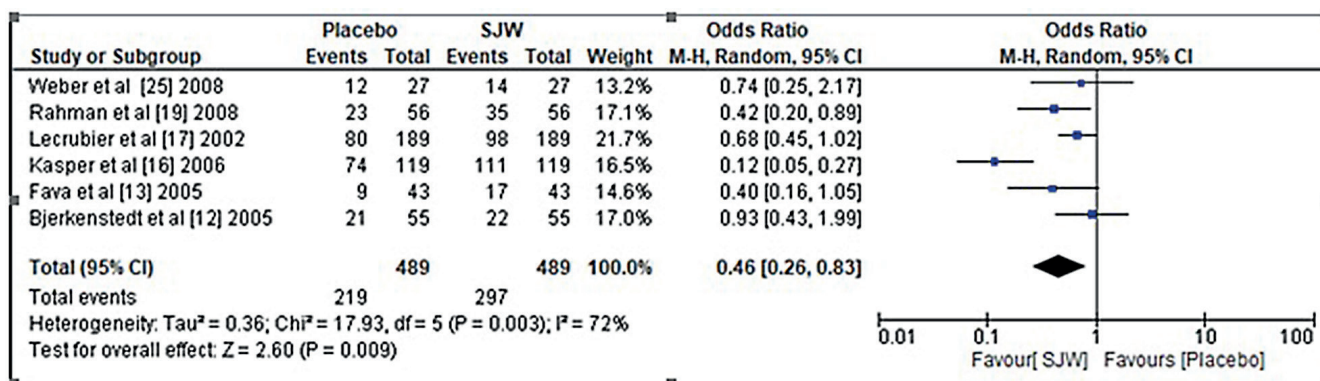


Fig. 6. Forest plot of odds ratio (OR) of St. John's wort (SJW) compared to placebo

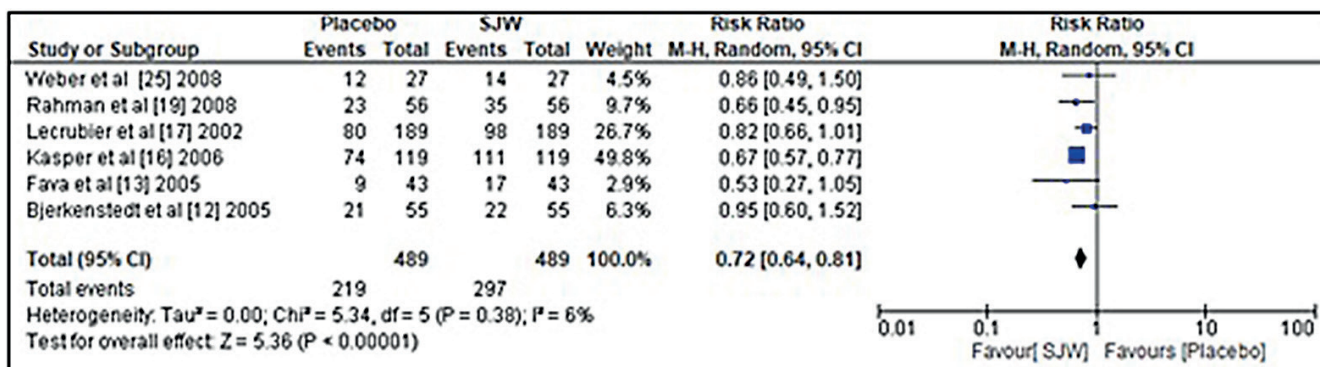


Fig. 7. Forest plot of risk ratio (RR) of St. John's wort (SJW) compared to placebo

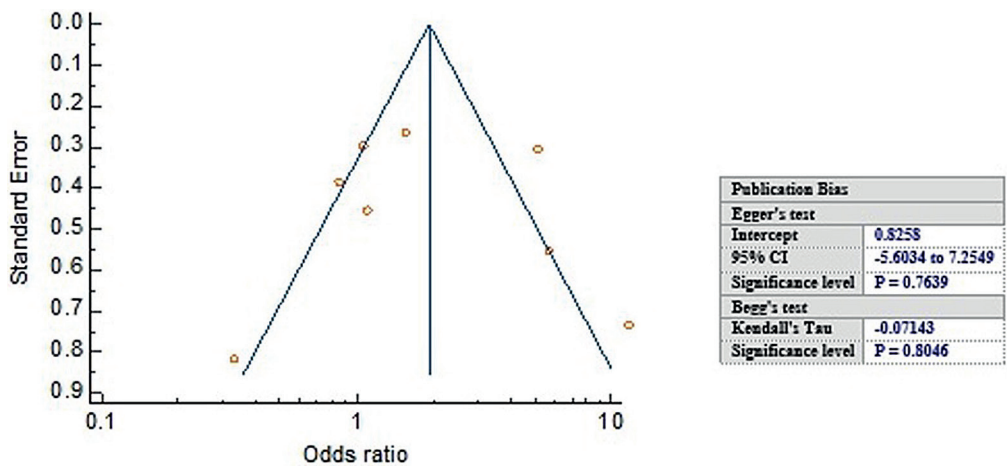


Fig. 8. Funnel plot of St. John's wort (SJW) compared to selective serotonin reuptake inhibitors (SSRIs)

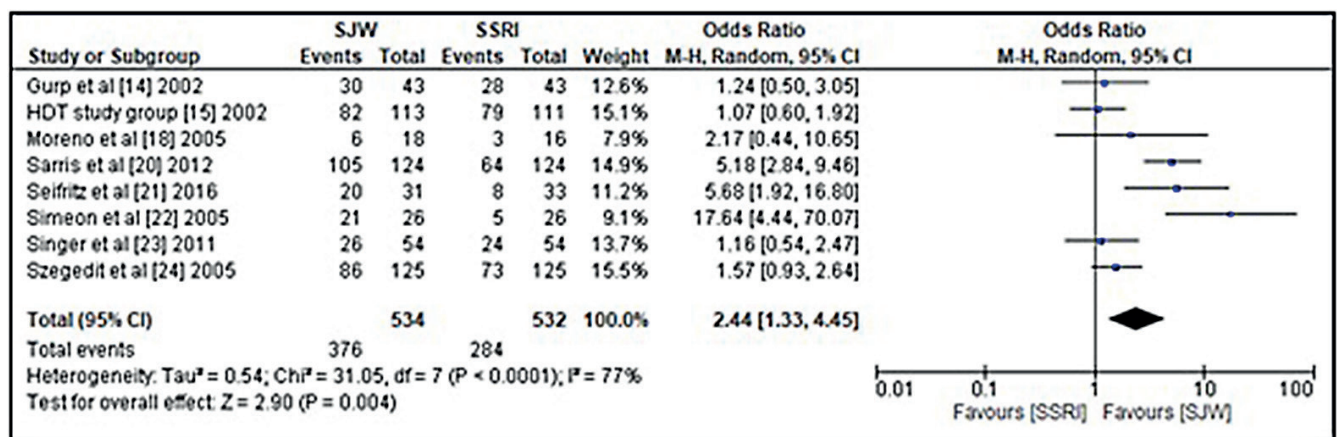


Fig. 9. Forest plot of odds ratio (OR) of St. John's wort (SJW) compared to selective serotonin reuptake inhibitors (SSRIs)

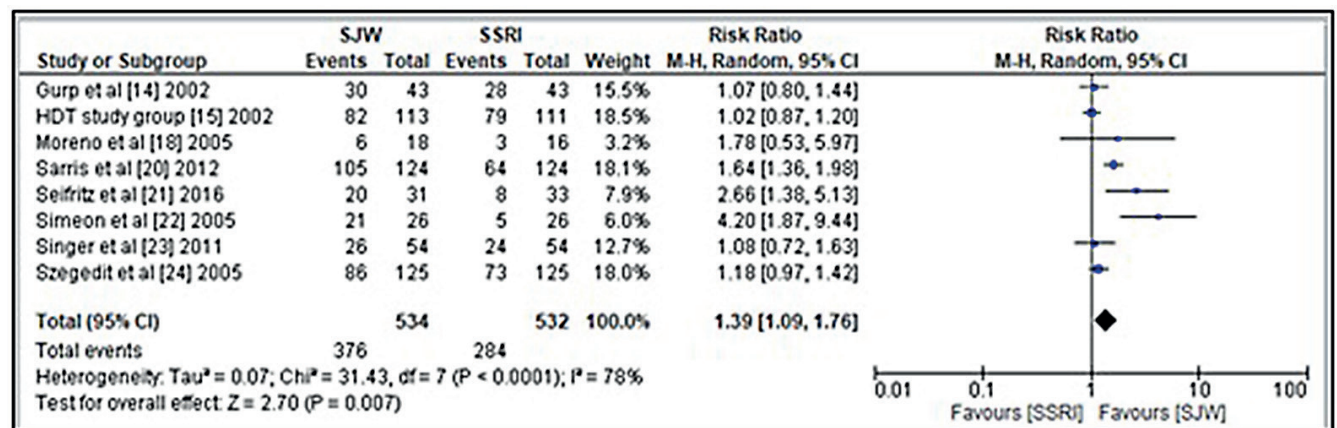


Fig. 10. Forest plot of risk ratio (RR) of St. John's wort (SJW) compared to selective serotonin reuptake inhibitors (SSRIs)

and an overall Z-value of 2.70 with p = 0.007. All of these results were statistically significant with p < 0.05.

The statistical data are summarized in Table 3 and these results showed SJW to be more effective than placebo and conventional SSRIs, with statistically significant sensitivity of 80%. A greater number of patients with reduced HAMD scores was observed in both comparisons of SJW to placebo and SJW to SSRIs. This proves that SJW is a more effective and better alternative for the treatment of depression.

Considering all of these statistically significant results with p < 0.05, this study supports the use of SJW in adults with mild to moderate depression.

## Discussion

Depression is a mental disorder that arises from a hectic and stressful lifestyle. When a person finds themselves



**Table 3.** Statistical summary of the included studies

Study, year, reference	OR (95% CI)	RR (95% CI)	Sensitivity (%)
Lecrubier et al., 2002 [17]	0.68 (0.45–1.02)	0.82 (0.66–1.01)	84.5
Van Gorp et al., 2002 [14]	1.24 (0.50–3.05)	1.07 (0.80–1.44)	93.8
Hypericum Depression Trial (HTD) group 2002 [15]	1.07 (0.60–1.92)	1.02 (0.87–1.20)	96.5
Simeon et al., 2005 [22]	17.64 (4.44–70.07)	4.20 (1.87–9.44)	56.8
Moreno et al., 2005 [18]	2.17 (0.44–10.65)	1.78 (0.53–5.97)	66.7
Fava et al., 2005 [13]	0.40 (0.16–1.05)	0.53 (0.27–1.05)	68.0
Bjerkstedt et al., 2005 [12]	0.93 (0.43–1.99)	0.95 (0.60–1.52)	95.7
Szegedi et al., 2005 [24]	1.57 (0.93–2.64)	1.18 (0.97–1.42)	92.5
Kasper et al., 2006 [16]	0.12 (0.05–0.27)	0.67 (0.57–0.77)	75.0
Weber et al., 2008 [25]	0.74 (0.25–2.17)	0.86 (0.49–1.50)	87.5
Ur-Rahman et al., 2008 [19]	0.42 (0.20–0.89)	0.66 (0.45–0.95)	74.5
Singer et al., 2011 [23]	1.16 (0.54–2.47)	1.08 (0.72–1.63)	92.9
Sarris et al., 2012 [20]	5.18 (2.84–9.46)	1.64 (1.36–1.98)	71.9
Seifritz et al., 2016 [21]	5.68 (1.92–16.80)	2.66 (1.38–5.13)	62.5

OR – odds ratio; 95% CI – 95% confidence interval; RR – risk ratio.

lagging behind others, their mental well-being becomes disturbed and they persistently remain sad and stressed.<sup>35,36</sup> Depression has a significant impact on a patient's social behavior and results in people separating themselves from society and indulging in bizarre thinking. If the problem escalates, it can lead to borderline personality disorder as well contemplating self-harm and suicide.<sup>37–40</sup>

For the treatment of mild to moderate depression, self-care and psychological therapies are initially suggested. If the condition of the patient does not improve, medications may be used. For this purpose, SSRIs like fluoxetine, paroxetine and sertraline are generally used, but their utility is limited by their adverse side effects such as dizziness, indigestion and diarrhea.<sup>41–45</sup> Therefore, to achieve an effective treatment of depression with minimum side effects, various studies reported the use of SJW extract, which is a well-known herbal remedy for various mental illnesses. It selectively regulates neurotransmitters in the brain and is useful in the treatment of neurological issues.

Various studies, including Fava et al.,<sup>13</sup> Van Gorp et al.,<sup>14</sup> Moreno et al.,<sup>18</sup> and Singer et al.,<sup>23</sup> reported that SJW can effectively reduce HAMD scores and clinical symptoms of depression. Linde and Mulrow performed a meta-analysis that included a total of 29 studies with 5489 patients, and concluded that SJW is equally safe in depression patients as compared to existing SSRIs, and superior to SSRIs in terms of safety and minimal adverse effects. Similarly, Apaydin et al.<sup>27</sup> evaluated 35 studies that included 6993 patients for the efficacy of SJW, Behnke et al.<sup>29</sup> performed a double-blind randomized controlled trial with SJW, and Hübner and Kirste<sup>28</sup> studied the effect of SJW in 101 children under 12 years of age. All found SJW to be a safer and more effective medicine than placebo, with mild side effects. Similarly to these studies, we also found a pooled OR of 2.44

with a 95% CI of 1.33–4.45 for SSRIs compared to SJW, and a pooled OR of 0.46 with a 95% CI of 0.26–0.83 for placebo compared to SJW. With these statistically significant ( $p < 0.05$ ) results and a statistically significant sensitivity of 80%, we recommend SJW as an efficient drug for adults with mild to moderate depression.

Unlike our results, case studies by Ferrara et al.<sup>30</sup> and Arold et al.<sup>31</sup> did not find SJW to be a perfectly safe and convincing alternative. They reported risks associated with its use, and suggested checking its dose and side effects more carefully before using it in the treatment of depression. Recent studies have reported that if depression occurs because of changes in neuronal networks or alterations in brain complexes, non-invasive brain surgery is a good treatment option instead of medications.<sup>46–50</sup> Battaglia et al. also reported that understanding the psychophysiology of fear can help in finding new insights for the characterization of mental illnesses.<sup>51</sup>

## Limitations


The limitation of the present study is the variability of control drugs used for the treatment of depression compared to SJW, which could have skewed our results. Similarly, the observation of HAMD score and clinical symptoms using different analytical tests performed by different persons also influences the risk of false negative results. Many studies failing to report on the comparable efficiency of SJW with conventional SSRI drugs may also affect the data to some extent. Data from other relevant studies showing the efficacy of SJW compared to SSRIs could have included more results to suggest its use more precisely. Detailed data from the patient's case history, physical examination and pathological tests can further improve the results supporting SJW as an effective treatment for mild to moderate depression.

## Conclusions

The SSRIs are widely used for the treatment of depression and can significantly alleviate clinical symptoms and lower the HAMD score. However, due to their strong side effects, they are not recommended. Instead, the use of SJW is preferred as it is a cheap, readily available and effective treatment strategy for mild to moderate depression, with mild side effects. Thus, based on the current meta-analysis and statistically significant results ( $p < 0.05$ ), the use of SJW for depression in adults is strongly recommended.

## ORCID iDs

Xin Zhao  <https://orcid.org/0000-0003-4357-9027>

Hong Zhang  <https://orcid.org/0000-0002-6109-8238>

Yanyan Wu  <https://orcid.org/0000-0002-8719-982X>

Chun Yu  <https://orcid.org/0000-0002-2596-5314>

## References

- Nicolussi S, Drewe J, Butterweck V, Meyer zu Schwabedissen HE. Clinical relevance of St. John's wort drug interactions revisited. *Br J Pharmacol*. 2020;177(6):1212–1226. doi:10.1111/bph.14936
- Kasper S, Gastpar M, Möller HJ, et al. Better tolerability of St. John's wort extract WS 5570 compared to treatment with SSRIs: A reanalysis of data from controlled clinical trials in acute major depression. *Int Clin Psychopharmacol*. 2010;25(4):204–213. doi:10.1097/YIC.0b013e328335dc1a
- Battaglia S, Garofalo S, di Pellegrino G, Starita F. Revaluing the role of vmPFC in the acquisition of Pavlovian threat conditioning in humans. *J Neurosci*. 2020;40(44):8491–8500. doi:10.1523/JNEUROSCI.0304-20.2020
- Battaglia S, Harrison BJ, Fullana MA. Does the human ventromedial prefrontal cortex support fear learning, fear extinction or both? A commentary on subregional contributions. *Mol Psychiatry*. 2022;27(2):784–786. doi:10.1038/s41380-021-01326-4
- Battaglia S. Neurobiological advances of learned fear in humans. *Adv Clin Exp Med*. 2022;31(3):217–221. doi:10.17219/acem/146756
- Battaglia S, Thayer JF. Functional interplay between central and autonomic nervous systems in human fear conditioning. *Trends Neurosci*. 2022;45(7):504–506. doi:10.1016/j.tins.2022.04.003
- Tanaka M, Vécsei L. Editorial of Special Issue "Crosstalk between Depression, Anxiety, and Dementia: Comorbidity in Behavioral Neurology and Neuropsychiatry." *Biomedicines*. 2021;9(5):517. doi:10.3390/biomedicines9050517
- Avan R, Sahebnaasagh A, Hashemi J, et al. Update on statin treatment in patients with neuropsychiatric disorders. *Life (Basel)*. 2021;11(12):1365. doi:10.3390/life11121365
- Kalb R, Trautmann-Sponzel RD, Kieser M. Efficacy and tolerability of hypericum extract WS 5572 versus placebo in mildly to moderately depressed patients: A randomized double-blind multicenter clinical trial. *Pharmacopsychiatry*. 2001;34(3):96–103. doi:10.1055/s-2001-14280
- Shelton RC, Keller MB, Gelenberg A, et al. Effectiveness of St John's wort in major depression: A randomized controlled trial. *JAMA*. 2001;285(15):1978–1986. doi:10.1001/jama.285.15.1978
- Forsdike K, Pirotta M. St John's wort for depression: Scoping review about perceptions and use by general practitioners in clinical practice. *J Pharm Pharmacol*. 2018;71(1):117–128. doi:10.1111/jphp.12775
- Bjerkenstedt L, Edman GV, Alken RG, Mannel M. Hypericum extract LI 160 and fluoxetine in mild to moderate depression: A randomized, placebo-controlled multicenter study in outpatients. *Eur Arch Psychiatry Clin Neurosci*. 2005;255(1):40–47. doi:10.1007/s00406-004-0532-z
- Fava M, Alpert J, Nierenberg AA, et al. A double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. *J Clin Psychopharmacol*. 2005;25(5):441–447. doi:10.1097/01.jcp.0000178416.60426.29
- Van Gurp G, Meterisian GB, Haiek LN, McCusker J, Bellavance F. St John's wort or sertraline? Randomized controlled trial in primary care. *Can Fam Physician*. 2002;48:905–912. PMID:12053635. PMCID: PMC2214056.
- Hypericum Depression Trial Study Group. Effect of *Hypericum perforatum* (St John's wort) in major depressive disorder: A randomized controlled trial. *JAMA*. 2002;287(14):1807. doi:10.1001/jama.287.14.1807
- Kasper S, Angheliescu IG, Szegedi A, Diemel A, Kieser M. Superior efficacy of St John's wort extract WS\*5570 compared to placebo in patients with major depression: A randomized, double-blind, placebo-controlled, multi-center trial [ISRCTN72727298]. *BMC Med*. 2006;4(1):14. doi:10.1186/1741-7015-4-14
- Leclubier Y, Clerc G, Didi R, Kieser M. Efficacy of St. John's wort extract WS 5570 in major depression: A double-blind, placebo-controlled trial. *Am J Psychiatry*. 2002;159(8):1361–1366. doi:10.1176/appi.ajp.159.8.1361
- Moreno RA, Teng CT, de Almeida KM, Tavares H, Jr. *Hypericum perforatum* versus fluoxetine in the treatment of mild to moderate depression: A randomized double-blind trial in a Brazilian sample. *Rev Bras Psiquiatr*. 2006;28(1):29–32. doi:10.1590/S1516-44462006000100007
- Ur-Rahman R, Ansari MA, Hayder Z, Siddiqui AA, Bukhari IA, Quayyum MA. Double blind placebo controlled clinical trial examining the effectiveness of St. John's wort (*Hypericum perforatum*) in mild to moderate depression. *J Pakistan Psychiatric Soc*. 2008;5(2):106–111. [http://www.jpss.com.pk/article/\\_2359.html](http://www.jpss.com.pk/article/_2359.html). Accessed December 15, 2021.
- Sarris J, Fava M, Schweitzer I, Mischoulon D. St John's wort (*Hypericum perforatum*) versus sertraline and placebo in major depressive disorder: Continuation data from a 26-week RCT. *Pharmacopsychiatry*. 2012;45(07):275–278. doi:10.1055/s-0032-1306348
- Seifritz E, Hatzinger M, Holsboer-Trachsler E. Efficacy of *Hypericum* extract WS\*5570 compared with paroxetine in patients with a moderate major depressive episode: A subgroup analysis. *Int J Psychiatry Clin Pract*. 2016;20(3):126–132. doi:10.1080/13651501.2016.1179765
- Simeon J, Nixon MK, Milin R, Jovanovic R, Walker S. Open-label pilot study of St. John's wort in adolescent depression. *J Child Adolesc Psychopharmacol*. 2005;15(2):293–301. doi:10.1089/cap.2005.15.293
- Singer A, Schmidt M, Hauke W, Stade K. Duration of response after treatment of mild to moderate depression with *Hypericum* extract STW 3-VI, citalopram and placebo: A reanalysis of data from a controlled clinical trial. *Phytomedicine*. 2011;18(8–9):739–742. doi:10.1016/j.phymed.2011.02.016
- Szegedi A, Kohnen R, Diemel A, Kieser M. Acute treatment of moderate to severe depression with hypericum extract WS 5570 (St John's wort): Randomised controlled double blind non-inferiority trial versus paroxetine. *BMJ*. 2005;330(7490):503. doi:10.1136/bmj.38356.655266.82
- Weber W, Vander Stoep A, McCarty RL, Weiss NS, Biederman J, McClellan J. *Hypericum perforatum* (St John's wort) for attention-deficit/hyperactivity disorder in children and adolescents: A randomized controlled trial. *JAMA*. 2008;299(22):2633. doi:10.1001/jama.299.22.2633
- Linde K, Mulrow C. St John's wort for depression. *Cochrane Database Syst Rev*. 2008;4:CD000448. doi:10.1002/14651858.CD000448
- Apaydin EA, Maher AR, Shanman R, et al. A systematic review of St. John's wort for major depressive disorder. *Syst Rev*. 2016;5(1):148. doi:10.1186/s13643-016-0325-2
- Hübner WD, Kirste T. Experience with St John's Wort (*Hypericum perforatum*) in children under 12 years with symptoms of depression and psychovegetative disturbances. *Phytother Res*. 2001;15(4):367–370. doi:10.1002/ptr.829
- Behnke K, Jensen GS, Graubaum HJ, Gruenwald J. *Hypericum perforatum* versus fluoxetine in the treatment of mild to moderate depression. *Adv Therapy*. 2002;19(1):43–52. doi:10.1007/BF02850017
- Ferrara M, Mungai F, Starace F. St John's wort (*Hypericum perforatum*)-induced psychosis: A case report. *J Med Case Rep*. 2017;11(1):137. doi:10.1186/s13256-017-1302-7
- Arold G, Donath F, Maurer A, et al. No relevant interaction with alprazolam, caffeine, tolbutamide, and digoxin by treatment with a low-hyperforin St John's wort extract. *Planta Med*. 2005;71(4):331–337. doi:10.1055/s-2005-864099

32. Tawfik GM, Dila KAS, Mohamed MYF, et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health*. 2019;47(1):46. doi:10.1186/s41182-019-0165-6
33. Jackson D, Law M, Barrett JK, et al. Extending DerSimonian and Laird's methodology to perform network meta-analyses with random inconsistency effects. *Stat Med*. 2016;35(6):819–839. doi:10.1002/sim.6752
34. Efthimiou O, Rücker G, Schwarzer G, Higgins JPT, Egger M, Salanti G. Network meta-analysis of rare events using the Mantel–Haenszel method. *Stat Med*. 2019;38(16):2992–3012. doi:10.1002/sim.8158
35. Kanter JW, Busch AM, Weeks CE, Landes SJ. The nature of clinical depression: Symptoms, syndromes, and behavior analysis. *Behav Analyst*. 2008;31(1):1–21. doi:10.1007/BF03392158
36. Kennedy SH. Core symptoms of major depressive disorder: Relevance to diagnosis and treatment. *Dialogues Clin Neurosci*. 2008;10(3):271–277. doi:10.31887/DCNS.2008.10.3/shkennedy
37. Schneiderman N, Ironson G, Siegel SD. Stress and health: Psychological, behavioral, and biological determinants. *Annu Rev Clin Psychol*. 2005;1(1):607–628. doi:10.1146/annurev.clinpsy.1.102803.144141
38. Salari N, Hosseini-Far A, Jalali R, et al. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: A systematic review and meta-analysis. *Globalization and Health*. 2020;16(1):57. doi:10.1186/s12992-020-00589-w
39. Gundogan S, Arpacı I. Depression as a mediator between fear of COVID-19 and death anxiety [Published online as ahead of print on April 26, 2022]. *Curr Psychol*. 2022. doi:10.1007/s12144-022-03120-z
40. Olfson M. National trends in the outpatient treatment of depression. *JAMA*. 2002;287(2):203. doi:10.1001/jama.287.2.203
41. Duval F, Lebowitz BD, Macher JP. Treatments in depression. *Dialogues Clin Neurosci*. 2006;8(2):191–206. doi:10.31887/DCNS.2006.8.2/fduval
42. Gabriel FC, de Melo DO, Fráguas R, Leite-Santos NC, Mantovani da Silva RA, Ribeiro E. Pharmacological treatment of depression: A systematic review comparing clinical practice guideline recommendations. *PLoS ONE*. 2020;15(4):e0231700. doi:10.1371/journal.pone.0231700
43. Cid-Jofré V, Moreno M, Reyes-Parada M, Renard GM. Role of oxytocin and vasopressin in neuropsychiatric disorders: Therapeutic potential of agonists and antagonists. *Int J Mol Sci*. 2021;22(21):12077. doi:10.3390/ijms222112077
44. Spekker E, Tanaka M, Szabó Á, Vécsei L. Neurogenic inflammation: The participant in migraine and recent advancements in translational research. *Biomedicines*. 2021;10(1):76. doi:10.3390/biomedicines10010076
45. Tanaka M, Tóth F, Polyák H, Szabó Á, Mándi Y, Vécsei L. Immune influencers in action: Metabolites and enzymes of the tryptophan-kynurenine metabolic pathway. *Biomedicines*. 2021;9(7):734. doi:10.3390/biomedicines9070734
46. Mutz J, Edgcumbe DR, Brunoni AR, Fu CHY. Efficacy and acceptability of non-invasive brain stimulation for the treatment of adult unipolar and bipolar depression: A systematic review and meta-analysis of randomised sham-controlled trials. *Neurosci Biobehav Rev*. 2018;92:291–303. doi:10.1016/j.neubiorev.2018.05.015
47. Tortella G, Selingardi P, Moreno M, Veronezi B, Brunoni A. Does non-invasive brain stimulation improve cognition in major depressive disorder? A systematic review. *CNS Neurol Disord Drug Targets*. 2015;13(10):1759–1769. doi:10.2174/1871527313666141130224431
48. Borgomaneri S, Battaglia S, Avenanti A, Pellegrino G di. Don't Hurt Me No More: State-dependent transcranial magnetic stimulation for the treatment of specific phobia. *J Affect Disord*. 2021;286:78–79. doi:10.1016/j.jad.2021.02.076
49. Tortella G, Selingardi P, Moreno M, Veronezi B, Brunoni A. Does non-invasive brain stimulation improve cognition in major depressive disorder? A systematic review. *CNS Neurol Disord Drug Targets*. 2015;13(10):1759–1769. doi:10.2174/1871527313666141130224431
50. Mutz J, Edgcumbe DR, Brunoni AR, Fu CHY. Efficacy and acceptability of non-invasive brain stimulation for the treatment of adult unipolar and bipolar depression: A systematic review and meta-analysis of randomised sham-controlled trials. *Neurosci Biobehav Rev*. 2018;92:291–303. doi:10.1016/j.neubiorev.2018.05.015
51. Battaglia S, Orsolini S, Borgomaneri S, Barbieri R, Diciotti S, di Pellegrino G. Characterizing cardiac autonomic dynamics of fear learning in humans [published online as ahead of print on June 7, 2022]. *Psychophysiology*. 2022. doi:10.1111/psyp.14122



# A meta-analysis to investigate the role of magnetic resonance spectroscopy in the detection of temporal lobe epilepsy

Xiangxiang Cui<sup>B,C</sup>, Dan Zhong<sup>D–F</sup>, Jinou Zheng<sup>A</sup>

Department of Neurology, First Affiliated Hospital of Guangxi Medical University, Minzu Hospital of Guangxi Zhuang Autonomous Region, Nanning, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):163–173

## Address for correspondence

Jinou Zheng

E-mail: jio-2006@outlook.com

## Funding sources

Guangxi Natural Science Foundation grant No. 2018GXNSFAA0150149.

## Conflict of interest

None declared

Received on January 30, 2022

Reviewed on June 22, 2022

Accepted on August 23, 2022

Published online on October 13, 2022

## Cite as

Cui X, Zhong D, Zheng J. A meta-analysis to investigate the role of magnetic resonance spectroscopy in the detection of temporal lobe epilepsy. *Adv Clin Exp Med.* 2023;32(2):163–173. doi:10.17219/acem/152943

## DOI

10.17219/acem/152943

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Temporal lobe epilepsy (TLE) is a complex neuropsychiatric illness that alters patient's behavior and social well-being. Early and accurate diagnosis of TLE is a key factor in its treatment. Currently, magnetic resonance spectroscopy (MRS) is recommended due to its high sensitivity and specificity in the detection of seizures. However, some studies have reported on its limited role in lesion detection.

**Objectives.** The present meta-analysis aims to analyze different studies thoroughly and investigate the role of MRS in the detection of TLE.

**Materials and methods.** A systematic literature review was conducted using Medline (via PubMed), Cinahl (via EBSCO), Scopus, and Web of Sciences databases for case-control, retrospective and prospective studies regarding the use of MRS in detecting TLE. Studies were included using the Population, Intervention, Comparison, Outcomes and Study (PICOS) criteria and relevant event data were extracted. The risk of publication bias was analyzed using the Begg and Mazumdar test. A risk of bias assessment was conducted using RevMan software. The Mantel–Haenszel method was used to calculate the sensitivity, pooled odds ratio (OR), and risk ratio (RR), also using RevMan software.

**Results.** A total of 16 studies published between 2000 and 2022 were included, which encompassed a total of 645 patients. We obtained a high sensitivity of 84.8%, which shows a high efficiency of MRS, and a pooled OR value of 0.37 (95% confidence interval (95% CI): 0.14–0.97) with a tau<sup>2</sup> value of 2.63,  $\chi^2$  value of 84.99, degrees of freedom (df) value of 15, I<sup>2</sup> value of 82%, Z-value of 2.03, and  $p < 0.05$ . The pooled RR was 0.82 (95% CI: 0.69–0.97) with a tau<sup>2</sup> value of 0.10,  $\chi^2$  value of 122.11, df of 15, I<sup>2</sup> value of 88%, Z-value of 2.25, and  $p < 0.05$ . These results were statistically significant for a low risk of publication bias.

**Conclusions.** The current meta-analysis highly recommends the use of MRS in the accurate detection of alterations seen in the brain in patients with TLE.

**Key words:** magnetic resonance spectroscopy (MRS), temporal lobe epilepsy (TLE), electroencephalography (EEG), magnetic resonance imaging (MRI)

## Background

A seizure is defined as a sudden, uncontrolled electrical disturbance in the brain that can alter feelings, behavior, movements, and level of consciousness. According to the new 2017 International League Against Epilepsy (ILAE) Seizure Classification,<sup>1</sup> a seizure originating in one hemisphere of the brain is classified as having a focal onset. If it appears simultaneously in both hemispheres, it is categorized as having a generalized onset. Two or more seizures within 24 h without an identifiable cause meet the criteria for epilepsy. Assuming that the patient is having epileptic seizures, the ILAE guidelines recommend the seizures be analyzed on 3 levels: level 1 – diagnosis of the seizure type; level 2 – diagnosis of epilepsy type (focal epilepsy, generalized epilepsy, or combined generalized and focal epilepsy); and level 3 – diagnosis of specific epilepsy syndromes. This new classification incorporates etiology of each stage and entails significant treatment implications.<sup>2</sup>

Temporal lobe epilepsy (TLE) is the most common form of localized or focal epilepsy, characterized by recurrent, unprovoked focal seizures originating in the temporal lobe of the brain. Due to this, patients have an altered sense of hearing, smell, taste, and touch, with a sudden sense of fear, panic, anxiety, confusion, and anger. Therefore, it is necessary to recognize this problem early, before it can spread. This condition can be caused by an infection in the brain, traumatic brain injury, meningitis, stroke, cerebral tumors, and febrile seizures, among other causes. It is treated with medications, diet control, electrical brain stimulators, and surgery in extreme cases.<sup>3</sup>

For its diagnosis, magnetic resonance imaging (MRI), positron emission tomography (PET), electroencephalography (EEG), and magnetic resonance spectroscopy (MRS) can be used. Clear imaging of suspected lesions is required for the accurate diagnosis and successful treatment of TLE. Among the various imaging techniques used for the detection of TLE, MRS is the most preferred technique owing to its high sensitivity (85–88%) and diagnostic accuracy. The MRS images can clearly demarcate the lateralization of TLE, mesial temporal sclerosis with hippocampal atrophy, and structural distortions. This technique can detect the dysfunction of neurons using the anatomic distribution of signals from metabolites like N-acetyl aspartate (NAA), creatine (Cr), phosphocreatine, and choline (Cho)-containing compounds. Among these, NAA is of major interest due to its neuronal distribution and the fact that any alteration or reduction in NAA signaling is indicative of TLE.<sup>4</sup>

Various studies have recommended its use in the detection of TLE – for example, papers by Someya et al.,<sup>5</sup> Li et al.,<sup>6</sup> Matton et al.,<sup>7</sup> Lundbom et al.,<sup>8</sup> and Doelken et al.<sup>9</sup> These studies reported that MRS is effective in the identification of an epileptic zone in epilepsy patients and has a high diagnostic accuracy. Mantoan et al. reported that a reduction in NAA levels on the contralateral side acts as a marker for the identification of epileptogenicity in the contralateral mesial temporal

structures.<sup>10</sup> Similarly, Shen et al.<sup>11</sup> and Azab et al.<sup>12</sup> reported that MRS is very useful in the preoperative investigation of patients with TLE and can detect lateralization and bilateral abnormalities. Similarly, Aun et al. reported that MRS enables multiparametric mapping of TLE in less time and with a high sensitivity, and is a promising tool in evaluating patients with TLE.<sup>13</sup> In a prospective study by Ercan et al.<sup>14</sup> evaluating 14 patients, and a multimodal imaging study by Voets et al.<sup>15</sup> evaluating 12 patients, MRS was reported to be the preferred tool for studying the lateralization of epileptogenic foci and evaluating patients preoperatively. Meanwhile, a study by Liao et al. reported that magnetic resonance (MR) fingerprinting is more sensitive and accurate compared to MRS in the detection of temporal lobe seizures.<sup>16</sup> In contrast, Saber et al.<sup>17</sup> and Ma et al.<sup>18</sup> highly recommend the use of MRS in TLE. Essam-el-dein Mohamed et al. suggested its limited use and recommends the combination of pulsed arterial spin-labeling (PASL) MRI and proton magnetic resonance spectroscopy (1H-MRS) for the proper identification of epileptogenic zones in patients.<sup>19</sup> Liu et al. conducted a retrospective study that highlights the diagnosing capacity of MRS and recommended it highly.<sup>20</sup> Considering the contrasting data of MRS in the detection and localization of TLE, the current meta-analysis aimed to statistically analyze the literature and predict diagnostic accuracy of MRS.

## Objectives

This study aimed to investigate the role of MRS in the detection of changes and abnormalities in the hippocampus and limbic system of patients with TLE.

## Materials and methods

We followed the guidelines and normative recommendations by Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA; registration No. NU#/IRB/2020/1022).

## Search strategy

This meta-analysis is based on an extensive search of the databases of Medline (via PubMed), Cinahl (via Ebsco), Scopus, and Web of Science between 2000 and 2022. The following keywords were used to search for relevant studies: “magnetic resonance spectroscopy”, “MRS”, “temporal lobe epilepsy”, “TLE”, “electroencephalography”, “EEG”, “magnetic resonance imaging”, and “MRI”. The included articles were selected using the PRISMA guidelines and randomly selected irrespective of the language, publication status, or type of study (prospective, retrospective, clinical trial). Demographic summary of the patients and event data of the included studies were collected and analyzed.<sup>5–20</sup> Two of the authors (JF and XZ) separately scanned

relevant sources for related studies. Mainly full texts articles were collected; abstracts were used only if they provided sufficient information for the meta-analysis. Obsolete references were excluded and useful studies were included if they met the inclusion criteria. Event data from useful variables were collected by 2 researchers (LC and SL) independently.

## Inclusion and exclusion criteria

The studies included reported on the use of MRS as an effective imaging method for patients with TLE. Studies were selected from the years 2000 to 2022. In the current study, only full-text data were included, while studies with insufficient data, studies reporting the use of other imaging methods, and studies published before 2000 were excluded.

## Evaluation of the analytical standard and source of heterogeneity

Two reviewers (JF and XZ) separately evaluated the methodological validity of the included studies and calculated the heterogeneity for the included studies, while author SL was responsible for resolving any type of disagreement between JF and XZ. To investigate the heterogeneity, a Cochran's Q test and  $I^2$  index were calculated with RevMan software v. 5 (The Nordic Cochrane Center, The Cochrane Collaboration, Copenhagen, Denmark).<sup>21</sup> The heterogeneity was investigated for the use of different study types (case-control, prospective and retrospective studies), age of patients, sex ratios, imaging instruments, and lab technicians' performance.

## Statistical analyses

For statistical analysis, a diagnostic odds ratio (OR) was calculated using the DerSimonian–Laird technique.<sup>22</sup> To accomplish this, a 2×2 table was constructed using event data and the meta-analysis was performed using RevMan software. A pooled diagnostic OR with 95% confidence interval (95% CI) and a risk ratio (RR) with 95% CI was calculated using the Mantel–Haenszel test.<sup>23</sup> Respective forest plots and the heterogeneity of the included studies ( $\chi^2$  value, Q value, degrees of freedom (df) value,  $I^2$  value, and p-value) were evaluated with RevMan software. To assess the efficiency, MRI sensitivities were also analyzed. A risk of bias summary was also performed using the RevMan software, while a Deek's funnel plot was created using MedCalc software v. 20.115 (MedCalc Software, Ostend, Belgium) to assess the risk of publication bias.<sup>24</sup>

## Results

### Literature search results

We found a total of 1128 studies through our electronic scans of different databases. Among these studies,

135 studies were excluded after reading their titles and abstracts, leaving 993 records to be screened. We excluded additional 743 studies because of improper study design and duplicity reasons. This left 250 studies for final screening. Out of the 250 studies, 209 studies were excluded because they did not fulfill the inclusion criteria, and thus 41 studies were eligible for further assessment. The key reasons for exclusion were inadequate evidence and comparison criteria inappropriate to create a 2×2 table for review. Finally, for this meta-analysis, 16 studies published between the years 2000 to 2022 fulfilled the inclusion criteria i.e., the use of MRS for the detection of TLE, as shown in Fig. 1.

The included studies encompassed a total of 645 epileptic patients of different ages who were chosen randomly and scanned with MRS along with an EEG or MRI. The demographic details of the studies were included in this meta-analysis and are shown in Table 1. The details collected included the author of the study, publishing year, type of study, method of assessment, duration of the study, total sample size, age and gender of the patients, type of instruments used, total number of epileptic cases diagnosed, and total number of cases tested. Event data of these studies (including the number of total samples studied and epileptic cases diagnosed using MRS and EEG) were collected and a meta-analysis was performed.

## Results

The risk of bias for the included studies was assessed as shown in Table 2. The risk of bias is summarized in Fig. 2 and a risk of bias graph is shown as Fig. 3.

The current meta-analysis had a low risk of publication bias, as apparent from the funnel plot (Fig. 4) and results of the Egger's and Begg and Mazumdar tests. The Egger's regression test determines the degree of funnel plot asymmetry by measuring the intercept from the regression of the standard normal that deviates against the precision, while the Begg and Mazumdar rank correlation test explains the correlation between ranks of effect sizes and their variances. If  $p > 0.05$ , then the study has low risk of publication bias. The Begg's test and Egger's test values were considered significant and with low risk of publication bias if their p-value is greater than 0.05. Since we also achieved the p-values of Begg's test and Egger's test greater than 0.05 as p-value of 0.78 for the Begg's test and 0.57 for the Egger's test, respectively, this confirms that present study has low risk of publication bias.<sup>25</sup>

We obtained a sensitivity of 84.8%, which showed a high efficiency of MRI in detecting all the neurobiological changes that occur during TLE. The OR of the included studies was calculated using RevMan software and the Mantel–Haenszel test with random effects to assess the association between the exposure to MRS in epileptic patients and its outcome in the accurate detection

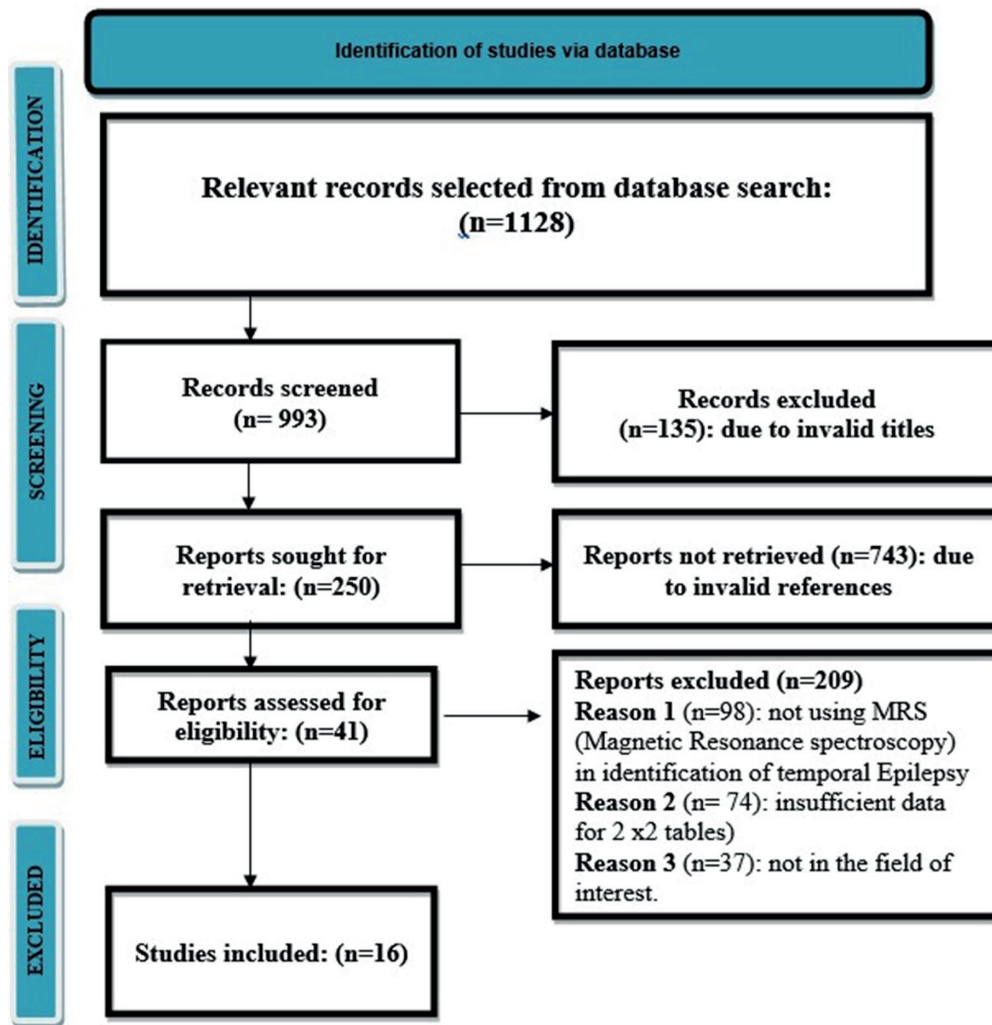


Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram of the study

Table 1. Demographic summary of included studies

Author, year, reference	Journal of publication	Study design	Study duration	Assessment	Patient's details			Instrument used	Epilepsy presence	Results	
					number	sex	age			MRS	EEG
Someya et al., 2000 [5]	<i>Seizure</i>	case-control study	–	quantitative	17	5 M 12 F	24–40 years	MRS (Gyrosan ACS2 operated at 1.5 T (Philips Medical Systems, Amsterdam, the Netherlands) EEG (Stellate Harmonie system; Stellate Systems, Westmount Canada)	total cases epileptic	17 15	17 17
Li et al., 2000 [6]	<i>Epilepsia</i>	case-control study	–	quantitative	150	64 M 86 F	12–64 years	MRS (Philips ACS I1 or 111 combined imaging and spectroscopy system (1.5 T; Philips Medical System). EEG (Stellate Harmonie system)	total cases epileptic	150 120	150 117
Maton et al., 2001 [7]	<i>Epilepsia</i>	case-control study	–	quantitative	31	15 M 16 F	25–48 years	MRS (Philips ACS I1 or 111 combined imaging and spectroscopy system (1.5 T; Philips Medical Systems). EEG (Stellate Harmonie system)	total cases epileptic	31 19	31 20

of unilateral or bilateral epilepsy in comparison to the standard method of EEG. A forest plot of the ORs and heterogeneity of data is presented in Fig. 5. We obtained a pooled OR value of 0.37 (95% CI: 0.14–0.97) with a tau<sup>2</sup> value

of 2.63,  $\chi^2$  value of 84.99, degrees of freedom (df) value of 15, I<sup>2</sup> value of 82%, Z-value of 2.03, and  $p < 0.05$ . The  $p$ -value less than 0.05 confirms the statistical significance of present meta-analysis and indicates the use of MRS for



Table 1. Demographic summary of included studies – cont.

Author, year, reference	Journal of publication	Study design	Study duration	Assessment	Patient's details			Instrument used	Epilepsy presence	Results	
					number	sex	age			MRS	EEG
Lundbom et al., 2001 [8]	<i>Epilepsia</i>	case-control study	1 year	quantitative	14	4 M 10 F	4–26 years	MRS 1.5-T MR imager (Magnetom Vision; Siemens Healthcare, Erlangen, Germany) using a standard circularly polarized (CP) head coil EEG (Stellate Harmonie system)	total cases epileptic	14 14	14 10
Doelken et al., 2008 [9]	<i>Seizure</i>	case-control study	1–2 weeks	quantitative	23	–	19–62 years	MRS (1.5 T Siemens Magnetom Sonata (Siemens Healthcare)) EEG (Glonner system (Munich, Germany)).	total cases epileptic	23 17	23 23
Mantoan et al., 2009 [10]	<i>Epilepsy and Behaviour</i>	prospective study	1 year	quantitative	29	10 M 19 W	20–61 years	1.5-T equipment (Magnetom Sonata (Maestro Class), Siemens Healthcare) 32-channel EEG recording equipment (Ceegraph software; Biologic Systems Corp., Seyssinet-Pariset, France)	total cases epileptic	29 10	29 29
Shen et al., 2009 [11]	<i>The Journal of International Medical Research</i>	case-control study	–	quantitative	38	21 M 17 F	13–62 years	MRS Intera Achieva 3.0 T MRI system (Philips Medical Systems) EEG (Stellate Harmonie system)	total cases epileptic	38 30	38 25
Azab et al., 2015 [12]	<i>Italian Journal of Paediatrics</i>	case-control study	3 years	quantitative	40	22 M 18 F	14 years	MRS (1.5-Tesla MR units-Achiva; Philips Medical Systems) EEG (Stellate Harmonie system)	total cases epileptic	40 38	40 32
Aun et al., 2016 [13]	<i>The Egyptian Journal of Radiology and Nuclear Medicine</i>	case-control study	2.5 years	quantitative	30	16 M 14 F	17–47 years	MRS (1.5 T super conducting system, Philips Achieva MRI machine) EEG (Stellate Harmonie system)	total cases epileptic	30 26	30 28
Ercan et al., 2016 [14]	<i>Disease Markers</i>	prospective study	–	quantitative	14	8 M 6 F	19–66 years	MRS (1.5 T MR scanner (Philips Achieva; slew rate 40 mT/m) EEG (Stellate Harmonie system)	total cases epileptic	14 10	14 14
Voets et al., 2017 [15]	<i>Scientific Reports Nature</i>	multimodal imaging study	–	quantitative	12	3 M 9 F	19–57 years	7 Tesla Siemens MRI system, using a 32-channel head coil EEG (Stellate Harmonie system)	total cases epileptic	12 5	12 12
Liao et al., 2018 [16]	<i>Neuroradiology</i>	case-control study	8 months	quantitative	33	10 M 23 F	16–60 years	MRS (3.0-T imager (Magnetom Prisma; Siemens Healthcare)) EEG (Stellate Harmonie system)	total cases epileptic	33 32	33 23
Saber et al., 2019 [17]	<i>Al-Azhar Journal of Paediatrics</i>	prospective study	2 years	quantitative	50	29 M 21 F	3 month–15 years	1.5 T super conducting system (Philips Achieve MRI machine) manufactured in 2009 EEG (Stellate Harmonie system)	total cases epileptic	50 12	50 50
Ma et al., 2019 [18]	<i>European Neurology</i>	retrospective study	4 years	quantitative	115	59 M 56 F	2–62 years	MRS (1.5 T super conducting system, Philips Achieva MRI machine) EEG (Stellate Harmonie system)	total cases epileptic	115 80	115 110
Essam-el-dein Mohamed et al., 2020 [19]	<i>The Egyptian Journal of Radiology and Nuclear Medicine</i>	case-control study	2 years	quantitative	26	12 M 14 F	16–57 years	<sup>1</sup> H-MRSI (Philips ACS I1 or I11 combined imaging and spectroscopy system (1.5 T, Philips Medical Systems)) EEG (Stellate Harmonie system)	total cases epileptic	26 21	26 19
Liu et al., 2021 [20]	<i>Quantitative Imaging in Medicine and Surgery</i>	retrospective study	1–2 weeks	quantitative	23	15 M 14 F	8–80 years	1.5T system (GE SIGNA HDe Medical System) EEG (Stellate Harmonie system)	total cases epileptic	23 9	23 23

MRI – magnetic resonance imaging; EEG – electroencephalography; M – male; F – female; MRS – magnetic resonance spectroscopy; <sup>1</sup>H-MRSI – proton magnetic resonance spectroscopy imaging.

Table 2. Risk assessment for included studies

Question	Someya et al., 2000 [5]	Li et al., 2000 [6]	Maton et al., 2001 [7]	Lundbom et al., 2001 [8]	Doelken et al., 2008 [9]	Mantoan et al., 2009 [10]	Shen et al., 2009 [11]	Azab et al., 2015 [12]	Aun et al., 2016 [13]	Ercan et al., 2016 [14]	Voets et al., 2017 [15]	Liao et al., 2018 [16]	Saber et al., 2019 [17]	Ma et al., 2019 [18]	Essam-el-dein Mohamed, 2020 [19]	Liu et al., 2021 [20]
Was a consecutive or random sample of patients enrolled?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Did the study avoid inappropriate exclusions?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Did all patients receive the same reference standard?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were all patients included in the analysis?	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Was the sample frame appropriate to address the target population?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were study participants sampled in an appropriate way?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were the study subjects and the setting described in detail?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Were valid methods used for the identification of the condition?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was the condition measured in a standard, reliable way for all participants?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Was there appropriate statistical analysis?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

detection of temporal lobe epilepsy. From forest plot also it is evident that results are in favor of MRS for detection of TLE as compared to EEG (Fig. 5).<sup>26</sup> This proves that detecting epileptic regions is much more accurate with MRS compared to EEG, with the latter diagnosing spike or sharp waves in the tip or front part of the temporal lobe.

The RR of included studies was calculated using RevMan software and a respective forest plot is shown in Fig. 6. The pooled RR was 0.82 (95% CI: 0.69–0.97) with a tau<sup>2</sup> value of 0.10,  $\chi^2$  value of 122.11, df value of 15, I<sup>2</sup> value of 88%, Z-value of 2.25, and  $p < 0.05$ . The RR was less than 1, which indicates a low risk and suggests that the use of MRS in epilepsy detection is effective and safe.<sup>27</sup> The heterogeneity value in our meta-analysis demonstrates the variations in study outcomes between the different studies selected for this meta-analysis; high heterogeneity shows that the chosen study results were random and different.<sup>28</sup> In our meta-analysis results, we obtained high heterogeneity, which is confirmed by a high I<sup>2</sup> index (above 80%) in both OR and RR. It shows the dispersion of effect sizes in this meta-analysis.<sup>29</sup> The Z-value shows the significant weighted average effect and is considered statistically significant with a  $p < 0.05$ .<sup>30</sup> The p-value shows the probability of obtaining a significant observed effect. We also obtained high Z-value ( $>2$ ) in both OR and RR calculations with  $p < 0.05$ , showing the statistical significance of our results.

These values suggest that a random sampling of data using categorical study variables and determining epileptic regions using EEG and MRS are comparable. The I<sup>2</sup> values above 80% suggest that broad heterogeneity was observed in the studies. A high sensitivity value of 84.8% indicates that the use of MRS for TLE detection is justified as it is highly sensitive and can detect even minor changes in the different regions of temporal lobe. P-value  $< 0.05$  means all these values are statistically significant and reflect a high diagnostic accuracy of MRS in the detection of TLE.

## Discussion

Temporal lobe epilepsy is a disease of faulty neuronal resonators and is the most common focal epilepsy, responsible for chronic neurological issues like mental confusion, mood swings and altered sensation, among other things.<sup>31</sup> It is caused by traumatic brain injuries, malformations in blood vessels of the brain, infections like encephalitis or meningitis, stroke, brain tumors, scarring (gliosis) in the hippocampus, or genetic syndromes.<sup>32,33</sup> Approximately 80% of TLE cases develop in or near the area of the hippocampus known as the mesial (“near the middle”), while neocortical or lateral TLE seizures start in the outer portion of the temporal lobe.<sup>34</sup> In a study

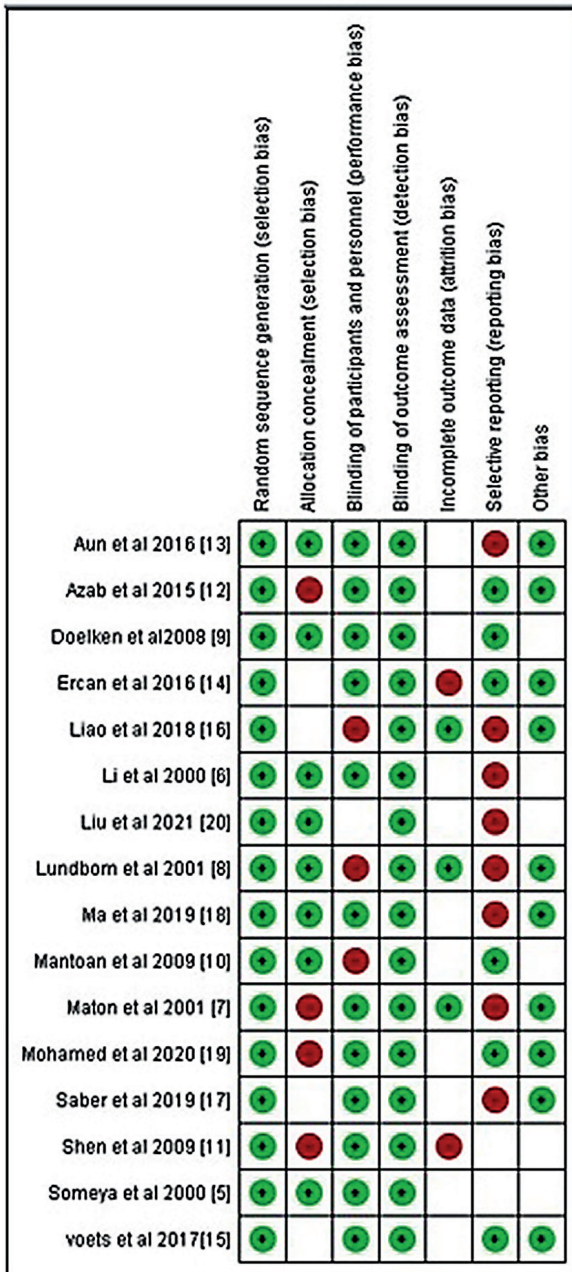


Fig. 2. Risk of bias summary for the studies included in the meta-analysis

by Navidhamidi et al., it was reported that limbic seizures begin in the mesial temporal lobe, and the hippocampal region shows a wide range of changes and epilepsy-related abnormalities in patients with temporal lobe epileptic seizures.<sup>35</sup> Epilepsy is one of the major reasons for paroxysmal alterations in the brain as it causes alterations in membrane excitability and leads to overstimulation of neurons.<sup>36</sup> Garofalo et al.<sup>37</sup> and Battaglia et al.<sup>38</sup> discussed the neurobiological signs of epilepsy that help in the accurate detection of TLE. Ahmed et al. reported that epilepsy symptoms are caused by co-occurring psychiatric disorders and antiepileptic medicines are essential for diagnosis and treatment of this epileptic disorder.<sup>39</sup> Tanaka et al. showed that disturbance in normal brain development during early life affects neuropsychiatric symptoms during later life and can be a reason for epilepsy.<sup>40</sup> Similarly, Battaglia et al. suggested that fear conditioning has a deep impact on the functioning of the central nervous system and is involved in the cardiac vagal dynamics observed in the psychophysiology of fear.<sup>41,42</sup>

Tanaka et al. reported that epilepsy along with neurodegenerative and immunological changes occur in temporal epileptic disorders and are responsible for the inflammation in these areas.<sup>43</sup> The high sensitivity of MRI help to accurately identify the inflamed areas in TLE. Approximately 66% of seizure patients can be successfully treated with medications, while in cases of mesial TLE surgery is recommended.<sup>44</sup> Borbély et al. evaluated the clinical impact of hybrid [<sup>18</sup>F]-FDG PET/MRI imaging on the detection of epilepsy in patients and reported its clinical significance.<sup>45</sup>

Zhao et al. mentioned in their review on TLE that it affects primarily cognitive functions related to memory, language and social judgment, so it is essential to detect its occurrence and cause timely and accurately since this condition affects the mental well-being of a patient and can disturb their social life.<sup>46</sup> Therefore, for diagnostic purposes, various imaging techniques like MRI, positron emission tomography (PET), EEG, computed tomography (CT) scan, and MRS are generally used.<sup>47,48</sup> However, for

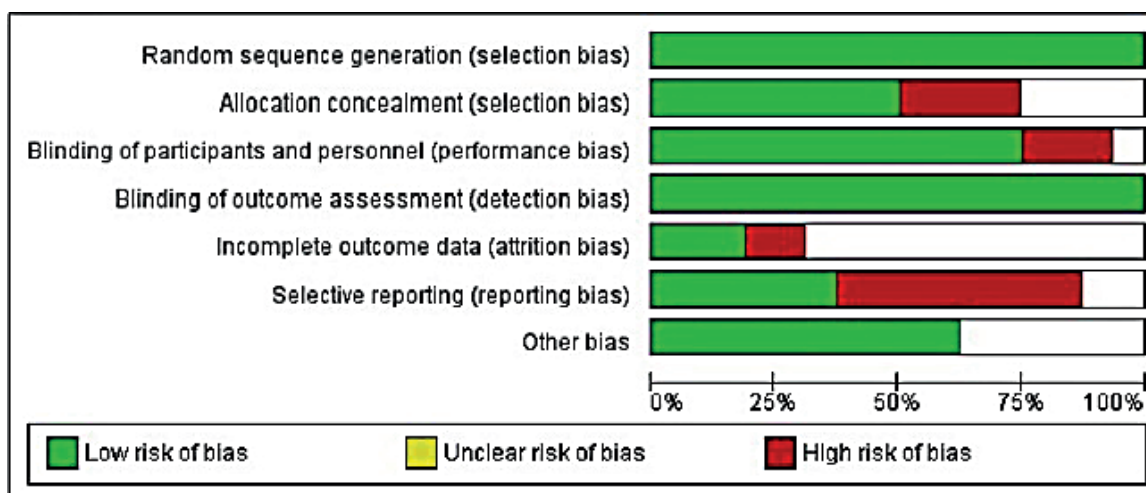


Fig. 3. Risk of bias graph for the included studies

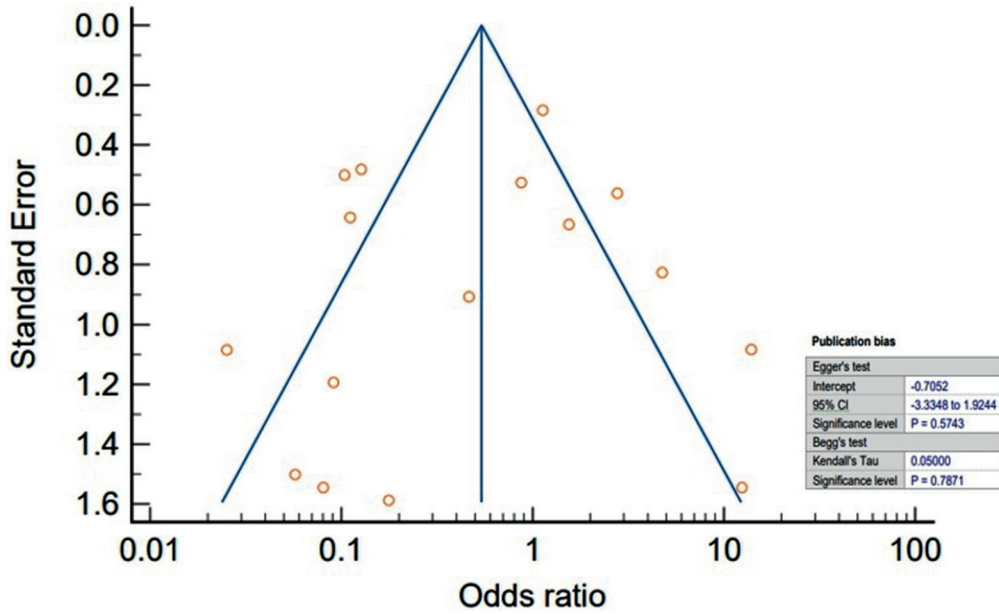


Fig. 4. Funnel plot for publication bias

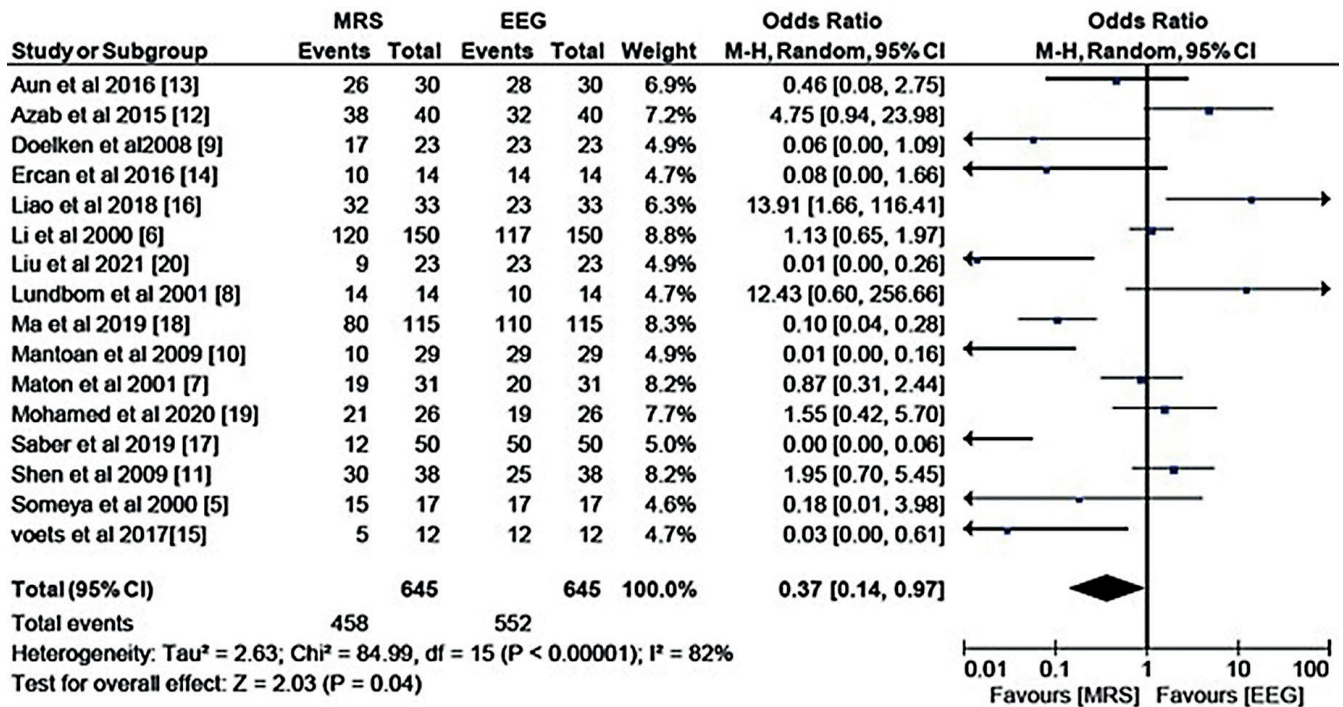


Fig. 5. Forest plot for diagnostic odds ratio (OR)

the accurate diagnosis of the cause and type of TLE, MRS is almost universally recommended on account of its high sensitivity, specificity and diagnostic accuracy.

For instance, Cendes et al.,<sup>49</sup> Lee et al.<sup>50</sup> and Wood et al.<sup>51</sup> reported that neuroimaging of epilepsy is the best way to diagnose TLE by detecting different effects epileptic seizures have on the temporal lobe. Similarly, Fan et al. reported that MRS is very useful in lateralizing and localizing the epileptogenic foci in TLE patients.<sup>52</sup> In a radiology article by Gaillard et al., it was shown that images obtained in different phases of TLE can act as markers for its detection using MRS.<sup>53</sup> However, some studies,

like Hellström et al.<sup>54</sup> and Carne et al.,<sup>55</sup> reported that MRS is useful in detecting TLE in selected cases, but not universally.

In the present study, we investigated the role of different diagnostic methods thoroughly, and similarly to the studies supporting MRS, we obtained a significant pooled OR value of 0.37 (95% CI: 0.14–0.97) with  $p < 0.05$ , which indicates high efficiency of MRS in the detection of epileptic regions. The low pooled RR of 0.82 (95% CI: 0.69–0.97) with a  $p < 0.05$  suggests that MRI is a low-risk imaging method for detection of temporal lobe epilepsy for patients of all age groups. On the basis of comparable determination rates

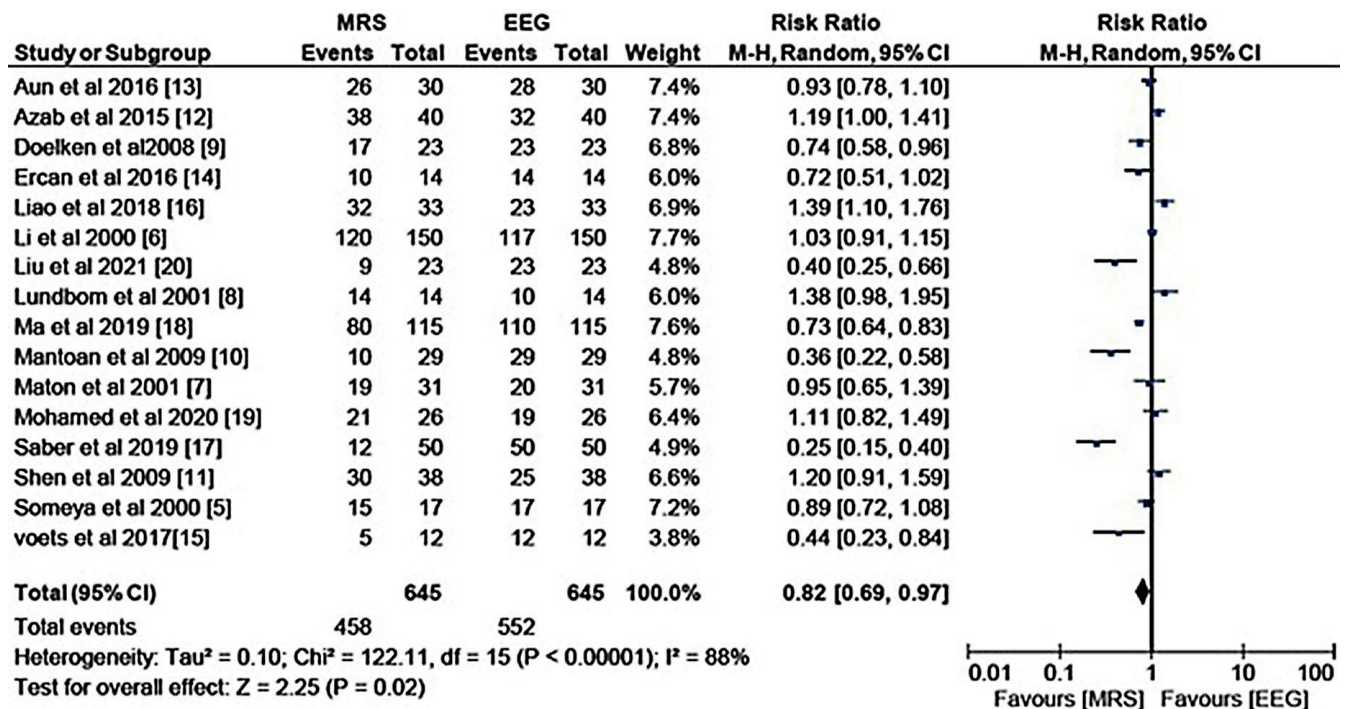


Fig. 6. Forest plot for risk ratio (RR)

of epileptic regions using MRS, similar to other neurodiagnostic techniques (EEG, MRI, etc.) and its high sensitivity, the use of MRS for detection of TLE is recommended.

Based on these statistically significant results and high sensitivity, specificity and diagnostic accuracy of this method, this meta-analysis strongly recommends the use of MRS in detecting the cause of TLE.

### Limitations

A limitation of the present study is the diversity of MRS instruments used, and the fact that the tests were performed by different radiographers, which influences the risk of false negative results. Many studies failed to report comparable diagnosis capability of MRI, EEG and PET, so assessing the comparable accuracy affects the data up to some extent. Data from other relevant studies showing the diagnostic accuracy of MRS in comparison to other diagnostic imaging methods were also included to signify the efficacy of MRS more clearly. Detailed data on a patient’s case history and physical examination and pathological tests results can further increase the diagnostic accuracy rate of MRS in the prediction of TLE.

### Conclusions

The focal or generalized onset of seizures can lead to TLE and altered behavior, sensations and social well-being. To restore mental fitness, it is essential to identify the root cause of the seizure in a timely and accurate manner. Generally, MRI, PET and EEG are common

imaging techniques used for the diagnosis of alterations in the membrane excitability in the hippocampus or the limbic system of the brain, and in the identification of the type of TLE. Still, among these methods, the high accuracy of seizure detection, high sensitivity towards any electrical stimulation and high specificity for the precise identification of epileptic zones of MRS make its use highly recommended. It is an effective imaging method for detecting the cause of TLE, and is preferred over other methods as it allows the accurate detection of epileptic zones and clearly identifies epilepsy as unilateral or bilateral, with a focal or generalized onset. It is very useful as a scanning guide in epileptic surgeries and remarkably increases the effectiveness of treatment. We performed a meta-analysis of the related available literature on the use of MRS in detecting the cause of TLE. Based on the high sensitivity (84.8%) and high efficiency of MRS in detecting all neurobiological changes that occur during TLE, and the statistically significant heterogeneity of our results, we recommend the use of MRS in determination of TLE.

### ORCID iDs

Xiangxiang Cui <https://orcid.org/0000-0003-0668-811X>  
 Dan Zhong <https://orcid.org/0000-0002-8585-4069>  
 Jinou Zheng <https://orcid.org/0000-0001-7876-0518>

### References

1. Scheffer IE, Berkovic S, Capovilla G, et al. ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia*. 2017;58(4):512–521. doi:10.1111/epi.13709
2. Sarmast ST, Abdullahi AM, Jahan N. Current classification of seizures and epilepsies: Scope, limitations and recommendations for future action. *Cureus*. 2020;12(9):e10549. doi:10.7759/cureus.10549

3. Bernhardt BC, Hong S, Bernasconi A, Bernasconi N. Imaging structural and functional brain networks in temporal lobe epilepsy. *Front Hum Neurosci.* 2013;7:624. doi:10.3389/fnhum.2013.00624
4. Park JE, Cheong EN, Jung DE, Shim WH, Lee JS. Utility of 7 Tesla magnetic resonance imaging in patients with epilepsy: A systematic review and meta-analysis. *Front Neurol.* 2021;12:621936. doi:10.3389/fneur.2021.621936
5. Someya Y, Obata T, Suhara T, et al. Seizure frequency and bilateral temporal abnormalities: A proton magnetic resonance spectroscopy of temporal lobe epilepsy. *Seizure.* 2000;9(4):274–279. doi:10.1053/seiz.2000.0396
6. Li LM, Caramanos Z, Cendes F, et al. Lateralization of temporal lobe epilepsy (TLE) and discrimination of TLE from extra-TLE using pattern analysis of magnetic resonance spectroscopic and volumetric data. *Epilepsia.* 2000;41(7):832–842. doi:10.1111/j.1528-1157.2000.tb00250.x
7. Maton B, Gilliam F, Sawrie S, Faught E, Hugg J, Kuzniecky R. Correlation of scalp EEG and 1H-MRS metabolic abnormalities in temporal lobe epilepsy. *Epilepsia.* 2002;42(3):417–422. doi:10.1046/j.1528-1157.2001.25999.x
8. Lundbom N, Gaily E, Vuori K, et al. Proton spectroscopic imaging shows abnormalities in glial and neuronal cell pools in frontal lobe epilepsy. *Epilepsia.* 2001;42(12):1507–1514. doi:10.1046/j.1528-1157.2001.15301.x
9. Doelken MT, Stefan H, Pauli E, et al. 1H-MRS profile in MRI-positive versus MRI-negative patients with temporal lobe epilepsy. *Seizure.* 2008;17(6):490–497. doi:10.1016/j.seizure.2008.01.008
10. Mantoan MAS, Caboclo LOSF, de Figueiredo Ferreira Guilhoto LM, et al. Correlation between memory, proton magnetic resonance spectroscopy, and interictal epileptiform discharges in temporal lobe epilepsy related to mesial temporal sclerosis. *Epilepsy Behav.* 2009;16(3):447–453. doi:10.1016/j.yebeh.2009.08.032
11. Shen J, Zhang L, Tian X, Liu J, Ge X, Zhang X. Use of short echo time two-dimensional <sup>1</sup>H-magnetic resonance spectroscopy in temporal lobe epilepsy with negative magnetic resonance imaging findings. *J Int Med Res.* 2009;37(4):1211–1219. doi:10.1177/147323000903700428
12. Azab SF, Sherief LM, Saleh SH, et al. Childhood temporal lobe epilepsy: Correlation between electroencephalography and magnetic resonance spectroscopy. A case-control study. *Ital J Pediatr.* 2015;41(1):32. doi:10.1186/s13052-015-0138-2
13. Aun AAK, Mostafa AA, Aboul Fotouh AM, et al. Role of magnetic resonance spectroscopy (MRS) in nonlesional temporal lobe epilepsy. *Egypt J Radiol Nucl Med.* 2016;47(1):217–231. doi:10.1016/j.ejrn.2015.09.008
14. Ercan K, Gunbey HP, Bilir E, Zan E, Arslan H. Comparative lateralizing ability of multimodality MRI in temporal lobe epilepsy. *Dis Markers.* 2016;2016:5923243. doi:10.1155/2016/5923243
15. Voets NL, Hodgetts CJ, Sen A, Adcock JE, Emir U. Hippocampal MRS and subfield volumetry at 7T detects dysfunction not specific to seizure focus. *Sci Rep.* 2017;7(1):16138. doi:10.1038/s41598-017-16046-5
16. Liao C, Wang K, Cao X, et al. Detection of lesions in mesial temporal lobe epilepsy by using MR fingerprinting. *Radiology.* 2018;288(3):804–812. doi:10.1148/radiol.2018172131
17. Saber MM, TahaElkeie M, Abdelhameid IA, Abdul Aziz MS. Correlation of metabolic alteration in magnetic resonance spectroscopy with EEG in diagnosis of childhood epilepsy. *Al-Azhar J Pediatr.* 2019;22(3):301–323. [https://journals.ekb.eg/article\\_70123.html](https://journals.ekb.eg/article_70123.html) Accessed on May 1, 2022.
18. Ma W, Li C, Liu L, Li S, Liu Y. Pre-operative interictal discharge patterns and magnetic resonance imaging findings affect prognosis of temporal lobe epilepsy surgery. *Eur Neurol.* 2019;81(3–4):152–162. doi:10.1159/000501002
19. Essam-el-dein Mohamed R, Aboelsafa AA, Dawoud RM. Arterial spin-labelling and magnetic resonance spectroscopy as imaging biomarkers for detection of epileptogenic zone in non-lesional focal impaired awareness epilepsy. *Egypt J Radiol Nucl Med.* 2020;51(1):200. doi:10.1186/s43055-020-00326-8
20. Liu D, Yang Y, Chen D, et al. Brain metabolic differences between temporal lobe epileptic seizures and organic non-epileptic seizures in postictal phase: A retrospective study with magnetic resonance spectroscopy. *Quant Imaging Med Surg.* 2021;11(8):3781–3791. doi:10.21037/qims-20-1147
21. Tawfik GM, Dila KAS, Mohamed MYF, et al. A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Trop Med Health.* 2019;47(1):46. doi:10.1186/s41182-019-0165-6
22. Jackson D, Law M, Barrett JK, et al. Extending DerSimonian and Laird's methodology to perform network meta-analyses with random inconsistency effects. *Stat Med.* 2016;35(6):819–839. doi:10.1002/sim.6752
23. Efthimiou O, Rücker G, Schwarzer G, Higgins JPT, Egger M, Salanti G. Network meta-analysis of rare events using the Mantel–Haenszel method. *Stat Med.* 2019;38(16):2992–3012. doi:10.1002/sim.8158
24. Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [Updated March 2011]*. London, UK: Cochrane Collaboration; 2011. <https://handbook-5-1.cochrane.org/>.
25. Buccheri S, Sodeck GH, Capodanno D. Statistical primer: Methodology and reporting of meta-analyses. *Eur J Cardiothorac Surg.* 2018;53(4):708–713. doi:10.1093/ejcts/ezy004
26. Alavi M, Hunt GE, Visentin DC, Watson R, Thapa DK, Cleary M. Seeing the forest for the trees: How to interpret a meta-analysis forest plot. *J Adv Nurs.* 2021;77(3):1097–1101. doi:10.1111/jan.14721
27. Viswanathan M, Patnode CD, Berkman ND, et al. Recommendations for assessing the risk of bias in systematic reviews of health-care interventions. *J Clin Epidemiol.* 2018;97:26–34. doi:10.1016/j.jclinepi.2017.12.004
28. Chyou PH. A simple and robust way of concluding meta-analysis results using reported P values, standardized effect sizes, or other statistics. *Clin Med Res.* 2012;10(4):219–223. doi:10.3121/cm.2012.1068
29. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. *BMC Med Res Methodol.* 2014;14(1):120. doi:10.1186/1471-2288-14-120
30. Kruschke JK, Liddell TM. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychon Bull Rev.* 2018;25(1):178–206. doi:10.3758/s13423-016-1221-4
31. Helmstaedter C, Kurthen M, Lux S, Reuber M, Elger CE. Chronic epilepsy and cognition: A longitudinal study in temporal lobe epilepsy. *Ann Neurol.* 2003;54(4):425–432. doi:10.1002/ana.10692
32. Vinti V, Dell'Isola GB, Tascini G, et al. Temporal lobe epilepsy and psychiatric comorbidity. *Front Neurol.* 2021;12:775781. doi:10.3389/fneur.2021.775781
33. Miró J, Ripollés P, López-Barroso D, et al. Atypical language organization in temporal lobe epilepsy revealed by a passive semantic paradigm. *BMC Neurol.* 2014;14(1):98. doi:10.1186/1471-2377-14-98
34. Engel J. Mesial temporal lobe epilepsy: What have we learned? *Neuroscientist.* 2001;7(4):340–352. doi:10.1177/107385840100700410
35. Navidhamidi M, Ghasemi M, Mehranfard N. Epilepsy-associated alterations in hippocampal excitability. *Rev Neurosci.* 2017;28(3):307–334. doi:10.1515/revneuro-2016-0059
36. Stafstrom CE, Carmant L. Seizures and epilepsy: An overview for neuroscientists. *Cold Spring Harb Perspect Med.* 2015;5(6):a022426. doi:10.1101/cshperspect.a022426
37. Garofalo S, Timmermann C, Battaglia S, Maier ME, di Pellegrino G. Medial frontal negativity signals unexpected timing of salient outcomes. *J Cogn Neurosci.* 2017;29(4):718–727. doi:10.1162/jocn\_a\_01074
38. Battaglia S. Neurobiological advances of learned fear in humans. *Adv Clin Exp Med.* 2022;31(3):217–221. doi:10.17219/acem/146756
39. Ahmed GK, Darwish AM, Khalifa H, Haridy NA. Relationship between attention deficit hyperactivity disorder and epilepsy: A literature review. *Egypt J Neurol Psychiatry Neurosurg.* 2022;58(1):52. doi:10.1186/s41983-022-00482-w
40. Tanaka M, Spekker E, Szabó Á, Polyák H, Vécsei L. Modelling the neurodevelopmental pathogenesis in neuropsychiatric disorders. Bioactive kynurenes and their analogues as neuroprotective agents: In celebration of 80<sup>th</sup> birthday of Professor Peter Riederer. *J Neural Transm.* 2022;129(5–6):627–642. doi:10.1007/s00702-022-02513-5
41. Battaglia S, Thayer JF. Functional interplay between central and autonomic nervous systems in human fear conditioning. *Trends Neurosci.* 2022;45(7):504–506. doi:10.1016/j.tins.2022.04.003
42. Battaglia S, Orsolini S, Borgomaneri S, Barbieri R, Diciotti S, di Pellegrino G. Characterizing cardiac autonomic dynamics of fear learning in humans. *Psychophysiology.* 2022;2022:e14122. doi:10.1111/psyp.14122

43. Tanaka M, Toldi J, Vécsei L. Exploring the etiological links behind neurodegenerative diseases: Inflammatory cytokines and bioactive kynurenines. *Int J Mol Sci*. 2020;21(7):2431. doi:10.3390/ijms21072431
44. Arifin MT, Bakhtiar Y, Andar EBPS, et al. Surgery for radiologically normal-appearing temporal lobe epilepsy in a centre with limited resources. *Sci Rep*. 2020;10(1):8144. doi:10.1038/s41598-020-64968-4
45. Borbély K, Emri M, Kenessey I, et al. PET/MRI in the presurgical evaluation of patients with epilepsy: A concordance analysis. *Biomedicines*. 2022;10(5):949. doi:10.3390/biomedicines10050949
46. Zhao F, Venkatesh D, Chandra M, Rastogi P, Kang H, You Li. Neuropsychological deficits in temporal lobe epilepsy: A comprehensive review. *Ann Indian Acad Neurol*. 2014;17(4):374. doi:10.4103/0972-2327.144003
47. Coito A, Michel CM, van Mierlo P, Vulliemoz S, Plomp G. Directed functional brain connectivity based on EEG source imaging: Methodology and application to temporal lobe epilepsy. *IEEE Trans Biomed Eng*. 2016;63(12):2619–2628. doi:10.1109/TBME.2016.2619665
48. Salmenpera TM. Imaging in epilepsy. *J Neurol Neurosurg Psychiatry*. 2005;76(Suppl 3):iii2–iii10. doi:10.1136/jnnp.2005.075135
49. Cendes F, Theodore WH, Brinkmann BH, Sulc V, Cascino GD. Neuroimaging of epilepsy. In: *Handbook of Clinical Neurology*. Vol. 136. Amsterdam, the Netherlands: Elsevier; 2016:985–1014. doi:10.1016/B978-0-444-53486-6.00051-X
50. Lee SK, Kim DW, Kim KK, Chung CK, Song IC, Chang KH. Effect of seizure on hippocampus in mesial temporal lobe epilepsy and neocortical epilepsy: An MRS study. *Neuroradiology*. 2005;47(12):916–923. doi:10.1007/s00234-005-1447-8
51. Wood SJ, Berger GE, Wellard RM, et al. Medial temporal lobe glutathione concentration in first episode psychosis: A 1H-MRS investigation. *Neurobiol Dis*. 2009;33(3):354–357. doi:10.1016/j.nbd.2008.11.018
52. Fan Z, Zhang Y, Ju W, Liang J, Li D, Wang X. MRS application in temporal lobe epilepsy without hippocampal sclerosis. *Chin J Med Imaging Technol*. 2017;12:1326–1330.
53. Gaillard F, Sharma R. Temporal lobe epilepsy: Reference article. 2008. <https://radiopaedia.org/articles/2152>. Accessed December 25, 2021.
54. Hellström J, Romanos Zapata R, Libard S, et al. The value of magnetic resonance spectroscopy as a supplement to MRI of the brain in a clinical setting. *PLoS One*. 2018;13(11):e0207336. doi:10.1371/journal.pone.0207336
55. Carne RP, O'Brien TJ, Kilpatrick CJ, et al. "MRI-negative PET-positive" temporal lobe epilepsy (TLE) and mesial TLE differ with quantitative MRI and PET: A case control study. *BMC Neurol*. 2007;7(1):16. doi:10.1186/1471-2377-7-1





# Evaluation of the impact of kangaroo mother care on neonatal mortality and hospitalization: A meta-analysis

Wanying Guo<sup>A–F</sup>

Department of Neonatology, Hangzhou Obstetrics and Gynecology Hospital, Zhejiang, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):175–183

## Address for correspondence

Wanying Guo

E-mail: Guowanying2020430@163.com

## Funding sources

None declared

## Conflict of interest

None declared

Received on June 16, 2022

Reviewed on July 13, 2022

Accepted on September 1, 2022

Published online on October 17, 2022

## Abstract

**Introduction.** The kangaroo mother care (KMC) technique for preterm and low-birthweight (LBW) neonates, which consists of skin-to-skin contact, is thought to have a beneficial impact on clinical outcomes. Hence, the current meta-analysis aims to evaluate the influence of KMC on neonatal mortality and length of hospitalization compared with conventional care.

**Materials and methods.** A systematic literature search of studies published between 1988 and 2021 found 24 trials involving 19,980 participants, of which 10,354 received KMC and 9626 were controls under conventional care. To measure the impact of applying KMC in preterm LBW neonates on mortality and the length of hospital stay, statistical analysis using dichotomous and continuous analysis methods was performed employing fixed and random models to calculate odds ratios (ORs) with 95% confidence intervals (95% CIs).

**Results.** Compared to the control group, the application of KMC in preterm LBW neonates resulted in significantly lower mortality (OR: 0.65, 95% CI: 0.44–0.97,  $p = 0.03$ ) in a short term (within 2 months,  $I^2 = 71%$ ) and long term (3–12 months) (OR: 0.72, 95% CI: 0.59–0.87,  $p = 0.0007$ ,  $I^2 = 0%$ ), and had no significant impact on the length of hospital stay (OR: –1.43, 95% CI: –2.88–0.02,  $p = 0.05$ ,  $I^2 = 86%$ ).

**Conclusions.** In comparison with the control group, the implementation of KMC in preterm LBW neonates resulted in significantly lower mortality but had no significant impact on the length of hospitalization. More studies are needed to confirm the current findings.

**Key words:** preterm neonates, hospital stay, mortality, kangaroo mother care

## Cite as

Guo W. Evaluation of the impact of kangaroo mother care on neonatal mortality and hospitalization: A meta-analysis. *Adv Clin Exp Med.* 2023;32(2):175–183. doi:10.17219/acem/153417

## DOI

10.17219/acem/153417

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Introduction

Millions of babies are born each year in high-risk conditions around the world. According to the World Health Organization (WHO), these conditions include premature deliveries, newborns who are underweight for their gestational age (SGA), and newborns at risk of illnesses, death or developmental disabilities. Low-birthweight (LBW) newborns are born weighing less than 2500 g, while the term 'prematurity' is used if they are born before 37 weeks of gestation.<sup>1</sup>

Many procedures have been proven to reduce the burden of newborn illnesses and mortality, and are referred to as conventional or modern neonatal care processes. It is expensive and time-consuming to provide conventional newborn care for LBW infants because it necessitates the presence of qualified staff and long-term logistics. A baby's ability to adapt to life outside the womb is complex. In low-income countries, many LBW babies are admitted to hospitals each year because of a lack of financial and human resources dedicated to their care. A significant advance in their care would be made if the new interventions for LBW newborns were found to lower neonatal morbidity and mortality, as well as related expenditures.

The majority of infant deaths occur during the newborn period. Death within the first 28 days of life is believed to have occurred in 2.5 million people last year. Most of these were LBW newborns, with 2/3 of them being premature. The Every Newborn Action Plan launched by the United Nations International Children's Emergency Fund aims to reduce newborn mortality to 12 or below per 1000 live births and to provide kangaroo mother care (KMC) or other humanized care techniques to at least 75% of eligible children.<sup>2</sup>

The KMC is an effective and safe option for clinically stable infants, especially in developing nations in Africa and South Asia, where 12% and 60% of preterm infants are born, respectively.<sup>1,2</sup>

In Bogota, Colombia, in 1978, Rey and Martinez established KMC to encourage the early discharge of LBW, preterm and SGA babies. Overcrowding, lack of equipment, absence of prepared specialists, and high cross-infection rates are some of the reasons for this concept.<sup>2</sup> The KMC is a skin-to-skin contact-based care method for LBW and/or preterm infants at birth that is structured and defined by a protocol. By transferring the ability and responsibility of being the primary caregivers of their newborns to the mothers and families, this program hopes to empower these mothers and families to meet their babies' physical and emotional needs. The KMC includes the following: skin-to-skin contact, breast-to-skin contact, early discharge from the hospital (ideally), social support, and follow-up when it comes to proving the efficacy of KMC.<sup>3-5</sup> Admission criteria for KMC units include the following; infants with low weight under 2500 g, under 37 weeks gestational age, oxygen-dependent infants with oxygen saturation >90%, or those receiving nasal prongs. Studies and clinical trials evaluating KMC have

several heterogeneity aspects such as weight of infants, which varied from <1500 g to 2500 g, time of KMC application, mode of application (continuous or intermittent), and length of the follow-up (1–12 months). A review of current evidence is needed to determine if KMC reduces LBW and preterm infants' mortality and the length of their hospital stay.

## Objectives

The current study aimed to assess and evaluate the impact of applying the KMC technique in preterm LBW infants care on clinical outcomes such as mortality rate and length of hospitalization.

## Materials and methods

### Study selection

The main goals of our meta-analysis were to evaluate the effect of KMC application in preterm neonates on neonatal mortality and length of hospitalization compared to controls. Mortality rates and hospital stays from published clinical trials were analyzed using statistical analysis tools such as frequency rate, odds ratio (OR), relative risk, or mean difference (MD) with a 95% confidence interval (95% CI).

Only papers in English were included. The criteria for the inclusion of articles in the current meta-analysis were not constrained by the size of the studies, while studies not describing practical interventions, such as letters to editors, and review articles were excluded from the study. The meta-analysis model is depicted in Fig. 1.

The effect of KMC in preterm LBW neonates on mortality and length of hospital stay was compared to controls by performing a sensitivity analysis for each subcategory.

The inclusion criteria were as follows: 1) Prospective, randomized controlled trials and retrospective investigations

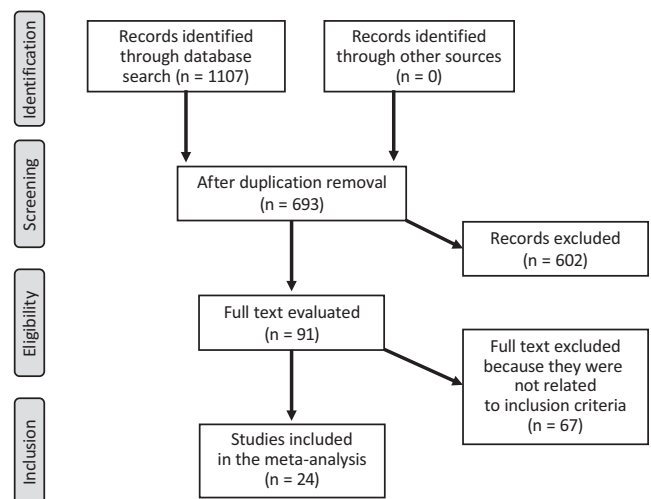


Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram reflecting the mode of meta-analysis

were all eligible to be included in the research; 2) Subjects of the studies included preterm/LBW neonates; 3) Interventions used were KMC compared to control receiving conventional care.

The exclusion criteria were as follows: 1) Trials that did not examine the influence of KMC on length of hospital stay or mortality in preterm neonates; 2) Studies lacking parameters such as weight of the neonates; 3) Studies that did not compare the impact of KMC on mortality or hospital stay to conventional care.

### Identification

A systematic deep literature search was accomplished using PubMed, Embase and Cochrane Library. The search was conducted among studies published between 1988 and 2021 using the keywords ‘kangaroo mother care’, ‘mortality’, ‘hospital stay’, and ‘preterm neonates’, as shown in Table 1. Studies were identified using PICOS framework as follows: P (population) – preterm/LBW neonates; I (intervention/exposure) – KMC; C (comparison): KMC compared to controls; O (outcome) – mortality and hospital stay; S (study design) – randomized clinical trials. To eliminate duplicates, the research studies were grouped using EndNote software (Clarivate Analytics, London, UK). Furthermore, all title and abstract data were subjected to a thorough review to eliminate any information that did not include any risk variables or the influence of KMC on preterm neonates.

### Screening

The subject-related and study data were collected using a standard format. The place of practice, first author’s surname, subject’s weight, design of the study, demographic data, sample size, primary outcome evaluation, treatment mode, duration of the study, categories, methods of statistical analysis, information source, and qualitative as well as quantitative evaluations were also collected using a traditional form.

Using the Cochrane Handbook for Systematic Reviews of Interventions v. 5.1, the Risk of Bias Tool was used to evaluate the quality of each study’s methodology.<sup>6</sup>

### Different levels of bias risk were used in the assessment

Three levels of bias risks were identified in the evaluation of the studies. When all quality criteria were met, the bias risk was judged low. When the criteria were not satisfied or only partially met, the risk was assessed as moderate. The risk was high in cases where all of the quality criteria were not met or included and when the paper also contained inconsistencies.

### Statistical analyses

Using the random and fixed-effect model, the OR with a 95% CI was calculated using a dichotomous technique for mortality at 2 months and 12 months, and

Table 1. Search strategy for each database

Database	Search strategy
PubMed	#1: "kangaroo mother care" [MeSH terms] OR "mortality" [MeSH terms] OR "hospital stay" [all fields] #2: "skin-to-skin contact" [MeSH terms] OR "preterm neonate" [all fields] #3: #1 AND #2
Embase	#1: 'kangaroo mother care'/exp OR 'hospital stay'/exp OR 'skin-to-skin contact'/exp #2: 'preterm neonate'/exp OR 'mortality'/exp #3: #1 AND #2
Cochrane Library	#1: (kangaroo mother care):ti,ab,kw OR (mortality):ti,ab,kw OR (hospital stay):ti,ab,kw (word variations have been searched) #2: (skin-to-skin contact):ti,ab,kw OR (preterm neonate):ti,ab,kw (word variations have been searched) #3: #1 AND #2

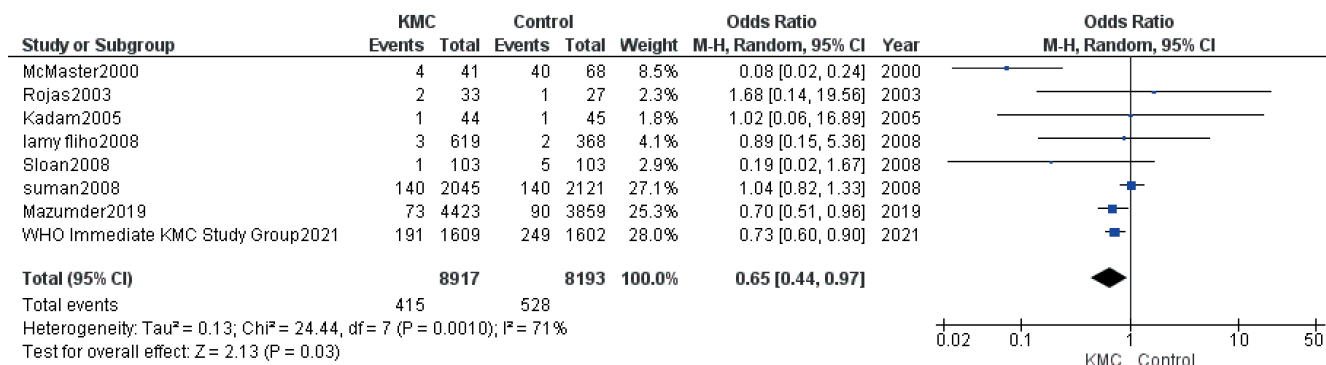


Fig. 2. A forest plot illustrating mortality rates after the application of kangaroo mother care (KMC) compared to controls in preterm neonates during the first 2 months of life

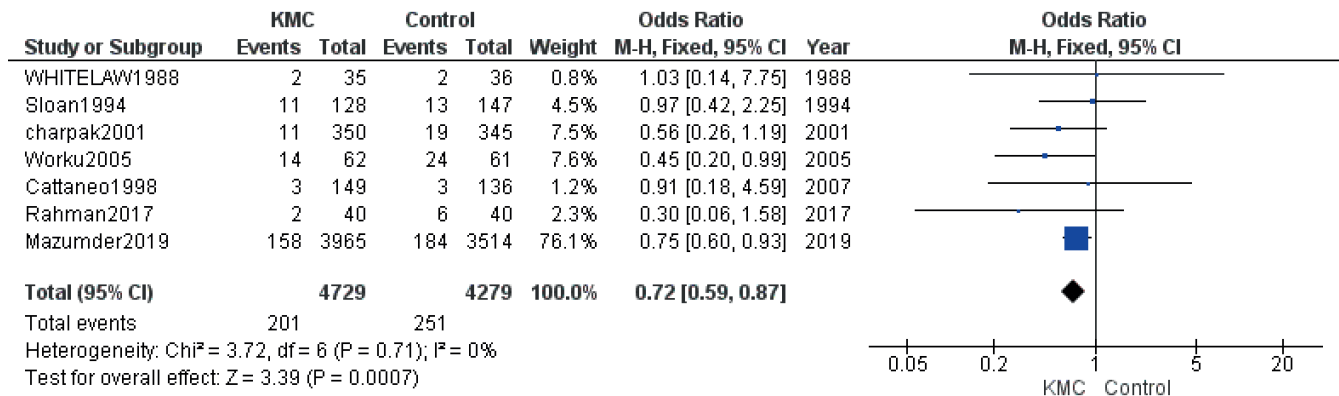


Fig. 3. A forest plot illustrating the mortality rate after the application of kangaroo mother care (KMC) compared to controls in preterm neonates during the 1<sup>st</sup> year of life

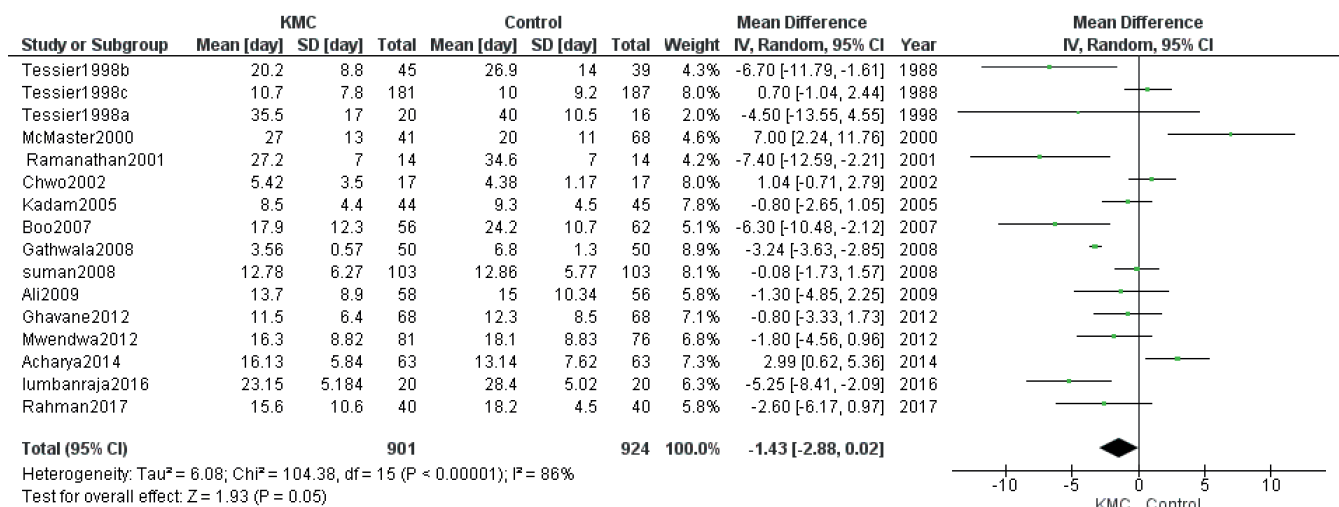


Fig. 4. A forest plot illustrating the length of hospital stay after the application of kangaroo mother care (KMC) compared to controls in preterm neonates

a continuous technique for assessing the length of hospital stay. The  $I^2$  index ranged from 0% to 100%. The heterogeneity scale was 4-point – 0%, 25%, 50%, and 75% – indicating no, low, moderate, and high heterogeneity, respectively. If the  $I^2$  value was 50% or higher, the random effect was taken into account, whereas the possibility of using a fixed influence increased in cases where  $I^2$  was less than 50%. However, other parameters that indicate a high similarity between included studies should be checked to ensure the usefulness of this model. The risk of publication bias was assessed using Begg's test in addition to visual evaluation of funnel plots. The p-values for the rank correlation and regression tests were generated using Jamovi v. 2.2.5 software (<https://www.jamovi.org/download.html>). Subgroup analysis was conducted to obtain more specific results for articles with 1 or more outcomes and a high similarity between the included articles. The statistical analysis was carried out using two-tailed p-values with Reviewer Manager v. 5.4.1 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark).

## Results

Among the 1107 unique reports, the current meta-analysis included 24 studies published between 1988 and 2021 (Table 2).<sup>7–30</sup> The studies chosen for this meta-analysis included 19,980 participants in total, of which 10,354 received KMC and 9626 were controls. The number of subjects included in each study ranged from 28 to 8282.

All studies evaluated the effect of KMC in preterm and LBW neonates on mortality and length of hospital stay. A total of 14 research studies compared data stratified by mortality rate (Fig. 2,3), and also in 14 studies, the data were stratified by length of hospital stay (Fig. 4). As shown in Fig. 2–4, when compared to the control group, the use of KMC in preterm LBW neonates resulted in significantly lower mortality (OR: 0.65, 95% CI: 0.44–0.97),  $p < 0.05$ ) in the short term (within 2 months,  $I^2 = 71%$ ) and long term (3–12 months,  $I^2 = 0%$ , OR: 0.72, 95% CI: 0.59–0.87,  $p < 0.001$ ), but had no significant impact on the length of hospital stay (OR: -1.43, 95% CI: -2.88–0.02,  $p = 0.05$ ,  $I^2 = 86%$ ).

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Acharya2014	+	?	-	?	+	+	+
Ali2009	+	?	-	?	+	-	-
Boo2007	+	+	-	?	-	+	+
Cattaneo2007	+	?	-	?	?	+	?
charpak2001	+	+	-	?	+	+	?
Chwo2002	+	+	-	?	+	+	?
Gathwala2008	+	?	-	?	?	+	+
Ghavane2012	+	+	-	?	+	+	+
Kadam2005	+	+	-	?	+	+	+
lamy filho2008	?	?	?	?	+	+	+
lumbanraja2016	+	?	-	?	?	+	+
Mazumder2019	+	+	-	?	+	+	+
McMaster2000	?	?	-	?	+	+	+
Mwendwa2012	+	?	-	-	+	+	?
Rahman2017	+	?	-	?	+	+	?
Ramanathan2001	+	?	-	?	+	?	+
Rojas2003	+	+	-	?	+	+	+
Sloan1994	+	?	-	?	+	?	?
Sloan2008	+	?	-	?	+	-	-
suman2008	+	+	-	?	-	+	?
Tessier1998a	+	+	-	?	+	+	?
WHITELAW1988	+	+	-	-	+	?	+
WHO Immediate KMC Study Group2021	+	+	?	+	+	+	+
Worku2005	+	?	-	-	?	-	+

Fig. 5. Risk of bias summary

Included studies did not report findings related to ethnicity and gender, hence these factors were not included in the current meta-analysis. Figure 5 depicts the possibility of bias in the included studies. We concluded that no study adequately covered all 7 domains. There was a wide range in quality between methods employed in the included studies. Insufficient methodological tools were found In a study carried out by Lamy Filho et al.<sup>23</sup> Bias was

a problem in 2 studies because of the selective reporting of outcomes.<sup>16,29</sup> We were unable to indentify clear findings in the vast majority of categories listed in the Methods section in the study by Ali et al., such as the severity of disease, sepsis, diarrhea (which can lead to pneumonia), aspiration (which can cause pneumonia), increase in body weight, and the feelings of the mother.<sup>16</sup> Primary outcomes such as neonatal mortality, weight, length, and head circumference at discharge and follow-up were also provided in that study; however, they were not sufficiently recorded. Some secondary outcomes specified in the Methods section of the remaining 3 trials were either not published<sup>10</sup> or were mentioned but not adequately described.<sup>7,28</sup> Block randomization was utilized in 1 study.<sup>16</sup> Using block randomization in an unblinded trial makes it possible to forecast future assignments if the assignments are provided to participants after they have been recruited. This is especially true when blocks are of a set size.

### Publication bias assessment

Publication bias was assessed for mortality and length of hospital stay. Regarding the length of hospital stay (Begg’s  $p = 0.09$ ), publication bias was evident as the  $p$ -value was low compared to the other groups. Funnel plots showed evidence of bias for both length of hospital stay and mortality at 2 months (Fig. 6B,C). On the other hand, no evidence of publication bias was noted for mortality during the first 2 months using the Begg’s test ( $p = 0.72$ ). Similarly, no evidence of publication bias was found for mortality at 12 months (Begg’s  $p = 0.77$ ) as seen on visual inspection of the funnel plot (Fig. 6B)

### Discussion

Our meta-analysis involved 24 studies with 19,980 participants consisting of 10,354 patients receiving KMC and 9626 as controls.<sup>7–30</sup> When compared to the control group, the application of KMC resulted in considerably lower mortality ( $p < 0.05$  and  $p < 0.001$  at 2 months and after 12 months, respectively). However, because some of the included studies had small sample sizes (7 studies had a sample size of less than 100 subjects), careful analysis of the results is required, and the need for future trials to confirm these findings and assess the impact of KMC is obvious.

The primary goal of the current study was to evaluate the impact of KMC in preterm and LBW neonates on mortality during the first 2 months and long-term mortality (3–12 months) as well as the length of hospital stay.

In 8 studies reporting on mortality during the first 2 months of the neonate’s life, a significant ( $p < 0.05$ ) decrease in mortality in the KMC group (OR: 0.65, 95 % CI: 0.44–0.97,  $I^2 = 71\%$ ) was recorded. In 7 studies reporting on mortality between 3 and 12 months of life, a greater

**Table 2.** Characteristics of the studies selected for the meta-analysis

Study and reference	Year	Country	Total	KMC	Control	Weight
Whitelaw et al. [7]	1988	UK	71	35	36	<1500
Tessier et al. [8]	1998	Colombia	488	246	242	all infants
McMaster et al. [9]	2000	Papua New Guinea	109	41	68	<1500
Ramanathan et al. [10]	2001	India	28	14	14	<1500
Chwo et al. [11]	2002	Taiwan	34	17	17	all infants
Kadam et al. [12]	2005	India	89	44	45	<1800
Boo et al. [13]	2007	Malaysia	118	56	62	<1501
Gathwala et al. [14]	2008	India	100	50	50	≤1800
Suman et al. [15]	2008	India	206	103	103	<2000
Ali et al. [16]	2009	India	114	58	56	≤1800
Ghavane et al. [17]	2012	India	136	68	68	<1500
Mwendwa et al. [18]	2012	Kenya	157	81	76	≤1750
Acharya et al. [19]	2014	Nepal	126	63	63	<2000
Lumbanraja et al. [20]	2016	Indonesia	40	20	20	<2500
Rahman et al. [21]	2017	Bangladesh	80	40	40	<1800
Rojas et al. [22]	2003	USA	60	33	27	≤1500
Lamy Filho et al. [23]	2008	Brazil	987	619	368	<1750
Sloan et al. [24]	2008	Ecuador	4166	2045	2121	all infants
Mazumder et al. [25]	2019	India	8282	4423	3859	1500–2250
WHO Immediate KMC Study Group [26]	2021	Ghana, India, Malawi, Nigeria, Tanzania	3211	1609	1602	<1800
Cattaneo et al. [27]	1998	Ethiopia, Indonesia, Mexico	285	149	136	<2000
Charpak et al. [30]	2001	Colombia	695	350	345	<2000
Worku et al. [29]	2005	Ethiopia	123	62	61	<2000
Sloan et al. [28]	1994	Bangladesh	275	128	147	<2000
Total			19,980	10,354	9626	–

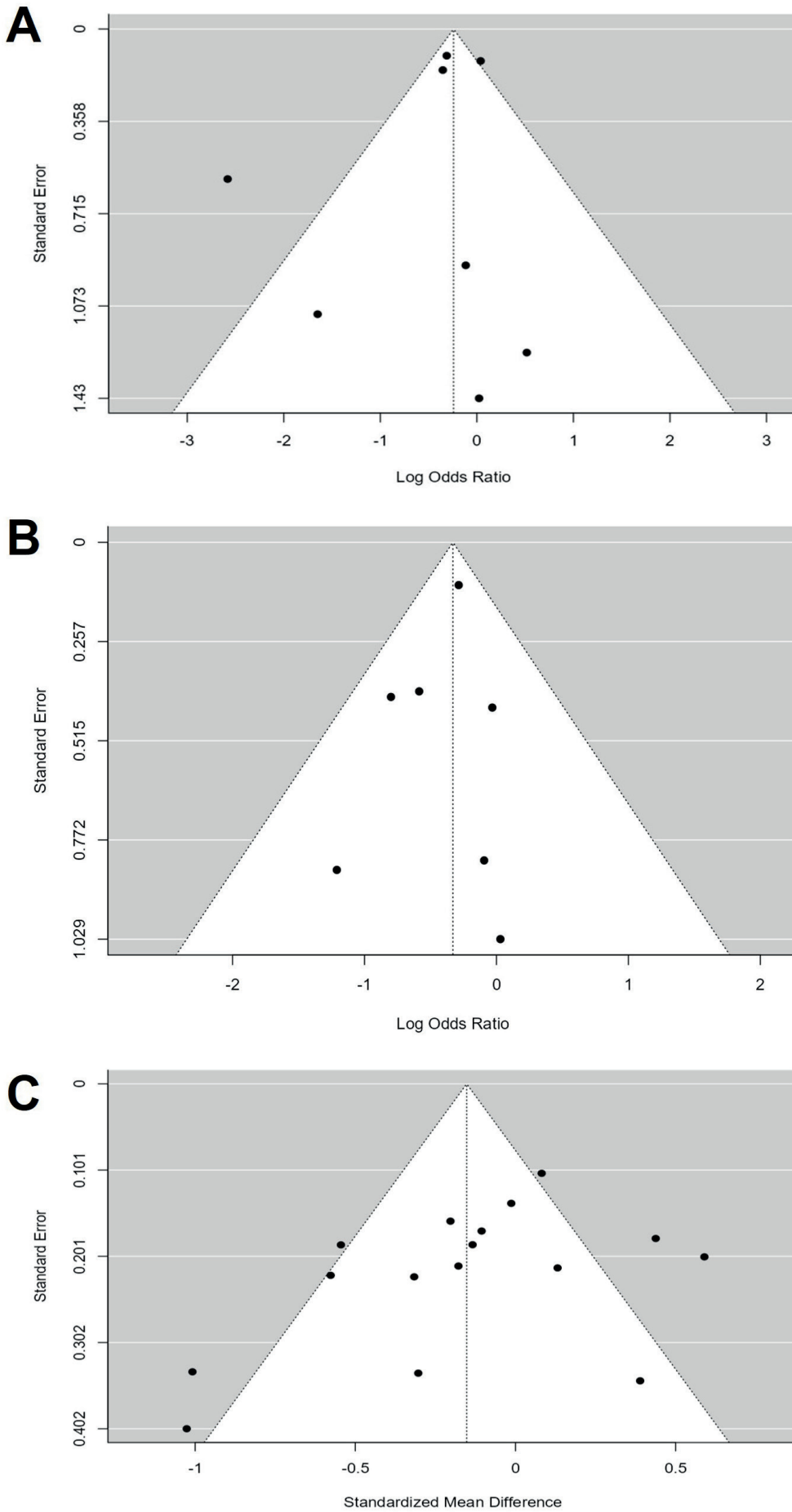
KMC – kangaroo mother care.

reduction in mortality ( $p < 0.001$ ) was seen in the KMC group compared to controls (OR: 0.72, 95% CI: 0.59–0.87,  $I^2 = 0\%$ , Fig. 2,3). Regarding the length of hospital stay, there was no significant difference ( $p = 0.05$ ) between both groups of neonates (OR: –1.43, 95% CI: –2.88–0.02,  $I^2 = 86\%$ ; Fig. 4). Although the current meta-analysis found no significant impact of KMC on the length of hospital stay, additional studies are needed to more accurately evaluate this factor. No significant variations in death rates were found in regard to the region of the world, country and level of economic development. The reduction in newborn mortality rate was found to be a result of KMC use.

A modest reduction in neonatal morbidity, improved quality of mother-to-child bonding, and shorter hospital stay with lower expenses are all reasons for the use of KMC in LBW children. If a newborn care facility is unavailable, some researchers believe that KMC is the best alternative, and in overpopulated nurseries, the use of KMC could free up incubators for sicker infants.<sup>31</sup>

Six studies tested the impact of KMC on the length of hospital stay ( $n = 512$ ) in LBW (<1501 g) neonates

and found no significant difference ( $p > 0.05$ ) between the KMC and the control group. This non-significant impact could be due to the high heterogeneity of the studies ( $I^2 = 86\%$ ). The KMC group had a longer hospital stay according to Acharya et al.<sup>19</sup> However, this could have been a result of the discharge criteria used in that study (weight >1600 g), as well as the fact that babies in the control group weighed more than those in the intervention group. Studies with longer follow-up periods reported more significant findings reflecting the positive impact of using KMC. The KMC employed for more than 6 h had no significant impact on the clinical status of included subjects. In addition, KMC was linked to a statistically non-significant decrease in the probability of apnea compared to standard treatment. The number of subjects per study had a great impact on the conclusion of 2 studies included in our review,<sup>25,26</sup> and had more than a 50% impact on KMC use on mortality during the 1<sup>st</sup> month after birth. Heterogeneity ( $I^2 = 0$ ) was absent when evaluating the impact of KMC compared to the control group on mortality during the 1<sup>st</sup> year after birth.



**Fig. 6.** Funnel plot showing the publication bias of the included studies for mortality during the first 2 months (A), 1<sup>st</sup> year (B) as well as for the length of hospital stay (C)

Finally, the KMC technique has been shown to have a beneficial impact on preterm/LBW neonates regarding clinical outcomes such as mortality rates, which were significantly lower than when conventional care technique was employed. However, no positive impact of KMC was observed on the length of hospitalization in preterm/LBW neonates.

## Limitations

One of the drawbacks of the present study was that there were many biases as numerous papers were eliminated from the current meta-analysis because they did not match the inclusion criteria. Furthermore, there was some skepticism about how to include ethnicity, economical level and gender of the neonates into this study. Because only data from prior studies were used, the analysis may have been skewed because of missing information. The meta-analysis included 24 studies, 7 of which were small (under 100 participants). The incomplete presentation of finding in the included articles could be the source of bias in studies include in the present meta-analysis. In addition, the participants' weight varied among different studies.

## Conclusions

This meta-analysis showed that the KMC technique has been proven to have a beneficial impact on preterm/LBW neonates regarding clinical outcomes such as mortality rates, which were significantly lower than in neonates receiving conventional care. However, no positive impact of KMC was noted on the length of hospitalization. Additionally, our meta-analysis was unable to demonstrate a connection between the results of analyzed studies and population characteristics such as economic level of the country of residence, ethnicity and gender. Because of the small sample size in some of the studies, additional investigations are needed to validate these findings and increase confidence in the effects of KMC.

## ORCID iDs

Wanying Guo  <https://orcid.org/0000-0002-5638-4686>

## References

- World Health Organization. *Survive and Thrive: Transforming Care for Every Small and Sick Newborn*. Geneva, Switzerland: World Health Organization; 2019. <https://apps.who.int/iris/handle/10665/326495>. Accessed September 1, 2022.
- Nyqvist K; Expert Group of the International Network on Kangaroo Mother Care. Kangaroo mother care: Application in a high-tech environment. *Acta Paediatr*. 2010;99(6):812–819. doi:10.1111/j.1651-2227.2010.01794.x
- Pillai Riddell RR, Racine NM, Turcotte K, et al. Non-pharmacological management of infant and young child procedural pain. *Cochrane Database Syst Rev*. 2011;5(10):CD006275. doi:10.1002/14651858.CD006275.pub2
- Johnston C, Campbell-Yeo M, Fernandes A, Inglis D, Streiner D, Zee R. Skin-to-skin care for procedural pain in neonates. *Cochrane Database Syst Rev*. 2014;23(1):CD008435. doi:10.1002/14651858.CD008435.pub2
- Conde-Agudelo A, Díaz-Rossello JL. Kangaroo mother care to reduce morbidity and mortality in low birthweight infants. *Cochrane Database Syst Rev*. 2014;16(3):CD002771. doi:10.1002/14651858.CD002771.pub2
- Higgins JP, Altman DG, Gøtzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. doi:10.1136/bmj.d5928
- Whitelaw A, Heisterkamp G, Sleath K, Acolet D, Richards M. Skin to skin contact for very low birthweight infants and their mothers. *Arch Dis Child*. 1988;63(11):1377–1381. doi:10.1136/adc.63.11.1377
- Tessier R, Cristo M, Velez S, et al. Kangaroo mother care and the bonding hypothesis. *Pediatrics*. 1998;102(2):e17. doi:10.1542/peds.102.2.e17
- McMaster P, Haina T, Vince JD. Kangaroo care in Port Moresby, Papua New Guinea. *Trop Doct*. 2000;30(3):136–138. doi:10.1177/004947550003000307
- Ramanathan K, Paul VK, Deorari AK, Taneja U, George G. Kangaroo mother care in very low birth weight infants. *Indian J Pediatr*. 2001;68(11):1019–1023. doi:10.1007/BF02722345
- Chwo MJ, Anderson GC, Good M, Dowling DA, Shiau SHH, Chu DM. A randomized controlled trial of early kangaroo care for preterm infants: Effects on temperature, weight, behavior, and acuity. *J Nurs Res*. 2002;10(2):129–142. doi:10.1097/01.JNR.0000347592.43768.46
- Kadam S, Binoy S, Kanbur W, Mondkar JA, Fernandez A. Feasibility of kangaroo mother care in Mumbai. *Indian J Pediatr*. 2005;72(1):35–38. doi:10.1007/BF02760578
- Boo NY, Jamli FM. Short duration of skin-to-skin contact: Effects on growth and breastfeeding. *J Paediatr Child Health*. 2007;43(12):831–836. doi:10.1111/j.1440-1754.2007.01198.x
- Gathwala G, Singh B, Balhara B. KMC facilitates mother baby attachment in low birth weight infants. *Indian J Pediatr*. 2008;75(1):43–47. doi:10.1007/s12098-008-0005-x
- Suman RPN, Udani R, Nanavati R. Kangaroo mother care for low birth weight infants: A randomized controlled trial. *Indian Pediatr*. 2008;45(1):17–23. PMID:18250500.
- Ali S, Sharma J, Sharma R, Alam S. Kangaroo mother care as compared to conventional care for low birth weight babies. *Dicle Med J*. 2009;36(3):155–160. <https://dergipark.org.tr/tr/download/article-file/54001>. Accessed March 12, 2022.
- Ghavane S, Murki S, Subramanian S, Gaddam P, Kandraj H, Thumalla S. Kangaroo mother care in kangaroo ward for improving the growth and breastfeeding outcomes when reaching term gestational age in very low birth weight infants. *Acta Paediatr*. 2012;101(12):e545–e549. doi:10.1111/apa.12023
- Simiyu D. Neonatal septicaemia in low birth weight infants at Kenyatta National Hospital, Nairobi. *East Afr Med J*. 2005;82(3):148–152. doi:10.4314/eamj.v82i3.9272
- Acharya N, Singh R, Bhatta N, Poudel P. Randomized control trial of kangaroo mother care in low birth weight babies at a tertiary level hospital. *J Nepal Paediatr Soc*. 2014;34(1):18–23. doi:10.3126/jnps.v34i1.8960
- Lumbanraja SN. Influence of maternal factors on the successful outcome of kangaroo mother care in low birth-weight infants: A randomized controlled trial. *J Neonatal Perinatal Med*. 2016;9(4):385–392. doi:10.3233/NPM-161628
- Rahman M. Kangaroo mother care for low birth weight babies: A randomized controlled trial in a tertiary care hospital of Bangladesh. *J Pediatr Neonatal Care*. 2017;7(2):00285. doi:10.15406/jpnc.2017.07.00285
- Rojas MA, Kaplan M, Quevedo M, et al. Somatic growth of preterm infants during skin-to-skin care versus traditional holding: A randomized controlled trial. *J Dev Behav Pediatr*. 2003;24(3):163–168. doi:10.1097/00004703-200306000-00006
- Lamy Filho F, Silva AAM da, Lamy ZC, et al. Evaluation of the neonatal outcomes of the kangaroo mother method in Brazil. *J Pediatr (Rio J)*. 2008;84(5):428–435. doi:10.2223/JPED.1821
- Sloan NL, Ahmed S, Mitra SN, et al. Community-based kangaroo mother care to prevent neonatal and infant mortality: A randomized, controlled cluster trial. *Pediatrics*. 2008;121(5):e1047–e1059. doi:10.1542/peds.2007-0076
- Mazumder S, Taneja S, Dube B, et al. Effect of community-initiated kangaroo mother care on survival of infants with low birthweight: A randomised controlled trial. *Lancet*. 2019;394(10210):1724–1736. doi:10.1016/S0140-6736(19)32223-8



26. Lakshamma VT, R PV, Thayavathi P, B KN. Survive and thrive: Transforming care for every premature newborn during Covid-19. *J Perinatal Pediatr Neonatal Nurs.* 2022;4(1):1–6. <http://matjournals.co.in/index.php/JPPNN/article/view/71>. Accessed March 12, 2022.
27. Cattaneo A, Davanzo R, Worku B, et al. Kangaroo mother care for low birthweight infants: A randomized controlled trial in different settings. *Acta Paediatr.* 2007;87(9):976–985. doi:10.1111/j.1651-2227.1998.tb01769.x
28. Sloan NL, Rojas EP, Stern C, Camacho LWL. Kangaroo mother method: Randomised controlled trial of an alternative method of care for stabilised low-birthweight infants. *Lancet.* 1994;344(8925):782–785. doi:10.1016/S0140-6736(94)92341-8
29. Worku B, Kassie A. Kangaroo mother care: A randomized controlled trial on effectiveness of early kangaroo mother care for the low birth-weight infants in Addis Ababa, Ethiopia. *J Trop Pediatr.* 2005;51(2): 93–97. doi:10.1093/tropej/fmh085
30. Charpak N, Ruiz-Peláez JG, Figueroa de CZ, Charpak Y. A randomized controlled trial of kangaroo mother care: Results of follow-up at 1 year of corrected age. *Pediatrics.* 2001;108(5):1072–1079. doi:10.1542/peds.108.5.1072
31. Ruiz-Peláez JG, Charpak N, Cuervo LG. Kangaroo mother care: An example to follow from developing countries. *BMJ.* 2004;329(7475): 1179–1181. doi:10.1136/bmj.329.7475.1179



# Machine learning model-based risk prediction of severe complications after off-pump coronary artery bypass grafting

Yang Zhang<sup>1,A–D,F</sup>, Lin Li<sup>1,A,B,D–F</sup>, Ye Li<sup>2,A,F</sup>, Zhihe Zeng<sup>1,B,C,F</sup>

<sup>1</sup> Department of Anesthesiology, General Hospital of Northern Theater Command, Shenyang, China

<sup>2</sup> Microsoft (China) Co., Ltd, Beijing, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):185–194

## Address for correspondence

Lin Li  
E-mail: lilinlashofmine@qq.com

## Funding sources

This work was supported by the Key R&D Guiding Program of Liaoning Province 2019010209 (grant No. 2019JH8/10300083).

## Conflict of interest

None declared

Received on April 14, 2022

Reviewed on June 10, 2022

Accepted on August 23, 2022

Published online on October 13, 2022

## Cite as

Zhang Y, Li L, Li Y, Zeng Z. Machine learning model-based risk prediction of severe complications after off-pump coronary artery bypass grafting. *Adv Clin Exp Med.* 2023;32(2):185–194. doi:10.17219/acem/152895

## DOI

10.17219/acem/152895

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Compared with coronary artery bypass grafting (CABG) under cardiopulmonary bypass, off-pump coronary artery bypass (OPCAB) is minimally invasive and reduces the risk of intraoperative blood transfusion and acute kidney injury. Nonetheless, OPCAB-related complications still pose a threat. Machine learning technology can analyze a large number of clinical data, establish risk prediction models and help clinicians make early and correct clinical decisions.

**Objectives.** Risk prediction models are available for mortality and morbidity after cardiac surgery, but they are not specific to OPCAB. This study aimed to develop a predictive model of severe complications after OPCAB, based on machine learning.

**Materials and methods.** Anesthesia records of OPCAB from the General Hospital of the Northern Theater Command (Shenyang, China) collected between January 1, 2019, and June 15, 2020, were analyzed. The endpoint of the study was the occurrence of serious complications after OPCAB (postoperative unplanned intra-aortic balloon pump, secondary surgery and death). The features entered into the models were as follows: intraoperative ventricular fibrillation, number of saphenous vein grafts, nerve block (NeB), venous oxygen saturation (SvO<sub>2</sub>), skin incision-bypass time, and hypertension. A total of 8 machine learning algorithms were tested: logistic regression analysis (LRA), k-nearest neighbor (KNN), naïve Bayes (NB), support vector machine (SVM), random forest (RF), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical features gradient boosting (CatBoost).

**Results.** Among the 506 patients found in the records, 27 met the endpoint. The highest area under the curve (AUC) value was achieved with the XGBoost model (AUC = 0.94), and the lowest with the SVM model (AUC = 0.75). The highest and lowest accuracy were observed with the XGBoost and NB models, respectively, while the highest and lowest precision were achieved using the SVM and NB models, respectively. Based on the receiver operating characteristic (ROC) curves, the XGBoost model was selected as the most useful in this study.

**Conclusions.** This study suggests using the XGBoost model to predict the risk of complications after OPCAB.

**Key words:** complications, machine learning, off-pump coronary artery bypass grafting, prediction model

## Background

Revascularization is paramount to the management of acute coronary syndrome (ACS). It aims to improve blood flow to the myocardium<sup>1</sup> and is performed using percutaneous coronary intervention or coronary artery bypass graft (CABG).<sup>1–3</sup> The latter can be performed either off-pump (i.e., without the assistance of a heart-lung machine) or on-pump. On-pump CABG is associated with more severe surgical trauma, while off-pump coronary artery bypass (OPCAB) can reduce perioperative bleeding and allogeneic blood transfusions, as well as reduce the risk of acute kidney injury (AKI) in patients with kidney dysfunction.<sup>4</sup> The OPCAB does not appear to increase 30-day mortality compared with on-pump CABG, but an extensive systematic review of observational studies suggested that OPCAB might reduce short-term mortality.<sup>5,6</sup> Therefore, OPCAB is probably a good option for selected patients.<sup>7,8</sup>

There are still some risks related to the use of OPCAB,<sup>4,9,10</sup> including perioperative complications such as mortality, stroke, kidney failure, respiratory failure, and blood loss.<sup>11–13</sup> Furthermore, OPCAB appears to be associated with higher 10-year rates of incomplete revascularization, repeat revascularization and mortality, compared with on-pump CABG.<sup>10</sup> Additionally, OPCAB is associated with increased adverse events at 1 year and mortality at 5 years.<sup>14,15</sup> Although OPCAB has a similar risk of myocardial infarction compared to on-pump CABG, the data are inconsistent for the risk of stroke.<sup>5,6,16</sup> A decreased left ventricular ejection fraction (LVEF) is observed in about 22% of the patients after OPCAB and can compromise their short- and long-term outcomes.<sup>17–19</sup>

Some tools are available for estimating the risk of mortality and morbidity after CABG. The Society of Thoracic Surgeons (STS) score can be used to determine the risk of mortality and morbidity after cardiac surgery, but it is not specific to OPCAB.<sup>20</sup> The European System for Cardiac Operative Risk Evaluation (EuroSCORE) can overestimate the risk of complications in the highest-risk and lowest-risk patients undergoing CABG,<sup>21</sup> as well as in patients undergoing OPCAB.<sup>22</sup> Other risk models are available but they are not specific to OPCAB.<sup>23–25</sup>

Machine learning algorithms can be used to analyze data and establish risk models more accurately than traditional statistical models.<sup>26,27</sup> Indeed, machine learning has been used to create models that predict mortality after cardiac surgery,<sup>28–31</sup> as well as estimate the length of hospital stay after CABG.<sup>32</sup> Regardless, these models are still not specific to OPCAB.

## Objectives

Using machine learning, this study aimed to build a predictive model for the detection of early postoperative serious complications after OPCAB. The results could

provide reference data for optimizing the clinical pathway for OPCAB, and anesthesia strategy to maintain vital signs, regulate the circulation, balance myocardial oxygen supply and demand, and reduce complications.

## Materials and methods

### Study design

All data were taken from the Do-care anesthesia record system (v. 5.0, MEDICAL SYSTEM Co., Ltd., Suzhou, China). All records of OPCABs performed from January 1, 2019, to June 15, 2020, at the 2<sup>nd</sup> Ward of the Department of Anesthesiology of the General Hospital of the Northern Theater Command (Shenyang, China) were included. At that hospital, OPCAB has been carried out for 20 years. In this study, 3 teams comprised of 15 surgeons were involved, who all had the qualification of chief surgeon, with an annual operation volume of 300–450 cases per surgeon.

This study was approved by the Ethics Committee of the hospital (Approval No. k(2020)01). The requirement for individual informed consent was waived by the Committee due to the retrospective nature of this study.

### Inclusion and exclusion criteria

All patients who underwent CABG were screened. The exclusion criteria were: 1) CABG under cardiopulmonary bypass; 2) CABG combined with other surgical procedures; 3) cancellation of the operation; or 4) intraoperative change of the initial surgical plan (e.g., intraoperative decision for valve replacement).

### Data collection and definitions

Demographic data (sex, age and body mass index (BMI)), data on comorbidities and intraoperative parameters (heart rate (HR), mean arterial pressure (MAP), respiratory rate, and mixed venous oxygen saturation (SvO<sub>2</sub>)) were collected retrospectively. The SvO<sub>2</sub> of the patients was continuously measured and recorded using a Swan–Ganz catheter (Edward Company, Irvine, USA) placed through the internal jugular vein.

The endpoint was the occurrence of serious complications after OPCAB, defined as postoperative unplanned intra-aortic balloon pump (IABP) assistance, secondary surgery (e.g., thoracotomy and repeat revascularization), intraoperative emergency conversion to on-pump CABG, and death. Revascularization was defined as revascularization for acute graft failure during the same hospital stay and emergency revascularization for bleeding (i.e., hemorrhagic shock caused by bleeding from the anastomotic site of the transplanted blood vessel during postoperative hospitalization, with repeat revascularization after 2 emergency operations). A patient in whom any of the above

events occurred after OPCAB and before discharge from the hospital was considered to have met the endpoint.

## Feature selection and model evaluation

The core principle of model feature screening was based on the feature importance of the machine learning model, combined with Pearson's correlation analysis and statistical analysis of the difference. The specific implementation was as follows:

1. Four algorithms with characteristically important parameters were selected: logical regression analysis (LRA), support vector machine (SVM), random forest (RF), and extreme gradient boosting (XGBoost).
2. The feature importance of the standard features was calculated and ranked based on the above 4 models.
3. The top 10 features of each model were selected as the feature groups (a total of 4 feature groups).
4. Pearson's correlation analysis was carried out on the standard features, and the features with the top 10 correlation coefficients were selected as the feature group.
5. The  $\chi^2$  test was performed on the standard features, and the features with statistically significant differences ( $p < 0.05$ ) were selected as the feature group.
6. All 6 feature groups were compared, and the features that appeared 4 times or more were selected as the main features.
7. Finally, Pearson's correlation coefficient was used to distinguish the variables that might affect the endpoint.

After removing meaningless features (Supplementary Table 1), interpolating missing values (Supplementary Table 2), discretizing the numerical variables, and selecting the features, the remaining 6 features were entered into the model. The features were intraoperative ventricular fibrillation (VF), number of saphenous vein grafts (SVG), nerve block (NeB), mixed venous oxygen saturation, skin incision-bypass time (T1), and hypertension (HBP). Results of the tests used for feature selection are shown in Supplementary File 1 and Supplementary Figure 3. Results of verifying the assumptions for the application of the preferred tests are shown in Supplementary Tables 3–10.

"Simpleimputer" in the "sklearn" module was used for the interpolation of the missing data. Mean interpolation was used for numerical variables and mode interpolation for binary and hierarchical variables. Meaningless features were first deleted, and the remaining features were interpolated one by one according to the characteristics and distribution of each feature, rather than based on the 54 features.

All numerical features were discretized by segmentation of continuous numerical data into discrete intervals. The segmentation principle was based on equal frequency, equal distance or optimization methods. Data discretization is also required by many algorithms, since discretization can speed up model training and enhance the robustness of the model by converting the continuous variables

into category variables through discretization. In order to unify the characteristic segmentation of different dimensions, this study used the mean  $\pm$  standard deviation ( $M \pm SD$ ) as the segmentation principle. Specifically, all values of the characteristic column were divided into 4 segments according to the nodes of  $M-1 \times SD$  or  $M$  or  $M+1 \times SD$ , and each segment was marked as 0, 1, 2, or 3. All features were defined as standard features after meaningless feature removal and numerical feature discretization.

Eight machine learning algorithms were tested in this study: LRA, k-nearest neighbor (KNN), naïve Bayes (NB), SVM, RF, XGBoost, light gradient boosting machine (LightGBM), and categorical features gradient boosting (Catboost).

## Statistical analyses

Statistical analysis was performed using the SciPy v. 1.4.1 scientific computing module within the Python 3.8 environment (<https://pypi.org/project/scipy/1.4.1/>). Data were assessed for normality using the Shapiro–Wilk test. Continuous data conforming to a normal distribution were presented as  $M \pm SD$  and analyzed using the independent samples t test. Those not conforming to a normal distribution were presented as median (range) and analyzed using the Mann–Whitney U test. Categorical data were presented as n (%) and analyzed using the  $\chi^2$  test. Correlation analyses were performed using the Pearson's analysis. Training and validation sets were divided using k-fold cross-validation. The k-fold module in sklearn ([https://scikit-learn.org/stable/modules/generated/sklearn.model\\_selection.KFold.html](https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html)) was used to randomly divide the database into 5 equal and non-overlapping groups, and the proportion of negative and positive samples in each group was the same. Each time, 4 groups were used as the training set, and 1 group was used as the validation set for model training verification. Precision, recall, F1-score (combining precision and recall into one metric by calculating the harmonic mean between those two<sup>33</sup>), and the area under the curve (AUC) were calculated. The above process was performed 5 times to ensure that each group was used as the validation set. Each time, the model was retrained and validated to avoid overfitting, and the average score of the 5 cross-validations was used as the final performance score of the model. The value of  $p < 0.05$  was considered statistically significant.

## Results

### Patient selection

Figure 1 presents the patient selection process. Among the 11,495 patients included in the database, 1238 underwent CABG, and 506 from them were selected based on the eligibility criteria. They were then divided into the training set ( $n = 405$ ) and the validation set ( $n = 101$ ). Table 1 presents the characteristics of the patients. Among the 506 patients

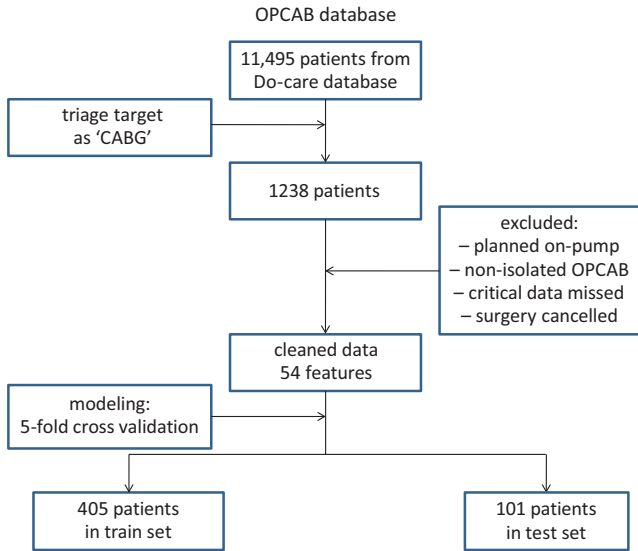


Fig. 1. Screening process of the patients. The patients were selected from the off-pump coronary artery bypass (OPCAB) graft database of the General Hospital of the Northern Theater Command (Shenyang, China)

CABG – coronary artery bypass grafting.

chosen, 27 met the endpoint (positive group), including postoperative emergency IABP assistance (n = 10), secondary surgery (n = 2), death without other outcomes (n = 3), postoperative emergency IABP assistance with secondary surgery (n = 2), death after postoperative emergency IABP (n = 6), death after secondary surgery (n = 2), and death after postoperative emergency IABP assistance and secondary surgery (n = 2). The in-hospital mortality rate was 2.6% (13/506). Compared with the controls, the patients who met the endpoint had a lower LVEF (52 ±7% compared to 55 ±6%, p = 0.027), lower fractional shortening (26 ±5% compared to 28 ±4%, p = 0.013), worse New York Heart Association (NYHA) classification score (p = 0.041), lower frequency of preoperative diabetes mellitus (DM; 22% compared to 46%, p = 0.016), lower intraoperative urine output

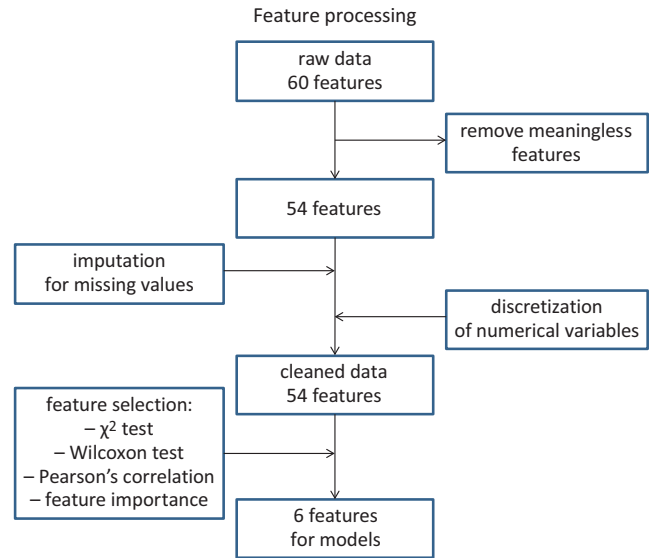


Fig. 2. Feature selection process

(639 ±2512 mL compared to 771 ±426 mL, p = 0.018), shorter T1 (62.0 ±17.8 min compared to 69.3 ±18.7 min, p = 0.047), higher SvO<sub>2</sub> values (74% compared to 53%, p = 0.036), and a smaller numbers of grafts (p < 0.001). Unilateral (left) internal mammary artery to the anterior descending artery anastomosis was performed in all patients. Radial arteries or other arteries were not used as graft vessels.

### Feature selection

Figure 2 presents the feature selection process. From the initial 60 features, 6 were removed due to meaninglessness. After missing data imputation and discretization of the continuous variables, 54 clean features were tested, and 6 were retained (intraoperative VF, number of SVG, NeB, SvO<sub>2</sub>, T1, and HBP). The results of the correlation analysis of the 6 features are presented in Fig. 3.

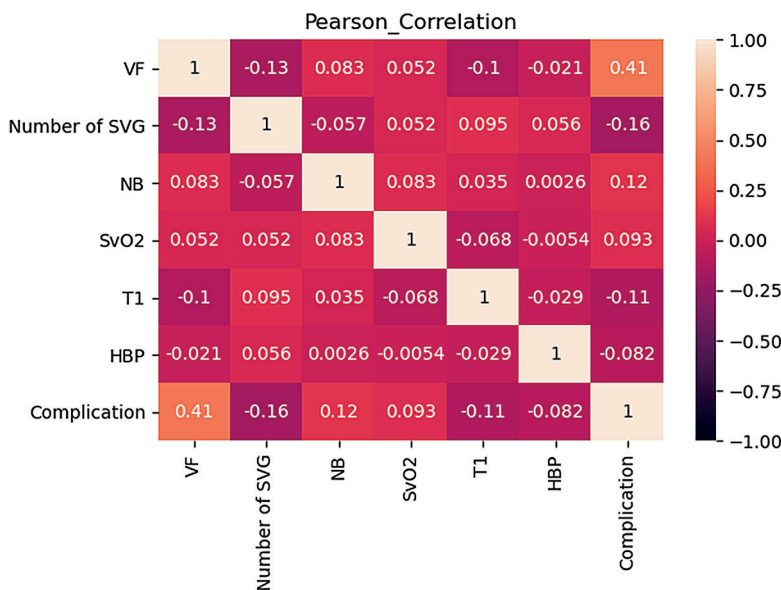


Fig. 3. Correlation analysis among the characteristic variables. Pearson's correlation coefficient was used to distinguish the variables that might affect the endpoint. Finally, these 6 variables were selected to be included in the final model after reordering

VF – ventricular fibrillation; NeB – nerve block; SvO<sub>2</sub> – mixed venous oxygen saturation; T1 – skin incision-bypass time; HBP – high blood pressure; SVG – saphenous vein grafts.

**Table 1.** Characteristics of the patients with and without serious complications after off-pump coronary artery bypass (OPCAB)

Characteristic	Patients with complications (n = 27)	Patients without complications (n = 479)	Test	Test result	p-value
Age [years], M ±SD	63.3 ±8.6	64.1 ±7.3	Student's t	0.534	0.594
Males, n (%)	22 (81)	362 (76)	χ <sup>2</sup>	0.440	0.507
BMI [kg/m <sup>2</sup> ], M ±SD	23.9 ±2.6	24.7 ±3.0	Student's t	1.295	0.196
LVD [mm], M ±SD	49.6 ±11.6	46.0 ±4.4	Student's t	-1.553	0.132
LVEF (%), M ±SD	52 ±7	55 ±6	Student's t	2.337	0.027
FS (%), M ±SD	26 ±5	28 ±4	Student's t	2.658	0.013
Emergency surgery, n (%)	0 (0)	22 (5)	Fisher's	0.427	0.513
NYHA					
I	0 (0)	5 (1)	Wilcoxon rank sum	5961.0	0.041
II	24 (89)	454 (95)			
III	3 (11)	18 (4)			
IV	0 (0)	2 (0)			
Preoperative comorbidities/risk factors, n (%)					
HBP	14 (52)	330 (69)	χ <sup>2</sup>	3.410	0.065
DM	6 (22)	220 (46)	χ <sup>2</sup>	5.812	0.016
CVD	4 (15)	98 (20)	χ <sup>2</sup>	0.216	0.642
AF	0 (0)	7 (1)	Fisher's	0.046	0.830
AMI	5 (19)	81 (17)	χ <sup>2</sup>	0.047	0.829
PCI	6 (22)	80 (17)	χ <sup>2</sup>	0.552	0.457
LM	9 (33)	114 (24)	χ <sup>2</sup>	1.263	0.261
RCA	24 (89)	435 (91)	χ <sup>2</sup>	0.000	0.996
Intraoperative parameters					
HR	62 ±14	61 ±13	Student's t	-0.473	0.636
MAP [mm Hg], M ±SD	76 ±12	75 ±12	Student's t	-0.571	0.568
mPAP [mm Hg]	20 ±5	19 ±5	Student's t	-1.460	0.160
TP [°C], M ±SD	36.2 ±0.4	36.2 ±0.6	Student's t	-0.079	0.937
CI [L/min × m <sup>2</sup> ], M ±SD	1.90 ±0.40	1.95 ±0.51	Student's t	1.495	0.606
ACCT [min], M ±SD	13.9 ±5.8	15.9 ±6.3	Student's t	1.649	0.100
Intraoperative transfusion [mL], M ±SD	1791 ±798	1652 ±671	Student's t	-1.037	0.300
Urine [mL], M ±SD	639 ±2512	771 ±426	Student's t	2.473	0.018
T1 [min]	62.0 ±17.8	69.3 ±18.7	Student's t	1.749	0.047
SvO <sub>2</sub> < 75%, n (%)	20 (74)	256 (53)	χ <sup>2</sup>	-0.079	0.036
SVG, n (%)					
1	1 (4)	1 (0)	Wilcoxon rank sum	4301.0	<0.001
2	11 (41)	76 (16)			
3	13 (48)	309 (64)			
4	2 (7)	90 (19)			
5	0 (0)	3 (1)			

BMI – body mass index; LVD – left ventricular diameter; LVEF – left ventricular ejection fraction; FS – fractional shortening; NYHA – New York Heart Association; HBP – high blood pressure; DM – diabetes mellitus; CVD – cardiovascular disease; AF – atrial fibrillation; AMI – acute myocardial infarction; PCI – percutaneous coronary intervention; LM – left main artery; RCA – right coronary artery; HR – heart rate; MAP – mean arterial pressure; mPAP – mean pulmonary arterial pressure; TP – temperature; CI – cardiac index; ACCT – aortic cross-clamp time; SvO<sub>2</sub> – mixed venous oxygen saturation; T1 – skin incision-bypass time; SVG – saphenous vein grafts; M ±SD – mean ± standard deviation.

## Algorithms

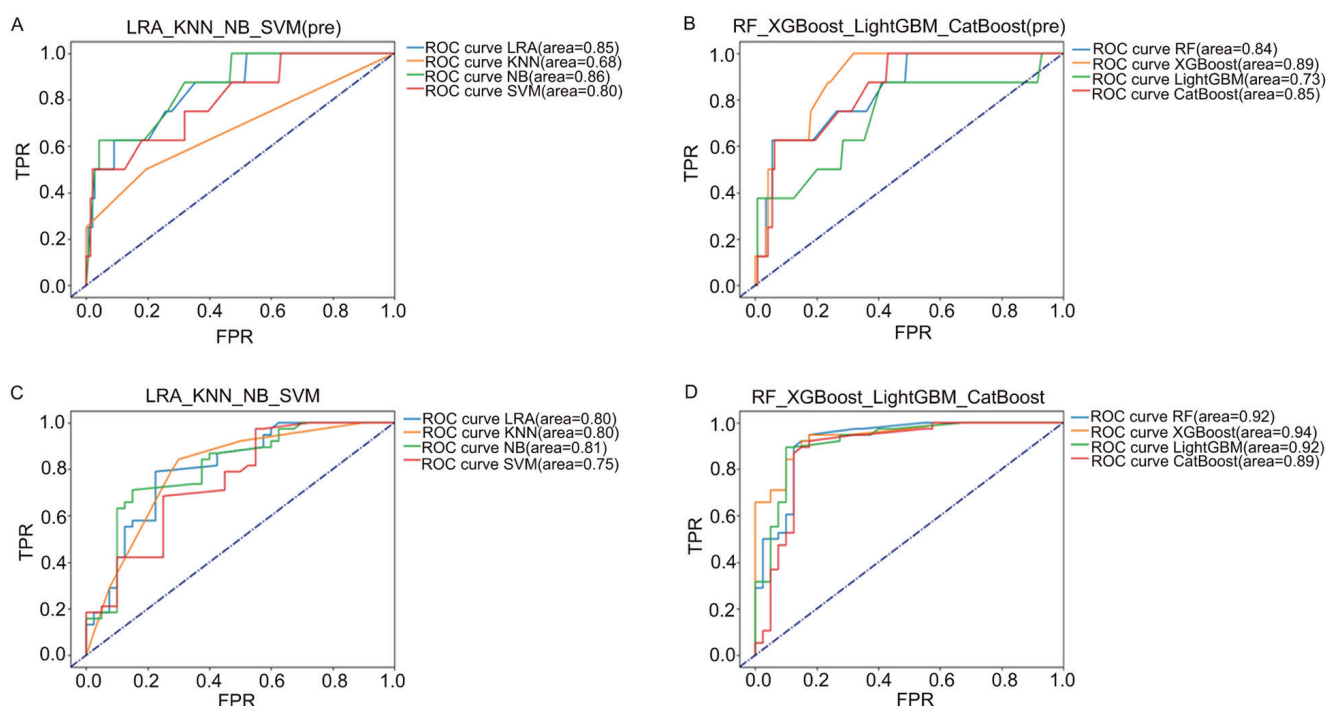
Prediction values for the 8 machine learning models are presented in Table 2. The highest AUC was achieved with

the XGBoost model (AUC = 0.94) and the lowest AUC with the SVM model (AUC = 0.75). The highest and lowest accuracy were observed with the XGBoost and NB models, respectively, while the highest and lowest precision were

**Table 2.** Cross-validation comparison between models

Model	Accuracy (95% CI)	Precision (95% CI)	Recall (95% CI)	F1 (95% CI)	AUC (95% CI)
LRA	0.81 (0.76–0.85)	0.89 (0.86–0.93)	0.70 (0.64–0.65)	0.76 (0.71–0.81)	0.80 (0.72–0.89)
KNN	0.81 (0.76–0.85)	0.86 (0.81–0.90)	0.74 (0.69–0.80)	0.79 (0.75–0.84)	0.80 (0.71–0.89)
NB	0.64 (0.58–0.70)	0.68 (0.62–0.74)	0.63 (0.57–0.68)	0.59 (0.53–0.65)	0.81 (0.72–0.89)
SVM	0.83 (0.79–0.88)	0.94 (0.91–0.97)	0.71 (0.65–0.76)	0.80 (0.75–0.85)	0.75 (0.65–0.85)
RF	0.84 (0.79–0.88)	0.92 (0.89–0.95)	0.75 (0.69–0.80)	0.82 (0.78–0.87)	0.92 (0.86–0.98)
XGBoost	0.84 (0.80–0.89)	0.92 (0.89–0.95)	0.75 (0.70–0.80)	0.83 (0.78–0.87)	0.94 (0.89–0.99)
LightGBM	0.82 (0.78–0.87)	0.87 (0.83–0.91)	0.77 (0.72–0.82)	0.81 (0.76–0.86)	0.92 (0.86–0.98)
CatBoost	0.84 (0.80–0.88)	0.92 (0.89–0.96)	0.75 (0.70–0.80)	0.83 (0.78–0.87)	0.89 (0.82–0.86)

AUC – area under the curve; 95% CI – 95% confidence interval; LRA – logistic regression analysis; KNN – k-nearest neighbor; NB – naïve Bayes; SVM – support vector machine; RF – random forest; XGBoost – extreme gradient boosting; LightGBM – light gradient boosting machine; CatBoost – categorical features gradient boosting. The F1-score combines precision and recall into one metric by calculating the harmonic mean between those two.<sup>35</sup>



**Fig. 4.** Receiver operating characteristic (ROC) curve of 2 types of machine learning models before and after data processing algorithm. A. LRA\_KNN\_NB\_SVM before data processing; B. RF\_XGBoost\_LightGBM\_CatBoost before data processing; C. LRA\_KNN\_NB\_SVM after data processing; D. RF\_XGBoost\_LightGBM\_CatBoost after data processing. The XGBoost was confirmed as the final model for this study

TPR – true positive rates; FPR – false positive rates; LRA – logistic regression analysis; KNN – k-nearest neighbor; NB – naïve Bayes; SVM – support vector machine; RF – random forest; XGBoost – extreme gradient boosting; LightGBM – light gradient boosting machine; CatBoost – categorical features gradient boosting.

achieved using the SVM and NB models, respectively. Based on receiver operating characteristics (ROC) curve analysis, the XGBoost model was selected as the final model for the study (Fig. 4). Figure 5 shows the importance of the different variables when analyzed by the different models. Table 3 displays all of the variables evaluated in this study.

## Discussion

Results suggest that it is possible to use machine learning algorithms to predict the risk of complications after OPCAB. The highest predictive value was achieved using

the XGBoost model, based on VE, SVG, NeB, SvO<sub>2</sub>, T1, and HBP, as revealed by the AUC, which can be used as the main metric to determine the optimal classifier.<sup>34</sup>

Previous studies used machine learning to predict the mortality and morbidity of cardiac surgery. In a study by Kartal, mortality risk was predicted using the EuroSCORE and the C4.5 algorithm: both the EuroSCORE and the C4.5 algorithm included age, serum creatinine, LVEF, and mean pulmonary hypertension (mPAP).<sup>28</sup> They used their algorithm to develop a web application for risk prediction after cardiac surgery. Castela Forte et al. used machine learning to evaluate 88 perioperative variables in order to predict 5-year mortality after cardiac surgery;



Table 3. Features taken into account in this study

No.	Name	Abbreviation	No.	Name	Abbreviation
1.	gender	/	29.*	baseline mean arterial pressure	MAP0
2.*	age	/	30.	skin incision-bypass mean arterial pressure variables	MAP1
3.*	body mass index	BMI	31.	bypass-end mean arterial pressure variables	MAP2
4.	high blood pressure	HBP	32.*	baseline mean pulmonary arterial pressure	mPAP0
5.	diabetes mellitus	DM	33.	skin incision-bypass mean pulmonary arterial pressure variables	mPAP1
6.	cerebrovascular disease	CVD	34.	bypass-end mean pulmonary arterial pressure variables	mPAP2
7.	chronic renal failure	CRF	35.*	temperature	TP
8.	bronchial asthma	BA	36.	mixed venous oxygen saturation	SvO <sub>2</sub>
9.	hyperlipidemia	HL	37.*	cardiac index	CI
10.	chronic obstructive pulmonary disease	COPD	38.	mean pulmonary capillary wedge pressure	mPCWP
11.	atrial fibrillation	AF	39.	American Society of Anesthesiologists score	ASA
12.	acute myocardial infarction	AMI	40.	morphine	/
13.	percutaneous coronary intervention	PCI	41.	etomidate	Etomi
14.	New York Heart Association classification score	NYHA	42.	nerve block	NeB
15.	left main	LM	43.	skin-bypass single push vasoactive agents	VD1
16.	root cause analysis	RCA	44.	bypass-end single push vasoactive agents	VD2
17.*	left ventricle diameter	LVD	45.	internal mammary artery	IMA
18.*	left ventricular ejection fraction	LVEF	46.	radial artery	RA
19.*	fraction shortening	FS	47.	saphenous vein	SV
20.	thickened ventricular septum	TVS	48.*	aortic cross-clamp time	ACCT
21.	valvular heart disease	VHD	49.*	fluid intake	FI
22.	complete left bundle branch block	CLBBB	50.*	urine	/
23.	emergency surgery	/	51.	number of saphenous vein grafts	SVG
24.*	skin incision-bypass time	T1	52.	potassium supplement (intraoperative)	PS
25.*	operation time	T2	53.	calcium supplement (intraoperative)	CS
26.*	baseline heart rate	HRO	54.	on-pump coronary artery bypass	ONCAB
27.	skin incision-bypass heart rate variability	HR1	55.	ventricular fibrillation	VF
28.	bypass-end heart rate variability	HR2	56.	preventive intra-aortic balloon pump	IABP_pre

\* continuous variable.

they observed that postoperative urea concentration, age and creatinine concentration, achieved the best predictive values across different cardiac surgery types.<sup>29</sup> Kim et al. examined deep neural network, GBM and a generalized linear model to predict major adverse cardiovascular events 1, 6 and 12 months after cardiac surgery, and achieved accuracies >95%.<sup>30</sup> Zhong et al. used deep learning to predict the risk of septic shock, thrombocytopenia and liver dysfunction after open-heart surgery.<sup>31</sup> They examined the performance of XGBoost, RF, KNN and logistic regression, and showed that the XGBoost model achieved the best predictive value for complications. Alshakhs et al. used machine learning to determine the length of hospital stay after CABG, which might be considered a surrogate for the occurrence of postoperative complications.<sup>32</sup> They also showed that an RF model including age, height, EuroSCORE II, and the use of IABP achieved the best predictive value.

In the present study, the Pearson's correlation analysis was used to consider the importance of extracting features from different directions (machine learning direction and statistical direction) to make the screened features more convincing. The data indicated that VF, SVG, NeB, SvO<sub>2</sub>, T1, and HBP in the XGBoost model achieved the best predictive value. The XGBoost is an advanced complex implementation of gradient boosting algorithms.<sup>35,36</sup> It can handle both regularization and over/underfitting issues.<sup>35,36</sup> The parameters selected by the user (i.e., the hyperparameters) usually have a strong effect on the performance of a machine learning algorithm.<sup>37,38</sup> Still, XGBoost can adapt to the selected hyperparameters to achieve the best fitting,<sup>35,36,39</sup> which explains its good performance in the present study.

Comparisons among studies are difficult. Indeed, various studies have examined different machine learning

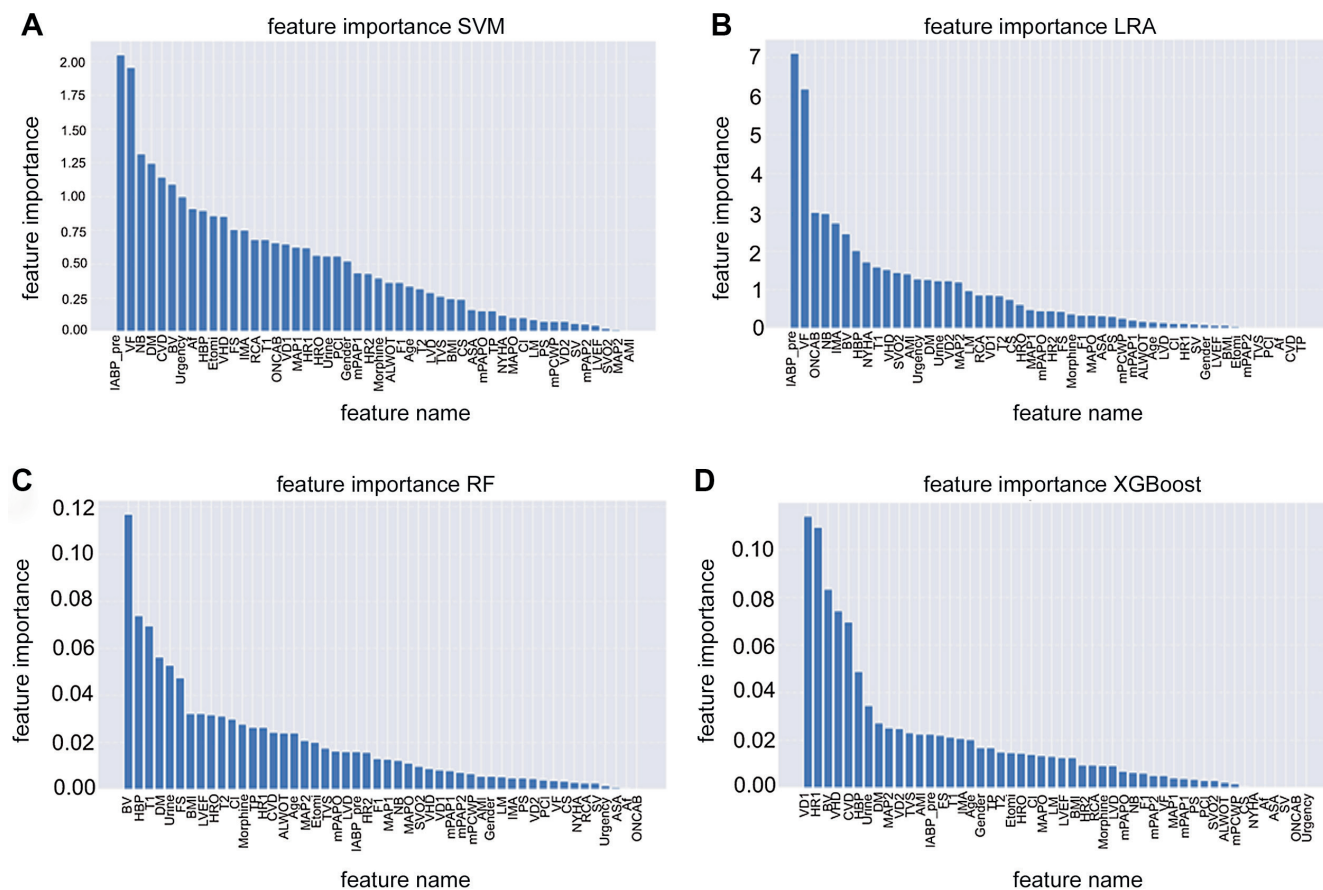


Fig. 5. The feature importance of different algorithms. A. Support vector machine (SVM); B. Logistic regression analysis (LRA); C. Random forest (RF); D. Extreme gradient boosting (XGBoost)

models based on a wide variety of different variables. In addition, the endpoints and the definitions of complications vary, and the study populations have included various types of surgery. In the present study, only patients who underwent OPCAB were included, and the endpoint was the occurrence of IABP assistance, secondary surgery and death. Nonetheless, various studies have shown that the XGBoost model achieved good predictive value. Indeed, similar to above, Zhong et al. used the XGBoost model for predicting complications after open-heart surgery.<sup>31</sup> Kilic et al. used the XGBoost model to predict the occurrence of operative mortality (AUC = 0.771), renal failure (AUC = 0.776), prolonged ventilation (AUC = 0.739), reoperation (AUC = 0.637), stroke (AUC = 0.684), and deep wound infection (AUC = 0.599) after aortic valve replacement.<sup>40</sup> Additionally, Lee et al. showed that the XGBoost model had the highest predictive power for AKI after cardiac surgery.<sup>41</sup>

Apart from comparison with other deep learning models, the model established here should be compared with well-known and recognized models. Indeed, the EuroSCORE II and the original EuroSCORE have been used for decades to predict mortality risk after cardiac surgery and help improve patient outcomes.<sup>42,43</sup> The STS score can also be used to determine the risk of CABG.<sup>20</sup> However, both scores are not specific to OPCAB. Furthermore,

the data used in the present study were taken directly from the anesthesia monitor system, and some components were not included in the EuroSCORE II and STS scores. Future studies should be set up to allow such direct comparisons using the same set of patients.

Patients who met the study endpoint had a low frequency of DM and high SvO<sub>2</sub>. Diabetes mellitus is associated with poor outcomes after CABG or cardiac surgery.<sup>44–47</sup> On the other hand, poor outcomes after CABG have been associated with either high SvO<sub>2</sub><sup>48</sup> or low SvO<sub>2</sub>.<sup>49</sup> Considering the small number of patients who met the endpoint in the present study, no conclusion can be drawn on these points.

## Limitations

This study has a number of limitations. The data were unbalanced, with the proportion of patients who met the complication endpoint being small. Although the category imbalance was corrected at the data and algorithm levels, it inevitably affected the fitting degree of the model. Follow-up studies are required to optimize the algorithm based on category imbalance characteristics, in order to reduce the impact of category imbalance on the model performance. Although the predictive factors selected in this study related to the endpoint as much as possible, some

predictive factors that had not been discovered or confirmed might have been omitted. In the future, more predictive factors could be added through an in-depth study of OPCAB-related risk factors to improve the performance of the model.

The sample size of this study was small, and it was a single-center retrospective study. The data were from a single center or a single physician team, which limited generalizability and probably introduced some bias caused by varying experience of the surgeons and anesthesiologists. Future studies should be extended to multiple centers. As a retrospective study, this investigation collected the data of all patients who met the criteria in our center during the study period. Relevant data from this period are relatively complete, and data quality cannot be guaranteed in earlier cases. After June 2020, the number of operations decreased due to the coronavirus pandemic, which might have led to bias. An independent validation dataset was also lacking. Therefore, the final model might have poor generalizability. Continuous iterations of the model, through large multicenter samples and prospective validation studies, should increase the generalizability of the model. Since this study only predicted specific, not all complications of OPCAB surgery, its purpose was not to compare the performance of the final model with the EuroSCORE. Data were insufficient to allow separate analyses of patients undergoing total artery bypass grafting.

## Conclusions

This study verified the effectiveness of different machine learning models and provided suggestions for the best mathematical model for predicting the risk of complications after OPCAB. This knowledge could be used to continuously optimize the model and introduce it into the clinical medical electronic system, which would allow clinicians to use optimizing treatment strategies in real-time.

## Data availability

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

## Supplementary files

The Supplementary Files are available at <https://doi.org/10.5281/zenodo.7063461>. The package contains the following files:

Supplementary Fig. 1. Results of the Pearson's correlation analysis.

Supplementary File 2. Features selection results of Wilcoxon and  $\chi^2$  tests).

Supplementary Table 1. The list of meaningless features.

Supplementary Table 2. The list of missing data.

Supplementary Table 3. Verification results of LRA.

Supplementary Table 4. Verification results of KNN.

Supplementary Table 5. Verification results of naive Bayes (NB).

Supplementary Table 6. Verification results of SVM.

Supplementary Table 7. Verification results of RF.


Supplementary Table 8. Verification results of XGBoost.


Supplementary Table 9. Verification results of LightGBM.


Supplementary Table 10. Verification results of CatBoost.

## ORCID iDs

Yang Zhang  <https://orcid.org/0000-0002-3246-1552>

Lin Li  <https://orcid.org/0000-0001-7262-5788>

Ye Li  <https://orcid.org/0000-0002-0154-2464>

Zhihe Zeng  <https://orcid.org/0000-0003-2335-2300>

## References

- Diodato M, Chedrawy EG. Coronary artery bypass graft surgery: The past, present and future of myocardial revascularisation. *Surg Res Pract*. 2014;2014:726158. doi:10.1155/2014/726158
- O'Gara PT, Kushner FG, Ascheim DD, et al. 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction: A report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2013;127(4):e362–e425. doi:10.1161/CIR.0b013e3182742cf6
- Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. *Eur Heart J*. 2018;39(2):119–177. doi:10.1093/eurheartj/ehx393
- Writing Committee Members, Hillis LD, Smith PK, et al. 2011 ACCF/AHA Guideline for Coronary Artery Bypass Graft Surgery: A report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation*. 2011;124(23):2610–2642. doi:10.1161/CIR.0b013e31823c074e
- Møller CH, Penninga L, Wetterslev J, Steinbrüchel DA, Gluud C. Off-pump versus on-pump coronary artery bypass grafting for ischaemic heart disease. *Cochrane Database Syst Rev*. 2012;14(3):CD007224. doi:10.1002/14651858.CD007224.pub2
- Kuss O, von Salviati B, Börgermann J. Off-pump versus on-pump coronary artery bypass grafting: A systematic review and meta-analysis of propensity score analyses. *J Thorac Cardiovasc Surg*. 2010;140(4):829.e13–835.e13. doi:10.1016/j.jtcvs.2009.12.022
- Matkovic M, Tutus V, Bilbija I, et al. Long term outcomes of the off-pump and on-pump coronary artery bypass grafting in a high-volume center. *Sci Rep*. 2019;9(1):8567. doi:10.1038/s41598-019-45093-3
- Diegeler A, Börgermann J, Kappert U, et al. Five-year outcome after off-pump or on-pump coronary artery bypass grafting in elderly patients. *Circulation*. 2019;139(16):1865–1871. doi:10.1161/CIRCULATIONAHA.118.035857
- Sousa-Uva M, Neumann FJ, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on Myocardial Revascularization. *Eur J Cardiothorac Surg*. 2019;55(1):4–90. doi:10.1093/ejcts/ezy289
- Chikwe J, Lee T, Itagaki S, Adams DH, Egorova NN. Long-term outcomes after off-pump versus on-pump coronary artery bypass grafting by experienced surgeons. *J Am Coll Cardiol*. 2018;72(13):1478–1486. doi:10.1016/j.jacc.2018.07.029
- Daviewala PM. Current outcomes of off-pump coronary artery bypass grafting: Evidence from real world practice. *J Thorac Dis*. 2016;8(S10):S772–S786. doi:10.21037/jtd.2016.10.102
- Wunsch H, Linde-Zwirble WT, C. Angus D. Methods to adjust for bias and confounding in critical care health services research involving observational data. *J Crit Care*. 2006;21(1):1–7. doi:10.1016/j.jccr.2006.01.004
- Mack MJ, Pfister A, Bachand B, et al. Comparison of coronary bypass surgery with and without cardiopulmonary bypass in patients with multivessel disease. *J Thorac Cardiovasc Surg*. 2004;127(1):167–173. doi:10.1016/j.jtcvs.2003.08.032

14. Shroyer AL, Grover FL, Hattler B, et al. On-pump versus off-pump coronary-artery bypass surgery. *N Engl J Med*. 2009;361(19):1827–1837. doi:10.1056/NEJMoa0902905
15. Shroyer AL, Hattler B, Wagner TH, et al. Five-year outcomes after on-pump and off-pump coronary-artery bypass. *N Engl J Med*. 2017;377(7):623–632. doi:10.1056/NEJMoa1614341
16. Afilalo J, Rasti M, Ohayon SM, Shimony A, Eisenberg MJ. Off-pump vs. on-pump coronary artery bypass surgery: An updated meta-analysis and meta-regression of randomized trials. *Eur Heart J*. 2012;33(10):1257–1267. doi:10.1093/eurheartj/ehs307
17. Caputti GM, Palma JH, Gaia DF, Buffolo E. Off-pump coronary artery bypass surgery in selected patients is superior to the conventional approach for patients with severely depressed left ventricular function. *Clinics (Sao Paulo)*. 2011;66(12):2049–2053. doi:10.1590/S1807-59322011001200009
18. Maltais S, Ladouceur M, Cartier R. The influence of a low ejection fraction on long-term survival in systematic off-pump coronary artery bypass surgery. *Eur J Cardiothorac Surg*. 2011;39(5):e122–e127. doi:10.1016/j.ejcts.2010.12.022
19. Xia L, Ji Q, Song K, et al. Early clinical outcomes of on-pump beating-heart versus off-pump technique for surgical revascularization in patients with severe left ventricular dysfunction: The experience of a single center. *J Cardiothorac Surg*. 2017;12(1):11. doi:10.1186/s13019-017-0572-x
20. Ragosta M. Left main coronary artery disease: Importance, diagnosis, assessment, and management. *Curr Probl Cardiol*. 2015;40(3):93–126. doi:10.1016/j.cpcardiol.2014.11.003
21. Grant SW, Hickey GL, Dimarakis I, et al. How does EuroSCORE II perform in UK cardiac surgery: An analysis of 23 740 patients from the Society for Cardiothoracic Surgery in Great Britain and Ireland National Database. *Heart*. 2012;98(21):1568–1572. doi:10.1136/heartjnl-2012-302483
22. Parolari A, Pesce LL, Trezzi M, et al. Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: Single institution experience and meta-analysis. *Eur Heart J*. 2008;30(3):297–304. doi:10.1093/eurheartj/ehn581
23. Hannan EL, Wu C, Bennett EV, et al. Risk stratification of in-hospital mortality for coronary artery bypass graft surgery. *J Am Coll Cardiol*. 2006;47(3):661–668. doi:10.1016/j.jacc.2005.10.057
24. Singh M, Gersh BJ, Li S, et al. Mayo Clinic Risk Score for percutaneous coronary intervention predicts in-hospital mortality in patients undergoing coronary artery bypass graft surgery. *Circulation*. 2008;117(3):356–362. doi:10.1161/CIRCULATIONAHA.107.711523
25. Birim Ö, van Gameren M, Bogers AJJC, Serruys PW, Mohr FW, Kappelein AP. Complexity of coronary vasculature predicts outcome of surgery for left main disease. *Ann Thorac Surg*. 2009;87(4):1097–1105. doi:10.1016/j.athoracsur.2008.11.079
26. Weller GB, Lovely J, Larson DW, Earnshaw BA, Huebner M. Leveraging electronic health records for predictive modeling of post-surgical complications. *Stat Methods Med Res*. 2018;27(11):3271–3285. doi:10.1177/0962280217696115
27. Soguero-Ruiz C, Hindberg K, Mora-Jiménez I, et al. Predicting colorectal surgical complications using heterogeneous clinical data and kernel methods. *J Biomed Inform*. 2016;61:87–96. doi:10.1016/j.jbi.2016.03.008
28. Kartal E. Machine learning techniques in cardiac risk assessment. *Turk Gogus Kalp Dama*. 2018;26(3):394–401. doi:10.5606/tgkdc.dergisi.2018.15559
29. Castela Forte J, Mungroop HE, de Geus F, et al. Ensemble machine learning prediction and variable importance analysis of 5-year mortality after cardiac valve and CABG operations. *Sci Rep*. 2021;11(1):3467. doi:10.1038/s41598-021-82403-0
30. Kim YJ, Saqlian M, Lee JY. Deep learning-based prediction model of occurrences of major adverse cardiac events during 1-year follow-up after hospital discharge in patients with AMI using knowledge mining. *Pers Ubiquit Comput*. 2022;26(2):259–267. doi:10.1007/s00779-019-01248-7
31. Zhong Z, Yuan X, Liu S, Yang Y, Liu F. Machine learning prediction models for prognosis of critically ill patients after open-heart surgery. *Sci Rep*. 2021;11(1):3384. doi:10.1038/s41598-021-83020-7
32. Alshakhs F, Alharthi H, Aslam N, Khan IU, Elasheri M. Predicting post-operative length of stay for isolated coronary artery bypass graft patients using machine learning. *Int J Gen Med*. 2020;13:751–762. doi:10.2147/IJGM.S250334
33. Orozco-Arias S, Piña JS, Tabares-Soto R, Castillo-Ossa LF, Guyot R, Isaza G. Measuring performance metrics of machine learning algorithms for detecting and classifying transposable elements. *Processes*. 2020;8(6):638. doi:10.3390/pr8060638
34. Provost F, Fawcett T. Robust classification for imprecise environments. *arXiv*. 2000;2000:arXiv:cs/0009007. doi:10.48550/ARXIV.CS/0009007
35. Xia Y, Liu C, Li Y, Liu N. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring. *Expert Syst Appl*. 2017;78:225–241. doi:10.1016/j.eswa.2017.02.017
36. Zięba M, Tomczak SK, Tomczak JM. Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Syst Appl*. 2016;58:93–101. doi:10.1016/j.eswa.2016.04.001
37. Brown SD, de Juan A. ICRM-2011 international chemometrics research meeting. *Chemom Intell Lab Syst*. 2012;111(1):66. doi:10.1016/j.chemolab.2011.12.002
38. Mantovani RG, Rossi ALD, Vanschoren J, Bischl B, de Carvalho ACPLF. Effectiveness of random search in SVM hyper-parameter tuning. In: *2015 International Joint Conference on Neural Networks (IJCNN)*. Killarney, Ireland: IEEE; 2015:1–8. doi:10.1109/IJCNN.2015.7280664
39. Budholiya K, Shrivastava SK, Sharma V. An optimized XGBoost based diagnostic system for effective prediction of heart disease. *J King Saud Univ Comput Inf Sci*. 2022;34(7):4514–4523. doi:10.1016/j.jksuci.2020.10.013
40. Kilic A, Goyal A, Miller JK, Gleason TG, Dubrawski A. Performance of a machine learning algorithm in predicting outcomes of aortic valve replacement. *Ann Thorac Surg*. 2021;111(2):503–510. doi:10.1016/j.athoracsur.2020.05.107
41. Lee HC, Yoon HK, Nam K, et al. Derivation and validation of machine learning approaches to predict acute kidney injury after cardiac surgery. *J Clin Med*. 2018;7(10):322. doi:10.3390/jcm7100322
42. Nashef SAM, Roques F, Sharples LD, et al. EuroSCORE II. *Eur J Cardiothorac Surg*. 2012;41(4):734–745. doi:10.1093/ejcts/ezs043
43. Biancari F, Vasques F, Mikkola R, Martin M, Lahtinen J, Heikkinen J. Validation of EuroSCORE II in patients undergoing coronary artery bypass surgery. *Ann Thorac Surg*. 2012;93(6):1930–1935. doi:10.1016/j.athoracsur.2012.02.064
44. van Straten AHM, Soliman Hamad MA, van Zundert AAJ, et al. Diabetes and survival after coronary artery bypass grafting: Comparison with an age- and sex-matched population. *Eur J Cardiothorac Surg*. 2010;37(5):1068–1074. doi:10.1016/j.ejcts.2009.11.042
45. Kogan A, Ram E, Levin S, et al. Impact of type 2 diabetes mellitus on short- and long-term mortality after coronary artery bypass surgery. *Cardiovasc Diabetol*. 2018;17(1):151. doi:10.1186/s12933-018-0796-7
46. Ram E, Sternik L, Klempfner R, et al. Type 2 diabetes mellitus increases the mortality risk after acute coronary syndrome treated with coronary artery bypass surgery. *Cardiovasc Diabetol*. 2020;19(1):86. doi:10.1186/s12933-020-01069-6
47. Järvinen O, Hokkanen M, Huhtala H. Diabetics have inferior long-term survival and quality of life after CABG. *Int J Angiol*. 2019;28(01):50–56. doi:10.1055/s-0038-1676791
48. Balzer F, Sander M, Simon M, et al. High central venous saturation after cardiac surgery is associated with increased organ failure and long-term mortality: An observational cross-sectional study. *Crit Care*. 2015;19(1):168. doi:10.1186/s13054-015-0889-6
49. Holm J, Håkanson E, Vánky F, Svedjeholm R. Mixed venous oxygen saturation predicts short- and long-term outcome after coronary artery bypass grafting surgery: A retrospective cohort analysis. *Br J Anesth*. 2011;107(3):344–350. doi:10.1093/bja/aer166

# Evaluation of pretreatment and early treatment changes in serum $\beta$ -hCG with methotrexate

Huijuan Zhou<sup>A–F</sup>, Shuangdi Li<sup>A–F</sup>

Department of Obstetrics and Gynecology, Shanghai First Maternity and Infant Hospital, School of Medicine, Tongji University, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

Adv Clin Exp Med. 2023;32(2):195–202

## Address for correspondence

Shuangdi Li  
E-mail: shuangdili@126.com

## Funding sources

None declared

## Conflict of interest

None declared

Received on May 6, 2022

Reviewed on July 18, 2022

Accepted on August 26, 2022

Published online on December 8, 2022

## Abstract

**Background.** Serum beta-human chorionic gonadotropin ( $\beta$ -hCG) is an important biomarker for the detection of ectopic pregnancies (EPs). The  $\beta$ -hCG levels between days 1 and 4 after methotrexate (MTX) treatment as an indicator of the success of the MTX in EP have been the focus of research.

**Objectives.** To determine whether the change in the  $\beta$ -hCG levels at day 1 and 4 and pretreatment at 48-hour increments can predict early treatment failure of single-dose MTX in EP.

**Materials and methods.** This was a retrospective study of 1120 EPs treated with a single dose of MTX. Treatment failure was defined as an obligation to proceed to surgery or the need for additional doses of MTX.

**Results.** A total 722 out of 1120 EPs had an increase in  $\beta$ -hCG on day 4 after MTX treatment. The logistic regression analysis indicated that 3 dependents were significantly associated with treatment failure: 1) a pretreatment 48-hour increase in  $\beta$ -hCG (odds ratio (OR): 1.249, 95% confidence interval (95% CI): 1.008–2.049,  $p < 0.001$ ); 2) a change in  $\beta$ -hCG between day 1 and 4 (OR: 1.384, 95% CI: 1.097–2.198,  $p < 0.001$ ); and 3) a history of EP (OR: 1.208, 95% CI: 1.041–2.011,  $p < 0.001$ ). The optimal cutoff point for the prediction of treatment failure was an increase of more than 19% in the 48 h before the treatment, and an increase of more than 36% between day 1 and day 4 in  $\beta$ -hCG concentrations. Patients with an increase in  $\beta$ -hCG levels of less than 36% on day 4 experienced MTX treatment failure in 4.2% ( $n = 25$ ), compared to 74.5% ( $n = 88$ ) of the patients with an increase above 36%.

**Conclusions.** A serum  $\beta$ -hCG increase of more than 36% on day 4 after the administration of MTX alongside a more than 19% increase in  $\beta$ -hCG concentration 48 h before the MTX treatment may predict the early failure of medical treatment for an EP.

**Key words:**  $\beta$ -hCG, methotrexate, ectopic pregnancy

## Cite as

Zhou H, Li S. Evaluation of pretreatment and early treatment changes in serum  $\beta$ -hCG with methotrexate. *Adv Clin Exp Med.* 2023;32(2):195–202. doi:10.17219/acem/153043

## DOI

10.17219/acem/153043

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Background

Ectopic pregnancy (EP) is a common acute gynecological issue. Its incidence has increased significantly all over the world in recent years. Tubal pregnancies are approx. 2% of all pregnancies. An available treatment of non-ruptured tubal pregnancy used all over the world is a single dose of methotrexate (MTX). It was first identified for the treatment of tubal pregnancies by Stovall et al.<sup>1</sup> The tubal EP and MTX treatment protocol have attracted the attention of researchers in recent years.<sup>2–15</sup> Jurkovic et al. followed up EP with low serum beta-human chorionic gonadotropin ( $\beta$ -hCG) levels ( $<1500$  IU/L) and confirmed the diagnose of EP by ultrasonography. They administered a single-dose injection of MTX at  $50$  mg/m<sup>2</sup> to the treatment group. In the group of patients who did not undergo a surgical protocol, a decrease in serum  $\beta$ -hCG to  $<20$  IU/L or a negative pregnancy test was determined as a successful treatment.<sup>2</sup> It has been stated that the success rate of MTX may be more effective in subgroups with lower values of serum  $\beta$ -hCG, especially below  $2000$  mIU/mL.<sup>14</sup> In another study, it has been suggested that MTX treatment administered as a single dose had increased success rates in EP with lower serum  $\beta$ -hCG at the start of treatment. They reported that changes in serum  $\beta$ -hCG within the first 4 days are important in determining the treatment protocol and efficacy. Similarly, if the serum  $\beta$ -hCG level is lower than 4% or increases in the first 4 days, an additional dose of MTX should be given on day 4.<sup>3,7</sup> In a study in which EP was evaluated retrospectively, the success rate of MTX administered at a dose of  $1$  mg/kg intramuscularly was assessed.<sup>4</sup> Similar to our study, serum  $\beta$ -hCG levels on days 0, 4 and 7 were evaluated after the administration of MTX, and it was suggested that differences in  $\beta$ -hCG levels between days 0 and 4 may serve as an early indicator for treatment outcome.<sup>5</sup> According to Hoyos et al., the morbidly obese were 5 times more likely to require an additional dose compared to non-morbidly obese.<sup>6</sup> Helvacioğlu and Doğan reported that MTX administered in 2 doses significantly affected serum  $\beta$ -hCG levels and achieved successful results.<sup>8</sup> It has been reported that a positive response in patients who are adequately treated with MTX, together with the low and decreasing serum  $\beta$ -hCG levels, rarely results in ruptured EP,<sup>9</sup> and it reminded again that it should not be overlooked in the surgical procedure.<sup>9</sup> It was stated that in cases where ovulation is high and patients underwent laparoscopic left salpingectomy, serum  $\beta$ -hCG should be followed up after the definitive treatment of tubal EP including surgical interventions.<sup>11</sup> Lin and Hsieh found the initial serum  $\beta$ -hCG level to be  $1,258$  mIU/mL. They reported that they applied laparoscopic salpingectomy at the end because the increase continued and did not fall into the acceptable range after 1 month despite the 2<sup>nd</sup> and 3<sup>rd</sup> doses of MTX given to the patients.<sup>10</sup> Grigoriu et al. reported that right salpingectomy was performed after unsuccessful MTX treatment.<sup>14</sup> Cohen et al. observed that

women with an increase in serum  $\beta$ -hCG  $>69\%$  within 48 h prior to MTX had an 85% probability of tubal rupture, and that women with serum  $\beta$ -hCG elevation  $<20\%$  had a lower absolute risk for tubal rupture.<sup>12</sup> They also reported that medical treatment with a single dose of oral mifepristone, as well as systemic multi-dose MTX with folinic acid were required when a living fetus and high serum  $\beta$ -hCG levels were present.<sup>13</sup>

The success rates of MTX treatment have been determined to be between 52% and 96%, and the fertility rate with delivery after medical treatment for tubal EP is 67–80.7%. These rates are comparable to the success rates of surgical procedures.<sup>16–20</sup> Serum  $\beta$ -hCG levels were measured on the 1<sup>st</sup>, 4<sup>th</sup> and 7<sup>th</sup> day after MTX treatment, and surgical intervention can be avoided when the levels decrease by more than 15% between the 4<sup>th</sup> and 7<sup>th</sup> day after a single dose of MTX.<sup>21</sup> In recent years, researchers have observed a decrease in serum  $\beta$ -hCG levels between days 1 and 4 after MTX treatment in patients with EP, which significantly affects the success rates of treatment. The proper treatment protocol in a situation when serum  $\beta$ -hCG level rises on the 4<sup>th</sup> day is still being discussed among researchers.

This study aimed to comprehensively examine serum  $\beta$ -hCG level changes in the 48 h prior to the administration of MTX and changes between days 1 and 4 after MTX treatment to predict the early failure of MTX therapy.

## Objectives

There are few studies reporting on increased serum  $\beta$ -hCG levels between days 1 and 4. Instead, many investigators wait until the 7<sup>th</sup> day to determine the treatment approach, while we aimed to determine it on day 4.

## Materials and methods

### Trial design

This retrospective study was conducted on patients in Shanghai First Birth and Infant Hospital, China, specifically women with EP who received  $50$  mg/m<sup>2</sup> of MTX. The study obtained approval No. KS1959 from the Ethics Committee of Shanghai First Maternity and Infant Hospital. The Chinese Clinical Trial Registry number is ChiCTR2000030658.

### Follow-up and procedures for patients

We reviewed patient records from the period from January 1, 2003, to December 31, 2019, using the hospital pharmacy MTX database. Ectopic pregnancy was diagnosed according to the standard criteria.<sup>21</sup> In total, 1135 patients were admitted with the diagnosis of EP and received MTX

treatment. Serum  $\beta$ -hCG values were measured on the 1<sup>st</sup>, 4<sup>th</sup> and 7<sup>th</sup> day in pregnant women administered MTX. Patients whose serum  $\beta$ -levels increased on day 4 were selected. An additional dose of MTX was given to the pregnant women when their serum  $\beta$ -hCG levels on the 7<sup>th</sup> day were less than 15% of the value measured between the 4<sup>th</sup> and 7<sup>th</sup> days. Study groups included a success group and a failure group, evaluated in terms of treatment outcomes. Successful treatment was defined as a decrease of more than 15% in serum  $\beta$ -hCG levels between the 4<sup>th</sup> and 7<sup>th</sup> day and no need for 2<sup>nd</sup> MTX treatment or surgery. Treatment failure was determined as the need for a 2<sup>nd</sup> MTX treatment or surgery in patients with EP. The values for these patients were not included in the statistical analysis.

Inclusion criteria were pretreatment serum  $\beta$ -hCG concentration  $\leq 2000$  U/L, no cardiac problems, gestational sac width  $< 4$  cm, biochemical values within expected values, and no problems with blood circulation. Exclusion criteria were ruptured tubal duct, non-tubal EP, or surgical intervention before the 4<sup>th</sup> day of MTX therapy. Methotrexate was administered to patients who met the specified criteria.

The day on which 50 mg/m<sup>2</sup> of MTX was administered was defined as the 1<sup>st</sup> day. Serum  $\beta$ -hCG levels were then measured on days 4 and 7. All eligible patients consented to receive MTX therapy. In cases where serum  $\beta$ -hCG levels measured between days 4 and 7 decreased by  $>15\%$ , serum  $\beta$ -hCG levels were not followed until negative values were detected. An additional dose of MTX was administered if serum  $\beta$ -hCG levels decreased by  $<15\%$  with a stable condition, as was a 3<sup>rd</sup> dose of MTX if the patient would like to receive the treatment once again. In this situation, the patients have the choice to decide whether get a 3<sup>rd</sup> dose or have the surgery, in case of no contraindication of MTX therapy. Surgery was performed when levels decreased by  $<15\%$ , there was an aggravated clinical syndrome or unstable condition, or the patient refused the 2<sup>nd</sup> dose of MTX. Additionally, during MTX treatment, we monitored the patients' condition, including blood pressure, pulse and vaginal bleeding. Furthermore, we advised the patients to abstain from sex during treatment.

## Statistical analyses

Statistical analyses were performed using IBM SPSS v. 19 (IBM Corp., Armonk, USA). The normal data distribution of continuous variables was verified with the Kolmogorov–Smirnov test. The Student's t-test was used to analyze continuous variables with a normal distribution, and the results are presented as mean and standard deviation ( $M \pm SD$ ). When the data distribution was non-normal, the comparison was performed using the Mann–Whitney test. Categorical variables were compared using the  $\chi^2$  test. In this study, we reported statistically significant positive and negative predictive values. The best cutoff value for changes in serum  $\beta$ -hCG levels between days 1 and 4

was determined using the receiver operating characteristic (ROC) curve method. When using the binary logistic regression model, the best model was built with the use of the Akaike information criterion (AIC), and the R-package “stats”, run in the R software v. 4.1.0 (R Core Team 2021, R Foundation for Statistical Computing, Vienna, Austria). Significant features in the logistic regression model evaluating MTX treatment were the ratio of day 4 to day 1 serum  $\beta$ -hCG levels, the ratio of day 1 to the day before the 48-hour treatment serum  $\beta$ -hCG levels, and day 1 serum  $\beta$ -hCG levels. The p-values  $< 0.05$  were considered statistically significant.

## Results

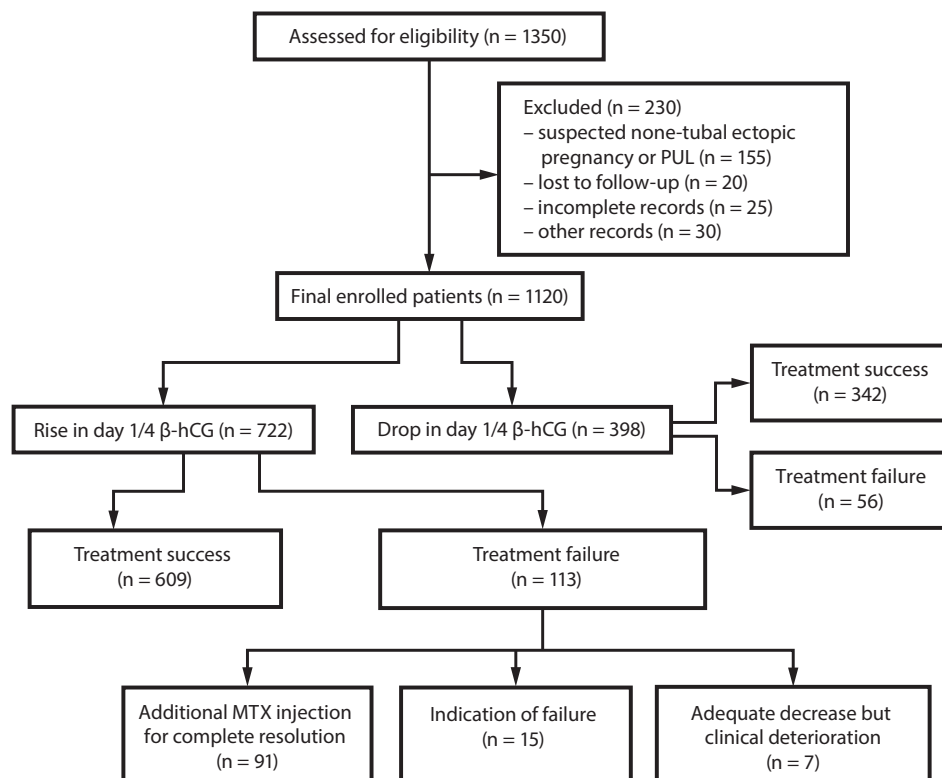
A total of 1350 patients who received MTX treatment for tubal EP at the Shanghai First Birth and Infant Hospital Department between January 1, 2003, and December 31, 2019, were enrolled into the study. A total of 230 patients were excluded because they did not meet our criteria, of which 155 were due to suspected non-tubal EP or pregnancy of unknown location (PUL), 20 due to loss to follow-up, 25 due to incomplete records, and 30 due to abnormal blood circulation or surgery before the 4<sup>th</sup> day. The records of 1120 patients met the criteria for this study.

### Baseline data

In total, 65% (722/1120) of patients had a serum  $\beta$ -hCG rise on day 4 and 35% (398/1120) had a drop on day 4 (Fig. 1). In patients whose serum  $\beta$ -hCG levels increased on the 4<sup>th</sup> day after a single dose of MTX treatment, successful results were obtained in 609 patients (84.3%) and treatment failure was observed in 113 patients (15.7%). An additional dose of MTX was administered as supplemental therapy to 91 patients (80.5%) who did not respond to treatment. Twenty-two female patients (19.5%) who initially responded positively to MTX required surgical intervention. The main characteristics of the participants are presented in Table 1. There were no statistically significant age-related differences between the evaluated groups, gravidity, parity, days since last menstrual period (LMP), ectopic size, history of cesarean section, history of EP, and initial progesterone values. Endometrial thickness [mm] was significantly higher in the failure group compared to the success group (6.69 (6.50–6.88) compared to 8.13 (8.51–9.75);  $p = 0.000$ ,  $Z = 7.952$ ).

### Evaluation of treatment results using serum $\beta$ -hCG levels

Serum  $\beta$ -hCG levels on day 1 were 1075 (836–1252.5) mIU/mL in the success group compared to 1180 (986–1484) mIU/mL in the failure group. The difference was



**Fig. 1.** Patient flow diagram of methotrexate (MTX) treatment for tubal ectopic pregnancy (EP) based on day 1 to day 4 serum beta-human chorionic gonadotropin ( $\beta$ -hCG) change  
PUL – pregnancy of unknown location.

**Table 1.** Baseline characteristics

Characteristics	Success	Failure	p-value	df	t
Age [years]	29.06 $\pm$ 4.69	29.65 $\pm$ 3.12	0.199 <sup>b</sup>	218	1.687
Weight (BMI) [kg/m <sup>2</sup> ]	23.41 $\pm$ 1.09	23.33 $\pm$ 1.41	0.566 <sup>b</sup>	137	-0.576
Gravidity, n	3.98 $\pm$ 1.99	3.84 $\pm$ 2.04	0.490 <sup>a</sup>	720	-0.69
Parity, n	0.97 $\pm$ 0.80	1.12 $\pm$ 0.78	0.760 <sup>a</sup>	720	1.776
Days since LMP	4.03 $\pm$ 4.86	53.44 $\pm$ 8.03	0.440 <sup>b</sup>	126	-0.766
Progesterone value day 1 [ng/mL]	13.07 $\pm$ 4.49	13.92 $\pm$ 3.32	0.061 <sup>b</sup>	169	1.909
Ectopic size [cm]	2.65 $\pm$ 0.71	2.60 $\pm$ 0.68	0.371 <sup>a</sup>	720	-0.896
History of cesarean section, n	1.04 (0.97, 1.10)	1.05 (0.89, 1.21)	0.135 <sup>c</sup>		-1.495 <sup>d</sup>
History of ectopic pregnancy, n	1.45 (1.36, 1.54)	1.88 (1.16, 1.99)	0.284 <sup>c</sup>		-1.071 <sup>d</sup>
Endometrial thickness [mm]	6.69 (6.50, 6.88)	13 (8.51, 9.75)	0.000 <sup>c</sup>		7.95 <sup>c</sup>

<sup>a</sup> values are given as mean  $\pm$  standard deviation (M  $\pm$  SD), independent sample t-test for equal variance was used; <sup>b</sup> values are given as M  $\pm$  SD, independent sample t-test for unequal variance was used; <sup>c</sup> values are given as the median (25<sup>th</sup> and 75<sup>th</sup> percentile (P25, P75)), Mann-Whitney U test was used; <sup>d</sup> Z-value; BMI – body mass index; LMP – last menstrual period; df – degrees of freedom.

statistically significant ( $p < 0.001$ ). This difference in serum  $\beta$ -hCG values was also found on day 4 ( $p < 0.001$ ) and day 7 ( $p < 0.001$ ). The median 48-hour pretreatment increment in  $\beta$ -hCG levels was lower in the success group (10% compared to 21%,  $p = 0.001$ ). The ratio between differences on days 1 and 4 and between days 4 and 7 was calculated for each group. When serum  $\beta$ -hCG levels were compared between days 1 and 4, an increase in serum  $\beta$ -hCG levels was more pronounced in the failure group ( $p < 0.001$ ). On the other hand, the decrease in serum  $\beta$ -hCG levels between days 4 and 7 was statistically significant in the success group ( $p < 0.001$ ). These data are shown in Table 2.

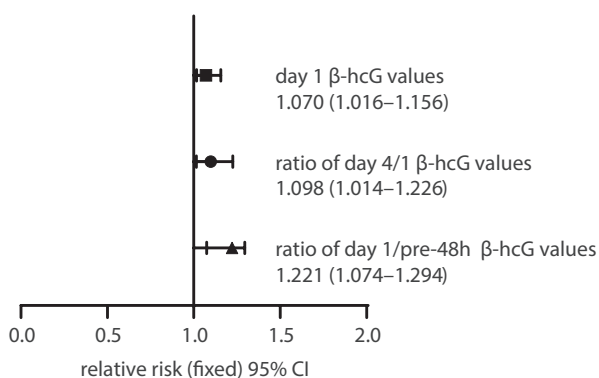
The best cutoff value for the rise in serum  $\beta$ -hCG levels from day 1 to day 4 was 36%. This cutoff value provided a sensitivity of 72.57%, a specificity of 86.86%, and the area under ROC curve (AUC) of 0.8083 (95% confidence interval (95% CI): 0.7480–8587). The most significant rate of increase in serum  $\beta$ -hCG before 48 h of treatment was 19%. This cutoff value provided a sensitivity of 61.95%, a specificity of 94.75% and an AUC of 0.7495 (95% CI: 0.6814–0.8177). The most significant cutoff value for the determined serum  $\beta$ -hCG value on day 1 was 18%. This cutoff value provided a sensitivity of 80.95%, a specificity of 93.98% and an AUC of 0.9182 (95% CI: 0.8540–0.9825) (Fig. 2).



**Table 2.** The relationship between serum  $\beta$ -hCG changes and MTX treatment results

Variable	Success (n = 609)	Failure (n = 113)	Z-value	p-value*
$\beta$ -hCG1	1075 (836, 1252.5)	1180 (986, 1484)	-5.939	0.000
$\beta$ -hCG4	1225 (977, 1455.5)	1616 (1366.5, 1885.5)	-9.887	0.000
$\beta$ -hCG7	704 (498, 904)	1555.5 (1797, 2220.5)	-14.583	0.000
Ratio $\beta$ -hCG1-4 (100%) <sup>§</sup>	24 (9, 26)	36 (22, 47)	-5.008	0.000
Ratio $\beta$ -hCG4-7 (100%) <sup>‡</sup>	48 (40, 55)	26 (17, 37)	-14.539	0.000
48-h pretreatment increment of $\beta$ -hCG (100%)	9.8 (6.1, 16.6)	21.4 (9.6, 23.8)	-17.897	0.000

\* values are given as mean (25<sup>th</sup> and 75<sup>th</sup> percentile (P25, P75)), Mann–Whitney U test was used; <sup>§</sup> ratio of beta-human chorionic gonadotropin ( $\beta$ -hCG) levels to day 4 serum  $\beta$ -hCG level minus day 1  $\beta$ -hCG level, divided by day 1 serum  $\beta$ -hCG level; <sup>‡</sup> ratio of  $\beta$ -hCG levels to day 4 serum  $\beta$ -hCG level minus day 7  $\beta$ -hCG level, divided by day 4 serum  $\beta$ -hCG level; MTX – methotrexate.



**Fig. 2.** Receiver operating characteristic (ROC) curves for the prediction of treatment failure of single-dose methotrexate (MTX) for tubal ectopic pregnancy (EP), showing the percentage change in day 1 to day 4, day 4 to day 7, and the 48-hour pretreatment increase in serum beta-human chorionic gonadotropin ( $\beta$ -hCG) levels

AUC – area under ROC curve; 95% CI – 95% confidence interval.

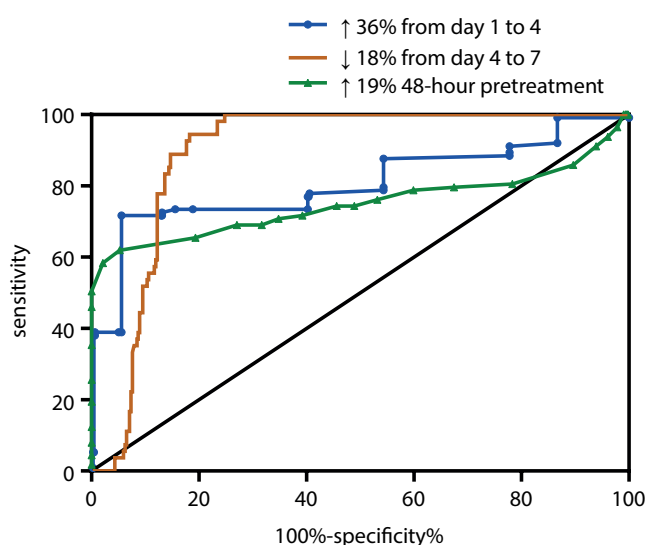
### Evaluation of serum $\beta$ -hCG levels after MTX treatment

Binary logistic regression was performed based on AIC using the R environment. The best logistic regression model of AIC was found to be 344.9. The important covariates were the ratio of serum  $\beta$ -hCG levels on day 4 to day 1, the ratio of serum  $\beta$ -hCG levels on day 1 to 48-hour pretreatment, and day 1 serum  $\beta$ -hCG levels, the coefficients of which were 0.0941, 0.2003 and 0.0016, respectively. The odds ratios (ORs) (95% CIs) of these 3 covariates are shown in Fig. 3.

In the group of patients whose  $\beta$ -hCG levels increased by less than 36% in the first 4 days of MTX treatment, the failure rate was 4.2%, while the success rate was 95.8% (579/604). On the other hand, when serum  $\beta$ -hCG level increased by more than 36% in the first 4 days, the success rate was 25% and the failure rate was 75% (Table 3).

### Discussion

Changes in serum  $\beta$ -hCG levels in the first 4 days after a single dose of MTX, along with the concomitant



**Fig. 3.** Multivariate binary logistic regression model performed using the R environment based on Akaike’s information criterion (AIC). The best model of AIC was 344.9. The statistically significant variables were the ratio of day 4 to day 1 serum beta-human chorionic gonadotropin ( $\beta$ -hCG) levels, the ratio of day 1 to 48-hour pretreatment serum  $\beta$ -hCG levels, and day 1 serum  $\beta$ -hCG levels

need for a 2<sup>nd</sup> dose of MTX, can help determine whether an emergency surgery is necessary. However, the question remains whether it is an efficient indicator. The purpose of our research was to identify the association between the changes in serum  $\beta$ -hCG levels in the first 4 days after a single dose of MTX and MTX treatment outcome. Recently, the role of day 4 serum  $\beta$ -hCG levels has become a topic of debate among researchers; however, there are only few studies on increased serum  $\beta$ -hCG levels on day 4. In clinic, the question whether the treatment was successful after single-dose MTX when the day 4 serum  $\beta$ -hCG level increased was often discussed. Often, we could only wait for the assessment of serum  $\beta$ -hCG level on day 7. Based on the increase in serum  $\beta$ -hCG level on the 4<sup>th</sup> day, we aimed to investigate whether a surgical protocol could be applied. The percentage of serum  $\beta$ -hCG change during the first 4 days was considered an important criterion for the need for additional doses of MTX. Akselim et al. determined the threshold value of percent serum  $\beta$ -hCG ratio during the first 4 days to be 4% (sensitivity 72.9%,

**Table 3.** A cutoff value for increased serum beta-human chorionic gonadotropin ( $\beta$ -hCG) percentage of 36% on day 4 in predicting methotrexate (MTX) treatment outcome

Increased serum $\beta$ -hCG percentage $\times$ MTX outcome		Success	Failure	Total
Increased serum $\beta$ -hCG below 36%	count, n	579	25	604
	within increased percentage	95.90%	4.10%	100%
	within MTX treatment outcome	95.10%	17.50%	89.20%
	total	80.20%	3.50%	89.20%
Increased serum $\beta$ -hCG above 36%	count, n	30	88	118
	within increased percentage	25.40%	74.60%	100%
	within MTX treatment outcome	4.90%	61.50%	10.70%
	total	4.10%	12.20%	10.70%
Total	count, n	609	113	722
	within increased percentage	84.30%	16.70%	100%
	within MTX treatment outcome	100%	100%	100%
	total	84.30%	16.70%	100%

specificity 78.9%, positive predictive value 88.6%, and negative predictive value 56.6%).<sup>7</sup> Some researchers found a transient increase in serum  $\beta$ -hCG levels of 26–60% due to a single dose of MTX.<sup>22–24</sup> On the other hand, Helvacioğlu and Doğan suggested that a 2-dose MTX protocol had a reasonable success rate on serum  $\beta$ -hCG levels.<sup>8</sup> In a study examining serum  $\beta$ -hCG, progesterone levels and patient characteristics, they observed that the success rates were similar in both groups. They achieved success with MTX in 83% of patients and with placebo in 76%, and found that this difference was not significant in their statistical analysis. They reported that the MTX treatment was successful in 125 patients and unsuccessful in 37. When the serum  $\beta$ -hCG values measured on days 0, 4 and 7 were compared in pairs, the difference was reported to be statistically significant ( $p < 0.001$ ). The mean serum  $\beta$ -hCG value was 783.0 in the success group and 1802.0 in the failure group ( $p < 0.001$ ). However, they reported that the increase in serum  $\beta$ -hCG values during the first 4 days was 21.6% in the success group and 25.7% in the failure group ( $p < 0.001$ ). This paper reported that MTX treatment was successful in 117 patients (73.6%) out of 159 women included in the study, while it was unsuccessful in 42 patients (26.4%). They observed that most of the patients (58.9%) had a tendency for decreased serum  $\beta$ -hCG values during the first 4 days, unlike the failure group, which had an increase in serum  $\beta$ -hCG values (76.2%). These researchers accepted a  $>18\%$  decrease in serum  $\beta$ -hCG levels as an indicator of treatment success.<sup>8</sup>

Therefore, the key to this problem was the percentage of serum  $\beta$ -hCG change. In our study, 64.9% (737/1135) of patients had an increased serum  $\beta$ -hCG level in the first 4 days, with an initial level of less than 2000 mIU/mL, and all patients were selected strictly according to the inclusion criteria. However, treatment success was achieved in 82.6% (609/737) of these cases. Furthermore, in our study, the most influential covariates were 1) the ratio of day 4 to day 1 serum  $\beta$ -hCG level and 2) the cutoff value of a 36%

increase on day 4, which predicted treatment success with a sensitivity of 72.57%, a specificity of 86.86% and an AUC of 0.8083 (95% CI: 0.7480–8587). For this reason, increased serum  $\beta$ -hCG levels on day 4 after the administration of MTX cannot be considered a criterion for treatment failure. This success could be explained by the theory that the vast majority of trophoblastic cells died after MTX therapy, released hCG, and then the MTX therapy failed. This was also demonstrated by our statistical data. Upon exceeding the cutoff value of the percentage increase, the greater the percentage increase on day 4, the lower the prediction sensitivity. This was the answer we had been looking for.

Other studies have shown decreased serum  $\beta$ -hCG levels within the first 4 days after the MTX treatment, with success rates ranging from 88% to 100%, which is consistent with our 84.82% success rate.<sup>22–24</sup> Brunello et al. determined the threshold value for a decrease in serum  $\beta$ -hCG after the MTX treatment to be 10–22%, whereas in our study this rate was determined as 13%.<sup>25</sup> The difference in cutoff values among these studies might be due to differences between patients (especially between different ethnic groups), inclusion/exclusion criteria, or definition of the treatment success. However, some studies have suggested that the lower the serum  $\beta$ -hCG level, the higher the probability of treatment success.<sup>26–28</sup> In contrast, Gabbur et al. suggested that  $\beta$ -hCG levels on the 4<sup>th</sup> day could not be evaluated as an indicator of treatment success.<sup>29</sup> Similarly, Cohen et al. concluded that neither absolute serum  $\beta$ -hCG levels nor percent changes in serum  $\beta$ -hCG levels within the first 4 days could predict treatment success after a single dose of MTX.<sup>30</sup> In other studies, serum  $\beta$ -hCG values on day 4 and 7 were a better diagnostic criterion than the percentage decrease of serum  $\beta$ -hCG values between days 1 and 4. Çelik et al. reported that a decreased percentage between days 1 and 4 did not seem to be a better predictor than that between days 4 and 7.<sup>31</sup> Girija et al. reported that the MTX treatment could be successful not only in patients whose serum  $\beta$ -hCG levels

decreased on the 4<sup>th</sup> day, but also those whose serum  $\beta$ -hCG level increased on the 4<sup>th</sup> day. They observed that a >50% increase in serum  $\beta$ -hCG levels in the first 4 days significantly increased the risk of MTX treatment failure, as in our study.<sup>32</sup> Although the sensitivity of the cutoff value for the rise in serum  $\beta$ -hCG levels from day 1 to day 4 was higher in our study, we should not make a decision arbitrarily. We took the 48-hour pretreatment increase in serum  $\beta$ -hCG into consideration. In our research, the best cutoff value for the rise in the serum  $\beta$ -hCG levels 48 h before treatment was 19%. This cutoff value provided a sensitivity of 61.95%, a specificity of 94.75% and an AUC of 0.7495 (95% CI: 0.6814–0.8177). The greater the increase 48 h before treatment, the higher the specificity.

If we had applied the 36% threshold level of serum  $\beta$ -hCG increase on the 4<sup>th</sup> day after a single dose of MTX, all patients who exceeded the 36% would have undergone surgery. However, a total of 55 patients would not have been included and the correct protocol would not have been applied. Based on this statistically significant result, we are aware that the most appropriate treatment in patients with serum  $\beta$ -hCG levels below 36% is conservative care. However, when serum  $\beta$ -hCG levels rise above 36%, the decision for surgery cannot be made definitively. Once we meet these criteria, we must pay attention to the 48-hour pretreatment increase in serum  $\beta$ -hCG. If an increase of more than 19% in 48-hour pretreatment serum  $\beta$ -hCG occurs, the possibility of surgery or a 2<sup>nd</sup> dose of MTX should be considered. By applying the increased serum  $\beta$ -hCG cutoff level, we were able to reduce the failure rate of patients with serum  $\beta$ -hCG levels that had an increase below 36% to 4.2%. However, we found in our study that 25.5% of patients with serum  $\beta$ -hCG levels above 36% were treated surgically and a single dose of MTX therapy was indeed successful. According to the results of our research, thanks to the increased serum  $\beta$ -hCG levels on the 4<sup>th</sup> day after a single dose of MTX treatment, the patient's anxiety can be minimized and unnecessary surgeries reduced.

## Limitations

The main limitations of our study are its retrospective nature and lack of case-control studies with a large number of patients for reference. It is obvious that there is a need for more comprehensive and prospective research on this population.

## Conclusions

An increment of more than 19% in the 48 h before MTX treatment and an increase of more than 36% between the day 1 and 4  $\beta$ -hCG concentrations were found to be good predictors of treatment failure. We think that we applied the treatment in the most accurate way in the follow-up evaluation after MTX treatment.

## ORCID iDs

Huijuan Zhou  <https://orcid.org/0000-0001-5358-8010>  
Shuangdi Li  <https://orcid.org/0000-0002-7524-5753>

## References

1. Stovall TG, Ling FW, Gray LA. Single-dose methotrexate for treatment of ectopic pregnancy. *Obstet Gynecol.* 1991;77(5):754–757. PMID:2014091.
2. Jurkovic D, Memtsa M, Sawyer E, et al. Single-dose systemic methotrexate vs expectant management for treatment of tubal ectopic pregnancy: A placebo-controlled randomized trial. *Ultrasound Obstet Gynecol.* 2017;49(2):171–176. doi:10.1002/uog.17329
3. Yıldız A, Bilge O. The importance of  $\beta$ -hCG values in prediction of the effectiveness of single dose methotrexate therapy in tubal ectopic pregnancy [published online as ahead of print on March 22, 2022]. *Ginekol Pol.* 2022. doi:10.5603/GP.a2021.0247
4. Racková J, Driák D, Neumannová H, Hurt K, Sehnal B, Halaška M. Use of methotrexate in the ectopic pregnancy and pregnancy of unknown location [in Czech]. *Ceska Gynekol.* 2016;81(2):140–146. PMID:27457397.
5. Shatkin Hamish N, Wolf M, Tendler R, Sharon A, Bornstein J, Odeh M. Early prediction of methotrexate treatment outcome in tubal ectopic pregnancy based on days 0 and 4 human chorionic gonadotropin levels. *J Obstet Gynaecol Res.* 2020;46(7):1104–1109. doi:10.1111/jog.14259
6. Hoyos LR, Malik M, Najjar M, et al. Morbid obesity and outcome of ectopic pregnancy following capped single-dose regimen methotrexate. *Arch Gynecol Obstet.* 2017;295(2):375–381. doi:10.1007/s00404-016-4229-0
7. Akselim B, Karaşin S, Sabir Y. Can the additional dose requirement be determined before the seventh day to successfully treat tubal ectopic pregnancy with a single-dose methotrexate protocol? *J Coll Physicians Surg Pak.* 2021;31(9):1046–1050. doi:10.29271/jcpsp.2021.09.1046
8. Helvacioğlu C, Doğan K. Predictive factors of treatment success in two-dose methotrexate regimen in ectopic tubal pregnancy: A retrospective study. *Pak J Med Sci.* 2021;37(5):1309–1312. doi:10.12669/pjms.37.5.4299
9. Prabhakaran M, Beesetty A. Ectopic pregnancy with low beta-human chorionic gonadotropin (HCG) managed with methotrexate and progressed to rupture. *Cureus.* 2021;13(10):e18749. doi:10.7759/cureus.18749
10. Lin C, Hsieh H. Early ectopic pregnancy refractory to methotrexate treatment: A case report. *Cureus.* 2021;13(11):e19686. doi:10.7759/cureus.19686
11. Bhat CS, Reddy NS, Vembu R, Pandurangi M. Dual extrauterine ectopic pregnancy: Double management. *BMJ Case Rep.* 2021;14(11):e244417. doi:10.1136/bcr-2021-244417
12. Cohen A, Baron S, Cohen Y, et al. Ruptured ectopic pregnancies following methotrexate treatment: Clinical course and predictors for improving patient counseling. *Reprod Sci.* 2022;29(4):1209–1214. doi:10.1007/s43032-022-00881-7
13. Stabile G, Romano F, Zinicola G, et al. Interstitial ectopic pregnancy: The role of mifepristone in the medical treatment. *Int J Environ Res Public Health.* 2021;18(18):9781. doi:10.3390/ijerph18189781
14. Grigoriu C, Bohiltea R, Mihai B, et al. Success rate of methotrexate in the conservative treatment of tubal ectopic pregnancies. *Exp Ther Med.* 2021;23(2):150. doi:10.3892/etm.2021.11073
15. Hayashi S, Takeda A. Gasless laparoendoscopic single-site surgery for management of unruptured tubal pregnancy in a woman with moderate COVID-19 pneumonia after administration of remdesivir and casirivimab-imdevimab: A case report. *Case Rep Womens Health.* 2022;33:e00368. doi:10.1016/j.crwh.2021.e00368
16. Fernandez H, Capmas P, Lucot JP, et al. Fertility after ectopic pregnancy: The DEMETER randomized trial. *Hum Reprod.* 2013;28(5):1247–1253. doi:10.1093/humrep/det037
17. Marret H, Fauconnier A, Dubernard G, et al. Overview and guidelines of off-label use of methotrexate in ectopic pregnancy: A report by CNGOF. *Eur J Obstet Gynecol Reprod Biol.* 2016;205:105–109. doi:10.1016/j.ejogrb.2016.07.489
18. Bonin L, Pedreiro C, Moret S, Chene G, Gaucherand P, Lamblin G. Predictive factors for the methotrexate treatment outcome in ectopic pregnancy: A comparative study of 400 cases. *Eur J Obstet Gynecol Reprod Biol.* 2017;208:23–30. doi:10.1016/j.ejogrb.2016.11.016

19. Baggio S, Garzon S, Russo A, et al. Fertility and reproductive outcome after tubal ectopic pregnancy: Comparison among methotrexate, surgery and expectant management. *Arch Gynecol Obstet.* 2021;303(1):259–268. doi:10.1007/s00404-020-05749-2
20. Naveed AK, Anjum MU, Hassan A, Mahmood SN. Methotrexate versus expectant management in ectopic pregnancy: A meta-analysis. *Arch Gynecol Obstet.* 2022;305(3):547–553. doi:10.1007/s00404-021-06236-y
21. Elson C, Salim R, Potdar N, Chetty M, Ross J, Kirk E. Diagnosis and management of ectopic pregnancy: Green-top Guideline No. 21. *BJOG.* 2016;123(13):e15–e55. doi:10.1111/1471-0528.14189
22. Brunello J, Guerby P, Cartoux C, et al. Can early  $\beta$ hCG change and baseline progesterone level predict treatment outcome in patients receiving single dose methotrexate protocol for tubal ectopic pregnancy? *Arch Gynecol Obstet.* 2019;299(3):741–745. doi:10.1007/s00404-019-05068-1
23. Levin G, Rottenstreich A. Earlier predictors for treatment outcome among single dose methotrexate for an ectopic pregnancy. *Arch Gynecol Obstet.* 2019;300(3):793–793. doi:10.1007/s00404-019-05250-5
24. Goh A, Karine P, Kirby A, Williams C, Kapurubandara S. Day 1 to day 4 serum hCG change in predicting single-dose methotrexate treatment failure for tubal ectopic pregnancies. *Eur J Obstet Gynecol Reprod Biol.* 2020;255:105–110. doi:10.1016/j.ejogrb.2020.10.036
25. Brunello J, Guerby P, Cartoux C, et al. Early predictive factors of single dose methotrexate outcome in patients with ectopic pregnancy. *Arch Gynecol Obstet.* 2019;300(3):795. doi:10.1007/s00404-019-05255-0
26. Levin G, Dior U, Shushan A, Gilad R, Benshushan A, Rottenstreich A. Early prediction of the success of methotrexate treatment success by 24-hour pretreatment increment in HCG and day 1–4 change in HCG. *Reprod BioMed Online.* 2019;39(1):149–154. doi:10.1016/j.rbmo.2019.02.005
27. Shatkin Hamish N, Wolf M, Tendler R, Sharon A, Bornstein J, Odeh M. Early prediction of methotrexate treatment outcome in tubal ectopic pregnancy based on days 0 and 4 human chorionic gonadotropin levels. *J Obstet Gynaecol Res.* 2020;46(7):1104–1109. doi:10.1111/jog.14259
28. Tasgoz FN, Temur M, Dundar B, Kartal E, Ustunyurt E. The role of day 0 and day 4  $\beta$ -human chorionic gonadotropin values and initial ultrasound findings in predicting the success of methotrexate treatment in ectopic pregnancy. *Ginekol Pol.* 2020;91(7):389–393. doi:10.5603/GP.2020.0071
29. Gabbur N, Sherer D, Hellmann M, Abdelmalek E, Phillip P, Abulafia O. Do serum beta-human chorionic gonadotropin levels on day 4 following methotrexate treatment of patients with ectopic pregnancy predict successful single-dose therapy? *Amer J Perinatol.* 2006;23(3):193–196. doi:10.1055/s-2006-934097
30. Cohen A, Bibi G, Almog B, Tsafrir Z, Levin I. Second-dose methotrexate in ectopic pregnancies: The role of beta human chorionic gonadotropin. *Fertil Steril.* 2014;102(6):1646–1649. doi:10.1016/j.fertnstert.2014.08.019
31. Çelik H, Tosun M, Işık Y, Avcı B. The role of early determination of  $\beta$ -human chorionic gonadotropin levels in predicting the success of single-dose methotrexate treatment in ectopic pregnancy. *Ginekol Pol.* 2016;87(7):484–487. doi:10.5603/GP.2016.0030
32. Girija S, Manjunath AP, Salahudin A, et al. Role of day 4 HCG as an early predictor of success after methotrexate therapy for ectopic pregnancies. *Eur J Obstet Gynecol Reprod Biol.* 2017;215:230–233. doi:10.1016/j.ejogrb.2017.06.020

# Wrist motion assessment using Microsoft Azure Kinect DK: A reliability study in healthy individuals

Aleksandra Królikowska<sup>1,A–F</sup>, Agnieszka Maj<sup>2,B,D–F</sup>, Maciej Dejneka<sup>2,3,B,D–F</sup>, Robert Prill<sup>4,C–F</sup>, Anna Skotowska-Machaj<sup>5,B,D–F</sup>, Anna Kołcz<sup>1,C–F</sup>

<sup>1</sup> Ergonomics and Biomedical Monitoring Laboratory, Department of Physiotherapy, Faculty of Health Sciences, Wrocław Medical University, Poland

<sup>2</sup> Clinical Department of Trauma and Hand Surgery, Jan Mikulicz-Radecki University Hospital, Wrocław, Poland

<sup>3</sup> Department of Trauma Surgery, Faculty of Medicine, Wrocław Medical University, Poland

<sup>4</sup> Center of Orthopaedics and Traumatology, University Hospital Brandenburg/Havel Brandenburg Medical School, Brandenburg an der Havel, Germany

<sup>5</sup> Department of Rehabilitation, Primate Cardinal Wyszyński Regional Hospital, Sieradz, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;

D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):203–209

## Address for correspondence

Aleksandra Królikowska

E-mail: [aleksandra.krolikowska@umw.edu.pl](mailto:aleksandra.krolikowska@umw.edu.pl)

## Funding sources

This research was financially supported by the Ministry of Health subvention according to the number of STM.E067.20.098 from the IT Simple system of Wrocław Medical University.

## Conflict of interest

None declared

Received on July 5, 2022

Reviewed on August 1, 2022

Accepted on August 18, 2022

Published online on September 22, 2022

## Cite as

Królikowska A, Maj A, Dejneka M, Prill R, Skotowska-Machaj A, Kołcz A. Wrist motion assessment using Microsoft Azure Kinect DK: A reliability study in healthy individuals.

*Adv Clin Exp Med.* 2023;32(2):203–209.

doi:10.17219/acem/152884

## DOI

10.17219/acem/152884

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Motion analysis systems have been widely used in orthopedics and rehabilitation for diagnostics, patient monitoring and outcome evaluation purposes. Since Microsoft Azure Kinect Developer Kit (DK) had been released, only a few studies were published concerning its usage. However, it has not been used for wrist motion assessments, even though the use of standardized examinations with known reliability, validity and responsiveness remains a constant challenge.

**Objectives.** This study aimed to examine the reliability of hand and forearm range of motion (ROM) measurements recorded using new software utilizing the Microsoft Azure Kinect DK.

**Materials and methods.** Twenty-eight healthy males and 28 healthy females participated in measurements of active ROM for wrist extension, wrist flexion, radial deviation, ulnar deviation, and forearm supination and pronation on 3 separate occasions. Sessions 1 and 2 were carried out on the same day with a 90-minute rest period between each session, while the 3<sup>rd</sup> session was conducted a week later. Data were recorded simultaneously in both limbs using a custom-made software developed by a software development company (Oleky Medical & Sports Sciences, Łańcut, Poland) for the purposes of the present study using Microsoft Azure Kinect DK. The assessment of intra-day and inter-day reliability was based on intraclass correlation coefficient (ICC) calculations and interpreted based on commonly used guidelines.

**Results.** In the group of males, the lowest ICC was 0.846 for intra-day comparisons and 0.816 for inter-day analyses. In the female group, the lowest ICC for intra-day comparisons was 0.826 and exceeded 0.833 for inter-day comparisons.

**Conclusions.** The developed software using Microsoft Azure Kinect DK demonstrated high reliability in measuring wrist and forearm active ROM. These promising results support the use of Microsoft Azure Kinect DK in a clinical capacity as a potential hand assessment and rehabilitation tool.

**Key words:** rehabilitation, orthopedics, hand, patient monitoring, motion

## Background

Motion analysis systems have been widely used in orthopedics and rehabilitation for diagnostics, monitoring progress and evaluating outcomes. Optoelectronic motion capture systems based on markers have generally been considered the gold standard for motion analysis research.<sup>1–3</sup> One proposed alternative to optoelectronic systems is human motion analysis using wearable inertial sensors.<sup>4,5</sup> Also, markerless motion capture systems have been validated and applied to functional diagnostics and rehabilitation.<sup>6–8</sup>

The common commercially available sensors used for markerless motion analysis are Leap Motion Controller (LMC; Leap Motion Inc., San Francisco, USA) and consecutive versions of Microsoft Kinect sensors (Kinect for Windows v. 1 (V1) and Kinect for Windows v. 2 (V2); Microsoft Corp., Redmond, USA).<sup>9,10</sup> For center-based and home rehabilitation purposes, Nintendo Wii Mote, Wii Balance Board (Nintendo Inc., Redmond, USA) and Microsoft Kinect motion-sensing game controllers have gained popularity over the past years.<sup>11–14</sup>

Since Microsoft had released Azure Kinect Developer Kit (DK) in 2019, only a few studies were published concerning its usage. The literature describes its use in rehabilitation-related purposes and hand gesture recognition.<sup>15–18</sup> However, it has not been used for hand motion assessments, even though standardized examinations with known reliability, validity and responsiveness remain a constant challenge in physiotherapy.<sup>19</sup> When evaluating the reliability of a new testing method, it is primarily performed on healthy individuals to make sure that repeated trials on the same participants are reproducible.<sup>20</sup>

## Objectives

This study aimed to evaluate the reliability of wrist and forearm range of motion (ROM) measurements recorded using new software developed for Microsoft Azure Kinect DK.

## Materials and methods

### Study design

The test–retest reliability was employed to determine the variation in the ROM measurements assessed using custom-developed software and Microsoft Azure Kinect DK in the same participant under the same conditions. The development of the software for the purposes of the present study was outsourced to a software development company (Oleky Medical & Sports Sciences, Łańcut, Poland) that designed it according to the predefined specification. The test–retest design was chosen as it assumes that in a software reliability assessment the effect of an examiner is rather

negligible. Therefore, levels of consistency were precisely noted between 2 measurements taken on the same day (intra-day reliability) and 2 measurements taken on different days (inter-day reliability).

### Setting

The present study was performed in 2020–2021 in an academic setting. Each participant took part in 3 measurement sessions. Sessions 1 and 2 were performed on the same day, separated by a 90-minute rest period. Session 3 was carried out exactly a week later. A flowchart of the study is presented in Fig. 1.

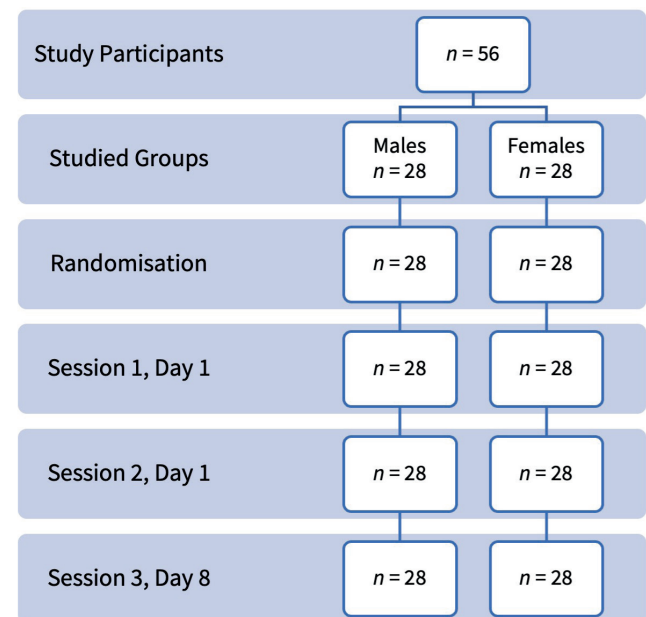


Fig. 1. Flowchart of the study

### Participants

The first 28 male and 28 female students from the university who volunteered to participate and met all the inclusion criteria comprised the study group. In all studied participants, an interview and physical examination were carried out to assess their general health and musculoskeletal system condition. The inclusion criteria encompassed a dominant right upper limb, lack of musculoskeletal disorders and injuries to the upper limbs, cervical spine and thoracic spine, and no diagnosed general diseases. Upper limb dominance was determined by asking the participants which hand they use to write. During the 1<sup>st</sup> session, all participants underwent body height and mass measurements and their body mass index (BMI) was calculated.

### Variables

The measured variable was a maximal value of wrist extension, wrist flexion, radial deviation, ulnar deviation,

forearm supination, and forearm pronation. The results were expressed in degrees [°], and recorded separately for dominant and non-dominant upper limbs during 3 separate measurement sessions.

## Measurement

The variables mentioned above were obtained during the measurements of active wrist ROM. The measurements were carried out on the wrists simultaneously and bilaterally. The 3 measurement sessions were performed by the same examiner who was trained and experienced in the use of the developed software. The examiner asked participants not to change their regular training regimens during the 8-day-long experiment. The sessions were conducted at precisely the same time in the morning.

Data recorded by the developed software for the present study using Microsoft Azure Kinect DK is presented in Fig. 2. Microsoft Azure Kinect DK is a device that uses infrared light to generate in-depth images that give a 3D representation after a process called reconstruction. The custom application was written using a Unity-based engine. A custom plugin was created to include the Microsoft Software Development Kit (SDK) for the Kinetic Azure device.

One Microsoft Azure Kinect DK was set up on the desk using a frame that faced down to the desk, as shown in Fig. 3. Metcalf et al. previously suggested a similar method.<sup>6</sup> The participants sat by the desk with both arms close to their bodies. The elbows were flexed 90°. Their forearms were placed on the desk, parallel to one another.

For measuring wrist extension and flexion, the participants placed their forearms in a midline position, also

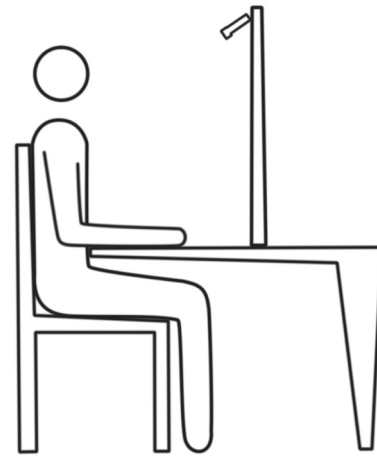


Fig. 3. The participant's position during measurements of wrist and forearm range of motion simultaneously in both upper limbs using Microsoft Azure Kinect DK

called neutral or zero position. At first, the participants extended their wrists maximally, and their fingers were allowed to flex passively. Next, the participants returned to the starting position and consecutively flexed their hands maximally.

The starting position for measuring radial and ulnar deviation was with the forearm in pronation and the hand in neutral. First, the participants performed maximal radial deviation. Then, from the starting position, they performed maximal ulnar deviation.

The forearm was then placed in the midline position for measuring rotational movements. First, the participants rotated their forearms to maximum supination. Then, after



Fig. 2. Front view (A), lateral view (B), top view (C), and general view (D) of Microsoft Azure Kinect DK

placing the forearm in a midline position, the participants rotated the forearm into maximum pronation.

Verbal commands were used during the assessment. The participant performed 3 repetitions of each movement within the full ROM; however, the parameter taken for analysis was the maximal value obtained during the 3 repetitions.

## Bias

To address a potential source of bias, the reliability of wrist and forearm ROM measurements were recorded separately for males and females. Additionally, the statistician was blinded regarding the gender of the participants, as well to the type of the analysis (intra-day and inter-day reliability).

## Study size

The minimal sample size was determined using the method described by Bujang and Baharum.<sup>21</sup> Two measurements were performed for intra-day and 2 for inter-day reliability assessment purposes, with a statistical power of 80%, and a minimum intraclass correlation coefficient (ICC) of 0.50 required a minimal number of participants of 22. An additional 20% of participants were added in case of dropouts between the sessions, also according to the Bujang and Baharum's guidelines.<sup>21</sup>

## Quantitative variables

The maximal values of wrist and forearm motion between sessions 1 and 2 were taken for comparative analysis for the intra-day comparison. The results from the 1<sup>st</sup> and 3<sup>rd</sup> session were considered to determine inter-day reliability.

## Statistical analyses

The statistical analysis was carried out using IBM SPSS Statistics Premium v. 28 (IBM Corp., Armonk, USA) and Microsoft Office Excel 365 Personal (Microsoft Corp.).

Calculation of the arithmetic mean ( $\bar{x}$ ) and standard deviation (SD) was performed for the participant's age, body height, body mass, and BMI. Using the Shapiro–Wilk test, the data of the maximal values of assessed movements, expressed in degrees [°], were found to be normally distributed. Two-way mixed effects, absolute agreement and single examiner/measurement form of the ICC, determined by McGraw and Wong, were used.<sup>22</sup> According to guidelines by Koo and Li, the ICC interpretation was based on an estimated ICC 95% confidence interval (95% CI).<sup>23</sup> The ICC values were also interpreted with reference to the criteria proposed by Cicchetti and Sparrow as follows: ICC < 0.40 – poor; ICC = 0.40–0.59 – fair; ICC = 0.60–0.74 – good, and ICC ≥ 0.75 – excellent.<sup>24</sup> One-way analysis of variance (ANOVA) and F tests were used for rating the significance of the ICC. A p-value <0.05 was considered statistically significant.

## Results

As presented in Fig. 1, 28 males (age  $\bar{x}$  = 22.32 ± 0.99 years, body height  $\bar{x}$  = 178.77 ± 4.10 cm, body mass  $\bar{x}$  = 76.36 ± 3.86 kg, and BMI  $\bar{x}$  = 23.88 ± 0.60 kg/m<sup>2</sup>) and 28 females (age  $\bar{x}$  = 21.73 ± 1.24 years, body height  $\bar{x}$  = 169.36 ± 5.06 cm, body mass  $\bar{x}$  = 60.82 ± 5.06 kg, and BMI  $\bar{x}$  = 21.18 ± 1.08 kg/m<sup>2</sup>) took part in 3 consecutive sessions. There were no dropouts from the study and no missing data to be addressed.

As presented in Table 1, in the group of males, the intra-day reliability in the dominant and non-dominant limbs was rated good to excellent. The lower bound of ICC values

**Table 1.** The intra-day test–retest reliability assessment results of the hand and wrist range of motion measurements recorded with the developed software and using Microsoft Azure Kinect DK

Measured motion	Studied limb	Males			Females		
		ICC (95% CI)	F	p-value	ICC (95% CI)	F	p-value
Hand extension	dominant	0.846 (0.659–0.934)	13.232	≤0.001	0.924 (0.827–0.968)	24.390	≤0.001
	non-dominant	0.871 (0.691–0.946)	16.910	≤0.001	0.886 (0.749–0.951)	16.462	≤0.001
Hand flexion	dominant	0.894 (0.701–0.959)	23.052	≤0.001	0.863 (0.701–0.941)	13.312	≤0.001
	non-dominant	0.911 (0.607–0.971)	34.578	≤0.001	0.826 (0.631–0.923)	10.593	≤0.001
Hand radial deviation	dominant	0.902 (0.677–0.964)	27.208	≤0.001	0.967 (0.922–0.986)	64.510	≤0.001
	non-dominant	0.959 (0.903–0.983)	45.193	≤0.001	0.975 (0.942–0.990)	78.415	≤0.001
Hand ulnar deviation	dominant	0.939 (0.860–0.974)	32.298	≤0.001	0.853 (0.679–0.936)	13.399	≤0.001
	non-dominant	0.940 (0.862–0.975)	31.124	≤0.001	0.852 (0.672–0.936)	13.749	≤0.001
Forearm supination	dominant	0.977 (0.881–0.933)	140.812	≤0.001	0.936 (0.824–0.975)	37.197	≤0.001
	non-dominant	0.988 (0.971–0.995)	162.303	≤0.001	0.943 (0.869–0.976)	35.165	≤0.001
Forearm pronation	dominant	0.944 (0.862–0.977)	39.247	≤0.001	0.929 (0.762–0.974)	37.297	≤0.001
	non-dominant	0.908 (0.737–0.964)	26.597	≤0.001	0.879 (0.735–0.948)	15.393	≤0.001

F – F test value; ICC – intraclass correlation coefficient and 95% confidence interval (95% CI) of the ICC.



**Table 2.** The inter-day test–retest reliability assessment results of the hand and wrist range of motion measurements recorded with the developed software and using Microsoft Azure Kinect DK

Measured motion	Studied limb	Males			Females		
		ICC (95% CI)	F	p-value	ICC (95% CI)	F	p-value
Hand extension	dominant	0.886 (0.741–0.951)	18.114	≤0.001	0.879 (0.731–0.948)	16.575	≤0.001
	non-dominant	0.865 (0.706–0.941)	13.588	≤0.001	0.907 (0.783–0.961)	22.687	≤0.001
Hand flexion	dominant	0.910 (0.793–0.962)	22.671	≤0.001	0.857 (0.688–0.938)	12.569	≤0.001
	non-dominant	0.871 (0.718–0.944)	8.843	≤0.001	0.833 (0.642–0.927)	10.718	≤0.001
Hand radial deviation	dominant	0.891 (0.727–0.955)	20.575	≤0.001	0.964 (0.917–0.985)	53.734	≤0.001
	non-dominant	0.883 (0.743–0.949)	16.468	≤0.001	0.917 (0.794–0.966)	26.799	≤0.001
Hand ulnar deviation	dominant	0.945 (0.616–0.984)	68.422	≤0.001	0.894 (0.686–0.960)	23.616	≤0.001
	non-dominant	0.863 (0.692–0.946)	17.476	≤0.001	0.879 (0.900–0.993)	141.537	≤0.001
Forearm supination	dominant	0.971 (0.926–0.988)	77.365	≤0.001	0.935 (0.848–0.973)	31.920	≤0.001
	non-dominant	0.964 (0.912–0.985)	61.065	≤0.001	0.942 (0.820–0.978)	43.472	≤0.001
Forearm pronation	dominant	0.822 (0.624–0.922)	10.664	≤0.001	0.893 (0.756–0.954)	19.241	≤0.001
	non-dominant	0.816 (0.602–0.920)	10.802	≤0.001	0.851 (0.674–0.935)	11.896	≤0.001

F – F test value; ICC – intraclass correlation coefficient and 95% confidence interval (95% CI) of the ICC.

was noted in wrist flexion of the non-dominant limb and exceeded 0.607. The highest upper bound of ICC values were observed in non-dominant limb forearm supination and exceeded 0.995. In the female group, the intra-day reliability was good to excellent, with a lower bound of ICC of 0.631 for wrist flexion in the non-dominant limb and the upper bound of ICC of 0.990 for radial deviation in the non-dominant limb.

As presented in Table 2, the inter-day reliability was good to excellent in both males and females. For the male group, the lower bound was noted for ulnar deviation in the dominant limb and exceeded 0.616. The upper bound exceeded 0.988 and was reported in forearm supination of the dominant limb. In females, the lower bound of ICC was 0.642 for wrist flexion in the non-dominant limb. The upper bound of ICC was 0.993 for ulnar deviation in the non-dominant limb.

## Discussion

Our study determined that the new software utilizing Microsoft Azure Kinect DK demonstrates high reliability in measuring wrist movements in the sagittal and frontal planes and forearm rotational movements.

In motion analysis, the quality and quantitative parameters of movement can be distinguished. The most common quantitative parameters used in clinical rehabilitation include biomechanical variables like postural angles, pressure distribution, moments, and forces produced by examined patients.<sup>5</sup> Motion capture has been utilized in numerous fields, including biomechanical analysis of clinical populations with hand and wrist dysfunctions. The so-called gold standard optoelectronic motion capture systems are based on camera tracking and data collection on the position of retroreflective markers placed

on the body of the examined patient. The positional data has been used to perform biomechanical analysis in static and dynamic conditions.<sup>1–3,25</sup> However, marked motion capture systems have some limitations. They are expensive and few clinical centers can afford them. The placement of the markers on the hand is also challenging. Apart from the fact that such measurements are time-consuming, there are patients with diseases in which the placement of markers can cause difficulties, such as in patients with complex regional pain syndrome (CRPS).<sup>26</sup>

Of course, there are some non-marker-based possibilities available to track wrist motion, like instrumented gloves. They are less expensive than marker-based systems and their usage is not limited to the laboratory setting.<sup>27,28</sup> Although the gloves come in different sizes, they cannot be perfectly adjusted to every patient, especially considering various deformities from hand diseases and injuries. Moreover, the gloves are in complete contact with the patient's skin, which can influence the performed movement. There are also conditions where the contact of the glove with the patient's skin is contraindicated.<sup>9</sup>

Another option for hand motion analysis is computer vision-based systems. An example of this system is Monochrome Egocentric Articulated Hand-Tracking for Virtual Reality (MEgATrack), which was introduced by Han et al.<sup>29</sup> It should be noted that studies using this system are promising; however, there is still a need for its validation in clinical settings.

The LMC, released in 2013, was supposed to allow the analysis of hand motion at a low cost.<sup>9</sup> However, its utility was initially criticized because the area of motion tracking was relatively limited. Moreover, the examination capacity itself was also limited because a lot depended on the placement of the device.<sup>30</sup> However, promising results appeared in studies by Houston et al., assessing an updated

LMC released in 2018, that allowed for data capturing using multiple cameras in one setup.<sup>9</sup>

In 2010, Microsoft released Kinect V1 with an integrated RGB and infrared (IR) camera, primarily designed as a gaming controller for Xbox. The Kinect V1 allowed users to interact with games without a controller device.<sup>31</sup> Four years later, an improved version of the Kinect V1 was launched; however, Microsoft halted the production of this device.<sup>32</sup>

Because the Microsoft Azure Kinect DK is a relatively new sensor option, the literature supporting its use for rehabilitation-related purposes is limited.<sup>15–18</sup> In a study by Albert et al., a comparative analysis between the Microsoft Azure Kinect DK and Microsoft Kinect V2 was carried out by performing gait analysis on a treadmill. Additionally, a gold standard optoelectronic Vicon motion capturing system was used. The study supported the usage of Microsoft Azure Kinect DK and Microsoft Kinect V2 for the analysis of temporal gait parameters. However, the Microsoft Azure Kinect DK superiority in spatial gait parameters was highlighted.<sup>17</sup> Also, Yeung et al. supported the use of Microsoft Azure Kinect DK for tracking gait patterns while walking on a treadmill, and compared the results to the Microsoft Kinect V2 and Orbbec Astra Pro v. 2.<sup>16</sup> Antico et al. examined the accuracy of Microsoft Azure Kinect DK in examining postural control and concluded it is a promising technique in the future of home rehabilitation.<sup>15</sup> Apart from direct rehabilitation-based purposes, Lee et al. developed a real-time system using Microsoft Azure Kinect DK for gesture recognition.<sup>18</sup>

As presented above, some studies support the usage of Microsoft Azure Kinect DK for rehabilitation-related purposes; however, none of the authors reported reliability measurements performed neither in different sessions on the same day nor in sessions performed on different days.<sup>15–17</sup> Therefore, it is impossible to compare the results obtained in the present study to other studies. Also, all the mentioned studies had very small study groups compared to this paper. Antico et al. conducted their study on 26 participants (18 males and 8 females), Yeung et al. on 10 healthy participants (2 females and 8 males), and Albert et al. on 5 participants,<sup>15–17</sup> while Lee et al. assessed the defined features in a group of 10 people.<sup>18</sup>

As the present study supports the utility of Microsoft Azure Kinect DK for diagnostic and monitoring purposes, another interesting issue that might attract researchers' attention is the possibility of using it to assess the functional direction of wrist joint motion. It has been determined that most activities of daily life are performed using functional wrist oblique motion that occurs in the plane of the dart-throwing motion (DTM), specifically from radial extension to ulnar flexion.<sup>33–38</sup> So far, inertial measurement units, optoelectronics systems with reflective skin markers and standard goniometry have been used to measure the range of the DTM.<sup>33,39,40</sup>

## Limitations





The main limitation of our study is that the accuracy of the Microsoft Azure Kinect DK has not been compared to an optoelectronic system or other motion sensors. So far, the Microsoft Azure Kinect DK has been compared to the Vicon 3D system and Microsoft Kinect V2 in the context of treadmill walking by Albert et al.,<sup>17</sup> and to the Vicon 3D system in the context of the assessment of postural control by Antico et al.<sup>15</sup> Therefore, when using Microsoft Azure Kinect DK, raw data should not be compared to raw data recorded by another sensor until research on this topic is available.

As in our study, reliability tests are performed on healthy individuals. The next step should be to validate the wrist and forearm ROM measurements using Microsoft Azure Kinect DK on a group of patients with hand and wrist dysfunctions to confirm its clinical capacity as a hand assessment device and hand therapy tool.

## Conclusions

The developed software using Microsoft Azure Kinect DK demonstrated high reliability in measuring hand and forearm active ROM. The promising results support Microsoft Azure Kinect becoming a potential hand assessment and rehabilitation tool. However, it should be confirmed in studies using a clinical population of patients with hand and wrist dysfunctions.

## ORCID iDs

Aleksandra Królikowska  <https://orcid.org/0000-0002-6283-5500>  
 Maciej Dejneka  <https://orcid.org/0000-0002-6675-0256>  
 Robert Prill  <https://orcid.org/0000-0002-4916-1206>  
 Anna Kołcz  <https://orcid.org/0000-0002-6646-3652>

## References

1. Wong WY, Wong MS, Lo KH. Clinical applications of sensors for human posture and movement analysis: A review. *Prosthet Orthot Int*. 2007;31(1):62–75. doi:10.1080/03093640600983949
2. Merriault P, Dupuis Y, Boutheau R, Vasseur P, Savatier X. A study of Vicon System Positioning performance. *Sensors*. 2017;17(7):1591. doi:10.3390/s17071591
3. Wieteci P, Pawik Ł, Fink-Lwów F, et al. Kinematic parameters following Pilon fracture treatment with the Ilizarov method. *J Clin Med*. 2022;11(10):2763. doi:10.3390/jcm11102763
4. Prill R, Walter M, Królikowska A, Becker R. A systematic review of diagnostic accuracy and clinical applications of wearable movement sensors for knee joint rehabilitation. *Sensors*. 2021;21(24):8221. doi:10.3390/s21248221
5. Lopez-Nava IH, Munoz-Melendez A. Wearable inertial sensors for human motion analysis: A review. *IEEE Sensors J*. 2016;16(22):7821–7834. doi:10.1109/JSEN.2016.2609392
6. Metcalf CD, Robinson R, Malpass AJ, et al. Markerless motion capture and measurement of hand kinematics: Validation and application to home-based upper limb rehabilitation. *IEEE Trans Biomed Eng*. 2013;60(8):2184–2192. doi:10.1109/TBME.2013.2250286
7. Mahmoud SS, Cao Z, Fu J, Gu X, Fang Q. Occupational therapy assessment for upper limb rehabilitation: A multisensor-based approach. *Front Digit Health*. 2021;3:784120. doi:10.3389/fgdth.2021.784120

8. Beshara P, Anderson DB, Pelletier M, Walsh WR. The reliability of the Microsoft Kinect and ambulatory sensor-based motion tracking devices to measure shoulder range-of-motion: A systematic review and meta-analysis. *Sensors*. 2021;21(24):8186. doi:10.3390/s21248186
9. Houston A, Walters V, Corbett T, Coppack R. Evaluation of a multi-sensor leap motion setup for biomechanical motion capture of the hand. *J Biomechanics*. 2021;127:110713. doi:10.1016/j.jbiomech.2021.110713
10. Bawa A, Banitsas K, Abbod M. A review on the use of Microsoft Kinect for gait abnormality and postural disorder assessment. *J Healthcare Eng*. 2021;2021:1–19. doi:10.1155/2021/4360122
11. Gatica-Rojas V, Méndez-Rebolledo G, Guzman-Muñoz E, et al. Does Nintendo Wii Balance Board improve standing balance? A randomized controlled trial in children with cerebral palsy. *Eur J Phys Rehabil Med*. 2017;53(4):535–544. doi:10.23736/S1973-9087.16.04447-6
12. da Silva Ribeiro NM, Ferraz DD, Pedreira É, et al. Virtual rehabilitation via Nintendo Wii® and conventional physical therapy effectively treat post-stroke hemiparetic patients. *Top Stroke Rehabil*. 2015;22(4):299–305. doi:10.1179/1074935714Z.0000000017
13. Karasu A, Batur E, Karataş G. Effectiveness of Wii-based rehabilitation in stroke: A randomized controlled study. *J Rehabil Med*. 2018;50(5):406–412. doi:10.2340/16501977-2331
14. Voon K, Silberstein I, Eranki A, Phillips M, Wood FM, Edgar DW. Xbox Kinect™ based rehabilitation as a feasible adjunct for minor upper limb burns rehabilitation: A pilot RCT. *Burns*. 2016;42(8):1797–1804. doi:10.1016/j.burns.2016.06.007
15. Antico M, Balletti N, Laudato G, et al. Postural control assessment via Microsoft Azure Kinect DK: An evaluation study. *Comput Methods Programs Biomed*. 2021;209:106324. doi:10.1016/j.cmpb.2021.106324
16. Yeung LF, Yang Z, Cheng KCC, Du D, Tong RKY. Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2. *Gait Posture*. 2021;87:19–26. doi:10.1016/j.gaitpost.2021.04.005
17. Albert JA, Owolabi V, Gebel A, Brahms CM, Granacher U, Arnrich B. Evaluation of the pose tracking performance of the Azure Kinect and Kinect v2 for gait analysis in comparison with a gold standard: A pilot study. *Sensors*. 2020;20(18):5104. doi:10.3390/s20185104
18. Lee C, Kim J, Cho S, Kim J, Yoo J, Kwon S. Development of real-time hand gesture recognition for tabletop holographic display interaction using Azure Kinect. *Sensors*. 2020;20(16):4566. doi:10.3390/s20164566
19. Bruton A, Conway JH, Holgate ST. Reliability: What is it, and how is it measured? *Physiotherapy*. 2000;86(2):94–99. doi:10.1016/S0031-9406(05)61211-4
20. Hopkins WG. Measures of reliability in sports medicine and science. *Sports Med*. 2000;30(1):1–15. doi:10.2165/00007256-200030010-00001
21. Bujang M, Baharum N. A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: A review. *Arch Orofac Sci*. 2017;12(1):1–11. [http://aos.usm.my/docs/Vol\\_12/aos-article-0246.pdf](http://aos.usm.my/docs/Vol_12/aos-article-0246.pdf). Accessed May 5, 2022.
22. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychol Methods*. 1996;1(1):30–46. doi:10.1037/1082-989X.1.1.30
23. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropract Med*. 2016;15(2):155–163. doi:10.1016/j.jcm.2016.02.012
24. Cicchetti DV, Sparrow SA. Developing criteria for establishing inter-rater reliability of specific items: Applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981;86(2):127–137. PMID:7315877.
25. Czamara A, Markowska I, Królikowska A, Szopa A, Domagalska-Szopa M. Kinematics of rotation in joints of the lower limbs and pelvis during gait: Early results of SB ACLR approach versus DB ACLR approach. *BioMed Res Int*. 2015;2015:707168. doi:10.1155/2015/707168
26. Jänig W, Baron R. Complex regional pain syndrome: Mystery explained? *Lancet Neurol*. 2003;2(11):687–697. doi:10.1016/S1474-4422(03)00557-X
27. Gracia-Ibáñez V, Vergara M, Buffi JH, Murray WM, Sancho-Bru JL. Across-subject calibration of an instrumented glove to measure hand movement for clinical purposes. *Comput Methods Programs Biomed*. 2017;20(6):587–597. doi:10.1080/10255842.2016.1265950
28. Yang SH, Koh CL, Hsu CH, et al. An instrumented glove-controlled portable hand-exoskeleton for bilateral hand rehabilitation. *Biosensors*. 2021;11(12):495. doi:10.3390/bios11120495
29. Han S, Liu B, Cabezas R, et al. MEgATrack: Monochrome egocentric articulated hand-tracking for virtual reality. *ACM Trans Graph*. 2020;39(4):1–13. doi:10.1145/3386569.3392452
30. Guna J, Jakus G, Pogačnik M, Tomažič S, Sodnik J. An analysis of the precision and reliability of the leap motion sensor and its suitability for static and dynamic tracking. *Sensors*. 2014;14(2):3702–3720. doi:10.3390/s140203702
31. Sarbolandi H, Lefloch D, Kolb A. Kinect range sensing: Structured-light versus Time-of-Flight Kinect. *Computer Vision and Image Understanding*. 2015;139:1–20. doi:10.1016/j.cviu.2015.05.006
32. Asaeda M, Kuwahara W, Fujita N, Yamasaki T, Adachi N. Validity of motion analysis using the Kinect system to evaluate single leg stance in patients with hip disorders. *Gait Posture*. 2018;62:458–462. doi:10.1016/j.gaitpost.2018.04.010
33. Reissner L, Politikou O, Fischer G, Calcagni M. In-vivo three-dimensional motion analysis of the wrist during dart-throwing motion after midcarpal fusion and radioscapholunate fusion. *J Hand Surg Eur Vol*. 2020;45(5):501–507. doi:10.1177/1753193420901462
34. Brigstocke GHO, Hearnden A, Holt C, Whatling G. In-vivo confirmation of the use of the dart thrower's motion during activities of daily living. *J Hand Surg Eur Vol*. 2014;39(4):373–378. doi:10.1177/1753193412460149
35. Werner FW, Green JK, Short WH, Masaoka S. Scaphoid and lunate motion during a wrist dart throw motion. *J Hand Surg*. 2004;29(3):418–422. doi:10.1016/j.jhsa.2004.01.018
36. Moritomo H, Apergis EP, Garcia-Elias M, Werner FW, Wolfe SW. International Federation of Societies for Surgery of the Hand 2013 Committee's Report on Wrist Dart-Throwing Motion. *J Hand Surg*. 2014;39(7):1433–1439. doi:10.1016/j.jhsa.2014.02.035
37. Crisco JJ, Coburn JC, Moore DC, Akelman E, Weiss APC, Wolfe SW. In vivo radiocarpal kinematics and the dart thrower's motion. *J Bone Joint Surg*. 2005;87(12):2729–2740. doi:10.2106/JBJS.D.03058
38. Braidotti F, Atzei A, Fairplay T. Dart-Splint: An innovative orthosis that can be integrated into a scapho-lunate and palmar midcarpal instability re-education protocol. *J Hand Ther*. 2015;28(3):329–335. doi:10.1016/j.jht.2015.01.007
39. Fischer G, Wirth MA, Balocco S, Calcagni M. In vivo measurement of wrist movements during the dart-throwing motion using inertial measurement units. *Sensors*. 2021;21(16):5623. doi:10.3390/s21165623
40. Mitsukane M, Tanabe H, Sugama K, Suzuki Y, Tsurumi T. Test–retest reliability of goniometric measurements of the range of dart-throwing motion. *J Phys Ther Sci*. 2019;31(3):236–241. doi:10.1589/jpts.31.236



# Impact of assisted exercises on skeletal muscle oxygenation levels in men with acutely decompensated heart failure

Olga Kisiel<sup>1,B,D</sup>, Agnieszka Ewa Siennicka<sup>2,C,D</sup>, Krystian Josiak<sup>1,3,C,E</sup>, Robert Zymlński<sup>3,B,E</sup>, Waldemar Banasiak<sup>1,E,F</sup>, Kinga Węgrzynowska-Teodorczyk<sup>1,4,A–D,F</sup>

<sup>1</sup> Centre for Heart Diseases, 4<sup>th</sup> Military Hospital, Wrocław, Poland

<sup>2</sup> Department of Physiology and Patophysiology, Wrocław Medical University, Poland

<sup>3</sup> Institute of Heart Diseases, Wrocław Medical University, Poland

<sup>4</sup> Faculty of Physiotherapy, University of Health and Sport Sciences in Wrocław, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;

D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

Adv Clin Exp Med. 2023;32(2):211–218

## Address for correspondence

Agnieszka Ewa Siennicka

E-mail: agnieszka.siennicka@umw.edu.pl

## Funding sources

The research was financially supported by the National Science Centre (grant No. Sonata 6 UMO-2013/11/D/NZ7/00922).

## Conflict of interest

None declared

Received on February 22, 2022

Reviewed on August 4, 2022

Accepted on August 23, 2022

Published online on November 14, 2022

## Cite as

Kisiel O, Siennicka AE, Josiak K, Zymlński R, Banasiak W, Węgrzynowska-Teodorczyk K. Impact of assisted exercises on skeletal muscle oxygenation levels in men with acutely decompensated heart failure. *Adv Clin Exp Med.* 2023;32(2):211–218. doi:10.17219/acem/152930

## DOI

10.17219/acem/152930

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** The complex clinical status of modern day patients hospitalized due to acute heart failure (AHF) results from their advanced age, comorbidities, frailty, heart failure symptoms (including massive swelling of the lower limb), and dramatic reduction of exercise tolerance. Hence, there is a need to implement physiotherapeutic procedures as early as possible, aiming to both accelerate the restoration of clinical stabilization and prevent post-hospital disability.

**Objectives.** We investigated whether assisted lower limb exercises have an impact on perfusion and oxygenation in skeletal muscle and if they are feasible in patients with AHF.

**Materials and methods.** We examined 34 men (age:  $66 \pm 11$  years; left ventricular ejection fraction (LVEF):  $34 \pm 11\%$ ; clinical presentation: 31 wet-warm and 3 wet-cold). The intervention (carried out on the 2<sup>nd</sup> day of hospitalization) included: 1) a 3-minute rest period; 2) an exercise phase (45 repetitions of assisted flexion and extension of the lower limb; and 3) a 10-minute relaxation period. We analyzed blood pressure (BP), heart rate (HR), respiratory rate (RR), tissue oxygenation (reflected by oxygen saturation measured with a pulse oximeter), and changes in peripheral tissue perfusion (reflected by the tissue oxygenation index (TOI) measured with near-infrared spectroscopy (NIRS)).

**Results.** The hemodynamic parameters (both  $\Delta$ HR and  $\Delta$ systolic BP) and oxygen saturation did not change (all  $p > 0.05$ ), whereas the RR declined ( $p < 0.001$ ). The exercises improved venous outflow (reflected by decreased oxygenated, deoxygenated and total hemoglobin, all  $p < 0.05$ ) and increased peripheral tissue perfusion, as reflected by the TOI ( $p < 0.05$ ). The patients reported relief and lack of dyspnea during and after the assisted exercises.

**Conclusions.** The physiotherapeutic intervention improved both venous outflow and muscle oxygenation in men with AHF. The presented protocol was safe, feasible and well-tolerated, and resulted in relief for the patients. We believe that such procedures might be recommended for the initial period of rehabilitation in this challenging subgroup of patients.

**Key words:** physiotherapy, heart failure, near-infrared spectroscopy, skeletal muscle oxygenation

## Background

Morbidity and mortality due to heart failure increase every year, affecting more than 10% of people over the age of 70.<sup>1</sup> A fundamental aspect of heart failure is periods of exacerbation, known as cardiac decompensation or acute heart failure (AHF).<sup>2</sup> Efficient therapeutic options for patients with AHF have been evolving over the last 2 decades and have resulted in an extension of life expectancy and significant changes to their clinical profile compared to 20 years ago. Therefore, patients hospitalized for AHF are now older, have more comorbidities, such as frailty syndrome, and a worse prognosis.<sup>3–5</sup>

It is noteworthy that hospitalization of patients due to AHF may have additional negative consequences, such as the so-called “post-hospital syndrome”, which is associated with an increased risk of disability.<sup>6,7</sup> Studies have demonstrated that up to 60% of AHF patients develop hospital-acquired disability, which is associated with a worse clinical outcome.<sup>8</sup>

Hospitalization due to AHF is associated with extended time spent lying in bed, especially for patients who are unable to walk independently. As a result, some experience a new major physical disability following hospital discharge.<sup>9,10</sup>

The data suggest that exercise and early rehabilitation protocols applied during acute hospitalization may prevent functional and cognitive decline in older patients. Additionally, these treatments are associated with shorter hospital stay and lower costs.<sup>11–13</sup> It is very likely that physiotherapeutic interventions implemented as early as possible during hospitalization due to AHF may mitigate these unfavorable outcomes.<sup>10,14–16</sup> However, cardiac physiotherapy implemented early in AHF has not been well established. It is generally accepted that such patients require special physiotherapeutic care.<sup>17,18</sup>

In AHF, the patient’s clinical picture can vary greatly. One major symptom is massive swelling of the lower limbs, which limits the patient’s mobility or even makes it impossible to move, forcing the patient to remain in a lying or sitting position. Additionally, AHF is associated with an increased level of fatigue and dyspnea associated with fear of exacerbation of symptoms during physical exercise.<sup>18</sup>

Having these considerations in mind, we selected assisted exercises for being investigated in this study. Importantly, there is currently no available literature describing the feasibility and effects of assisted exercises in patients in the acute stages of heart failure.

## Objectives

The aim of this study was to evaluate the feasibility of assisted lower limb exercises and to examine their impact on the levels of perfusion and oxygenation in the skeletal muscles of patients with AHF. We decided to analyze this

at 3 levels: level related to the patient’s subjective perception, hemodynamic level (heart rate (HR) and blood pressure (BP) response) and tissue level (i.e., tissue oxygenation).

## Materials and methods

### Study design and participants

The study included patients with AHF. Acute heart failure was defined according to the current 2021 European Society of Cardiology (ESC) Guidelines<sup>2</sup> as a rapid or gradual onset of symptoms and/or signs of heart failure requiring unplanned hospitalization of the patient. The use of an intravenous loop diuretic and/or vasodilator was not necessary to participate in this study.

This study was registered at Clinicaltrials.gov (ID No. NCT03102164). All study participants provided written informed consent to participate in the study.

Patients were excluded if they met any of the following exclusion criteria: acute coronary syndrome, bacterial infection, pre-existing chronic respiratory failure, need for invasive mechanical ventilation, significant arrhythmia (arrhythmia that was hemodynamically unstable and/or needed urgent targeted treatment) or conductivity disorders, anemia (hemoglobin <9 g%), active neoplastic process, liver injury (aspartate aminotransferase (AST) or alanine aminotransferase (ALT) more than 3 times the reference level), or chronic kidney failure with creatinine clearance <30 mL/min. Patients with acute deep vein thrombosis or significant orthopedic diseases of the lower limb were also excluded. Finally, patients with massive swelling and/or shortness of breath making it impossible to lie down were excluded.

### Procedure and exercise settings

Patients were recruited into the study by doctors who exercised medical supervision of the rehabilitation process. A total of 34 men with AHF participated in the study. The clinical characteristics of the study group are presented in Table 1.

The tests were carried out in the patient’s room at the cardiology department on their 2<sup>nd</sup> day of hospitalization. No noninvasive ventilation was used during the study.

The study protocol was divided into 3 stages:

1) Recording of the resting parameters (3 min), during which the accuracy of measurement and recordings were simultaneously monitored.

2) Assisted exercise phase. Assisted exercise refers to exercise in which a physiotherapist alone facilitates the movement by relieving the exercised limb. The movement is performed with a minimal use of the patient’s muscle strength.<sup>19</sup> During this phase, the physiotherapist performed 45 repetitions of assisted exercise, involving flexion of the studied lower limb at the knee and hip joints. For the hip joint, the angle of flexion was in the 45–90°

**Table 1.** Clinical characteristics of the study group and the pharmacological treatment used at the time of the inclusion in the study

Clinical variable	Men with AHF (n = 34)
Age [years]	66 ± 11
BMI [kg/m <sup>2</sup> ]	29.7 ± 6.4
AHF etiology, ischemic, n (%)	18 (53)
HFrEF, n (%)	23 (68)
HFmEF, n (%)	6 (18)
HFpEF, n (%)	5 (14)
LVEF [%]	33.8 ± 11
LVEDD [%]	64 ± 11.4
NYHA IV, n (%)	34 (100)
SBP [mm Hg]	128 ± 21
SpO <sub>2</sub> [%]	95 ± 2
Hemoglobin [g/dL]	13.2 ± 2
Creatinine [mg/dL]	1.5 ± 0.8
NTproBNP [pg/mL]	6522 ± 4750
PaO <sub>2</sub> [mm Hg]	67 ± 11
MR (n, %)	
Mild	23 (68)
Moderate/severe	11 (32)
Hypertension, n (%)	20 (59)
Diabetes mellitus, n (%)	17 (50)
Chronic obstructive pulmonary disease, n (%)	6 (18)
ICD, n (%)	9 (27)
CRT-D, n (%)	5 (15)
CABG, n (%)	4 (12)
AF, n (%)	21 (62)
QRS > 120, n (%)	13 (38)
LBBB, n (%)	5 (15)
Clinical presentation (n, %)	
Wet and warm	31 (91)
Wet and cold	3 (9)
Peripheral congestion	29 (85)
Pulmonary congestion	34 (100)
Pharmacological treatment (n, %)	
ACE-I and/or ARB	29 (85)
β-blocker	34 (100)
Digoxin	3 (9)
Furosemide	34 (100)
Vasodilator	23 (68)
Dobutamine	1 (3)
Noradrenaline	0 (0)

BMI – body mass index; AHF – acute heart failure; HFrEF – heart failure with reduced ejection fraction; HFmEF – heart failure with mid-range ejection fraction; HFpEF – heart failure with preserved ejection fraction; LVEF – left ventricular ejection fraction; LVEDD – left ventricular end-diastolic diameter; NYHA – New York Heart Association; AS – aortic stenosis; SBP – systolic blood pressure; SpO<sub>2</sub> – blood oxygen saturation; NTproBNP – N-terminal pro B-type natriuretic peptide; PaO<sub>2</sub> – partial pressure of oxygen in arterial blood; MR – mitral regurgitation; CRT-D – cardiac resynchronization therapy defibrillator; CABG – coronary artery bypass graft; AF – atrial fibrillation; LBBB – left bundle branch block; ACE-I – angiotensin-converting enzyme inhibitor; ARB – angiotensin receptor blockers; ICD – implantable cardioverter-defibrillator; QRS – QRS complex (Q wave R wave S wave).

range, and the knee joint was expected to be at an obtuse angle of 90–100° (heel raised above the knee). The exercise duration was approx. 1.5 min.

3) Relaxation phase. Relaxation phase comprised of a 10-minute measurement. Care was taken to ensure the patient lay still during this time and does not make any movements.

## Variables and data sources

The impact of the exercises on the following factors was assessed:

1) Subjective feelings of the patient. After the exercise, the patient was asked to provide information (feedback) as to whether the exercise: affected their discomfort levels, tissue tension and lower limb pain (unchanged ↔/increased ↑/decreased ↓); provoked or increased their feeling of breathlessness (yes/no).

2) Hemodynamic parameters. Each examination was conducted with supervision and monitoring of the following hemodynamic parameters: HR, BP and respiratory rate (RR).

3) Tissue oxygenation. In addition to standard saturation assessment measured with a pulse oximeter, near-infrared spectroscopy (NIRS) was used to assess tissue oxygenation. This was important considering the reduced peripheral perfusion and frequent swelling of the lower limbs experienced by patients with AHF. Near-infrared spectroscopy is a tool used for continuous noninvasive monitoring of changes in peripheral tissue perfusion.<sup>20</sup> Near-infrared spectroscopy is based on the absorption of near-infrared photons by chromophores (e.g., hemoglobin). Differences in the intensity of NIRS absorption by oxygenated and reduced hemoglobin were analyzed using algorithms based on the Beer–Lambert law. The NIRS light emitted by the sensor's diode penetrated locally into the tissue underneath, and the hemoglobin level in the microcirculation was measured with 2 detectors.<sup>21–23</sup> The photosensor was placed on the subject's quadriceps muscle, 1/3 of the way to the proximal lateral head. The distance between the sensors in the photosensor was 4 cm; thus, the sensor's penetration depth was up to 4 cm. Commonly, NIRS is used to check the dynamic balance between the delivery and extraction of oxygen in the tissue during surgery or exercise.<sup>20</sup> We used NIRS to assess leg muscle perfusion and to indirectly determine venous outflow during the assisted exercises. The following parameters were measured and subsequently analyzed: tissue oxygenation index (TOI, i.e., the percentage ratio of oxygenated hemoglobin to total hemoglobin) and changes in oxygenated hemoglobin ( $\Delta\text{O}_2\text{Hb}$ ), deoxygenated hemoglobin ( $\Delta\text{HHb}$ ) and total hemoglobin ( $\Delta\text{total Hb}$ ).

The research was carried out as part of the Sonata 6 UMO-2013/11/D/NZ7/00922 research grant awarded by the National Science Centre, Kraków, Poland. It was conducted in accordance with the Declaration of Helsinki and approved by the Bioethics Committee of Wrocław Medical University, Poland (approval No. 204/2015).

## Statistical analyses

Continuous variables with normal distribution were described using mean  $\pm$  standard deviation ( $M \pm SD$ ). Variables with skewed distribution were described as medians with upper and lower quartiles. Categorical variables were given as numbers and percentages. In the case of a nonparametric test for 2 trials of dependent variables, a Wilcoxon signed-rank test was used to analyze changes in the hemodynamic parameters before and after the exercise. The differences between particular phases of exercise were analyzed using the Friedman's analysis of variance (ANOVA) with determination of the Kendall's compliance factor. A value of  $p < 0.05$  was considered statistically significant. Statistical analyses were performed using Statistica v. 12 (StatSoft, Tulsa, USA).

## Results

The assisted exercises were well tolerated by the patients. After the series of 45 repetitions of the assisted exercise, BP and HR remained unchanged, while a decrease in RR was observed (Table 2). No adverse symptoms were reported during the exercises or directly after (Table 3).

The exercises did not change the saturation levels ( $p = 0.11$ ); their value remained at 95% (93–96%). The percentage changes in selected oxygenation parameters of the quadriceps muscle during the various phases of testing are presented in the Figures. Figures 1–4 present the values of changes in the  $\Delta O_2Hb$ ,  $\Delta HHb$ ,  $\Delta total Hb$ , and TOI at the following test phases: at rest, during the assisted

exercise, and after the 1<sup>st</sup> and 10<sup>th</sup> min of relaxation following exercise completion.

A decrease in the level of oxygenated hemoglobin was achieved as a result of the assisted exercises. In the 1<sup>st</sup> minute of relaxation, an upward trend was observed, only to return to the resting value in the final phase (Fig. 1). A significant decrease in the levels of deoxygenated hemoglobin and total hemoglobin was achieved as a result of the assisted exercises. In the 1<sup>st</sup> minute of relaxation, an upward trend was observed, only to return to the resting value in the final phase (Fig. 2,3). After performing the assisted exercises, a slight increase in oxygenation in the quadriceps muscle of the thigh was observed in the later phase of rest ( $p < 0.03$ ) (Fig. 4).

## Discussion

The increased frequency of peripheral swelling and resting dyspnea during a decompensated heart failure episode significantly limits the individual's ability to independently perform basic activities. Consequently, the management of AHF by a physiotherapist remains an ongoing challenge.

This research demonstrated that the assisted exercises studied were safe and accepted by AHF patients. Also, no adverse symptoms were observed. The majority of patients reported relief due to reduced discomfort and lower tissue tension in the lower limb. We also observed a decrease in the oxygenated hemoglobin, deoxygenated hemoglobin and total hemoglobin levels, accompanied by a slight increase in oxygenation in the quadriceps muscle of the thigh in the later phase of rest.

**Table 2.** Hemodynamic parameters before and after assisted exercise

Parameters	Resting median (upper-lower quartile)	After exercise median (upper-lower quartile)	p-value
SBP [mm Hg]	117 (101–136)	115 (100–133)	0.068
DBP [mm Hg]	71 (65–80)	71 (63–80)	0.088
HR [beats/min]	88 (65–99)	86 (66–94)	0.82
RR [breaths/min]	22 (20–26)	20 (20–24)	0.0002

SBP – systolic blood pressure; DBP – diastolic blood pressure; HR – heart rate; RR – respiratory rate.

**Table 3.** Subjective feelings of increased breathlessness or discomfort in the lower limb following exercise

Subjective feelings of symptoms	Answer	Patients (n (%))
Increased feelings of breathlessness	yes	5 (15) There was a slight difficulty at the beginning (breath holding or shallow breathing), which passed after a few repetitions and stabilized in the correct respiratory rhythm, coordinated with the movement of the lower limb
	no	29 (85)
Discomfort, tissue tension or pain in the lower limb	unchanged	10 (29) (6 of these patients had peripheral congestion)
	increased	0
	decreased	24 (71) (23 of these patients had peripheral congestion)



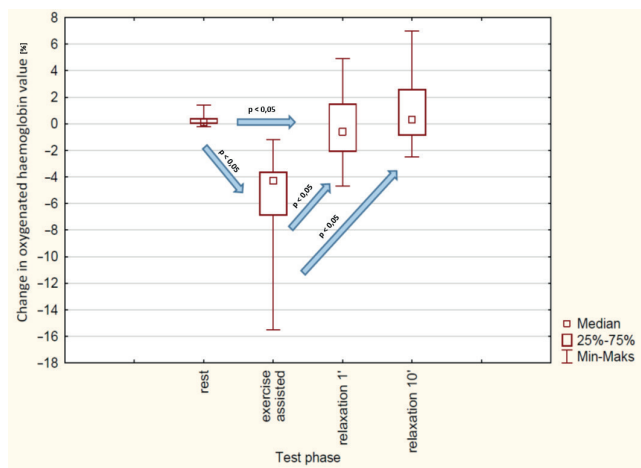


Fig. 1. Changes in oxygenated hemoglobin levels at different phases of the test

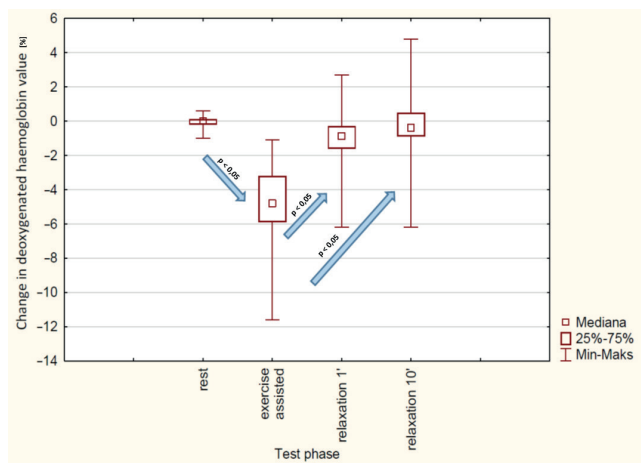


Fig. 2. Changes in deoxygenated hemoglobin levels at different phases of the test

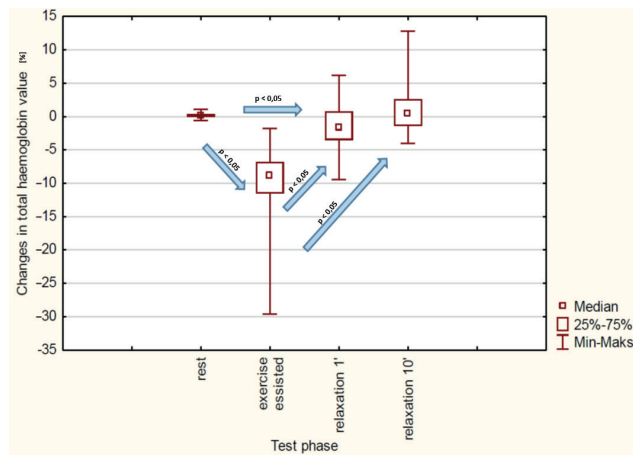


Fig. 3. Changes in total hemoglobin levels at different phases of the test

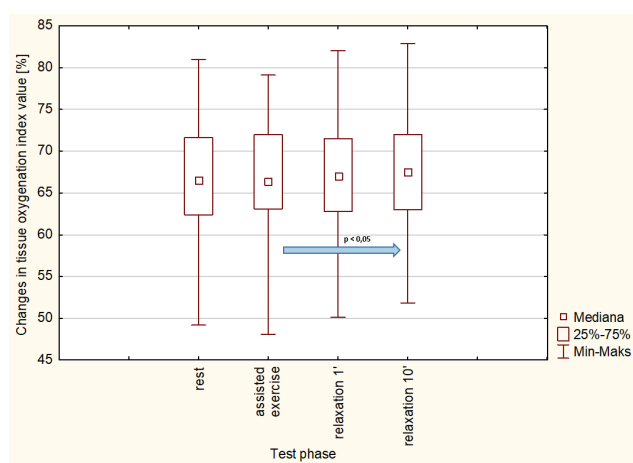


Fig. 4. Changes in tissue oxygenation index levels at different phases of the test

Most patients were of the warm-wet profile (i.e., with peripheral stagnation and lower limb swelling), which is a group of patients to which these exercises are dedicated.

Additionally, we showed that the assisted exercises applied in this study are safe because they do not adversely affect the hemodynamic parameters (no changes in HR and RR). Unfavorable tendencies of decreased systolic BP or HR changes were not observed; these have an adverse impact on the prognosis in AHF.<sup>24,25</sup> In evaluating the supply and use of oxygen by the skeletal muscles, we used NIRS, whose suitability has previously been demonstrated.<sup>26–28</sup> Near-infrared spectroscopy has commonly been used to evaluate the dynamic balance between the delivery and extraction of oxygen in the tissue during surgery or exercise.<sup>20,29</sup>

In a study by Belardinelli et al., changes in skeletal muscle oxygen saturation were evaluated during incremental exercise, both in patients with chronic heart failure (CHF) and in healthy subjects. It was observed that patients with CHF had reduced muscle oxygenation during exercise at any level of intensity compared to the control group.

Additionally, patients with CHF had earlier acceleration in muscle deoxygenation, indicating a premature onset of anaerobic metabolism.<sup>30,31</sup> Although the abovementioned publications evaluated lower limb skeletal muscle oxygenation in patients with heart failure, several disparities prevent us from comparing them with our results. The study by Belardinelli et al. involved patients with stable heart failure (New York Heart Association (NYHA) class II), whereas our project included subjects with AHF (NYHA class IV). Moreover, the studies differ in the type of exercise analyzed: dynamic exercise with a constant or increasing workload of several minutes on an ergometer cannot be compared with assisted exercises performed in clinically unstable patients. Assisted exercises last approx. 2 min and are performed by a physiotherapist, with minimal work performed by the patient.<sup>30</sup> We were unable to identify any studies in the available literature that had evaluated the impact of assisted exercises on the level of skeletal muscle oxygenation in patients with AHF. Thus, our results cannot be directly compared with the results of other studies.

The purpose of using NIRS in our study was not to assess oxygen extraction and metabolism, but blood flow and fluid distribution in the venous circulation, as well as to evaluate perfusion changes in the lower limb muscles under the influence of elevation and assisted movement. The NIRS tissue oxygen saturation measurement is made on a mixture of both venous and arterial blood with a mean ratio of 70% venous to 30% arterial. It has been empirically determined that the same mean ratio of 70% venous to 30% arterial blood volume applies to the whole body.<sup>32</sup> This suggests that increased venous outflow from the peripheral tissues leads to a reduction in peripheral stasis, which may contribute to relief of the right ventricle. We can also assume that an increased venous outflow results in a better outflow of metabolites (such as H<sup>+</sup>, lactate, and prostaglandins PGE<sub>2</sub> and PGF<sub>1</sub>), which may in turn positively affect a response from the ergoreceptors.<sup>33</sup> In the pathogenesis of heart failure symptoms, an increased response from the ergoreceptors is of major importance. It causes increased stimulation of the sympathetic nervous system, increased peripheral resistance, excessive ventilation, and increased dyspnea.<sup>34</sup>

In our study, we observed increased venous outflow as a result of the assisted exercises with elevation of the lower limbs taken into account; however, the exercises did not (as one might expected) increase dyspnea in patients. On the contrary, the RR decreased. Based on these results, we can assume that the increased outflow of venous blood may contribute to an improvement in outflow of metabolites, reduction of the ergoreflex, and, consequently, reduction of sympathetic activation and dyspnea.

Patients with CHF have a decreased functional efficiency due to age, comorbidities and skeletal muscle dysfunction.<sup>35–37</sup> Additionally, frailty syndrome, a consequence of decompensation, hospitalization and bed rest, leads to a sharp worsening of functional deficits, disease progression and a poorer prognosis.<sup>4,5</sup> In order to increase their chances of returning to baseline functional physical fitness or to minimize the deficits resulting from AHF, it is necessary to start the rehabilitation as soon as possible.<sup>15,38</sup>

The assumption of our research was to commence physical exercises, making it the first step as an introduction to the rehabilitation program of patients who are weak and bed-bound, helping them regain independence and mobility. Providing a uniform rehabilitation program and simple recommendations during an AHF episode is not possible. This is due to the varied clinical picture and diverse decompensation etiologies of the underlying disease or comorbidities.<sup>5,39,40</sup>

Patients with AHF are often placed in intensive cardiac supervision wards where their rehabilitation commences; this resembles early mobilization of patients staying in the intensive care unit (ICU). It involves a gradual increase in functional activities, beginning with passive and active assisted exercises while the patient is lying

in bed, through active exercises in a sitting and standing position, progressing to walking in one spot, and then gait walking.<sup>16,41–43</sup> Taking into account the absolute contraindications to exercise, the physiotherapist should develop an individualized, monitored and supervised rehabilitation program.<sup>17,40,44</sup>

When choosing exercises for our research, the following principle was adopted: the more severe the heart failure, the lower the intensity and duration of exercise, and the higher the exercise frequency. Assisted exercises have been included in the Polish “HF model” of cardiac rehabilitation of patients with AHF.<sup>40</sup>

In order to prevent progressive impairment of numerous body functions, Massimo Piepoli, an expert on heart failure and then the Chairman of the European Guidelines on cardiovascular disease prevention in clinical practice, has been recommending early mobilization of patients for years. Following a period of immediate threat and clinical instability, he encourages simple and slow movements performed as soon as possible.<sup>17</sup> Piepoli emphasizes the important role of early activation of patients after stable clinical AHF status has been achieved. Exercises should be individualized based on simple movements, engaging body weight only.<sup>17,37</sup>

Researchers have demonstrated the safety and efficacy of multidomain rehabilitation, which improves physical function and shortens rehospitalization. In their studies, they included independent and mobile patients who could walk at least 4 m at enrollment, were functionally independent before admission and were expected to be discharged from hospital.<sup>13,14</sup> In the REHAB-HF project and in a multi-center study by Kitzman et al., the 30-minute rehabilitation sessions commenced in the hospital during AHF and continued at home and in the outpatient clinic for the next 12 weeks.<sup>14,45,46</sup>

Our research stems from practice and from the daily needs of those working with elderly patients with disabilities who are bed-bound in hospital with severe disease symptoms. It was challenging to decide which exercises to start off with, and how to perform them or to estimate which groups of patients will benefit from them the most. In this project, we wanted to determine whether early activation is safe, if it has a positive impact on the patient's well-being, and, above all, whether it has an objective, measurable impact on the oxidation of skeletal muscles during a decompensated heart failure stage.

During AHF, patients often experience massive swelling associated with an increased tension of the skin integument of the entire lower limb, occasionally expanding the skin pores through which serous fluid may seep out. These symptoms are often accompanied by pain and discomfort, which hinder or prevent independent, active movements of the lower limb. With these ailments and our own experiences in mind, assisted exercises were selected. These types of exercises do not require effort from the patient or significant involvement of their own

muscles. Furthermore, during the early period of the disease, assisted exercises intend to gradually introduce other physiotherapeutic activities to the exercise therapy program with greater patient involvement. The additional benefit of assisted exercises is the reduction of risks involving thromboembolism and the reduction or prevention of edema.<sup>47,48</sup> Peripheral edema of the lower extremities occurs in more than 50% of patients with stable heart failure and in about 75% of patients with decompensated heart failure.<sup>49,50</sup> The role of assisted exercises was to enhance flow in the venous vascular bed by raising the heel and assisting the outflow of accumulated fluid, which can reduce stasis of the peripheral circulation.

When assessing oxygenation of the quadriceps muscle as a reaction to the exercises and in the rest phase 10 min after the end of the exercise, we observed a slight increase in the TOI. This is certainly a beneficial effect that requires further in-depth research. We presume that the improvement in venous outflow contributed to an increase in the arteriovenous difference and increased blood flow and oxygenation throughout the muscles.

The assisted lower limb movements also helped reduce the pain connected with swelling, making it more pleasant for the patient. This result may have been attained because the fluid pressure exerted on the peripheral tissue was reduced. This was reflected in the feeling of reduced discomfort due to the relieved skin tension reported by the patients. This may have the added benefit of reducing reluctance to exercise. The beneficial effects of early activation of the patient as a way of preventing intensification of the symptoms of FS, post-hospital syndrome or rehospitalization should not be underestimated.

## Limitations


Our research is not free from limitations, and this needs to be acknowledged. First, the study included a relatively small number of patients. This is due to the unique nature of our study population, which is comprised of elderly patients with many different chronic conditions. Due to the small study group, we decided not to include both genders to avoid heterogeneity within our sample, which could introduce another limitation. However, we felt that gender diversity was particularly important in the context of our study in terms of the effects of hormones on muscle tissue and, moreover, significant differences in muscle mass, strength and endurance. Another limitation is the lack of a control group. It is due to the fact that, at this stage of the research, we focused on the effects of exercise on hemodynamic parameters and muscle oxygenation, while the observed changes and their dynamics were compared to resting values. The methods used for diagnosis and assessment of pulmonary congestion have had limited accuracy and reliability. We endeavor to eliminate these limitations by continuing this research in the future.

## Conclusions

We have demonstrated that assisted exercises of the lower limb are safe because they do not burden the circulatory system. This is because they did not cause changes in BP, HR and oxygen saturation or increase dyspnea; therefore, they can be performed during the initial rehabilitation period of male patients hospitalized for AHF. Furthermore, the presented data suggest that such physiotherapeutic interventions are beneficial, particularly in patients with peripheral congestion and lower limb edema, as they improve venous outflow and muscle oxygenation. Finally, assisted exercises helped reduce the severe and uncomfortable symptoms of AHF, such as feelings of tension and pain in the lower extremities, that limit the patient's mobility.

## ORCID iDs

Olga Kisiel  <https://orcid.org/0000-0002-5398-5654>


Agnieszka Ewa Siennicka  <https://orcid.org/0000-0003-0988-5821>

Krzysztof Josiak  <https://orcid.org/0000-0003-2099-7453>

Robert Zymliński  <https://orcid.org/0000-0003-1483-7381>

Waldemar Banasiak  <https://orcid.org/0000-0003-0702-7508>

Kinga Węgrzynowska-Teodorczyk

 <https://orcid.org/0000-0001-9189-748X>

## References

- Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol*. 2016;13(6):368–378. doi:10.1038/nrcardio.2016.25
- McDonagh TA, Metra M, Adamo M, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J*. 2021;42(36):3599–3726. doi:10.1093/eurheartj/ehab368
- Takara Y, Saitoh M, Morisawa T, et al. Clinical characteristics of older heart failure patients with hospital-acquired disability: A preliminary, single-center, observational study. *Cardiol Res*. 2021;12(5):293–301. doi:10.14740/cr1306
- Reeves GR, Whellan DJ, Patel MJ, et al. Comparison of frequency of frailty and severely impaired physical function in patients ≥60 years hospitalized with acute decompensated heart failure versus chronic stable heart failure with reduced and preserved left ventricular ejection fraction. *Am J Cardiol*. 2016;117(12):1953–1958. doi:10.1016/j.amjcard.2016.03.046
- Murad K, Kitzman DW. Frailty and multiple comorbidities in the elderly patient with heart failure: Implications for management. *Heart Fail Rev*. 2012;17(4–5):581–588. doi:10.1007/s10741-011-9258-y
- Krumholz HM. Post-hospital syndrome: An acquired, transient condition of generalized risk. *N Engl J Med*. 2013;368(2):100–102. doi:10.1056/NEJMp1212324
- Zisberg A, Shadmi E, Gur-Yaish N, Tonkikh O, Sinoff G. Hospital-associated functional decline: The role of hospitalization processes beyond individual risk factors. *J Am Geriatr Soc*. 2015;63(1):55–62. doi:10.1111/jgs.13193
- Rossello X, Miró Ò, Llorens P, et al. Effect of Barthel index on the risk of thirty-day mortality in patients with acute heart failure attending the emergency department: A cohort study of nine thousand ninety-eight patients from the Epidemiology of Acute Heart Failure in Emergency Departments Registry. *Ann Emerg Med*. 2019;73(6):589–598. doi:10.1016/j.annemergmed.2018.12.009
- Gill TM, Allore HG, Gahbauer EA, Murphy TE. Change in disability after hospitalization or restricted activity in older persons. *JAMA*. 2010;304(17):1919. doi:10.1001/jama.2010.1568
- Martínez-Velilla N, Casas-Herrero A, Zambom-Ferraresi F, et al. Effect of exercise intervention on functional decline in very elderly patients during acute hospitalization: A randomized clinical trial. *JAMA Intern Med*. 2019;179(1):28. doi:10.1001/jamainternmed.2018.4869
- de Morton N, Keating JL, Jeffs K. Exercise for acutely hospitalised older medical patients. *Cochrane Database Syst Rev*. 2007;2007(1):CD005955. doi:10.1002/14651858.CD005955.pub2

12. Yasmeen I, Krewulak KD, Grant C, Stelfox HT, Fiest KM. The effect of caregiver-mediated mobility interventions in hospitalized patients on patient, caregiver, and health system outcomes: A systematic review. *Arch Rehabil Res Clin Transl*. 2020;2(3):100053. doi:10.1016/j.arrct.2020.100053
13. Semsar-Kazerooni K, Dima D, Valiquette J, Berube-Dufour J, Goldfarb M. Early mobilization in people with acute cardiovascular disease. *Can J Cardiol*. 2021;37(2):232–240. doi:10.1016/j.cjca.2020.03.038
14. Kitzman DW, Whellan DJ, Duncan P, et al. Physical rehabilitation for older patients hospitalized for heart failure. *N Engl J Med*. 2021;385(3):203–216. doi:10.1056/NEJMoa2026141
15. Munir H, Morais JA, Goldfarb M. Health-related quality of life in older adults with acute cardiovascular disease undergoing early mobilization. *CJC Open*. 2021;3(7):888–895. doi:10.1016/j.cjco.2021.02.013
16. Kakutani N, Fukushima A, Kinugawa S, et al. Progressive mobilization program for patients with acute heart failure reduces hospital stay and improves clinical outcome. *Circ Rep*. 2019;1(3):123–130. doi:10.1253/circrep.CR-19-0004
17. Piepoli MF, Conraads V, Corrà U, et al. Exercise training in heart failure: From theory to practice. A consensus document of the Heart Failure Association and the European Association for Cardiovascular Prevention and Rehabilitation. *Eur J Heart Fail*. 2011;13(4):347–357. doi:10.1093/eurjhf/hfr017
18. Reeves GR, Whellan DJ, Duncan P, et al. Rehabilitation Therapy in Older Acute Heart Failure Patients (REHAB-HF) trial: Design and rationale. *Am Heart J*. 2017;185:130–139. doi:10.1016/j.ahj.2016.12.012
19. Pollock A, Baer G, Campbell P, et al. Physical rehabilitation approaches for the recovery of function and mobility following stroke. *Cochrane Database Syst Rev*. 2014;2014(4):CD001920. doi:10.1002/14651858.CD001920.pub3
20. Trafidło T, Gaszyński T. NIRS: Spektroskopia bliskiej podczerwieni jako wielofunkcyjna metoda monitorowania miejscowej oksygenacji tkankowej w anestezjologii i ratownictwie. *Anestezjologia i Ratownictwo*. 2009;3:351–359.
21. Barroco AC, Sperandio PA, Reis M, Almeida DR, Neder JA. A practical approach to assess leg muscle oxygenation during ramp-incremental cycle ergometry in heart failure. *Braz J Med Biol Res*. 2017;50(12):e6327. doi:10.1590/1414-431x20176327
22. Jones S, Chiesa ST, Chaturvedi N, Hughes AD. Recent developments in near-infrared spectroscopy (NIRS) for the assessment of local skeletal muscle microvascular function and capacity to utilize oxygen. *Artery Res*. 2016;16:25–33. doi:10.1016/j.artres.2016.09.001
23. De Blasi RA, Palmisani S, Alampi D, et al. Microvascular dysfunction and skeletal muscle oxygenation assessed by phase-modulation near-infrared spectroscopy in patients with septic shock. *Intensive Care Med*. 2005;31(12):1661–1668. doi:10.1007/s00134-005-2822-y
24. Hiki M, Iwata H, Takasu K, et al. Elevated heart rate in combination with elevated blood pressure predicts lower cardiovascular mortality in acute decompensated heart failure. *Int Heart J*. 2020;61(2):308–315. doi:10.1536/ihj.19-521
25. Aalders M, Kok W. Comparison of hemodynamic factors predicting prognosis in heart failure: A systematic review. *J Clin Med*. 2019;8(10):1757. doi:10.3390/jcm8101757
26. Carlier PG, Bertoldi D, Baligand C, Wary C, Fromes Y. Muscle blood flow and oxygenation measured by NMR imaging and spectroscopy. *NMR Biomed*. 2006;19(7):954–967. doi:10.1002/nbm.1081
27. Barroco AC, Sperandio PA, Reis M, Almeida DR, Neder JA. A practical approach to assess leg muscle oxygenation during ramp-incremental cycle ergometry in heart failure. *Braz J Med Biol Res*. 2017;50(12):e6327. doi:10.1590/1414-431x20176327
28. Kravari M, Vasileiadis I, Gerovasili V, et al. Effects of a 3-month rehabilitation program on muscle oxygenation in congestive heart failure patients as assessed by NIRS. *Int J Ind Ergon*. 2010;40(2):212–217. doi:10.1016/j.ergon.2009.03.006
29. La Monaca M, David A, Gaeta R, Lentini S. Near-infrared spectroscopy for cerebral monitoring during cardiovascular surgery [in Italian]. *Clin Ter*. 2010;161(6):549–553. PMID:21181086.
30. Belardinelli R, Georgiou D, Barstow TJ. Near infrared spectroscopy and changes in skeletal muscle oxygenation during incremental exercise in chronic heart failure: A comparison with healthy subjects. *G Ital Cardiol*. 1995;25(6):715–724. PMID:7649420.
31. Belardinelli R. Monitoring skeletal muscle oxygenation during exercise by near infrared spectroscopy in chronic heart failure. *Congest Heart Fail*. 1999;5(3):116–119. PMID:12189315.
32. Pang CCY. Measurement of body venous tone. *J Pharmacol Toxicol Methods*. 2000;44(2):341–360. doi:10.1016/S1056-8719(00)00124-6
33. Scott AC, Wensel R, Davos CH, et al. Chemical mediators of the muscle ergoreflex in chronic heart failure: A putative role for prostaglandins in reflex ventilatory control. *Circulation*. 2002;106(2):214–220. doi:10.1161/01.CIR.0000021603.36744.5E
34. Kaczmarek A, Banasiak W, Ponikowski P. The role of skeletal muscle ergoreceptors in chronic heart failure [in Polish]. *Kardiologia Pol*. 2004;60(4):374–376. PMID:15226790.
35. Middlekauff HR. Making the case for skeletal myopathy as the major limitation of exercise capacity in heart failure. *Circ Heart Failure*. 2010;3(4):537–546. doi:10.1161/CIRCHEARTFAILURE.109.903773
36. Coats AJS. What causes the symptoms of heart failure? *Heart*. 2001;86(5):574–578. doi:10.1136/heart.86.5.574
37. Piepoli MF, Coats AJS. The 'skeletal muscle hypothesis in heart failure' revised. *Eur Heart J*. 2013;34(7):486–488. doi:10.1093/eurheartj/ehs463
38. Mentz RJ, Whellan DJ, Reeves GR, et al. Rehabilitation intervention in older patients with acute heart failure with preserved versus reduced ejection fraction. *JACC Heart Fail*. 2021;9(10):747–757. doi:10.1016/j.jchf.2021.05.007
39. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Kardiologia Pol*. 2016;74(10):1037–1147. doi:10.5603/KP.2016.0141
40. Rehabilitacja pacjentów z niewydolnością serca. In: *Rekomendacje w zakresie realizacji kompleksowej rehabilitacji kardiologicznej – Stanowisko Ekspertów Sekcji Rehabilitacji Kardiologicznej i Fizjologii Wysiłku Polskiego Towarzystwa Kardiologicznego*. Gdańsk, Poland: Wydawnictwo Asteria Med;65–72. <https://kif.info.pl/file/2018/07/Rekomendacje-w-zakresie-realizacji-kompleksowej-rehabilitacji-kardiologicznej-1.pdf>.
41. Hodgson CL, Stiller K, Needham DM, et al. Expert consensus and recommendations on safety criteria for active mobilization of mechanically ventilated critically ill adults. *Crit Care*. 2014;18(6):658. doi:10.1186/s13054-014-0658-y
42. Connolly B, Salisbury L, O'Neill B, et al. Exercise rehabilitation following intensive care unit discharge for recovery from critical illness. *Cochrane Database Syst Rev*. 2015;2015(6):CD008632. doi:10.1002/14651858.CD008632.pub2
43. Goldfarb M, Afilalo J, Chan A, Herscovici R, Cercsek B. Early mobility in frail and non-frail older adults admitted to the cardiovascular intensive care unit. *J Crit Care*. 2018;47:9–14. doi:10.1016/j.jcrrc.2018.05.013
44. Węgrzynowska-Teodorczyk K, Siennicka A, Josiak K, et al. Evaluation of skeletal muscle function and effects of early rehabilitation during acute heart failure: Rationale and study design. *BioMed Res Int*. 2018;2018:1–8. doi:10.1155/2018/6982897
45. Flint KM, Forman DE. Lessons from the first 202 REHAB-HF participants: Will adherence be the Achilles' heel of exercise rehabilitation among older adults hospitalized for heart failure? *Circ Heart Fail*. 2018;11(11):e005611. doi:10.1161/CIRCHEARTFAILURE.118.005611
46. Reeves GR, Whellan DJ, O'Connor CM, et al. A novel rehabilitation intervention for older patients with acute decompensated heart failure. *JACC Heart Fail*. 2017;5(5):359–366. doi:10.1016/j.jchf.2016.12.019
47. Martin-Du Pan RC, Benoit R, Girardier L. The role of body position and gravity in the symptoms and treatment of various medical diseases. *Swiss Med Wkly*. 2004;134(37–38):543–551. doi:2004/37/smw-09765
48. Mehra MR, Stewart GC, Uber PA. The vexing problem of thrombosis in long-term mechanical circulatory support. *J Heart Lung Transplant*. 2014;33(1):1–11. doi:10.1016/j.healun.2013.12.002
49. Allen LA, Tang F, Jones P, Breeding T, Ponirakis A, Turner SJ. Signs, symptoms, and treatment patterns across serial ambulatory cardiology visits in patients with heart failure: Insights from the NCDR PINNACLE® registry. *BMC Cardiovasc Disord*. 2018;18(1):80. doi:10.1186/s12872-018-0808-2
50. Dahn R, Walker S. New medications in the treatment of acute decompensated heart failure. *Hosp Pharm*. 2018;53(2):85–87. doi:10.1177/001857817750096

# Effects of ferroptosis in myocardial ischemia/reperfusion model of rat and its association with Sestrin 1

Feng Yang<sup>A,C,D,F</sup>, Wei Wang<sup>A,E,F</sup>, Yiling Zhang<sup>B,C</sup>, Jifei Nong<sup>B</sup>, Longdan Zhang<sup>B,C</sup>

Department of Emergency, First Affiliated Hospital of Guangxi Medical University, Nanning, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):219–231

## Address for correspondence

Wei Wang

E-mail: weiwanggx@163.com

## Funding sources

This study was supported by the National Nature Science Foundation of China (grant No. 81860346 and 81560318); Training Project of '139' Program for High-level Medical Talents in Guangxi (grant No. G201903034); Nanning Qingxiu District Science and Technology Project (grant No. 2020046); and Guangxi Medical and Health Suitable Technology Development Project (grant No. S2017021).

## Conflict of interest

None declared

Received on January 5, 2022

Reviewed on August 1, 2022

Accepted on September 7, 2022

Published online on November 22, 2022

## Cite as

Yang F, Wang W, Zhang Y, Nong J, Zhang L.

Effects of ferroptosis in myocardial ischemia/reperfusion model of rat and its association with Sestrin 1.

*Adv Clin Exp Med.* 2023;32(2):219–231.

doi:10.17219/acem/153599

## DOI

10.17219/acem/153599

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Ferroptosis is a type of iron-dependent programmed cell death. The inhibition of ferroptosis has been reported to alleviate myocardial ischemia/reperfusion injury (IRI). However, it is unknown whether this protective effect occurs in the ischemia or reperfusion phase. Sestrin 1 (Sesn1) possesses remarkable cytoprotective functions to diverse cellular stresses. However, whether Sesn1 is involved in the regulatory process of ferroptosis during myocardial IRI is unknown.

**Objectives.** This study aimed to simulate an acute myocardial infarction (AMI) that occurs in rats within 6 h, verify the occurrence and effects of ferroptosis in the phases of ischemia and reperfusion, and further explore the relationship between ferroptosis, IRI and Sesn1.

**Materials and methods.** The hearts of Sprague Dawley (SD) rats undergoing ischemia for varying lengths of time or having undergone ischemia followed by varying lengths of reperfusion were examined. The occurrence of ferroptosis was verified by detecting changes in ferroptosis biomarkers. In addition, ferritin-1 (Fer-1) was administered to demonstrate the effect of ferroptosis in myocardial IRI and to detect changes in Sesn1.

**Results.** The results showed that the myocardial damage was more severe with more prolonged myocardial ischemia. There were no significant changes in ferroptosis biomarkers in cardiac tissues during the ischemia phase, the levels of iron and malondialdehyde (MDA) were elevated, and the expression of glutathione peroxidase 4 (GPX4) and ferritin heavy chain 1 (FTH1) were decreased after myocardial IRI. Compared to the ischemia/reperfusion (I/R) group, the treatment with Fer-1 before reperfusion can attenuate myocardial IRI, reverse the decrease in GPX4 and FTH1 expression, and decrease the rise in iron content and MDA. In addition, we found that the expression of Sesn1 was reduced in hearts that suffered IRI; however, the treatment with Fer-1 can reverse this situation.

**Conclusions.** Ferroptosis occurred during the myocardial reperfusion phase but not ischemia. The inhibition of ferroptosis exerted beneficial effects on myocardial IRI, providing a theoretical basis for targeted therapy in patients with AMI. Sestrin 1, regulated by ferroptosis, may play an important role in myocardial IRI.

**Key words:** ferroptosis, myocardial IRI, Fer-1, Sesn1

## Background

Acute myocardial infarction (AMI) has become one of the most dangerous diseases worldwide. The high medical costs and complications severely affect the quality of life. Currently, reperfusion strategies, such as percutaneous coronary intervention (PCI), are recommended as the optimal treatment for AMI.<sup>1</sup> The treatment within the first 6 h is considered optimal. With the development of interventional therapy and the rise in people's awareness, the treatment rate of patients with AMI within 6 h has improved, meaning that patients are receiving timely treatment. However, the consequent reperfusion will likely precipitate paradoxical cardiomyocyte dysfunction, namely ischemia/reperfusion injuries (IRIs). Studies in animal models of AMI suggested that lethal reperfusion injury accounted for up to 50% of the final size of the myocardial infarct.<sup>2</sup> Several preventive strategies have been proposed in these models to ameliorate reperfusion injuries. However, the results were disappointing. Over the past decade, researchers have proposed that many mechanisms contribute to IRI, such as calcium overload, oxidative stress, inflammation, and energy metabolism disorders, ultimately leading to myocardial cell death.<sup>3</sup>

Previously, apoptosis and necrosis were considered the main types of cardiomyocyte death. Apoptosis is a form of programmed cell death, while necrosis is a form of an accidental and uncontrolled pathological cell death. Recently, studies have demonstrated that necrosis can be regulated by different signaling pathways, namely regulated cell death (RCD).<sup>4,5</sup> According to the recommendations of the 2018 Nomenclature Committee on Cell Death (NCCD), RCD can take various forms, such as necrotic apoptosis, pyroptosis and ferroptosis, among others.<sup>6</sup> Ferroptosis is a relatively recently defined form of iron-dependent programmed cell death, first proposed by Dixon et al. in 2012.<sup>7</sup> According to the existing literature, ferroptosis is characterized by 2 aspects that distinguish it from apoptosis, necroptosis and autophagy. Regarding cell morphology, ferroptosis reduces cell mitochondria, increases mitochondrial membrane density and decreases mitochondrial cristae. Regarding cell composition, ferroptosis results in lipid peroxidation and the accumulation of reactive oxygen species (ROS). Iron is an essential mineral involved in different biological processes within organisms. Iron metabolism disorders participate in the pathological processes associated with many diseases, including type 2 diabetes, obesity, nonalcoholic fatty liver disease, and cardiovascular disease.<sup>8–10</sup> Ferroptosis leads to lipid peroxidation and high levels of ROS. Oxidative stress was one of the mechanisms involved in myocardial injury. Therefore, we wanted to elucidate whether ferroptosis was involved in myocardial ischemia and/or reperfusion injuries.

Sestrin 1 (Sesn1) belongs to the sestrin family, a group of highly conserved, stress-induced proteins that

participate in various pathophysiological processes.<sup>11–15</sup> The p53 tumor suppressor protein induces the production of sestrins. Their effects are associated with their antioxidant defense ability and regulation of intracellular signaling pathways, such as double direction regulating and controlling of the target of rapamycin kinase complexes 1 and 2 (mTORC1 and mTORC2), regulating autophagy, and activating the AMP-activated protein kinase (AMPK) pathway.<sup>16,17</sup> Furthermore, because of their high sensitivity to stress, sestrins can quickly respond to various damages, protecting the organism from stress. Recently, studies have reported Sesn1 to play a significant role in cardiovascular and cerebrovascular diseases by reducing neuronal injury caused by oxygen-glucose deprivation/reoxygenation in vitro and doxorubicin cardiotoxicity in vivo.<sup>14,18</sup> However, whether Sesn1 alleviated myocardial IRI was unknown. In addition, *Sesn1* was regulated by *p53*, a key gene in ferroptosis.<sup>19,20</sup> Therefore, we suspected Sesn1 to play a crucial cytoprotective role in ferroptosis involving myocardial IRI.

Although some studies have indicated that ferroptosis is involved in myocardial injury caused by AMI, we do not know whether it plays a role in the ischemia phase or the reperfusion phase. Furthermore, extensive studies have verified the role of Sesn1 in cell survival and oxidative stress under various conditions, but little is known about whether Sesn1 is involved in ferroptosis regulating myocardial IRI. Therefore, experimental research is required to clarify this matter.

## Objectives

This study aimed to simulate clinical AMI occurring within 6 h in rats, investigate the differences in ferroptosis between the phases of ischemia and reperfusion, and further explore the relationship between ferroptosis, IRI and Sesn1.

## Materials and methods

### Animals

Adult male Sprague Dawley (SD) rats of specific pathogen-free (SPF) grade (250–280 g) were provided by the Animal Laboratory Center, Guangxi Medical University, Nanning, China. The rats were raised by professional breeders and their living environment was strictly controlled. The rats were adaptively fed for 3–5 days before the operation. The Institutional Animal Care and Research Ethics Committee of Guangxi Medical University approved the plan for the animal experiment (approval No. 2022-KY-E-(227)). All procedures complied with the regulations specified by the National Institutes of Health Guide for the Care and Use of Laboratory Animals.

## Ischemia/reperfusion injury model

The rats were anesthetized by intraperitoneal sodium pentobarbital (40 mg/kg) injection and fixed on the operating table in the supine position. They were intubated through the larynx, and the endotracheal tube was connected to an animal ventilator to assist breathing. The mode of the ventilator was 80 bpm in respiratory rate, 20 mL/kg in tidal volume, and a respiratory ratio of 1:1.5. The myocardial infarction model was created by ligating the left anterior descending coronary artery (LAD) with a 6-0 silk suture.

The first part of the experiments was to confirm whether ferroptosis was involved in ischemia-induced myocardial injury. To simulate clinical AMI presenting within 6 h, the animals were randomly divided into 4 groups ( $n = 6$  per group), including sham group (only threading without ligation), ischemia 2-hour group, ischemia 4-hour group, and ischemia 6-hour group. After the operation was finished, blood was collected and centrifuged, and serum was taken for the detection of creatine kinase-MB (CK-MB) and malondialdehyde (MDA). In addition, the myocardium was removed from the rats to measure the expression of glutathione peroxidase 4 (GPX4), ferritin heavy chain 1 (FTH1) and iron content.

The purpose of the 2<sup>nd</sup> part of the experiments was to observe the occurrence and changes of ferroptosis in hearts of rats undergoing different reperfusion periods after suffering ischemia. Ninety rats were randomly divided into 3 clusters ( $n = 30$  per cluster), and each cluster was subsequently divided into 5 groups ( $n = 6$  per group). In briefly, the 1<sup>st</sup> cluster included the sham group, the ischemia 2-hour plus reperfusion groups (reperfusion 3-hour (R3h) group, reperfusion 6-hour (R6h) group, reperfusion 12-hour (R12h) group, and reperfusion 24-hour (R24h) group); the 2<sup>nd</sup> cluster was divided into the sham group and ischemia 4-hour plus reperfusion groups (R3h group, R6h group, R12h group, and R24h group); the 3<sup>rd</sup> cluster consisted of the sham group and the ischemia 6-hour plus reperfusion groups (R3h group, R6h group, R12h group, and R24h group). After 2 h, 4 h or 6 h of myocardial ischemia, the ligation wires were loosened and reperfusion was performed for another 3 h, 6 h, 12 h, or 24 h, respectively. When the experiment was over, serum and myocardial tissue were saved to detect the indexes of CK-MB, MDA, GPX4, and FTH1, as well as iron content. The time points of 2 h of ischemia plus 12 h of reperfusion were selected for the 3<sup>rd</sup> part of the experiments.

The 3<sup>rd</sup> part of the experiments evaluated the effects of ferroptosis on myocardial IRI. The time points of 2 h of ischemia plus 12 h of reperfusion were selected for the 3<sup>rd</sup> part of the experiments. It was a new cluster consisting of 3 groups (the sham group, the ischemia/reperfusion (I/R) group and the + Fer-1 group). Rats were treated with Fer-1 (3 mg/kg, intraperitoneal injection) before reperfusion. Finally, the serum of the rats was taken for CK-MB and MDA detection, and the myocardium was removed

to measure the expression of GPX4, FTH1 and Sesn1 as well as iron content, to calculate the myocardial infarction area, and to detect the histopathological injury. The morphological changes of the mitochondria in each group were observed using transmission electron microscope (TEM).

## Measurements of serum biochemistry and iron content

Creatine kinase-MB is a creatine kinase isoenzyme found primarily in heart tissue. An increase in CK-MB indicates myocardial cell injury or necrosis. Malondialdehyde is the stable metabolite of lipid peroxidation products that indirectly reflects the degree of myocardial cell damage. Iron content is a common biomarker of ferroptosis. The serum MDA level in the myocardial tissues was measured using an MDA assay (No. A003-1; Nanjing Jiancheng Bio-engineering Institute, Nanjing, China), serum CK-MB level was measured using a CK-MB assay (No. C060; Changchun Huili Biotech Co., Ltd., Changchun, China) and iron content was measured using a iron content assay (Solarbio, Beijing, China), all according to manufacturers' instructions.

## Histological observation

Hematoxylin and eosin (H&E) staining detected histopathological injury to the myocardium. After the heart was isolated, the tissue was fixed, embedded, sliced, dewaxed, and stained. The sections were then observed under a light microscope (Olympus BX53F; Olympus Corp., Tokyo, Japan).

## Infarct area assessment

The assessment of the infarct area directly revealed the death of cardiomyocytes. The myocardial infarction area was calculated by staining the heart slice with 2% 2,3,5-triphenyltetrazolium chloride (TTC; Solarbio), according to the manufacturer's instructions. The judgment criteria were that the normal tissue remained red and the infarct area was pale. The infarct area of the processed cardiac tissue section was calculated and analyzed using ImageJ software (National Institutes of Health, Bethesda, USA). The infarct area was expressed as the ratio of the infarction area to the total myocardial area.

## Western blot

Both GPX4 and FTH1 are well-recognized biomarkers of ferroptosis. Focusing on Sesn1 was one of the key objectives during the 3<sup>rd</sup> experiment. Therefore, western blot was used to detect the difference of GPX4, FTH1 and Sesn1 protein levels among groups. It was carried out using the standard protocol. First, the myocardial tissue was lysed with radioimmunoprecipitation assay (RIPA) buffer containing protease inhibitors (Solarbio). Next, the total

protein concentration was determined using the BCA protein assay kit (Beyotime, Nanjing, China) and the supernatant from centrifugation. Next, an equal quantity of protein samples was used to perform electrophoresis (10% or 7.5% sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE)), membrane transfer (0.22  $\mu\text{m}$  polyvinylidene difluoride (PVDF) membranes), blocking (5% skim milk), and incubation with primary antibodies, such as GPX4 (No. 67763-1-Ig; Proteintech, Rosemont, USA), FTH1 (No. A19544; Abclonal, Woburn, USA) and Sesn1 (No. ab134091; Abcam, Cambridge, UK), in a 4°C refrigerator overnight, and was finally treated with secondary antibodies (No. SA5-35521 and No. SA5-35571; Invitrogen, Waltham, USA). The endogenous control was  $\alpha$ -tubulin (No. 66031-1-Ig; Proteintech). The results were visualized and analyzed using ImageJ software. The results of the targeting proteins were normalized against control, and the results were expressed as a fold of the control.

### Quantitative real-time polymerase chain reaction assay

Quantitative real-time polymerase chain reaction assay (qRT-PCR) was carried out to detect the gene expression of *GPX4*, *FTH1* and *Sesn1* according to the standard protocol. The TriZol reagent (No. 9108; TaKaRa Bio Inc., Kusatsu, Japan) was used to collect the total RNA. The cDNA was synthesized from 1  $\mu\text{g}$  of total RNA using MonScript™ RTIII All-in-one Mix Kit (No. MR05101; Monad Biotech, Wuhan, China), according to the manufacturer's protocol. Subsequently, qRT-PCR was performed using the MonAmp™ SYBR Green qPCR Mix Kit (No. MQ10301S; Monad Biotech) on Applied Biosystems 7500 Real-Time PCR Systems (Thermo Fisher Scientific, Waltham, USA). The total reaction volume was 20  $\mu\text{L}$  per well. The amplification protocol comprised of the following PCR cycles: a single cycle of 30 s at 95°C, 40 cycles of 10 s at 95°C, a single cycle of 30 s at 60°C, and a final cycle of 15 s at 95°C. The expression levels of  $\beta$ -actin normalized the threshold cycle values for target genes. The relative gene expression data were calculated using the  $2^{-\Delta\Delta\text{CT}}$  method. The primers (Sangon Biotech, Shanghai China) were as follows:

1.  *$\beta$ -actin*: forward (5' > 3') TGTCACCAAACCTGGGACGATA; reverse (5' > 3') GGGGTGTTGAAGGTCTCAA.
2. *GPX4*: forward (5' > 3') GATACGCCGAGTGTGGTTT; reverse (5' > 3') CTTGGGCTGGACTTTCATC.
3. *FTH1*: forward (5' > 3') GCCAAATACTTTCTCCATCAA; reverse (5' > 3') TCATCACGGTCAGGTTTCT.
4. *Sesn1*: forward (5' > 3') ACCTCGTGACCCCTGACTTT; reverse (5' > 3') GCTCATTTACCCCAAACCT.

### TEM observation

The mitochondrial reduction is one of the characteristics of ferroptosis. Therefore, TEM was selected to observe the morphological changes of the myocardial mitochondria.

At the end of the experiment, the rats were euthanized by excessive anesthesia, and their hearts were removed. Then, a 1-mm<sup>3</sup> piece was cut from the cardiac apex and fixed in an electron microscope fixative solution. After preparing the specimens, the samples were observed and imaged with a TEM system (model HT7800; Hitachi, Tokyo, Japan).

### Statistical analyses

The IBM SPSS v. 23.0 software (IBM Corp., Armonk, USA) was used to perform statistical analysis. Intergroup data having independent, normal distribution and variance homogeneity were analyzed using one-way analysis of variance (ANOVA), followed by Tukey's honestly significant difference (HSD) test. If the data did not meet these conditions, the differences between groups were determined using the nonparametric Kruskal–Wallis test, followed by Dunn's post hoc test. If the data had normal distribution and homogeneity of variance, independent sample t-tests were performed between the groups. A p-value <0.05 indicated statistical significance.

## Results

### Ferroptosis did not occur during myocardial ischemia in rats

To verify whether ferroptosis was involved in myocardial ischemia injuries, rat hearts were subjected to ischemia for 2 h, 4 h and 6 h, respectively. Ischemia led to myocardial injury in rats and such injury was aggravated by prolonged periods of ischemia, manifested as a gradual increase in CK-MB ( $p < 0.001$ ; Fig. 1A). These results suggested that the myocardial ischemia injury model was established successfully. Although ferroptosis was iron-dependent, there were no differences in iron content in myocardial tissues between the sham group and ischemia groups (sham group compared to ischemia 2-hour group:  $p = 0.992$ ; compared to ischemia 4-hour group:  $p = 0.427$ ; compared to ischemia 6-hour group:  $p = 0.834$ ; Fig. 1B). Similarly, compared with the sham group, the concentration lipid peroxidation product, MDA, was not significantly higher in the ischemia groups (sham group compared to ischemia 2-hour group:  $p = 0.083$ ; compared to ischemia 4-hour group:  $p = 0.756$ ; compared to ischemia 6-hour group:  $p = 0.085$ ; Fig. 1B). There were no significant changes in protein expression of GPX4 and FTH1, well-recognized biomarkers of ferroptosis, between the sham group and ischemia groups (GPX4: sham group compared to ischemia 2-hour group:  $p = 0.808$ ; compared to ischemia 4-hour group:  $p = 0.894$ ; compared to ischemia 6-hour group:  $p = 0.994$ ; FTH1: sham group compared to ischemia 2-hour group:  $p = 0.957$ ; compared to ischemia 4-hour group:  $p = 0.979$ ; compared to ischemia 6-hour group:  $p = 0.988$ ; Fig. 1C,D). Based on the above evidence, we concluded that ferroptosis did not occur during myocardial ischemia.



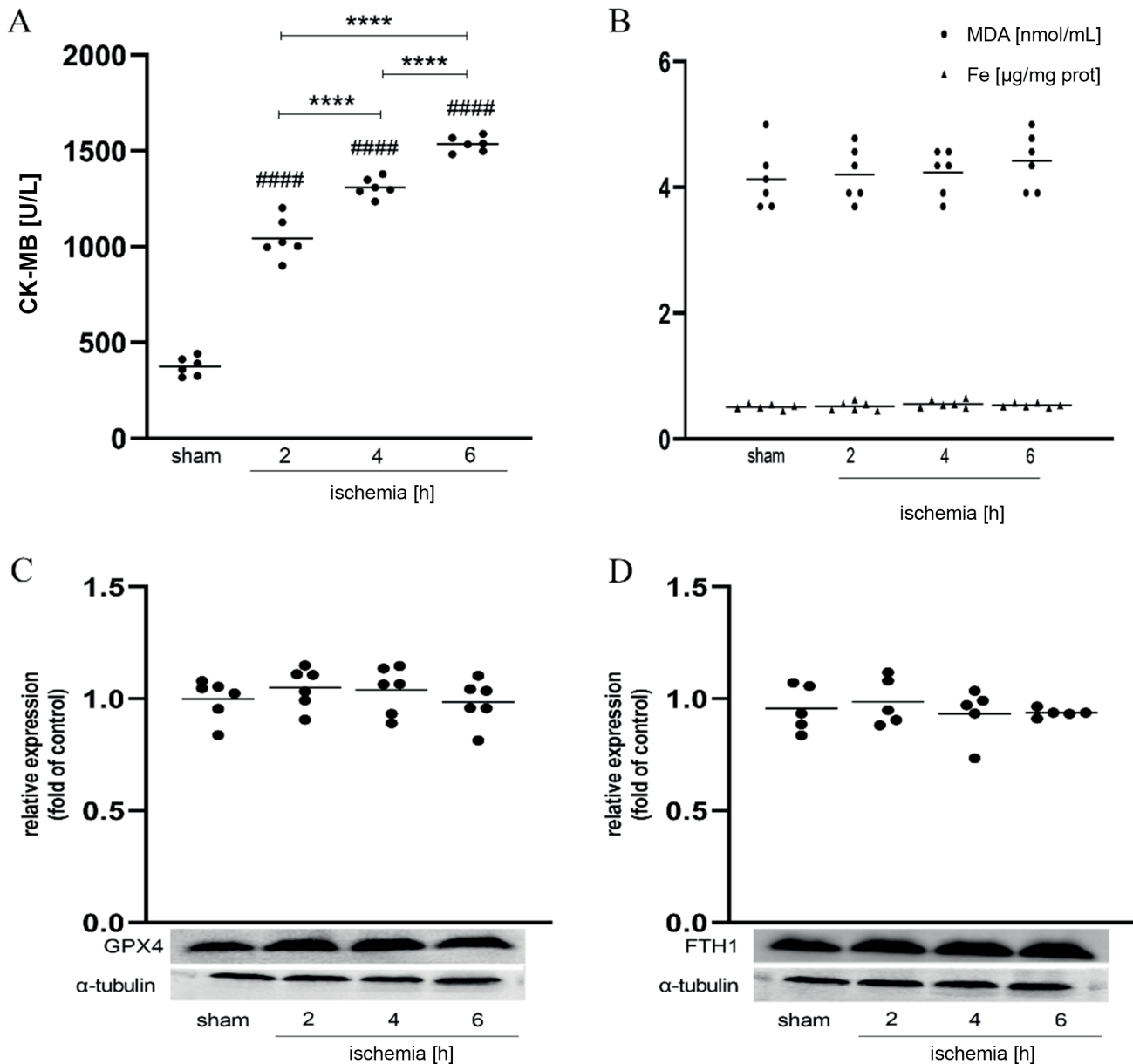


Fig. 1. No significant changes were observed regarding ferroptosis biomarkers in rat hearts subjected to ischemia. A. Creatine kinase-MB (CK-MB) levels in different ischemia groups; B. Malondialdehyde (MDA) and iron content in different groups; C,D. The protein expression of glutathione peroxidase 4 (GPX4) and ferritin heavy chain 1 (FTH1) measured using western blot. Data are presented using the mean (n = 5 or 6 per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test

#### p < 0.0001 when compared to the sham group; \*\*\*\* p < 0.0001 when compared to other ischemia groups.

### Ferroptosis participated in myocardial IRI

This part of the research aimed to observe the changes of ferroptosis in the hearts of rats that underwent ischemia for different times (2 h, 4 h and 6 h) and different reperfusion periods (R3 h, R6 h, R12 h, and R24 h). In the ischemia 2-hour plus reperfusion groups, CK-MB level (Fig. 2A) increased in the reperfusion groups (compared to the sham group: p < 0.001), and peaked at R12h (compared to R3h: p < 0.001; compared to R6h: p = 0.008; compared to R24h: p < 0.001). As illustrated in Fig. 2B, there was a significant difference in MDA level between the sham group and

reperfusion groups (compared to the sham group: for R3h p = 0.006, for R6h, R12h and R24h p < 0.001). The peak formed at R12h (compared to the R3h: p = 0.049; compared to R6h: p = 0.421; compared to R24h: p = 0.538). The iron content (Fig. 2B) was significantly different between the sham group and the reperfusion groups (p < 0.001) and there were no statistical differences between reperfusion groups. In addition, as shown in Fig. 2C,D, the protein levels of GPX4 and FTH1 decreased in the reperfusion groups; the decreases began 3 h after myocardial reperfusion (compared to sham group: p < 0.001). There was no significant difference in GPX4 protein level along with

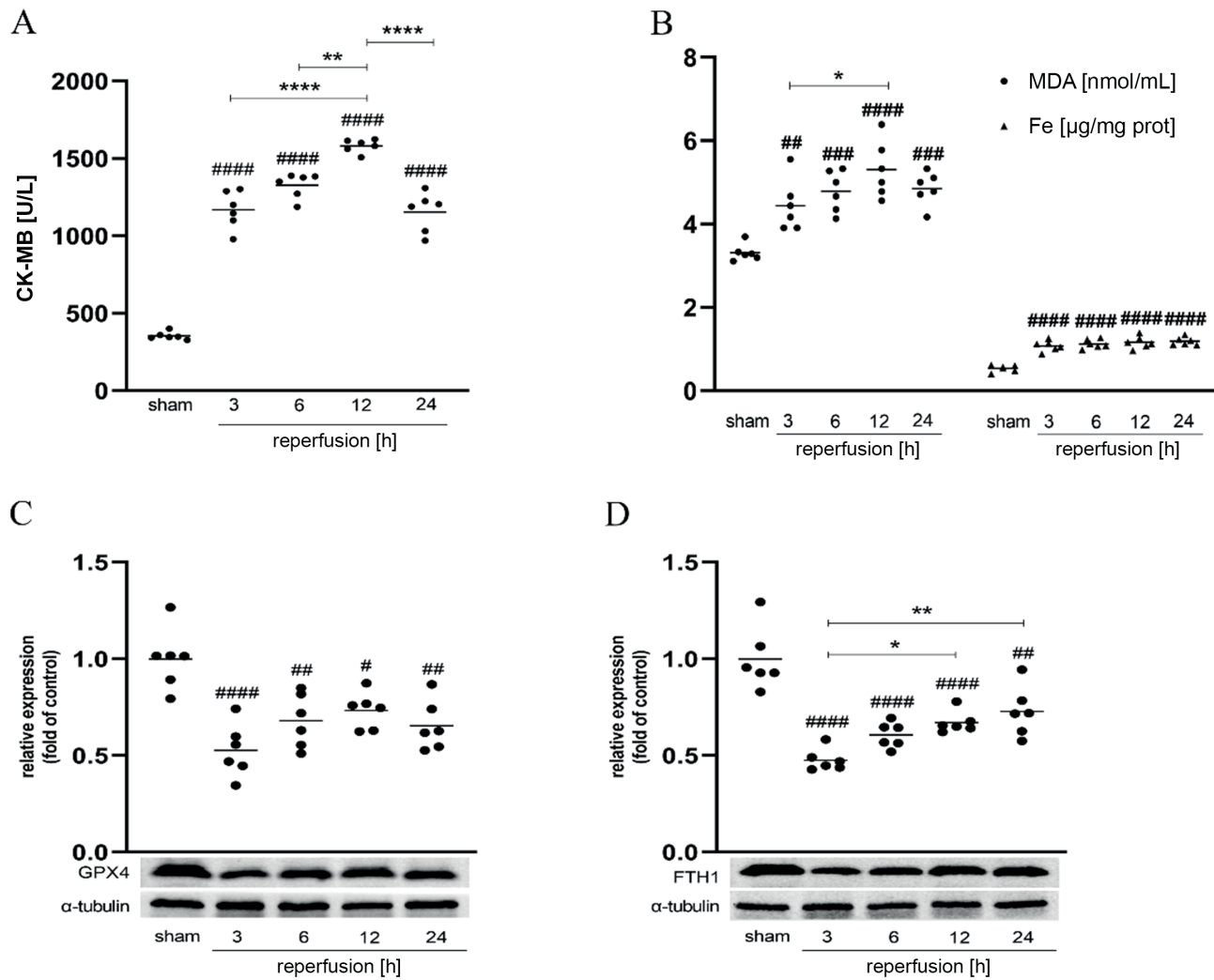


Fig. 2. Ferroptosis occurred in rat hearts subjected to ischemia for 2 h plus reperfusion. A. Creatine kinase-MB (CK-MB) levels in different groups; B. Malondialdehyde (MDA) and iron content in different groups. C,D. The protein expression of glutathione peroxidase 4 (GPX4) and ferritin heavy chain 1 (FTH1) measured using western blot. Data are presented using the mean ( $n = 6$  per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test

#  $p < 0.05$ , ##  $p < 0.01$ , ###  $p < 0.001$ , ####  $p < 0.0001$  when compared to sham; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*\*  $p < 0.0001$  when compared to other reperfusion groups.

myocardial reperfusion prolongation. An upward trend of FTH1 could be observed at R3h (R3h compared to R12h:  $p = 0.025$ ; R3h compared to R24h:  $p = 0.003$ ).

As demonstrated in Fig. 3A, in the ischemia 4-hour plus reperfusion groups, CK-MB level increased in the reperfusion groups (compared to sham group:  $p < 0.001$ ) and peaked at R12h (compared to R3h and R6h:  $p < 0.001$ ; compared to R24h:  $p = 0.172$ ). Compared to the sham group, both MDA level and iron content (Fig. 3B) increased in the reperfusion groups (MDA: for R12h  $p = 0.008$ , for R24h  $p = 0.022$ ; iron content: for R3h, R6h, R12h, and R24h  $p < 0.001$ ). There were no statistical differences between the reperfusion groups. Compared to the sham group, the protein levels of GPX4 and FTH1 (Fig. 3C,D) decreased significantly, which began at R3h (for GPX4:  $p = 0.023$ , for FTH1:  $p < 0.001$ ). No significant differences in GPX4 and FTH1 were found between the reperfusion groups.

In the ischemia 6-hour plus reperfusion groups, compared with the sham group, CK-MB level (Fig. 4A) also increased in the reperfusion groups ( $p < 0.001$ ) and peaked at R24h (compared to R3h and R6h:  $p < 0.001$ ; compared to R12h:  $p = 0.004$ ). As shown in Fig. 4B, an upward trend of MDA level can be observed among the sham group and reperfusion groups (compared to the sham group: for R3h  $p = 0.12$ , for R6h  $p = 0.048$ , for R12h  $p = 0.005$ , for R24h  $p < 0.001$ ). However, there were no statistical differences between the reperfusion groups. Compared to the sham group, the iron content (Fig. 4B) increased in reperfusion groups (for R3h:  $p = 0.001$ ; for R6h, R12h and R24h:  $p < 0.001$ ) and peaked at R6h (compared to R3h:  $p < 0.001$ ; compared to R12h and R24h:  $p = 1.000$ ). There were significant differences in the protein expression of GPX4 and FTH1 (Fig. 4C,D) between the sham group and reperfusion groups. The decrease of GPX4 and FTH1 began at R6h ( $p < 0.001$ ). There was no significant difference

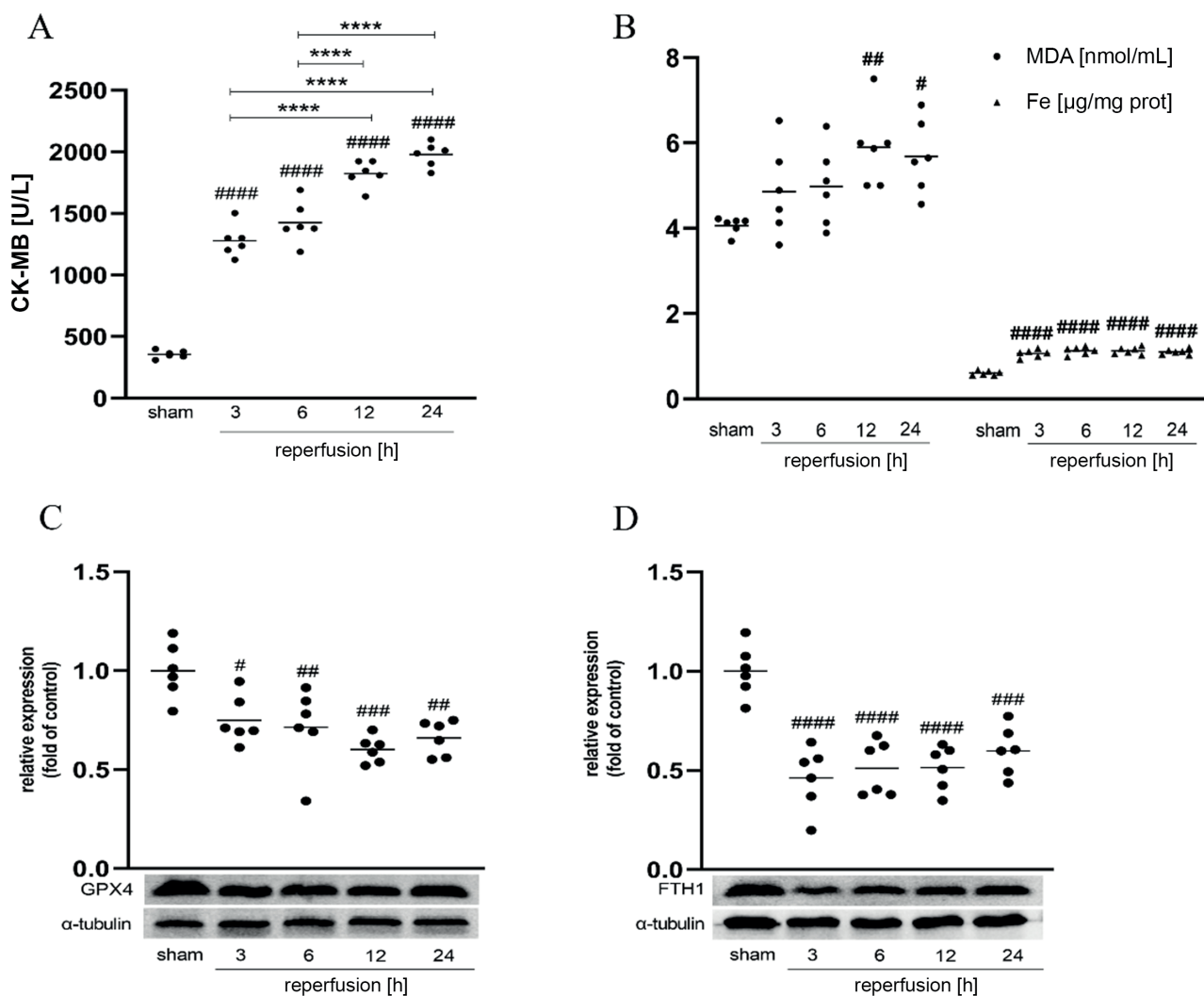


Fig. 3. Ferroptosis occurred in rat hearts subjected to ischemia for 4 h plus reperfusion. A. Creatine kinase-MB (CK-MB) levels in different groups; B. Malondialdehyde (MDA) and iron content in different groups. C,D. The protein expression of glutathione peroxidase 4 (GPX4) and ferritin heavy chain 1 (FTH1) measured using western blot. Data are presented using the mean (n = 6 per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test

# p < 0.05, ## p < 0.01, ### p < 0.001, #### p < 0.0001 when compared to sham; \*\*\*\* p < 0.0001 when compared to other reperfusion groups.

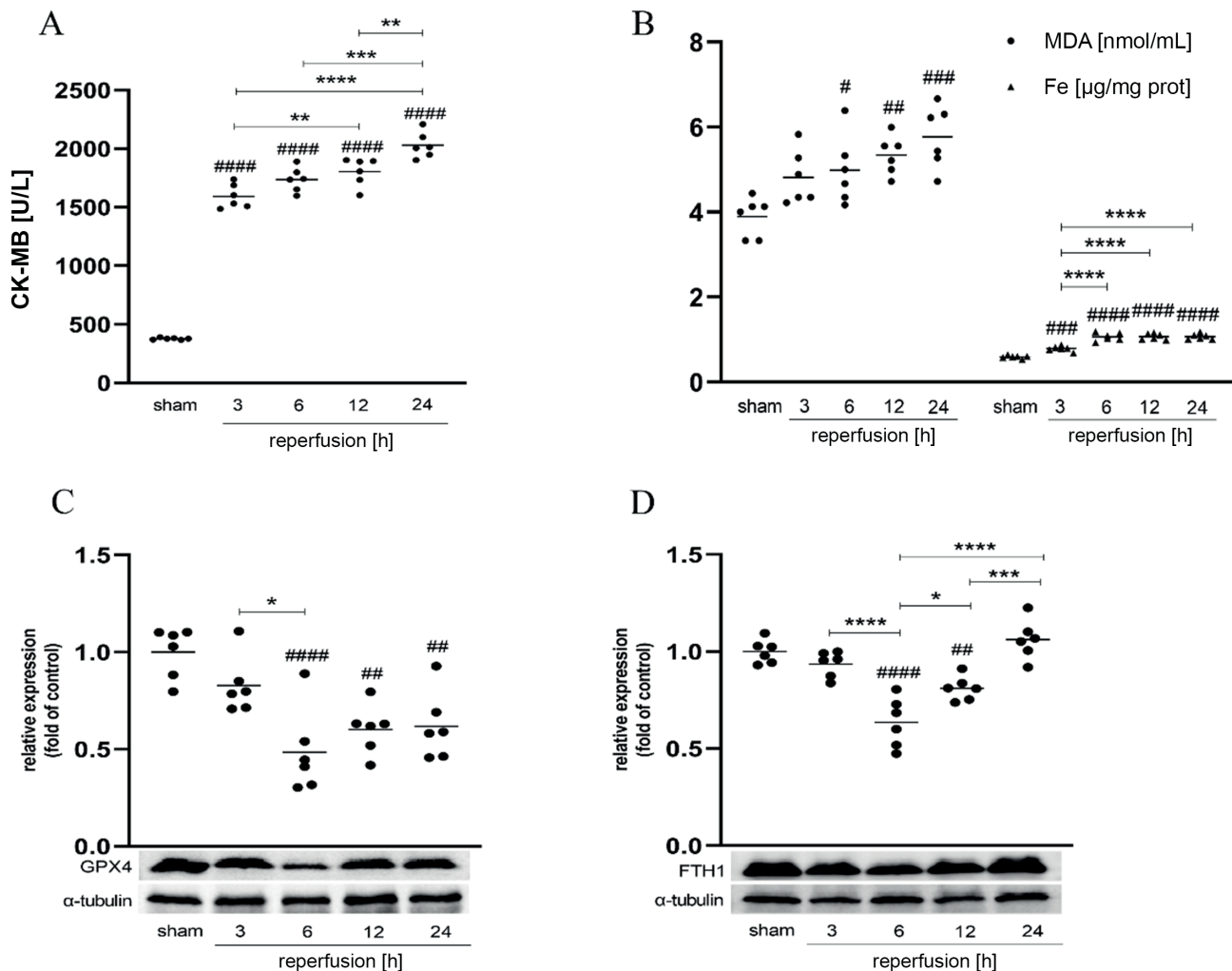
in GPX4 along with myocardial reperfusion prolongation. An upward trend of FTH1 was observed in the R6h group (R6h compared to R12h: p = 0.015; R6h compared to R24h: p < 0.001; R12h compared to R24h: p < 0.001).

### Administration of Fer-1 alleviated myocardial IRI in rats

The Fer-1, a ferroptosis inhibitor, was intraperitoneally injected into rats before reperfusion to evaluate the contribution of ferroptosis to myocardial IRI. Compared to the I/R group, Fer-1 significantly reduced myocardial IRI in the +Fer-1 group. The rise of serum CK-MB activity reversed (p < 0.001; Fig. 5A), the area of myocardial infarction decreased (p = 0.008; Fig. 5B,C) and the injury of myocardial pathological tissue was alleviated (Fig. 5D). The results verified that inhibition of ferroptosis played a therapeutic role in myocardial IRI.

### Treatment with Fer-1 alleviated myocardial IRI in an iron-dependent manner and influenced the expression of *Sesn1*

The levels of MDA, iron content, GPX4, FTH1, *Sesn1*, and mitochondrial ultrastructure in the myocardium were all measured simultaneously to detect the effects of Fer-1. Compared to the sham group, MDA level and iron content (Fig. 6A) increased in the I/R group (for MDA: p < 0.001; for iron content: p = 0.004). The increased effect of MDA induced by I/R can be significantly blocked by Fer-1 (for I/R group compared to +Fer-1 group: p = 0.008); however, Fer-1 had no effect on the iron content (for I/R group compared to +Fer-1 group: p = 0.589). As shown in Fig. 6B, the transcription level of *GPX4*, *FTH1* and *Sesn1* decreased in the I/R group (compared to the sham group: for *GPX4* p < 0.001, for *FTH1* p = 0.001 and for *Sesn1* p < 0.001); in addition, treating with Fer-1 can reduce the decline of *Sesn1*,



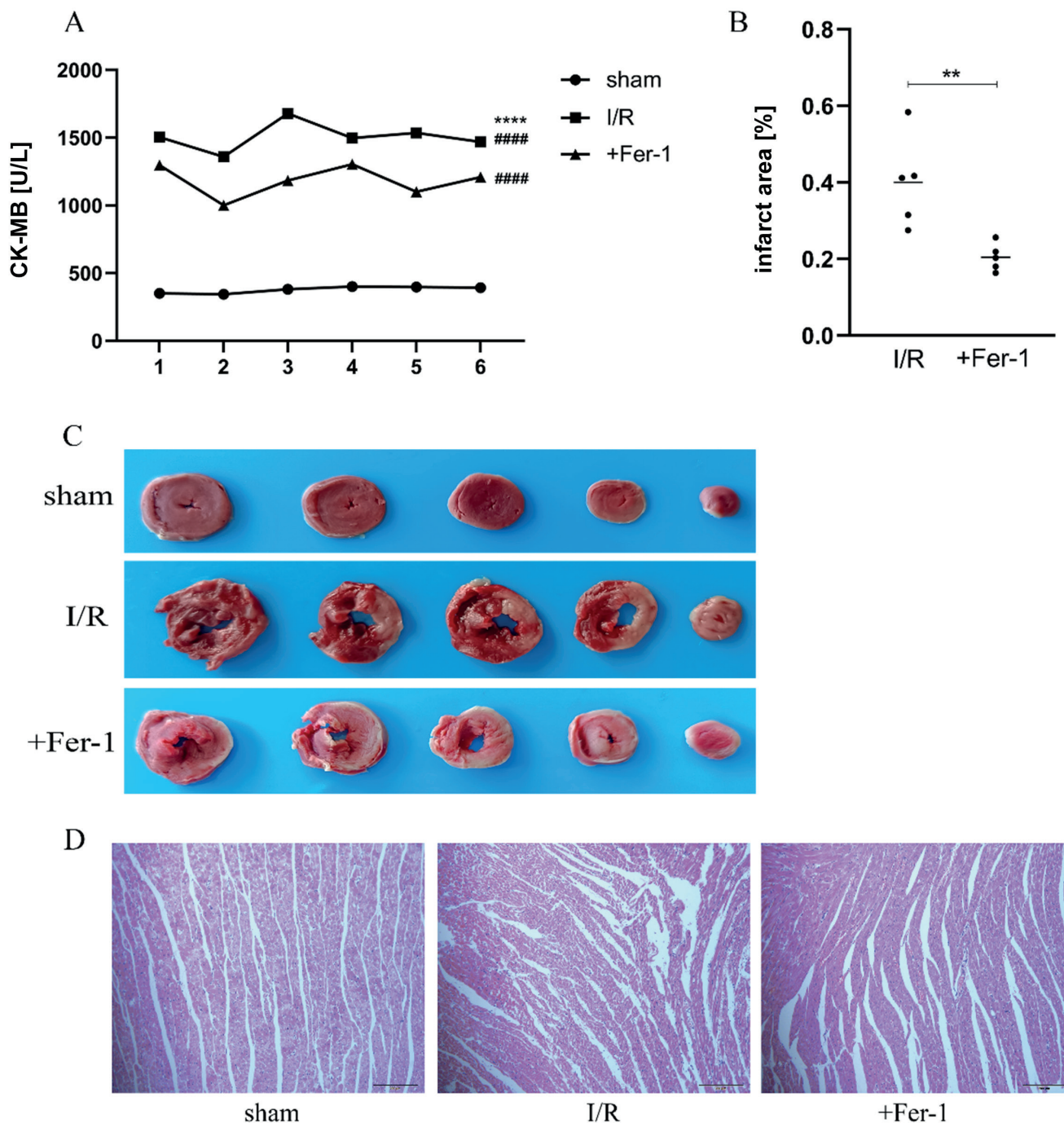
**Fig. 4.** Ferroptosis occurred in rat hearts subjected to ischemia for 6 h plus reperfusion. **A.** Creatine kinase-MB (CK-MB) levels in different groups; **B.** Malondialdehyde (MDA) and iron content in different groups; **C,D.** The protein expression of glutathione peroxidase 4 (GPX4) and ferritin heavy chain 1 (FTH1) measured using western blot. Data are presented using the mean ( $n = 6$  per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test

#  $p < 0.05$ , ##  $p < 0.01$ , ###  $p < 0.001$ , ####  $p < 0.0001$  when compared to the sham group; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ , \*\*\*\*  $p < 0.0001$  when compared to other reperfusion groups.

but not *GPX4* and *FTH1* (I/R group compared to +Fer-1 group: for *GPX4*  $p = 0.142$ , for *FTH1*  $p = 0.557$  and for *Sesn1*  $p = 0.025$ ). The trends for *GPX4*, *FTH1* and *Sesn1* protein expression levels were consistent with the transcription levels. As demonstrated in Fig. 6C–E, owing to the occurrence of ferroptosis during myocardial I/R in rats, the protein expression level of *GPX4*, *FTH1* and *Sesn1* decreased (compared to the sham group:  $p < 0.001$ ); however, this influence can be reversed by Fer-1 (I/R group compared to +Fer-1 group: for *GPX4* and *FTH1*  $p < 0.001$ , for *Sesn1*  $p = 0.017$ ). In addition, characteristic mitochondrial changes, such as smaller mitochondria, increased membrane density and decreased or fractured cristae, occurred in the I/R group due to ferroptosis. At the same time, Fer-1 alleviated these mitochondrial changes (Fig. 6F). In summary, the treatment with Fer-1 can alleviate myocardial IRI in an iron-dependent manner. Sestrin 1 was differentially expressed in ferroptosis involving myocardial IRI model.

## Discussion

In this research, our primary objective was to explore whether ferroptosis is involved in myocardial IRI in vivo. Our findings were as follows: 1) Ferroptosis was not observed in rat hearts that underwent ischemia from 2 h to 6 h, although the ischemic injury worsened with prolonged ischemia. 2) Ferroptosis occurred in rat hearts that suffered ischemia for 2 h, 4 h or 6 h plus reperfusion from 3 h to 24 h, and the duration of ischemia determined the appearance time of ferroptosis during myocardial reperfusion. 3) The administration of Fer-1 significantly attenuated myocardial IRI by inhibiting ferroptosis. 4) Treating rats with Fer-1 reversed the decrease in the expression of *Sesn1* caused by myocardial IRI. This study demonstrated the role ferroptosis plays in the ischemic and reperfusion phases of myocardial IRI. In addition, this study proposed that *Sesn1* was associated with ferroptosis in myocardial IRI.

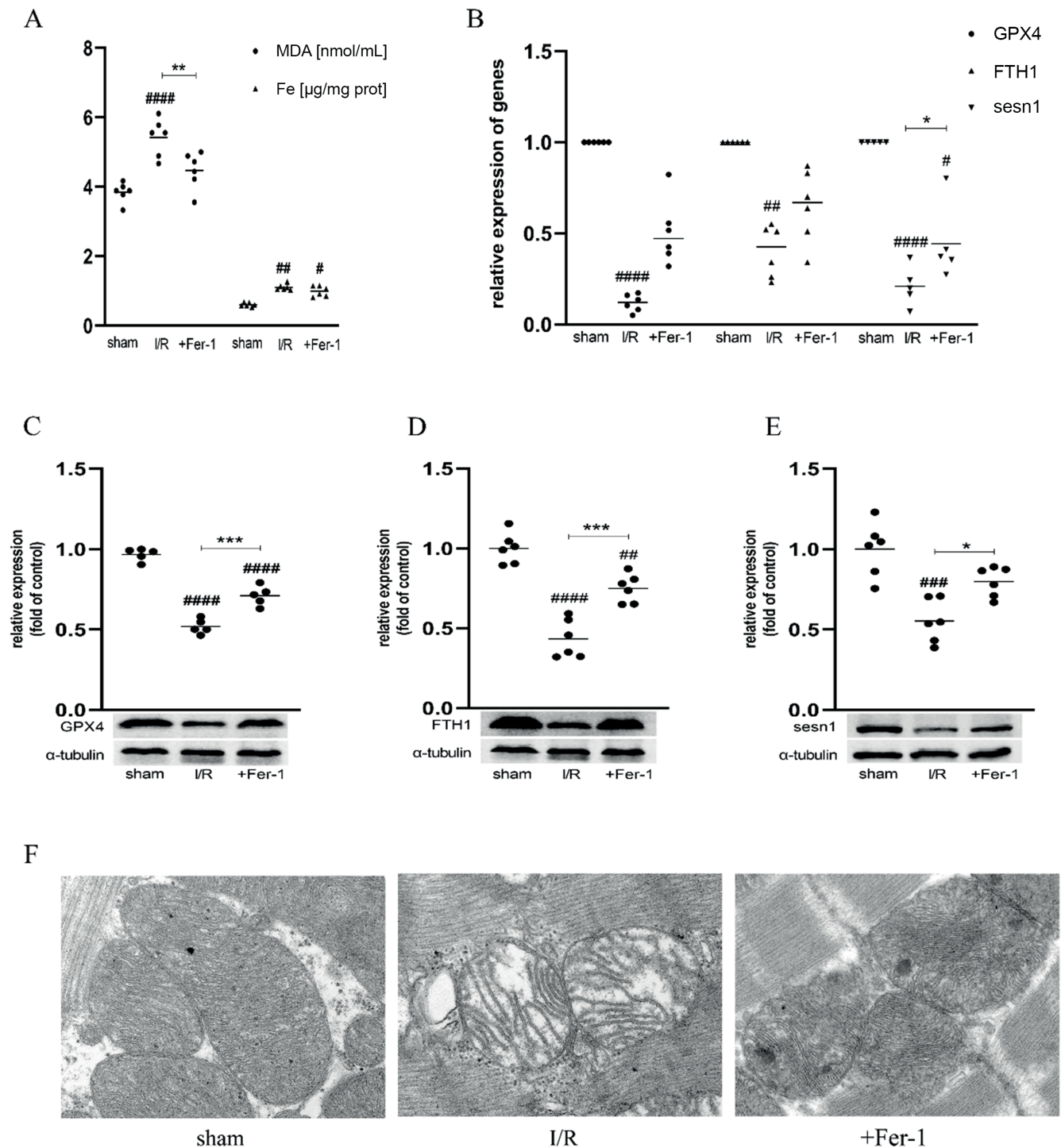


**Fig. 5.** The inhibition of ferroptosis alleviated myocardial ischemia/reperfusion injury (IRI). A. Creatine kinase-MB (CK-MB) levels in different groups. Data are presented using the mean (n = 6 per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test; B,C. Myocardial infarction size detected using 2,3,5-triphenyltetrazolium chloride (TTC) in different groups. Data are presented using the mean (n = 5 per group) and independent sample t-tests; D. Histopathological images of different groups

#### p < 0.0001 when compared to sham group; \*\* p < 0.01, \*\*\*\* p < 0.0001 in the ischemia/reperfusion (I/R) group when compared to +Fer-1 group. Fer-1 – ferrstatin-1.

It is without a doubt that ischemia can lead to myocardial injury, and the fact that reperfusion causes injury cannot be ignored. The study was initially designed based on the following considerations: in clinical practice, it can be observed that some patients refuse PCI for personal reasons or are not suitable candidates for PCI. In these cases, limited effects of other treatment result in permanent myocardial injury. On the other hand, with

growing awareness of AMI, more patients see a doctor in time to open the infarct-related arteries. However, further aggravation of myocardial injury occurs following reperfusion. Although many studies have tried to explain the mechanisms of myocardial ischemia injury and reperfusion injury, the difference between them has not been fully clarified, and the effects of clinical therapies have been disappointing. The current research hotspot, ferroptosis,



**Fig. 6.** Changes in ferroptosis biomarkers in rat hearts after treatment with Fer-1. **A.** Malondialdehyde (MDA) and iron content in different groups. The data on MDA levels are presented using the mean ( $n = 6$  per group) and one-way analysis of variance (ANOVA) followed by Tukey's honest significant difference (HSD) test. The data of iron content are presented using the median ( $n = 6$  per group), and nonparametric Kruskal–Wallis test followed by Dunn's post hoc test; **B.** The transcriptional expression levels of glutathione peroxidase 4 (GPX4), ferritin heavy chain 1 (FTH1) and Sestrin 1 (Sesn1) in cardiac tissue in different groups. The data of GPX4 and FTH1 are presented using the median ( $n = 6$  per group) and nonparametric Kruskal–Wallis test followed by Dunn's post hoc test. The data of Sesn1 is presented using the mean ( $n = 5$  per group) and one-way ANOVA followed by Tukey's HSD test; **C–E.** The protein expression of GPX4, FTH1 and Sesn1 in cardiac tissue in different groups. Data are presented using the mean ( $n = 5$  or 6 per group) and one-way ANOVA followed by Tukey's HSD testing; **F.** The changes in the mitochondria of myocardial tissue were observed using transmission electron microscope (TEM) in different groups

#  $p < 0.05$ , ##  $p < 0.01$ , ###  $p < 0.001$ , ####  $p < 0.0001$  when compared to the sham group; \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  in the ischemia/reperfusion (I/R) group when compared to +Fer-1 group. Fer-1 – ferritin-1.

has aroused our interest to contribute to the distinction between myocardial ischemia injury and reperfusion injury, laying a basis for precision medicine.

Next, we would like to discuss ischemic time and why we chose persistent ischemia for 2 h, 4 h and 6 h. At present, the ligation of the LAD is a mature model of AMI in animals.<sup>21</sup> According to most studies, the ischemic time of AMI models is less than 1 h. The rodent heart is in a state of reversible injury when the myocardial ischemia time is too short and will not cause apparent infarction. However, once the ischemia time is too long, transmural infarction occurs, which can induce the death of model animals during the experiment. In the pre-experiment phase, we found that rats can tolerate myocardial ischemia well. The longest ischemia time was 1 week, and still the mortality of these rats was low. It was found that ferroptosis occurred distinctly in rats after 1 h of myocardial ischemia followed by reperfusion. However, no differences were observed in iron and MDA content between the reperfusion and ischemia groups, except for a visible, increasing trend in the reperfusion group.<sup>22</sup> The researchers explained that ferroptosis might occur slightly in rats with cardiac ischemia. Since the time of myocardial ischemia they chose was less than 1 h, we wondered whether the apparent ferroptosis could be observed by prolonging the time of ischemia.

On the other hand, the optimum treatment time for myocardial reperfusion is within 6 h. Therefore, this study aimed to simulate clinical AMI within 6 h using a rat model. Significant differences exist between humans and animals. Although we do not know the ideal treatment time of AMI in rats, our purpose was to explore whether ferroptosis was involved in myocardial injury resulting from ischemia within or beyond the treatment time window or from reperfusion, which would contribute to precision medicine and save the damaged cardiomyocytes to the greatest extent. Therefore, we designed 3 time points (2 h, 4 h and 6 h) to dynamically observe the change during myocardial infarction. In the 2<sup>nd</sup> experiment, 4 reperfusion subgroups according to the reperfusion time (R3h, R6h, R12h, and R24h) were constructed. Our design aimed to observe the changes in ferroptosis during the extension of reperfusion.

Iron participates in various metabolic processes, such as electron transport, DNA synthesis and oxygen transport. Several studies have shown ferroptosis to be an important form of cell death in cardiomyocytes.<sup>23–25</sup> Evidence demonstrates that ferroptosis occurs primarily due to the lethal accumulation of lipid peroxidation products, which are iron-dependent due to glutathione (GSH) depletion or inhibition of system Xc<sup>-</sup> (a cystine/glutamate antiporter system).<sup>26</sup> Several vital metabolic pathways are involved in this process. The 1<sup>st</sup> is the iron metabolic pathway – iron overload produces lipid peroxidation products through the Fenton reaction. The 2<sup>nd</sup> is GSH production and the GPX4 pathway, which plays a protective

role to reduce lipid peroxidation. The 3<sup>rd</sup> is the glucose metabolism, which generates nicotinamide adenine dinucleotide phosphate (NADPH) through pentose phosphate pathways. NADPH, as the coenzyme of GSH, plays an important role in maintaining the content of GSH in cells.<sup>23</sup> Ferroptosis was reported to be related to the pathogenesis of many diseases, such as cancer, neurological diseases and sepsis.<sup>27–29</sup> Recent studies found that ferroptosis is involved in myocardial IRI. The levels of iron, MDA and the expression of ACSL4 in reperused rat hearts were gradually increased, with a decreased GPX4 level compared to the control group.<sup>22</sup> Similarly, in our study, we found that ferroptosis was stimulated in the myocardial I/R group, which was evidenced by the decreased expression of GPX4 and FTH1, and the increased levels of iron and MDA. In particular, we were surprised to find that the duration of ischemia determined the appearance time of ferroptosis during myocardial reperfusion. In the myocardial ischemia (2 h and 4 h) plus reperfusion groups, ferroptosis biomarkers changed significantly at R3h, which meant that ferroptosis occurred in rat hearts. However, when the myocardial ischemia was prolonged to 6 h, ferroptosis emerged later – at R6h. In our opinion, this result can be explained as follows: Since ferroptosis only occurs during the myocardial reperfusion phase rather than the ischemia phase, ferroptosis was found to be closely related to blood flow. The longer the coronary artery occlusion, the more serious the damage to the vascular structure and function. The total time of myocardial ischemia was frequently reported to be associated with an increased incidence of the no-reflow phenomenon.<sup>30,31</sup> Research revealed that prolonged total ischemic time was an independent predictor of the no-reflow phenomenon.<sup>32</sup> The present study depicted that no matter how long the time of myocardial ischemia (2–6 h), ferroptosis only occurred during the reperfusion phase, and the myocardial injury was alleviated after the administration of a ferroptosis inhibitor – Fer-1. Our research provided a definite time orientation for ferroptosis in myocardial IRI and a basis for clinical treatment.

Sestrin 1 is essential for maintaining normal function during oxidative injury. Its overexpression suppresses inflammation and apoptosis induced by oxidized low-density lipoproteins (ox-LDL) in human umbilical vein endothelial cells.<sup>33</sup> Silencing *Sesn1* can guide the inflammation and lipid accumulation in RAW264.7 cells exposed to ox-LDL.<sup>34</sup> Research showed that after being fed a Western diet for 8 weeks, *Sesn1* knockout mice exhibited higher levels of hepatic lipotoxicity and oxidative stress than their wild-type counterparts.<sup>15</sup> These findings suggested an antioxidant function of *Sesn1*. Over the past few years, extensive research has revealed that the essence of ferroptosis is an iron-dependent accumulation of lipid peroxidation induced by oxidative stress.<sup>35–38</sup> Recently, researchers have verified that the interactions between *Sesn1*, *p53* and *mtf2* participate in many pathological processes.<sup>14,39–41</sup>

While *p53* and *nrf2* are the hub genes for ferroptosis, we speculated that there was a connection between *Sesn1* and ferroptosis. Oxidative stress was one of the crucial mechanisms of myocardial IRI. However, the relationship between *Sesn1*, ferroptosis and myocardial IRI was unclear. In our study, the expression of *Sesn1* in rat hearts decreased significantly after ischemia (2 h)/reperfusion (12 h), but this effect was reversed after the administration of Fer-1. In addition, we confirmed that the inhibition of ferroptosis can alleviate myocardial IRI. In summary, *Sesn1* is involved in the mechanism by which ferroptosis regulates myocardial IRI.

## Limitations

There were some limitations of this study that should be put forward for discussion to guide follow-up research. First, whether the inhibition of ferroptosis can protect against myocardial ischemia is unknown; however, there were no signs of ferroptosis in the ischemia group. Second, in this study, we verified that *Sesn1* was differentially expressed in myocardial IRI. However, we failed to confirm the effects of its overexpression or knockdown on myocardial IRI. Lastly, the upstream and downstream pathways of *Sesn1* involved in myocardial IRI were not elucidated. However, these limitations will help direct future research and our further studies.

## Conclusions

Our findings verified that ferroptosis occurred in the phase of myocardial reperfusion but not ischemia. We are the first to report that a longer duration of myocardial ischemia resulted in a later occurrence of ferroptosis during reperfusion. Moreover, the inhibition of ferroptosis alleviated reperfusion-induced myocardial damage. Our findings provide a theoretical basis for specific interventions targeting ferroptosis in AMI. Furthermore, we discovered *Sesn1* to be involved in myocardial IRI, laying the groundwork for future research.

## ORCID iDs

Feng Yang  <https://orcid.org/0000-0003-4216-4225>  
 Wei Wang  <https://orcid.org/0000-0003-3941-4860>  
 Yiling Zhang  <https://orcid.org/0000-0001-5148-3880>  
 Jifei Nong  <https://orcid.org/0000-0002-2623-4003>  
 Longdan Zhang  <https://orcid.org/0000-0002-5189-5022>

## References

- Ibanez B, James S, Agewall S, et al. 2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation. *Eur Heart J*. 2018;39(2):119–177. doi:10.1093/eurheartj/ehx393
- Yellon DM, Hausenloy DJ. Myocardial reperfusion injury. *N Engl J Med*. 2007;357(11):1121–1135. doi:10.1056/NEJMra071667
- Chen J, Luo Y, Wang S, Zhu H, Li D. Roles and mechanisms of SUMOylation on key proteins in myocardial ischemia/reperfusion injury. *J Mol Cell Cardiol*. 2019;134:154–164. doi:10.1016/j.jmcc.2019.07.009
- Gong P, Zhang Z, Zou C, et al. Hippo/YAP signaling pathway mitigates blood–brain barrier disruption after cerebral ischemia/reperfusion injury. *Behav Brain Res*. 2019;356:8–17. doi:10.1016/j.bbr.2018.08.003
- Pefanis A, Ierino FL, Murphy JM, Cowan PJ. Regulated necrosis in kidney ischemia–reperfusion injury. *Kidney Int*. 2019;96(2):291–301. doi:10.1016/j.kint.2019.02.009
- Galluzzi L, Vitale I, Aaronson SA, et al. Molecular mechanisms of cell death: Recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ*. 2018;25(3):486–541. doi:10.1038/s41418-017-0012-4
- Dixon SJ, Lemberg KM, Lamprecht MR, et al. Ferroptosis: An iron-dependent form of nonapoptotic cell death. *Cell*. 2012;149(5):1060–1072. doi:10.1016/j.cell.2012.03.042
- Sha W, Hu F, Xi Y, Chu Y, Bu S. Mechanism of ferroptosis and its role in type 2 diabetes mellitus. *J Diabetes Res*. 2021;2021:9999612. doi:10.1155/2021/9999612
- Luo Y, Chen H, Liu H, et al. Protective effects of ferroptosis inhibition on high fat diet-induced liver and renal injury in mice. *Int J Clin Exp Pathol*. 2020;13(8):2041–2049. PMID:32922599.
- Fan Z, Cai L, Wang S, Wang J, Chen B. Baicalin prevents myocardial ischemia/reperfusion injury through inhibiting ACSL4-mediated ferroptosis. *Front Pharmacol*. 2021;12:628988. doi:10.3389/fphar.2021.628988
- Budanov AV, Lee JH, Karin M. Sestrin proteins take an aging fight. *EMBO Mol Med*. 2010;2(10):388–400. doi:10.1002/emmm.201000097
- Ho A, Cho CS, Namkoong S, Cho US, Lee JH. Biochemical basis of sestrin physiological activities. *Trends Biochem Sci*. 2016;41(7):621–632. doi:10.1016/j.tibs.2016.04.005
- Kozak J, Wdowiak P, Maciejewski R, Torres A. Interactions between microRNA-200 family and sestrin proteins in endometrial cancer cell lines and their significance to anoikis. *Mol Cell Biochem*. 2019;459(1–2):21–34. doi:10.1007/s11010-019-03547-2
- Yang F, Chen R. Sestrin1 exerts a cytoprotective role against oxygen-glucose deprivation/reoxygenation-induced neuronal injury by potentiating Nrf2 activation via the modulation of Keap1. *Brain Res*. 2021;1750:147165. doi:10.1016/j.brainres.2020.147165
- Fang Z, Kim HG, Huang M, et al. Sestrin proteins protect against lipotoxicity-induced oxidative stress in the liver via suppression of C-Jun N-terminal kinases. *Cell Mol Gastroenterol Hepatol*. 2021;12(3):921–942. doi:10.1016/j.jcmgh.2021.04.015
- Budanov AV, Karin M. p53 target genes sestrin1 and sestrin2 connect genotoxic stress and mTOR signaling. *Cell*. 2008;134(3):451–460. doi:10.1016/j.cell.2008.06.028
- Dalina AA, Kovaleva IE, Budanov AV. Sestrins are gatekeepers in the way from stress to aging and disease. *Mol Biol*. 2018;52(6):823–835. doi:10.1134/S0026893318060043
- Li R, Huang Y, Semple I, Kim M, Zhang Z, Lee JH. Cardioprotective roles of sestrin 1 and sestrin 2 against doxorubicin cardiotoxicity. *Am J Physiol Heart Circ Physiol*. 2019;317(1):H39–H48. doi:10.1152/ajpheart.00008.2019
- Tang LJ, Zhou YJ, Xiong XM, et al. Ubiquitin-specific protease 7 promotes ferroptosis via activation of the p53/TfR1 pathway in the rat hearts after ischemia/reperfusion. *Free Radic Biol Med*. 2021;162:339–352. doi:10.1016/j.freeradbiomed.2020.10.307
- Chen W, Jiang L, Hu Y, et al. Ferritin reduction is essential for cerebral ischemia-induced hippocampal neuronal death through p53/SLC7A11-mediated ferroptosis. *Brain Res*. 2021;1752:147216. doi:10.1016/j.brainres.2020.147216
- Riehle C, Bauersachs J. Small animal models of heart failure. *Cardiovasc Res*. 2019;115(13):1838–1849. doi:10.1093/cvr/cvz161
- Tang LJ, Luo XJ, Tu H, et al. Ferroptosis occurs in phase of reperfusion but not ischemia in rat heart following ischemia or ischemia/reperfusion. *Naunyn Schmiedebergs Arch Pharmacol*. 2021;394(2):401–410. doi:10.1007/s00210-020-01932-z
- Chen X, Li X, Xu X, et al. Ferroptosis and cardiovascular disease: Role of free radical-induced lipid peroxidation. *Free Radic Res*. 2021;55(4):405–415. doi:10.1080/10715762.2021.1876856
- Zhang Y, Xin L, Xiang M, et al. The molecular mechanisms of ferroptosis and its role in cardiovascular disease. *Biomed Pharmacother*. 2022;145:112423. doi:10.1016/j.biopha.2021.112423
- Li N, Jiang W, Wang W, Xiong R, Wu X, Geng Q. Ferroptosis and its emerging roles in cardiovascular diseases. *Pharmacol Res*. 2021;166:105466. doi:10.1016/j.phrs.2021.105466



26. Han C, Liu Y, Dai R, Ismail N, Su W, Li B. Ferroptosis and Its potential role in human diseases. *Front Pharmacol*. 2020;11:239. doi:10.3389/fphar.2020.00239
27. Mou Y, Wang J, Wu J, et al. Ferroptosis, a new form of cell death: Opportunities and challenges in cancer. *J Hematol Oncol*. 2019;12(1):34. doi:10.1186/s13045-019-0720-y
28. Ren JX, Sun X, Yan XL, Guo ZN, Yang Y. Ferroptosis in neurological diseases. *Front Cell Neurosci*. 2020;14:218. doi:10.3389/fncel.2020.00218
29. Liu Q, Wu J, Zhang X, Wu X, Zhao Y, Ren J. Iron homeostasis and disorders revisited in the sepsis. *Free Radic Biol Med*. 2021;165:1–13. doi:10.1016/j.freeradbiomed.2021.01.025
30. Bessonov IS, Kuznetsov VA, Gorbatenko EA, Dyakova AO, Sapozhnikov SS. Influence of total ischemic time on clinical outcomes in patients with ST-segment elevation myocardial infarction. *Kardiologija*. 2021;61(2):40–46. doi:10.18087/cardio.2021.2.n1314
31. Scarpone M, Cenko E, Manfrini O. Coronary no-reflow phenomenon in clinical practice. *Curr Pharm Des*. 2018;24(25):2927–2933. doi:10.2174/1381612824666180702112536
32. Nair Rajesh G, Jayaprasad N, Madhavan S, et al. Predictors and prognosis of no-reflow during primary percutaneous coronary intervention. *Proc (Bayl Univ Med Cent)*. 2019;32(1):30–33. doi:10.1080/08998280.2018.1509577
33. Gao F, Zhao Y, Zhang B, et al. SESN1 attenuates the Ox-LDL-induced inflammation, apoptosis and endothelial-mesenchymal transition of human umbilical vein endothelial cells by regulating AMPK/SIRT1/LOX1 signaling. *Mol Med Rep*. 2022;25(5):161. doi:10.3892/mmr.2022.12678
34. Gao F, Zhao Y, Zhang B, et al. Forkhead box protein 1 transcriptionally activates sestrin1 to alleviate oxidized low-density lipoprotein-induced inflammation and lipid accumulation in macrophages. *Bioengineered*. 2022;13(2):2917–2926. doi:10.1080/21655979.2021.2000228
35. Chen GH, Song CC, Pantopoulos K, Wei XL, Zheng H, Luo Z. Mitochondrial oxidative stress mediated Fe-induced ferroptosis via the NRF2-ARE pathway. *Free Radic Biol Med*. 2022;180:95–107. doi:10.1016/j.freeradbiomed.2022.01.012
36. Liu M, Kong XY, Yao Y, et al. The critical role and molecular mechanisms of ferroptosis in antioxidant systems: A narrative review. *Ann Transl Med*. 2022;10(6):368–368. doi:10.21037/atm-21-6942
37. Leng Y, Luo X, Yu J, Jia H, Yu B. Ferroptosis: A potential target in cardiovascular disease. *Front Cell Dev Biol*. 2022;9:813668. doi:10.3389/fcell.2021.813668
38. Zeng YY, Luo YB, Ju XD, et al. Solasonine causes redox imbalance and mitochondrial oxidative stress of ferroptosis in lung adenocarcinoma. *Front Oncol*. 2022;12:874900. doi:10.3389/fonc.2022.874900
39. Cai B, Ma M, Chen B, et al. MiR-16-5p targets SESN1 to regulate the p53 signaling pathway, affecting myoblast proliferation and apoptosis, and is involved in myoblast differentiation. *Cell Death Dis*. 2018;9(3):367. doi:10.1038/s41419-018-0403-6
40. Cordani M, Butera G, Dando I, et al. Mutant p53 blocks SESN1/AMPK/PGC-1 $\alpha$ /UCP2 axis increasing mitochondrial O $_2^{\cdot-}$  production in cancer cells. *Br J Cancer*. 2018;119(8):994–1008. doi:10.1038/s41416-018-0288-2
41. Rhee SG, Bae SH. The antioxidant function of sestrins is mediated by promotion of autophagic degradation of Keap1 and Nrf2 activation and by inhibition of mTORC1. *Free Radic Biol Med*. 2015;88:205–211. doi:10.1016/j.freeradbiomed.2015.06.007



# Comprehensive analysis of cell death genes in hepatocellular carcinoma based on multi-omics data

Lanlan Zhang<sup>1,A–C</sup>, Xiaomin Chen<sup>1,C,D</sup>, Xiangchai Guo<sup>1,B,C</sup>, Huajuan Shen<sup>1,B,E</sup>, Danying Qiu<sup>1,A,C</sup>, Weiqun Jin<sup>1,A,C</sup>, Dongjie Hou<sup>2,A,E,F</sup>

<sup>1</sup> Nursing Division, Zhejiang Provincial People's Hospital, Hangzhou Medical College, China

<sup>2</sup> Plastic & Reconstructive Surgery Center, Department of Hand and Reconstructive Surgery, Zhejiang Provincial People's Hospital, Hangzhou Medical College, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;

D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):233–244

## Address for correspondence

Dongjie Hou

E-mail: djhou@stu.suda.edu.cn

## Funding sources

This research was funded by the Zhejiang Provincial Natural Science Foundation of China, Hangzhou, China (grant No. 2021KY478, LL. Z.).

## Conflict of interest

None declared

Received on May 26, 2022

Reviewed on July 29, 2022

Accepted on August 11, 2022

Published online on February 8, 2023

## Abstract

**Background.** Hepatocellular carcinoma (HCC) is a common tumor of the digestive system. Cell death is an essential process in normal tissue that consists of 3 classical pathways: apoptosis, necrosis and autophagy.

**Objectives.** To perform a comprehensive analysis of the impact of cell death on liver cancer.

**Materials and methods.** The Kyoto Encyclopedia of Genes and Genomes (KEGG) database and the Cancer Genome Atlas (TCGA) datasets were used to analyze the relationships between mutations in cell death-related genes and clinical variables of HCC. Then, we applied the DESeq2 package to identify aberrantly expressed genes in HCC and their related biological functions through a Pearson correlation analysis. Finally, a cell death-related signature of HCC was constructed using the single-factor Cox regression.

**Results.** We identified the genes involved in apoptosis, necrosis and autophagy, of which *TP53* and *SPTA1* had the highest frequency of mutations. The results revealed that cell death-related tumor mutational burden (TMB) was significant for the pathologic stage and had a strong relationship with the prognosis. Moreover, 53 cell death-related genes that are differentially expressed in HCC were screened, and 3 of them were correlated with HCC prognosis. Harvey rat sarcoma viral oncogene homolog (HRAS) affected the infiltration of immune cells and was closely correlated with ferroptosis. Peptidylprolyl isomerase A (PPIA) played a significant role in mitochondrial pathways. At last, we constructed a cell death-related signature of HCC using 10 prognosis-related genes and a nomogram based on 3 variables (expression, group and stage).

**Conclusions.** This study provided a comprehensive analysis of cell death-related genes in HCC based on multi-omics data, identified the contribution of each variable to clinical outcome and predicted the survival probability of HCC patients more directly.

**Key words:** apoptosis, necrosis, autophagy, liver hepatocellular carcinoma, multi-omics

## Cite as

Zhang L, Chen X, Guo X, et al. Comprehensive analysis of cell death genes in hepatocellular carcinoma based on multi-omics data. *Adv Clin Exp Med.* 2023;32(2):233–244. doi:10.17219/acem/152737

## DOI

10.17219/acem/152737

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Background

Hepatocellular carcinoma (HCC) is a malignant tumor of the digestive system, which ranks first in cancer incidence and mortality around the world.<sup>1,2</sup> Although the overall survival of patients with HCC has increased due to improved surgery, chemotherapy options and targeted therapy, a few HCC patients still experience distant metastasis leading to a poor prognosis.<sup>3</sup> However, the molecular mechanisms of HCC are not fully understood; thus, it is urgent to explore key factors and new biomarkers in the progression and prognosis of HCC.

Cell death is an essential process in the development of normal tissue, consisting of 3 classical pathways: apoptosis, necrosis and autophagy.<sup>4,5</sup> Apoptosis is a process in which cells cease to grow and divide. It is also called programmed cell death or cell suicide. Apoptosis can be initiated by intracellular sensors activating the intrinsic pathway or damaged receptors in the extrinsic pathway. It is characterized by nuclear fragmentation, membrane blebbing and chromosomal condensation.<sup>6–8</sup> Autophagy is a cellular stress-mediated mechanism that promotes the degradation of cellular macromolecules to protect or kill stressed cells.<sup>9</sup> Necrosis, unlike the 2 previous pathways, is an uncontrolled form of cell death, where the cell is damaged by sudden external injury (hypoxia, chemicals, heat, etc.).<sup>10</sup> The dysregulation of the above 3 cell death-related pathways can play a significant role in the initiation and development of a tumor.<sup>11</sup>

The central role of genetic abnormalities is highlighted in the process of cancer metabolism described in recent studies.<sup>12–14</sup> An increased mutation rate is a well-characterized feature of human cancers that contributes to the loss of function, including somatic mutations, copy number variations, genomic rearrangements, etc. In this study, we focused on the genetic and transcriptomic characteristics of cell death in HCC using bioinformatics analysis, along with their relationship with clinical variables, the immune response and prognosis. At last, we constructed a cell death-related signature and prognostic nomogram of HCC that could be regarded as a reference for clinical treatment.

## Objectives

Multi-omics data were used to study the impact of 3 cell death-related pathways on HCC.

## Materials and methods

### Data collection

The pathways related to cell death, including apoptosis, necrosis and autophagy, were accessed using the Kyoto Encyclopedia of Genes and Genomes (KEGG)

database. We downloaded the RNA sequencing, copy number variants, mutation variants, expression data, and clinical data of patients with HCC from the Cancer Genome Atlas (TCGA) datasets. Gene expression levels were analyzed using transcripts per million (TPM) reads. The clinical data obtained included age, sex, clinical stages, TNM stage, survival status, and survival time.

### Somatic mutations and copy number variations

Genes with mutation samples greater than 5 were selected for prognostic analysis to ensure a sufficient sample size. The Kaplan–Meier log rank tests were performed to evaluate whether the gene mutations affected the overall survival of HCC patients.

To further explore the impact of cell death-related gene mutations on HCC patients, we calculated the tumor mutation burden (TMB) and the fraction genome altered (FGA) of each sample according to gene mutation information. The calculation of FGA is based on copy number variation (CNV) segment data. The main calculation method used the number of bases in segments with a  $\log_2R > 0.2$  divided by the number of bases in all segments. The FGA included the fraction genome gain (FGG) and the fraction genome loss (FGL). Next, we evaluated the associations of FGG and FGL with the prognosis and clinical characteristics of HCC.

### Gene expression analysis

In order to determine cell death-related genes that affect the development and progression of HCC, we applied the DESeq2 package, a method for differential analysis of count data using shrinkage estimation for dispersions and fold changes to improve stability and interpretability of estimates and to perform the differentially expressed analysis. We used a false discovery rate (FDR)  $< 0.05$  and a  $\log_2$  (fold change)  $> 1$  as the screening criteria. To explore the associations between cell death-related genes and prognosis in HCC, we used the median value of gene expression as the cutoff value and divided the dataset into 2 groups. The genes identified as HCC-specific for subsequent analysis were meaningful in the differential expression and prognostic analyses.

### Functional analysis

To gain further understanding of the additional influence that cell death-related genes have on HCC, we evaluated the scores of 24 immune cells in HCC samples using the gene set variation analysis (GSVA) algorithm. The Pearson's correlation analysis was further performed to observe the relationships between cell death-related genes and the infiltration of immune cells. At the same time, we also observed whether there was a co-expression

correlation between the abovementioned genes and ferroptosis-related genes. Since cell death is generally considered to be an effect of the mitochondria, we downloaded mitochondrial-related pathway information from MitoCarta3.0 (<https://www.broadinstitute.org/files/shared/metabolism/mitocarta/human.mitocarta3.0.html>) and analyzed whether the above genes had an impact on mitochondrial-related pathways using the gene set enrichment analysis (GSEA) algorithm.

## Construction of a cell death-related signature for HCC

First, a single-factor Cox regression with cell death-related genes was performed in order to screen genes that affect the prognosis of HCC. Furthermore, because of the collinearity between the genes, we used the least absolute shrinkage and selection operator (LASSO) to screen independent genes. A signature related to cell death was constructed by combining clinical parameters of HCC-related cases.

## Statistical analyses

Statistical analysis and visualization of data were performed using R software (v. 4.1.2; R Foundation for Statistical Computing, Vienna, Austria). Measurement data were expressed as mean (standard deviation (SD)). If the data met a normal distribution, we used t-tests to assess the differences between groups and analysis of variance (ANOVA) to assess the differences between multiple groups. If the data did not conform to a normal distribution, then the differences between the 2 groups were estimated using the Wilcoxon rank-sum test, and the differences between multiple groups were assessed using the Kruskal–Wallis test. A value of  $p < 0.05$  was considered statistically significant.

## Results

### Cell-death gene set

From the KEGG database, we identified 137 autophagy-related genes, 136 apoptosis-related genes and 159 necrosis-related genes. Five genes had functions in all 3 pathways (B-cell lymphoma-2 (*BCL2*), CASP8 and FADD-like apoptosis regulator (*CFLAR*), mitogen-activated protein kinase (*MAPK*) 10, *MAPK8*, and *MAPK9*; Fig. 1). In addition, 23 genes were involved in autophagy and apoptosis, 18 genes in apoptosis and necroptosis, and 1 gene in autophagy and necroptosis. To further explore the interactive relationships among genes from the 3 cell death-related pathways, the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (<https://cn.string-db.org>) was used to confirm interactions between the genes instead of them being mutually independent (Fig. 1). Using a network diagram of protein interactions we discovered that the relationships between the 3 forms of cell death indicate strong interactions between the genes involved in the 3 types of cell death. Therefore, those 3 types may jointly affect the occurrence and development of HCC.

### Genetic mutations of cell death-related genes

After screening, it was found that 81 autophagy-related genes, 97 apoptosis-related genes and 89 necroptosis-related genes were mutated in HCC (Fig. 2A). Tumor protein 53 (*TP53*) had the highest mutation frequency in 32% of patients, and the 2<sup>nd</sup> most frequent mutation occurred in spectrin alpha erythrocytic 1 (*SPTA1*) at 8% (Fig. 2B). However, the mutation frequencies of all the other genes were less than 5%. The 2 most frequently mutated genes (*TP53* and *SPTA1*) were apoptosis-related, indicating that the functions of apoptosis-related genes in HCC may

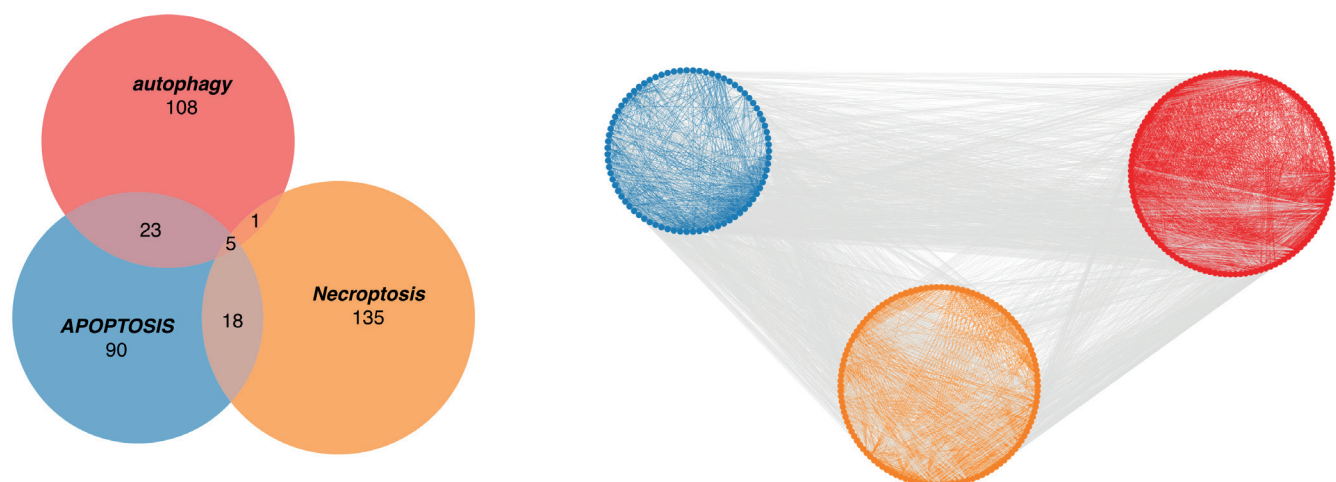
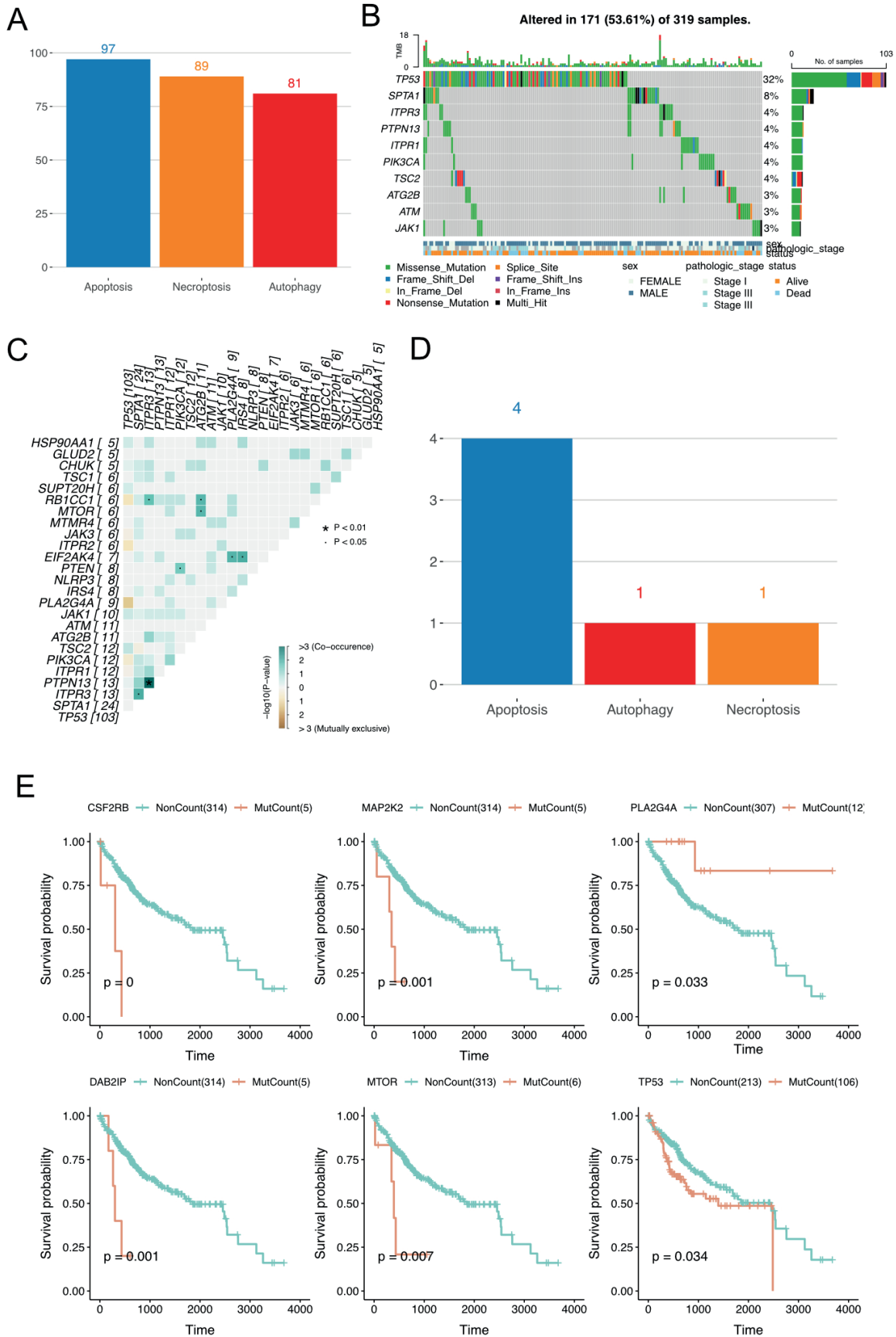


Fig. 1. Summary of gene information related to cell death based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. The left panel is a Venn diagram of the overlapping relationship among the 3 pathways: apoptosis, necrosis and autophagy. The right panel is a Venn diagram of the protein–protein interactions/relationships between the 3 pathways



**Fig. 2.** Mutation characteristics of cell death-related genes in hepatocellular carcinoma (HCC). **A.** The number of mutant genes based on the 3 pathways; **B.** OncoPrint of the top 10 cell death-related gene mutations in HCC; **C.** Mutual occurrence/mutual exclusion analysis between mutant genes; **D.** The number of genes with genetic mutations that affect the prognosis; **E.** Survival plot of the top 3 genes most influential in the prognosis of HCC. NonCount – non-mutation sample count; MutCount – mutation sample count.

indeed be related to gene mutations. In order to further explore the associations among mutated genes, we performed a co-occurrence analysis. In this analysis, only protein tyrosine phosphatase non-receptor type 13 (*PTPN13*) and inositol 1,4,5-trisphosphate receptor type 3 (*ITPR3*) were found to have a co-occurrence relationship, both of which are apoptosis-related genes (Fig. 2C).

Moreover, after a survival analysis of mutated genes, the mutation of 6 genes was correlated with the prognosis of HCC, consisting of 4 apoptosis-related genes, 1 necroptosis-related gene and 1 autophagy-related gene (Fig. 2D). Among the 6 mutated genes associated with prognosis, the mutations in apoptosis-related genes and autophagy genes (colony stimulating factor 2 receptor subunit beta (*CSF2RB*), mitogen-activated protein kinase 2 (*MAP2K2*), DAB2 interacting protein (*DAB2IP*), mechanistic target of rapamycin kinase (*MTOR*), and *TP53*) led to a decreased survival time in patients, while the mutation in the necrosis gene (phospholipase A2 group IVA (*PLA2G4A*)) led to an increased survival time (Fig. 2E). In order to observe the relationship between the 3 cell death pathway gene mutations and HCC as a whole, we further aggregated the gene mutations for prognostic analysis. The results were consistent with the prognosis of previous analyses, finding that gene mutations in the apoptotic pathway led to a decreased patient survival and gene mutations in the necrosis pathway led to an increased patient survival (Table 1).

We calculated the association of TMB with cell death in HCC to analyze the relationship between TMB and clinical parameters. Patients with a high cell death rate related to TMB have a shorter survival time, indicating that mutations in the 3 pathways make the tumor more difficult to control (Fig. 3A). Moreover, the results revealed that the cell death-related TMB was significantly associated with gender ( $p < 0.001$ ) with male patients having higher TMB than females (Fig. 3B and Table 2).

### Copy number variations of cell death-related genes

After the CNV analysis, we found apoptosis-related genes to have the most changes (Fig. 3D). Moreover, we found higher FGA, FGG and FGL values to be significant for an unfavorable prognosis in HCC ( $p = 0.008$ ,  $p = 0.007$  and  $p = 0.002$ , respectively, Fig. 3C). In addition, the associations between copy number variations and clinical characteristics were explored and the results showed

distinct correlations. Grade 4 HCC had the highest values of FGA, FGG and FGL ( $p < 0.005$ , Fig. 3E and Table 3). The more severe the HCC, the higher the number of CNV and the shorter the survival time. This is consistent with our finding that high scores in FGA, FGG and FGL lead to a poor prognosis in HCC. Additionally, radiation therapy also causes changes in FGG. A low FGG was shown in genomes not exposed to radiation therapy (Fig. 3F).

### Transcriptomes of cell death-related genes

Based on RNA-sequencing (RNA-seq) data of HCC, we identified 53 cell death-related genes that were differentially expressed in HCC compared to normal tissues, including 31 apoptosis-related genes, 15 necroptosis-related genes and 7 autophagy-related genes (Fig. 4A). Most

**Table 2.** Relationship between cell death-related tumor mutational burden (TMB) and clinical phenotype of hepatocellular carcinoma (HCC)

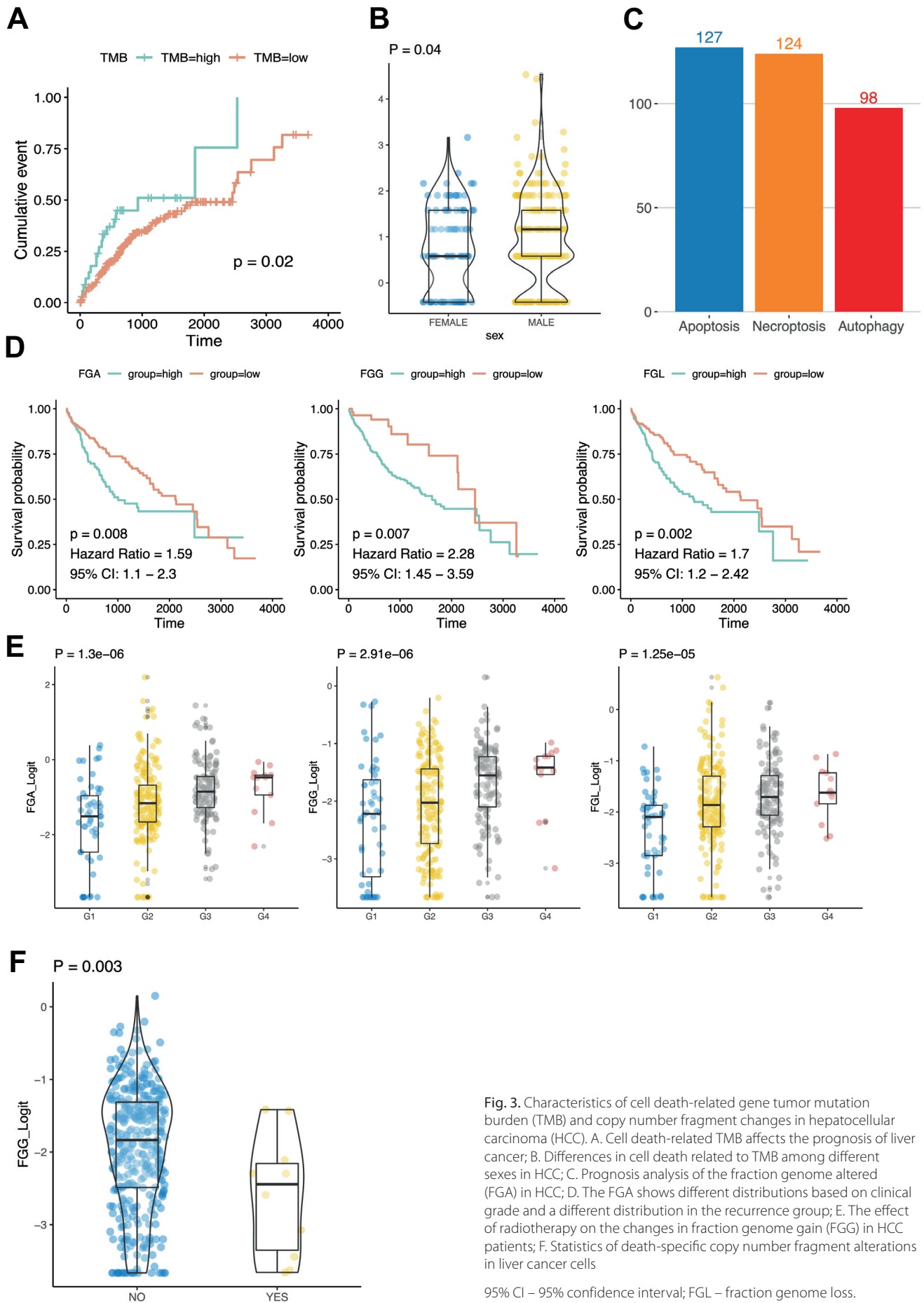
Group	n	mean ±SD	p-value
Sex			
Female	102	0.741 ±0.913	0.04
Male	217	0.974 ±1	
Radiation therapy			
No	312	0.903 ±0.984	0.552
Yes	7	0.747 ±0.644	
Pharmaceutical therapy			
No	297	0.908 ±0.977	0.573
Yes	22	0.783 ±0.993	
Tumor grade			
G1	44	0.71 ±1.008	0.203
G2	154	0.846 ±0.92	
G3	108	1.047 ±1.056	
G4	13	0.958 ±0.764	
Pathologic stage			
Stage I	145	0.882 ±0.942	0.484
Stage II	85	1.021 ±0.999	
Stage III	85	0.799 ±1.024	
Stage IV	4	1.122 ±0.747	
T stage			
T1-3	305	0.917 ±0.971	0.169
T4	13	0.464 ±1.102	

SD – standard deviation.

**Table 1.** Prognostic analysis of hepatocellular carcinoma (HCC) based on gene mutations summarized by 3 pathways

Pathway	MutCount	NonCount	DeathMuta	DeathNon	HR	95% CI	p-value
Apoptosis	110	203	44 (40%)	66 (60%)	1.718	1.133–2.606	0.004
Necroptosis	12	301	1 (0.91%)	109 (99.09%)	0.157	0.069–0.356	0.033
Autophagy	6	307	4 (3.64%)	106 (96.36%)	3.613	0.569–22.926	0.007

95% CI – 95% confidence interval; HR – hazard ratio; NonCount – non-mutation sample count; MutCount – mutation sample count; DeathMuta – mutation death sample count; DeathNon – non-mutation death sample count.



**Fig. 3.** Characteristics of cell death-related gene tumor mutation burden (TMB) and copy number fragment changes in hepatocellular carcinoma (HCC). **A.** Cell death-related TMB affects the prognosis of liver cancer; **B.** Differences in cell death related to TMB among different sexes in HCC; **C.** Prognosis analysis of the fraction genome altered (FGA) in HCC; **D.** The FGA shows different distributions based on clinical grade and a different distribution in the recurrence group; **E.** The effect of radiotherapy on the changes in fraction genome gain (FGG) in HCC patients; **F.** Statistics of death-specific copy number fragment alterations in liver cancer cells

95% CI – 95% confidence interval; FGL – fraction genome loss.



**Table 3.** Association of cell death-associated copy number alterations with clinical information in hepatocellular carcinoma (HCC) patients

Group	FGA		FGG		FGL	
	mean $\pm$ SD	p-value	mean $\pm$ SD	p-value	mean $\pm$ SD	p-value
Sex						
Female	-1.003 $\pm$ 1.103	0.24	-1.906 $\pm$ 0.827	0.381	-1.775 $\pm$ 0.918	0.209
Male	-1.141 $\pm$ 0.922		-1.987 $\pm$ 0.852		-1.896 $\pm$ 0.73	
Radiation						
No	-1.089 $\pm$ 0.99	0.208	-1.944 $\pm$ 0.838	0.037	-1.859 $\pm$ 0.8	0.802
Yes	-1.4 $\pm$ 0.711		-2.594 $\pm$ 0.836		-1.808 $\pm$ 0.616	
Pharmaceutical						
No	-1.092 $\pm$ 0.993	0.695	-1.942 $\pm$ 0.846	0.096	-1.865 $\pm$ 0.794	0.521
Yes	-1.163 $\pm$ 0.854		-2.223 $\pm$ 0.781		-1.755 $\pm$ 0.822	
Tumor grade						
G1	-1.672 $\pm$ 1.143	<0.001	-2.314 $\pm$ 1.023	<0.001	-2.335 $\pm$ 0.788	<0.001
G2	-1.143 $\pm$ 1.001		-2.072 $\pm$ 0.838		-1.845 $\pm$ 0.825	
G3	-0.826 $\pm$ 0.792		-1.691 $\pm$ 0.69		-1.7 $\pm$ 0.7	
G4	-0.765 $\pm$ 0.659		-1.594 $\pm$ 0.638		-1.633 $\pm$ 0.533	
Pathologic stage						
Stage I	-1.237 $\pm$ 1.023	0.059	-2.073 $\pm$ 0.854	0.081	-1.938 $\pm$ 0.805	0.269
Stage II	-0.922 $\pm$ 0.889		-1.828 $\pm$ 0.837		-1.75 $\pm$ 0.741	
Stage III	-1.012 $\pm$ 0.952		-1.897 $\pm$ 0.812		-1.808 $\pm$ 0.796	
Stage IV	-0.935 $\pm$ 1.389		-1.633 $\pm$ 0.876		-1.932 $\pm$ 1.272	
T stage						
T1–T3	-1.095 $\pm$ 0.993	0.92	-1.969 $\pm$ 0.852	0.161	-1.849 $\pm$ 0.801	0.257
T4	-1.117 $\pm$ 0.728		-1.725 $\pm$ 0.547		-2.065 $\pm$ 0.611	

SD – standard deviation; FGA – fraction genome altered; FGG – fraction genome gain; FGL – fraction genome loss.

of these differentially expressed genes were upregulated in HCC (Fig. 4B). Furthermore, there were 18 cell death-related genes with significant effects on the prognosis of HCC, and no gene was found to have protective factors (Fig. 4C). Through differential expression and prognostic analysis, we found that the effects of cell death-related genes on tumors were consistent, that is, high gene expression leads to tumorigenesis and poor survival rates. We conducted a cross analysis on the above differentially expressed genes and prognosis-related genes. It was found that 3 of the genes, namely peptidylprolyl isomerase A (*PPIA*), Harvey rat sarcoma viral oncogene homolog (*HRAS*) and baculoviral IAP repeat containing 5 (*BIRC5*), were not only differentially expressed in HCC, but also affected the overall survival of HCC patients, which belong to the necroptosis, apoptosis and autophagy pathways, respectively (Fig. 4D).

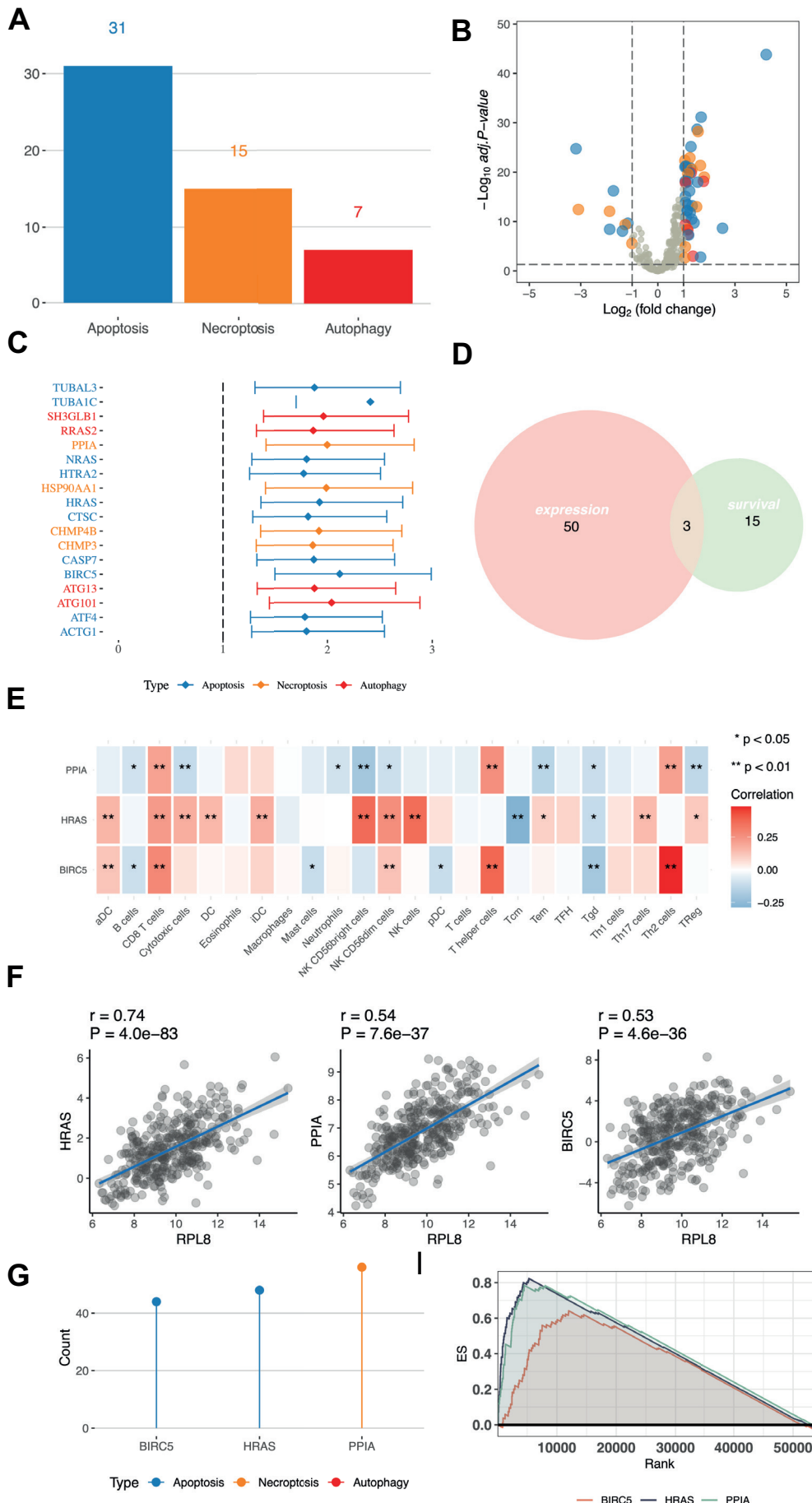
With the infiltration scores of 24 immune cells calculated by the GSVA algorithm, the correlations between the above 3 genes and immune infiltration were investigated to characterize the immunologic landscape. Through correlation analysis, we can evaluate whether there is a correlation between the expression of these 3 genes and the immune cell score. If the correlation coefficient is high, the gene may affect immune cells. The results showed that *HRAS* affected

the infiltration of immune cells (Fig. 4E), thus, it may play a significant role in the progression and prognosis of HCC by affecting the death of immune cells. Since ferroptosis is an emerging way of cell death, we analyzed the relationships between the expressions of the above 5 genes and iron death-related genes. Most genes had no association with the iron death-related genes; however, ribosomal protein L8 (*RPL8*) was co-expressed in most of them (Fig. 4F).

Using the GSEA algorithm, the effects of the above 5 genes on mitochondria-related pathways were explored, and the *PPIA*, *HRAS* and *BIRC5* genes were all significant. Among them, *PPIA* was closely correlated with 56 mitochondrial pathways (Fig. 4G). The mitochondrial complex III pathway was influenced by the expressions of *PPIA*, *HRAS* and *BIRC5* in HCC (Fig. 4H).

## Construction of cell death-related signature

We performed the Cox regression to identify the genes that affect the survival of HCC, using LASSO to screen out genes with collinearity (Fig. 5A). As a result, a total of 10 genes (4 apoptosis genes, 4 necroptosis genes and 2 autophagy genes) could be used to construct a cell death-related signature for HCC. A prognostic analysis



**Fig. 4.** Cell death-related gene expression analysis. A. Number of differentially expressed genes based on the 3 pathways; B. Volcano plot for the differential expression analysis of cell death-related genes. Blue represents apoptosis-related genes, orange represents necrosis-related genes and red represents autophagy-related genes; C. Forest plot for prognostic analysis of cell death-related genes. Blue represents apoptosis-related genes, orange represents necrosis-related genes and red represents autophagy-related genes. The prismatic point represents the hazard ratio (HR) value for the prognostic analysis, and the distribution at both ends of the line is the 95% confidence interval (95% CI) for the prognostic analysis; D. Venn diagram of cross-analysis of differentially expressed genes and prognosis-related genes; E. Correlation between hepatocellular carcinoma (HCC) cell death characteristic genes and immune cells; F. Correlation between HCC cell death genes and iron death genes. The p-value in the graph represents whether there is a correlation between the 2 types of genes. The r value represents the degree of correlation seen in the correlation analysis; G. The number of HCC characteristic genes and mitochondrial-related pathways; H. Gene set enrichment analysis (GSEA) plot of the 4 characteristic genes in complex III

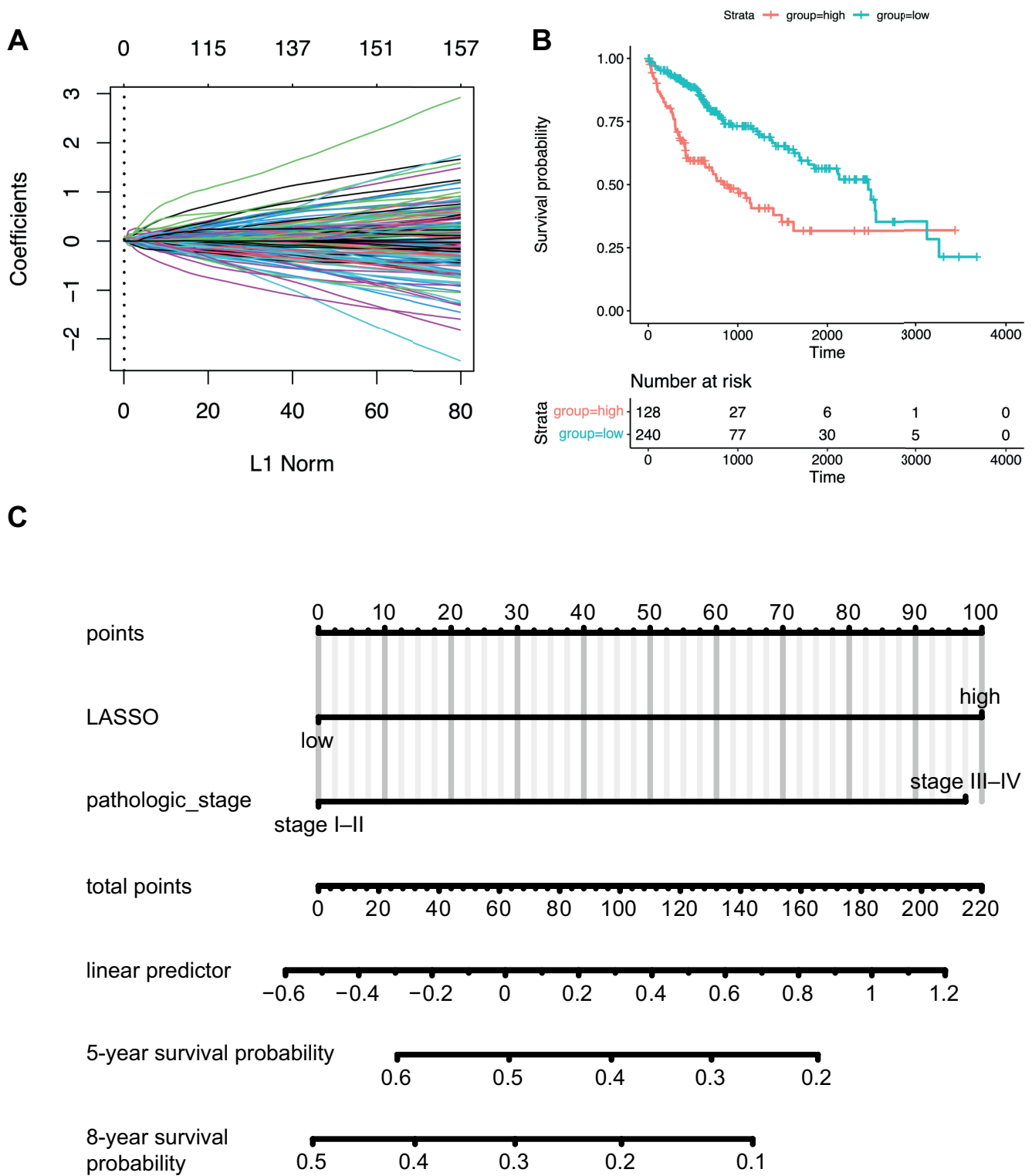


Fig. 5. Construction of the hepatocellular carcinoma (HCC)-related cell death model. A. Least absolute shrinkage and selection operator (LASSO) regression analysis selects the variables included in the model; B. Prognostic analysis of cell death-related models; C. The nomogram of HCC-related features

of the signature confirmed that the high expression groups are closely correlated with a worse survival (Fig. 5B). Finally, a multi-factor Cox regression analysis on the cell death-related signature and clinical parameters of HCC was studied. We found that the cell death-related signature, age and T stage were the most significant for the prognosis

of HCC (Table 4). Therefore, based on these 3 variables, a nomogram was constructed to indicate the contribution of each variable to the outcome and to predict the survival probability more directly (Fig. 5C). In the final nomogram, it can be seen that when the cell death model had a high score, the stage was high and the patient survival was poor.

**Table 4.** Prognostic analysis of hepatocellular carcinoma (HCC)-related clinical features

Variable	Sample number	Univariate analysis		Multivariate analysis	
Phenotype number		HR (95% CI)	p-value	HR (95% CI)	p-value
Age	367	1.01 (1.00–1.03)	0.091	1.03 (1.01–1.05)	0.09
Sex					
Female	118	reference	N/A	N/A	N/A
Male	249	0.79 (0.56–1.13)	0.201	N/A	N/A
Stage					
Stage I–II	265	reference	N/A	N/A	N/A
Stage III–IV	102	2.42 (1.71–3.43)	<0.001	3.24 (0.98–10.66)	<0.001
T stage					
T1–T3	354	reference	N/A	N/A	N/A
T4	13	3.75 (1.95–7.21)	<0.001	N/A	0.064
LASSO score					
Low	239	reference	N/A	N/A	N/A
High	128	0.41 (0.29–0.58)	<0.001	3.23 (2.06–5.07)	<0.001

LASSO – least absolute shrinkage and selection operator; HR – hazard ratio; 95% CI – 95% confidence interval; N/A – not applicable.

## Discussion

Hepatocellular carcinoma is one of the most common and lethal tumors of the digestive system, leading to an unfavorable prognosis.<sup>15</sup> Therefore, it is urgent to uncover tumor microenvironment variations resulting in the development of HCC. Cell death pathways are an essential process of maintaining the growth of an organism.<sup>16,17</sup> They mainly include apoptosis, necrosis and autophagy. Our study aimed to evaluate the dysregulation of the cell death program and the related gene signatures with the clinical characteristics of HCC. Furthermore, cell death-related signatures that could help predict the overall survival of HCC patients were built.

In this study, we identified genes related to apoptosis, necrosis and autophagy pathways in HCC and their interactions with each other. There were 5 genes seen in all 3 pathways, namely *BCL2*, *CFLAR*, *MAPK10*, *MAPK8*, and *MAPK9*. Thus, they might play significant roles in the development of HCC by acting on various cell death pathways. The *BCL2* family are considered important regulators of apoptosis and have been implicated in the initiation and progression of liver cancer.<sup>18</sup> The activation of autophagy may also be initiated by a reduction in *BCL2* levels,<sup>19</sup> which was validated in experiments on mice with colorectal cancer. The mitogen-activated protein kinase (MAPK) pathway was reported to be required for cell proliferation and survival, of which inhibition could contribute to cell cycle arrest and apoptosis, as well as autophagy-mediated cell death.<sup>20–22</sup> For example, Zhang and Jiang showed that odd-skipped related transcription factor 1 (OSR1) affected HCC progression and induced apoptosis by inhibiting MAPK pathways.<sup>23</sup> They also mentioned that SN38 could induce the activation

of the MAPK pathway and promote autophagy, thereby reducing the chemoresistance of tumor cells.<sup>24</sup>

Hepatocellular carcinoma exhibited an elevated mutational burden compared to normal tissues.<sup>25</sup> The sequence analysis of cell death-related coding genes in HCC helps to understand their potential roles in carcinogenesis.<sup>26</sup> Thus, we studied somatic mutations of genes related to the 3 cell death pathways. Among them, *TP53* presented the highest mutation frequency, with *SPTA1* having the 2<sup>nd</sup> highest frequency. Besides, *TP53* and *SPTA1* were mutually exclusive with other genes, which explains the reason for the less frequent occurrence of the mutations of other genes. The *TP53* and *SPTA1* were commonly associated with liver cancer and are considered early driver genes in malignant tumors.<sup>27,28</sup> A mutation in *TP53* may cause a loss of function resulting in resistance to apoptosis, which is characterized by inhibited cell cycle arrest and invasiveness of liver cancer.<sup>29</sup> In addition, *TP53* mutations may contribute to the deregulation of the cell cycle and antitumor immunity, which could serve as potential biomarkers for predicting the response to immunotherapy.<sup>30</sup> The *SPTA1* mutations were significant predictors of shorter survival and shorter recurrence time in liver cancer patients.<sup>31</sup> After performing a survival analysis on mutated genes, 6 genes correlated with the prognosis of HCC, with the most significant being *CSF2RB*, *MAP2K2* and *PLA2G4A*. Although their relationships with the overall survival of HCC patients have not been studied, it could be the focus of further research.

Tumor mutation burden is the number of mutations that exist within a tumor<sup>32,33</sup>; it has been identified as a potential predictor for therapy with immune checkpoint inhibitors. We calculated the TMB associated with cell death and analyzed its correlation with clinical parameters based on the cell death-related load. In regard to changes

in the genome, there are also changes in the copy number of genes. We found that the CNVs of apoptosis-related genes were observed most often. Higher FGA, FGG and FGL values were all significant predictors of an unfavorable prognosis in HCC.

In this study, we further explored RNA-seq data in HCC patients for variations in cell death-related gene expression and their relationship with the prognosis. We identified 54 cell death-related genes that were differentially expressed in HCC compared to normal tissues, and 3 of the differentially expressed genes were correlated with the prognosis: *PPIA*, *HRAS* and *BIRC5*. A decrease in the nerve growth factor (NGF) level can affect the antiproliferative function of gastrin-releasing peptide receptor antagonists in liver cancer cells,<sup>34</sup> resulting in poor overall survival. Glycogen phosphorylase, muscle associated (*PYGM*), charged multivesicular body protein 2B (*CHMP2B*), and caspase 9 (*CASP9*) may also be potential biomarkers for the diagnosis and prognosis of HCC that need to be further validated in future studies.

We explored specific functions that are under the influence of the above 5 genes in HCC. In immune infiltration, *HRAS* was investigated to affect the infiltration of mostly immune cells. The *PYGM* was shown to play a significant role in regulating T cell migration and proliferation by binding to a small GTPase in the RAS family.<sup>35</sup> The *BIRC5* is required for activation, trafficking and antitumor activity of CD8<sup>+</sup>T cells.<sup>36</sup> Therefore, we speculated that *HRAS* and *BIRC5* could be associated with the progression and prognosis of HCC by affecting the proliferation and death of immune cells. Iron is essential for most biological processes, and the regulation of ferroptosis is dependent on iron.<sup>37</sup> Ferroptosis is also a form of programmed necrotic cell death.<sup>38</sup> Since ferroptosis is an emerging form of necrosis, we also analyzed the relationships between the expression of the above 5 genes and ferroptosis-related genes. Most of the genes did not correlate with ferroptosis-related genes, except *RPL8*, which was of interest because of the influence *RPL8* has on the expression of ferroptosis in the initiation and progression of HCC. The mitochondria are known to modulate Ca<sup>2+</sup> signaling, and the massive accumulation of Ca<sup>2+</sup> in the mitochondria can contribute to apoptosis.<sup>39,40</sup> Using the GSEA algorithm, the effects of the above 5 genes on mitochondria-related pathways were explored, and *PPIA* was found to closely correlate with 58 mitochondrial pathways. The interaction between p38 MAPK and *CASP9* can regulate mitochondria-mediated apoptosis in liver carcinoma cells.<sup>41</sup>

Additionally, we constructed a cell death-related signature for HCC using 10 prognosis-related genes. Among these, the high-expression group was validated to be significant for worse survival. Age and stage are also correlated with the prognosis of HCC. Finally, a nomogram was built based on these 3 variables to more directly guide the contributions of each variable in the clinical outcomes and prediction of the survival in HCC patients.

## Limitations

This study has some limitations. First, the present results were analyzed based on high-throughput sequencing data. Specific experiments are required for verification. Therefore, special attention should be paid to the follow-up research. Second, the genes of the 3 cell death methods studied are based on the classical genes studied in the past, and there is no way to analyze genes that may have similar mechanisms yet to be studied. Finally, because the TCGA database was used, its specific ethnicity can have an impact on cell death-related models. Therefore, it needs to be expanded to other races and ethnicities for verification.

## Conclusions

In the present study, we focused on the genetic and transcriptomic characteristics of cell death in HCC, along with their relationships with clinical variables, prognosis and biological functions. Finally, we constructed a cell death-related signature and prognostic nomogram for HCC that could be considered a reference for clinical use.

### ORCID iDs

Lanlan Zhang  <https://orcid.org/0000-0001-5097-4186>  
 Xiaomin Chen  <https://orcid.org/0000-0002-1862-7851>  
 Xiangchai Guo  <https://orcid.org/0000-0002-4407-2060>  
 Huajuan Shen  <https://orcid.org/0000-0003-1300-1750>  
 Danying Qiu  <https://orcid.org/0000-0002-0031-5315>  
 Weiqun Jin  <https://orcid.org/0000-0001-7596-7812>  
 Dongjie Hou  <https://orcid.org/0000-0003-1112-6488>

### References

1. Tang J, Chen JX, Chen L, et al. Metastasis associated in colon cancer 1 (MACC1) promotes growth and metastasis processes of colon cancer cells. *Eur Rev Med Pharmacol Sci.* 2016;20(13):2825–2834. PMID:27424982.
2. Stein U, Walther W, Arlt F, et al. MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis. *Nat Med.* 2009;15(1):59–67. doi:10.1038/nm.1889
3. Chang Z, Huang R, Fu W, et al. The construction and analysis of ceRNA network and patterns of immune infiltration in liver hepatocellular carcinoma metastasis. *Front Cell Dev Biol.* 2020;8:688. doi:10.3389/fcell.2020.00688
4. D'Arcy MS. Cell death: A review of the major forms of apoptosis, necrosis and autophagy. *Cell Biol Int.* 2019;43(6):582–592. doi:10.1002/cbin.11137
5. Chen Q, Kang J, Fu C. The independence of and associations among apoptosis, autophagy, and necrosis. *Signal Transduct Target Ther.* 2018; 3:18. doi:10.1038/s41392-018-0018-5
6. Xu X, Lai Y, Hua ZC. Apoptosis and apoptotic body: Disease message and therapeutic target potentials. *Biosci Rep.* 2019;39(1):BSR20180992. doi:10.1042/BSR20180992
7. Mariño G, Niso-Santano M, Baehrecke EH, Kroemer G. Self-consumption: The interplay of autophagy and apoptosis. *Nat Rev Mol Cell Biol.* 2014;15(2):81–94. doi:10.1038/nrm3735
8. Fink SL, Cookson BT. Apoptosis, pyroptosis, and necrosis: Mechanistic description of dead and dying eukaryotic cells. *Infect Immun.* 2005;73(4):1907–1916. doi:10.1128/IAI.73.4.1907-1916.2005
9. He C, Klionsky DJ. Regulation mechanisms and signaling pathways of autophagy. *Annu Rev Genet.* 2009;43:67–93. doi:10.1146/annurev-genet-102808-114910

10. Karsch-Bluman A, Feiglin A, Arbib E, et al. Tissue necrosis and its role in cancer progression. *Oncogene*. 2019;38(11):1920–1935. doi:10.1038/s41388-018-0555-y
11. Degterev A, Yuan J. Expansion and evolution of cell death programmes. *Nat Rev Mol Cell Biol*. 2008;9(5):378–390. doi:10.1038/nrm2393
12. Song J, Ye A, Jiang E, et al. Reconstruction and analysis of the aberrant lncRNA-miRNA-mRNA network based on competitive endogenous RNA in CESC. *J Cell Biochem*. 2018;119(8):6665–6673. doi:10.1002/jcb.26850
13. Ojesina AI, Lichtenstein L, Freeman SS, et al. Landscape of genomic alterations in cervical carcinomas. *Nature*. 2014;506(7488):371–375. doi:10.1038/nature12881
14. Zhou N, Yuan Y, Long X, Wu C, Bao J. Mutational signatures efficiently identify different mutational processes underlying cancers with similar somatic mutation spectra. *Mutat Res*. 2017;806:27–30. doi:10.1016/j.mrfmmm.2017.07.004
15. Das V, Kalita J, Pal M. Predictive and prognostic biomarkers in colorectal cancer: A systematic review of recent advances and challenges. *Biomed Pharmacother*. 2017;87:8–19. doi:10.1016/j.biopha.2016.12.064
16. Majno G, Joris I. Apoptosis, oncosis, and necrosis. An overview of cell death. *Am J Pathol*. 1995;146(1):3–15. PMID:7856735.
17. Fricker M, Tolkovsky AM, Borutaite V, Coleman M, Brown GC. Neuronal cell death. *Physiol Rev*. 2018;98(2):813–880. doi:10.1152/physrev.00011.2017
18. Ramesh P, Medema JP. BCL-2 family deregulation in colorectal cancer: Potential for BH3 mimetics in therapy. *Apoptosis*. 2020;25(5–6):305–320. doi:10.1007/s10495-020-01601-9
19. Talero E, Alcaide A, Ávila-Román J, García-Mauriño S, Vendramini-Costa D, Motilva V. Expression patterns of sirtuin 1-AMPK-autophagy pathway in chronic colitis and inflammation-associated colon neoplasia in IL-10-deficient mice. *Int Immunopharmacol*. 2016;35:248–256. doi:10.1016/j.intimp.2016.03.046
20. Toit-Kohn JL, Louw L, Engelbrecht AM. Docosahexaenoic acid induces apoptosis in colorectal carcinoma cells by modulating the PI3 kinase and p38 MAPK pathways. *J Nutr Biochem*. 2009;20(2):106–114. doi:10.1016/j.jnutbio.2007.12.005
21. Lei Y, Yuan H, Gai L, Wu X, Luo Z. Uncovering active ingredients and mechanisms of spica prunellae in the treatment of liver hepatocellular carcinoma: A study based on network pharmacology and bioinformatics. *Comb Chem High Throughput Screen*. 2021;24(2):306–318. doi:10.2174/1386207323999200730210536
22. Remacle-Bonnet MM, Garrouste FL, Heller S, André F, Marvaldi JL, Pommier GJ. Insulin-like growth factor-I protects colon cancer cells from death factor-induced apoptosis by potentiating tumor necrosis factor alpha-induced mitogen-activated protein kinase and nuclear factor kappaB signaling pathways. *Cancer Res*. 2000;60(7):2007–2017. PMID:10766192.
23. Zhang F, Jiang Z. Downregulation of OSR1 promotes liver hepatocellular carcinoma progression via FAK-mediated Akt and MAPK signaling. *Oncotargets Ther*. 2020;13:3489–3500. doi:10.2147/OTT.S242386
24. Paillas S, Causse A, Marzi L, et al. MAPK14/p38 $\alpha$  confers irinotecan resistance to TP53-defective cells by inducing survival autophagy. *Autophagy*. 2012;8(7):1098–1112. doi:10.4161/autophagy.20268
25. Lee-Six H, Olafsson S, Ellis P, et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature*. 2019;574(7779):532–537. doi:10.1038/s41586-019-1672-7
26. Wolff RK, Hoffman MD, Wolff EC, et al. Mutation analysis of adenomas and carcinomas of the colon: Early and late drivers. *Genes Chromosomes Cancer*. 2018;57(7):366–376. doi:10.1002/gcc.22539
27. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: Origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol*. 2010;2(1):a001008. doi:10.1101/cshperspect.a001008
28. Buscail L, Bournet B, Cordelier P. Role of oncogenic KRAS in the diagnosis, prognosis and treatment of pancreatic cancer. *Nat Rev Gastroenterol Hepatol*. 2020;17(3):153–168. doi:10.1038/s41575-019-0245-4
29. Watanabe S, Tsuchiya K, Nishimura R, et al. TP53 mutation by CRISPR system enhances the malignant potential of liver cancer. *Mol Cancer Res*. 2019;17(7):1459–1467. doi:10.1158/1541-7786.MCR-18-1195
30. Li L, Li M, Wang X. Cancer type-dependent correlations between TP53 mutations and antitumor immunity. *DNA Repair (Amst)*. 2020;88:102785. doi:10.1016/j.dnarep.2020.102785
31. Taieb J, Le Malicot K, Shi Q, et al. Prognostic value of BRAF and SPTA1 mutations in MSI and MSS stage III liver cancer. *J Natl Cancer Inst*. 2016;109(5):djw272. doi:10.1093/jnci/djw272
32. Ritterhouse LL. Tumor mutational burden. *Cancer Cytopathol*. 2019;127(12):735–736. doi:10.1002/cncy.22174
33. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med*. 2017;377(25):2500–2501. doi:10.1056/NEJMc1713444
34. de Farias C, Stertz L, Lima R, Kapczinski F, Schwartzmann G, Roesler R. Reduced NGF secretion by HT-29 human colon cancer cells treated with a GRPR antagonist. *Protein Pept Lett*. 2009;16(6):650–652. doi:10.2174/092986609788490177
35. Migocka-Patrzałek M, Elias M. Muscle glycogen phosphorylase and its functional partners in health and disease. *Cells*. 2021;10(4):883. doi:10.3390/cells10040883
36. Wei Z, Du Q, Li P, et al. Death-associated protein kinase 1 (DAPK1) controls CD8<sup>+</sup>T cell activation, trafficking, and antitumor activity. *FASEB J*. 2021;35(1):e21138. doi:10.1096/fj.201903067RR
37. Nakamura T, Naguro I, Ichijo H. Iron homeostasis and iron-regulated ROS in cell death, senescence and human diseases. *Biochim Biophys Acta Gen Subj*. 2019;1863(9):1398–1409. doi:10.1016/j.bbagen.2019.06.010
38. Park S, Oh J, Kim M, Jin EJ. Bromelain effectively suppresses Kras-mutant colorectal cancer by stimulating ferroptosis. *Anim Cells Syst (Seoul)*. 2018;22(5):334–340. doi:10.1080/19768354.2018.1512521
39. Jeong SY, Seol DW. The role of mitochondria in apoptosis. *BMB Rep*. 2008;41(1):11–22. doi:10.5483/bmbrep.2008.41.1.011
40. Bauer TM, Murphy E. Role of mitochondrial calcium and the permeability transition pore in regulating cell death. *Circ Res*. 2020;126(2):280–293. doi:10.1161/CIRCRESAHA.119.316306
41. Wang Y, Xia C, Lun Z, Lv Y, Chen W, Li T. Crosstalk between p38 MAPK and caspase-9 regulates mitochondria-mediated apoptosis induced by tetra- $\alpha$ -(4-carboxyphenoxy) phthalocyanine zinc photodynamic therapy in LoVo cells. *Oncol Rep*. 2018;39(1):61–70. doi:10.3892/or.2017.6071

# Prognostic relevance of clinicopathological factors in sporadic and syndromic odontogenic keratocysts: A comparative study

Konrad Kisielowski<sup>1,A–D,F</sup>, Bogna Drozdowska<sup>2,A,C,D,F</sup>, Mariusz Szuta<sup>1,C,E,F</sup>, Tomasz Kaczmarzyk<sup>1,A–D,F</sup>

<sup>1</sup> Chair of Oral Surgery, Jagiellonian University Medical College, Kraków, Poland

<sup>2</sup> Department of Pathomorphology, Faculty of Medical Science in Zabrze, Medical University of Silesia, Poland

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online)

*Adv Clin Exp Med.* 2023;32(2):245–259

## Address for correspondence

Tomasz Kaczmarzyk  
E-mail: tomasz.kaczmarzyk@uj.edu.pl

## Funding sources

Grant from the Medical University of Silesia in Katowice, Poland (grant No. KNW-1-142/N/5/0).

## Conflict of interest

None declared

## Acknowledgements

The authors would like to express their gratitude to Mr. Łukasz Deryło, MSc., for his professional assistance in the statistical analysis.

Received on May 22, 2022

Reviewed on July 20, 2022

Accepted on August 31, 2022

Published online on October 12, 2022

## Cite as

Kisielowski K, Drozdowska B, Szuta M, Kaczmarzyk T. Prognostic relevance of clinicopathological factors in sporadic and syndromic odontogenic keratocysts: A comparative study. *Adv Clin Exp Med.* 2023;32(2):245–259. doi:10.17219/acem/153390

## DOI

10.17219/acem/153390

## Copyright

Copyright by Author(s)

This is an article distributed under the terms of the Creative Commons Attribution 3.0 Unported (CC BY 3.0) (<https://creativecommons.org/licenses/by/3.0/>)

## Abstract

**Background.** Current evidence suggests that nevoid basal cell carcinoma syndrome (NBCCS)-associated odontogenic keratocysts (OKCs) exhibit more aggressive clinical behavior and a higher tendency to relapse. The prognostic efficacy of various markers in sporadic and syndromic OKCs is unclear, and so are the results of studies on the usefulness of immunohistochemistry in distinguishing syndromic from sporadic OKCs.

**Objectives.** This retrospective study aimed to compare the prognostic relevance of various clinicoradiological and histopathological features, as well as the immunoexpression of COX-2, Bcl-2, proliferating cell nuclear antigen (PCNA), p53, Ki-67, osteoprotegerin (OPG), receptor activator of nuclear factor  $\kappa$  B (RANK) and receptor activator of nuclear factor  $\kappa$  B ligand (RANKL), as well as RANKL/OPG balance between sporadic and syndromic OKCs, and to test their utility in distinguishing the 2 types of OKC.

**Materials and methods.** We compared the immunoexpression of the aforementioned markers between 31 sporadic and 12 syndromic OKCs, and tested clinicopathological findings and levels of immunostaining against recurrence.

**Results.** We found a significant association between NBCCS and OKC recurrence. There were significant differences in PCNA, p53 and OPG immunoexpression between sporadic and syndromic OKCs. We also found that recurrent sporadic OKCs were significantly larger and markedly more often associated with cortical perforation. Recurrent sporadic OKCs exhibited COX-2 upregulation, but we failed to demonstrate its prognostic relevance. Recurrent syndromic OKCs showed a markedly higher RANKL > OPG ratio.

**Conclusions.** The NBCCS-associated OKCs are significantly more prone to recur than their sporadic counterparts. Larger size and radiological signs of cortical perforation in sporadic OKCs may indicate a higher risk of recurrence. The COX-2 is upregulated in recurrent sporadic OKCs, whereas recurrent syndromic OKCs exhibit higher RANKL and lower OPG expression; however, these findings have no prognostic relevance. The immunoexpression of p53, PCNA and OPG may help to distinguish syndromic from sporadic OKCs.

**Key words:** oral surgery, odontogenic cyst, keratocystic odontogenic tumor, odontogenic keratocysts

## Background

Odontogenic keratocyst (OKC) is a benign intraosseous lesion of the jaw with a tendency toward aggressive growth and a relatively high recurrence rate. There is no agreement among authors as to whether OKC should be considered a cyst or a neoplasm.<sup>1,2</sup> Although it has been recently reclassified as a cyst,<sup>3</sup> for over a decade OKC was defined by the World Health Organization (WHO) as an intraosseous tumor due to *PTCH* gene mutations and infiltrative growth.<sup>4</sup> The WHO consensus panel does not necessarily state that OKC is not neoplastic, but rather that currently there is a lack of evidence to justify its classification as a tumor.<sup>3,5,6</sup> These discrepancies may lead to confusion among clinicians, particularly since the lesion is the 3<sup>rd</sup> most common cyst of the jaw.<sup>3,7</sup> The results of a recent meta-analysis advocate the concept that OKC is prone to behave as an odontogenic tumor in view of its p53 expression.<sup>7</sup>

As many as 5–6.4% of all OKCs occur as part of the nevoid basal cell carcinoma syndrome (NBCCS).<sup>8</sup> The NBCCS is an autosomal dominant inherited disorder related to a germline mutation of the patched gene 1 (*PTCH1*), whose alteration results in a carcinogenic process activated by an altered cell cycle and cellular proliferation.<sup>9</sup> The OKC is one of the most common signs of NBCCS, along with basal cell carcinomas (BCCs), palmar pits, skeletal abnormalities, and calcified falx cerebri. Syndromic cases of OKC tend to be multiple and occur in younger patients.<sup>3</sup> It has been suggested that syndromic OKCs have more aggressive clinical behavior and a higher tendency to relapse.<sup>10,11</sup> An early recognition of NBCCS is of vital importance, as affected patients are prone to develop neoplasms, such as BCC, medulloblastoma and ovarian fibroma.<sup>12</sup>

As genetic confirmation of NBCCS is not routinely conducted due to its high cost, the diagnosis is currently based on a combination of major and minor clinical criteria.<sup>13,14</sup> Researchers have tried to elucidate whether immunohistochemistry may allow them to distinguish between sporadic and syndromic OKC, but their results are inconclusive. A recent meta-analysis determined that several markers of epithelial cell proliferation and apoptosis are indistinguishable between sporadic and syndromic OKCs; however, the quality assessment of the included studies revealed notable inconsistencies regarding the diagnostic criteria, including demographic and clinical characteristics, as well as the histopathological description of OKC.<sup>12</sup> On the other hand, various studies have demonstrated that immunohistochemistry is a valuable method to assess the biological profile and prognosis of OKC, as several markers, including Ki-67,<sup>1</sup> Bcl-2,<sup>15</sup> cyclin D1,<sup>15</sup> p53,<sup>15</sup> and proliferating cell nuclear antigen (PCNA),<sup>15,16</sup> may be consistent with the local aggressiveness and propensity for the recurrence of sporadic OKC cases. However, data on such correlations with regard to syndromic OKCs are scarce.

An association with NBCCS is not the only reason believed to be responsible for the higher recurrence of OKC, as some clinicoradiological features have also been shown to be negative prognostic factors of recurrence. In our previous reports, we demonstrated that larger size, multilocularity and cortical perforation may be related to a relapse in sporadic cases of OKCs. We did not, however, reveal any prognostic significance of immunoexpression of Bcl-2, cyclin D1, p53, osteoprotegerin (OPG), PCNA, receptor activator of nuclear factor  $\kappa$  B (RANK), and receptor activator of nuclear factor  $\kappa$  B ligand (RANKL) for OKCs not associated with NBCCS.<sup>17,18</sup> Hitherto, little is known about whether clinicoradiological and pathological features or expression of immunohistochemical markers are also associated with recurrence in cases of syndromic OKCs.

## Objectives

We aimed to compare the prognostic relevance of clinicoradiological and histopathological features, as well as the immunoexpression of COX-2, Bcl-2, PCNA, p53, Ki-67, OPG, RANK, and RANKL, and RANKL/OPG balance, between sporadic and syndromic OKCs. The secondary objective was to test whether immunoexpression of the aforementioned proteins is useful in distinguishing between the 2 types of OKC.

## Materials and methods

### Study design and patients

Forty-three cases of OKC who were not lost to follow-up were selected for this retrospective study. Patients were treated with simple enucleation in the Chair of Oral Surgery at the Jagiellonian University Medical College, Kraków, Poland, between 1997 and 2015. The surgical technique did not include any adjunct procedures (e.g., Carnoy's solution, liquid nitrogen, peripheral ostectomy, or application of regenerative graft materials) and was standardized among board-certified specialists in oral surgery with at least 5 years of practical experience. Among the 43 subjects, 31 were diagnosed with sporadic OKC and 12 with syndromic OKC. The NBCCS cases were diagnosed using the features described by Kimonis et al.<sup>13</sup> and were subsequently found to be in conformity with the Consensus Statement from the First International Colloquium on Basal Cell Nevus Syndrome.<sup>14</sup> All cases were confirmed to meet the current WHO criteria for OKC (i.e., presence of a fibrous wall lined by a folded, thin, regular parakeratinized epithelium 5–8-cell layers thick, without rete ridges, with corrugation of the parakeratin surface and a well-defined, palisaded basal layer, hyperchromatic nuclei, and focal areas showing reversed nuclear polarity).<sup>3</sup> Ethical approval was obtained from



the Bioethical Committee of Jagiellonian University (approval No. 1072.6120.73.2019).

To evaluate the potential clinicoradiological, histopathological and immunohistochemical prognostic factors, we employed the methods used in our previous studies.<sup>17,18</sup> At the time of diagnosis, panoramic and occlusal radiographs were taken, supplemented in most cases with computed tomography (CT). Follow-up appointments took place quarterly in the 1<sup>st</sup> postoperative year, and semi-annually starting from the 2<sup>nd</sup> postoperative year. Upon appointments, panoramic radiographs were performed. In the event of any signs of recurrence (e.g., expansion of bone and/or radiographic radiolucent zone in the area of the previous surgery), CT scans were performed. The interval between surgery and the detection of relapse was defined as the recurrence period. Each relapse was confirmed to meet the 2017 WHO microscopic criteria for OKC.<sup>3</sup>

The following clinicoradiological features were measured: age, gender, follow-up period, evidence of recurrence, size of the cyst in panoramic radiographs (determined by multiplying the major and minor axes), outline of the cyst on radiography (uni- or multilocular), and cortical perforation observed on occlusal radiographs or CT. The following histopathological features were investigated: number of satellite (daughter) cysts, inflammation intensity and presence of epithelial dysplasia. The position of the cyst was divided into anterior maxilla, posterior maxilla, anterior mandible, and posterior mandible. The posterior surface of the second premolar served as the division line between the anterior and posterior parts of the jaw.

## Specimen characteristics and assay methods

Formalin-fixed, paraffin-embedded archival blocks were sectioned (5 µm in thickness) and stained with hematoxylin and eosin (H&E). Slides were used to corroborate a diagnosis of OKC and to appraise its histopathological features using a light microscope (Olympus BX40; Olympus Corp., Tokyo, Japan). The inflammatory score was calculated by counting the inflammatory cells and categorized with the 4-grade scoring system according to Kuroyanagi et al.:<sup>1</sup> grade 0 – no inflammation, grade 1 – fewer than 15 cells, grade 2 – 15–50 cells, and grade 3 – more than 50 cells in 10 high-power fields (HPFs).

For immunohistochemical analysis, 3-µm thick tissue sections were used. They were deparaffinized with xylene, rehydrated in graded alcohol and washed in deionized water. Retrieval of antigen was accomplished with Heat-Induced Epitope Retrieval Buffer (Thermo Fisher Scientific, Fremont, USA) at pH 6 or pH 9 for 20 min at 95°C. Subsequently, sections were blocked by incubation with 3% H<sub>2</sub>O<sub>2</sub> and protein block (Thermo Fisher Scientific), followed by the incubation of the slides in a humidified chamber with one of the following antibodies:

- rabbit monoclonal COX-2 (SP21; Cell Marque, Rocklin, USA; 1:200) at room temperature for 45 min;
- rabbit monoclonal Bcl-2 (E17; Cell Marque; 1:300) at 4°C overnight;
- mouse monoclonal p53 (DO7; Cell Marque; 1:100) at room temperature for 50 min;
- rabbit polyclonal PCNA (RB-9055-P0; Thermo Fisher Scientific; 1:700) at room temperature for 30 min;
- mouse monoclonal Ki-67 (SP6; Cell Marque; 1:150) at room temperature for 30 min;
- mouse monoclonal anti-RANK (ab13918; Abcam, Waltham, USA; 1:200) at 4°C overnight;
- rabbit polyclonal anti-RANKL (ab169966; Abcam; 1:400) at 4°C overnight;
- rabbit polyclonal anti-OPG (ab183910; Abcam; 1:400) at 4°C overnight.

Next, the sections were washed in Tris-buffered saline (TBS) and treated in compliance with the manufacturer's instructions using the Primary Antibody Amplifier Quanto system, followed by the HRP Polymer Quanto system (both from Thermo Fisher Scientific). Then, they were stained using a 3-3'-diaminobenzidine Quanto kit (Thermo Fisher Scientific). Finally, the sections were counterstained with hematoxylin, dehydrated and coverslipped for further analyses. With regard to Bcl-2, p53 and PCNA, tonsil tissue was used as a positive control. For RANK, RANKL and OPG, a case of central giant cell granuloma was used as a positive control. For the negative control, the section treatment was similar, but there was no primary antibody exposure. For COX-2, RANK, RANKL, and OPG, a cytoplasmic staining pattern was detected. For Bcl-2, a membranous-cytoplasmic pattern was detected, and for p53, Ki-67 and PCNA, a nuclear pattern was detected.

For COX-2, Bcl-2, RANK, RANKL, and OPG, a semi-quantitative analysis with a 4-grade scoring system was employed, as follows: grade 0 – negative reaction (no stained cells), grade 1 – staining of 1–25% of cells, grade 2 – staining of 26–50% of cells, and grade 3 – more than 50% of stained cells in the epithelium. For PCNA, Ki-67 and p53, a quantitative assessment was applied. The number of positive-staining nuclear cells was counted in 10 HPFs (×400 magnification), and the mean value was used for the analysis. All histopathological and immunohistochemical analyses were completed by a board-certified specialist in pathomorphology (B. D.).

## Statistical analyses

The clinicopathological and immunohistochemical data sets were evaluated against the frequency of relapse. The results were reported as mean ± standard deviation (M ±SD) and median (1<sup>st</sup> quartile–3<sup>rd</sup> quartile (Q1–Q3)). In the case of qualitative data, the χ<sup>2</sup> test or Fisher's exact test (when the χ<sup>2</sup> test assumption of large expected value was not met) were employed to analyze the differences in clinicoradiological and pathological data sets. With

regard to quantitative data, the Mann–Whitney U test was used. Nonparametric tests were used because of deviations from normality in quantitative variable distribution, as was shown by the results of the Shapiro–Wilk test ( $p < 0.05$ ). To evaluate the hazard ratios (HRs) and 95% confidence intervals (95% CIs) as estimates of risk for recurrence potential, the Cox proportional hazard model for time-dependent variables was utilized. The  $p$ -values less than 0.05 were considered statistically significant. All analyses were performed using R software v. 4.1.3<sup>19</sup> (R Foundation for Statistical Computing, Vienna, Austria; <https://www.r-project.org/>) with the “survival” package.<sup>20,21</sup>

## Results

The follow-up period for all analyzed OKCs was  $104.09 \pm 46.97$  months ( $110.71 \pm 52.28$  months and  $87 \pm 22.95$  months for sporadic and syndromic lesions, respectively;  $U = 234$ ,  $p = 0.123$ ; Mann–Whitney U test). The recurrence period for the entire analyzed group was  $77.14 \pm 31.88$  months ( $69.82 \pm 36.72$  months and  $85.2 \pm 24.95$  months for sporadic and syndromic lesions, respectively;  $U = 43.5$ ,  $p = 0.434$ ; Mann–Whitney U test). Of all the OKC cases, 48.84% recurred; however, syndromic cases were significantly more prone to recur than sporadic cases (83.33% compared to 35.48%;  $p = 0.013$ ; Fisher’s exact test). An association with NBCCS significantly increased the risk of OKC recurrence (HR: 9.091, 95% CI: 1.682–49.123;  $p = 0.01$ ).

There were no gender differences between sporadic and syndromic OKCs, but the patients with syndromic OKCs were significantly younger than those with sporadic lesions (Table 1). Also, there were no significant differences between sporadic and syndromic OKCs in regard to radiological size of the lesion, locularity, cortical perforation, histological number of daughter cysts, inflammatory score, and epithelial dysplasia grading (Table 1). However, there were significant differences in PCNA, p53 and OPG immunorexpression between sporadic and syndromic lesions (Table 2). Details of immunorexpression of PCNA, p53 and OPG are shown in Fig. 1–3.

The clinicoradiological, histopathological and immunohistochemical characteristics of sporadic cases are presented in Table 2. Recurrent sporadic OKCs were significantly larger ( $p = 0.002$ ) and markedly more often associated with cortical perforation ( $p = 0.038$ ). In fact, sporadic lesions with radiological signs of cortical perforation were 6.8 times more likely to recur. With regard to immunostaining, recurrent OKCs had significantly increased the immunorexpression of COX-2. However, by using the Cox proportional hazard model, we did not find COX-2 significance to be a prognostic marker (Table 2). Details of COX-2 immunorexpression are shown in Fig. 4. Because of the fact that all of the non-recurrent sporadic OKCs were negative for epithelial dysplasia, the HR for this group was indefinite (Table 2).

The clinicoradiological, histopathological and immunohistochemical characteristics of the syndromic cases are presented in Table 3. Of the 12 syndromic cases, only 2 did not relapse. Among the analyzed factors, we did not find any significant differences between non-recurrent and recurrent lesions, with the exception of RANKL/OPG ratio, where RANKL predominance was particularly marked in recurrent OKCs ( $p = 0.015$ ). Details of RANKL immunorexpression are shown in Fig. 5. Due to the small number of non-recurrent cases and the lack of non-recurrent cases for some groups, the HRs for several clinicoradiological (gender, cyst location and cortical perforation), histopathological (epithelial dysplasia) and immunohistochemical (RANK, RANKL, RANKL/OPG ratio) features were indefinite (Table 3). The complete characteristics of the recurrent sporadic and syndromic cases are presented in Table 4,5. Details of immunorexpression of Bcl-2, Ki-67 and RANK are shown in Fig. 6–8.

## Discussion

The OKC is the main feature observed in patients with NBCCS. The possible reasons for its high recurrence rate have been a subject of debate for several decades. Authors have linked the recurrence rate to treatment methods<sup>22,23</sup> and association with NBCCS.<sup>24</sup> Recent evidence has attributed OKC relapse to some clinicoradiological features, such as association with dentition,<sup>25</sup> larger size, multilocularity, and cortical perforation,<sup>17</sup> or to some pathological features, mainly the presence of daughter cysts.<sup>26</sup> Moreover, the expression of various markers have been tested against the recurrence rate, and a few have proved to be significantly associated with cyst relapse.<sup>1,15,16</sup> In the present study, we endeavored to compare the clinicopathological features of sporadic and syndromic cases of OKC, and test whether the expression of some previously considered markers was related to recurrence.

Most authors agree that OKCs associated with NBCCS generally exhibit greater aggressiveness and higher propensity for recurrence.<sup>27</sup> The results of several studies support the existence of distinct biological behaviors of syndromic and sporadic OKCs.<sup>28–32</sup> However, only Titinchi and Nortje<sup>24</sup> demonstrated that there is a statistically significant higher recurrence rate in syndromic cases of OKC.<sup>24</sup> Our results indicate that there is a strong association between NBCCS and OKC recurrence. In fact, in the current study, the association of OKC with NBCCS increased the risk of relapse by 9.091 times. However, this observation should be considered with caution, as OKCs in patients affected with NBCCS are typically multifocal and tend to appear metachronously. Thus, it may be difficult to unequivocally determine whether the new lesion is, in fact, a recurrence of a previously treated cyst or a new development. Other authors have also pointed out that it may be difficult to correctly estimate the recurrence rate of syndromic cases of OKC

**Table 1.** Comparative clinicoradiological, histopathological and immunohistochemical characteristics of the sporadic and syndromic cases of OKC

Variables		Sporadic OKC (n = 31)	Syndromic OKC (n = 12)	Total (n = 43)	p-value
Gender	female	13 (41.94%)	5 (41.67%)	18 (41.86%)	$\chi^2 = 0.000, p = 1^{***}$
	male	18 (58.06%)	7 (58.33%)	25 (58.14%)	
Age at the time of diagnosis [years]	M ±SD	40.23 ±17.18	18.42 ±5.79	34.14 ±17.82	U = 329, <b>p &lt; 0.001*</b>
	median	39	17	30	
	Q1–Q3	26.0–51.5	13.5–23.0	19–51	
Cyst location	posterior maxilla	2 (6.45%)	5 (41.67%)	7 (16.28%)	p = 0.053**
	anterior maxilla	2 (6.45%)	0 (0.00%)	2 (4.65%)	
	posterior mandible	20 (64.52%)	5 (41.67%)	25 (58.14%)	
	anterior mandible	7 (22.58%)	2 (16.67%)	9 (20.93%)	
Size of lesion [mm <sup>2</sup> ]	M ±SD	504.06 ±276.37	487.08 ±282.23	499.33 ±274.74	U = 182.5, p = 0.935*
	median	396	397.5	396	
	Q1–Q3	315–647	345–531.25	315–594	
Lesion type	unilocular	21 (67.74%)	7 (58.33%)	28 (65.12%)	p = 0.723**
	multilocular	10 (32.26%)	5 (41.67%)	15 (34.88%)	
Cortical perforation	negative	22 (70.97%)	6 (50.00%)	28 (65.12%)	p = 0.287**
	positive	9 (29.03%)	6 (50.00%)	15 (34.88%)	
Number of daughter cysts	M ±SD	0.77 ±1.57	7.08 ±9.89	2.57 ±6.02	U = 129.5, p = 0.081*
	median	0	0	0	
	Q1–Q3	0–0	0–11.5	0–1.75	
Inflammatory score	grade 0	14 (45.16%)	2 (16.67%)	16 (37.21%)	p = 0.084**
	grade 1	9 (29.03%)	3 (25.00%)	12 (27.91%)	
	grade 2	7 (22.58%)	6 (50.00%)	13 (30.23%)	
	grade 3	1 (3.23%)	1 (8.33%)	2 (4.65%)	
Epithelial dysplasia	negative	29 (93.55%)	9 (75.00%)	38 (88.37%)	p = 0.063**
	low-grade	1 (3.23%)	3 (25.00%)	4 (9.30%)	
	high-grade	1 (3.23%)	0 (0.00%)	1 (2.33%)	
PCNA	M ±SD	19.07 ±8.26	27.17 ±10.91	21.38 ±9.69	U = 97, <b>p = 0.022*</b>
	median	18.5	30.5	19.75	
	Q1–Q3	13.62–24	19.12–34	14–27.25	
p53	M ±SD	11.8 ±6.59	17.08 ±5.93	13.31 ±6.78	U = 98.5, <b>p = 0.024*</b>
	median	11.25	17	12.25	
	Q1–Q3	7.0–17.5	14.12–19	9–18	
Ki-67	M ±SD	13.92 ±6.69	13.2 ±3.7	13.71 ±5.96	U = 175, p = 0.9*
	median	12	12.75	12.25	
	Q1–Q3	10.50–17.38	11.25–14.25	10.50–17.38	
Bcl-2	grade 0	10 (32.25%)	0 (0.00%)	10 (23.25%)	p = 0.215**
	grade 1	15 (48.39%)	9 (75.00%)	24 (55.81%)	
	grade 2	5 (16.13%)	3 (25.00%)	8 (18.60%)	
	grade 3	1 (3.23%)	0 (0.00%)	1 (2.33%)	
COX-2	grade 0	6 (19.35%)	0 (0.00%)	6 (13.95%)	p = 0.26**
	grade 1	5 (16.13%)	5 (41.67%)	10 (23.25%)	
	grade 2	10 (32.26%)	2 (16.67%)	12 (27.91%)	
	grade 3	10 (32.26%)	5 (41.67%)	15 (34.88%)	
RANK	grade 0	17 (54.83%)	6 (50.00%)	23 (53.48%)	p = 0.694**
	grade 1	12 (38.71%)	4 (33.33%)	16 (37.21%)	
	grade 2	2 (6.45%)	2 (16.67%)	4 (9.30%)	
	grade 3	0 (0.00%)	0 (0.00%)	0 (0.00%)	
RANKL	grade 0	1 (3.23%)	0 (0.00%)	1 (2.33%)	p = 0.438**
	grade 1	7 (22.58%)	1 (8.33%)	8 (18.60%)	
	grade 2	9 (29.03%)	3 (25.00%)	12 (27.91%)	
	grade 3	14 (45.16%)	8 (66.67%)	22 (51.16%)	
OPG	grade 0	3 (9.67%)	1 (8.33%)	4 (9.30%)	<b>p = 0.001**</b>
	grade 1	19 (61.29%)	1 (8.33%)	20 (46.51%)	
	grade 2	3 (9.67%)	6 (50.00%)	9 (20.93%)	
	grade 3	6 (19.35%)	4 (33.33%)	10 (23.25%)	
RANKL/OPG ratio	RANKL < OPG	1 (3.23%)	2 (16.67%)	3 (6.98%)	p = 0.367**
	RANKL = OPG	14 (45.16%)	4 (33.33%)	18 (41.86%)	
	RANKL > OPG	16 (51.61%)	6 (50.00%)	22 (51.16%)	

\*Mann–Whitney U test, \*\*Fisher's exact test, \*\*\* $\chi^2$  test. Statistically significant values are bolded. Q1 – lower quartile; Q3 – upper quartile; M ±SD – mean ± standard deviation; OKC – odontogenic keratocyst; RANKL – receptor activator for nuclear factor  $\kappa$  B ligand; OPG – osteoprotegerin; PCNA – proliferating cell nuclear antigen; RANK – receptor activator of nuclear factor  $\kappa$  B.

**Table 2.** Comparative clinicoradiological, histopathological and immunohistochemical characteristics of non-recurrent and recurrent sporadic OKC cases with risk for recurrence

Variables		Non-recurrent sporadic OKC (n = 20)	Recurrent sporadic OKC (n = 11)	p-value	Cox proportional hazard model	
					HR (95% CI)	p-value
Gender	female (n = 13)	9 (45.00%)	4 (36.36%)	p = 0.718**	1.00	0.642
	male (n = 18)	11 (55.00%)	7 (63.64%)		1.432 (0.316–6.492)	
Age [years]	M ±SD	40.5 ±16.02	39.73 ±19.94	U = 117, p = 0.788*	0.997 (0.955–1.042)	0.903
	median	46	36			
	Q1–Q3	28.5–51.25	26–55			
Cyst location	posterior maxilla (n = 2)	1 (5.00%)	1 (9.09%)	p = 0.923**	1.00	–
	anterior maxilla (n = 2)	2 (10.00%)	0 (0.00%)		NR	–
	posterior mandible (n = 20)	13 (65.00%)	7 (63.64%)		1.615 (0.14–18.581)	0.7
	anterior mandible (n = 7)	4 (20.00%)	3 (27.27%)		2.25 (0.149–33.933)	0.558
Size of lesion [mm <sup>2</sup> ]	M ±SD	370.75 ±130.55	746.45 ±310.79	U = 34.5, <b>p = 0.002*</b>	1.007 (1.002–1.012)	<b>0.005</b>
	median	367.5	760			
	Q1–Q3	270–420	522–987.5			
Lesion type	unilocular (n = 21)	15 (75.00%)	6 (54.55%)	p = 0.423**	1.00	0.25
	multilocular (n = 10)	5 (25.00%)	5 (45.45%)		2.5 (0.525–11.894)	
Cortical perforation	negative (n = 22)	17 (85.00%)	5 (45.45%)	<b>p = 0.038**</b>	1.00	<b>0.028</b>
	positive (n = 9)	3 (15.00%)	6 (54.55%)		6.8 (1.233–37.497)	
Number of daughter cysts	M ±SD	1 ±1.73	0.36 ±1.21	U = 127, p = 0.201*	0.721 (0.387–1.341)	0.301
	median	0	0			
	Q1–Q3	0–1.5	0–0			
Inflammatory score	grade 0 (n = 15)	10 (50.00%)	5 (45.45%)	p = 1**	1.00	–
	grade 1 (n = 9)	6 (30.00%)	3 (27.27%)		0.9 (0.154–5.258)	0.907
	grade 2 (n = 7)	4 (20.00%)	3 (27.27%)		1.35 (0.211–8.617)	0.751
	grade 3 (n = 0)	0 (0.00%)	0 (0.00%)		NR	–
Epithelial dysplasia	negative (n = 30)	20 (100.00%)	10 (90.91%)	p = 0.367**	indefinite	
	low-grade (n = 1)	0 (0.00%)	1 (9.09%)			
	high-grade (n = 0)	0 (0.00%)	0 (0.00%)			
PCNA	M ±SD	18.55 ±9.4	19.95 ±6.14	U = 89, p = 0.518*	1.021 (0.932–1.12)	0.65
	median	18	23			
	Q1–Q3	13.5–22.75	13.75–24.5			
p53	M ±SD	11.08 ±7	13.05 ±5.92	U = 86.5, p = 0.451*	1.048 (0.933–1.178)	0.427
	median	10.5	11.5			
	Q1–Q3	6–17	9.5–15.5			
Ki-67	M ±SD	12.61 ±5.23	16.18 ±8.48	U = 83, p = 0.365*	1.09 (0.961–1.237)	0.179
	median	12	16.5			
	Q1–Q3	10–15.75	10.5–18.5			
Bcl-2	grade 0 (n = 10)	8 (40.00%)	2 (18.18%)	p = 0.362**	1.00	–
	grade 1 (n = 15)	9 (45.00%)	6 (50.00%)		4 (0.379–42.177)	–
	grade 2 (n = 5)	3 (15.00%)	2 (18.18%)		6 (0.422–85.248)	0.249
	grade 3 (n = 1)	0 (0.00%)	1 (9.09%)		6 (0.422–85.248)	0.186
COX-2	grade 0 (n = 6)	5 (25.00%)	1 (9.09%)	<b>p = 0.042**</b>	1.00	–
	grade 1 (n = 5)	5 (25.00%)	0 (0.00%)		NR	–
	grade 2 (n = 10)	7 (35.00%)	3 (27.27%)		0.857 (0.055–13.479)	0.913
	grade 3 (n = 10)	3 (15.00%)	7 (63.64%)		4.667 (0.297–73.384)	0.273
RANK	grade 0 (n = 17)	13 (65.00%)	4 (36.36%)	p = 0.126**	1.00	–
	grade 1 (n = 12)	7 (35.00%)	5 (45.45%)		3 (0.642–14.023)	0.163
	grade 2 (n = 2)	0 (0.00%)	2 (18.18%)		3 (0.642–14.023)	0.163
	grade 3 (n = 0)	0 (0.00%)	0 (0.00%)		NR	–
RANKL	grade 0 (n = 1)	1 (5.00%)	0 (0.00%)	p = 0.304**	NR	–
	grade 1 (n = 7)	6 (30.00%)	1 (9.09%)		1.00	–
	grade 2 (n = 9)	6 (30.00%)	3 (27.27%)		3 (0.239–37.672)	0.395
	grade 3 (n = 14)	7 (35.00%)	7 (63.64%)		6 (0.565–63.676)	0.137
OPG	grade 0 (n = 3)	2 (10.00%)	1 (9.09%)	p = 0.635**	1.00	–
	grade 1 (n = 19)	13 (65.00%)	6 (54.55%)		1.00	–
	grade 2 (n = 3)	2 (10.00%)	1 (9.09%)		0.929 (0.071–12.136)	0.955
	grade 3 (n = 6)	3 (15.00%)	3 (27.27%)		1.857 (0.293–11.756)	0.511
RANKL/OPG ratio	RANKL < OPG	2 (10.00%)	1 (9.09%)	p = 0.24**	1.00	0.476
	RANKL = OPG	9 (45.00%)	3 (27.27%)		1.00	
	RANKL > OPG	9 (45.00%)	7 (63.64%)		1.75 (0.376–8.14)	

\*Mann–Whitney U test, \*\*Fisher's exact test. Statistically significant values are bolded. Q1 – lower quartile; Q3 – upper quartile; M ±SD – mean ± standard deviation; NR – no recurrence; HR – hazard ratio; 95% CI – 95% confidence interval; OKC – odontogenic keratocyst; RANKL – receptor activator for nuclear factor κ B ligand; OPG – osteoprotegerin; PCNA – proliferating cell nuclear antigen; RANK – receptor activator of nuclear factor κ B.

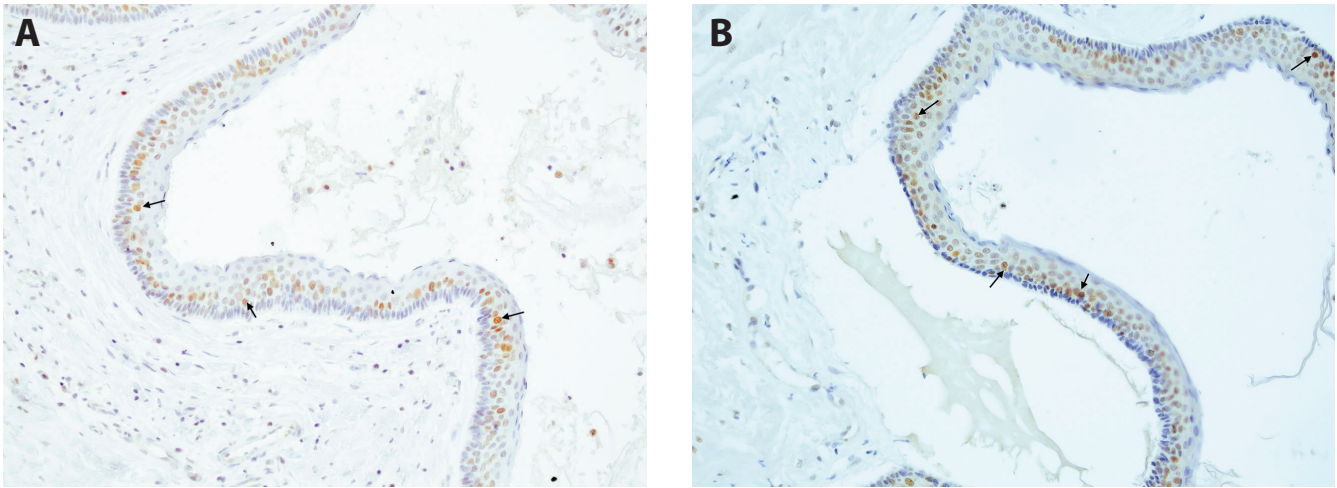


Fig. 1. Proliferating cell nuclear antigen (PCNA) – nuclear brown staining (arrows) in the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x200 magnification)

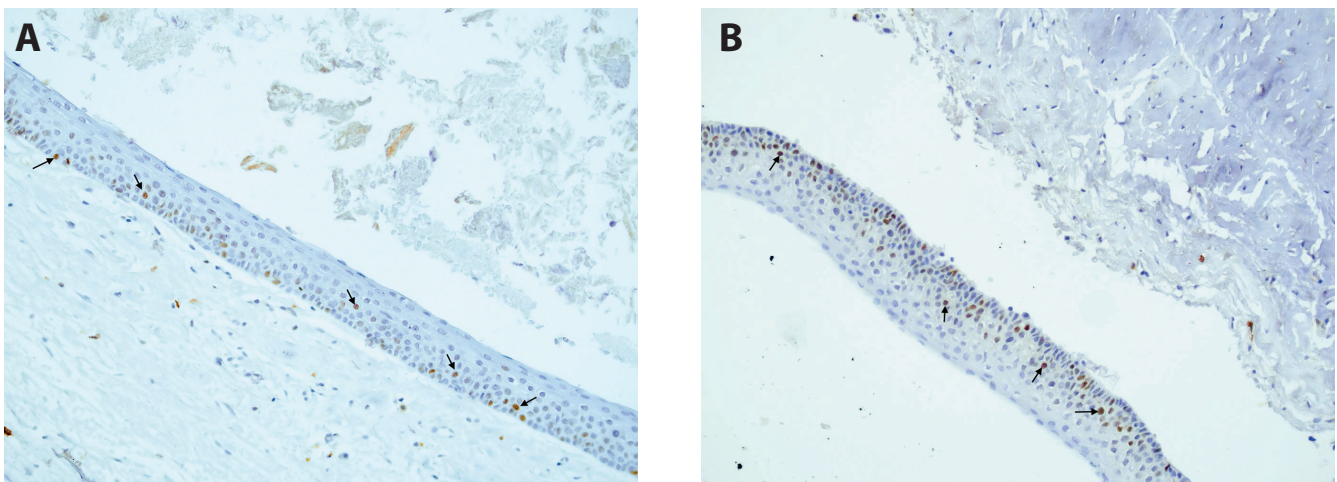


Fig. 2. p53 – nuclear brown staining (arrows) in the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x200 magnification)

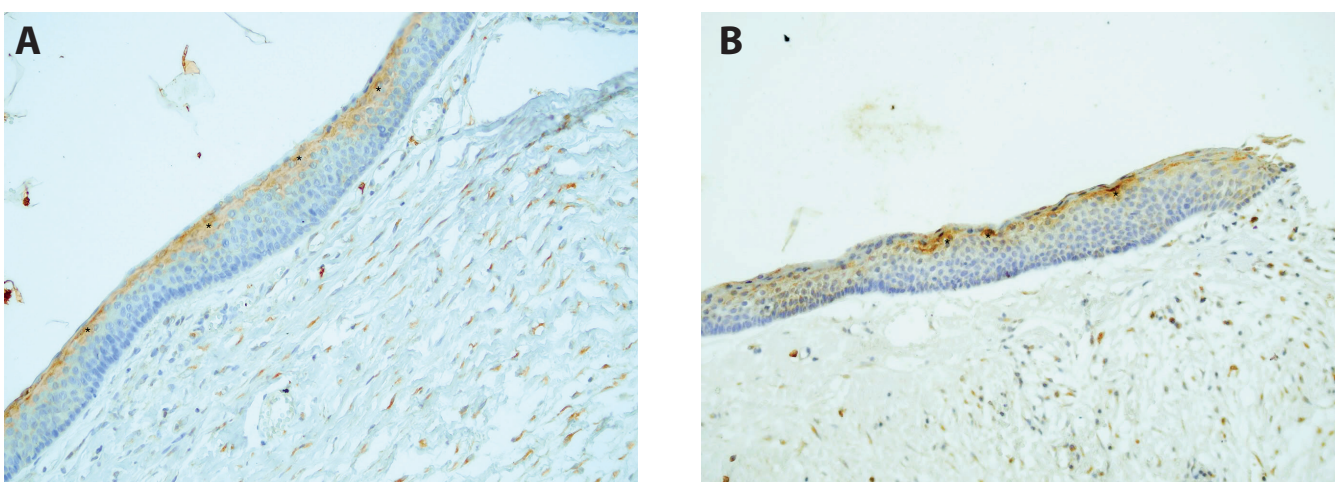


Fig. 3. Osteoprotegerin (OPG) – cytoplasmic brown staining in the superficial layer (asterisks) of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x200 magnification)

for the above reasons.<sup>24,33,34</sup> Accordingly, it would appear reasonable to conduct further research into the molecular background of the distinct pathobiology of syndromic OKCs.

In our previous report, we demonstrated that larger size, multilocularity and radiological evidence of cortical perforation were correlated to a higher risk of recurrence

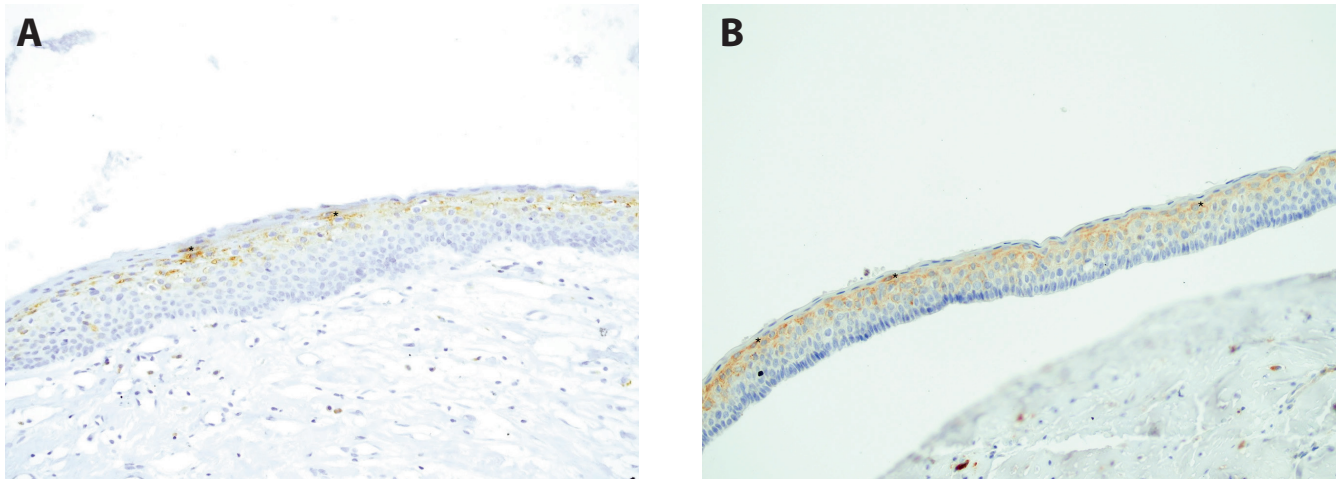


Fig. 4. COX-2 – cytoplasmic brown staining in the superficial layer (asterisks) of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x200 magnification)

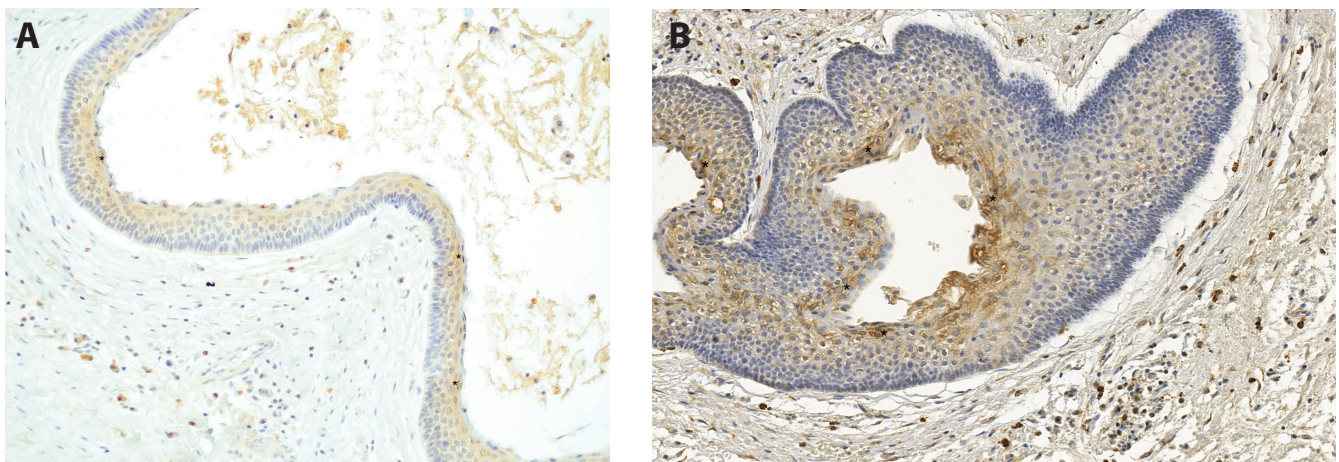


Fig. 5. Receptor activator for nuclear factor  $\kappa$  B ligand (RANKL) – cytoplasmic brown staining in the superficial layer (asterisks) of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x50 magnification)

of sporadic OKCs.<sup>17</sup> The current results partially support these findings. The issue of cyst size and multilocularity have been addressed by many authors, and it has been shown that smaller lesions are usually unilocular, but with their expansion, they tend to become multilocular.<sup>34,35</sup> Larger lesions are more difficult to access, and epithelial residue may be a starting point of relapse. Recently, Fidele et al. suggested that the lining of larger OKCs is more thin and fragile and, thus, more prone to recur.<sup>36</sup> This is in contrast with smaller lesions, which are easier to enucleate in one piece. The authors also found a significantly higher recurrence rate for cysts whose diameter exceeded 4 cm.<sup>36</sup> With regard to syndromic cases, however, we did not find a significant influence of cyst size on recurrence rate, which may be due to the limited number of syndromic cases in the current series, as well as the fact that there might be disparate biological factors responsible for the relapse of NBCCS-associated counterparts of OKCs.

In the current study, sporadic OKCs with cortical perforation were 6.8 times more likely to recur. Our previous hypothesis<sup>17</sup> stating that this radiological sign

is a significant prognostic factor has thus far been supported by the results of the study by Fidele et al.,<sup>36</sup> who found that cortical perforation, particularly combined with teeth involvement, is markedly associated with a high recurrence rate. The presence of either may lead to an incomplete surgical removal and subsequent relapse, since cortical perforation suggests the involvement of adjacent soft tissues, and close proximity to dentition (especially when the treatment plan entails teeth preservation) may be related to the involvement of the periodontal space. However, in the current study, we did not find significant differences in cortical perforation between recurrent and non-recurrent syndromic OKCs. In fact, all of the non-recurrent syndromic cases were negative for cortical perforation, which is why the HR for recurrence was indefinite. Interestingly, 50% of the syndromic cases exhibited cortical perforation (in contrast to 29.03% of the sporadic cases), which may be another reason for the high recurrence rate of NBCCS-associated cases. Contrary to the results of our previous report,<sup>17</sup> we did not demonstrate that multilocularity is associated with higher recurrence rates

**Table 3.** Comparative clinicoradiological, histopathological and immunohistochemical characteristics of non-recurrent and recurrent syndromic OKC cases with a risk for recurrence

Variables		Non-recurrent syndromic OKC (n = 2)	Recurrent syndromic OKC (n = 10)	p-value	Cox proportional hazard model	
					HR (95% CI)	p-value
Gender	female (n = 5)	0 (0.00%)	5 (50.00%)	p = 0.47**	indefinite	
	male (n = 7)	2 (100.0%)	5 (50.00%)			
Age [years]	M ±SD	17 ±0	18.7 ±6.36	U = 9, p = 0.913*	1.062 (0.787–1.432)	0.695
	median	17	18			
	Q1–Q3	17–17	12.5–23.0			
Cyst location	posterior maxilla (n = 5)	0 (0.00%)	5 (50.00%)	p = 0.621**	indefinite	
	anterior maxilla (n = 0)	0 (0.00%)	0 (0.00%)			
	posterior mandible (n = 5)	2 (100.0%)	3 (30.00%)			
	anterior mandible (n = 12)	0 (0.00%)	2 (20.00%)			
Size of lesion [mm <sup>2</sup> ]	M ±SD	397.5 ±31.82	505 ±308.38	U = 9.5, p = 1*	1.003 (0.992–1.014)	0.637
	median	397.5	412.5			
	Q1–Q3	386.25–408.75	315–543.75			
Lesion type	unilocular (n = 7)	1 (50.00%)	6 (60.00%)	p = 1**	1.00 0.667 (0.032–14.033)	0.794
	multilocular (n = 5)	1 (50.00%)	4 (40.00%)			
Cortical perforation	negative (n = 6)	2 (100.0%)	4 (40.00%)	p = 0.455**	indefinite	
	positive (n = 6)	0 (0.00%)	6 (60.00%)			
Number of daughter cysts	M ±SD	12.5 ±17.68	6 ±8.77	U = 13, p = 0.548*	0.936 (0.805–1.09)	0.395
	median	12.5	0			
	Q1–Q3	6.25–18.75	0–9			
Inflammatory score	grade 0 (n = 2)	0 (0.00%)	2 (20.00%)	p = 1**	1.00 1.00 1.5 (0.071–31.575) 1.5 (0.071–31.575)	0.794
	grade 1 (n = 3)	1 (50.00%)	2 (20.00%)			
	grade 2 (n = 6)	1 (50.00%)	5 (50.00%)			
	grade 3 (n = 0)	0 (0.00%)	1 (10.00%)			
		0 (0.00%)	0 (0.00%)			
Epithelial dysplasia	negative (n = 9)	2 (100.0%)	7 (70.00%)	p = 1**	indefinite	
	low-grade (n = 3)	0 (0.00%)	3 (30.00%)			
	high-grade (n = 0)	0 (0.00%)	0 (0.00%)			
PCNA	M ±SD	26.5 ±9.9	27.3 ±11.6	U = 9.5, p = 1*	1.007 (0.872–1.164)	0.921
	median	26.5	30.5			
	Q1–Q3	23–30	18.75–35.00			
p53	M ±SD	16.25 ±2.47	17.25 ±6.49	U = 10.5, p = 1*	1.032 (0.784–1.36)	0.821
	median	16.25	17			
	Q1–Q3	15.38–17.12	13.38–20.75			
Ki-67	M ±SD	10.75 ±1.77	13.69 ±3.84	U = 4, p = 0.234*	1.38 (0.73–2.608)	0.321
	median	10.75	13			
	Q1–Q3	10.12–11.38	11.75–16.75			
Bcl-2	grade 0 (n = 0)	0 (0.00%)	0 (0.00%)	p = 0.455**	NR 1.00 0.25 (0.01–5.985) NR	0.392
	grade 1 (n = 9)	1 (50.00%)	8 (80.00%)			
	grade 2 (n = 3)	1 (50.00%)	2 (20.00%)			
	grade 3 (n = 0)	0 (0.00%)	0 (0.00%)			
COX-2	grade 0 (n = 0)	0 (0.00%)	0 (0.00%)	p = 0.621**	NR 1.00 1.5 (0.071–31.575) 1.5 (0.071–31.575)	0.794
	grade 1 (n = 5)	1 (50.00%)	4 (40.00%)			
	grade 2 (n = 2)	1 (50.00%)	1 (10.00%)			
	grade 3 (n = 5)	0 (0.00%)	5 (50.00%)			
RANK	grade 0 (n = 6)	2 (100.0%)	4 (40.00%)	p = 0.636**	indefinite	
	grade 1 (n = 4)	0 (0.00%)	4 (40.00%)			
	grade 2 (n = 2)	0 (0.00%)	2 (20.00%)			
	grade 3 (n = 0)	0 (0.00%)	0 (0.00%)			
RANKL	grade 0 (n = 0)	0 (0.00%)	0 (0.00%)	p = 0.091**	indefinite	
	grade 1 (n = 1)	1 (50.00%)	0 (0.00%)			
	grade 2 (n = 3)	1 (50.00%)	2 (20.00%)			
	grade 3 (n = 8)	0 (0.00%)	8 (80.00%)			
OPG	grade 0 (n = 1)	0 (0.00%)	1 (10.00%)	p = 1**	1.00 1.00 1.00 0.429 (0.02–9.364)	0.59
	grade 1 (n = 1)	0 (0.00%)	1 (10.00%)			
	grade 2 (n = 6)	1 (50.00%)	5 (50.00%)			
	grade 3 (n = 4)	1 (50.00%)	3 (30.00%)			
RANKL/OPG ratio	RANKL < OPG	2 (100.0%)	0 (0.00%)	p = 0.015**	indefinite	
	RANKL = OPG	0 (0.00%)	4 (40.00%)			
	RANKL > OPG	0 (0.00%)	6 (60.00%)			

\*Mann–Whitney U test; \*\*Fisher's exact test. Statistically significant values are bolded. Q1 – lower quartile; Q3 – upper quartile; M ±SD – mean ± standard deviation; NR – no recurrence; HR – hazard ratio; 95% CI – 95% confidence interval; OKC – odontogenic keratocyst; RANKL – receptor activator for nuclear factor κ B ligand; OPG – osteoprotegerin; PCNA – proliferating cell nuclear antigen; RANK – receptor activator of nuclear factor κ B.

**Table 4.** Clinicoradiological, histopathological and immunohistochemical characteristics of the recurrent sporadic cases of OKC

Patient No.	Age [years]	Sex	Location	Recurrence period [months]	Lesion type	Size of lesion [mm <sup>2</sup> ]	Cortical perforation	Inflammatory score	Number of daughter cysts	Epithelial dysplasia	PCNA [%]	p53 [%]	Ki-67 [%]	Bcl-2 [grade]	COX-2 [grade]	RANK [grade]	RANKL [grade]	OPG [grade]
1.	48	F	A-Mn	24	UL	950	(+)	2	0	(-)	10.5	6	10.5	1	2	0	2	1
2.	36	F	P-Mn	36	UL	760	(-)	0	0	(-)	13	13	16.5	2	2	0	3	1
3.	26	M	P-Mn	120	UL	450	(-)	2	0	(-)	14	25	36.5	2	3	1	2	3
4.	15	M	P-Mn	72	ML	1148	(+)	0	0	(-)	23	11.5	17	1	3	0	1	1
5.	26	F	P-Mx	108	UL	330	(-)	1	4	(-)	26	12	17.5	3	3	1	3	1
6.	19	M	A-Mn	108	ML	950	(+)	1	0	(+)	24	18	18.5	1	2	1	3	3
7.	62	F	P-Mn	108	UL	240	(-)	0	0	(-)	25	10	10	0	3	1	3	2
8.	36	M	P-Mn	36	ML	594	(+)	1	0	(-)	13.5	11	10.5	1	3	0	3	1
9.	73	M	A-Mn	24	ML	1025	(+)	1	0	(-)	27.5	7	10.5	0	3	2	3	1
10.	67	M	P-Mn	60	ML	1064	(+)	0	0	(-)	19	21	6	1	0	1	2	0
11.	29	M	P-Mn	72	UL	700	(-)	0	2	(-)	24	9	19.5	0	3	2	3	3

A-Mn – anterior mandible; P-Mn – posterior mandible; P-Mx – posterior maxilla; UL – unilocular; ML – multilocular; M – male; F – female; OKC – odontogenic keratocyst; RANKL – receptor activator for nuclear factor  $\kappa$  B ligand; OPG – osteoprotegerin; PCNA – proliferating cell nuclear antigen; RANK – receptor activator of nuclear factor  $\kappa$  B.

**Table 5.** Clinicoradiological, histopathological and immunohistochemical characteristics of the recurrent syndromic cases of OKC

Patient No.	Age [years]	Sex	Location	Recurrence period [months]	Lesion type	Size of lesion [mm <sup>2</sup> ]	Cortical perforation	Inflammatory score	Number of daughter cysts	Epithelial dysplasia	PCNA [%]	p53 [%]	Ki-67 [%]	Bcl-2 [grade]	COX-2 [grade]	RANK [grade]	RANKL [grade]	OPG [grade]
1.	25	F	P-Mx	96	UL	300	(-)	1	0	(-)	7.5	7.5	7.4	1	3	0	2	2
2.	12	M	P-Mn	24	UL	600	(+)	2	0	(+)	14	10.5	19.5	1	3	1	3	3
3.	14	M	P-Mx	96	UL	270	(-)	1	9	(+)	21	27	18	2	3	2	3	2
4.	19	F	A-Mn	72	ML	1320	(+)	2	0	(-)	18	22	11.5	1	1	1	2	1
5.	12	M	P-Mx	96	UL	375	(-)	0	0	(-)	33.5	17	10.5	1	1	1	3	0
6.	30	M	P-Mx	84	UL	360	(-)	2	9	(-)	31	13	18.5	2	3	2	3	3
7.	17	M	P-Mx	96	ML	450	(+)	2	19	(+)	41.5	17	13	1	1	0	3	2
8.	12	F	P-Mn	120	ML	550	(+)	3	0	(-)	35.5	17	13	1	1	0	3	2
9.	23	F	P-Mn	84	UL	300	(+)	0	0	(-)	30	14.5	13	1	3	1	3	3
10.	23	F	A-Mn	84	ML	525	(+)	2	23	(-)	41	27	12.5	1	2	0	3	2

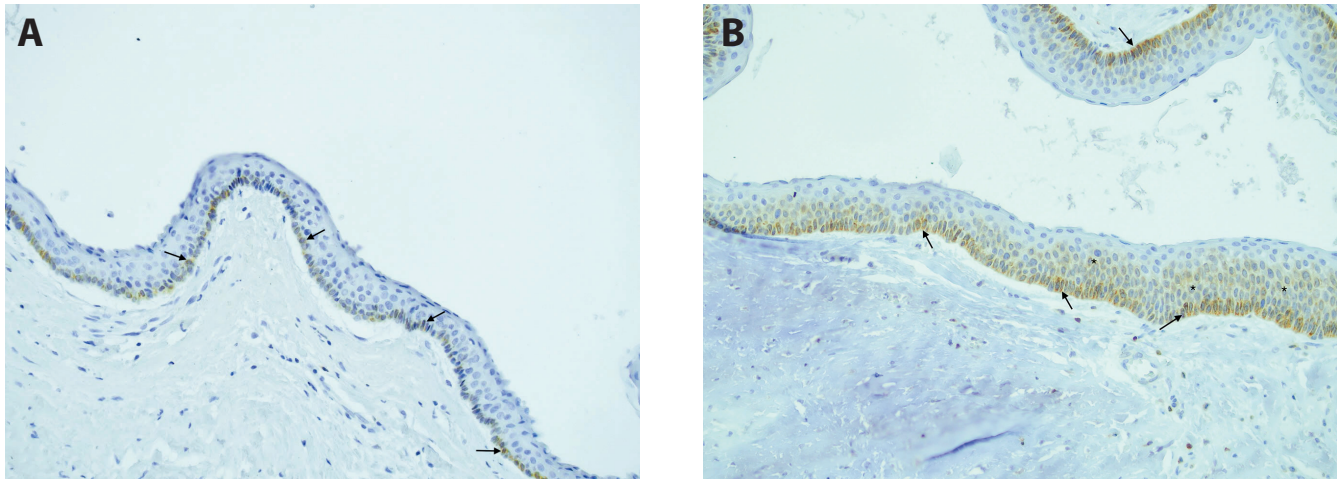
A-Mn – anterior mandible; P-Mn – posterior mandible; P-Mx – posterior maxilla; UL – unilocular; ML – multilocular; M – male; F – female; OKC – odontogenic keratocyst; RANKL – receptor activator for nuclear factor  $\kappa$  B ligand; OPG – osteoprotegerin; PCNA – proliferating cell nuclear antigen; RANK – receptor activator of nuclear factor  $\kappa$  B.

of OKCs. The published results in this regard are ambiguous. On the one hand, Yagyuu et al. found a significantly higher propensity for relapse of multilocular OKC compared to its unilocular counterpart (recurrence rate 35.7% compared to 7.4%).<sup>37</sup> Similar results were demonstrated in a study by Tabrizi et al., who calculated the recurrence rate of multilocular OKC at 62.16%, compared to 5.5% for unilocular lesions.<sup>38</sup> On the other hand, de França et al. found a similar frequency of relapse between unilocular and multilocular OKCs (44.4% compared to 55.6%), and did not show that locularity of the cyst has any significant

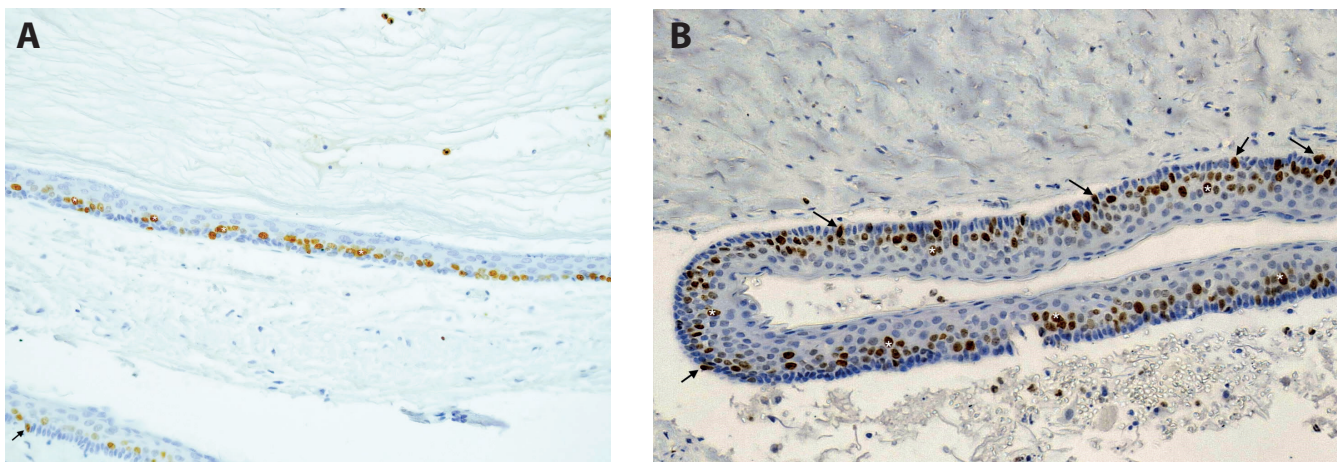
influence on relapse.<sup>39</sup> The ambiguity of our own results and of the published data suggests that, in essence, there may be a combination of several factors having an impact on OKC recurrence, including size, locularity and association with dentition. Thus, it seems reasonable that future research should focus on testing these factors – taken collectively or in different combinations.

The results of some previous studies suggest that COX-2 is overexpressed in OKC<sup>40</sup> and may serve as an important marker involved in local aggressiveness and recurrence of the lesion.<sup>41,42</sup> Herein, we showed that COX-2

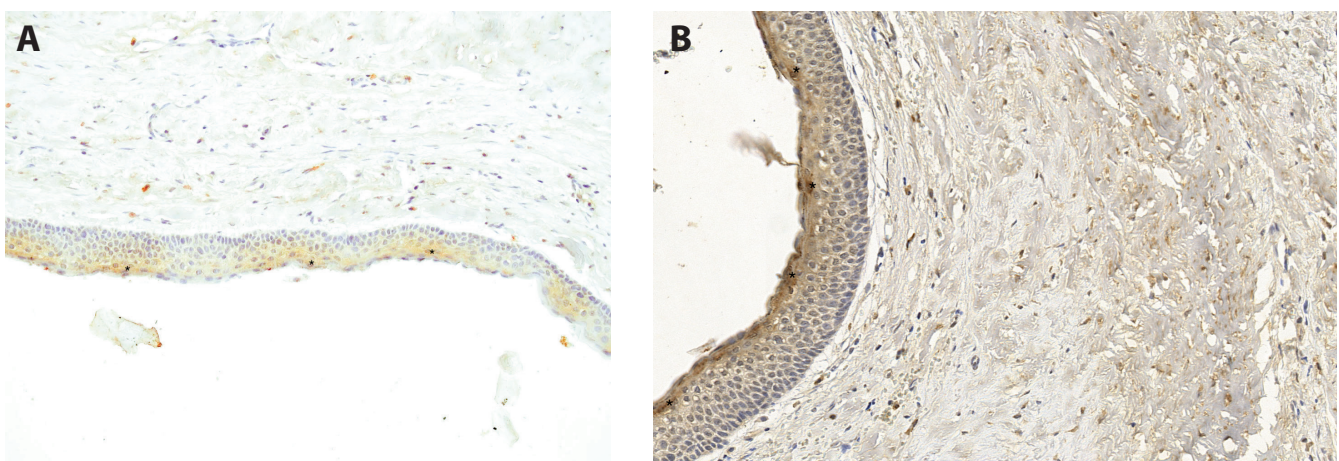




**Fig. 6.** Bcl-2 – membranous-cytoplasmic brown staining in the basal (arrows) and suprabasal (asterisks) layers of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x200 magnification)



**Fig. 7.** Ki-67 – nuclear brown staining in the basal (arrows) and suprabasal/superficial (asterisks) layers of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x60 magnification)



**Fig. 8.** Receptor activator of nuclear factor  $\kappa$  B (RANK) – cytoplasmic brown staining in the superficial layer (asterisks) of the epithelial lining of sporadic (A) and syndromic (B) odontogenic keratocyst (OKC) (x60 magnification)

immunostaining was significantly more pronounced in recurrent compared to non-recurrent sporadic cases of OKC. In fact, over 90% of recurrent cysts exhibited a moderate to strong reaction to COX-2, whereas 50%

of the non-recurrent lesions showed a negative or merely weak stain. By juxtaposing these results with the evidence that overexpression of the COX-2 gene alters adhesion, inhibits apoptosis and alters the response to growth regulatory

signals,<sup>43,44</sup> one might speculate that at least some of the recurrent OKCs are in fact neoplastic. However, in the current study, the immunoeexpression of COX-2 had no prognostic relevance, as it did not influence the time to recurrence in a survival model; in this respect, it is consistent with our previous findings.<sup>17</sup> Although all of the syndromic cases exhibited COX-2 expression, we did not find any difference between recurrent and non-recurrent lesions in this regard. However, we found that 50% of the recurrent and 0% of the non-recurrent syndromic cases exhibited a strong reaction to COX-2. In light of this, it may be assumed that COX-2 upregulation is relevant to the growth and progression of OKCs. Previous studies have demonstrated that COX-2 plays an important role in the biological regulation of OKC epithelial lining,<sup>41</sup> and it has been shown to be partially involved in the mechanism of OKC progression.<sup>39</sup> The COX-2 is known to increase the level of Bcl-2, thus suppressing apoptosis.<sup>44</sup> Moreover, COX-2-dependent prostaglandins have been implicated in the induction of matrix metalloproteinases (MMPs), which are important in matrix degradation during tumor growth and invasion, and induction of angiogenesis.<sup>45,46</sup> Very few studies have elucidated an association between COX-2 immunoeexpression and OKC recurrence. Furthermore, current evidence does not indicate that COX-2 has any prognostic relevance in OKC.<sup>41</sup> Nevertheless, our results hinting at COX-2 upregulation in recurrent sporadic cases coupled with data on its overexpression in OKCs suggest that COX-2 should be considered a potential marker to analyze the biological behavior of OKCs.<sup>40</sup> However, further investigation on larger samples is necessary. Recently, Alsaegh et al.<sup>46</sup> proposed that active human papillomavirus (HPV) infection may be associated with COX-2 expression in the odontogenic epithelium of OKC, and that HPV infection may have a role in the pathogenesis and aggressiveness of the cyst. However, this hypothesis has not been supported by any further research.

The RANKL and OPG are critical molecules for the control of osteoclastogenesis and pathophysiological bone remodeling.<sup>27,47</sup> In short, binding of RANKL to RANK leads to proosteoclast recruitment, and osteoclast activation and survival, whereas OPG is a decoy receptor for RANKL that blocks osteoclast formation by inhibiting RANKL from binding to RANK.<sup>18,27,48</sup> Not surprisingly, the vast majority of lesions in the current series exhibited RANKL > OPG or, to a lesser extent, a RANKL = OPG ratio, irrespective of the association with NBCCS. Many previous studies have demonstrated an elevated expression of RANKL in comparison with OPG in various osteolytic diseases, including aggressive odontogenic tumors like ameloblastoma, odontogenic myxoma and ameloblastic fibroma.<sup>49–51</sup> In our previous report, we showed that the majority of sporadic OKCs exhibited a RANKL > OPG ratio or, to a lesser degree, a RANKL = OPG ratio.<sup>18</sup> In the current study, we showed that there is a significantly increased expression of RANKL in recurrent syndromic lesions. In fact, 80% of the recurrent syndromic OKCs exhibited a strong

reaction to RANKL, and 60% of the recurrent syndromic OKCs had a RANKL > OPG ratio. These findings suggest that there is an elevated osteoclast activity in recurrent OKCs. However, we are not certain whether it has any prognostic relevance. First, our previous<sup>18</sup> and current results do not show that the expression of RANK, RANKL or OPG, or RANKL/OPG balance in sporadic OKCs has any influence on the risk and time to recurrence in the survival model. It was also not possible to calculate the HR for syndromic cases in this regard. Finally, the results of some studies on RANKL/OPG ratio in OKC indicate that the majority of OKC cases exhibit a predominance of OPG over RANKL.<sup>27,51</sup> De Matos et al. demonstrated that the majority of OKC cases exhibited an OPG = RANKL or OPG > RANKL ratio and suggested that this may stem from the cystic but not solid architecture of OKCs.<sup>51</sup> However, the authors did not test their findings against recurrence. In turn, Nonaka et al., having achieved similar results of RANKL/OPG ratio, additionally found that there were significant differences in the balance of the 2 molecules between recurrent and non-recurrent cases.<sup>27</sup> However, their analysis addressed this issue only in regard to sporadic cases, whereas syndromic lesions were not divided into recurrent and non-recurrent. In view of the above, the current results and data from the published material indicate that there is a disturbance of the functional equilibrium in the RANKL/OPG system in both sporadic and syndromic OKCs; however, this dysregulation may be in either direction, and its prognostic significance is unclear.

A recent systematic review and meta-analysis failed to identify immunohistochemical markers that could accurately distinguish syndromic from sporadic cases of OKC.<sup>12</sup> In the current study, however, we have shown that there is a significant difference in the expression of PCNA, p53 and OPG between the 2 types of OKC. These 3 markers are associated with the expansion of the lesion. In addition, PCNA and p53 are directly related to cellular proliferation. Previous studies found that the proliferation of odontogenic epithelium in syndromic OKCs is significantly more common than in its sporadic counterpart; these conclusions were mainly based on the results of PCNA and p53 expression in the cystic epithelium.<sup>16,52,53</sup> Lo Muzio et al. found the overexpression of p53 protein in syndromic OKCs and, together with the increase of PCNA positivity, deduced that it constitutes a valid background for the existence of a more aggressive phenotype of lesions associated with NBCCS.<sup>53</sup> Recently, Slusarenko da Silva et al. in their systematic review and meta-analysis found that OKCs are more prone to express the p53 marker than clinically benign dentigerous cysts, but quite similar to clinically aggressive ameloblastomas.<sup>7</sup> This is another argument in favor of the concept that OKC is prone to behave as a destructive odontogenic tumor. The significantly higher immunoeexpression of p53 revealed in syndromic cases in the present study may suggest that at least this subset of OKCs might be considered tumors rather than cysts.

We also found that there is a significant difference in the expression of OPG between sporadic and syndromic OKCs. The OPG upregulation may suggest decreased osteolytic activity, since OPG is an inhibitor of RANK/RANKL interaction, thus preventing osteoclastogenesis. In the current series, the majority of the lesions exhibited positive OPG immunoreaction; surprisingly, we found that most of the sporadic cases showed weak reactions, whereas in syndromic cysts the reaction was mainly moderate to strong. In light of the preceding considerations, we believe that OPG expression should be assessed in conjunction with RANKL expression, since the balance between both molecules regulates the process of bone resorption. As there was no significant difference in RANKL expression and RANKL/OPG balance, we do not think that our discovery of OPG upregulation in syndromic OKC cases indicates their decreased osteolytic activity.

## Limitations

A weak point of this study is the small sample size, particularly in the subset of syndromic OKCs. Moreover, having only 2 cases of non-recurrent syndromic OKC prevented us from calculating the HR for recurrence potential with respect to certain variables. The shortcomings of this study preclude us from drawing any binding conclusions, but our results suggest that immunohistochemistry may play a role in discriminating syndromic from sporadic OKC cases and has some prognostic relevance. Contrary to expensive and not widely available genetic testing for NBCCS, immunohistochemistry is a cost-effective technique that can be routinely applied in fixed tissue samples and may offer the potential to distinguish syndromic from sporadic OKCs.<sup>12</sup> An early recognition of NBCCS is of vital importance, as patients with NBCCS have an elevated risk of developing malignancies, such as BCC or medulloblastoma, and OKCs associated with NBCCS are more aggressive and have an increased risk of recurrence.<sup>27</sup> In their meticulous meta-analysis on differences in immunoprofile between sporadic and syndromic OKCs, Kalogirou et al. pointed out that there is a remarkable inconsistency in the diagnostic criteria employed for the selection of patients with NBCCS in various studies.<sup>12</sup> There are no universally accepted criteria for the diagnosis of NBCCS, and, together with the retrospective nature of most studies and their incomplete histopathological description of OKC (reporting merely the presence of a parakeratinized epithelium, which is found in other odontogenic cysts, without mentioning pathognomonic characteristics), there is increased bias in most published evidence.<sup>12</sup>

## Conclusions

In conclusion, the results of this study show that NBCCS-associated OKCs are significantly more prone to recur than their sporadic counterparts. Larger size and radiological

signs of cortical perforation in sporadic OKCs may indicate a higher risk of recurrence. The significance of these features, together with the cyst's locularity and association with dentition, should be confirmed in large-scale studies, especially by testing them in different configurations. The COX-2 is upregulated in recurrent sporadic OKCs, but it has no prognostic relevance. Recurrent syndromic OKCs exhibited higher RANKL and lower OPG expression, but we failed to demonstrate their prognostic significance. The immunoexpression of p53, PCNA and OPG may distinguish syndromic from sporadic OKCs, but these results require a confirmation with a larger sample.

## ORCID iDs

Konrad Kisielowski  <https://orcid.org/0000-0001-7894-0105>  
 Bogna Drozdowska  <https://orcid.org/0000-0002-2287-6842>  
 Mariusz Szuta  <https://orcid.org/0000-0002-7182-4811>  
 Tomasz Kaczmarzyk  <https://orcid.org/0000-0003-3024-9926>

## References

1. Kuroyanagi N, Sakuma H, Miyabe S, et al. Prognostic factors for keratocystic odontogenic tumor (odontogenic keratocyst). Analysis of clinico-pathologic and immunohistochemical findings in cysts treated by enucleation: Prognostic factors for keratocystic odontogenic tumour. *J Oral Pathol Med.* 2009;38(4):386–392. doi:10.1111/j.1600-0714.2008.00729.x
2. Titinchi F. Novel recurrence risk stratification of odontogenic keratocysts: A systematic review [published online as ahead of print on June 1, 2021]. *Oral Dis.* 2021. doi:10.1111/odi.13931
3. Speight P, Devilliers P, Li T, Odell T, Wright JM. Odontogenic keratocyst. In: El-Naggar A, Chan J, Grandis J, Takata T, Sloatweg PJ, eds. *WHO Classification of Head and Neck Tumours.* 4th ed. Lyon, France: International Agency for Research on Cancer (IARC); 2017:235–236. ISBN:978-92-832-2438-9.
4. Philipsen H. Keratocystic odontogenic tumour. In: Barnes L, Eveson J, Reichart P, Sidransky D, eds. *Pathology and Genetics of Head and Neck Tumours.* Lyon, France: International Agency for Research on Cancer (IARC); 2007:306–307. ISBN:978-92-832-2417-4.
5. Wright JM, Vered M. Update from the 4th Edition of the World Health Organization Classification of Head and Neck Tumours: Odontogenic and maxillofacial bone tumors. *Head Neck Pathol.* 2017;11(1):68–77. doi:10.1007/s12105-017-0794-1
6. Hoyos Cadavid AM, Kaminagakura E, Rodrigues MFSD, Pinto CAL, Teshima THN, Alves FA. Immunohistochemical evaluation of Sonic Hedgehog signaling pathway proteins (Shh, Ptch1, Ptch2, Smo, Gli1, Gli2, and Gli3) in sporadic and syndromic odontogenic keratocysts. *Clin Oral Invest.* 2019;23(1):153–159. doi:10.1007/s00784-018-2421-2
7. Slusarenko da Silva Y, Stoelinga PJW, Grillo R, da Graça Naclério-Homem M. Cyst or tumor? A systematic review and meta-analysis on the expression of p53 marker in odontogenic keratocysts. *J Craniomaxillofac Surg.* 2021;49(12):1101–1106. doi:10.1016/j.jcms.2021.09.015
8. MacDonald-Jankowski DS. Keratocystic odontogenic tumour: Systematic review. *Dentomaxillofac Radiol.* 2011;40(1):1–23. doi:10.1259/dmfr/29949053
9. Cesinero AM, Burtini G, Maiorana A, Rossi G, Migaldi M. Expression of calretinin in odontogenic keratocysts and basal cell carcinomas: A study of sporadic and Gorlin–Goltz syndrome-related cases. *Ann Diagn Pathol.* 2020;45:151472. doi:10.1016/j.anndiagpath.2020.151472
10. Woolgar JA, Rippin JW, Browne RM. A comparative study of the clinical and histological features of recurrent and nonrecurrent odontogenic keratocysts. *J Oral Pathol Med.* 1987;16(3):124–128. doi:10.1111/j.1600-0714.1987.tb01478.x
11. Antonoglou GN, Sándor GK, Koidou VP, Papageorgiou SN. Non-syndromic and syndromic keratocystic odontogenic tumors: Systematic review and meta-analysis of recurrences. *J Craniomaxillofac Surg.* 2014;42(7):e364–e371. doi:10.1016/j.jcms.2014.03.020

12. Kalogirou EM, Thermos G, Zogopoulos V, et al. The immunohistochemical profile of basal cell nevus syndrome-associated and sporadic odontogenic keratocysts: A systematic review and meta-analysis. *Clin Oral Invest*. 2021;25(6):3351–3367. doi:10.1007/s00784-021-03877-w
13. Kimonis VE, Goldstein AM, Pastakia B, et al. Clinical manifestations in 105 persons with nevoid basal cell carcinoma syndrome. *Am J Med Genet*. 1997;69(3):299–308. PMID:9096761.
14. Bree AF, Shah MR; for the BCNS Colloquium Group. Consensus statement from the First International Colloquium on Basal Cell Nevus Syndrome (BCNS). *Am J Med Genet*. 2011;155(9):2091–2097. doi:10.1002/ajmg.a.34128
15. Ibrahim N, Nazimi AJ, Ajura AJ, Nordin R, Latiff ZA, Ramli R. The clinical features and expression of Bcl-2, cyclin D1, p53, and proliferating cell nuclear antigen in syndromic and nonsyndromic keratocystic odontogenic tumor. *J Craniofac Surg*. 2016;27(5):1361–1366. doi:10.1097/SCS.0000000000002792
16. El Murtadi A, Grehan D, Toner M, McCartan BE. Proliferating cell nuclear antigen staining in syndrome and nonsyndrome odontogenic keratocysts. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol*. 1996;81(2):217–220. doi:10.1016/S1079-2104(96)80418-5
17. Kaczmarzyk T, Kisielowski K, Koszowski R, et al. Investigation of clinicopathological parameters and expression of COX-2, Bcl-2, PCNA, and p53 in primary and recurrent sporadic odontogenic keratocysts. *Clin Oral Invest*. 2018;22(9):3097–3106. doi:10.1007/s00784-018-2400-7
18. Kisielowski K, Drozdowska B, Koszowski R, et al. Immunoexpression of RANK, RANKL and OPG in sporadic odontogenic keratocysts and their potential association with recurrence. *Adv Clin Exp Med*. 2021;30(3):301–307. doi:10.17219/acem/130907
19. R Core Team. R: A language and environment for statistical computing. 2020. <https://www.R-project.org/>. Accessed April 15, 2022.
20. Therneau T. A package for survival analysis in R. 2022. <https://cran.r-project.org/web/packages/survival/vignettes/survival.pdf>. Accessed April 25, 2022.
21. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model*. New York-Berlin-Heidelberg: Springer; 2000. ISBN:978-0-387-98784-2.
22. Stoelinga PJW. Long-term follow-up on keratocysts treated according to a defined protocol. *Int J Oral Maxillofac Surg*. 2001;30(1):14–25. doi:10.1054/ijom.2000.0027
23. Kaczmarzyk T, Mojsa I, Stypulkowska J. A systematic review of the recurrence rate for keratocystic odontogenic tumour in relation to treatment modalities. *Int J Oral Maxillofac Surg*. 2012;41(6):756–767. doi:10.1016/j.ijom.2012.02.008
24. Titinchi F, Nortje CJ. Keratocystic odontogenic tumor: A recurrence analysis of clinical and radiographic parameters. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2012;114:136–142. doi:10.1016/j.oooo.2012.01.032
25. Chirapathomsakul D, Sastravaha P, Jansisanont P. A review of odontogenic keratocysts and the behavior of recurrences. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol*. 2006;101(1):5–9. doi:10.1016/j.tripleo.2005.03.023
26. Naruse T, Yamashita K, Yanamoto S, et al. Histopathological and immunohistochemical study in keratocystic odontogenic tumors: Predictive factors of recurrence. *Oncol Lett*. 2017;13(5):3487–3493. doi:10.3892/ol.2017.5905
27. Nonaka CFW, Cavalcante RB, Nogueira RLM, de Souza LB, Pinto LP. Immunohistochemical analysis of bone resorption regulators (RANKL and OPG), angiogenic index, and myofibroblasts in syndrome and non-syndrome odontogenic keratocysts. *Arch Oral Biol*. 2012;57(3):230–237. doi:10.1016/j.archoralbio.2011.08.002
28. Kimi K, Kumamoto H, Ooya K, Motegi K. Immunohistochemical analysis of cell-cycle- and apoptosis-related factors in lining epithelium of odontogenic keratocysts: Cell cycle and apoptosis in odontogenic keratocysts. *J Oral Pathol Med*. 2001;30(7):434–442. doi:10.1034/j.1600-0714.2001.300709.x
29. Amorim R, Godoy G, Galvao H, Souza L, Freitas R. Immunohistochemical assessment of extracellular matrix components in syndrome and non-syndrome odontogenic keratocysts. *Oral Dis*. 2004;10(5):265–270. doi:10.1111/j.1601-0825.2004.01023.x
30. Kolar Z, Geierova M, Bouchal J, Pazdera J, Zboril V, Tvrdy P. Immunohistochemical analysis of the biological potential of odontogenic keratocysts. *J Oral Pathol Med*. 2006;35(2):75–80. doi:10.1111/j.1600-0714.2006.00382.x
31. Cavalcante RB, Pereira KMA, Nonaka CFW, Maia Nogueira RL, de Souza LB. Immunohistochemical expression of MMPs 1, 7, and 26 in syndrome and nonsyndrome odontogenic keratocysts. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol*. 2008;106(1):99–105. doi:10.1016/j.tripleo.2007.12.028
32. Leonardi R, Matthews J, Caltabiano R, et al. MMP-13 expression in keratocyst odontogenic tumour associated with NBCCS and sporadic keratocysts: MMP13 expression in KCOTs. *Oral Dis*. 2010;16(8):795–800. doi:10.1111/j.1601-0825.2010.01690.x
33. Manfredi M, Vescovi P, Bonanini M, Porter S. Nevoid basal cell carcinoma syndrome: A review of the literature. *Int J Oral Maxillofac Surg*. 2004;33(2):117–124. doi:10.1054/ijom.2003.0435
34. Chrcanovic BR, Gomez RS. Recurrence probability for keratocystic odontogenic tumors: An analysis of 6427 cases. *J Craniofac Surg*. 2017;45(2):244–251. doi:10.1016/j.jcms.2016.11.010
35. Avril L, Lombardi T, Ailianou A, et al. Radiolucent lesions of the mandible: A pattern-based approach to diagnosis. *Insights Imaging*. 2014;5(1):85–101. doi:10.1007/s13244-013-0298-9
36. Fidele N, Yueyu Z, Zhao Y, et al. Recurrence of odontogenic keratocysts and possible prognostic factors: Review of 455 patients. *Med Oral*. 2019;24(4):e491–e501. doi:10.4317/medoral.22827
37. Yagyu T, Kirita T, Sasahira T, Moriwaka Y, Yamamoto K, Kuniyasu H. Recurrence of keratocystic odontogenic tumor: Clinicopathological features and immunohistochemical study of the hedgehog signaling pathway. *Pathobiology*. 2008;75(3):171–176. doi:10.1159/000124977
38. Tabrizi R, Omidi M, Dehbozorgi M, Hekmat M. Correlation of radiographic features and treatments with the frequency of recurrence in odontogenic keratocysts of the mandible. *J Craniofac Surg*. 2014;25(5):e413–e417. doi:10.1097/SCS.0000000000000910
39. de França GM, da Silva LBA, Mafra RP, da Silva WR, de Lima KC, Galvão HC. Recurrence-free survival and prognostic factors of odontogenic keratocyst: A single-center retrospective cohort. *Eur Arch Otorhinolaryngol*. 2021;278(4):1223–1231. doi:10.1007/s00405-020-06229-8
40. Nimmanagoti R, Nandan SK, Kulkarni P, Reddy S, Keerthi M, Pupala G. Protein 53, B-cell lymphoma-2, cyclooxygenase-2, and CD105 reactivity in keratocystic odontogenic tumors: An immunohistochemical analysis. *Int J App Basic Med Res*. 2019;9(1):27–31. doi:10.4103/ijabmr.IJABMR\_138\_18
41. Mendes RA, Carvalho JFC, van der Waal I. A comparative immunohistochemical analysis of COX-2, p53, and Ki-67 expression in keratocystic odontogenic tumors. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol*. 2011;111(3):333–339. doi:10.1016/j.tripleo.2010.10.004
42. Wang J, Zhang X, Ding X, et al. Cyclooxygenase-2 expression in keratocystic odontogenic tumour decreased following decompression. *Mol Clin Oncol*. 2013;1(6):982–986. doi:10.3892/mco.2013.169
43. Tsujii M, DuBois RN. Alterations in cellular adhesion and apoptosis in epithelial cells overexpressing prostaglandin endoperoxide synthase 2. *Cell*. 1995;83(3):493–501. doi:10.1016/0092-8674(95)90127-2
44. Mendes RA, Carvalho JFC, van der Waal I. An overview on the expression of cyclooxygenase-2 in tumors of the head and neck. *Oral Oncol*. 2009;45(10):e124–e128. doi:10.1016/j.oraloncology.2009.03.016
45. Gomes CC, Duarte AP, Diniz MG, Gomez RS. Current concepts of ameloblastoma pathogenesis. *J Oral Pathol Med*. 2010;39(8):585–591. doi:10.1111/j.1600-0714.2010.00908.x
46. Alsaegh MA, Miyashita H, Zhu SR. Expression of human papillomavirus is correlated with Ki-67 and COX-2 expressions in keratocystic odontogenic tumor. *Pathol Oncol Res*. 2015;21(1):65–71. doi:10.1007/s12253-014-9789-3
47. Baud'huin M, Lamoureux F, Duplomb L, Rédini F, Heymann D. RANKL, RANK, osteoprotegerin: Key partners of osteoimmunology and vascular diseases. *Cell Mol Life Sci*. 2007;64(18):2334–2350. doi:10.1007/s00018-007-7104-0
48. Tekkesin MS, Mutlu S, Olgac V. The role of RANK/RANKL/OPG signalling pathways in osteoclastogenesis in odontogenic keratocysts, radicular cysts, and ameloblastomas. *Head Neck Pathol*. 2011;5(3):248–253. doi:10.1007/s12105-011-0271-1
49. Kumamoto H, Ooya K. Expression of parathyroid hormone-related protein (PTHrP), osteoclast differentiation factor (ODF)/receptor activator of nuclear factor-kappaB ligand (RANKL) and osteoclastogenesis inhibitory factor (OCIF)/osteoprotegerin (OPG) in ameloblastomas. *J Oral Pathol Med*. 2004;33(1):46–52. doi:10.1111/j.1600-0714.2004.00204.x

50. Andrade FR, Sousa DP, Mendonça EF, Silva TA, Lara VS, Batista AC. Expression of bone resorption regulators (RANK, RANKL, and OPG) in odontogenic tumors. *Oral Surg Oral Med Oral Pathol Oral Radiol Endodontol.* 2008;106(4):548–555. doi:10.1016/j.tripleo.2008.05.042
51. de Matos FR, de Moraes M, das Neves Silva EB, Galvão HC, de Almeida Freitas R. Immunohistochemical detection of receptor activator nuclear  $\kappa$ B ligand and osteoprotegerin in odontogenic cysts and tumors. *J Oral Maxillofac Surg.* 2013;71(11):1886–1892. doi:10.1016/j.joms.2013.05.023
52. Li TJ, Browne RM, Prime SS, Paterson IC, Matthews JB. p53 expression in odontogenic keratocyst epithelium. *J Oral Pathol Med.* 1996; 25(5):249–255. doi:10.1111/j.1600-0714.1996.tb01380.x
53. Lo Muzio L, Staibano S, Pannone G, et al. Expression of cell cycle and apoptosis-related proteins in sporadic odontogenic keratocysts and odontogenic keratocysts associated with the nevoid basal cell carcinoma syndrome. *J Dent Res.* 1999;78(7):1345–1353. doi:10.1177/00220345990780070901

