

SÃO PAULO Medical Journal

E V I D E N C E F O R H E A L T H C A R E

May 7 - Volume 144 - Number 3

Editorial

- Safety in the transport of substances of human origin: a critical challenge for the quality of transplants

Cross-sectional study

- Clinical features of chronic tension-type headache combined with hyperacusis

Prospective cohort study

- Connecting care—telemedicine breaks barriers for low-income Brazilians battling breast cancer

Medline, LILACS,
SciELO, Science
Citation Index Expanded,
Journal Citation Reports/
Sciences Edition
(impact factor 1.838),
EBSCO Publishing and
PubMed Central
(PMC)



José Parrila Square, Adamantina city, State of São Paulo
Photo: Alf Ribeiro - expressão studio

XXV

CONGRESSO BRASILEIRO DE LASER
EM MEDICINA E CIRURGIA

SBLMC - 30 ANOS

LASER MÉDICO E TECNOLOGIAS CORRELATAS

O congresso foi pensado para quem quer se aprofundar nos avanços da medicina a laser. Ao longo de três dias – **1, 2 e 3 de outubro de 2026, em formato híbrido** –, especialistas de referência conduzem debates sobre tendências, desafios atuais e novas possibilidades na área.

PRINCIPAIS TEMAS

- ✓ Laser na Dor
- ✓ Laser na Oncologia
- ✓ Ortopedia e medicina esportiva
- ✓ Tecnologias multiprofissionais
- ✓ Vascular e endovascular
- ✓ Urologia
- ✓ Cirurgia plástica e dermatologia
- ✓ Ginecologia e saúde íntima
- ✓ Proctologia
- ✓ Oftalmologia e otorrinolaringologia
- ✓ Cardiologia

1, 2 e 3 de outubro de 2026

01/10 Pré-congresso · Evento híbrido

Local: Sede da APM – São Paulo

Inscreva-se

Acesse

congressobrasileirolaser.com.br
ou aponte a câmera para o QR code.



Médico técnico-científico responsável pelo evento: Dr. Moisés Wolfenson – CRM-PE 2935 | RQE 1993

Realização



Comercialização e promoção



Patrocínio Keypartner



Editorial

- e20261443 Safety in the transport of substances of human origin: a critical challenge for the quality of transplants
Bartira de Aguiar Roza, Andre Ibrahim David, Flavio Pola dos Reis, Paulo Manuel Pêgo-Fernandes

Original Article

- e20253116 Connecting care—telemedicine breaks barriers for low-income Brazilians battling breast cancer: a prospective cohort study
Gabriel Salum D'Alessandro, Yedda Nunes Reis, Mila Trementosa Garcia, Daniela Zaros Guimarães, Roberta Amparado Mizziara, Flavia Colsomagno Silveira Berlofa, Mayara de Cássia Benedito Andrade, Emily Rie Yoshizato, Jonathan Yugo Maesaka
- e20253135 Translation, cultural adaptation, and evidence of content and face validity of a questionnaire to assess REDs knowledge in Brazil: a cross-cultural validation study
Uyara Pereira de Maria, Claudia Ridel Juzwiak
- e20253184 Psychometric properties of the Positive Aspects of Caregiving Scale for unpaid caregivers of people living with dementia in the Brazilian context: a validation study
Sofia Cristina Iost Pavarini, Aline Cristina Martins Gratao, Camila Rafael Ferreira Campos, Diana Quirino Monteiro, Elizabeth Joan Barham, Fabiana de Souza Orlandi, Gabriela Martins, Gustavo Carrijo Barbosa, Keila Cristianne Trindade da Cruz, Larissa Corrêa, Luana Aparecida Rocha, Ludmyla Caroline de Souza Alves, Ana Carolina Ottaviani
- e20253205 Clinical features of chronic tension-type headache combined with hyperacusis: a cross-sectional study
Zenghui Fu, Yan Jin, Zaihong Lin, Yang Liu, Kedong Guo
- e20253373 Predictive modeling with machine learning for low birth weight in the state of São Paulo: a retrospective population-based study
Allbert Velleniche de Aquino Almeida, Adriana de Oliveira Ribeiro dos Santos, Luiz Fernando Costa Nascimento
- e20253392 The effects of agmatine treatment on cisplatin-induced hepatotoxicity: an experimental rat study
Murat Yeniceri, Mustafa Can Senoyamak, Süleyman Bas, Musa Salmanoglu, Alpaslan Tanoglu
- e20253465 Assessment of balance, sleep quality, and depression in relation to attention in Parkinson's disease: A cross-sectional P300 study
Gulfem Ezgi Ozaltin, Agit Simsek, Busra Candiri, Fatma Ebru Algul, Mehmet Tecellioglu, Burcu Talu
- e20253529 Family functionality and social support in Brazilian older people with chronic kidney disease in hemodialysis: a cross-sectional study
Ana Beatriz Murari Thomazini, Diana Gabriela Mendes dos Santos, Ynggrid Larisse Teodoro de Souza, Daiane Baldo Apolinário, Ana Luiza Barbosa, Marisa Silvana Zazzetta, Fabiana de Souza Orlandi
- e20263635 Multilevel auditory and cognitive processing in post-stroke aphasia: associations with language performance using ABR, LLR, and P300: a cross-sectional observational study
Agit Şimşek, Nihal Sümeyye Ulutaş, Feyza Deniz Saman
- e20263641 Prevalence of melasma in pregnant women and its relationship with body perception: a cross-sectional study in Türkiye
Demet Toktaş, Rabia Atilla
- e20263647 Treatment patterns and clinical outcomes in chronic myeloid leukemia treated with second-line nilotinib—association of molecular milestones and prognostic scores with deep molecular response: a real-world retrospective cohort study
Irena Čojbašić, Ivan Tijanić, Žarko Čojbašić
- e20263631 The effect of BIS-guided sevoflurane anesthesia on postoperative nausea and vomiting in pediatric adenoidectomy: a prospective randomized controlled trial
Engin Cetin, Ayse Sencan, Ahmet Yuksek, Bedirhan Gunel, Fatihhan Zeytun, Mehmet Yilmaz



Correspondence to:

ASSOCIAÇÃO PAULISTA DE MEDICINA

Publicações Científicas

Av. Brig. Luís Antônio, 278 - 7ª andar -
São Paulo (SP) - Brasil - CEP 01318-901

Tel. (+55 11) 3188-4310/3188-4311

E-mail: revistas@apm.org.br

www.scielo.br/spmj

Founded in 1932, a bimonthly publication of the Associação Paulista de Medicina e-mail: revistas@apm.org.br

Editors: Paulo Manuel Pêgo Fernandes, Marianne Yumi Nakai and Álvaro Nagib Atallah.
Editorial assistants: Thiago Silva and Fellipe Cotrim.
Associate editors: Adriana Seber, Ailton Tetelbom Stein, Alexander Wagner Silva de Souza, Antonio José Gonçalves, Aytan Miranda Sipahi, Cristina Muccioli, Delcio Matos, Edina Mariko Koga da Silva, Fernando Antonio de Almeida, Flávio Faloppa, Heráclito Barbosa de Carvalho, José Antônio Rocha Gontijo, José Carlos Costa Baptista-Silva, José Maria Soares Júnior, José Roberto Lapa e Silva, Laércio Joel Franco, Maria do Patrocínio Tenório Nunes, Milton de Arruda Martins, Moacir Fernandes de Godoy, Olavo Pires de Camargo, Renato Corrêa Baena, Sergio Tufik, Vania dos Santos Nunes.
Proofreading: Editage.
Desktop publishing: Zeppelini Publishers (www.zeppelini.com.br).
Listed in: Medline, Lilacs, SciELO, Science Citation Index Expanded and Journal Citation Reports/Sciences Edition, EBSCO publishing and PubMed Central.
International Board: Alexandre Wagner Silva de Souza (University Medical Center Groningen, Groningen, Netherlands), Charles J. Menkes (Cochin Hospital, Paris, France), José Fragata (CUF Infante Santo Hospital, Lisbon), Luiz Dratcu (Guy's Hospital, London, and Maudsley NHS Trust, York Clinic, London), Marcelo Cypel (University Health

Network, Toronto, Canada), Karla Soares-Weiser (Enhance Reviews Ltd, Wantage, United Kingdom), Tirone Espiridião David (Toronto General Hospital, Toronto, Canada), Mário Viana de Queiroz (Hospital de Santa Maria, Lisbon), Wadih Arap (MD Anderson Cancer Center, University of Texas, Houston, United States), Wellington V. Cardoso (Boston University, Boston, United States).
• All articles published, including editorials and letters, represent the opinions of the authors and do not reflect the official policy of the Associação Paulista de Medicina or the institution with which the authors are affiliated, unless this is clearly specified.
• All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publisher. Copyright © 2026 by Associação Paulista de Medicina.
• SPMJ website: access to the entire São Paulo Medical Journal/Revista Paulista de Medicina website is free to all. We will give at least six months notice of any change in this policy. SPMJ printed version: six issues/year; 1 volume/year, beginning on first Thursday in January.

Scientific Council

Abrão Rapoport – *Hospital Heliópolis, São Paulo*
Adriana Costa e Forti – *Faculdade de Medicina, Universidade Federal do Ceará*
Alexandre Fogaça Cristante – *Faculdade de Medicina da Universidade de São Paulo*
Álvaro Nagib Atallah – *Escola Paulista de Medicina, Universidade Federal de São Paulo*
Auro del Giglio – *Faculdade de Medicina da Fundação ABC*
Carmen Cabanelas Pazos de Moura – *Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro*
Cármio Antonio de Souza – *Faculdade de Ciências Médicas, Universidade Estadual de Campinas*
Dario Birolini – *Faculdade de Medicina, Universidade de São Paulo*
Eduardo Maia Freese de Carvalho – *Faculdade de Medicina, Universidade Federal de Pernambuco, Centro de Pesquisas Aggeu Magalhães - CpqAM/FIOCRUZ*
Egberto Gaspar de Moura – *Instituto de Biologia Roberto Alcântara Gomes, Universidade Estadual do Rio de Janeiro*
Eliézer Silva – *Hospital Israelita Albert Einstein, São Paulo*
Emílio Antonio Francischetti – *Faculdade de Medicina da Universidade Estadual do Rio de Janeiro*
Emmanuel de Almeida Burdmann – *Faculdade de Medicina da Universidade de São Paulo*
Fabio Bessa Lima – *Instituto de Ciências Biomédicas, Universidade de São Paulo*
Florence Kerr-Corrêa – *Faculdade de Medicina de Botucatu, Universidade Estadual de São Paulo*
Francisco José Penna – *Faculdade de Medicina Universidade Federal de Minas Gerais*
Geraldo Rodrigues de Lima – *Escola Paulista de Medicina, Universidade Federal de São Paulo*
Irineu Tadeu Velasco – *Faculdade de Medicina da Universidade de São Paulo*
João Renato Rebello Pinho – *Hospital Israelita Albert Einstein e Faculdade de Medicina da Universidade de São Paulo*
Joel Spadaro – *Faculdade de Ciências Médicas de Botucatu, Universidade Estadual de São Paulo*
Jorge Sabbaga – *Hospital Alernão Oswaldo Cruz, São Paulo*
José Antonio Marin-Neto – *Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo*

José Carlos Nicolau – *Instituto do Coração, Universidade de São Paulo*
José Geraldo Mill – *Faculdade de Medicina, Universidade Federal do Espírito Santo*
José Mendes Aldrighi – *Faculdade de Saúde Pública, Universidade de São Paulo*
José Roberto Lapa e Silva – *Instituto de Doenças do Tórax, Universidade Federal do Rio de Janeiro*
Leonardo Roeber – *Universidade Federal de Uberlândia*
Leopoldo Soares Piegas – *Instituto Dante Pazzanese de Cardiologia, São Paulo*
Luiz Paulo Kowalski – *Hospital AC Camargo, São Paulo*
Márcio Abrahão – *Escola Paulista de Medicina, Universidade Federal de São Paulo*
Maria Inês Schmidt – *Faculdade de Medicina, Universidade Federal do Rio Grande do Sul*
Maurício Mota de Avelar Alchorne – *Universidade Nove de Julho, São Paulo*
Mauro Schechter – *Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro*
Milton de Arruda Martins – *Faculdade de Medicina, Universidade de São Paulo*
Nelson Hamerschlag – *Hospital Israelita Albert Einstein, São Paulo*
Noedir Antônio Groppo Stolf – *Faculdade de Medicina, Universidade de São Paulo*
Paulo Manuel Pêgo Fernandes – *Instituto do Coração, Hospital das Clínicas HCFMUSP, Faculdade de Medicina, Universidade de São Paulo*
Pêrsio Roxo Júnior – *Faculdade de Medicina de Ribeirão Preto*
Raul Cutait – *Hospital Sírio-Libanês, São Paulo*
Raul Marino Junior – *Faculdade de Medicina, Universidade de São Paulo*
Ricardo Brandt de Oliveira – *Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo*
Roberto Alexandre Franken – *Faculdade de Ciências Médicas da Santa Casa de Misericórdia de São Paulo*
Souhbi Kahhale – *Faculdade de Medicina, Universidade de São Paulo*
Wilson Roberto Catapani – *Faculdade de Medicina do ABC, Santo André*
Wilson Cossermelli – *Reclin Reumatologia Clínica, São Paulo*

Diretoria Executiva da Associação Paulista de Medicina (Triênio 2023-2026)

Presidente: Antonio José Gonçalves
1º Vice-Presidente: João Sobreira de Moura Neto
2º Vice-Presidente: José Luiz Gomes do Amaral
3º Vice-Presidente: Akira Ishida
4º Vice-Presidente: Roberto Lofti Júnior
Secretário Geral: Paulo Cezar Mariani
1º Secretário: Paulo Cezar Mariani
Secretária Geral Adjunta: Maria Rita de Souza Mesquita
Diretor Administrativo: Lacildes Rovella Júnior
Diretor Administrativo Adjunto: Ademar Anzai
1º Diretor de Patrimônio e Finanças: Flórida Meinão
2º Diretor de Patrimônio e Finanças: Clóvis Acúrcio Machado
Diretor Científico: Paulo Manuel Pêgo Fernandes
Diretora Científica Adjunta: Marianne Yumi Nakai
Diretor de Defesa Profissional: José Eduardo Paciência Rodrigues
Diretor de Defesa Profissional Adjunto: Marun David Cury
Diretor de Comunicações: Marcos Cabello dos Santos
Diretor de Comunicações Adjunto: Renato Azevedo Júnior
Diretor de Marketing: Nicolau D'Amico Filho
Diretor de Marketing Adjunto: David Alves de Souza Lima
Diretor de Eventos: Fernando Sabia Tallo
Diretor de Eventos Adjunto: Geovanne Furtado Souza
Diretor de Tecnologia de Informação: Júlio Leonardo Barbosa Pereira
Diretora de Tecnologia de Informação Adjunta: Zilda Maria Tosta Ribeiro
Diretor de Previdência e Mutualismo: Antônio Carlos Endrigo

Diretor de Previdência e Mutualismo Adjunto: Clóvis Francisco Constantino
Diretora Social: Ana Beatriz Soares
Diretor Social Adjunto: Leonardo da Silva
Diretor de Responsabilidade Social: Jorge Carlos Machado Curi
Diretor de Responsabilidade Social Adjunto: Paulo Celso Nogueira Fontão
Diretora Cultural: Cleusa Cascaes Dias
Diretor Cultural Adjunto: Guido Arturo Palomba
Diretora de Serviços aos Associados: Diana Lara Pinto de Santana
Diretora de Serviços aos Associados Adjunta: Alice Antunes Mariani
Diretor de Economia Médica e Saúde Baseada em Evidências: Álvaro Nagib Atallah
Diretor de Economia Médica e Saúde Baseada em Evidências Adjunto: Paulo De Conti
1ª Diretora Distrital: Tereza Cristina Machado de Godoy
2º Diretor Distrital: Edemilson Cavalheiro
3º Diretor Distrital: Othon Mercadantes Becker
4º Diretor Distrital: Eduardo Luís Cruells Vieira
5ª Diretora Distrital: Fátima Ferreira Bastos
6º Diretor Distrital: João Carlos Sanches Anêas
7º Diretor Distrital: José Eduardo Marques
8º Diretor Distrital: Leandro Freitas Colturato
9º Diretor Distrital: Paulo Gil Katsuda
10ª Diretora Distrital: Juliana Cristina Kuhn Medina
11º Diretor Distrital: Eder Carvalho Sousa
12º Diretor Distrital: Luís Henrique Brandão Falcão
13º Diretor Distrital: Cezar Antônio Roselino Secchieri
14º Diretor Distrital: Ricardo Tedeschi Matos

Safety in the transport of substances of human origin: a critical challenge for the quality of transplants

Bartira de Aguiar Roza^I, Andre Ibrahim David^{II}, Flavio Pola dos Reis^{III}, Paulo Manuel Pêgo-Fernandes^{IV}

Instituto do Coração, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP)

^IMD, PhD. Associate Professor, Department of Surgical Nursing, Escola Paulista de Enfermagem, Universidade Federal de São Paulo (UNIFESP), São Paulo (SP), Brazil; WHO Expert Advisory Panel/Committee on Donation and Transplantation (ECDT). Collaborating Professor PAHO/WHO.

 <https://orcid.org/0000-0002-6445-6846>

^{II}MD, PhD. Adjunct Professor, Head of the discipline of operating techniques and experimental surgery, Escola Paulista de Medicina, Universidade Federal de São Paulo (UNIFESP), São Paulo (SP), Brazil.

 <https://orcid.org/0009-0000-0530-8918>

^{III}MD. Attending Physician, Thoracic Surgery Department, Instituto do Coração, Hospital das Clínicas, Faculdade de Medicina, Universidade de São Paulo (USP), São Paulo (SP), Brazil.

 <https://orcid.org/0000-0002-4510-1070>

^{IV}MD, PhD. Vice-Director, Faculdade de Medicina, Universidade de São Paulo (USP), São Paulo (SP), Brazil; Professor, Department of Cardiopneumology, Faculdade de Medicina, Universidade de São Paulo (USP), São Paulo (SP), Brazil; Director, Scientific Department, Associação Paulista de Medicina (APM), São Paulo (SP), Brazil.

 <https://orcid.org/0000-0001-7243-5343>

The safe transport of substances of human origin, particularly organs and tissues for transplantation, remains one of the most sensitive and least visible aspects of the transplantation chain. Although Brazil occupies a prominent position in the global transplantation milieu, avoidable losses still occur due to logistic weaknesses that compromise graft viability, system efficiency, and above all the lives of patients on the waiting list.

Notably, a proportion of organs offered for donation are not transplanted due to problems associated with their packaging and transport, such as a lack of temperature control, risk of freezing, or excessive warming.¹ This scenario becomes even more critical in the context of rising global temperatures linked to climate change, which increase the risk of exposure to elevated temperatures during transport, particularly over long distances or in regions with limited infrastructure. International reports issued by the World Health Organization and United Nations warn that global warming is already affecting the safety of healthcare supply chains, with the adoption of resilient strategies and secure pathways being required to preserve substances of human origin.^{2,3} These factors affect graft viability and also pose important ethical dilemmas: Each organ lost represents an opportunity denied to a patient on the waiting list for an organ or tissue.

Therefore, temperature variation during the transport of solid organs remains a major challenge, with evidence showing that both cold ischemia time and thermal stability along the route directly influence transplant outcomes.⁴

In lung transplantation, strict temperature control during graft preservation is crucial to reduce ischemic injury. Studies indicate that, compared with the conventional ice method, lungs maintained in ice-free preservation systems at a stable temperature between 6 °C and 8 °C can undergo prolonged periods of cold ischemia without worsening of clinical outcomes.⁵ Conversely, temperature fluctuations outside this range are associated with a high risk of primary graft dysfunction and reperfusion edema.

The possibility of extending safe cold ischemia time without increasing organ damage has a direct impact on transplant logistics. Such an extension would facilitate organ procurement from more distant centers, thus expanding the geographic reach of donation; it would also potentially contribute to an increase in the number of transplants performed. Furthermore, the extension of cold ischemia time would facilitate semi-elective scheduling of the transplantation, thereby optimizing intrahospital logistics, expanding the range of possible interventions such as intra-operative plasmapheresis in sensitized recipients, and providing greater predictability for recipient arrival and preparation.⁶

Initial studies in heart transplantation using temperature-controlled systems have shown promising results, enabling ischemia times to be lengthened without increasing graft dysfunction.⁷

A similar situation occurs in liver transplantation. Prolonged exposure to uncontrolled hypothermia increases the risk of ischemia-reperfusion injury. Normothermic *ex situ* perfusion has proven to be an effective strategy for mitigating the deleterious effects of temperature variations, with liver function being maintained even after prolonged periods of preservation.⁸

Therefore, traditional transport models based on insulated boxes with ice, without continuous monitoring, do not meet the requirements of a modern, quality-oriented system. Thermal variability is associated with a high risk of cell injury, poor posttransplant outcomes,^{4,9} and impaired traceability and management of adverse events.¹⁰

In this context, analyses of adverse events related to the donation–transplantation process must be strengthened, highlighting the role of biovigilance and quality management in ensuring the safety of substances of human origin.^{1,11} From an ethical and regulatory standpoint, the World Health Organization confirms that substances of human origin must be collected, stored, and transported under validated, safe, and traceable conditions.¹²

Therefore, adopting transport technologies with active thermal monitoring, that are aligned with international recommendations for the traceability and safety of substances of human origin, represents a technical advance as well as an ethical and public health requirement.¹³ Investing in innovation, scientific validation, and process standardization is essential to reduce losses, expand access to transplantation, and uphold the social and ethical value of donations.

REFERENCES

1. Roza BA, Paim SMS, Leite R, et al. Safe transport of organs and tissues for transplants: technological innovation product validation method. *Rev Assoc Med Bras.* 2023;69(6):e20221537. PMID: 37377283; <https://doi.org/10.1590/1806-9282.20221537>.
2. World Health Organization. Climate change and health: health sector resilience and supply chains. Geneva: WHO; 2023. Available from: <https://www.who.int/publications/i/item/climate-change-and-health-health-sector-resilience-and-supply-chains>. Accessed on 2026 (Feb. 26).
3. United Nations. Climate change and health: the critical role of global health systems. New York: UN; 2022. Available from: <https://www.un.org/en/climatechange/health>. Accessed on 2026 (Feb. 26).
4. Moers C, Smits JM, Maathuis MHJ, et al. Machine perfusion or cold storage in deceased-donor kidney transplantation. *N Engl J Med.* 2009;360(1):7-19. PMID: 19118301; <https://doi.org/10.1056/nejmoa0802289>.
5. Deitz R, Chan EG, Ryan J, et al. Adoption of a semi-elective lung transplantation practice by safely extending cold ischemic times. *J Thorac Cardiovasc Surg.* 2025;169(3):813-21. PMID: 39393628; <https://doi.org/10.1016/j.jtcvs.2024.09.038>.
6. Hoetzenecker K, Benazzo A, Schwarz S, Keshavjee S, Cypel M. The advent of semi-elective lung transplantation-prolonged static cold storage at 10°C. *Transpl Int.* 2024;37:12310. PMID: 38317690; <https://doi.org/10.3389/ti.2024.12310>.
7. Trahanas JM, Harris T, Petrovic M, et al. Out of the ice age: preservation of cardiac allografts with a reusable 10°C cooler. *JTCVS Open.* 2024;21:197-209. PMID: 39534335; <https://doi.org/10.1016/j.xjon.2024.08.005>.
8. Shamaa T, Crombez C, Konel J, et al. The use of normothermic liver preservation in combined liver and lung transplantation. *Am J Transplant.* 2022;22(9):2261-4. PMID: 35384271; <https://doi.org/10.1111/ajt.17053>.
9. Watson CJE, Dark JH. Organ transplantation: historical perspective and current practice. *Br J Anaesth.* 2012;18(Suppl 1):i29-42. PMID: 22194428; <https://doi.org/10.1093/bja/aer384>.
10. Guibert EE, Petrenko AY, Balaban CL, et al. Organ preservation: current concepts and new strategies for the next decade. *Transfus Med Hemother.* 2011;38(2):125-142. PMID: 21566713; <https://doi.org/10.1159/000327033>.
11. Roza BA, Lazarini LF, Treviso P, et al. Biosurveillance and reporting of adverse events in organ donation and transplantation: a systematic review. *Acta Paul Enferm.* 2023;36:eAPE00101. <http://dx.doi.org/10.37689/acta-ape/2023AR001011>.
12. World Health Organization. WHO guiding principles on human cell, tissue and organ transplantation. Geneva: WHO; 2010. Available from: <https://iris.who.int/handle/10665/341814>. Accessed on 2026 (Feb. 24).
13. Schuantes-Paim SM, Leite RF, Gonçalves VAC, et al. Static cold package for transporting organs for transplants: validation method and pilot test. *Sao Paulo Med J.* 2025;143(6):e20252930. PMID: 41172401; <https://doi.org/10.1590/1516-3180.2025.2930.29042025>.



Connecting care—telemedicine breaks barriers for low-income Brazilians battling breast cancer: a prospective cohort study

Gabriel Salum D'Alessandro^I, Yedda Nunes Reis^{II}, Mila Trementosa Garcia^{III}, Daniela Zarus Guimarães^{IV}, Roberta Amparado Mizziara^V, Flavia Colsomagno Silveira Berlofa^{VI}, Mayara de Cássia Benedito Andrade^{VII}, Emily Rie Yoshizato^{VIII}, Jonathan Yugo Maesaka^{IX}

Hospital Israelita Albert Einstein, São Paulo (SP), Brazil

^IMD, PhD. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0000-0003-0757-4106>

^{II}MD, MSc. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0000-0002-1928-6972>

^{III}MD, MSc. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0000-0003-4728-9684>

^{IV}MD. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0009-0008-2304-028X>

^VMD. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0000-0003-4762-7188>

^{VI}MD. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0009-0002-1351-1953>

^{VII}MD. Residency in Gynecology and Obstetrics, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0009-0004-4968-9643>

^{VIII}Undergraduate Student in Medicine, Albert Einstein Medical College, São Paulo (SP), Brazil.
<https://orcid.org/0009-0001-8583-3487>

^{IX}MD, PhD. Physician, Breast Surgery, Hospital Israelita Albert Einstein, São Paulo (SP), Brazil.
<https://orcid.org/0000-0002-1725-8732>

KEYWORDS (MeSH terms):

Telemedicine.
 Breast neoplasms.
 Developing countries.
 Healthcare disparities.
 Health equity.

AUTHOR'S KEYWORDS:

Breast cancer.
 Healthcare access.
 Low-income population.
 Tele-oncology.
 Patient satisfaction.
 Public health.

ABSTRACT

BACKGROUND: Telemedicine may help advance healthcare equity, particularly in low-income communities, by reducing economic burdens and mitigating geographic barriers to specialized care, including transportation costs, time away from work, and long travel distances to referral centers.

OBJECTIVES: This study aims to evaluate the impact of a remote consultation protocol on healthcare access among predominantly low-income patients with breast cancer in Brazil.

DESIGN AND SETTING: This prospective cohort study was conducted at a tertiary public cancer center and included 946 women with breast cancer.

METHODS: Sociodemographic characteristics, teleconsultation-related data, and patient satisfaction measured using a validated Likert-based questionnaire were collected. The study was approved by the local ethics committee, and all participants provided informed consent.

RESULTS: The mean age was 57.3 years. Overall, 51.6% of participants identified as white, and 64.4% reported a family income below two Brazilian minimum wages. Over one-third (34.3%) had not completed elementary school. Despite this socioeconomic vulnerability, 93.5% reported internet access, 90.3% owned a mobile phone, and 52.7% had an email address. During the study period, 946 appointments were scheduled and screened for teleconsultation eligibility. Among these, 229 patients met predefined eligibility criteria and were offered teleconsultation, and 117 completed the remote visit, corresponding to an uptake of 51.1% among eligible patients. Regarding hospital visits, 57% relied on public transportation, with a mean cost of \$ 4 per visit, representing a substantial financial burden. Travel time exceeded 90 minutes for 19.4% of patients. Satisfaction with teleconsultation was high, with a mean Likert score of 4.65/5 and a Net Promoter Score of 79.8.

CONCLUSION: Among patients eligible for telemedicine, more than half completed teleconsultation, indicating that implementation is feasible even in socially vulnerable populations with limited digital literacy. The remote care protocol improved access, reduced financial burden, and achieved high patient satisfaction, supporting telemedicine as a strategy to promote equity in breast cancer care in Brazil.

INTRODUCTION

Achieving equitable healthcare remains a major global challenge, particularly in predominantly low-income communities, where economic and geographic barriers impede access to medical services.¹ Telemedicine has emerged as a promising approach to address these inequities by enabling remote access to healthcare professionals and potentially improving the inclusivity and accessibility of care, especially for patients undergoing breast cancer treatment.²

As an important component of medical informatics, telemedicine uses information and communication technologies to deliver healthcare in a more cost-effective, efficient, and accessible manner, particularly in economically disadvantaged settings.^{3,4} Evidence indicates that telemedicine can be especially beneficial for vulnerable populations, including older adults, individuals with limited mobility, and ethnic minorities, thereby helping to reduce disparities in access to care.⁵

Equity in healthcare is grounded in the principle that individuals who face greater barriers to accessing services should receive additional support to achieve comparable access, treatment, and outcomes.⁶ In this context, strengthening healthcare delivery for socioeconomically disadvantaged populations is essential to improving health outcomes and advancing broader health system objectives.^{7,8}

In Brazil, breast cancer represents a major public health concern and disproportionately affects vulnerable populations with limited access to healthcare services. This limited access creates important challenges and potential risks to optimal patient management across a prolonged and complex treatment pathway.⁹ Telemedicine has emerged as a potential strategy to mitigate these difficulties by providing remote clinical support and reducing the financial and time burdens associated with in-person care.¹⁰

Territorial disparities in the distribution of healthcare services have long affected access to care in Brazil, reinforcing the need for context-specific strategies and policy solutions.^{11,12} In response to these challenges, some public healthcare facilities in Brazil have begun integrating telemedicine into the national public health system as a means of expanding access and addressing persistent inequities.^{1,13–15}

This study evaluated the impact of a telemedicine protocol for remote consultations on healthcare equity among a predominantly low-income population of patients with breast cancer treated at a Brazilian public hospital. Specifically, by implementing telephone-based telemedicine consultations, the study examined their effects on patient satisfaction and access to care.

OBJECTIVE

The primary objective was to assess patient satisfaction with remote care provided by the breast surgery team using a telemedicine satisfaction questionnaire. The secondary objectives included evaluating the socioeconomic profile of participating patients, identify the reasons for eligibility for remote care, and determine the proportion of scheduled outpatient appointments considered eligible for telemedicine.

METHODS

Study design and population

This prospective cohort study evaluated the implementation of telemedicine in the clinical practice of a breast surgery team at a municipal public hospital in São Paulo, Brazil, between July 2023 and April 2024. The study was approved by the institutional Research Ethics Committee (Approval No. CAAE 68518523.0.0000.0071) and was conducted in accordance with Resolution No. 466/12 of the Brazilian National Health Council. Written informed consent was obtained from all participants before enrollment. Consent was obtained during a prior in-person visit to enable subsequent teleconsultation. The study was reported in accordance with the STROBE guidelines for observational studies.¹⁶

Eligible participants were adult patients aged ≥ 18 years attending the breast surgery outpatient clinic. Although men were eligible for inclusion, no male patients were enrolled in the study.

As a public tertiary care service, the hospital did not provide screening consultations or second opinions. All participants had a confirmed diagnosis before referral and were managed within a predefined treatment pathway.

Patients undergoing neoadjuvant systemic therapy under shared follow-up by the clinical oncology and breast surgery teams were eligible for inclusion. During neoadjuvant treatment, surgical visits were scheduled for predefined purposes, including review of treatment response based on imaging findings and oncology reports, surgical planning, coordination of multidisciplinary care, and patient counseling. These visits did not involve physical examination, diagnostic assessment, or delivery of sensitive results. Accordingly, this neoadjuvant subgroup comprised established patients at intermediate to late stages of a predefined treatment pathway.

Exclusion criteria included any condition requiring in-person evaluation, such as clinical instability, need for physical examination (e.g., new breast complaints), pending diagnostic imaging or laboratory results requiring face-to-face discussion, or situations involving sensitive communication, such as suspected disease recurrence or disclosure of unfavorable diagnostic findings. Early postoperative follow-up visits were excluded by design because they routinely require physical examination. Patients attending their first consultation with the breast surgery team were also excluded. Additional exclusion criteria were significant cognitive impairment without a legal guardian and refusal to provide informed consent.

For participant recruitment, the outpatient schedule was reviewed weekly to identify patients eligible for remote care. Eligible patients were contacted by telephone at the same day and time as their originally scheduled in-person consultation and were offered teleconsultation. Following the teleconsultation, participants completed the Telemedicine Satisfaction Questionnaire (TSQ), which had been translated and validated for use in Brazil.¹⁷ This questionnaire comprises 14 items addressing multiple aspects of remote care.

Data collection and definitions

All data were collected, securely stored, and extracted using the REDCap electronic data capture platform.¹⁸ Anonymization procedures were rigorously applied to protect patient confidentiality.

Statistical analysis

Descriptive analyses were performed using absolute frequencies and percentages for categorical variables. Numerical variables were summarized using means and standard deviations, medians and quartiles, and minimum and maximum values, as appropriate. The distributions of numerical variables were assessed using histograms, boxplots, quantile comparison plots, and the Shapiro–Wilk normality test.

Differences in satisfaction scores between the first and second teleconsultations were assessed using the Wilcoxon signed-rank

test for paired data. Associations between categorical variables were examined using the chi-square test. Relationships between patients' sociodemographic characteristics and scores on the Telemedicine Patient Satisfaction Assessment Questionnaire were evaluated using the Mann-Whitney test or Kruskal-Wallis test for categorical variables and Spearman correlation coefficients for numerical variables.

All analyses were performed using IBM SPSS Statistics for Windows version 29.0.2022 (IBM Corp., Armonk, New York), with a two-sided significance level of 5%.

RESULTS

During the study period, a total of 946 patients were scheduled for appointments at the breast surgery clinic. All scheduled visits were screened for teleconsultation eligibility according to pre-defined criteria. Of these, 229 patients met the eligibility criteria and were offered teleconsultation. Among the eligible patients, 117 (51.1%) completed the teleconsultation, whereas 112 did not. Reasons for non-completion included failure of the back-office team to establish contact ($n = 18$), unanswered calls ($n = 15$), patient preference for an in-person consultation ($n = 70$), and other reasons ($n = 9$) (Figure 1).

Over the 43-week study period, 117 teleconsultations were conducted. Of the patients who received remote care, 71 had one teleconsultation, 20 had two, and two had three teleconsultations. Among the 22 patients who received two remote consultations, the interval between appointments ranged from 14 to 301 days, with a median of 46 days. For the two patients who underwent three teleconsultations, the interval between the second and third consultations was 70 days for one patient, and 126 days for the other.

The main reasons for referral to remote care were the unavailability of a relevant examination required for decision-making

(40.2%) and ongoing neoadjuvant chemotherapy with active follow-up by the clinical oncology team (29.9%).

Based on the first teleconsultation for each patient, the socio-demographic characteristics and teleconsultation-related data of 93 patients are presented in Tables 1 and 2.

Table 1. Socioeconomic characteristics ($n = 93$)

Variable	Category	n (%)
Age (years)	Mean (SD)	57.3 (13.8)
	Median (Q1; Q3)	57 (46; 67)
	Minimum; maximum	29; 86
Ethnicity	White	50 (53.8%)
	Black	9 (9.7%)
	Brown	31 (33.3%)
	Asian	3 (3.2%)
Marital Status	Single	23 (24.7%)
	Married	33 (35.5%)
	Divorced	14 (15.1%)
	Widowed	11 (11.8%)
	Stable union/Cohabiting	12 (12.9%)
Education Level	No formal education	6 (6.5%)
	Incomplete elementary education	32 (34.4%)
	Complete elementary education	16 (17.2%)
	Incomplete high school education	6 (6.5%)
	Complete high school education	21 (22.6%)
	Incomplete higher education	3 (3.2%)
	Complete higher education	9 (9.7%)
Monthly Family Income*	Up to two minimum wages	62 (66.7%)
	More than two minimum wages	31 (33.3%)
Work Status	Unemployed	36 (38.7%)
	Formal/informal work	21 (22.6%)
	Government assistance income	36 (38.7%)

SD, standard deviation; Q1, first quartile; Q3, third quartile; * During the study period (July 2023–April 2024), the national minimum wage was R\$ 1,320 in 2023 and R\$ 1,412 in 2024 (approximately \$ 250–270 per month, depending on exchange rates).

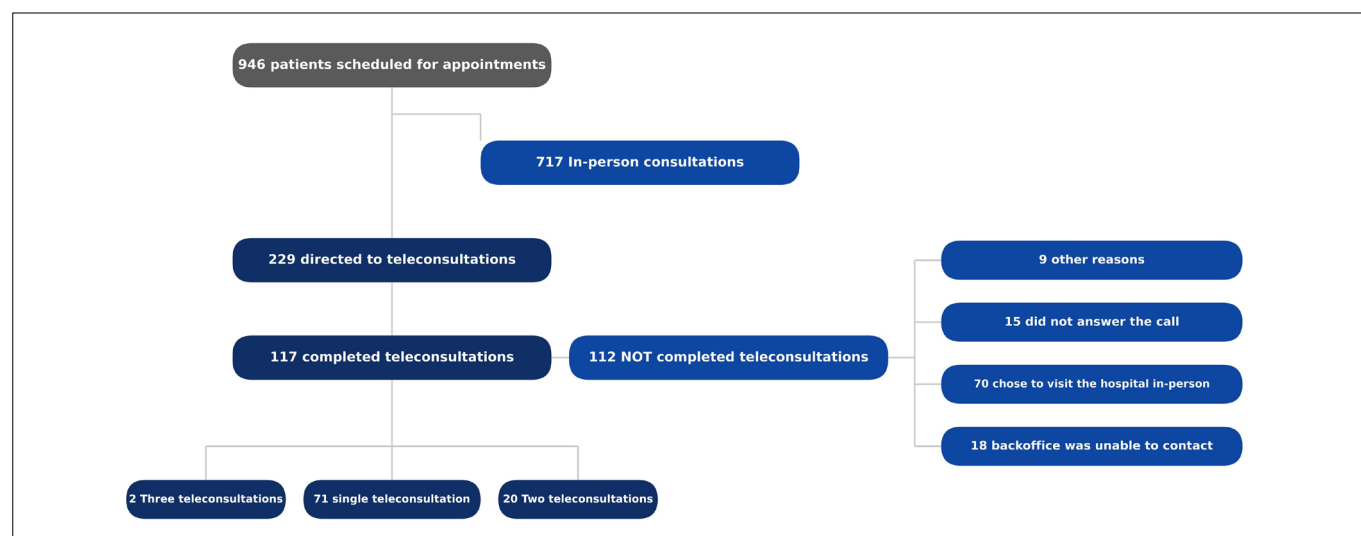


Figure 1. Flow-chart of the study.

Table 2. Teleconsultation-related data (n = 93)

Variable	Category	n (%)
Teleconsultation prevented workday loss*	No	13 (61.9%)
	Yes	8 (38.1%)
Internet access	No	6 (6.5%)
	Yes	87 (93.5%)
Cell phone ownership	No	9 (9.7%)
	Yes	84 (90.3%)
Email ownership	No	44 (47.3%)
	Yes	49 (52.7%)
Distance from residence to hospital (km)	Mean (SD)	12.8 (7.9)
	Median (Q1; Q3)	13 (5.6; 18)
	Minimum; maximum	0.6; 39
Mode of transport to the hospital	Public transport	53 (57%)
	Private car	28 (30.1%)
	Taxi/app-based driver	11 (11.8%)
	On foot	1 (1.1%)
Transportation cost (round-trip)**	Up to \$ 2	33 (35.5%)
	\$ 2–4	25 (26.9%)
	\$ 4–10	20 (21.5%)
	More tan \$ 10	15 (16.1%)
Travel time (one way) to the hospital	Up to 30 minutes	18 (19.4%)
	30–45 minutes	16 (17.2%)
	45 minutes–1 hour	16 (17.2%)
	1–1.5 hours	17 (18.3%)
	> 1.5 hours	26 (28%)
Attends consultations with a companion	Alone	20 (21.5%)
	With a companion	73 (78.5%)
Reason for referral to teleconsultation	Unavailable exam***	47 (40.2%)
	Undergoing neoadjuvant chemotherapy	35 (29.9%)
	Unperformed	15 (12.8%)
	interconsultation	15 (12.8%)
	Only for exam request	9 (7.7%)
	Check test results	8 (6.8%)
	Incorrect outpatient scheduling	1 (0.9%)
	Other	2 (1.7%)

SD, standard deviation; Q1, first quartile; Q3, third quartile; * Only those employed (n = 21); ** At the time of analysis, during July 2024, the US dollar exchange rate fluctuated between approximately R\$ 5.40 and R\$ 5.50; *** e.g., missed imaging appointment, delayed results, test not performed, or results unavailable in the system at the time of the visit.

The mean age was 57.3 years. Most patients were white, married, and had not completed elementary school (**Table 1**). Overall, 66.7% reported a monthly family income of up to two minimum wages, 38.7% were unemployed, and another 38.7% were dependent on government assistance. Among patients engaged in formal or informal employment, 38.1% reported avoiding the loss of a workday because of remote care (**Tables 1 and 2**).

Regarding digital connectivity, 93.5% reported having internet access, 90.3% owned a mobile phone, and 52.7% had a personal email account (**Table 2**).

The mean distance from the patient's residence to the hospital was 12.8 km. Most patients relied on public transportation, spent up

to \$ 2 to reach the hospital, and required over 90 minutes for travel, often accompanied by another person during consultations (**Table 2**).

The Patient Satisfaction Assessment Questionnaire for Telemedicine was completed in 94 of the 117 teleconsultations (80.3%). Among these, 76 questionnaires were completed after the first teleconsultation, 16 after the second, and two after the third. TSQ scores ranged from 2.4 to 5.0, with a median of 4.8 (**Tables 3, 4, and Figure 2**).

Based on the isolated response to question 14, "I would recommend telemedicine to other patients," the Net Promoter Score (NPS) was calculated to assess satisfaction with teleconsultation. The NPS was defined as the percentage of promoters minus the percentage of detractors, with scores between 75% and 100% generally considered excellent. Among the 94 questionnaires, 3.2% of respondents were classified as detractors and 83% as promoters, resulting in an NPS of 79.8%.

Patients who underwent more than one teleconsultation could complete the TSQ after each appointment. Fifteen patients completed the questionnaire twice, including the two patients who underwent three teleconsultations and responded after each visit. When satisfaction scores from the first and second teleconsultations were compared, seven patients had higher scores after the second consultation, five had lower scores, and three had unchanged scores (**Figure 3**). After the first teleconsultation, scores ranged from 3.86 to 5.00, with a median of 4.71 (first quartile, 4.29; third quartile, 5). After the second teleconsultation, scores ranged from 4.07 to 5.00, with a median of 4.93 (first quartile, 4.64; third quartile, 5). No significant difference was observed between the first and second teleconsultation scores ($p = 0.254$).

Additionally, no significant differences were found in the distribution of TSQ scores according to age group ($p = 0.861$), educational level ($p = 0.530$), monthly family income ($p = 0.242$), employment status ($p = 0.832$), transportation cost to the hospital ($p = 0.762$), or travel time to the hospital ($p = 0.287$). In addition, weak and non-significant correlations were observed between TSQ score and patient age (correlation coefficient = 0.024; $p = 0.839$) and between TSQ score and distance from residence to hospital (correlation coefficient = 0.123; $p = 0.291$) (**Table 5**).

DISCUSSION

This study evaluated the sociodemographic profile of patients, the feasibility of teleoncology implementation, and patient satisfaction with remote care in a socially vulnerable population of patients with breast cancer treated within the public health system. The findings indicate that telemedicine can be incorporated into routine breast cancer care with high patient satisfaction and meaningful practical benefits, including reduced travel-related costs and less time spent attending consultations. Notably, 83% of patients reported that they would recommend telemedicine to others.

Table 3. Distribution of patient responses on the TSQ instrument regarding teleconsultations (n = 94)

Questions	Responses				
	No, definitely not	Probably not	Maybe	Probably yes	Yes, for sure
Q1. Telemedicine is better than I expected.	2 (2.1%)	1 (1.1%)	3 (3.2%)	12 (12.8%)	76 (80.9%)
Q2. I am satisfied with my telemedicine care.	0 (0%)	1 (1.1%)	3 (3.2%)	12 (12.8%)	78 (83%)
Q3. I was worried about my privacy during the telemedicine consultation.#	76 (80.9%)	7 (7.4%)	2 (2.1%)	3 (3.2%)	6 (6.4%)
Q4. The care I received by telemedicine was as good as in-person care.	3 (3.2%)	4 (4.3%)	11 (11.7%)	19 (20.2%)	57 (60.6%)
Q5. Telemedicine saved me travel time.	0 (0%)	0 (0%)	2 (2.1%)	7 (7.4%)	85 (90.4%)
Q6. Telemedicine saved me money.	0 (0%)	0 (0%)	3 (3.2%)	7 (7.4%)	84 (89.4%)
Q7. I felt comfortable talking to the doctor on the phone.	1 (1.1%)	1 (1.1%)	4 (4.3%)	2 (2.1%)	86 (91.5%)
Q8. I felt that my care was complete.	1 (1.1%)	2 (2.1%)	5 (5.3%)	17 (18.1%)	69 (73.4%)
Q9. I prefer spending more time traveling for my next appointment in person rather than telemedicine.#	39 (41.5%)	23 (24.5%)	20 (21.3%)	4 (4.3%)	8 (8.5%)
Q10. I had difficulty hearing the doctor on the phone.#	83 (88.3%)	4 (4.3%)	4 (4.3%)	2 (2.1%)	1 (1.1%)
Q11. I was able to build a trusting relationship with the doctor.	0 (0%)	1 (1.1%)	6 (6.4%)	13 (13.8%)	74 (78.7%)
Q12. I was able to explain my problems clearly to the doctor during telemedicine care.	0 (0%)	2 (2.1%)	3 (3.2%)	15 (16%)	74 (78.7%)
Q13. Telemedicine care was good for me.	0 (0%)	1 (1.1%)	1 (1.1%)	12 (12.8%)	80 (85.1%)
Q14. I would recommend telemedicine care to other patients (NPS).	0 (0%)	1 (1.1%)	2 (2.1%)	13 (13.8%)	78 (83%)

Questions with negatively worded statements.

Table 4. Description of Results from the Application of the Patient TSQ for Telemedicine Consultations (n=94)

Questions from the instrument	Mean (SD)
Q1	4.69 (0.78)
Q2	4.78 (0.55)
Q3	4.53 (1.12)
Q4	4.31 (1.05)
Q5	4.88 (0.38)
Q6	4.86 (0.43)
Q7	4.82 (0.66)
Q8	4.61 (0.78)
Q9	3.86 (1.25)
Q10	4.77 (0.72)
Q11	4.70 (0.64)
Q12	4.71 (0.63)
Q13	4.82 (0.49)
Q14	4.79 (0.53)
Instrument score	
Mean (SD)	4.65 (0.43)
Median (Q1; Q3)	4.79 (4.43; 5)
Minimum; maximum	2.36; 5

SD: Standard deviation; Q1: First quartile; Q3: Third quartile

To our knowledge, this is the first prospective study conducted in Brazil outside the pandemic setting to examine telemedicine implementation in the clinical management of patients with breast cancer while also providing a sociodemographic characterization of the population and a structured evaluation of teleconsultation quality using a validated instrument.

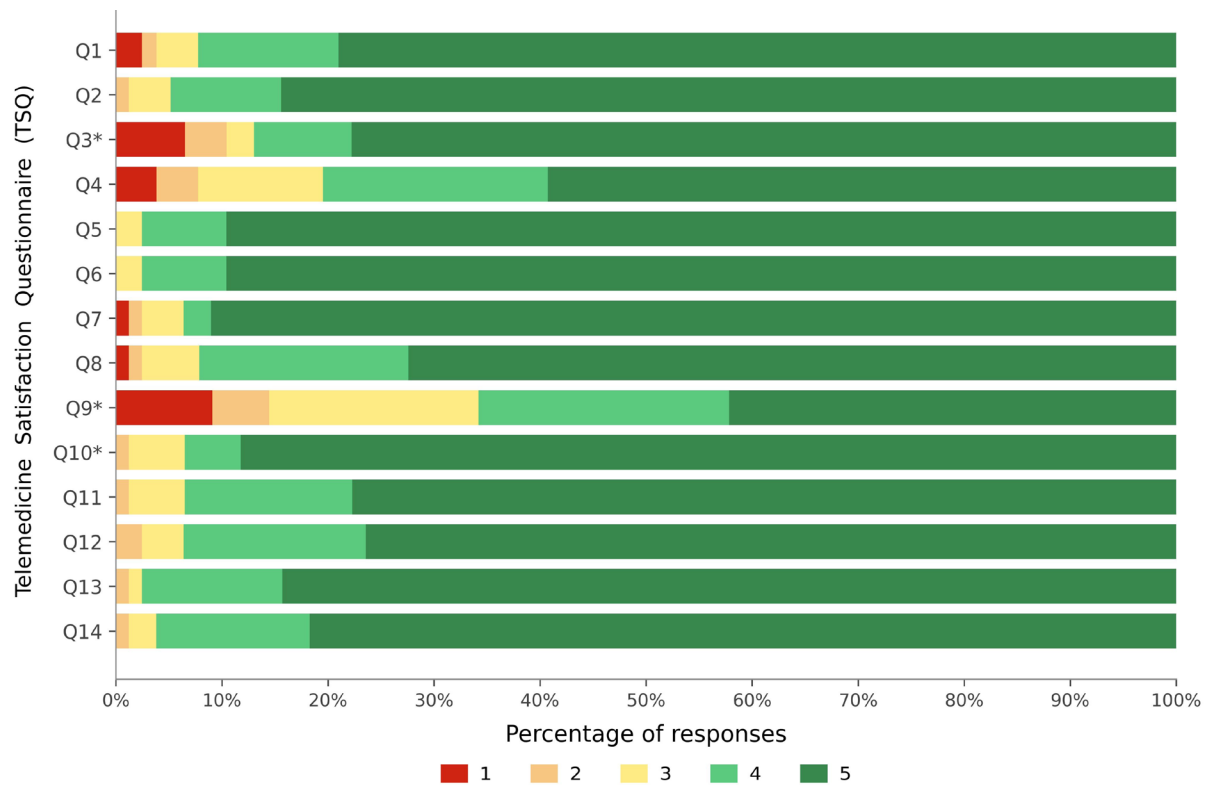
The sociodemographic characteristics of this cohort are consistent with those reported for predominantly low-income patients

with breast cancer in Brazil, a population that often faces substantial economic and social challenges.¹⁹ This population often relies on the public healthcare system and experiences barriers to accessing oncology care, including transportation difficulties, limited financial resources, and competing priorities, such as work and family responsibilities.^{20,21}

A central aim of this study was to evaluate the potential advantages of remote care, including convenience, safety, and effectiveness in meeting the needs of each clinical case. In particular, the study examined whether teleoncology could help overcome barriers associated with in-person visits, including transportation demands, financial costs, time constraints, and work absenteeism, all of which are particularly relevant for socioeconomically vulnerable patients undergoing breast cancer treatment.

The findings underscore the burden imposed by in-person care in this population. Most patients (66.7%) reported a monthly family income of up to two minimum wages, 57% relied on public transportation, 35.5% spent up to \$ 2 to reach the hospital, and 28% required over 90 minutes of travel time. These data illustrate the logistical and financial challenges associated with routine face-to-face consultations. In this context, the studies support the role of telemedicine in reducing both direct and indirect costs related to cancer care, including transportation expenses, lodging, and income lost because of missed work.²²

This practical benefit was also reflected in our findings. Among patients with formal or informal employment, 38.1% reported avoiding the loss of a workday because of teleconsultation. This suggests that remote care may meaningfully affect both financial stability and



* Questions with negatively worded statements.

Figure 2. Distribution of scores for questions answered by patients on the TSQ instrument (n = 94).

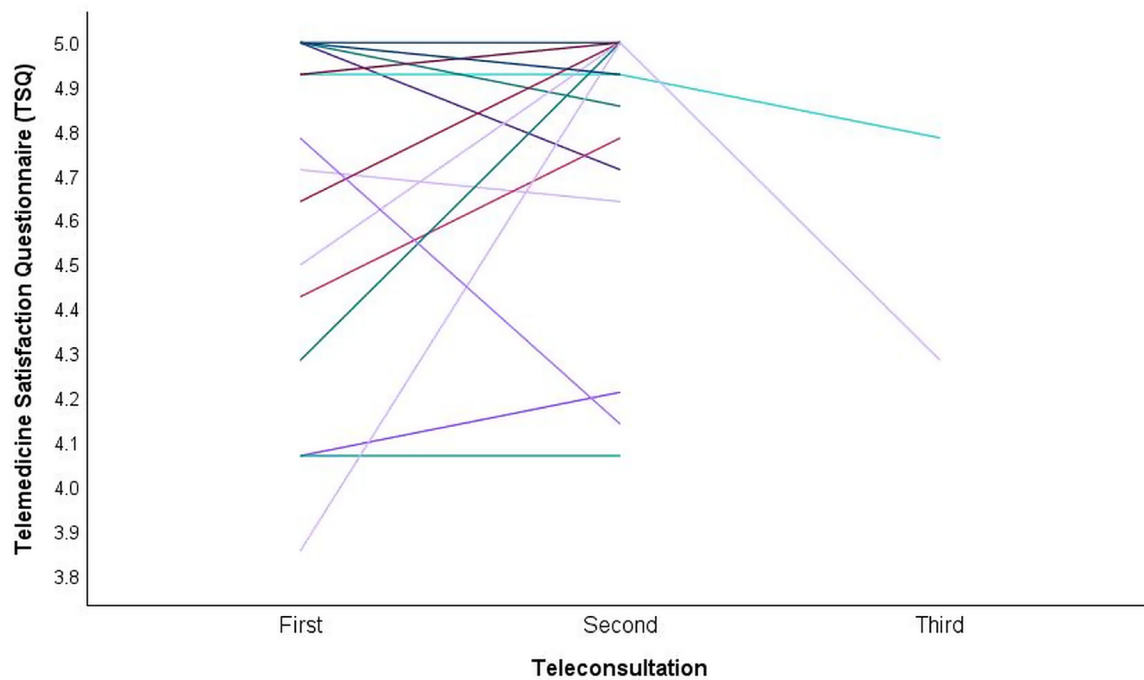


Figure 3. Individual TSQ scores of patients who underwent more than one teleconsultation (n = 15).

Table 5. Associations between sociodemographic data and TSQ after the first teleconsultation (n = 76)

Variable	Median (Q1; Q3)	Min; max	n	p value
Age (years)				
≤ 39	4.50 (4.14; 5)	3.93; 5	7	0.861 ^b
40–49	4.93 (4.21; 5)	2.36; 5	15	
50–60	4.79 (4.57; 4.93)	3.86; 5	27	
≥ 61	4.79 (4.43; 5)	3.64; 5	27	
Education				
No schooling to primary	4.79 (4.54; 5)	3.79; 5	40	0.53 ^b
Incomplete/complete high school	4.79 (4.14; 5)	3.64; 5	25	
Incomplete/complete higher education	4.71 (4.43; 4.93)	2.36; 4.93	11	
Monthly family income (IBGE)				
Up to 2 minimum wages	4.79 (4.25; 4.96)	2.36; 5	48	0.242 ^a
More than 2 minimum wages	4.86 (4.57; 4.96)	3.86; 5	28	
Employment status				
Unemployed	4.79 (4.43; 4.93)	3.93; 5	30	0.832 ^b
Formal/informal employment	4.79 (4.43; 5)	3.86; 5	19	
Government assistance	4.86 (4.57; 5)	2.36; 5	27	
Transport cost to hospital				
Up to \$ 2	4.79 (4.21; 4.93)	2.36; 5	26	0.762 ^b
\$ 2–4	4.82 (4.64; 4.93)	4.07; 5	20	
\$ 4–10	4.79 (4.14; 5)	3.79; 5	17	
More than \$ 10	4.86 (4.50; 5)	3.64; 5	13	
Travel Time to Hospital				
< 30 minutes	4.79 (4.43; 4.93)	3.93; 5	17	0.287 ^b
30–45 minutes	4.57 (4.43; 4.79)	3.64; 5	13	
45 minutes–1 hour	4.93 (4.71; 5)	3.79; 5	14	
1–1.5 hours	4.86 (4.43; 4.96)	3.86; 5	12	
> 1.5 hours	4.75 (4.32; 4.96)	2.36; 5	20	

Q1, first quartile; Q3, third quartile; n, number of patients; ^a Mann–Whitney test; ^b Kruskal–Wallis test.

daily functioning in this vulnerable population. In addition, a substantial proportion of patients (38.7%) depended on government assistance, including welfare or disability benefits, further emphasizing the economic fragility of the study population and the value of care models that reduce the burden of in-person visits.

Contrary to the assumption that socially vulnerable populations may have limited access to digital technologies, most patients in this cohort reported having the basic means required to participate in teleconsultations, including internet access and mobile phones. These findings suggest that technological access may be less restrictive than often presumed, even in low-income settings. Previous studies similarly indicate that telemedicine can be a feasible and effective approach for delivering oncology care to vulnerable populations.^{20,21,23,24}

Most patients reported a positive experience with teleconsultations, achieving an NPS score of 79.8%, which is classified as excellent. Other studies suggest that the average NPS for in-person healthcare typically falls between 60% and 80%, while telehealth appointments generally range from 70% to 90%.^{25,26} Furthermore, repeated use of teleconsultations did not lead to a decline in patient satisfaction, and some patients even reported increased satisfaction

over time.²⁴ Thus, the high satisfaction scores observed in this study suggest that teleconsultation can be a viable and well-received approach for delivering oncology care to predominantly low-income patients with breast cancer.

Additionally, most patients were satisfied with the telehealth experience, with a median TSQ score of 4.8 out of 5. These findings are consistent with previous reports showing high patient satisfaction with telehealth, particularly regarding convenience, reduced travel time, and improved access to care.²⁴ Other studies have reported similarly high satisfaction with teleoncology, particularly during the COVID-19 pandemic when in-person visits were restricted.^{21,23} Patient satisfaction remained high even after multiple teleconsultations, with no significant differences in satisfaction scores between the first and second appointments.

Reported drivers of telehealth satisfaction commonly include convenience, lower cost, ease of use, effective communication, and elimination of travel time,^{27,27} all of which align with the experience reported by patients in the present study.

Importantly, repeated exposure to teleconsultation did not appear to reduce patient satisfaction. Satisfaction scores remained high across successive consultations, and some patients reported

higher scores after subsequent remote visits. Although no significant difference was observed between the first and second teleconsultations, these findings suggest that teleoncology remains acceptable over time for patients appropriately selected for remote follow-up.²⁴

At the same time, telemedicine was not universally accepted. Among clinically eligible patients, only 51.1% completed teleconsultation, and patient preference for in-person care was the most common reason for non-completion. Inability to establish contact also contributed to non-completion in a smaller proportion of cases. Compared with a prior study conducted by our group during the COVID-19 pandemic, uptake in the present study was lower (51.1% vs. 71%),²¹ suggesting that social isolation may have increased acceptance of remote care during the pandemic period.

Sonagli et al. conducted a retrospective study at a tertiary cancer center in São Paulo during the COVID-19 pandemic, and showed that telemedicine was adopted to mitigate an almost 50% reduction in outpatient visits and was feasible for follow-up breast cancer, screening, and selected benign breast conditions.²⁸

The preference for in-person visits among a substantial subgroup of clinically eligible patients indicates that telemedicine is not universally accepted, even when remote care is clinically appropriate. This finding underscores the need for patient education regarding the scope and benefits of teleconsultation, while ensuring that clinical management is not compromised. Refusal of teleconsultation may also reflect diverse barriers, including the perceived value of physical examination, communication preferences, privacy concerns, mistrust of remote care, technological constraints, or perceived clinical seriousness. Future studies should incorporate targeted surveys of patients who decline teleconsultation to better characterize these barriers and inform patient-centered implementation strategies.

Collectively, teleoncology can be a viable option for delivering care to predominantly low-income and low-literacy patients with breast cancer, particularly when the aim is to reduce barriers related to travel, time, cost, and work absenteeism. However, its role is best understood within a selective and clinically appropriate model of care in which telemedicine complements, rather than replaces, in-person consultations.

This study has some limitations. It was conducted at a single institution within a specific geographic region, which may limit the generalizability of the findings. The absence of a control group receiving in-person care also restricted the ability to compare outcomes more rigorously. Although the sample of 93 patients provided important insight into patient characteristics and experiences with teleoncology, a larger sample might have offered greater statistical power to detect subtler associations across various outcome measures. Additionally, the lack of access to video-based consultations at our institution may have influenced the quality of the consultations.

Future studies should examine the long-term effects of teleoncology on clinical outcomes, treatment adherence, and overall quality of life among socioeconomically vulnerable patients with breast cancer. Further studies are warranted to better understand the barriers and facilitators influencing teleoncology adoption in this population and to identify strategies for sustainable implementation and scale-up within public health systems.

CONCLUSION

This study shows that teleoncology is a feasible and acceptable strategy for follow-up care in a clinically selected subgroup of predominantly low-income patients with breast cancer. Among patients considered eligible for teleconsultation, more than half completed the remote visit, indicating meaningful uptake in an appropriate clinical context. Although a substantial proportion of eligible patients preferred in-person care, these findings support integrating telemedicine alongside conventional consultations to reduce sociodemographic barriers and promote more equitable access to breast cancer management.

REFERENCES

1. Matheus R, Ribeiro MM. Telemedicine in Brazilian public policy management. ICEGOV 2009: Proceedings of the 3rd international conference on Theory and practice of electronic governance. 2009;1:274. <https://doi.org/10.1145/1693042.1693098>.
2. Bayard S, Fasano G, Gillot T, et al. Breast cancer disparities and the digital divide. *Curr Breast Cancer Rep*. 2022;14(4):205–12. PMID: 36467667; <https://doi.org/10.1007/s12609-022-00468-w>.
3. Hassibian MR, Hassibian S. Telemedicine acceptance and implementation in developing countries: benefits, categories, and barriers. *Razavi International Journal of Medicine* 2016;4. <https://doi.org/10.17795/rijm38332>.
4. Gupta PP, Khatoon S, Khatiwada N, Mishra A, Bhatta N, Geissbuhler A. Implementation, effectiveness and monitoring of telemedicine program in Bhutanese refugees camp in Eastern Nepal. *J Family Med Prim Care*. 2022 Jan;11(1):256–9. PMID: 35309651; https://doi.org/10.4103/jfmpc.jfmpc_1801_21.
5. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth technology use and implications in historically underserved and minority populations in the United States: systematic literature review. *JMIR Mhealth Uhealth*. 2018 Jun 18;6(6):e128. PMID: 29914860; <https://doi.org/10.2196/mhealth.8383>.
6. Waidmann TA, Rajan S. Race and ethnic disparities in health care access and utilization: an examination of state variation. *Med Care Res Rev*. 2000;57 Suppl 1:55–84. PMID: 11092158; <https://doi.org/10.1177/1077558700057001504>.
7. Bitton A, Ratcliffe HL, Veillard JH, et al. Primary health care as a foundation for strengthening health systems in low- and middle-income countries. *J Gen Intern Med*. 2017 May;32(5):566–71. PMID: 27943038; <https://doi.org/10.1007/s11606-016-3898-5>.

8. Kruk ME, Porignon D, Rockers PC, Van Lerberghe W. The contribution of primary care to health and health systems in low- and middle-income countries: a critical review of major primary care initiatives. *Soc Sci Med*. 2010 Mar;70(6):904–11. PMID: 20089341; <https://doi.org/10.1016/j.socscimed.2009.11.025>.
9. Azeredo-da-Silva AF, Zanotto BS, Martins F, Navarro N, Alencar R, Medeiros C. Health care accessibility and mobility in breast cancer: a Latin American perspective. *BMC Health Serv Res*. 2024 Jun 25;24(1):764. PMID: 38918823; <https://doi.org/10.1186/s12913-024-11222-6>.
10. Reis FJJ, Fernandes LG, Saragiotto BT. Telehealth in low- and middle-income countries: bridging the gap or exposing health disparities? *Health Policy Technol*. 2021 Dec;10(4):100577. PMID: 34745854; <https://doi.org/10.1016/j.hlpt.2021.100577>.
11. Pereira RHM, Braga CKV, Servo LM, et al. Geographic access to COVID-19 healthcare in Brazil using a balanced float catchment area approach. *Soc Sci Med*. 2021 Mar;273:113773. PMID: 33609968; <https://doi.org/10.1016/j.socscimed.2021.113773>.
12. Fonseca BP, Albuquerque PC, Saldanha RF, Zicker F. Geographic accessibility to cancer treatment in Brazil: a network analysis. *Lancet Reg Health Am*. 2021 Dec 23;7:100153. PMID: 36777653; <https://doi.org/10.1016/j.lana.2021.100153>.
13. Gundim RS, Chao WL. A graphical representation model for telemedicine and telehealth center sustainability. *Telemed J E Health*. 2011 Apr;17(3):164–8. PMID: 21375411; <https://doi.org/10.1089/tmj.2010.0064>.
14. Oliveira GL, Cardoso CS, Ribeiro ALP, Caiaffa WT. Satisfação de médicos com o cuidado às doenças cardiovasculares em municípios de Minas Gerais: Escala Cardiosatis-EQUIPE. *Rev Bras Epidemiol*. 2011;14:240–52. <https://doi.org/10.1590/S1415-790X2011000200006>.
15. Santos AF, Mata-Machado ATGD, Melo MDCB, Fonseca D Sobrinho, Araújo LL, Silva ÉA, Lima AMLD, Abreu DMX, Rocha HAD. Implementation of telehealth resources in primary care in Brazil and its association with quality of care. *Telemed J E Health*. 2019 Oct;25(10):996–1004. PMID: 30592699; <https://doi.org/10.1089/tmj.2018.0166>.
16. Von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007 Oct 20;335(7624):806–8. PMID: 17947786; <https://doi.org/10.1136/bmj.39335.541782.AD>.
17. Moreira TC, Constant HM, Faria AG, et al. Tradução, adaptação transcultural e validação de questionário de satisfação em telemedicina. *Revista Brasileira de Medicina de Família e Comunidade*. 2022 Mar 4;17(44):2837. [http://doi.org/10.5712/rbmfc17\(44\)2837](http://doi.org/10.5712/rbmfc17(44)2837).
18. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap): a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009 Apr;42(2):377–81. PMID: 18929686; <https://doi.org/10.1016/j.jbi.2008.08.010>.
19. Caleffi M, Crivelatti I, Burchardt NA, et al. Breast cancer survival in Brazil: how much health care access impact on cancer outcomes? *Breast*. 2020 Dec;54:155–9. PMID: 33120081; <https://doi.org/10.1016/j.breast.2020.10.001>.
20. Carneiro AC, Pinho GS, Belo JV, et al. Outcomes of telemedicine care during the COVID-19 pandemic: experience from an intervention program designed for vulnerable population in Brazil. *J Telemed Telecare*. 2024 May;30(4):715–21. PMID: 35321612; <https://doi.org/10.1177/1357633X221089151>.
21. Miziara RA, Maesaka JY, Matsumoto DRM, et al. Teleoncology orientation of low-income breast cancer patients during the COVID-19 pandemic: feasibility and patient satisfaction. *Rev Bras Ginecol Obstet*. 2021 Nov;43(11):840–6. PMID: 34872142; <https://doi.org/10.1055/s-0041-1739425>.
22. Patel KB, Turner K, Alishahi Tabriz A, et al. Estimated indirect cost savings of using telehealth among nonelderly patients with cancer. *JAMA Netw Open*. 2023 Jan 3;6(1):e2250211. PMID: 36626174; <https://doi.org/10.1001/jamanetworkopen.2022.50211>.
23. Yildiz F, Oksuzoglu B. Teleoncology or telemedicine for oncology patients during the COVID-19 pandemic: the new normal for breast cancer survivors? *Future Oncol*. 2020 Oct;16(28):2191–5. PMID: 32857603; <https://doi.org/10.2217/fon-2020-0714>.
24. Polinski JM, Barker T, Gagliano N, Sussman A, Brennan TA, Shrank WH. Patients' satisfaction with and preference for telehealth visits. *J Gen Intern Med*. 2016 Mar;31(3):269–75. PMID: 26269131; <https://doi.org/10.1007/s11606-015-3489-x>.
25. Gustke SS, Balch DC, West VL, Rogers LO. Patient satisfaction with telemedicine. *Telemedicine Journal*. 2000 May;6(1):5–13. <https://doi.org/10.1089/107830200311806>.
26. Orlando JF, Beard M, Kumar S. Systematic review of patient and caregivers' satisfaction with telehealth videoconferencing as a mode of service delivery in managing patients' health. *PLoS One*. 2019 Aug 30;14(8):e0221848. PMID: 31469865; <https://doi.org/10.1371/journal.pone.0221848>.
27. Shaver J. The state of telehealth before and after the COVID-19 pandemic. *Prim Care*. 2022 Dec;49(4):517–30. PMID: 36357058; <https://doi.org/10.1016/j.pop.2022.04.002>.
28. Sonagli M, Cagnacci R Neto, Leite FPM, Makdissi FBA. The use of telemedicine to maintain breast cancer follow-up and surveillance during the COVID-19 pandemic. *J Surg Oncol*. 2021 Feb;123(2):371–4. PMID: 33333581; <https://doi.org/10.1002/jso.26327>.

Authors' contributions: D'Alessandro GS: conceptualization (equal), data curation (equal), methodology (equal), writing – original draft, review and editing; Reis YN: investigation (equal); Garcia MT: investigation (equal); Guimarães DZ: investigation (equal); Miziara RA: investigation (equal); Berlofa FCS: investigation (equal); Andrade MCB: investigation (equal); Yoshizato ER: investigation (equal); Maesaka JY: conceptualization (equal), data curation (equal),

methodology (equal), validation. All authors reviewed and approved the final version submitted for publication.

Acknowledgments: The authors would like to acknowledge the technical and clinical support provided by the nursing staff, administrative back-office team, statisticians, and medical residents, whose contributions were essential to this study.

Events or congresses: This study was presented at the 2024 San Antonio Breast Cancer Symposium.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: The datasets generated and/or analyzed during the current study are not publicly available due to patient privacy and institutional restrictions but are available from the corresponding author upon reasonable request.

Declaration of generative AI in scientific writing: The authors used generative artificial intelligence tools solely for language refinement. All content, interpretations, and conclusions are the responsibility of the authors.

Date of first submission: April 23, 2025

Last received: February 25, 2026

Accepted: March 9, 2026

Address for correspondence:

Gabriel S. D'Alessandro, MD, PhD.
Hospital Municipal Vila Santa Catarina "Dr. Gilson de Cassia Marques de Carvalho"/Hospital Israelita Albert Einstein
Av. Santa Catarina, 2785
Vila Santa Catarina — São Paulo (SP) — Brasil
CEP 04378-500
Tel. (+55 11) 2151-5654
Email: gabriel.dalessandro@einstein.br

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Translation, cultural adaptation, and evidence of content and face validity of a questionnaire to assess REDs knowledge in Brazil: a cross-cultural validation study

Uyara Pereira de Maria^I, Claudia Ridel Juzwiak^{II}

Universidade Federal de São Paulo (Unifesp), Santos (SP), Brazil

^IMSc. Doctoral Student, Interdisciplinary Postgraduate Program in Health Sciences, Universidade Federal de São Paulo (Unifesp), Santos (SP), Brazil.

 <https://orcid.org/0000-0002-6903-073X>

^{II}PhD. Professor, Department of Human Movement Sciences, Universidade Federal de São Paulo (Unifesp), Santos (SP), Brazil.

 <https://orcid.org/0000-0003-1101-0063>

KEYWORDS (MeSH terms):

Validation study.
Athletes.
Female athlete triad syndrome.

AUTHOR'S KEYWORDS:

Validation studies.
Translating.
Cross-cultural adaptation.
Athlete.
Relative energy deficiency in sport.
Knowledge.

ABSTRACT

BACKGROUND: Relative energy deficiency in sport (REDs) is a syndrome that impairs physiological function because of low energy availability. Awareness and knowledge of REDs among athletes and health professionals are essential for early identification and prevention.

OBJECTIVE: To translate into Brazilian Portuguese and culturally adapt a questionnaire that assesses knowledge of the signs, symptoms, and consequences of REDs, originally created by Pai et al., and to obtain evidence of content and face validity.

DESIGN AND SETTING: Translation, cultural adaptation, and validation study conducted remotely using online tools and digital data collection.

METHODS: The translation and cultural adaptation consisted of the following steps: two independent translations, synthesis of the translations by a committee of health professionals, two back-translations, and review by a committee of experts to determine the best version for testing. For content validity, the content validity coefficient (CVC) was calculated based on the evaluation of the experts committee on the practical pertinence, theoretical relevance, and linguistic clarity of the items of the pretest version. For face validity, the pretest version was applied to 30 individuals who assessed the items for linguistic clarity, adequacy, and comprehensibility.

RESULTS: The CVC for the scale was 0.97. The CVC values for clarity, pertinence, and relevance were 0.97, 1 and 0.99, respectively. All items had a CVC > 0.8, confirming acceptable content validity. The pretest version was tested on 12 female runners, eight coaches, and 10 dietitians. Items were clear and comprehensible to 93% and adequate to 96% of respondents. No changes were needed, and the pretest version was defined as the Brazilian version of the questionnaire to assess REDs knowledge (QC-REDs).

CONCLUSION: The QC-REDs obtained strong content and face validity. It provides a rigorously adapted tool for Brazil, supported by positive feedback from the target population, demonstrating its applicability and usability. This instrument fills a crucial gap and offers a valuable resource for researchers and professionals working with athletes at risk of REDs. Future research should focus on psychometric validation to further solidify its effectiveness in the Brazilian context.

INTRODUCTION

Relative Energy Deficiency in Sport (REDs) was first described in 2014 in a position statement by the International Olympic Committee (IOC) as a syndrome characterized by low energy availability (LEA) as its central component.¹

LEA refers to insufficient energy availability to support physiological functions after considering energy expended during exercise,² and has been suggested to occur in women when it is equal to or less than 30 kcal/kg FFM/day and in men when equal to or less than 9–25 kcal/kg FFM/day.^{3–7} LEA's effects can vary based on the severity, duration, and frequency of exposure, with adaptable LEA causing minimal or no impact, whereas problematic LEA leads to persistent disruptions across various body systems.⁸ Therefore, the more severe and prolonged the LEA, the greater the risk of leading to a persistent disruption of hormonal, metabolic, and physiological regulations,^{9–11} acting as an etiological factor for REDs.

REDs has deleterious effects on multiple physiological functions, including menstrual dysfunction, impaired reproductive function, and impairments in several systems, such as bone health, cardiovascular and neurocognitive function, energy metabolism, immunity, and skeletal muscle function. Sports performance is also reduced, manifested as decreased endurance and

power, reduced muscle strength, impaired recovery, motivation, and training response, in addition to decreased cognitive performance.⁸

The identification and management of REDs are critical aspects of athlete health; however, research shows a concerning knowledge gap among sports coaches and other health professionals, such as physiotherapists and physicians.¹² A study found that fewer than 30% of health professionals, including certified athletic trainers, were familiar with the term REDs after attending a sports medicine conference.¹² In addition, even among physicians trained in sports medicine and other health professionals, few expressed confidence in treating patients with REDs.¹² This lack of awareness is particularly troubling, as these professionals are uniquely positioned within the technical team to identify athletes at risk and educate them on the importance of maintaining adequate energy availability.⁸ Moreover, athletes often have limited knowledge about REDs, with many unaware of the connection between inadequate nutrition and menstrual dysfunction, for example, which can lead to serious long-term health consequences.¹³

In the absence of a validated questionnaire designed to evaluate the level of knowledge about REDs among stakeholders in the sports environment, Pai et al.¹⁴ developed and validated a questionnaire with satisfactory validity evidence to assess knowledge of the signs and symptoms of REDs. The instrument includes items that identify awareness of the female athlete triad, LEA, and REDs, as well as the recognition of REDs-related symptoms and their health and performance consequences. When applying this questionnaire, Pai et al.¹⁴ observed that health professionals exhibited higher levels of knowledge about REDs than physically active individuals. This finding suggests that the instrument can be a valuable tool for assessing current knowledge on REDs and identifying gaps that need to be addressed.¹⁴

To the authors' knowledge, no validated questionnaire is available to assess knowledge of the signs, symptoms, and consequences of REDs among Brazilian health professionals and athletes. Given the severity of the consequences of LEA-related consequences, particularly in the development of REDs, raising awareness about this syndrome should be the first step toward prevention. Therefore, adapting and validating a questionnaire to assess knowledge of REDs within the Brazilian sports stakeholder community is crucial for advancing awareness and improving strategies for disseminating information about REDs in both sports and academic settings. In this context, this study aims to translate the questionnaire designed by Pai et al.¹⁴ into Brazilian Portuguese, culturally adapt it, and obtain content and face validity evidence for its use in Brazil.

MATERIAL AND METHODS

This cross-sectional study for linguistic and cross-cultural validation was reviewed and approved by the Ethics and Research Committee

of the Universidade Federal de São Paulo (No. 67851223.1.0000.5505), and all participants provided informed consent.

The questionnaire to assess REDs knowledge is a self-administered instrument that was first developed in English and consists of 18 items assessing the theoretical and practical understanding of REDs' signs, symptoms, and consequences.¹⁴ The original questionnaire was validated in a sample ($n = 174$) composed of health professionals and physically active individuals from New Zealand, aged between 18 and 70 years.¹⁴ The instrument developed by Pai et al.¹⁴ has a dichotomous format, with results classified as correct (score = 1) or incorrect (score = 0). Thus, comprising 18 scoring items, the instrument's total score ranges from 0 to 18. Higher scores indicate a higher level of knowledge about REDs.

To initiate the translation and adaptation process, permission to use the questionnaire to assess REDs knowledge was first obtained by email from the corresponding author of the original version. The process of translating, culturally adapting, and validating the content of the questionnaire was carried out based on the recommendations of Guillemin et al.¹⁵ and Beaton et al.¹⁶ (Figure 1).

The original version of the questionnaire was translated into Portuguese by two independent Brazilian translators fluent in English. Translator 1 was a registered dietitian with a master's degree in the LEA area and an expert on instrument construct. Translator 2 was an accredited translator with command of both English and Portuguese and no knowledge of the questionnaire's construct. The two versions in Portuguese (T1 and T2) were independently analyzed by three registered sports dietitians knowledgeable about REDs (two with a master's degree and one PhD professor) to determine the best wording for the target population. The experts assessed the equivalence between the translations

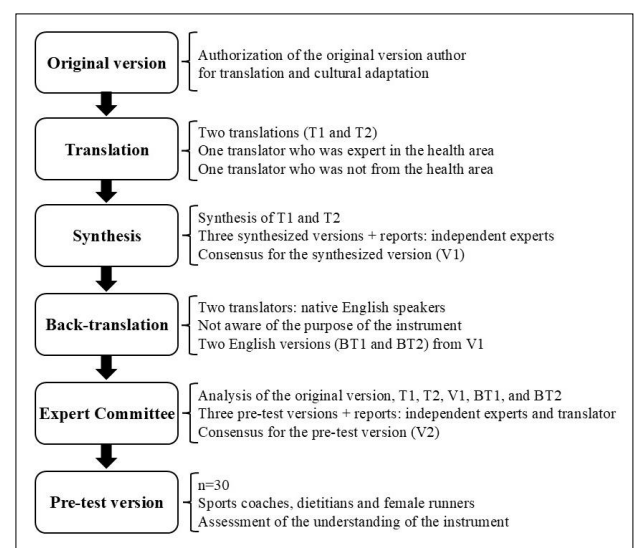


Figure 1. Translation and adaptation process, followed by evidence of validity.

and the original instrument across four aspects: (1) semantic equivalence, to determine whether the wording had the same meaning and to identify any grammatical errors; (2) idiomatic equivalence, to evaluate whether expressions that were difficult to translate were well adapted without changing the cultural meaning of the item; (3) experiential equivalence, to assess the applicability of the items in the new cultural context and, if necessary, replace them with equivalent items; and (4) conceptual equivalence, to ensure that terms and expressions, even if accurately translated, measured the same construct across different cultures.¹⁷ The three versions were synthesized by the authors into a single Portuguese version (V1).

Subsequently, version V1 was subjected to back-translation (translation from Portuguese to English) by two native English-speaking translators with command in Portuguese and no prior knowledge of the instrument's construct, resulting in two English versions, BT1 and BT2. Next, a committee of experts – comprising four health professionals and/or researchers from the health area with knowledge of the construct and fluency in English, as well as one of the translators involved in the process – independently evaluated all the materials produced during the process (original, T1, T2, V1, BT1, and BT2) to develop the pretest version. The committee included three sports registered dietitians and one sports gynecologist. Each member created a pretest version and provided a report on the process. These versions and reports were then compared by the authors, and based on theoretical experts' consensus, the final pretest version (V2) was structured.

The content validity process consisted of the evaluation of the materials generated from the translation, synthesis, and back-translation stages by an expert committee. The experts assessed the items of the synthesis version (V1) based on practical pertinence, theoretical relevance, and linguistic clarity, ensuring that the instrument effectively measured its intended construct.¹⁸ Each item was rated using a five-point Likert scale for each criterion (clarity, pertinence, and relevance), where “1” indicated that the item was not at all clear, pertinent, or relevant, whereas “5” indicated that the item was completely clear, pertinent, or relevant. These ratings enabled the calculation of the content validity coefficient (CVC).¹⁹

To calculate the CVC of the instrument, as well as the CVC for each item and each evaluated aspect related to the scale construct, the equations proposed by Filgueiras et al.²⁰ were applied. According to Hernández-Nieto,¹⁹ instrument items should have a minimum CVC value of 0.8. However, values above 0.7 are considered acceptable.¹⁹

Furthermore, face validity concerns whether the target population considers the items understandable, relevant, and easy to answer.²¹ Therefore, face validity was assessed by testing the pretest version with female runners, running coaches, and registered dietitians using a non-probabilistic snowball sampling method.

Participants were invited via social media and the researchers' professional networks. As recommended by Beaton et al.,¹⁶ 30 individuals participated in the face validity assessment. While completing the questionnaire, participants were asked to evaluate the clarity, adequacy, and comprehensibility of each item and to suggest modifications when necessary to improve understanding. Items that elicited doubts or were not understood by more than 15% of respondents were flagged for revision, and any modifications required would necessitate reapplication of the updated version to new respondents.²² The questionnaire was applied online, and data were collected and managed using the Research Electronic Data Capture (REDCap) platform hosted at Universidade Federal de São Paulo.^{23,24}

RESULTS

The Content Validity Coefficient (CVC) for the entire scale, calculated by subtracting the mean experts' CVC for the scale as a whole from the experts' standard error of polarization, was 0.97. For the clarity of the items, the CVC was 0.97. For the pertinence of the items, that is, their importance in composing the instrument, the CVC value was 1. Finally, for the relevance of the items in measuring the construct, the CVC was 0.99. **Table 1** presents the pretest version and the CVC value of each item according to each aspect judged by the five experts. However, the translator on the expert committee preferred to evaluate only the clarity of the items, considering himself unqualified to assess items within the construct that were not part of his area of expertise. Therefore, the CVC values for pertinence and relevance were calculated based on the judgments of four experts. All items had a CVC > 0.8, indicating that the pretest version of the instrument demonstrated acceptable content validity.

The pretest version (V2) of the questionnaire was applied to 12 female runners (mean age: 37.7 ± 11.3 years) who frequently participated in running events, eight running coaches (mean age: 38.8 ± 6.9 years), and 10 registered dietitians (mean age: 37.4 ± 11.4 years). Their feedback confirmed that the items were clear and understandable to 93% of respondents and adequate according to 96% of the sample. Given these findings, no modifications were necessary, and the pretest version (V2), presented in **Table 1**, was defined as the Brazilian version of the questionnaire to assess REDs knowledge (QC-REDs; see **Appendix 1**) to be tested using psychometric measures to accumulate validity evidence.

DISCUSSION

The Brazilian version of the questionnaire to assess REDs knowledge demonstrated acceptable evidence of content and face validity, standing as the first tool in Brazilian Portuguese designed to assess knowledge about the signs, symptoms, and consequences of REDs. The main strength of this study lies in the testing of a

Table 1. Calculation of CVC of the pretest version items

Pretest version (V2)	CVC		
	Clarity	Pertinence	Relevance
Você já ouviu falar de Baixa Disponibilidade de Energia?			
• Sim			
• Não	0,96	1	1
Você já ouviu falar de Deficiência Relativa de Energia nos Esportes?			
• Sim			
• Não	1	1	1
Você já ouviu falar de Tríade da Mulher Atleta?			
• Sim			
• Não	1	1	1
Ter um ciclo menstrual irregular costuma ser um sinal de que mulheres atletas/fisicamente ativas estão no auge da forma competitiva			
• Verdadeiro			
• Falso	0,96	1	1
• Não tenho certeza			
Você acha normal mulheres atletas/ fisicamente ativas não menstruarem? (excluindo gravidez ou ausência de menstruação por uso de contraceptivo)?			
• Sim			
• Não			
• Não tenho certeza	0,96	1	1
• Depende da situação			
Você acha que não consumir energia suficiente pode resultar na ausência de menstruação?			
• Sim			
• Não	1	1	1
• Não tenho certeza			
Você acha que fraturas (fissuras ou pequenas quebras) ocorrem mais frequentemente em meninas/mulheres que não menstruam por 3 ou mais meses do que naquelas que menstruam regularmente? (excluindo gravidez ou ausência de menstruação por uso de contraceptivo)			
• Sim			
• Não	0,92	1	1
• Não tenho certeza			
Você acha que a irregularidade ou ausência de menstruação está associada ao desenvolvimento de ossos mais fracos?			
• Sim			
• Não	0,96	1	1
• Não tenho certeza			
Quais das opções a seguir podem aumentar o risco de um atleta contrair infecções, tais como gripe e resfriado? (por favor, marque todas as que se aplicam)			
• Treino intenso com descanso inadequado	1	1	1
• Consumo insuficiente de líquidos	0,88	1	0,95
• Consumo insuficiente de energia	1	1	1
• Não sei	1	1	1
Em sua opinião, quais das seguintes opções poderiam ser resultado da insuficiência crônica de energia? (por favor, marque todas as que se aplicam)			
• Aumento do desempenho na corrida de velocidade	0,88	1	0,9
• Diminuição do desempenho na corrida de velocidade	0,88	1	0,9
• Mudanças no peso	1	1	1
• Mudanças na composição corporal	1	1	1
• Diminuição da força muscular	1	1	1
• Aumento da força muscular	1	1	1
• Diminuição da massa muscular	1	1	1
• Aumento de lesões	1	1	1

needed, easy-to-apply, and easy-to-interpret instrument to measure knowledge about REDs.

The process of translation and cultural adaptation of the questionnaire to assess REDs knowledge was conducted carefully, following methodological recommendations widely recognized in the health field.^{15,16} The objective was to ensure that the Brazilian version maintained the quality and content of the original version proposed by Pai et al.¹⁴ Recommendations for conducting translation and cultural adaptation reflect best practices and raise expectations that the translated and culturally adapted instrument will be very similar to the original.²⁵

Based on the judgment of the expert committee, an excellent CVC of 0.97 was identified for the studied scale, with all items scoring above 0.8, supporting the retention of the questionnaire structure without modifications. Similarly, the original version of the questionnaire achieved a CVC greater than 0.8 based on expert evaluation.¹⁴ Content validity of a scale refers to whether the items are adequate to measure the construct it purports to measure in a relevant and representative way.²⁶ The feedback provided by the evaluators is crucial for refining items before the scale is applied to larger samples.²⁶

According to the International Test Commission,²⁷ an expert committee should be composed of individuals with different qualifications to ensure that all relevant aspects are addressed, including those that may be overlooked by others. In this context, a translator without expertise in health and sports participated in the committee. While this committee member focused solely on evaluating the linguistic clarity of the items, the practical pertinence and theoretical relevance were assessed exclusively by experts with knowledge of the construct, thus ensuring methodological rigor in the evaluation process.

Throughout the process, the original format of the questionnaire was preserved without removing or rearranging items. The adaptations focused on linguistic adjustments to enhance comprehension by the Brazilian population without affecting the real meaning of the expressions. The Brazilian version was developed following a content validity assessment by a committee of experts and was later validated by the target population. Given that content validity is fundamental to establishing other forms of validity, it should be prioritized in the instrument development process.²⁸ Validity does not refer to the instrument itself but rather to the scores it produces when applied to a specific group of respondents for a particular purpose.²⁸

The fact that the pretest version underwent no modifications compared with the original version reinforces its fidelity to the original instrument. The involvement of the expert committee in the adaptation process is crucial for standardizing the items, ensuring clarity and ease of understanding, and maintaining the semantic equivalence of terms in relation to the original version.²⁹

The language and structure of the pretest version of the questionnaire were also validated by the target population (running coaches, dietitians, and female runners), who found the instrument to be clear, comprehensible, and adequate for their context. The absence of necessary modifications suggests that the adaptation was effective and well received by the participants. The importance of a test having high face validity is linked to the quality of the data collected and the participants' experience.²¹ A questionnaire with high face validity naturally enhances data quality.²¹

According to Magasi et al.,³⁰ the perspective of the target population is important because they have direct, lived experience with the concept. In contrast, experts offer a broader view, having observed it across diverse individuals and contexts.³⁰ Therefore, strong validity evidence requires balancing both perspectives (content and face validity) rather than favoring one over the other.³⁰

One limitation of this study is the lack of exploration of responses from health professionals and athletes beyond those included in the sample. While the participation of dietitians, running coaches, and female runners provided valuable insights into the clarity and adequacy of the QC-REDs, expanding the sample to include a broader range of professionals, such as physicians, physiotherapists, psychologists, and athletes from various disciplines, could offer a more comprehensive evaluation of the instrument's applicability across different contexts. Additionally, future studies should investigate how the adapted version of the QC-REDs performs in diverse populations to ensure its validity across various sporting backgrounds.

CONCLUSION

The findings of this study indicate that the Brazilian version of the QC-REDs demonstrated strong content and face validity.

This study contributes to the field by providing a rigorously adapted version of the QC-REDs for use in Brazil. The strong content validity and positive feedback from the target population support the questionnaire's applicability and usability. In addition, this instrument fills a significant gap in the field by providing a valuable tool for researchers and professionals working with athletes and individuals at risk of REDs.

Future research focusing on psychometric validation to establish the robustness of this instrument in assessing REDs knowledge within the Brazilian context is in progress and will be reported soon.

REFERENCES

1. Mountjoy M, Sundgot-Borgen J, Burke L, et al. The IOC consensus statement: beyond the Female Athlete Triad—Relative Energy Deficiency in Sport (RED-S). *Br J Sports Med*. 2014 Apr;48(7):491–7. <https://doi.org/10.1136/bjsports-2014-093502>.
2. Mountjoy M, Sundgot-Borgen JK, Burke LM, et al. IOC consensus statement on relative energy deficiency in sport (RED-S): 2018 update.

- Br J Sports Med. 2018 Jun;52(11):687–97. <https://doi.org/10.1136/bjsports-2018-099193>.
3. Koehler K, Hoerner NR, Gibbs JC, et al. Low energy availability in exercising men is associated with reduced leptin and insulin but not with changes in other metabolic hormones. *Journal of Sports Sciences*. 2016 Oct 17;34(20):1921–9. <https://doi.org/10.1080/02640414.2016.1142109>.
 4. Fagerberg P. Negative Consequences of Low Energy Availability in Natural Male Bodybuilding: A Review. *International Journal of Sport Nutrition and Exercise Metabolism*. 2018 Jul 1;28(4):385–402. <https://doi.org/10.1123/ijsnem.2016-0332>.
 5. Lane AR, Hackney AC, Smith-Ryan A, et al. Prevalence of Low Energy Availability in Competitively Trained Male Endurance Athletes. *Medicina*. 2019 Oct 1;55(10):665. <https://doi.org/10.3390/medicina55100665>.
 6. Langan-Evans C, Germaine M, Artukovic M, et al. The Psychological and Physiological Consequences of Low Energy Availability in a Male Combat Sport Athlete. *Medicine & Science in Sports & Exercise*. 2021 Apr;53(4):673–83. <https://doi.org/10.1249/MSS.00000000000002519>.
 7. Jurov I, Keay N, Rauter S. Reducing energy availability in male endurance athletes: a randomized trial with a three-step energy reduction. *Journal of the International Society of Sports Nutrition*. 2022 Dec 31;19(1):179–95. <https://doi.org/10.1080/15502783.2022.2065111>.
 8. Mountjoy M, Ackerman KE, Bailey DM, et al. 2023 International Olympic Committee's (IOC) consensus statement on Relative Energy Deficiency in Sport (REDs). *Br J Sports Med*. 2023 Sep;57(17):1073–98. <https://doi.org/10.1136/bjsports-2023-106994>.
 9. Low energy availability, not stress of exercise, alters LH pulsatility in exercising women | *Journal of Applied Physiology* | American Physiological Society. *Journal of Applied Physiology* [Internet]. [cited 2026 Apr 14]. Available from: <https://journals.physiology.org/doi/10.1152/jappl.1998.84.1.37>.
 10. Loucks AB, Thuma JR. Luteinizing Hormone Pulsatility Is Disrupted at a Threshold of Energy Availability in Regularly Menstruating Women. *The Journal of Clinical Endocrinology & Metabolism*. 2003 Jan;88(1):297–311. <https://doi.org/10.1210/jc.2002-020369>.
 11. Ihle R, Loucks AB. Dose-Response Relationships Between Energy Availability and Bone Turnover in Young Exercising Women. *Journal of Bone and Mineral Research*. 2004 Aug 1;19(8):1231–40. <https://doi.org/10.1359/JBMR.040410>.
 12. Tenforde A, Beauchesne A, Borg-Stein J, et al. Awareness and comfort treating the female athlete triad and relative energy deficiency in sport among healthcare providers. *Dtsch Z Sportmed*. 2020 Mar 1;71(3):76–80. <https://doi.org/10.5960/dzsm.2020.422>.
 13. Brown KN, Wengreen HJ, Beals KA. Knowledge of the Female Athlete Triad, and Prevalence of Triad Risk Factors among Female High School Athletes and their Coaches. *Journal of Pediatric and Adolescent Gynecology*. 2014 Oct;27(5):278–82. <https://doi.org/10.1016/j.jpag.2013.11.014>.
 14. Pai NN, Brown RC, Black KE. The development and validation of a questionnaire to assess relative energy deficiency in sport (RED-S) knowledge. *Journal of Science and Medicine in Sport*. 2022 Oct;25(10):794–9. <https://doi.org/10.1016/j.jsams.2022.07.004>.
 15. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. *Journal of Clinical Epidemiology*. 1993 Dec;46(12):1417–32. [https://doi.org/10.1016/0895-4356\(93\)90142-N](https://doi.org/10.1016/0895-4356(93)90142-N).
 16. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures: Spine. 2000 Dec;25(24):3186–91. <https://doi.org/10.1097/00007632-200012150-00014>.
 17. Borsa JC, Damásio BF, Bandeira DR. Adaptação e validação de instrumentos psicológicos entre culturas: algumas considerações. *Paidéia (Ribeirão Preto)*. 2012;22:423–32. <https://doi.org/https://doi.org/10.1590/S0103-863X2012000300014>.
 18. Oliveira MAM, Sousa WPS, Maia EMC. Adaptação e validade de conteúdo da versão brasileira da Cambridge Worry Scale. *Rev Enferm UFPE On Line*. 2017;11(5):2083–9. <https://doi.org/10.5205/reuol.9302-81402-1-RV.1105sup201713>.
 19. Hernández-Nieto R. Contributions to statistical analysis. Mérida: Universidad de Los Andes; 2002.
 20. Filgueiras A, Galvão BDO, Pires P, et al. Tradução e adaptação semântica do Questionário de Controle Atencional para o Contexto Brasileiro. *Estud psicol (Campinas)*. 2015 Jun;32(2):173–85. <https://doi.org/10.1590/0103-166X2015000200003>.
 21. Allen MS, Robson DA, Iliescu D. Face Validity: A Critical but Ignored Component of Scale Construction in Psychological Assessment. *European Journal of Psychological Assessment*. 2023 May;39(3):153–6. <https://doi.org/10.1027/1015-5759/a000777>.
 22. Ciconelli RM. Tradução para o português e validação do questionário genérico de avaliação de qualidade de vida medical outcomes study 36-item short-form health survey (SF-36) [Internet]. 1997 [cited 2026 Apr 14]. Available from: <http://repositorio.unifesp.br/handle/11600/15360>.
 23. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009 Apr;42(2):377–81. PMID: 18929686; <https://doi.org/10.1016/j.jbi.2008.08.010>.
 24. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*. 2019 Jul;95:103208. <https://doi.org/10.1016/j.jbi.2019.103208>.
 25. Hall DA, Zaragoza Domingo S, Hamdache LZ, et al. A good practice guide for translating and adapting hearing-related questionnaires for different languages and cultures. *International Journal of Audiology*. 2018 Mar 4;57(3):161–75. <https://doi.org/10.1080/14992027.2017.1393565>.
 26. Roebianto A, Savitri SI, Aulia I, Suciyan A, Mubarak L. Content validity: Definition and procedure of content validation in psychological research. *TPM-Testing, Psychometrics, Methodology in Applied Psychology*. 2023;30(1):5–18.

27. ITC Guidelines for Translating and Adapting Tests (Second Edition). International Journal of Testing. 2018 Apr 3;18(2):101–34. <https://doi.org/10.1080/15305058.2017.1398166>.
28. Zamanzadeh V, Ghahramanian A, Rassouli M, et al. Design and Implementation Content Validity Study: Development of an instrument for measuring Patient-Centered Communication. J Caring Sci. 2015 Jun 1;4(2):165–78. <https://doi.org/10.15171/jcs.2015.017>.
29. Nora CRD, Zoboli E, Vieira MM. Validação por peritos: importância na tradução e adaptação de instrumentos. Revista Gaúcha de Enfermagem [Internet]. 2017 Nov 16 [cited 2026 Apr 14];38(3). Available from: <https://seer.ufrgs.br/index.php/produto&producao/ojs/include/index.php/rgenf/article/view/64851>.
30. Magasi S, Ryan G, Revicki D, et al. Content validity of patient-reported outcome measures: perspectives from a PROMIS meeting. Qual Life Res. 2012 Jun;21(5):739–46. <https://doi.org/10.1007/s11136-011-9990-8>.

Authors' contributions: Maria UP: conceptualization, methodology, investigation, data curation, formal analysis, writing – original draft, writing – review and editing; Juzwiak CR: conceptualization, methodology, supervision, writing – original draft. All authors reviewed and approved the final version of the manuscript for publication.

Acknowledgments: The authors would like to thank the translators, experts, health professionals, and athletes who participated in this research.

Sources of funding: Uyara Pereira de Maria received a doctoral scholarship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), which provided financial support during the research process.

Conflicts of interest: None.

Data availability statement: Data supporting the findings of this study are available from the corresponding author upon reasonable request. Uyara Pereira de Maria.

Declaration of generative AI in scientific writing: During the preparation of this study, the authors did not use generative AI or AI-assisted technologies for content generation, data analysis, or interpretation.

Date of first submission: April 30, 2025

Last received: February 25, 2026

Accepted: April 8, 2026

Address for correspondence:

Uyara Pereira de Maria
Programa de Pós-Graduação Interdisciplinar em Ciências da Saúde,
Universidade Federal de São Paulo (PPGICS-Unifesp)
Rua Silva Jardim, 136, sala 2 A
Vila Mathias — Santos (SP) — Brasil
CEP 11015-020
Tel. (+55 11) 9 7504-7466
E-mail: uyara.maria@unifesp.br

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Psychometric properties of the Positive Aspects of Caregiving Scale for unpaid caregivers of people living with dementia in the Brazilian context: a validation study

Sofia Cristina Iost Pavarini^I, Aline Cristina Martins Gratão^{II}, Camila Rafael Ferreira Campos^{III}, Diana Quirino Monteiro^{IV}, Elizabeth Joan Barham^V, Fabiana de Souza Orlandi^{VI}, Gabriela Martins^{VII}, Gustavo Carrijo Barbosa^{VIII}, Keila Cristianne Trindade da Cruz^{IX}, Larissa Corrêa^X, Luana Aparecida Rocha^{XI}, Ludmyla Caroline de Souza Alves^{XII}, Ana Carolina Ottaviani^{XIII}

Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil

^IPhD. Nurse, Associate professor, Department of Nursing, Department of Gerontology, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil
<https://orcid.org/0000-0001-9359-8600>

^{II}PhD. Nurse, Associate professor, Department of Nursing, Department of Gerontology, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil
<https://orcid.org/0000-0002-8508-0251>

^{III}PhD. Psychologist, Postgraduate Program in Psychology, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil
<https://orcid.org/0000-0001-8700-153X>

^{IV}PhD. Gerontologist, Professor in Graduate Program in Gerontology, Universidade Católica de Brasília (UCB), Brasília (DF), Brazil
<https://orcid.org/0000-0002-3130-4096>

^VPhD. Psychologist, Associate professor, Department of Psychology, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil
<https://orcid.org/0000-0002-7270-4918>

^{VI}PhD. Nurse, Associate professor, Department of Nursing, Department of Gerontology, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil
<https://orcid.org/0000-0002-5714-6890>

^{VII}MSc. Gerontologist; PhD student, Postgraduate Program in Nursing, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0002-3637-4763>

^{VIII}MSc. Physiotherapist; PhD student, Postgraduate Program in Nursing, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0001-8670-1227>

^{IX}PhD. Nurse, Associate professor, Department of Nursing, Universidade de Brasília (UnB), Brasília (DF), Brazil.
<https://orcid.org/0000-0001-8146-8323>

^XMSc. Gerontologist; PhD student, Postgraduate Program in Nursing, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0002-4596-4845>

^{XI}MSc. Gerontologist; PhD student, Postgraduate Program in Nursing, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0002-6707-4982>

^{XII}MSc. Gerontologist; PhD student, Postgraduate Program in Nursing, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0001-5507-702X>

^{XIII}PhD. Gerontologist; Professor in the Department of Medicine, Universidade Federal de São Carlos (UFSCar), São Carlos (SP), Brazil.
<https://orcid.org/0000-0003-4037-4587>

KEYWORDS (MeSH terms):

Validation study.
 Caregivers.
 Dementia.
 Mental health.

AUTHOR'S KEYWORDS:

Psychometric properties.
 Family caregivers.
 Positive aspects of caregiving.
 Major neurocognitive disorder.

ABSTRACT

BACKGROUND: The Positive Aspects of Caregiving Scale (PAC) assesses the satisfaction and psychosocial benefits perceived by caregivers in the caregiving experience.

OBJECTIVE: To analyze the psychometric properties of the PAC among unpaid caregivers of people living with dementia in the Brazilian context.

DESIGN AND SETTING: Validation study of the instrument conducted online with caregivers from different regions of Brazil.

METHODS: The sample comprised 100 unpaid caregivers of people living with dementia. Sociodemographic characteristics, positive aspects of caregiving, burden, depressive symptoms, anxiety symptoms, and quality of life were assessed using an online form. Evidence of validity based on internal structure was examined using confirmatory factor analysis, reliability testing, and criterion validity analysis.

RESULTS: Most participants were women (95%), daughters of the care recipient (85%), with an average age of 54.5 years, a high educational level, and residence in the Southeast region of Brazil. Depressive symptoms (70%), anxiety (65%), moderate to severe burden (42%), and low quality of life (62%) were prevalent. Factor analysis indicated satisfactory fit for the PAC, and the high correlation between factors supported consideration of a one-dimensional scoring structure. The scale showed high internal consistency ($\alpha = 0.89$) and significant correlations with burden, anxiety symptoms, depressive symptoms, and quality of life.

CONCLUSION: The PAC showed good psychometric properties in the Brazilian context, supporting its validity and reliability for assessing positive aspects of caregiving among unpaid caregivers of people living with dementia.

INTRODUCTION

Globally, over 55 million people live with dementia, with nearly 10 million new cases expected annually, making dementia a public health problem and one of the leading causes of disability and dependency.¹ Most people with dementia are cared for by unpaid caregivers, particularly in low- and middle-income countries.²

Studies indicate that unpaid caregivers of people living with dementia often face considerable challenges, including potentially deleterious effects on well-being, such as burden, depression, anxiety, decline in physical capacity, and reduced quality of life.³⁻⁵ However, caregivers may also report positive experiences when caring for family members, and these experiences may play an important role in their well-being and quality of life.^{3,6}

Positive aspects of caregiving refer to the gains or satisfaction that caregivers derive from the caregiving experience.⁷ Satisfaction, rewards, perceived gains, competence, meaning, resilience, self-efficacy, personal growth, and a sense of mastery are commonly described as positive aspects in the literature.^{6,8,9}

A study involving 1,222 unpaid caregivers of people living with dementia in the United States found that positive aspects of caregiving were inversely associated with depressive and anxiety symptoms. Caregivers who reported higher levels of positive aspects of caregiving had fewer depressive and anxiety symptoms than those with lower levels. Thus, positive aspects of

caregiving may make the burden of care more tolerable and may help alleviate caregiver distress over time.¹⁰

The Positive Aspects of Caregiving (PAC) Scale is widely recognized as an instrument for assessing the psychological benefits of caregiving among caregivers.^{6,8} The PAC was originally developed by Tarlow et al.⁷ in 2004 and initially consisted of 11 items. After exploratory and confirmatory factor analyses, the instrument was reduced to nine items. The scale comprises two factors (self-affirmation and perspective in life) and yields a total score ranging from 9 to 45 points, with higher scores indicating greater positive perception and benefits derived from the caregiving experience. The internal consistency of the overall instrument was $\alpha = 0.89$.⁷

Although the PAC has been used among unpaid caregivers of people living with dementia in the United States,^{7,11} Europe,^{12,13} and Asia,^{14–16} evidence on its use in Brazil remains limited. Recently, the PAC was translated and culturally adapted for use with unpaid caregivers in the Brazilian context, following internationally recommended guidelines for cross-cultural adaptation, including translation, back-translation, expert committee review, and pre-testing.¹⁷ However, evidence regarding its psychometric properties in Brazil remains limited, reinforcing the need for validation studies to support its use in research and clinical practice.¹⁷

OBJECTIVE

To analyze the psychometric properties of the Positive Aspects of Caregiving Scale among unpaid caregivers of people living with dementia in the Brazilian context.

METHODS

This validation study examined the PAC Scale, which had previously been translated and culturally adapted into Brazilian Portuguese.¹⁷ This study is part of a larger umbrella project entitled iSupport-Brazil,¹⁸ which evaluates the implementation and effects of the World Health Organization's iSupport program, an online psychosocial intervention designed to support unpaid caregivers of people living with dementia¹⁸. All scientific and ethical guidelines were followed, and the study was approved by the Human Research Ethics Committee of the Universidade Federal de São Carlos (CAAE No. 88157118.0.1001.5504).

The study sample comprised 100 unpaid caregivers of people living with dementia from different regions of Brazil. The sample size was determined according to the recommendations of Hair et al.¹⁹ based on the number of items in the scale. The inclusion criteria were defined according to the iSupport-Brazil umbrella project and included being at least 18 years of age, self-identifying as an unpaid caregiver of a person living with dementia, having provided care for at least six months, residing in Brazil, and having access to a smartphone, computer, or tablet with internet access.

In addition, only caregivers who scored equal to or above the minimum cut-off in at least one of the following measures were included: general perception of burden (≥ 4), anxiety symptoms (≥ 3), or depressive symptoms (≥ 3). These cut-off points were adopted to ensure the inclusion of caregivers experiencing at least a minimal level of emotional burden or psychological distress, in accordance with the eligibility criteria of the iSupport-Brazil intervention,¹⁸ which targets caregivers at higher psychosocial risk. Caregivers of institutionalized individuals were excluded.

Participants were recruited through online dissemination linked to the iSupport-Brazil umbrella project.¹⁸ Interested caregivers completed a pre-registration form and were subsequently invited to complete an online screening questionnaire to assess the eligibility criteria. Caregivers who met the inclusion criteria received a new link to the full data collection form. In this form, they provided digital informed consent and answered questions on sociodemographic and caregiving characteristics, positive aspects of caregiving, burden, depressive and anxiety symptoms, and quality of life. Data were collected online between January and May 2024.

In addition to the PAC, the following variables of interest were assessed:

- **Sociodemographic and caregiving characteristics:** sex (female or male), age (continuous), marital status (with or without partner), ethnicity/skin color (white, brown, black, or yellow), region of the country (Southeast, Northeast, Midwest, South and North), number of years in school (5–8 years, 9–11 years, or ≥ 12 years), relationship to the care recipient (spouse, sibling, grandchild, or other), and diagnosis of the care recipient's type of dementia (Alzheimer's disease, vascular dementia, frontal-temporal dementia, Lewy body dementia, or I do not remember/do not know). All information was self-reported by participants through the online questionnaire.
- **Burden:** Burden was assessed using the Zarit Burden Interview (ZBI), which consists of 22 items rated on a Likert scale from 0 to 4 points and measures subjective burden. Scores between 21 and 40 indicate mild burden, scores between 41 and 60 indicate moderate to severe burden, and scores above 61 indicate intense burden.^{20,21}
- **Anxiety and depressive symptoms:** Anxiety and depressive symptoms were assessed using the Hospital Anxiety and Depression Scale (HADS), which consists of 14 items, with seven questions for each subscale, rated on a Likert scale from 0 to 3 points. Scores >8 suggest the possible or probable presence of anxiety or depressive symptoms.^{22,23}
- **Quality of life:** Caregivers' quality of life was assessed using the caregiver version of the Quality of Life in Alzheimer's Disease scale (QoL-AD), which consists of 13 items rated on a Likert scale from 1 to 4. Based on the median score of 32 points, higher values indicate better quality of life.^{24,25}

Statistical analyses were performed using SAS software version 9.4, except for confirmatory factor analysis, which was conducted using MPLUS software version 6.12. After data collection, descriptive analyses were calculated, including measures of frequency (relative and absolute), position (minimum, mean, and maximum), and dispersion (standard deviation). To assess the fit of the original models to the data, the following fit indices were used: chi-squared test (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Weighted Root Mean Square Residual (WRMR). The adequacy criteria considered were CFI and TLI (≥ 0.9), RMSEA (< 0.06), and WRMR (< 1), as recommended in the literature.^{26–28}

Cronbach's alpha was calculated to assess the internal consistency of the instruments, with values > 0.6 considered satisfactory.²⁹ Spearman's correlation coefficient was used to assess the relationship between scores, according to the following criteria: weak correlation (< 0.3), moderate correlation ($0.3–0.5$), and strong correlation (> 0.5).²⁹ A significance level of 5% was adopted for all analyses.

RESULTS

Among the 100 unpaid caregivers of people living with dementia, 95% were female, with an average age of 54.5 ($SD\ 9.61$) years. Most participants were white (56%), living with a partner (56%), had ≥ 12 years of schooling (89%) and living in the Southeast region of Brazil (65%). Most were daughters of the person living with dementia and had provided care for 1–3 years (38%). Among care recipients, 72% had Alzheimer's disease. Moderate to severe burden (42%), anxiety symptoms (65%), depressive symptoms (70%), and lower quality of life (62%) predominated among caregivers. **Table 1** presents the characteristics of the participants.

Analysis of the standardized estimates of factor loadings indicates that the PAC Scale items were adequately represented by factors F1 (self-affirmation) and F2 (perspective in life), showing good factor loadings and high statistical significance. All items showed a strong correlation with their respective factors, suggesting that the factors were well defined. However, the high correlation between F1 and F2 (0.872) exceeded the recommended limit of 0.8–0.85, which may indicate poor discriminant validity. This value suggests that the F1 and F2 factors are not fully distinct and may reflect overlapping constructs. The strong correlation between these factors suggests that a single-factor interpretation may be appropriate to reflect the construct measured by the scale (**Table 2**).

Table 3 presents the results of the confirmatory factor analysis conducted to assess the fit of the PAC scale in two models: bifactor and unidimensional models. The CFI and TLI indexes were above 0.9, indicating good model fit. However, the RMSEA value exceeded the recommended reference value of ≤ 0.06 , suggesting a reasonable but not optimal fit. Conversely, the WRMR values were

Table 1. Descriptive analysis of participants' sociodemographic, care, and health characteristics, 2024

Variable	n (%)
Sociodemographic characteristics	
Gender	
Female	95 (95%)
Male	5 (5.05)
Age	
18–59 years	70 (70%)
≥ 60 years	30 (30%)
Schooling	
5–8 years	3 (3%)
9–11 years	8 (8%)
≥ 12 years	89 (89%)
Marital status	
With a partner	56 (56%)
Without a partner	44 (44%)
Ethnicity	
White	56 (56%)
Brown	33 (33%)
Black	10 (10%)
Yellow	1 (1%)
Country region	
Southeast	65 (65%)
Northeast	16 (16%)
Midwest	10 (10%)
South	5 (5%)
North	4 (4%)
Characteristics of care	
Degree of kinship	
Child	85 (85%)
Spouse	9 (9%)
Sibling	2 (2%)
Grandchild	1 (1%)
Other	3 (3%)
Care time	
Less than a year	12 (12%)
1–3 years	38 (38%)
4–6 years	21 (21%)
7–10 years	23 (23%)
≥ 11 years	6 (6%)
Type of dementia	
Alzheimer's disease	72 (72%)
Vascular dementia	10 (10%)
Frontal-temporal dementia	5 (5%)
Lewy Body dementia	5 (5%)
I don't remember/I don't know	16 (16%)
Health characteristics	
Burden	
Absent (≤ 20)	5 (5%)
Mild (21–40)	38 (38%)
Moderate to severe (41–60)	42 (42%)
Intense (≥ 61)	15 (15%)
Anxiety symptoms	
Possible/likely (≥ 8)	65 (65%)
No symptoms (≤ 7)	35 (35%)
Depressive symptoms	
Possible/likely (≥ 8)	70 (70%)
Absence of symptoms (≤ 7)	30 (30%)
Quality of life	
Higher levels (≥ 32)	38 (38%)
Lower levels (< 32)	62 (62%)

Table 2. Standardized factor loading estimates for the PAC Scale. São Carlos, 2024

	Estimate	Standard error	p value*
F1 BY			
PAC1 – Caring has made me feel more useful	0.761	0.045	< 0.001
PAC2 – Caring has made me feel good about myself	0.784	0.043	< 0.001
PAC3 – Caring has made me feel that someone needs me	0.794	0.044	< 0.001
PAC4 – Caring has made me feel valued	0.853	0.036	< 0.001
PAC5 – Caring made me feel important	0.737	0.054	< 0.001
PAC6 – Caring has made me feel strong and confident	0.829	0.033	< 0.001
F2 BY			
PAC7 – Caring has allowed me to value life more	0.873	0.042	< 0.001
PAC8 – Caring has enabled me to develop a more positive attitude towards life	0.842	0.037	< 0.001
PAC9 – Caring has strengthened my relationships with other people	0.831	0.044	< 0.001
F1 WITH			
F2	0.872	0.039	< 0.001

Table 3. Confirmatory factor analysis of the PAC scale. São Carlos, SP, Brasil, 2024

Adjustment index	Model – two factors	Model – single factor
RMSEA	0.117 (CI 90%: 0.077–0.156)	0.126 (CI 90%: 0.088–0.165)
CFI	0.985	0.983
TLI	0.978	0.974
Chi-square	56.621 (24)	62.168 (24)
WRMR	0.637	0.649

RMSEA, root mean square error of approximation; CI, confidence interval; CFI, Bentler's comparative fit index; TLI, Tucker-Lewis index; WRMR, weighted root mean square residual.

below 1, indicating good fit with the observed data. The two-factor model showed marginally better fit indices than the unidimensional model. Nevertheless, the high correlation between self-affirmation and perspective in life suggests that a one-dimensional scoring structure may also be appropriate for representing the overall construct.

Regarding internal consistency, the Cronbach's alpha was 0.89 for the total scale, 0.86 for self-affirmation, and 0.80 for perspective in life. These findings indicate satisfactory reliability for the total scale and its domains, with the total score showing high internal consistency (**Table 4**).

The PAC showed significant negative correlations with burden, anxiety symptoms, and depressive symptoms, suggesting that higher PAC scores were associated with lower levels of these indicators. In addition, the scale showed significant positive correlations with quality of life, indicating that higher PAC scores were associated with better quality of life. These results suggest that the scale captures caregivers' positive experiences in a manner consistent with other relevant and theoretically related constructs (**Table 5**).

DISCUSSION

The results of this study indicate that the PAC Scale has adequate psychometric properties for the Brazilian population of unpaid caregivers of people living with dementia. The participants was

Table 4. Analysis of the internal consistency of the PAC scale, São Carlos, 2024

Item	Item-total correlation	Alpha if item is deleted
PAC Total		
PAC1	0.695	0.9
PAC2	0.737	0.897
PAC3	0.656	0.903
PAC4	0.785	0.894
PAC5	0.674	0.902
PAC6	0.694	0.9
PAC7	0.653	0.903
PAC8	0.663	0.903
PAC9	0.676	0.902
Self-Affirmation		
PAC1	0.711	0.871
PAC2	0.737	0.867
PAC3	0.658	0.879
PAC4	0.827	0.851
PAC5	0.707	0.871
PAC6	0.615	0.886
Perspective in Life		
PAC7	0.691	0.744
PAC8	0.723	0.706
PAC9	0.623	0.806

Cronbach's Alpha.

predominantly composed of women, daughters, and middle-aged caregivers, most of whom cared for family members with Alzheimer's disease. This profile reflects the characteristics of unpaid caregivers described in the literature.^{30,31} Most participants had over 12 years of schooling and lived in the Southeast region of Brazil, which may have influenced their perceptions of caregiving and the experiences reported in the PAC, considering that cultural differences may affect the relationships between caregiving experiences and caregiver roles.¹⁴

The internal consistency analysis showed that the PAC had satisfactory Cronbach's alpha values for both the total scale and

Table 5. Analysis of evidence for the construct validity of the PAC scale, São Carlos, 2024

	PAC – total	Self-affirmation	Perspective in life
Zarit	–0.36 (< 0.01)	–0.32 (< 0.01)	–0.35 (< 0.01)
HADS – Depression	–0.32 (< 0.01)	–0.29 (< 0.01)	–0.36 (< 0.01)
HADS – Anxiety	–0.26 (0.01)	–0.24 (0.01)	–0.26 (< 0.01)
QoL-AD	0.39 (< 0.01)	0.35 (< 0.01)	0.4 (< 0.01)

Spearman's Correlation: ρ (p value).

its dimensions, demonstrating adequate reliability. In the original study, in which the scale was applied to 1,299 family caregivers of people with Alzheimer's disease, Cronbach's alpha was 0.86 for self-affirmation, 0.80 for perspective in life, and 0.89 for the total scale.⁷ The findings of the present study are consistent with international validation studies conducted in different cultural contexts.^{11,12,14}

Confirmatory factor analysis revealed that the scale items were well represented by the original factors, self-affirmation and perspective in life. However, the high correlation between these factors suggests possible conceptual overlap, indicating that the constructs underlying the PAC are strongly interrelated. This finding in the Brazilian context corroborates the structural simplification observed in the original version of the scale,⁷ in which the strong bivariate association justified the use of a single global score.

In the present study, the high inter-factor correlation and high reliability of the total score support the use of a global score, reinforcing the hypothesis of a unidimensional interpretation of the scale. In Japan, Furukawa et al.¹⁵ also found satisfactory validity and reliability for the PAC, with Cronbach's alpha values of 0.895 for the total scale and 0.896 and 0.823 for the subscales. Exploratory factor analysis revealed a structure different from the original proposal, with two new factors, namely "Living an enriched life" and "Self-usefulness", demonstrating cross-cultural variations in the conceptual structure of the scale. Despite these structural differences, the Japanese validation also demonstrated high internal consistency and satisfactory psychometric performance, which is consistent with the findings of the present study and supports the overall robustness of the PAC across cultural contexts.

The possibility of a unidimensional interpretation in the Brazilian version, supported by the high correlation between factors and the reliability of the total score, is also consistent with the PAC validation study conducted in Greece. In that study, which included 247 family caregivers, the PAC showed excellent internal reliability ($\alpha = 0.898$ for the total scale), with components consistent with those observed in the present study. However, Items 7, giving meaning to life, and 8, learning new skills, although originally attributed only to the perspective in life factor, also showed significant loadings in the self-Affirmation factor, indicating possible ambiguity and interrelationship between domains.¹³ These findings

reinforce the notion that the domains assessed by the PAC may represent closely interrelated constructs rather than entirely distinct dimensions.

The need to consider alternative models becomes even more evident when comparing the present findings with the PAC validation study conducted in Singapore.¹⁶ In a national sample of 1,132 caregivers of individuals with functional limitations, a unidimensional model was initially suggested for the nine-item version. However, difficulties in model fit led to the proposal of a bifactor model, including a general and additional specific factors. High inter-item correlations indicated redundancy, resulting in the exclusion of Item 8 due to its overlap with Item 7 and its stronger correlation with Item 9. Consequently, a shortened version of the scale, the S-PAC, was developed and showed adequate psychometric properties.¹⁶ These results reinforce the importance of considering alternative models to the traditional bifactor structure, particularly in diverse cultural contexts.

The construct validity of the PAC was supported by its significant correlations with measures of burden, anxiety and depressive symptoms, and quality of life. The negative association between PAC scores and levels of burden, anxiety, and depression suggests that perceiving positive aspects of caregiving may mitigate the emotional impact of the caregiver role. Similarly, the positive correlations between PAC scores and quality of life indicate that caregivers who perceive more benefits from caregiving tend to report better quality of life.

A scoping review that included 230 studies on positive aspects of dementia caregiving confirmed these associations.⁶ The authors identified negative correlations between positive aspects of caregiving and depressive symptoms and burden, as well as positive correlations with well-being variables, such as mental health, quality of life, and life satisfaction.⁶ In addition, a systematic review indicated that higher PAC scores were predicted by factors, such as greater self-efficacy, continuity in the caregiving relationship, spirituality or religiosity, use of professional services, and less discomfort with daily care. Conversely, higher levels of burden and stress were associated with lower PAC scores.⁸

These findings reinforce the relevance of the PAC as a tool for assessing the positive impact of caregiving and highlight the importance of careful cross-cultural adaptation. They also indicate the need for future studies to further investigate the factor structure of the scale in different sociocultural contexts, with particular attention to the possibility of a unidimensional structure that adequately represents the experience of Brazilian unpaid caregivers.

Despite the contributions of this study, some limitations should be considered. The participants comprised mostly of caregivers from the Southeast region, which may limit the generalizability of the findings to other regions of Brazil, particularly given the linguistic variations in Portuguese spoken across the country. In addition, the

predominance of highly educated participants may have influenced the understanding and interpretation of the scale. Future studies should consider including participants with more diverse scores on the HADS and Zarit scales to examine the performance of the PAC across different levels of psychological distress and burden.

CONCLUSION

The PAC showed good psychometric properties in the Brazilian context and represents a valid and reliable tool for assessing positive aspects of caregiving among unpaid caregivers of people living with dementia. However, the high correlation between its factors suggests that future research should consider the possibility of a unidimensional structure that may better reflect the experience of these caregivers. This study contributes robust evidence regarding the applicability of the PAC in Brazil and supports its use in future research and interventions aimed at supporting unpaid caregivers, promoting coping strategies, and improving well-being in this population.

REFERENCES

- World Health Organization (WHO). WHO [Internet]. Geneva; 2025 [cited 2026 Apr 7]. Dementia. Available from: <https://www.who.int/news-room/fact-sheets/detail/dementia>
- Wimo A, Seeher K, Cataldi R, et al. The worldwide costs of dementia in 2019. *Alzheimers Dement*. 2023;9(7):2865–73. PMID: 36617519; <https://doi.org/10.1002/alz.12901>.
- Lindeza P, Rodrigues M, Costa J, Guerreiro M, Rosa MM. Impact of dementia on informal care: a systematic review of family caregivers' perceptions. *BMJ Support Palliat Care*. 2020 Oct 14;bmjspcare-2020-002242. PMID: 33055092; <https://doi.org/10.1136/bmjspcare-2020-002242>.
- Brini S, Hodkinson A, Davies A, et al. In-home dementia caregiving is associated with greater psychological burden and poorer mental health than out-of-home caregiving: a cross-sectional study. *Aging Ment Health*. 2022;26(4):709–15. PMID: 33554655; <https://doi.org/10.1080/13607863.2021.1881758>.
- Asfia SKBM, Bucholc J, McCaffrey N, Mihalopoulos C, Muldowney A, Engel L. Understanding the quality of life impacts of providing informal care to people with dementia: a systematic review of qualitative studies. *Journal of Alzheimer's Disease*. 2022;88(4):1293–309. PMID: 35754273; <https://doi.org/10.3233/JAD-220219>.
- Wang J, Li X, Liu W, Yang B, Zhao Q, Lü Y, Xiao M. The positive aspects of caregiving in dementia: A scoping review and bibliometric analysis. *Front. Public Health*. 2022;10(14):985391. PMID: 36187613; <https://doi.org/10.3389/fpubh.2022.985391>.
- Tarlow BJ, Wisniewski SR, Belle SH, Rubert M, Ory MG, Gallagher-Thompson D. Positive aspects of caregiving: Contributions of the REACH project to the development of new measures for Alzheimer's caregiving. *Res Aging*. 2004;26(4):429–53. <https://doi.org/10.1177/0164027504264493>.
- Lee Y, Li L. Evaluating the positive experience of caregiving: a systematic review of the positive aspects of caregiving scale. *Gerontologist*. 2022;62(9):e493–e507. PMID: 34216215; <https://doi.org/10.1093/geront/gnab092>.
- Pysklywec A, Plante M, Auger C, et al. The positive effects of caring for family carers of older adults: a scoping review. *International Journal of Care and Caring*. 2020;4(3):349–75. <https://doi.org/10.1332/239788220X15925902138734>.
- Cheng ST. Positive aspects of caregiving attenuate the relationship between behavioral bother and anxiety and depressive symptoms in dementia family caregivers. *Epidemiology, Clinical Practice And Health*. 2023;23(5):366–70. PMID: 37070249; <https://doi.org/10.1111/ggi.14581>.
- Roth DL, Dilworth-Anderson P, Huang J, Gross AL, Gitlin LN. Positive aspects of family caregiving for dementia: differential item functioning by race. *J Gerontol B Psychol Sci Soc Sci*. 2015;70(6):813–9. PMID: 26033356; <https://doi.org/10.1093/geronb/gbv034>.
- Gonçalves-Pereira M, Carmo I, Silva JA, Papoila AL, Mateos R, Zarit SH. Caregiving experiences and knowledge about dementia in Portuguese clinical outpatient settings. *Int Psychogeriatr*. 2010;22(2):270–80. PMID: 19891814; <https://doi.org/10.1017/S1041610209991050>.
- Tsatali M, Egkiazarova M, Toumpalidou M, Karagiozi K, Margaritidou P, Tsolaki M. Greek adaptation of the Positive Aspects of Caregiving (PAC) scale in dementia caregivers. *Clin Gerontol*. 2022;45(3):538–47. PMID: 31684839; <https://doi.org/10.1080/07317115.2019.1685047>.
- Devi F, Yuan Q, Wang P, et al. Positive aspect of caregiving among primary informal dementia caregivers in Singapore. *PLoS One*. 2020;15(8):e0237677. PMID: 32817648; <https://doi.org/10.1371/journal.pone.0237677>.
- Furukawa H, Greiner C. Reliability and validation of the Positive Aspects of Caregiving scale among Japanese caregivers of people with dementia. *Int J Nurs Sci*. 2021;8(2):210–4. PMID: 33997136; <https://doi.org/10.1016/j.ijnss.2021.03.007>.
- Siow JYM, Chan A, Østbye T, Cheng GH, Malhotra R. Validity and reliability of the Positive Aspects of Caregiving (PAC) scale and development of its shorter version (S-PAC) among family caregivers of older adults. *Gerontologist*. 2017;57(4):e75–e84. PMID: 28082275; <https://doi.org/10.1093/geront/gnw198>.
- Pavarini SCI, Gratão ACM, Campos CRF, et al. Translation and cultural adaptation of the Positive Aspects of Caregiving Scale for caregivers of people living with dementia in Brazilian context: a methodological study. *Sao Paulo Med J*. 2024;142(2):e2023325. PMID: 38511815; <https://doi.org/10.1590/1516-3180.2023.0325.R1.23012024>.
- Gratão ACM, Cardoso AM, Ottaviani AC, et al. Efficacy of iSupport-Brasil for unpaid caregivers of people living with dementia: protocol for a randomized and controlled clinical trial. *Dementia Neuropsychol*. 2023;24:17:e20230040. PMID: 38053642; <https://doi.org/10.1590/1980-5764-DN-2023-0040>.

19. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. *Análise multivariada de dados*. 6th ed. Porto Alegre: Bookman; 2009. 688 p. ISBN 978-85-7780-534-1.
20. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist*. 1980;20(6):649–55. PMID: 7203086; <https://doi.org/10.1093/geront/20.6.649>.
21. Taub A, Andreoli SB, Bertolucci PH. Dementia caregiver burden: reliability of the Brazilian version of the Zarit caregivers burden interview. *Cad de Saude Publica*. 2004;20(2):372–6. PMID: 15073616; <https://doi.org/10.1590/S0102-311X2004000200004>.
22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67(6):361–70. PMID: 6880820; <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>.
23. Botega NJ, Bio MR, Zomignani MA, Garcia C Jr, Pereira WA. Transtornos do humor em enfermaria de clínica médica e validação de escala de medida (HAD) de ansiedade e depressão [Mood disorders among inpatients in ambulatory and validation of the anxiety and depression scale HAD]. *Rev Saude Publica*. 1995 Oct;29(5):355–63. PMID: 8731275; <https://doi.org/10.1590/s0034-89101995000500004>.
24. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. *Psychosom Med*. 2002;64(3):510–9. PMID: 12021425; <https://doi.org/10.1097/00006842-200205000-00016>.
25. Novelli MM, Nitrini R, Caramelli P. Validation of the Brazilian version of the quality of life scale for patients with Alzheimer's disease and their caregivers (QOL-AD). *Aging Ment Health*. 2010;14(5):624–31. PMID: 20480421; <https://doi.org/10.1080/13607861003588840>.
26. Bentler PM. Comparative fit indexes in structural models. *Psychological bulletin*. 1990;107(2):238–46. PMID: 2320703; <https://doi.org/10.1037/0033-2909.107.2.238>.
27. Brown TA. *Confirmatory factor analysis for applied research*. 1st ed. New York: Guilford; 2006.
28. Yu CY. *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* [PhD Thesis]. Los Angeles: University of California; 2002.
29. Dancey C, Reidy J. *Estatística sem matemática para psicologia*. 7th ed. Porto Alegre: Penso; 2018. 624 p.
30. Jiménez García-Tizón S, Navarro Prados AB, Bueno Martínez MB. Towards a comprehensive psychosocial profile of the family caregiver of people with dementia. *Psicol Reflex Crit*. 2025 Mar 20;38(1):6. PMID: 40111622; <https://doi.org/10.1186/s41155-025-00342-5>.
31. 1. Brasil. Ministério da Saúde. *Sumário executivo ReNaDe: relatório nacional sobre a demência no Brasil: necessidades de cuidado, custos, produção científica e investimento em pesquisa* [Internet]. Brasília (DF): Ministério da Saúde; 2023. Available from: <https://static.poder360.com.br/2023/12/pesquisa-cuidados-demencia.pdf>.

Authors' contributions: Pavarini SCI: conceived and reviewed the manuscript; Grato ACM: revised the manuscript; Campos CRF: designed, drafted, and revised the manuscript; Monteiro DQ: designed, drafted, and revised the manuscript; Barham EJ: revised the manuscript; Orlandi FS: revised the manuscript; Martins G: analyzed and interpreted the data for the work; Barbosa GC: analyzed and interpreted the data for the work; Cruz KCT: revised the manuscript; Corrêa L: analyzed and interpreted the data for the work; Rocha LA: analyzed and interpreted the data for the work; Alves LCS: analyzed and interpreted the data for the work; Ottaviani AC: designed, drafted, and revised the manuscript. All authors reviewed and approved the final version of the manuscript for publication.

Sources of funding: Ministério da Saúde (TED 098/2021), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Process No. 307417/2021-2).

Conflicts of interest: None.

Data availability statement: Data supporting the findings of this study are available upon request from the corresponding author, Ana Carolina Ottaviani.

Declaration of generative AI in scientific writing: During the preparation of this work, the authors did not use any generative AI or AI-assisted technologies.

Date of first submission: May 5, 2025

Last received: February 20, 2026

Accepted: April 6, 2026

Address for correspondence:

Ana Carolina Ottaviani
Departamento de Medicina, Universidade Federal de São Carlos (UFSCar)
Rod. Washington Luís, km 235
Monjolinho — São Carlos (SP) — Brasil
CEP 13565-905
Tel. (+55 16) 9 9177-8704
E-mail: carol.ottaviani@gmail.com

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Clinical features of chronic tension-type headache combined with hyperacusis: a cross-sectional study

Zenghui Fu^I, Yan Jin^{II}, Zaihong Lin^{III}, Yang Liu^{IV}, Kedong Guo^V

Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University, Qiqihar, China

^IMSc. Lecturer, Physician, Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University, Qiqihar, China.
 ID <https://orcid.org/0000-0002-6591-3471>

^{II}MSc. Lecturer, Physician, Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University, Qiqihar, China.
 ID <https://orcid.org/0009-0005-6007-396X>

^{III}MSc. Associate Professor, Physician, Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University, Qiqihar, China.
 ID <https://orcid.org/0009-0007-5850-3742>

^{IV}BS. Nurse Practitioner, Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University, Qiqihar, China.
 ID <https://orcid.org/0009-0004-4664-5637>

^VMSc. Lecturer, Pathology School, Qiqihar Medical University, Qiqihar, China.
 ID <https://orcid.org/0009-0002-4281-9826>

KEYWORDS (MeSH terms):

Tension-type headache.
 Hyperacusis.
 Sleep quality.
 Anxiety.
 Cross-sectional studies.

AUTHOR'S KEYWORDS:

Pericranial tenderness.
 Loudness discomfort level.
 Sound intolerance.
 Chronic pain.

ABSTRACT

BACKGROUND: Chronic tension-type headache (CTTH) is a prevalent primary headache disorder that significantly affects quality of life. Hyperacusis, an abnormal intolerance to everyday sounds, is frequently observed in patients with CTTH, suggesting potential shared neural mechanisms. However, the clinical characteristics of patients with CTTH and comorbid hyperacusis remain poorly characterized.

OBJECTIVES: This study aimed to investigate the clinical characteristics of CTTH patients with comorbid hyperacusis, focusing on headache profiles, auditory sensitivity patterns, and psychological correlates, and to identify factors independently associated with hyperacusis in this population.

DESIGN AND SETTING: A cross-sectional study was conducted at the outpatient department of the Third Affiliated Hospital of Qiqihar Medical College from January 2022 to December 2024.

METHODS: We consecutively enrolled 234 patients with CTTH diagnosed according to International Classification of Headache Disorders, 3rd edition (ICHD-3) criteria were consecutively enrolled. Hyperacusis was confirmed via otolaryngological evaluation and audiometric testing (pure-tone thresholds and loudness discomfort levels [LDL]). Demographic data, headache features (frequency, intensity, and duration), hyperacusis distress severity, and psychological measures (Generalized Anxiety Disorder-7 [GAD-7], Self-rating Somatic Symptom Scale [SSS], and Pittsburgh Sleep Quality Index [PSQI]) were analyzed. Univariate analyses (Student's *t*-test and Chi-squared test) and multivariate logistic regression adjusted for potential confounders were performed to compare patients with (*n* = 47) and without (*n* = 187) hyperacusis.

RESULTS: Hyperacusis prevalence was 20.09% (47/234). Compared to non-hyperacusis patients, hyperacusis patients exhibited longer headache duration (13.9 ± 3.15 versus 8.11 ± 2.74 years, *P* = 0.01), higher weekly headache frequency (7.30 ± 1.99 versus 4.19 ± 1.2 , *P* = 0.005), more pericranial tenderness (95.74% versus 77.54%, *P* = 0.008), and poorer sleep quality (PSQI: 14.01 ± 1.13 versus 9.17 ± 0.51 , *P* = 0.004). After adjusting for age, sex, headache duration, and pain intensity in multivariate logistic regression, higher headache frequency (odds ratio [OR] = 1.42, 95% confidence interval [CI]: 1.12–1.79, *P* = 0.003) and higher PSQI scores (odds ratios = 1.38, 95% CI: 1.15–1.66, *P* = 0.001) remained independently associated with hyperacusis. Common triggering sounds included children crying (61.7%) and traffic noise (42.55%), with predominant emotional responses being irritability (76.6%) and anxiety (65.96%). LDL classification (analyzed as an ordinal variable) was correlated with headache duration (*P* < 0.05), while hyperacusis distress severity was linked to headache frequency and PSQI scores (*P* < 0.05).

CONCLUSION: CTTH with hyperacusis represents a distinct subgroup characterized by a history of prolonged headaches, increased attack frequency, and significant sleep disturbances. The independent associations identified suggest that sound sensitivity is closely linked to headache burden and sleep quality; however, causal relationships require longitudinal investigation. Integrated management addressing both headache and auditory hypersensitivity is warranted.

INTRODUCTION

Chronic tension-type headache (CTTH) is one of the most prevalent primary headache disorders, characterized by frequent bilateral pressing pain that significantly impacts quality of life.¹ Although the pathophysiology of CTTH remains incompletely understood, peripheral and central sensitization mechanisms are thought to play key roles.² Notably, a subset of patients with CTTH experience comorbid hyperacusis—an abnormal intolerance to everyday sounds—suggesting potential shared neural mechanisms involving sensory processing dysregulation.³ However, the clinical features and implications of hyperacusis in patients with CTTH have not been systematically investigated.⁴

Emerging evidence suggests that hyperacusis may not merely be an accompanying symptom but rather may be associated with more severe central sensitization in headache disorders.⁵

Patients with coexisting CTTH and hyperacusis often report a greater headache burden, emotional distress, and functional impairment.⁶ However, critical questions remain unanswered: What distinguishes patients with CTTH and hyperacusis from those without? How does hyperacusis severity correlate with headache profiles? And what are the most troublesome sound triggers in this population? Addressing these questions could provide insights into phenotypic variability and guide targeted therapeutic strategies.⁷

This study aimed to comprehensively characterize patients with CTTH harboring hyperacusis by examining (1) demographic and headache profile differences between patients with and without hyperacusis; (2) auditory sensitivity patterns and their emotional impacts; and (3) relationships between hyperacusis measures and headache characteristics. Given the observational nature of this study, we focused on describing associations rather than inferring causal mechanisms. Our findings may help identify a clinically relevant CTTH subtype that requires multidisciplinary management approaches.⁸

METHODS

Participants

Consecutive patients diagnosed with CTTH who visited the outpatient department of the Third Affiliated Hospital of Qiqihar Medical College from January 2022 to December 2024 were included in this study. The diagnostic criteria for CTTH were based on the International Classification of Headache Disorders, 3rd edition (ICHD-3),⁹ demonstrating recurrent headache episodes lasting 30 min to 7 days, characterized by bilateral pressing or tightening pain of mild to moderate intensity, not aggravated by routine physical activity, and without nausea or vomiting (photophobia or phonophobia may be present). Additionally, headaches must occur on ≥ 15 days per month (≥ 180 days per year) for ≥ 3 months, meeting the essential features of tension-type headache while excluding medication-overuse headache.

All participants underwent medical history collection and routine otolaryngological examinations, with no history of otitis media or ear surgery, normal external ear canals and tympanic membranes, and with negative central nervous system physical examination results. Brain magnetic resonance imaging (MRI) ruled out intracranial and middle ear lesions, superior semicircular canal dehiscence, and brain tumors. Tympanometry revealed Type A curves, and stapedia reflexes were elicited with thresholds within the normal range.

Hyperacusis was diagnosed based on the following criteria:¹⁰ (1) patient-reported intolerance to everyday sounds not attributable to peripheral auditory pathology; (2) otolaryngological confirmation after excluding external, middle, and inner ear disorders; (3) mean loudness discomfort level (LDL) across tested

frequencies ≤ 90 decibel hearing level (dB HL); and (4) symptoms are present for at least 3 months. Patients with a mean LDL > 90 dB HL who reported significant sound intolerance underwent a clinical interview to determine hyperacusis diagnosis based on functional impact.

The exclusion criteria were as follows: other primary or secondary headaches; medication overuse headache; history of head or neck trauma; systemic degenerative diseases such as rheumatoid arthritis or systemic lupus erythematosus; fibromyalgia syndrome; alcohol dependence or substance abuse; pregnancy; hyperacusis accompanied by tinnitus, where tinnitus was the most severe symptom affecting daily life; and severe psychiatric disorders.

All participants provided written informed consent before enrollment. The study was approved by the Ethics Committee of Qiqihar Medical College (2022-2037) and complied with the ethical requirements for clinical research in China and the Declaration of Helsinki.

Collection of demographic and clinical data

Demographic and clinical data, including sex, age, height, weight, education level, smoking history, alcohol consumption history, diabetes, and hypertension, were collected. All participants maintained an 8-week headache diary to confirm their diagnosis and record headache clinical features. In the diary, patients recorded headache frequency (days per week), numeric pain rating scale (0 = no pain, 10 = worst pain),¹¹ and duration of each headache episode (hours per day).

Participants confirmed by otolaryngologists to have hyperacusis were further assessed for the onset of hyperacusis, types of triggering sounds, emotional responses to triggering sounds, and the impact of hyperacusis. The degree of distress caused by hyperacusis was graded as follows:¹² No distress: Not bothered at all. Mild distress: Slight discomfort, not particularly bothered by surrounding sounds, no impact on daily life. Moderate distress: Frequently aware of surrounding sounds, reduced concentration, occasional use of earplugs or avoidance of sounds, slight impact on daily life and work. Severe distress: Significant impact, frequent use of earplugs, avoidance of noisy environments, frequent irritability, moderate impact on daily life and work. Extreme distress: Severe impact, inability to tolerate any sound, fear of going outside, frequent anxiety or fear, near inability to maintain normal daily life, work, or study.

Questionnaire assessments

The researchers involved in the study received training regarding the relevant assessment tools and selected appropriate times to administer the questionnaires. Questionnaires were distributed to all participants, collected on site, and immediately checked for completeness. If any omissions or issues affecting quality were

found, participants were asked to promptly supplement or revise their responses.

1. Generalized Anxiety Disorder-7 (GAD-7):¹³ A Chinese-validated version with good reliability was used to screen and assess the severity of anxiety symptoms. The scale consists of seven items, each scored from 0–3, with higher scores indicating greater anxiety. A total score of 0–4 indicates no anxiety, 5–9 mild anxiety, 10–14 moderate anxiety, and ≥ 15 severe anxiety.
2. Self-rating Somatic Symptom Scale (SSS):¹⁴ A Chinese-validated version was used for self-assessment of somatic symptoms. The scale includes 20 items, each scored from 1–4, with higher scores indicating more severe somatic symptoms. Based on national norms, a total score ≥ 40 was considered positive for somatic symptom screening.
3. Pittsburgh Sleep Quality Index (PSQI):¹⁵ A Chinese-validated version was used to assess sleep quality. The scale consists of 18 items grouped into seven components: sleep latency, sleep frequency, sleep quality, sleep disturbances, sleep duration, sleep medication use, and daytime dysfunction. Total scores range from 0–21, with higher scores indicating worse sleep quality.

Pure-tone audiometry and LDL testing

For CTTH patients with confirmed hyperacusis, pure-tone audiometry was performed using the Tinni Test[®] tinnitus diagnostic and therapeutic system (Sichuan Micro-Digital Co., Ltd., Chengdu, China, model Tinni Test TTS-1000A). Air conduction thresholds were measured at 0.25, 0.5, 1, 2, 4, and 8 kHz. LDL testing was conducted using pure-tone stimuli at the same frequencies, starting at 70 dB above the threshold and adjusted in 5 dB steps until the patient reported discomfort. The mean LDL was calculated for each patient. Based on Goldstein's classification,¹⁶ LDL was divided into four ordinal categories: Level 1 (≥ 95 dB at all frequencies), Level 2 (80–90 dB at ≥ 2 frequencies), Level 3 (65–75 dB at ≥ 2 frequencies), and Level 4 (≤ 60 dB at ≥ 2 frequencies).

Statistical analysis

Statistical analyses were performed using GraphPad Prism v.6. Normally distributed continuous data are expressed as means $\pm$ standard deviation (SD) and compared using independent Student's *t*-tests or analysis of variance (ANOVA). Categorical data are presented as counts and were compared using Chi-squared tests. To identify factors independently associated with hyperacusis, multivariate logistic regression analysis was performed with hyperacusis (present/absent) as the dependent variable. Independent variables included demographic factors (age, gender) and clinical variables that showed significant differences in univariate analyses (headache duration, headache frequency, pericranial tenderness, and PSQI scores). Variables were entered

using forward stepwise selection with an entry criterion $P < 0.05$ and a removal criterion $P > 0.1$. Results are presented as odds ratios (ORs) with 95% CIs. Spearman's rank correlation was used to assess relationships between variables, with LDL classification treated as an ordinal variable. Spearman's ρ was calculated with distress levels treated as ordered categories for correlations involving hyperacusis distress severity. $P < 0.05$ was considered statistically significant.

RESULTS

Comparison of demographic and clinical characteristics between CTTH patients with and without hyperacusis

A total of 234 patients with CTTH were enrolled, of whom 47 (20.09%) had hyperacusis. Compared to patients without hyperacusis, those with hyperacusis had longer headache durations ($P = 0.01$), higher headache frequencies ($P = 0.005$), more frequent pericranial tenderness ($P = 0.008$), and higher PSQI scores ($P = 0.004$). No significant differences were observed in age, sex, body mass index, smoking, alcohol consumption, education level, diabetes, hypertension, headache pain intensity, headache duration, GAD-7 scores, or SSS scores ($P > 0.05$) (Table 1).

Multivariate logistic regression analysis was performed to identify factors independently associated with hyperacusis. After adjusting for age, sex, headache duration, and pain intensity, two variables remained significantly associated with hyperacusis: higher headache frequency (OR = 1.42, 95% CI = 1.12–1.79, $P = 0.003$) and higher PSQI scores (OR = 1.38, 95% CI = 1.15–1.66, $P = 0.001$). Pericranial tenderness exhibited a trend toward significance (OR = 3.21, 95% CI = 0.98–10.52, $P = 0.014$). The model demonstrated good fit (Hosmer–Lemeshow $\chi^2 = 6.84$, $P = 0.554$) (Table 2).

The dependent variable was the presence of hyperacusis (coded as 1 for present and 0 for absent). Independent variables were entered as follows: pericranial tenderness (1 for present and 0 for absent), headache frequency (times per week; continuous), and PSQI score (continuous, 0–21). Variables included in the model were selected based on P values < 0.1 in univariate analyses and clinical relevance. A Hosmer–Lemeshow test P value > 0.05 indicated good model fit. Bold type indicates statistical significance ($P < 0.05$).

Characteristics of hyperacusis

Among the 47 CTTH patients with hyperacusis, the mean duration of hyperacusis was 2.84 ± 0.31 years. The distribution of hyperacusis duration was as follows: < 6 months, 7 (14.89%); 7–12 months, 10 (21.28%); 1–5 years, 22 (46.81%); and ≥ 6 years, 8 (17.02%). Hyperacusis symptoms worsened progressively in eight (17.02%) patients, whereas 25 (53.19%) reported

fluctuating symptoms. Additionally, 17 (36.17%) patients experienced ear fullness, pressure, itching, or echo sensations triggered by sounds, and 9 (19.15%) reported periauricular numbness or burning sensations.

Hearing was normal (speech-frequency pure-tone average ≤ 25 dB HL) in 35 (74.47%) patients, whereas 12 (25.53%) had unilateral sensorineural hearing loss (> 25 dB HL), with nine (19.15%) affecting the left ear and three (6.38%) the right ear. Based on hyperacusis distress severity, no distress: 4 (8.51%), mild distress: 17 (36.17%), moderate distress: 21 (44.68%), and severe distress: 5 (10.63%). The LDL classification (analyzed as an ordinal variable) was as follows: level 1 (≥ 95 dB at all frequencies): 8 (17.02%), level 2 (80–90 dB at ≥ 2 frequencies): 12 (25.53%), level 3 (65–75 dB at ≥ 2 frequencies): 12 (25.53%), and level 4 (≤ 60 dB at ≥ 2 frequencies): 10 (21.28%). Spearman's correlation analysis revealed no significant relationship between the LDL classification and hyperacusis distress severity ($\rho = 0.04$, $P > 0.05$).

Types of triggering sounds and emotional responses

The 47 patients reported discomfort to 10 types of sounds, most commonly: children crying/screaming: 29 (61.7%), crowded city noise: 26 (55.32%), traffic noise: 20 (42.55%), sudden high-pitched sounds: 18 (38.3%), loud laughter/cheering: 14 (29.79%), television sounds: 13 (27.66%). Each patient reported at least three negative emotional responses to triggering sounds, including: irritability: 36 (76.6%), anxiety: 31 (65.96%), reduced concentration: 20 (42.55%), and anger/irritability: 12 (25.53%).

Correlation between hyperacusis and CTTH

Spearman's correlation analysis was used to assess relationships between LDL classification (as an ordinal variable), hyperacusis distress severity, and CTTH clinical features (headache duration, frequency, pericranial tenderness, and PSQI scores). The results showed that LDL classification (ordinal) was positively correlated with headache duration and pericranial tenderness, and

Table 1. Comparison of demographic and clinical characteristics between CTTH patients with and without hyperacusis.

Item	CTTH patients without hyperacusis (n = 187)	CTTH patients with hyperacusis (n = 47)	Test values	P values
Age (years, mean $\pm$ SD)	33.51 $\pm$ 11.4	32.56 $\pm$ 13.91	1.302	0.194
Female [n (%)]	112 (59.89)	30 (63.82)	0.107	0.734
BMI (kg/m ² , mean $\pm$ SD)	24.36 $\pm$ 2.51	23.97 $\pm$ 3.51	1.2	0.161
Smoking [n (%)]	19 (10.16)	4 (8.51)	0.004	0.948
Alcohol consumption [n (%)]	31 (16.58)	10 (21.28)	0.295	0.587
Duration of education (years, mean $\pm$ SD)	13.91 $\pm$ 3.51	14.56 $\pm$ 4.69	0.91	0.105
Diabetes [n (%)]	2 (1.07)	0 (0)	0.03	0.861
Hypertension [n (%)]	6 (18.38)	2 (20.51)	0.009	0.924
Headache duration (years, mean $\pm$ SD)	8.11 $\pm$ 2.74	13.90 $\pm$ 3.15	4.19	0.01
Headache pain score (mean $\pm$ SD)	5.19 $\pm$ 1.25	5.31 $\pm$ 1.68	1.301	0.069
Headache attack frequency (times/week, mean $\pm$ SD)	4.19 $\pm$ 1.2	7.30 $\pm$ 1.99	6.194	0.005
Headache duration (hours, mean $\pm$ SD)	6.17 $\pm$ 1.45	6.31 $\pm$ 1.73	1.71	0.061
Headache with pericranial tenderness [n (%)]	145 (77.54)	45 (95.74)	7.004	0.008
GAD-7 score (mean $\pm$ SD)	7.10 $\pm$ 1.69	7.19 $\pm$ 2.11	1.819	0.094
SSS score (mean $\pm$ SD)	8.18 $\pm$ 1.93	8.31 $\pm$ 1.7	1.104	0.18
PSQI score (mean $\pm$ SD)	9.17 $\pm$ 2.4	14.01 $\pm$ 3.91	6.194	0.004

Table 2. Multivariate logistic regression analysis of factors independently associated with hyperacusis in CTTH patients.

Variable	β	S.E.	Wald χ^2	OR	95% CI	P value
Headache frequency (times/week)	0.351	0.118	8.85	1.42	1.12–1.79	0.003
PSQI score	0.322	0.095	11.49	1.38	1.15–1.66	0.001
Pericranial tenderness (present vs. absent)	1.167	0.605	3.72	3.21	0.98–1.52	0.014
Constant	–5.873	1.402	17.55	0.003	–	< 0.001

Table 3. Correlation between hyperacusis and CTTH.

	Headache durations	Pericranial tenderness	Headache frequency	PSQI scores
LDL classification	$\rho = 0.612$ $P = 0.013$	$\rho = 0.854$ $P = 0.001$	$\rho = 0.266$ $P = 0.184$	$\rho = 0.195$ $P = 0.372$
Hyperacusis distress severity	$\rho = 0.304$ $P = 0.184$	$\rho = 0.269$ $P = 0.092$	$\rho = 0.704$ $P = 0.011$	$\rho = 0.761$ $P = 0.008$

hyperacusis distress severity was correlated with headache frequency and PSQI scores (Table 3).

DISCUSSION

The present study provides novel insights into the clinical characteristics of patients with CTTH harboring comorbid hyperacusis and revealed several important findings. First, our results demonstrate that approximately 20% of patients with CTTH exhibit hyperacusis, suggesting that this comorbidity is relatively common in this population. This prevalence rate is notably higher than that reported in the general population (estimated at 8–15%),¹⁷ suggesting a potential association between CTTH and auditory hypersensitivity.

The observed differences between patients with and without hyperacusis in the CTTH group are particularly noteworthy. Patients with hyperacusis had significantly longer headache duration, higher attack frequency, and more frequent pericranial tenderness, all of which are indicators of more severe and chronic disease. Multivariate analysis revealed that higher headache frequency and poorer sleep quality were independently associated with hyperacusis after adjusting for potential confounders. This suggests that these factors may represent key features of the CTTH-hyperacusis phenotype, rather than being mere correlates. These findings align with the concept of central sensitization, in which prolonged nociceptive input from pericranial muscles may lead to heightened sensitivity in both the trigeminal and auditory systems.¹⁸ However, given the cross-sectional design of this study, we cannot determine whether central sensitization precedes or results from the observed associations. Longitudinal studies are needed to establish temporal relationships. The strong association between pericranial tenderness and hyperacusis (95.74% versus 77.54%) further supports the possibility of a shared mechanism.¹⁹

Our study also revealed that hyperacusis in patients with CTTH is associated with significantly worse sleep quality, as evidenced by higher PSQI scores. The independent association between PSQI scores and hyperacusis in the multivariate analysis strengthens this finding. This finding is clinically relevant, as poor sleep is known to exacerbate pain sensitivity and may contribute to a vicious cycle of headache chronification.²⁰ Notably, the direction of this association cannot be determined from our data; poor sleep may increase susceptibility to hyperacusis, hyperacusis may disrupt sleep, or both may share common underlying mechanisms. Interestingly, although anxiety and somatic symptom scores did not differ between the groups, the specific emotional responses to triggering sounds (notably irritability and anxiety in > 65% of patients) suggest that hyperacusis may represent a distinct dimension of psychological distress in CTTH patients.²¹ This apparent discrepancy may reflect the difference between state-like emotional responses to specific sound triggers versus trait-like anxiety as measured by the GAD-7 scaling.

The pattern of sound sensitivity observed in our cohort, particularly to high-pitched and sudden sounds, such as children crying (61.7%) and traffic noise (42.55%), mirrors the findings in migraine-related hyperacusis,²² suggesting possible commonalities in auditory processing abnormalities across primary headache disorders. However, the lack of correlation between LDL measures and subjective hyperacusis distress highlights the complex nature of sound intolerance, in which psychological factors may modulate perceptual sensitivity independent of pure physiological thresholds.²³ This dissociation underscores the importance of assessing both objective auditory measures and subjective distress in clinical evaluations.

Several limitations should be considered. First, the cross-sectional study design precludes causal inferences concerning the relationship between CTTH and hyperacusis.²⁴ Our findings represent associations that require confirmation in longitudinal studies to establish temporal sequences and potential causal pathways. Second, we did not assess other potential contributors to central sensitization, such as cutaneous allodynia.²⁵ Third, the single-center nature of the study may limit its generalizability.²⁶ Fourth, although multivariate analysis adjusted for several potential confounders, residual confounding from unmeasured variables (e.g., medication use, stress levels, and psychiatric comorbidities) cannot be excluded. Future longitudinal studies incorporating quantitative sensory testing and neuroimaging could help to elucidate the neural mechanisms underlying this comorbidity.²

From a clinical perspective, our findings suggest that assessment of sound sensitivity should be routinely included in the evaluation of patients with CTTH, particularly those with longstanding symptoms and pericranial tenderness.²⁷ The identification of hyperacusis may help identify a distinct CTTH subgroup that could benefit from targeted interventions addressing both headache and auditory hypersensitivity, potentially including cognitive-behavioral approaches for sound intolerance alongside conventional headache therapies.²⁸ However, interventional studies are needed to determine whether such integrated approaches improve outcomes in this patient population.

CONCLUSION

This cross-sectional study demonstrated that approximately 20% of patients with CTTH present with comorbid hyperacusis, characterized by longer headache duration, higher attack frequency, more frequent pericranial tenderness, and significantly worse sleep quality compared to those without hyperacusis. Multivariate analysis identified higher headache frequency and poorer sleep quality as independent factors associated with hyperacusis in this population. The pattern of sound sensitivity predominantly involves high-pitched and sudden sounds, with irritability and anxiety being the most common emotional

responses. These findings suggest that CTTH with hyperacusis represents a clinically distinct subgroup that requires integrated management strategies. Future longitudinal studies are warranted to establish causal relationships and evaluate the efficacy of targeted interventions addressing both headache and auditory hypersensitivity.

REFERENCES

- Exposto CR, Mansoori M, Bech BH, Baad-Hansen L. Prevalence of painful temporomandibular disorders and overlapping primary headaches among young adults. *Eur J Pain*. 2025 May;29(5):e70013. PMID: 40152614; <https://doi.org/10.1002/ejp.70013>.
- Bendtsen L, Evers S, Linde M, Mitsikostas DD, Sandrini G, Schoenen J; EFNS. EFNS guideline on the treatment of tension-type headache: report of an EFNS task force. *Eur J Neurol*. 2010 Nov;17(11):1318–25. PMID: 20482606; <https://doi.org/10.1111/j.1468-1331.2010.03070.x>.
- Wilcox SL, Veggeberg R, Lemme J, et al. Increased functional activation of limbic brain regions during negative emotional processing in migraine. *Front Hum Neurosci*. 2016 Jul 26;10:366. PMID: 27507939; <https://doi.org/10.3389/fnhum.2016.00366>.
- Paulin J, Andersson L, Nordin S. Characteristics of hyperacusis in the general population. *Noise Health*. 2016 Jul–Aug;18(83):178–84. PMID: 27569405; <https://doi.org/10.4103/1463-1741.189244>.
- Demarquay G, Royet JP, Mick G, Rylvlin P. Olfactory hypersensitivity in migraineurs: a H(2)(15)O-PET study. *Cephalalgia*. 2008 Oct;28(10):1069–80. PMID: 18727640; <https://doi.org/10.1111/j.1468-2982.2008.01672.x>.
- Ralli M, Lobarinas E, Fetoni AR, Stolzberg D, Paludetti G, Salvi R. Comparison of salicylate- and quinine-induced tinnitus in rats: development, time course, and evaluation of audiologic correlates. *Otol Neurotol*. 2010 Jul;31(5):823–31. PMID: 20502380; <https://doi.org/10.1097/MAO.0b013e3181de4662>.
- Meylakh N, Henderson LA. Exploring alterations in sensory pathways in migraine. *J Headache Pain*. 2022 Jan 12;23(1):5. PMID: 35021998; <https://doi.org/10.1186/s10194-021-01371-y>.
- Jastreboff PJ, Gray WC, Gold SL. Neurophysiological approach to tinnitus patients. *Am J Otol*. 1996 Mar;17(2):236–40. PMID: 8723954.
- Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders, 3rd edition. *Cephalalgia*. 2018 Jan;38(1):1–211. <https://doi.org/10.1177/0333102417738202>.
- Schecklmann M, Landgrebe M, Langguth B; TRI Database Study Group. Phenotypic characteristics of hyperacusis in tinnitus. *PLoS One*. 2014 Jan 31;9(1):e86944. PMID: 24498000; <https://doi.org/10.1371/journal.pone.0086944>.
- Hjermstad MJ, Fayers PM, Haugen DF, et al; European Palliative Care Research Collaborative (EPCRC). Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011 Jun;41(6):1073–93. PMID: 21621130; <https://doi.org/10.1016/j.jpainsymman.2010.08.016>.
- Jahn KN. Clinical and investigational tools for monitoring noise-induced hyperacusis. *J Acoust Soc Am*. 2022 Jul;152(1):553. PMID: 35931527; <https://doi.org/10.1121/10.0012684>.
- Löwe B, Decker O, Müller S, et al. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care*. 2008 Mar;46(3):266–74. PMID: 18388841; <https://doi.org/10.1097/MLR.0b013e318160d093>.
- Kroenke K, Spitzer RL, Williams JB, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. *Gen Hosp Psychiatry*. 2010 Jul–Aug;32(4):345–59. PMID: 20633738; <https://doi.org/10.1016/j.genhosppsych.2010.03.006>.
- Buyssse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res*. 1989 May;28(2):193–213. PMID: 2748771; [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4).
- Tyler RS, Pienkowski M, Roncancio ER, et al. A review of hyperacusis and future directions, part 1. Definitions and manifestations. *Am J Audiol*. 2014 Dec;23(4):402–19. PMID: 25104073; https://doi.org/10.1044/2014_AJA-14-0010.
- Rosing SN, Schmidt JH, Wedderkopp N, Baguley DM. Prevalence of tinnitus and hyperacusis in children and adolescents: a systematic review. *BMJ Open*. 2016 Jun 3;6(6):e010596. PMID: 27259524; <https://doi.org/10.1136/bmjopen-2015-010596>.
- Burstein R, Nosedá R, Borsook D. Migraine: multiple processes, complex pathophysiology. *J Neurosci*. 2015 Apr 29;35(17):6619–29. PMID: 25926442; <https://doi.org/10.1523/JNEUROSCI.0373-15.2015>.
- Ashina S, Bendtsen L, Lyngberg AC, Lipton RB, Hajjiyeva N, Jensen R. Prevalence of neck pain in migraine and tension-type headache: a population study. *Cephalalgia*. 2015 Mar;35(3):211–9. PMID: 24853166; <https://doi.org/10.1177/0333102414535110>.
- Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. *J Pain*. 2013 Dec;14(12):1539–52. PMID: 24290442; <https://doi.org/10.1016/j.jpain.2013.08.007>.
- Main A, Dowson A, Gross M. Photophobia and phonophobia in migraineurs between attacks. *Headache*. 1997 Sep;37(8):492–5. PMID: 9329231; <https://doi.org/10.1046/j.1526-4610.1997.3708492.x>.
- Aazh H, Knipper M, Danesh AA, et al. Insights from the third international conference on hyperacusis: causes, evaluation, diagnosis, and treatment. *Noise Health*. 2018 Jul–Aug;20(95):162–70. PMID: 30136676; https://doi.org/10.4103/nah.NAH_2_18.
- Sedgwick P. Bias in observational study designs: cross sectional studies. *BMJ*. 2015 Mar 6;350:h1286. PMID: 25747413; <https://doi.org/10.1136/bmj.h1286>.
- Lipton RB, Bigal ME, Ashina S, et al; American Migraine Prevalence Prevention Advisory Group. Cutaneous allodynia in the migraine population. *Ann Neurol*. 2008 Feb;63(2):148–58. PMID: 18059010; <https://doi.org/10.1002/ana.21211>.
- Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials*. 2006 May;1(1):e9. PMID: 16871331; <https://doi.org/10.1371/journal.pctr.0010009>.

26. De Ridder D, Elgoyhen AB, Romo R, Langguth B. Phantom percepts: tinnitus and pain as persisting aversive memory networks. *Proc Natl Acad Sci USA*. 2011 May 17;108(20):8075–80. PMID: 21502503; <https://doi.org/10.1073/pnas.1018466108>.
27. Silvia M, Smith AM. Development and feasibility of the headache-related light and sound sensitivity inventories in youth. *Children (Basel)*. 2021 Sep 28;8(10):861. PMID: 34682126; <https://doi.org/10.3390/children8100861>.
28. Jüris L, Andersson G, Larsen HC, Ekselius L. Cognitive behaviour therapy for hyperacusis: a randomized controlled trial. *Behav Res Ther*. 2014 Mar;54:30–7. PMID: 24508581; <https://doi.org/10.1016/j.brat.2014.01.001>.

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)

Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)

Authors' contributions: Fu Z: protocol or project development, data collection and analysis, conception and design, manuscript writing and editing, critical revision of the manuscript; Jin Y: data collection and analysis, statistical analysis; Lin Z: data and statistical analysis; Liu Y: data collection; Guo K: manuscript writing.

Sources of funding: Scientific research project of the Heilongjiang Provincial Health Commission (No. 20220303070603). Fundamental Research Funds for Undergraduate Universities in Heilongjiang Province (No. 2022-KYYWF-0817).

Conflicts of interest: None.

Data availability statement: The data used in this study contain sensitive or private information that cannot be shared publicly in accordance with the Ethics Committee of Qiqihar Medical College. Data are available from the authors upon approval by the relevant committee.

Declaration of generative AI in scientific writing: The authors used generative artificial intelligence tools to improve the English expression and native fluency of this manuscript. The AI was only applied for language editing and polishing purposes. The intellectual content, study design, data analysis, and conclusions remain the sole work of the authors. All authors have checked and approved the final text.

Date of first submission: May 29, 2025

Last received: March 10, 2026

Accepted: March 9, 2026

Address for correspondence:

Zenghui Fu

Department of Neurology, The Third Affiliated Hospital of Qiqihar Medical University

3 Taishun Street

Tiefeng — Qiqihar — China

PO Box 161002

Tel. (+86) 0452-212-0351

E-mail: fzh92@foxmail.com



Predictive modeling with machine learning for low birth weight in the state of São Paulo: a retrospective population-based study

Allbert Velleniche de Aquino Almeida^I, Adriana de Oliveira Ribeiro dos Santos^{II}, Luiz Fernando Costa Nascimento^{III}

Departamento de Engenharia, Faculdade de Engenharia e Ciências, Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), Guaratinguetá (SP), Brazil

^IMSc. Doctoral Student; Departamento de Engenharia, Faculdade de Engenharia e Ciências, Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), Guaratinguetá (SP), Brazil.

^{ID} <https://orcid.org/0000-0001-7482-529X>

^{II}PhD. Professor; Departamento de Medicina, Universidade de Taubaté, Taubaté (SP), Brazil.

^{ID} <https://orcid.org/0000-0002-5941-9522>

^{III}PhD. Professor; Departamento de Engenharia, Faculdade de Engenharia e Ciências, Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp), Guaratinguetá (SP), Brazil.

^{ID} <https://orcid.org/0000-0001-9793-750X>

KEYWORDS (MeSH terms):

Infant, low birth weight.
Machine learning.
Predictive learning model.
Logistic models.
Random forest.

AUTHOR'S KEYWORDS:

Infant, newborn.
Pregnancy.
Prenatal care.
Health policy.
Data mining.
Artificial intelligence.

ABSTRACT

BACKGROUND: Low birth weight (LBW) is a major public health concern associated with increased neonatal morbidity and mortality.

OBJECTIVE: This study was aimed to develop a predictive learning model using machine learning techniques to identify LBW from live birth data.

METHODS: A retrospective population-based study was conducted using data from the Brazilian Live Birth Information System (SINASC), which includes all live births registered in the state of São Paulo between 2019 and 2023. After data curation and exclusion of records with missing or unknown information, a total of 2,548,570 live births were analyzed. Predictors included maternal sociodemographic characteristics, obstetric history, prenatal care indicators, and newborn characteristics. Logistic regression and random forest models were trained using an 80/20 train-test split, with class imbalance addressed through class weighting. Model performance was evaluated using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).

RESULTS: The analysis showed that the balanced logistic regression model performed best in terms of sensitivity (0.688), F1-score (0.567), and AUC-ROC (0.850). A proportional analysis of the variables between the LBW and normal-weight groups revealed a higher prevalence of LBW among mothers with low education levels, without a partner, of Black or Brown ethnicity, aged ≤ 19 years or ≥ 35 years, with multiple pregnancies, with a reduced number of prenatal consultations, and with late initiation of prenatal care.

CONCLUSIONS: The findings highlight the importance of large-scale data analysis supported by machine learning techniques to inform public policies aimed at preventing LBW and promoting maternal and child health.

INTRODUCTION

Low birth weight (LBW), defined by the World Health Organization (WHO) as birth weight $< 2,500$ g, is widely recognized as a key indicator of neonatal morbidity and mortality and is associated with short- and long-term complications in child development.¹ Approximately 15% of births worldwide² and approximately 10% in Brazil³ are LBW, underscoring the magnitude of the problem and the importance of preventive strategies.

Several factors are associated with LBW, including unfavorable socioeconomic conditions, extreme maternal age, low educational level, obstetric history, prematurity, multiple pregnancies, and inadequate prenatal care.⁴⁻⁷ Recent studies on prematurity and cesarean delivery by Abdi et al. and Rezaei et al.^{8,9} highlight advanced maternal age, obstetric history, and maternal comorbidities as relevant risk factors.⁴⁻⁷

Many of these factors can be effectively addressed during pregnancy if identified early, reinforcing the importance of vigilant and qualified prenatal monitoring.^{5,10,11}

In recent years, the application of artificial intelligence (AI) techniques, especially machine learning (ML), has emerged as an innovative and effective approach in public health, enabling the analysis of large volumes of data and the development of robust predictive models.^{8,9,12-15} Models based on algorithms such as decision trees, logistic regression, artificial neural networks, and random forest have been successfully used to predict clinical outcomes, including LBW.^{8,9,16-18}

Recent studies have shown that predictive models using population data can achieve good accuracy in the early identification of pregnancies at risk for LBW, aiding in the planning of targeted interventions and efficient resource allocation.^{19,20} In Brazil, the Live Birth Information System (SINASC) provides a comprehensive, nationwide database on births, which includes relevant information regarding the mother, the pregnancy, the delivery, and the newborn. This makes it a valuable source for epidemiological studies and predictive modeling.³

Therefore, this study aimed to develop and validate ML models capable of predicting LBW occurrences in the state of São Paulo using SINASC data.

METHODOLOGY

This study was conducted using data from the state of São Paulo, which comprises 645 municipalities and has the largest population among Brazilian states, with approximately 45.9 million inhabitants according to estimates by the Instituto Brasileiro de Geografia e Estatística (IBGE).²¹ This territorial and demographic diversity makes the state a representative sample of the national population with varying socioeconomic contexts, levels of healthcare access, and human development indicators. These characteristics justify the choice of São Paulo as the study setting, allowing for broader analyses with the potential to be generalized to other regions of the country.

This applied research followed a methodological procedure consisting of problem definition, literature review, research design, data collection, data analysis, interpretation of results, and presentation of conclusions. Each stage required careful attention and methodological rigor to ensure the quality and relevance of the results.

The methodology adopted a quantitative and qualitative approach, combining the analysis of documentary data obtained from public databases with the use of ML techniques for prediction.

Data from SINASC referring to São Paulo covering five full years of records (from 2019 to 2023) were used. With approximately 600,000 births per year in the state, the sample included a wide range of cases, ensuring representativeness and diversity.

The SINASC data include numerous variables regarding live births, some of which are repeated (e.g., maternal education, which is recorded as three different variables with historical and categorized formats). To optimize processing and avoid redundant information, highly correlated variables were eliminated. The selection of the variables considered in the study was based on those commonly used in previous studies on LBW prediction and in those that were effectively available in SINASC records. This strategic selection contributed to the development of leaner, more interpretable models with good predictive performance.

Many studies have employed variables from SINASC to analyze LBW, including maternal age, marital status, education, race/color,

occupation, parity, pregnancy type (single or multiple), delivery type (cesarean or normal), previous fetal loss, number of prenatal visits, month when prenatal care started, gestational age, delivery location, and newborn sex and weight. Other common variables include the Apgar score and the municipality of residence. These sociodemographic, obstetric, and perinatal factors directly influence birth outcomes and have been widely used in LBW predictive models.

This study selected variables related to maternal medical history, socioeconomic factors, and prenatal health indicators to ensure the inclusion of relevant predictors for LBW. Selected variables include:

- a. **Maternal medical history:** age at delivery, number of previous child deaths, parity (birth order), and pregnancy type (single, twin, triplet, or more).
- b. **Socioeconomic factors:** education level (categories 0 to 5 according to SINASC), marital status (married/stable union and widowed/single/separated), and maternal race/color (White, Black, Yellow, Brown, Indigenous).
- c. **Prenatal health indicators:** gestational age category (preterm, term, post-term); number of prenatal visits; month prenatal care began; detected congenital anomalies (yes/no); and newborn sex (male/female).

Prior to model development, a rigorous data curation process was performed to ensure data quality and analytical consistency. All records containing missing values, inconsistent entries, or including categories such as “ignored” or “not informed” for key variables were excluded. After this selection process, 2,548,570 valid live birth records from an initial dataset of 2,677,170 records were retained.

Feature selection was guided by a combination of epidemiological relevance and data availability within the SINASC. Variables with high redundancy or strong collinearity—such as multiple representations of maternal education—were evaluated, and only the most informative one was retained. This approach was aimed to reduce dimensionality, improve model interpretability, and avoid overfitting. No automated feature selection algorithms were applied; instead, selection was based on prior literature and domain knowledge regarding established risk factors for LBW.

In the preprocessing step, a “low weight” binary variable was defined by assigning a 0 to births with normal weight ($\geq 2,500$ g) and a 1 to those with LBW ($< 2,500$ g).

SINASC data were accessed via DATASUS, converted into comma-separated values (.csv) files using TabWin version 4.15 (Departamento de Informática do Sistema Único de Saúde – DATASUS, Brasília, DF), and stored in Google Drive to facilitate collaborative analyses use through Google Colab.

Data processing and analysis were performed using Python in Google Colab, which can handle large datasets with virtual machines. Packages such as NumPy and Pandas supported data processing, and scikit-learn was used to implement ML algorithms.

Several ML algorithms were tested, including logistic regression and random forest. Logistic regression is a widely used statistical model for binary classification, whereas random forest combines multiple independent decision trees to generate a final prediction. These models were selected because they are commonly used for classification tasks in epidemiological and health-related studies.

The curated dataset was randomly partitioned into a training set (80%) and an independent holdout test set (20%). The training set was used exclusively for model fitting and internal cross-validation procedures, whereas the test set was reserved for final performance evaluation. In ML models, especially when working with imbalanced datasets, it is important to apply balancing techniques to prevent the majority class from dominating the results. Therefore, balancing techniques were applied to the algorithmic models to ensure that the model learns to identify relevant features from both classes, thereby improving precision and sensitivity, and making it possible to apply the results to other datasets.

ML model evaluation was conducted using metrics widely employed in classification tasks, with emphasis on the confusion matrix, which allows for a detailed analysis of performance in terms of true and false positives and negatives.

Metrics such as accuracy, precision, recall, and the F1-score are obtained from the confusion matrix, which allow the evaluation of the model predictive performance. Accuracy indicates the proportion of correct predictions, precision evaluates the quality of positive predictions, and sensitivity measures the ability to correctly identify true LBW cases. Finally, the F1-score balances precision and sensitivity, and is particularly useful when class imbalance is present.

Additionally, the complementary metric area under the receiver operating characteristic curve (AUC-ROC) was used to evaluate model performance. Values closer to 1 indicate excellent discriminatory ability, whereas those near 0.5 indicate a prediction ability equivalent to random chance.

All reported performance metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, were calculated exclusively using the holdout test set. Receiver operating characteristic (ROC) curves were generated using predicted probabilities obtained from the test data, ensuring unbiased assessment of the discriminative performance of the model. These methodological approaches were aimed to ensure a comprehensive evaluation of the performance that considers not only the overall accuracy but also the effectiveness in correctly identifying LBW.

Finally, an exploratory data analysis was conducted to investigate the distribution of the variables used in the study. This step was aimed to evaluate the behavior of maternal and neonatal characteristics present in the dataset, allowing for the identification of possible associations with LBW occurrence. To this end, graphical representations accompanied by proportional analyses between

comparison groups were used to address specific questions related to the phenomenon studied.

Research ethics committee approval

This study was approved by the Research Ethics Committee of Universidade de Taubaté on February 17, 2011, under protocol number 009/11.

RESULTS

Of the 2,548,570 valid records of live births included in the analysis, 2,307,505 (90.54%) corresponded to newborns with normal weight and 241,065 (9.46%) to cases of LBW (< 2,500 g).

The training set consisted of 2,038,856 records, whereas the remaining 509,714 records were used for testing/validation.

After model training, they were applied to the test/validation dataset. The following results were obtained for the random forest and logistic regression models: **Figure 1 (A and B)** presents the confusion matrices for each model, showing their ability to achieve the highest accuracy in predicting all cases. Training was then performed with balanced classes using the command `class_weight = 'balanced'`, in which the algorithm automatically adjusts class weights based on the inverse frequency of each class in the dataset. The results show a balance between the classes, enabling the prediction of specific LBW cases for each model used (**Figure 1, C and D**).

After training, the performance of the random forest and logistic regression models was evaluated using widely accepted metrics in the ML literature, including accuracy, precision, recall, and the F1-score, both before and after balancing classes (**Table 1**). Although both models showed higher accuracy without balancing, they did not demonstrate good recall performance under this condition.

All evaluated models showed AUC ROC values considered relevant to the objectives of the study, both with and without balancing. Among them, the logistic regression model achieved the

A. Confusion Matrix Random Forest				B. Confusion Matrix Logistic Regression			
True Label	Low Weight	20,024	27,987	True Label	Low Weight	19,224	28,787
	Normal Weight	13,298	448,405		Normal Weight	9,198	452,505
Predicted Label				Predicted Label			
True Label	Low Weight	26,598	21,413	True Label	Low Weight	33,066	14,945
	Normal Weight	34,323	427,380		Normal Weight	35,886	425,817
Predicted Label				Predicted Label			

Figure 1. Confusion matrices for the random forest and logistic regression models applied to the prediction of low birth weight (LBW) in the State of São Paulo from 2019 to 2023.

best performance, outperforming the random forest model in the evaluated metrics (Figure 2).

Table 2 presents the distribution of variables corresponding to maternal medical history, socioeconomic factors, and prenatal health indicators in the LBW and normal-weight groups.

Initially, the models were trained without dataset balancing, reflecting the distribution of real data, which is typically imbalanced between classes (low weight and normal weight). Under this condition, logistic regression achieved the highest accuracy (92.5%) and highest precision (67.8%), indicating good ability to correctly identify positive cases. However, the sensitivity (recall) was relatively low (39.9%), demonstrating that the model had difficulty identifying all true LBW cases, which is common when imbalanced is present.

The random forest model, in turn, showed a slightly lower accuracy (91.9%), lower precision (60.1%), and a sensitivity of 41.7% without balancing. This indicates that the overall performance was similar to that of the logistic regression model; however, it still had difficulties in the adequate identification of cases from the minority class (LBW).

A significant change was observed in the metrics after balancing, especially regarding sensitivity, which is crucial in public health contexts where detecting positive cases (LBW in this case) is a priority.

Table 1. Evaluation metrics for the random forest and logistic regression models before and after balancing

Model	Accuracy	Precision	Recall	F1-score
Random forest	0.919	0.601	0.417	0.492
Random forest (balanced)	0.89	0.436	0.554	0.488
Logistic regression	0.925	0.678	0.399	0.502
Logistic regression (balanced)	0.901	0.483	0.688	0.567

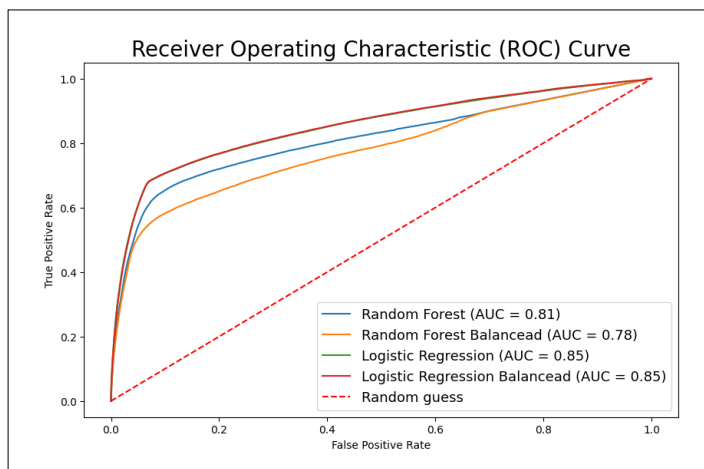


Figure 2. Areas under the receiver operating characteristic curve (AUC-ROC) for the random forest and logistic regression models under different balancing scenarios used to predict low birth weight occurrence in the State of São Paulo between 2019 and 2023.

The balanced logistic regression model showed a significant increase in recall (68.8%), indicating an enhanced ability to correctly identify LBW cases. Despite decreases in precision (48.3%) and accuracy (90.1%), the balanced model achieved the highest F1-score (56.7%), reflecting a more efficient balance between precision and recall.

The balanced random forest model also showed improvement in recall (55.4%) compared to the original model, although accuracy (89%) and precision (43.5%) were decreased. The F1-score was 48.8%, slightly lower than that of the balanced logistic regression model.

Table 2. Distribution of low birth weight and normal birth weight cases for each of the variables considered in the model

Variables	< 2,500 g (%)	≥ 2,500 g (%)
Maternal age		
≤ 19 years	10.457	9.181
20–34 years	65.489	70.710
≥ 35 years	24.054	20.109
Type of pregnancy		
Singleton	82.610	99.004
Twin or more	17.390	0.996
Parity		
Multiparous	59.479	62.463
Nulliparous	40.521	37.537
Previous child deaths		
None	76.242	80.029
One or more	23.758	19.971
Maternal marital status		
With partner	53.017	56.515
Without partner	46.983	43.485
Maternal race		
White	51.960	53.711
Black	8.297	7.259
Asian	0.536	0.537
Brown	39.111	38.380
Indigenous	0.097	0.113
Maternal education		
No education or elementary	16.600	13.297
High school	56.839	58.526
Higher education	26.561	28.176
Month of prenatal care start		
0–3 months	86.290	88.203
≥ 4 months	13.710	11.797
Prenatal consultations		
0–6	37.835	16.010
7–12	57.473	75.803
≥ 12	4.692	8.186
Newborn sex		
Female	53.154	48.429
Male	46.846	51.571
Identified anomaly		
No	96.571	98.898
Yes	3.429	1.102

Regarding AUC-ROC logistic regression stood out once again, both with and without balancing, achieving a value (85%) that reflects good discriminative ability. The AUC-ROC values for the random forest model without (81%) and with balancing (78%), in contrast, suggest a slight loss of discriminative capacity after balancing.

The precision–recall curves confirm these findings. The balanced logistic regression model achieved the best performance (60%), followed by the logistic regression model without balancing (57%). The random forest models showed worse performances (54% and 52%, respectively).

These results demonstrate that the balancing technique positively impacts sensitivity, and that the balanced logistic regression model proved to be the most appropriate in the analyzed context by achieving the best balance among the metrics. Moreover, the consistent performance in terms of AUC-ROC confirms the robustness of this model for LBW prediction.

DISCUSSION

The present study demonstrated the usefulness of machine learning approaches for identifying factors associated with low birth weight. Among the evaluated models, logistic regression showed the most favorable overall performance, particularly after class balancing, suggesting that simpler and more interpretable algorithms may perform as well as or better than more complex methods in population-based epidemiological datasets. The improvement observed after balancing reinforces the importance of addressing class imbalance when predicting relatively infrequent outcomes such as low birth weight.

Several maternal, socioeconomic, and prenatal care factors were associated with low birth weight and were consistent with findings reported in previous studies.

Extreme maternal ages (≤ 19 years and ≥ 35 years) were associated with a higher occurrence of LBW, corroborating the results reported by Vale et al.,⁵ Victor et al.,²⁰ Abdi et al.,⁸ and Rezaei et al.,⁹ who identified maternal age as one of the main predictors of LBW. Parity, represented by the first pregnancy, was also observed to be associated to a higher proportion of LBW, as described by Vale et al.⁵ Another relevant factor was a history of previous child deaths, indicating that women who had experienced the death of a previous child were more likely to have newborns with low birth weights, as pointed out by Ranjbar et al.⁷ Similarly, multiple pregnancies (twins or triplets) were associated with a higher proportion of LBW, as previously reported by Morais et al.²² and in agreement with prematurity-related outcomes described by Abdi et al.⁸ The analysis of socioeconomic factors demonstrated that pregnant women with low levels of education, especially those with no formal schooling or elementary education only, had a higher proportion of LBW, which aligns with the findings of Pollob et al.,¹⁹ Vale et al.,⁵ and Morais et al.²² Additionally, marital status revealed that pregnant women without partners are more likely to have newborns

with LBW, in agreement with the results reported by Vale et al.⁵ and Victor et al.²⁰ Regarding maternal self-declared race, pregnant women who identified as Black and Brown had a higher proportion of LBW, presumably due to social inequalities that directly impact maternal and child health, as previously highlighted by Vale et al.⁵

Regarding variables related to prenatal care, it was noted that LBW cases were significantly more frequent in pregnant women who had had up to six consultations. This result aligns with those reported by Nascimento et al.,¹⁰ Victor et al.,²⁰ and Suarez and Santana,¹¹ and is further supported by Abdi et al.,⁸ who emphasized the role of inadequate prenatal follow-up in adverse neonatal outcomes. Beginning prenatal care by the third month of pregnancy was associated with a higher proportion of newborns with normal weight. Another relevant finding was that female newborns had a higher prevalence of LBW, which has also been reported by Nascimento et al.,¹⁰ Vale et al.,⁵ and Suarez and Santana.¹¹ Additionally, the presence of anomalies identified during pregnancy was strongly associated with adverse outcomes, as highlighted in studies focused on identifying prenatal risk factors, such as that by Morais et al.²²

This study has some limitations that should be considered when interpreting the results. As it is based on a secondary data source (SINASC), the possibility of potential typing errors, incomplete or incorrect data entry, underreporting, or inaccurate classification of gestational age or birth weight cannot be ruled out, and this is also the case for errors in the measurement of clinical variables or the omission of relevant information by the professionals responsible for data recording.

Internal validity was strengthened through the use of a large, population-based dataset, standardized data collection procedures inherent to SINASC, and strict separation between the training and test datasets. The application of class balancing techniques and the use of multiple evaluation metrics further contributed to a robust assessment of model performance, particularly in the identification of LBW cases.

External validity is supported by the comprehensive coverage of SINASC, which includes virtually all live births that took place in the state of São Paulo during the study period. Given the demographic, socioeconomic, and healthcare diversity of this state, the findings may be generalizable to other Brazilian regions with similarly structured health systems. However, caution is warranted when extrapolating the results to settings with substantially different levels of healthcare access or data quality, or a different population profile. Future studies should assess model performance using external datasets from other states or countries to further validate its generalizability.

CONCLUSION

The results showed that the balanced logistic regression model achieved the best overall performance, standing out as a particularly appropriate approach for scenarios with class imbalance, as observed in the present study.

In parallel, the analysis of variable distribution confirmed evidence from the previous literature by identifying factors such as extreme maternal age, low educational attainment, absence of a partner, Black or Brown maternal race/ethnicity, multiple pregnancies, history of previous child deaths, female sex for the newborn, pregnancy anomalies, reduced number of prenatal visits, and late initiation of prenatal care as being associated with a higher proportion of LBW deliveries.

Notwithstanding the limitations of the study, the use of ML techniques proved effective in predicting LBW and analyzing relevant patterns in the data. The present study contributes to strengthening maternal and child health surveillance efforts by providing evidence-based insight that can be used to guide preventive strategies aimed at reducing LBW rates and promoting better health for mothers and newborns.

REFERENCES

- World Health Organization (WHO). WHO [Internet]. Geneva; 2025 [cited 2026 May 11]. ICD-11 for Mortality and Morbidity Statistics. Available from: <https://icd.who.int/browse/2025-01/mms/en>.
- Okwaraji YB, Krasevec J, Bradley E, et al. National, regional, and global estimates of low birthweight in 2020, with trends from 2000: a systematic analysis. *The Lancet*. 2024 Mar;403(10431):1071–80. [https://doi.org/10.1016/S0140-6736\(23\)01198-4](https://doi.org/10.1016/S0140-6736(23)01198-4).
- Brasil. Ministério da Saúde. Informações de Saúde (TABNET) – DATASUS [Internet]. 2023 [cited 2026 May 15]. Available from: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>.
- Gaíva MAM, Lopes FSP, Mufato LF, Ferreira SMB. Fatores associados à mortalidade neonatal em recém-nascidos de baixo peso ao nascer. *Revista Eletrônica Acervo Saúde*. 2020 Aug 13;12(11):e4831. <https://doi.org/10.25248/reas.e4831.2020>.
- Vale CCR, Almeida NKDO, Almeida RMVRD. Association between Prenatal Care Adequacy Indexes and Low Birth Weight Outcome. *Rev Bras Ginecol Obstet*. 2021 Apr;43(04):256–63. <https://doi.org/10.1055/s-0041-1728779>.
- Morais FL, Silva Rocha E, Masson G, et al. Utilization of tree-based machine learning models for predicting low birth weight cases. *BMC Pregnancy Childbirth*. 2025 Feb 26;25(1):207. <https://doi.org/10.1186/s12884-025-07303-x>.
- Ranjbar A, Montazeri F, Farashah MV, Mehrnough V, Darsareh F, Roozbeh N. Machine learning-based approach for predicting low birth weight. *BMC Pregnancy Childbirth*. 2023 Nov 20;23(1):803. <https://doi.org/10.1186/s12884-023-06128-w>.
- Abdi F, Roozbeh N, Darsareh F, Mehrnough V, Vahidi Farashah MS, Montazeri F. Developing a prognostic model for predicting preterm birth using a machine learning algorithm. *BMC Pregnancy Childbirth*. 2025 Sep 30;25(1):974. <https://doi.org/10.1186/s12884-025-08136-4>.
- Rezaei N, Amani M, Asgharpoor H, Mehrnough V, Darsareh F, Nikmanesh A. Machine Learning Approach to Predict Emergency Cesarean Sections Among Nulliparous Women. *Cureus*. 2025 Aug 19. <https://doi.org/10.7759/cureus.90501>.
- Nascimento RC, Barbosa MCR, Corrêa MM. Baixo-peso ao nascer: estudo de fatores associados em um hospital terciário da Grande Vitória. *DEMETER: Alimentação, Nutrição & Saúde*. 2019 Oct 28;14:e43508–e43508. <https://doi.org/10.12957/demeter.2019.43508>.
- Suarez TOF, Santana ABC. Fatores de risco para baixo peso ao nascer em recém-nascidos vivos no Amazonas 2016–2020. *Revista Eletrônica Acervo Saúde*. 2024 Jul 3;24(7):e16472. <https://doi.org/10.25248/reas.e16472.2024>.
- Ludermir TB. Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências. *Estudos Avançados*. 2021 Apr 30;35(101):85–94. <https://doi.org/10.1590/s0103-4014.2021.35101.007>.
- Paixão GMM, Santos BC, Araujo RM, et al. Machine Learning na Medicina: Revisão e Aplicabilidade. *Arq Bras Cardiol [Internet]*. 2022 Jan 10;118(1):95–102. Available from: <https://abccardiol.org/article/machine-learning-na-medicina-revisao-e-aplicabilidade/>.
- Askin L, Tanriverdi O, Cetin M. Algoritmos Artificiais Superam os Modelos Tradicionais na Predição de Doença Arterial Coronariana. *Arquivos Brasileiros de Cardiologia*. 2021 Nov 22;117(6):1071–2. <https://doi.org/10.36660/abc.20210823>.
- Carvalho ACPDLFD. Inteligência Artificial: riscos, benefícios e uso responsável. *Estud Av*. 2021 Apr;35(101):21–36. <https://doi.org/10.1590/s0103-4014.2021.35101.003>.
- Kingsford C, Salzberg SL. What are decision trees? *Nat Biotechnol*. 2008 Sep;26(9):1011–3. <https://doi.org/10.1038/nbt0908-1011>.
- Fernandes AAT, Filho DBF, Rocha EC, Nascimento WS. Read this paper if you want to learn logistic regression. *Revista de Sociologia e Política*. 2020 Dec 15;28(74):1–19. <https://doi.org/10.5380/rsocp.v28i74.78515>.
- Vellido A. The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural Comput & Applic*. 2020 Dec 1;32(24):18069–83. <https://doi.org/10.1007/s00521-019-04051-w>.
- Islam Pollob SMA, Abedin MM, Islam MT, Islam MM, Maniruzzaman M. Predicting risks of low birth weight in Bangladesh with machine learning. *PLoS One*. 2022 May 26;17(5):e0267190. PMID: 35617201; <https://doi.org/10.1371/journal.pone.0267190>.
- Victor A, Almeida F, Xavier SP, Rondó PHC. Predicting low birth weight risks in pregnant women in Brazil using machine learning algorithms: data from the Araraquara cohort study. *BMC Pregnancy Childbirth*. 2025 Mar 19;25(1):320. <https://doi.org/10.1186/s12884-025-07351-3>.
- Instituto Brasileiro de Geografia e Estatística (IBGE). IBGE [Internet]. Brasília (DF); 2025 [cited 2026 May 11]. Estimativas da população residente para os municípios e para as unidades da federação. Available from: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html>.
- Morais FL, Neri ABT, Silva Rocha E, et al. Predicting Low Birth Weight Using Machine Learning Models. In: 2023 18th Iberian Conference on Information Systems and Technologies (CISTI) [Internet]. Aveiro, Portugal: IEEE; 2023 [cited 2026 May 11]. p. 1–7. Available from: <https://ieeexplore.ieee.org/document/10211576>. <https://doi.org/10.23919/CISTI58278.2023.10211576>.

Authors' contributions: Almeida AVA, Santos AOR, Nascimento LFC: conceptualization (equal), supervision (equal), project administration (equal), resources (equal), validation (equal), methodology (equal), investigation (equal), data curation (equal), formal analysis (equal), writing – original draft (equal), writing – review and editing (equal). All authors reviewed and approved the final version of the manuscript for publication.

Events or congresses: Part of this study was published at the 16th São Paulo Congress of Pediatrics, where it was awarded first place in the Best Free Paper category.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: The datasets analyzed in the current study are publicly available from the Sistema de Informações sobre Nascidos Vivos database, maintained by the Brazilian Ministry of Health, through the DATASUS.

Declaration of generative AI in scientific writing: Generative AI tools, including ChatGPT (OpenAI), were used solely to assist with language editing, grammar refinement, and improvement of manuscript readability. All AI-assisted content was carefully reviewed, edited, and validated by the authors to ensure accuracy, originality, and scientific integrity. No generative AI tools were used for study design, data analysis, model development, interpretation of results, or scientific decision-making. The machine learning methods described in this study were independently developed and applied by the authors as part of the research methodology.

Date of first submission: August 27, 2025

Last received: January 22, 2026

Accepted: May 4, 2026

Address for correspondence:

Luiz Fernando Costa Nascimento
Departamento de Engenharia, Faculdade de Engenharia e Ciências,
Universidade Estadual Paulista "Júlio de Mesquita Filho" (Unesp)
Av. Ariberto Pereira da Cunha, 333
Portal das Colinas – Guaratinguetá (SP) — Brasil
CEP 12516-410
Tel. (+55 12) 3123-2161
E-mail: fernando.nascimento@unesp.br

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



The effects of agmatine treatment on cisplatin-induced hepatotoxicity: an experimental rat study

Murat Yeniceri^I, Mustafa Can Senoyamak^{II}, Süleyman Bas^{III}, Musa Salmanoglu^{IV}, Alpaslan Tanoglu^V

Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Türkiye

^IMD. Physician, Department of Rheumatology, University of Health Sciences, Kartal Dr. Lutfi Kirdar City Hospital, Istanbul, Türkiye.

 <https://orcid.org/0000-0003-3046-5784>

^{II}MD. Physician, Department of Endocrinology, University of Health Sciences, Sultan 2. Abdulhamid Han Training and Research Hospital, Istanbul, Türkiye.

 <https://orcid.org/0000-0002-1977-5127>

^{III}MD. Assistant Professor, Department of Internal Medicine, University of Health Sciences, Sancaktepe Sehit Prof. Dr. Ilhan Varank Training and Research Hospital, Istanbul, Türkiye.

 <https://orcid.org/0000-0002-5883-445X>

^{IV}MD. Assistant Professor, Department of Internal Medicine, University of Health Sciences, Sultan 2. Abdulhamid Han Training and Research Hospital, Istanbul, Türkiye.

 <https://orcid.org/0000-0002-5050-531X>

^VMD. Professor, Division of Gastroenterology, Department of Internal Medicine, Bahçeşehir University, Istanbul, Türkiye.

 <https://orcid.org/0000-0002-7477-6640>

KEY WORDS (MeSH terms):

Agmatine.
Cisplatin.
Chemical and drug-induced liver injury.
Oxidative stress.
Rats.

AUTHORS' KEY WORDS:

Agmatine treatment.
Cisplatin toxicity.
Oxidative liver damage.
Inflammation.
Experimental rat study.
Hepatoprotection.

ABSTRACT

BACKGROUND: Cisplatin is a widely used chemotherapeutic agent whose clinical utility is limited by dose-dependent hepatotoxicity, primarily driven by oxidative stress and inflammation. Agmatine, an endogenous polyamine, has demonstrated antioxidant and anti-inflammatory properties in various models; however, its hepatoprotective effects in cisplatin-induced liver injury remain underexplored.

METHODS: This experimental study investigated the protective effects of agmatine against cisplatin-induced hepatotoxicity in 28 male Sprague Dawley rats. The animals were randomized into four groups: Sham, Cisplatin, Cisplatin + Agmatine (20 g/kg/day, orally), and Placebo. Biochemical markers, including ALT, AST, MDA (malondialdehyde), SOD (superoxide dismutase), GPx (glutathione peroxidase), TNF- α , IL-1 β , and TGF- β , were assessed using ELISA (enzyme-linked immunosorbent assay). Histopathological evaluation was performed to assess liver architecture and injury severity.

RESULTS: Cisplatin administration significantly increased serum levels of ALT, AST, MDA, and proinflammatory cytokines (IL-1 β , TNF- α , and TGF- β), while reducing antioxidant enzyme activities (SOD and GPx) ($p < 0.01$). Agmatine co-administration significantly reversed these biochemical alterations and ameliorated histopathological damage, as evidenced by reduced inflammatory infiltration, sinusoidal dilatation, and hepatocellular degeneration.

CONCLUSION: Agmatine demonstrates significant hepatoprotective effects against cisplatin-induced liver injury by attenuating oxidative stress and inflammatory responses. These findings suggest its potential role as an adjunctive agent to improve the hepatic safety profile of cisplatin chemotherapy.

INTRODUCTION

Cisplatin (cis-diamminedichloroplatinum) is a platinum-based chemotherapeutic agent that is extensively used in the treatment of various malignancies because of its potent antineoplastic activity.¹ However, its clinical application is often hampered by severe adverse effects, among which hepatotoxicity holds significant clinical relevance. This toxicity is characterized by elevated serum liver enzyme levels and histopathological changes that compromise hepatic function and may adversely affect therapeutic outcomes.²

The hepatotoxic effects of cisplatin are primarily mediated by oxidative stress, resulting in excessive production of reactive oxygen species (ROS), lipid peroxidation, and protein peroxidation, as well as inhibition of endogenous antioxidant defense mechanisms.³ These events lead to hepatocellular injury, inflammation, and apoptosis. Therefore, there is an ongoing need for adjunctive agents that can mitigate these adverse hepatic effects without diminishing the chemotherapeutic efficacy of cisplatin.

Agmatine, an endogenous polyamine derived from the decarboxylation of L-arginine by arginine decarboxylase, exhibits multiple pharmacological properties, including modulation of nitric oxide synthesis, interaction with imidazoline and α 2-adrenergic receptors, and antagonism of NMDA receptors.⁴⁻⁶ Its antioxidant, anti-inflammatory, and anti-apoptotic effects have been demonstrated in various organ systems; however, its hepatoprotective potential, particularly in the context of cisplatin-induced toxicity, remains inadequately explored.^{7,8}

This study was designed to evaluate the protective effects of agmatine against cisplatin-induced hepatotoxicity in an experimental rat model and to elucidate its potential therapeutic role by employing comprehensive biochemical assays and histopathological analyses.

MATERIALS AND METHODS

Study design

This study was conducted at the Experimental Animal Laboratory, University of Health Sciences, Istanbul, Türkiye, following approval from the Institutional Animal Ethics Committee (Protocol No 2020-05-06) and in accordance with ARRIVE 3.0 guidelines.

A total of 28 male Sprague-Dawley rats, with body weights ranging from 250 to 350 g, were included in the study. The animals were housed under controlled laboratory conditions (12-h light/dark cycle, ambient temperature of approximately 24°C) and had ad libitum access to food and water. Food intake was restricted for 12-h prior to drug administration to standardize the baseline metabolic state.

The experimental timeline included agmatine administration, cisplatin injection, and euthanasia, which were performed in accordance with the predefined study protocol.

The rats were randomly divided into four experimental groups ($n = 7$ per group) as follows:

Group 1 – Sham: no intervention.

Group 2 – Cisplatin: single intraperitoneal (i.p.) injection of cisplatin (12 mg/kg) on Day 5; animals were euthanized on Day 8.

Group 3 – Cisplatin + Agmatine: oral agmatine (20 mg/kg/day) for 7 days; cisplatin administered on Day 5; euthanasia on Day 8.

Group 4 – Placebo: oral saline (1 mL, 0.9% NaCl) for 7 days; cisplatin (12 mg/kg) administered on Day 5; euthanasia on Day 8.

Biochemical and histopathological evaluation

On Day 8 of the study, all animals were anesthetized through an intraperitoneal injection of xylazine (10 mg/kg) and ketamine (75 mg/kg). Following midline laparotomy, blood samples were taken from the inferior vena cava. After excision of the liver tissue, the animals were euthanized humanely.

Serum biomarkers were evaluated to reflect systemic oxidative stress and inflammatory responses, which are widely used indicators in experimental hepatotoxicity studies.

Blood samples were immediately centrifuged at 3000 rpm for 10 min, and the resulting serum was stored at -70°C until subsequent analysis. The following biochemical parameters were quantified using rat-specific solid-phase enzyme-linked immunosorbent assay (ELISA) kits (BT Lab, China) in accordance with the manufacturer's protocols: alanine aminotransferase (ALT), aspartate aminotransferase (AST), tumor necrosis factor- α (TNF- α), transforming growth factor- β (TGF- β), malondialdehyde (MDA), superoxide dismutase (SOD), glutathione peroxidase (GPx), and interleukin-1 β (IL-1 β).

For histopathological evaluation, liver specimens were fixed in 10% neutral-buffered formalin, embedded in paraffin, and cut into sections. The tissue sections were stained with hematoxylin and eosin (H&E) and examined under a light microscope (Nikon

Eclipse E600). Liver injury was assessed semi-quantitatively based on four histological parameters: inflammatory cell infiltration, sinusoidal dilatation, hepatocyte damage, and vascular congestion. Each parameter was scored according to the following scale: The severity of liver tissue injury was graded as follows: score – (no histopathological damage), score + (mild damage affecting < 25% of the liver tissue), score ++ (moderate damage affecting 25–50% of the tissue), and score +++ (severe damage affecting > 50% of the tissue). All histopathological evaluations were conducted by an experienced pathologist blinded to the experimental groups to minimize observer bias.

Statistical Analyses

All statistical analyses were performed using SPSS version 25.0 (IBM Corp., Armonk, New York). Descriptive statistics, including means and standard deviations (SDs), were calculated for all variables. Normality of the data distribution was assessed using the Shapiro–Wilk test and visual inspection of histograms and Q–Q plots. For data exhibiting a normal distribution, comparisons between two groups were conducted using an independent samples t-test, whereas one-way analysis of variance (ANOVA) was used for comparisons involving more than two groups. Post hoc multiple comparisons were performed using the Bonferroni test. Homogeneity of variances was assessed prior to parametric analyses. Statistical significance was set at $p < 0.05$, and all results were reported with a 95% confidence interval (95% CI).

RESULTS

The serum levels of ALT, AST, GPx, IL-1 β , SOD, TGF- β , MDA, and TNF- α measured in all experimental groups are summarized in **Table 1**.

Cisplatin administration (Group 2) resulted in a significant elevation in serum ALT, AST, IL-1 β , TGF- β , TNF- α , and MDA levels compared to the Sham group (Group 1; $p < 0.01$). Conversely, cisplatin exposure caused a notable reduction in the levels of the antioxidant enzymes SOD and GPx. No statistically significant differences were observed between the Cisplatin-only group and the Placebo group (Group 4), indicating that oral saline administration did not modify the effects of cisplatin (**Figures 1 and 2**).

In contrast, rats in the Cisplatin + Agmatine group (Group 3) exhibited significantly lower levels of ALT, AST, IL-1 β , TGF- β , TNF- α , and MDA than those in the Cisplatin-only and Placebo groups ($p < 0.01$).

Additionally, GPx and SOD levels were significantly elevated in the agmatine-treated group, demonstrating the restoration of antioxidant defense mechanisms.

Histopathological examination corroborated the biochemical findings. In the sham group, liver tissues retained normal architecture with intact hepatocyte morphology. In contrast, the cisplatin group displayed severe hepatic damage, including hydropic

degeneration of hepatocytes, sinusoidal dilatation, hyperemia, focal necrosis, and inflammatory cell infiltration. Similar pathological alterations were observed in the placebo group (Table 2).

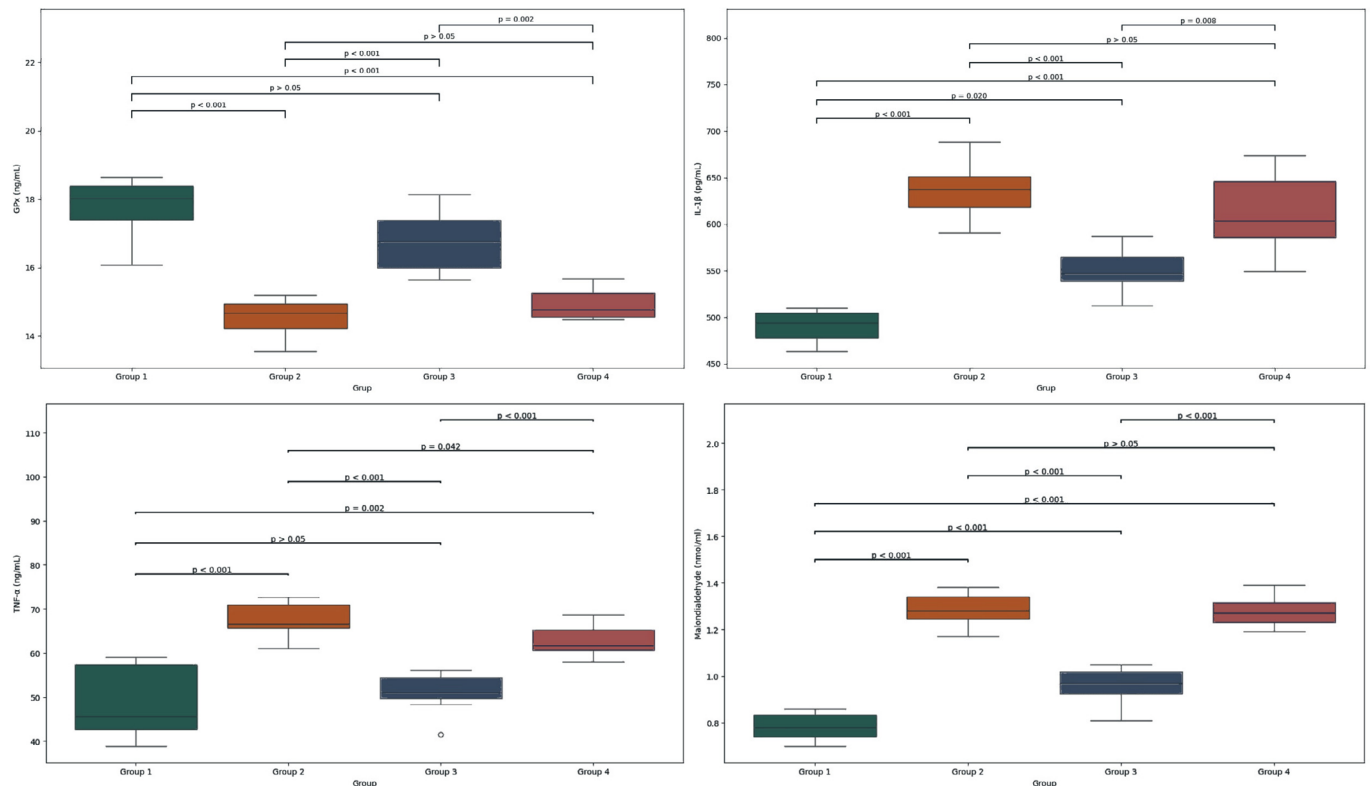
However, in the Cisplatin + Agmatine group, significant histological improvements were observed. Notably, the extent of

hydropic degeneration, sinusoidal dilatation, and inflammatory infiltration was markedly reduced compared to that in the Cisplatin-only and Placebo groups (Figure 3). These results further support the hepatoprotective effect of agmatine in alleviating cisplatin-induced liver injury.

Table 1. Biochemical parameters of study groups

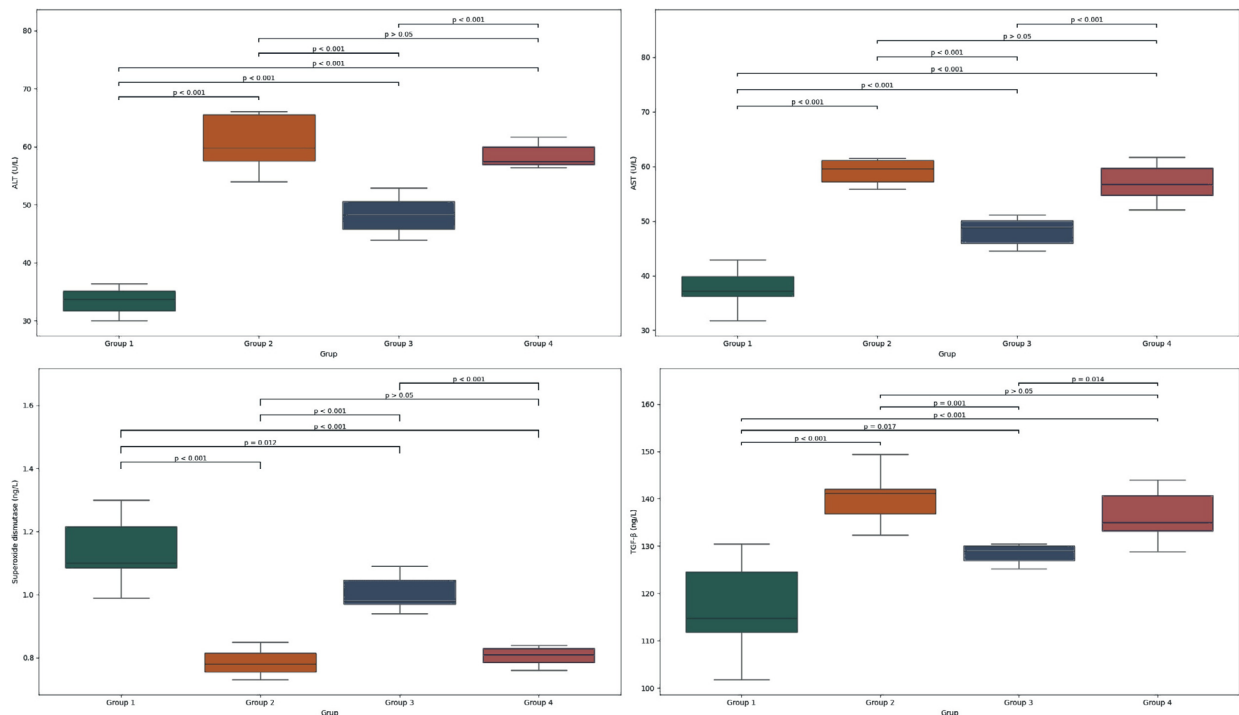
Parameter	Group 1 (SaM)	Group 2 (Hepatotoxic)	Group 3 (Hep + Treatment)	Group 4 (Placebo)	p value
ALT (U/L)	33.4 ± 2.3*	60.87 ± 4.87*	48.25 ± 3.33*	58.14 ± 4.72*	< 0.001 ^a
AST (U/L)	37.69 ± 3.57*	59.94 ± 3.93*	48.04 ± 2.58*	57.02 ± 3.53*	< 0.001 ^a
GPx (ng/mL)	17.93 ± 1.2*	14.53 ± 0.62*	16.76 ± 0.99*	14.66 ± 1.02*	< 0.001 ^a
IL-1β (pg/mL)	501.44 ± 41.24*	636.69 ± 32.27*	550.74 ± 25.57*	612.88 ± 45.31*	< 0.001 ^a
Superoxide dismutase (ng/L)	1.14 ± 0.1*	0.78 ± 0.04*	1 ± 0.06*	0.82 ± 0.06*	< 0.001 ^a
TGF-β (ng/L)	117.07 ± 10.03*	140.08 ± 5.51*	128.7 ± 4.71*	136.51 ± 5.41*	< 0.001 ^a
TNF-α (ng/mL)	49.05 ± 8.52*	67.65 ± 4.09*	50.97 ± 4.96*	62.88 ± 3.75*	< 0.001 ^a
Malondialdehyde (nmol/mL)	0.78 ± 0.06*	1.28 ± 0.07*	0.96 ± 0.08*	1.27 ± 0.07*	< 0.001 ^a

* $p < 0.01$; ^a Significantly different compared with the control group; ^b Significantly different compared with the CP group; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GPx, glutathione peroxidase; IL-1β, interleukin-1 beta; SOD, superoxide dismutase; TGF-β, transforming growth factor beta; TNF-α, tumor necrosis factor alpha; MDA, malondialdehyde.



The Kruskal–Wallis test (nonparametric ANOVA) was used for all statistical comparisons. IL-1β and TNF-α levels were markedly increased in Group 2 ($p < 0.001$), whereas levels in Group 4 were also higher than those in Groups 1 and 3. MDA levels were highest in Groups 2 and 4 and lowest in Group 1 ($p < 0.001$). ANOVA, analysis of variance; Group 1, sham group; Group 2, cisplatin group; Group 3, cisplatin + agmatine group; Group 4, control group; Gpx, glutathione peroxidase; IL-1β, interleukin-1 beta; TNF-α, tumor necrosis factor alpha; MDA, malondialdehyde.

Figure 1. Comparison of (A) Gp-x, (B) IL-1β, (C) TNF-α, and (D) MDA levels among Group 1, Group 2, Group 3 and Group 4.



The Kruskal–Wallis test (nonparametric ANOVA) was used for all statistical comparisons. ALT and AST levels were significantly higher in Group 2 than in all other groups ($p < 0.001$). SOD activity was reduced in Groups 2 and 4 compared to Group 1 ($p < 0.001$). TGF- β levels were elevated in Group 2 ($p < 0.001$). ANOVA, analysis of variance; Group 1, sham group; Group 2, cisplatin group; Group 3, cisplatin + agmatine group; Group 4, control group; ALT, alanine aminotransferase; AST, aspartate aminotransferase; SOD, superoxide dismutase; TGF- β , transforming growth factor beta.

Figure 2. Comparison of (A) ALT, (B) AST, (C) SOD and (D) TGF- β levels among Group 1, Group 2, Group 3 and Group 4.

Table 2. Grading of the histopathological changes in the liver sections of rats. Scoring was performed as follows: none (–), mild (+), moderate (++) and severe (+++)

Groups	Cell Infiltration	Dilation of Sinusoids	Hepatocyte Injury	Vascular Congestion
Group 1. Sham	–	–	–	–
Group 2. Cisplatin	+++	+++	+++	+++
Group 3. Cisplatin + Agmatine	+	+	+	++
Group 4. Placebo	+++	+++	+++	+++

DISCUSSION

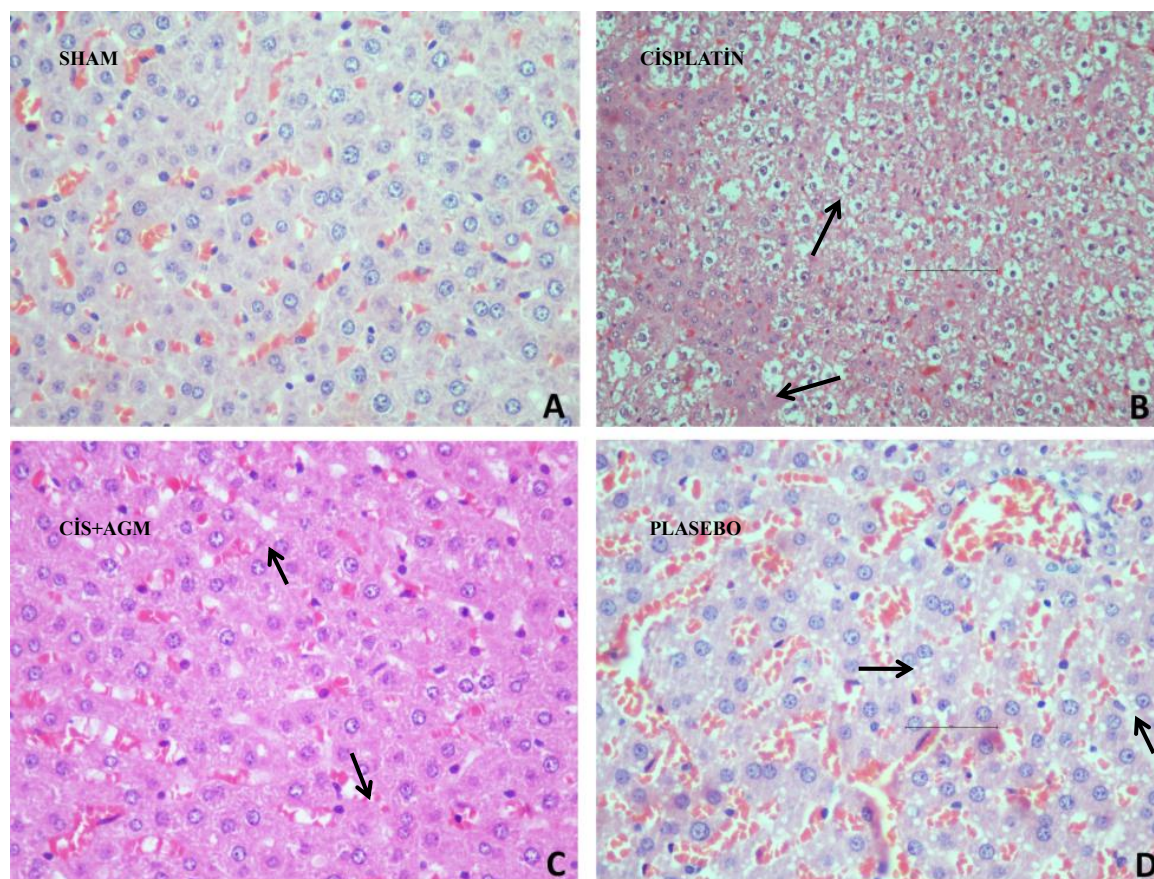
To the best of our knowledge, this is the first experimental study to assess the protective potential of agmatine against cisplatin-induced hepatotoxicity. These findings suggest that agmatine may serve as a promising adjunctive therapy to enhance the safety profile of cisplatin-based chemotherapy regimens.

Cisplatin-induced hepatotoxicity is primarily attributed to oxidative stress and inflammation.^{9,10} Consistent with previous studies, our findings show that cisplatin exposure led to elevated levels of MDA and proinflammatory cytokines (IL-1 β , TNF- α , and TGF- β), along with a reduction in the activities of antioxidant enzymes, such as SOD and GPx.^{11–13} These alterations indicate enhanced lipid peroxidation and a compromised antioxidant defense system, ultimately contributing to hepatocellular apoptosis and structural deterioration. Histopathological analysis corroborated these biochemical findings, revealing substantial hepatocyte damage, sinusoidal

dilatation, and inflammatory infiltration in cisplatin-treated rats. The biochemical findings were consistent with histopathological observations, supporting the protective effect of agmatine against hepatocellular injury.

The role of oxidative stress in cisplatin-induced hepatotoxicity has been well established. Previous studies have demonstrated that cisplatin administration increases ROS production, leading to lipid peroxidation and oxidative damage to cellular proteins and nucleic acids.^{13–15} This cascade of events activates inflammatory pathways, exacerbating tissue injury and impairing liver function. Our study is consistent with this mechanistic understanding, as evidenced by the elevated MDA levels and decreased SOD and GPx activities observed in the cisplatin-treated group. Furthermore, the increased levels of IL-1 β , TNF- α , and TGF- β in the cisplatin group support the inflammatory response associated with liver injury.

Given the pathophysiology of cisplatin-induced liver injury, the use of agents with both antioxidant and anti-inflammatory



Light microscopy images of the liver tissues stained with hematoxylin and eosin at $\times 100$ magnification. (A) Sham group (Group 1) with normal liver tissue. (B) Cisplatin group (Group 2) with hepatocyte injury and cell infiltration. (C) Cisplatin + Agmatine group (Group 3) with reduced cell infiltration and mild injury. (D) Placebo group (Group 4) with sinusoidal dilatation.

Figure 3. Histopathological evaluation of study groups.

properties has been proposed as a viable therapeutic strategy. Agmatine was selected in this study because of its well-documented ability to modulate oxidative stress and inflammation. In our experiment, agmatine significantly attenuated the expression of proinflammatory cytokines, decreased serum ALT and AST levels, and improved antioxidant enzyme profiles. These biochemical improvements were accompanied by a marked reduction in histopathological damage, underscoring agmatine's dual protective mechanism. The reduction in ALT and AST levels, which are indicative of hepatocellular injury, further supports the notion that agmatine can mitigate liver damage. Additionally, the restoration of antioxidant enzyme activity, particularly SOD and GPx, suggests that agmatine's hepatoprotective effects are, at least in part, mediated through the restoration of the hepatic antioxidant defense system.

Our findings are consistent with previous studies that have investigated the protective role of agmatine in hepatic and extra-hepatic tissues. For instance, Ahmed et al. observed similar protective effects of agmatine in a model of valproic acid-induced liver injury, where agmatine reduced serum aminotransferase levels,

attenuated proinflammatory cytokine levels, and improved histological parameters.¹⁶ Although their study focused on tissue SOD levels rather than serum levels, as in our study, the directionality of the response was consistent, reinforcing the broad therapeutic potential of agmatine in preventing various forms of hepatic injury. Moreover, Alharbi et al. demonstrated that agmatine significantly mitigated oxidative stress and inflammation in rat testicular tissue exposed to tacrolimus, a potent immunosuppressant.¹⁷ Similarly, Yeniceri et al.¹⁸ showed that agmatine treatment reduced inflammation and oxidative damage in pancreatic tissue, further highlighting its systemic anti-inflammatory and antioxidant properties. These findings collectively support agmatine's broad therapeutic potential as an agent capable of attenuating drug-induced toxicities across multiple organ systems.

In addition to its hepatoprotective effects, agmatine has been explored in models of cisplatin-induced toxicity in other organ systems.^{19–22} Donertas et al.¹⁹ demonstrated that agmatine alleviated degeneration of the dorsal root ganglion and sciatic nerve in a rat model of cisplatin-induced peripheral neuropathy. This finding further supports the neuroprotective potential of agmatine.

Similarly, studies on cisplatin-induced nephrotoxicity have shown that agmatine improves renal function parameters and histopathological outcomes, highlighting its systemic protective capacity. These studies emphasize the multifaceted nature of agmatine's protective effects and its potential as a therapeutic agent for mitigating cisplatin-induced toxicities in various organ systems.¹⁰

This study has several limitations. First, the sample size was relatively limited, which may restrict the generalizability of the findings. Second, although biochemical and histopathological evaluations were comprehensively performed, detailed molecular analyses investigating the underlying mechanisms were not conducted. Future studies with larger sample sizes and molecular investigations are needed to further clarify the protective mechanisms of agmatine in cisplatin-induced hepatotoxicity.

In conclusion, agmatine effectively mitigates cisplatin-induced hepatotoxicity by attenuating oxidative stress and inflammatory cytokine responses while preserving hepatic morphology and function. These findings propose agmatine as a promising adjunctive agent in oncology to safeguard liver function during cisplatin chemotherapy. Further studies are warranted to explore its translational applicability and systemic protective profile in other cisplatin-affected organ systems.

REFERENCES

- Dasari S, Tchounwou PB. Cisplatin in cancer therapy: molecular mechanisms of action. *Eur J Pharmacol*. 2014 Oct 5;740:364–78. PMID: 25058905; <https://doi.org/10.1016/j.ejphar.2014.07.025>.
- Chovanec M, Abu Zaid M, Hanna N, El-Kouri N, Einhorn LH, Albany C. Long-term toxicity of cisplatin in germ-cell tumor survivors. *Ann Oncol*. 2017 Nov 1;28(11):2670–9. PMID: 29045502; <https://doi.org/10.1093/annonc/mdx360>.
- Qi L, Luo Q, Zhang Y, Jia F, Zhao Y, Wang F. Advances in toxicological research of the anticancer drug cisplatin. *Chem Res Toxicol*. 2019 Aug 19;32(8):1469–86. PMID: 31353895; <https://doi.org/10.1021/acs.chemrestox.9b00204>.
- Christova TY, Gorneva GA, Taxirov SI, Duridanova DB, Setchenska MS. Effect of cisplatin and cobalt chloride on antioxidant enzymes in the livers of Lewis lung carcinoma-bearing mice: protective role of heme oxygenase. *Toxicol Lett*. 2003 Mar 3;138(3):235–42. PMID: 12565200; [https://doi.org/10.1016/s0378-4274\(02\)00416-2](https://doi.org/10.1016/s0378-4274(02)00416-2).
- Waseem M, Bhardwaj M, Tabassum H, Raisuddin S, Parvez S. Cisplatin hepatotoxicity mediated by mitochondrial stress. *Drug Chem Toxicol*. 2015 Oct;38(4):452–9. PMID: 25678195; <https://doi.org/10.3109/01480545.2014.992437>.
- Xu W, Gao L, Li T, Shao A, Zhang J. Neuroprotective role of agmatine in neurological diseases. *Curr Neuropharmacol*. 2018;16(9):1296–305. PMID: 28786346; <https://doi.org/10.2174/1570159X15666170808120633>.
- Kotagale NR, Taksande BG, Inamdar NN. Neuroprotective offerings by agmatine. *Neurotoxicology*. 2019 Jul;73:228–45. PMID: 31063707; <https://doi.org/10.1016/j.neuro.2019.05.001>.
- Galea E, Regunathan S, Eliopoulos V, Feinstein DL, Reis DJ. Inhibition of mammalian nitric oxide synthases by agmatine, an endogenous polyamine formed by decarboxylation of arginine. *Biochem J*. 1996 May 15;316(Pt 1):247–9. PMID: 8645212; <https://doi.org/10.1042/bj3160247>.
- Donertas B, Unel CC, Erol K. Cannabinoids and agmatine as potential therapeutic alternatives for cisplatin-induced peripheral neuropathy. *J Exp Pharmacol*. 2018 Jun 22;10:19–28. PMID: 29950907; <https://doi.org/10.2147/JEP.S162059>.
- Salihoglu YS, Elri T, Gulle K, et al. Evaluation of the protective effect of agmatine against cisplatin nephrotoxicity with 99mTc-DMSA renal scintigraphy and cystatin-C. *Ren Fail*. 2016 Oct;38(9):1496–502. PMID: 27604130; <https://doi.org/10.1080/0886022X.2016.1227919>.
- Xiang Y, Zhou X, Zhou H, et al. Limonin ameliorates cisplatin-induced acute liver injury by inhibiting 11 β -hydroxysteroid dehydrogenase type 1. *Biomed Pharmacother*. 2023 Dec;168:115680. PMID: 37832403; <https://doi.org/10.1016/j.biopha.2023.115680>.
- Bekhit AA, Beshay ON, Fawzy MA, et al. Curative effect of AD-MSCs against cisplatin-induced hepatotoxicity in rats is potentiated by azilsartan: targeting oxidative stress, MAPK, and apoptosis signaling pathways. *Stem Cells Int*. 2023 Oct 23;2023:6767735. PMID: 37908315; <https://doi.org/10.1155/2023/6767735>.
- Omar HA, Mohamed WR, Arafa SA, et al. Hesperidin alleviates cisplatin-induced hepatotoxicity in rats without inhibiting its antitumor activity. *Pharmacol Rep*. 2016 Apr;68(2):349–56. PMID: 26922538; <https://doi.org/10.1016/j.pharep.2015.09.007>.
- Waseem M, Pandey P, Tomar B, Raisuddin S, Parvez S. Ameliorative action of curcumin in cisplatin-mediated hepatotoxicity: an in vivo study in Wistar rats. *Arch Med Res*. 2014 Aug;45(6):462–8. PMID: 25080283; <https://doi.org/10.1016/j.arcmed.2014.07.006>.
- Niazvand F, Ashtari A, Chamkouri N, Azari M. Hepatoprotective effects of alpinia officinarum rhizome extract on cisplatin-induced hepatotoxicity in rats: a biochemical, histopathological and immunohistochemical study. *J Trace Elem Med Biol*. 2023 Dec;80:127306. PMID: 37757646; <https://doi.org/10.1016/j.jtemb.2023.127306>.
- Ahmed N, Aljuhani N, Al-Hujaili HS, et al. Agmatine protects against sodium valproate-induced hepatic injury in mice via modulation of nuclear factor- κ B/inducible nitric oxide synthetase pathway. *J Biochem Mol Toxicol*. 2018 Dec;32(12):e22227. PMID: 30273971; <https://doi.org/10.1002/jbt.22227>.
- Alharbi N, Nour O, Makled MN, Nader M. Agmatine abrogates tacrolimus-induced testicular injury in rats. *Pharmaceutics*. 2025 May 27;17(6):703. PMID: 40574016; <https://doi.org/10.3390/pharmaceutics17060703>.
- Yeniceri M, Tanoglu A, Salmanoglu M, et al. Efficacy of agmatine treatment in experimental acute pancreatitis rat model. *J Gastroenterol*. 2024 Jan;35(1):27–31. PMID: 38454275; <https://doi.org/10.5152/tjg.2024.23017>.

19. Donertas B, Cengelli Unel C, Aydin S, et al. Agmatine co-treatment attenuates allodynia and structural abnormalities in cisplatin-induced neuropathy in rats. *Fundam Clin Pharmacol*. 2018 Jun;32(3):288–96. <https://doi.org/10.1111/fcp.12351>. Erratum in: *Fundam Clin Pharmacol*. 2018 Dec;32(6):678. PMID: 29377255; <https://doi.org/10.1111/fcp.12429>.
20. Izadi S, Rezaei A, Esmaili Z, Soukhaklari R, Moosavi M. Agmatine prevents the memory impairment and the dysfunction of hippocampal GSK-3 β and ERK signaling induced by aluminum nanoparticle in mice. *Behav Pharmacol*. 2023 Aug 1;34(5):299–305. PMID: 37401404; <https://doi.org/10.1097/FBP.0000000000000735>.
21. Ommati MM, Niknahad H, Najibi A, et al. Cholestasis-associated pulmonary inflammation, oxidative stress, and tissue fibrosis: the protective role of the biogenic amine agmatine. *Pharmacology*. 2023;108(4):379–93. PMID: 37253339; <https://doi.org/10.1159/000530307>.
22. Raghavan SA, Dikshit M. Vascular regulation by the L-arginine metabolites, nitric oxide and agmatine. *Pharmacol Res*. 2004 May;49(5):397–414. PMID: 14998549; <https://doi.org/10.1016/j.phrs.2003.10.008>.

Authors' contributions: Yeniceri M: conceptualization, writing – original draft, writing – review and editing; Senoyamak MC: investigation, methodology; Bas S: data curation, methodology, formal analysis, writing – review and editing; Salmanoglu M: project administration, resources, validation, supervision, writing – review and editing; Tanoglu A: project administration, resources, validation, supervision, writing – review and editing. All authors reviewed and approved the final version of the manuscript for publication.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: Data supporting the findings of this study are available from the corresponding author, Murat Yeniceri, upon request.

Declaration of generative AI in scientific writing: During the preparation of this study, the authors did not use generative AI or AI-assisted technologies.

Date of first submission: September 5, 2025

Last received: February 13, 2026

Accepted: March 2, 2026

Address for correspondence:

Murat Yeniceri

Department of Rheumatology, University of Health Sciences, Kartal

Dr. Lütfi Kırdar City Hospital

Kartal — İstanbul — Türkiye

P.K. 34785

Tel. (+90 501) 207 98 01

E-mail: 4256.gatf@gmail.com

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)

Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Assessment of balance, sleep quality, and depression in relation to attention in Parkinson's disease: A cross-sectional P300 study

Gulfem Ezgi Ozaltin^I, Agit Simsek^{II}, Busra Candirli^{III}, Fatma Ebru Algul^{IV}, Mehmet Tecellioglu^V, Burcu Talu^{VI}

Department of Physiotherapy and Rehabilitation, Inonu University (U), Malatya (UF), Türkiye

^IPhD. Assistant Professor, Department of Physiotherapy and Rehabilitation, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0003-1591-4844>

^{II}PhD. Assistant Professor, Department of Speech and Language Therapy, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0003-1572-1188>

^{III}PhD. Assistant Professor, Department of Physiotherapy and Rehabilitation, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0001-7413-6371>

^{IV}MD. Associate Professor, Department of Neurology, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0003-0318-7571>

^VMD. Professor, Department of Neurology, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0002-8381-9630>

^{VI}PhD. Professor, Department of Physiotherapy and Rehabilitation, Inonu University (U), Malatya (UF), Türkiye.

<https://orcid.org/0000-0002-5623-8291>

KEYWORDS (MeSH terms):

Attention.
Cognition.
Parkinson disease.
Postural balance.
Neurologic Rehabilitation.

AUTHOR'S KEYWORDS:

Physiotherapy And Rehabilitation.
Balance Impairment.
Event-related potentials.
Sleep disturbances.
Cognitive processing.

ABSTRACT

BACKGROUND: Parkinson's disease (PD) is a progressive neurodegenerative disease characterized by motor and non-motor symptoms. Attention deficits are common in patients with PD, and balance can be affected by sleep quality and depression. The auditory event-related potential (ERP) P300 tool is associated with attention and discrimination processes and is a sensitive indicator of attention.

OBJECTIVES: This study aimed to investigate the effects of balance, sleep quality, and depression on attention in individuals with PD using P300 latency.

DESIGN AND SETTING: In this cross-sectional study, 28 individuals with PD and 27 healthy controls were assessed.

METHODS: Attention was measured using the auditory ERP P300, balance with the Berg Balance Scale (BBS), sleep quality with the Pittsburgh Sleep Quality Index (PSQI), and depression with the Beck Depression Inventory (BDI). Multiple linear regression analysis was used to examine the effects of these variables on P300 latency.

RESULTS: Significant differences were found between the groups in terms of attention, balance, sleep quality, and depression (all $P < 0.001$). Regression analysis indicated that balance ($B = -4.909$, $P = 0.001$) and sleep quality ($B = 12.704$, $P = 0.007$) were significantly associated with P300 latency, whereas depression ($B = 3.472$, $P = 0.084$) was not.

CONCLUSION: PD is associated with impaired attention, balance deficits, poor sleep quality, and increased depressive symptoms. The significant association between balance performance and sleep quality with attentional processing highlights their potential relevance in clinical assessment and management.

CLINICAL TRIAL REGISTRATION: The study was registered with Clinical Trials (NCT04687371 – <https://clinicaltrials.gov/study/NCT04687371>).

INTRODUCTION

Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized primarily by dopaminergic degeneration in the basal ganglia. In addition to motor symptoms, such as tremors, rigidity, gait disturbances, and postural instability, PD is also associated with widespread neurodegeneration affecting the serotonergic, noradrenergic, and cholinergic systems. These changes contribute to a wide range of non-motor symptoms, including cognitive impairment, attention deficits, depression, and sleep disturbances.¹⁻³ These non-motor features significantly contribute to disease burden and may interact with motor dysfunction. Cognitive impairment due to subcortical lesion formation is one of the most common symptoms of PD.³

The P300 component of auditory event-related potentials (ERP) is widely used as an objective neurophysiological marker of attentional and cognitive processing. Unlike conventional neuropsychological tests, the P300 provides objective data that are less influenced by motor impairment and patient compliance.⁴ Previous studies investigating P300 latency in patients with PD have reported inconsistent findings. While most studies have demonstrated prolonged P300 latencies, suggesting cognitive slowing, a limited number have reported no significant differences compared to healthy individuals.^{5,6} In addition, the relationships between motor and non-motor symptoms and attentional processing in patients with PD have generally been examined separately, and their combined associations remain unclear.⁷⁻⁹ Although impairments in attention are well recognized in patients with PD, the clinical factors associated with these changes

have not been fully elucidated. In light of this information, we hypothesized that there is a difference in attention levels between the PD and healthy groups, and that dynamic balance, sleep, and depression may be related to attention (P300) in patients with PD.

OBJECTIVE

This study aimed to investigate differences in P300 latency between individuals with PD and healthy controls and to examine the associations of balance, sleep quality, and depressive symptoms with attentional processing in patients with PD.

METHODS

Study Design

This study was planned as a cross-sectional study. Following institutional and ethics committee approval (Approval No:2022/3749, Date: 25.10.2022), it was conducted in the neurology outpatient clinic of a tertiary research hospital. The study was registered with Clinical Trials. It was conducted in accordance with the principles of the Declaration of Helsinki, and informed consent was obtained from the participants.

Sample size and participants

In the power analysis conducted before the study, it was calculated that at least 27 patients should be included in each group for the change in balance score to be 1 in patients with PD with $\alpha = 0.05$ and $1-\beta$ (power) = 0.80; this calculation was based on balance outcomes from a previous study.¹⁰

The study population consisted of patients who applied to the neurology outpatient clinic and were diagnosed with PD by a specialist physician and their relatives. The inclusion criteria for the patients in the study were age between 40 and 85 years, Hoehn and Yahr I, II, or III, literate, and a Mini-Mental State Examination scores > 24. Prior to the P300 assessment, auditory brainstem responses (ABRs) were recorded in all participants. Individuals with a clearly identifiable wave V at stimulus intensities of ≤ 20 dB nHL were considered to have normal hearing and were included in the study, whereas those who did not meet this criterion were excluded. Patients with severe hearing/visual impairment, orthopedic, neurological, or metabolic diseases that could prevent study participation, dementia, those undergoing deep brain stimulation, or those taking anticholinergic, antidepressant, or anxiolytic medications that could affect cognition were excluded.

Measurements

Attention was assessed using a neuro-audio/P300 device. The P300 (P3) wave is a cortical evoked potential that reflects the cognitive process by distinguishing variable stimuli. It was performed with the “oddball paradigm (surprising stimulus

sequence)” by randomly placing different stimuli into a standard stimulus sequence. After the patient was in a comfortable sitting position and the bilateral mastoid, forehead region, and area between the two eyebrows were cleaned with gel, reference electrodes (M1, M2) were placed on the bilateral mastoid processes, the positive electrode on the vertex (Cz), and the ground electrode (Fpz) on the forehead region. The patient was then instructed, “Two different sounds will come to your two ears at the same time randomly. One of the sounds will be high-pitched, and the other will be low-pitched. I want you to pay attention to the high-pitched sounds and count them. I want you to report the number to me at the end of the test.” After the information was provided, the headphones were placed bilaterally, and the impedances were checked. P300 latency was calculated as the time from the stimulus onset to the point where the amplitude had a positive or negative peak and was recorded in milliseconds (ms).⁴

Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI). This is a test that evaluates sleep quality in the preceding month. A maximum of 21 points can be obtained from the test. A score of ≥ 5 indicates poor sleep quality.¹¹

Depression was assessed with the Beck Depression Inventory (BDI). It is a 3-point likert-type scale with 21 items. Higher scores are associated with higher depression severity.¹²

Balance was assessed using the Berg Balance Scale (BBS). The 14-item scale evaluates balance during sitting, standing, and position changes. The best score is recorded as 56.¹³

Statistical Analyses

Data were analyzed using the IBM SPSS Statistics for Windows, version 25.0 (IBM Inc., Armonk, NY, USA). The Shapiro–Wilk test was used to evaluate compliance with the normal distribution. Descriptive statistics are presented as means (standard deviations), frequencies (percentages), medians, and min-max values. An independent sample t-test was used to compare data between groups. Multiple linear regression analysis was used to examine the effects of balance, depression, and sleep quality on the P300, and the P300 was determined as the dependent variable. The statistical significance level of all parameters was accepted as $P < 0.05$.

RESULTS

We included 55 participants: 28 individuals with PD and 27 healthy controls. The groups did not differ significantly in terms of age, height, weight, sex distribution, educational status, or dominant side (all $P > 0.05$) (Table 1).

The mean P300 latency was significantly longer in the PD group (342.42 ± 45.31 ms) than in the control group (267.99 ± 20.72 ms, $P < 0.001$) (Table 1). Figure 1 shows representative P300 waveforms from both groups.

BBS scores were significantly lower in the PD group (32.85 ± 4.45) than in the controls (44.07 ± 3.14 , $P < 0.001$). BDI

scores were higher in patients with PD (26.28 ± 1.95) than in controls (12.33 ± 2.14 , $P < 0.001$). PSQI scores were also higher in patients with PD (10.85 ± 1.40) than in controls (6.40 ± 1.39 , $P < 0.001$) (Table 1).

Multiple linear regression analysis revealed that BBS ($\beta = -0.482$, $P = 0.001$) and PSQI ($\beta = 0.394$, $P = 0.007$) scores were significantly associated with P300 latency, together explaining approximately 89% of the variance (adjusted $R^2 = 0.893$). Depression (BDI scores)

Table 1. Comparison of demographic and clinical characteristics between groups

	Parkinson Group (n = 28)	Control Group (n = 27)	P
Age (years)	64.67 $\pm$ 9.98	61.55 $\pm$ 6.97	0.186 ^a
Height (cm)	167.67 $\pm$ 8.48	167.59 $\pm$ 7.121	0.968 ^a
Weight (kg)	78.85 $\pm$ 11.08	80.48 $\pm$ 8.47	0.545 ^a
Sex (Female/Male)	6/22	11/16	0.121 ^b
Educational status (primary school/high school/university)	11/12/5	12/11/4	0.914 ^b
Time of diagnosis			
0–6 months	1		
6–12 months	4		
1–5 years	23		
Dominant side (Right/Left)	25/3	25/2	0.336 ^b
Hoehn and Yahr (I/II/III)	5/15/8		
P300 (ms)	342.42 $\pm$ 45.31 (250.70–422.00)	267.99 $\pm$ 20.72 (235.60–343.60)	0.001 ^a
BBS	32.85 $\pm$ 4.45 (26–40)	44.07 $\pm$ 3.14 (36–52)	0.001 ^a
BDI	26.28 $\pm$ 1.95 (22–30)	12.33 $\pm$ 2.14 (9–17)	0.001 ^a
PSQI	10.85 $\pm$ 1.40 (8–13)	6.40 $\pm$ 1.39 (5–9)	0.001 ^a

Mean (standard deviation); ^aIndependent sample t-test; ^bChi-Squared test; BBS = Berg Balance Scale; BDI = Beck Depression Inventory; PSQI = Pittsburgh Sleep Quality Index.

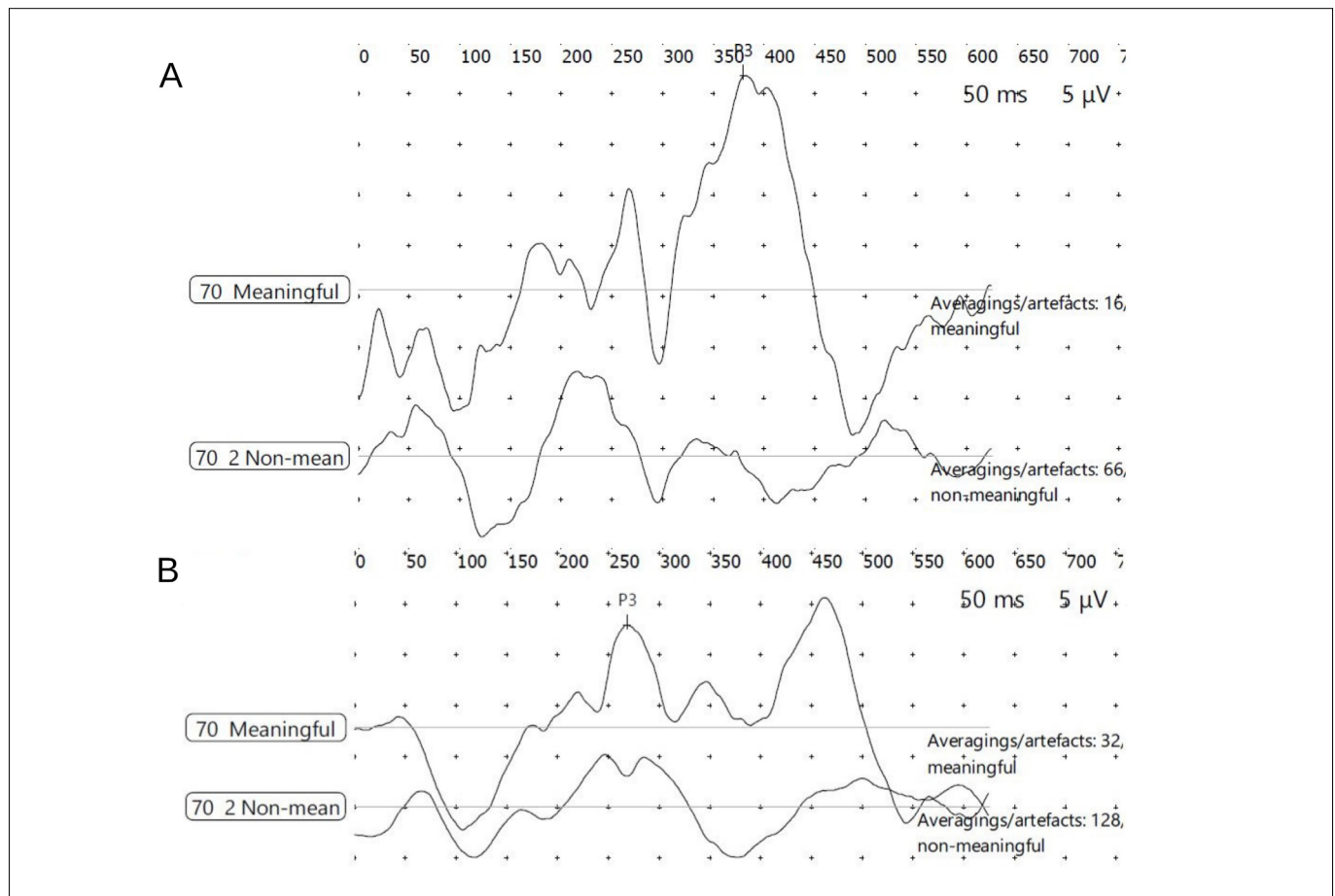


Figure 1. P300 latencies: A. Parkinson's disease, B. Healthy individual.

was not significantly associated with P300 latency ($P = 0.084$) (Table 2). Collinearity diagnostics were performed, and the variance inflation factor (VIF) values ranged from 1.74 to 4.45, indicating no severe multicollinearity among the independent variables.

Sensitivity analyses using separate linear regression models showed that balance performance ($R^2 = 0.837$), sleep quality ($R^2 = 0.827$), and depressive symptoms ($R^2 = 0.484$) were significantly associated with P300 latency.

DISCUSSION

This study demonstrated that P300 latencies were prolonged in individuals with PD compared to healthy controls. In addition, patients with PD exhibited differences in balance, sleep quality, and depressive symptoms. Balance performance and sleep quality were significantly associated with P300 latency, whereas no significant association was observed for depressive symptoms.

PD not only causes motor dysfunction but also affects cognitive function, attention, sleep, and depression. P300 is widely used as a neural marker of attentional processing,⁴ and most previous studies have reported prolonged P300 latencies in individuals with PD,⁵ suggesting cognitive slowing. However, a limited number of studies have found no significant differences when compared with healthy controls.⁶ Our findings are consistent with the majority of the literature, demonstrating delayed P300 latency and extending this evidence to individuals with PD without dementia.

The relationship between motor and non-motor symptoms in attentional processes has not been fully clarified. In the present study, balance performance was strongly associated with P300 latency. This finding is in line with previous studies, which have demonstrated a close relationship between motor performance, gait, and attention.^{14,15} It may reflect shared neural mechanisms involving fronto-subcortical circuits and dopaminergic pathways that contribute to both motor and cognitive functions.^{16,17} Although previous studies have examined the relationship between balance and attention using objective measures, to our knowledge, this study is among the first to investigate this association in Parkinson's disease by combining dynamic balance assessment with P300 latency analysis.

Depression and sleep disorders are common non-motor symptoms in patients with PD. This study showed that there were

differences in sleep quality and depression compared to healthy controls. Sleep quality was significantly associated with P300 latency, suggesting that worse sleep may be associated with worse attentional processing in patients with PD. In contrast, depressive symptoms demonstrated a more moderate association with P300 latency. However, this finding should be interpreted with caution. The borderline P value ($P = 0.084$) and the relatively small sample size suggest that the study may have been underpowered to detect a potential association, and a type II error cannot be excluded. Therefore, the absence of a statistically significant association does not necessarily indicate the absence of a true relationship. Previous studies have reported inconsistent findings regarding the relationship between depression and cognitive function in Parkinson's disease.^{4,8,20} For example, Norman et al. reported that while patients with depression showed worse performance in memory domains, no differences were observed in other cognitive functions. These inconsistencies suggest that the relationship between depression and cognition in PD is complex and may be influenced by multiple clinical and neurobiological factors.¹⁸ In addition, the heterogeneous nature of depression, the presence of multiple contributing factors, and the underlying pathophysiology of PD may have contributed to the lack of a clear association observed in this study.

Other non-motor symptoms affecting the quality of life in patients with PD are sleep disorders. Impaired serotonin and dopaminergic systems also lead to sleep difficulties. It can be suggested that there is an association between sleep disorders and attention in patients with PD.¹⁹ However, more comprehensive studies are needed to understand whether this situation is due to the pathology of the disease or only to changes in cognitive function. Stavitski et al. investigated the relationship between sleep quality and cognitive function in PD and showed that poor sleep is associated with attention and executive function.²⁰ The basal ganglia and cortical structures project to the hypothalamus, which regulates sleep-wake regulation.²¹ This may explain sleep disturbance in patients with PD. The relationship between attention levels and sleep may be due to some common pathological pathways, such as the involvement of the mesolimbic dopamine circuit that occurs with the disease.²⁰ Latreille et al. demonstrated that sleep spindle density can be used as a potential biomarker of cognitive function.²²

Given the strong interrelationships among balance performance, sleep quality, and depressive symptoms, the observed multicollinearity likely reflects overlapping clinical and neurobiological mechanisms inherent to PD. Rather than indicating statistical instability, this interrelatedness highlights the complex interaction between motor and non-motor symptoms influencing attentional processing, as reflected by P300 latency.

This study has several limitations that should be considered when interpreting the findings. First, the relatively small sample size may limit the stability of the regression coefficients and the

Table 2. Multiple linear regression analysis

	P300 (adjusted $R^2 = 0.893$)			
	B	S. E.	Beta	P value
Constant	274.546	92.887		0.007
BDI	3.472	1.923	0.150	0.084
PSQI	12.704	4.287	0.394	0.007
BBS	-4.909	1.294	-0.482	0.001

BBS = Berg Balance Scale; BDI = Beck Depression Inventory; PSQI = Pittsburgh Sleep Quality Index; SE = Standard Error.

generalizability of the findings. Second, the presence of multicollinearity among balance performance, sleep quality, and depressive symptoms restricts the interpretation of fully independent effects in multiple regression analyses. However, this interrelatedness likely reflects overlapping clinical and neurobiological mechanisms inherent to PD rather than a statistical artifact. To address this issue, sensitivity analyses using separate regression models were performed, which confirmed the robustness of the associations between P300 latency and balance performance, sleep quality, and depressive symptoms. Third, although disease duration and Hoehn and Yahr stages were recorded, they were not included in the regression analysis. In addition, PD-specific clinical variables, such as medication status (e.g., levodopa dosage and ON/OFF state), and non-motor symptoms, including fatigue, anxiety, and apathy, were not systematically assessed. Owing to the relatively small sample size, the number of predictors included in the regression model was intentionally limited to avoid overfitting and instability of the estimates. The absence of these covariates may have confounded the observed associations and should be considered when interpreting the results. Future studies with larger sample sizes and longitudinal designs are warranted to further clarify the relative and potentially causal contributions of motor and non-motor symptoms to attentional processing in PD.

CONCLUSION

This study demonstrated prolonged P300 latency in individuals with PD, along with differences in balance, sleep quality, and depressive symptoms compared to healthy controls. The findings further indicate that balance performance and sleep quality are significantly associated with attentional processing in PD. To our knowledge, this is one of the first studies to examine the combined associations of motor and non-motor features with attention using P300 measures in PD. These findings contribute to a better understanding of the complex interplay between motor and non-motor symptoms and their relationship with attentional processing in PD. However, the findings regarding depression should be interpreted with caution, as the lack of a significant association may be related to limited statistical power. Future studies with larger sample sizes and longitudinal designs are needed to further clarify these relationships.

REFERENCES

- Ferrazoli N, Donadon C, Rezende A, Skarzynski PH, Sanfins MD. The Application of P300-Long-Latency Auditory-Evoked Potential in Parkinson Disease. *International archives of otorhinolaryngology*. 2021;26(1):e158-66. PMID: 35096174; <https://doi.org/10.1055/s-0040-1722250>.
- Liu C, Zhang Y, Tang W, et al. Evoked potential changes in patients with Parkinson's disease. *Brain Behav*. 2017;7(5):e00703. PMID: 28523237; <https://doi.org/10.1002/brb3.703>.
- Rajendran D, Bandhu R, Gautam S, K Dhamija R, Mondal S. Auditory Evoked P300 Potential in Patients With Parkinson's Disease. *Cureus*. 2023;15(9):e45933. PMID: 37885554; <https://doi.org/10.7759/cureus.45933>.
- Sarikaya S, Yoldas TK, Yavasoglu NG. Evaluation of cognitive functions in Parkinson's patients without dementia with auditory event related potential (P300). *Dusunen Adam J Psychiatr Neurol Sci*. 2014;27(02):132-7. <https://dx.doi.org/10.5350/DAJPN2014270205>.
- Xu H, Gu L, Zhang S, et al. N200 and P300 component changes in Parkinson's disease: a meta-analysis. *Neurol Sci*. 2022;43(12):6719-30. PMID: 35982362; <https://doi.org/10.1007/s10072-022-06348-6>.
- Silva Lopes Md, Souza Melo Ad, Nóbrega AC. Delayed latencies of auditory evoked potential P300 are associated with the severity of Parkinson's disease in older patients. *Arq Neuropsiquiatr*. 2014;72(4):296-300. PMID: 24760094; <https://doi.org/10.1590/0004-282x20140005>.
- Wang Q, Meng L, Pang J, Zhu X, Ming D. Characterization of EEG data revealing relationships with cognitive and motor symptoms in Parkinson's disease: a systematic review. *Front Aging Neurosci*. 2020;12:587396. PMID: 33240076; <https://doi.org/10.3389/fnagi.2020.587396>.
- Gunn DG, Naismith SL, Terpening Z, Lewis SJG. The relationships between poor sleep efficiency and mild cognitive impairment in Parkinson disease. *J Geriatr Psychiatry Neurol*. 2014;27(2):77-84. PMID: 24196660; <https://doi.org/10.1177/0891988713509135>.
- Park JH, Lee SH, Kim Y, et al. Depressive symptoms are associated with worse cognitive prognosis in patients with newly diagnosed idiopathic Parkinson disease. *Psychogeriatrics*. 2020;20(6):880-90. PMID: 32840032; <https://doi.org/10.1111/psyg.12601>.
- Lee NY, Lee DK, Song HS. Effect of virtual reality dance exercise on the balance, activities of daily living, and depressive disorder status of Parkinson's disease patients. *J Phys Ther Sci*. 2015;27(1):145-7. PMID: 25642060; <https://doi.org/10.1589/jpts.27.145>.
- Ahn JH, Kim M, Park S, et al. Prolonged-release melatonin in Parkinson's disease patients with a poor sleep quality: a randomized trial. *Parkinsonism Relat Disord*. 2020;75:50-4. PMID: 32480307; <https://doi.org/10.1016/j.parkreldis.2020.03.029>.
- de Hemptinne C, Chen W, Racine CA, et al. Prefrontal Physiometers of Anxiety and Depression in Parkinson's Disease. *Front Neurosci*. 2021;15:748165. PMID: 34744613; <https://doi.org/10.3389/fnins.2021.748165>.
- Feng H, Li C, Liu J, et al. Virtual Reality Rehabilitation Versus Conventional Physical Therapy for Improving Balance and Gait in Parkinson's Disease Patients: A Randomized Controlled Trial. *Med Sci Monit*. 2019;25:4186-92. PMID: 31165721; <https://doi.org/10.12659/msm.916455>.
- Stuart S, Galna B, Delicato LS, Lord S, Rochester L. Direct and indirect effects of attention and visual function on gait impairment in Parkinson's disease: influence of task and turning. *Eur J Neurosci*. 2017;46(1):1703-16. PMID: 28444834; <https://doi.org/10.1111/ejn.13589>.
- Dana A, Shams A, Allafan N, Bahrami A. The relationship between attention and static balance disturbance in patients with Parkinson's disease. *Neurol Sci*. 2021;42(12):5107-15. PMID: 33774761; <https://doi.org/10.1007/s10072-021-05184-4>.

16. Stefani A, Lozano AM, Peppe A, et al. Bilateral deep brain stimulation of the pedunculo-pontine and subthalamic nuclei in severe Parkinson's disease. *Brain*. 2007;130(Pt 6):1596-607. PMID: 17251240; <https://doi.org/10.1093/brain/awl346>.
17. Souza CO, Voos MC, Barbosa AF, et al. Relationship Between Posturography, Clinical Balance and Executive Function in Parkinson's Disease. *J Mot Behav*. 2019;51(2):212-21. PMID: 29683777; <https://doi.org/10.1080/00222895.2018.1458279>.
18. Norman S, Tröster AI, Fields JA, Brooks R. Effects of depression and Parkinson's disease on cognitive functioning. *J Neuropsychiatry Clin Neurosci*. 2002;14(1):31-6. PMID: 11884652; <https://doi.org/10.1176/jnp.14.1.31>.
19. Sewell KR, Erickson KI, Rainey-Smith SR, et al. Relationships between physical activity, sleep and cognitive function: A narrative review. *Neurosci Biobehav Rev*. 2021;130:369-78. PMID: 34506842; <https://doi.org/10.1016/j.neubiorev.2021.09.003>.
20. Stavitsky K, Neargarder S, Bogdanova Y, McNamara P, Cronin-Golomb A. The impact of sleep quality on cognitive functioning in Parkinson's disease. *J Int Neuropsychol Soc*. 2012;18(1):108-17. PMID: 22152279; <https://doi.org/10.1017/s1355617711001482>.
21. Gunn DG, Naismith SL, Lewis SJ. Sleep disturbances in Parkinson disease and their potential role in heterogeneity. *J Geriatr Psychiatry Neurol*. 2010;23(2):131-7. PMID: 20101072; <https://doi.org/10.1177/0891988709358591>.
22. Latreille V, Carrier J, Lafortune M, et al. Sleep spindles in Parkinson's disease may predict the development of dementia. *Neurobiol Aging*. 2015;36(2):1083-90. PMID: 25442116; <https://doi.org/10.1016/j.neurobiolaging.2014.09.009>.

Authors' contributions: Ozaltin GE: conceptualization (equal), investigation (equal), methodology (equal); data curation (equal), project administration (equal), writing – review and editing (equal); Surname NN: investigation (equal), methodology (equal); Simsek A: conceptualization (equal), investigation (equal), data curation (equal), methodology (equal), writing – review and editing (equal); Candiri B: conceptualization (equal), investigation (equal), methodology (equal); data curation (equal), project administration (equal), writing – review and editing (equal); Algul FE: conceptualization (equal), investigation (equal), methodology (equal); project administration (equal), writing – review and editing (equal); Tecellioglu M: conceptualization (equal), investigation (equal), methodology (equal); project administration (equal), writing – review and editing (equal); supervision (equal); Talu B: conceptualization (equal), investigation (equal), methodology (equal); project administration (equal), writing – review and editing (equal); supervision (equal). All authors reviewed and approved the final version submitted for publication

Acknowledgments: We would like to thank all the participants who voluntarily participated in this study

Sources of funding: None

Conflicts of interest: None

Data availability statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. Owing to ethical considerations, the data are not publicly available

Declaration of generative AI in scientific writing: The authors did not use artificial intelligence tools for data analysis or production

Date of first submission: October 17, 2025

Last received: February 26, 2026

Accepted: March 10, 2026

Address for correspondence:

Gulfem Ezgi Ozaltin

Department of Physiotherapy and Rehabilitation, Faculty of Health Sciences, Inonu University

İnönü University Main Campus, 44280 Battalgazi, Malatya, Türkiye.

Phone: +9053 0285628

E-mail: gulfemezgi@gmail.com

Editor responsible for the evaluation process:

Paulo Manuel Pêgo-Fernandes, MD, PhD



Family functionality and social support in Brazilian older people with chronic kidney disease in hemodialysis: a cross-sectional study

Ana Beatriz Murari Thomazini^I, Diana Gabriela Mendes dos Santos^{II}, Yngrid Larisse Teodoro de Souza^{III}, Daiane Baldo Apolinário^{IV}, Ana Luiza Barbosa^V, Marisa Silvana Zazzetta^{VI}, Fabiana de Souza Orlandi^{VII}

Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil

^IUndergraduate student, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0009-0003-0435-9088>

^{II}MSc. Doctoral student, Department of Nursing, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil

<https://orcid.org/0000-0003-1572-5322>

^{III}Undergraduate student, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0009-0001-7421-6213>

^{IV}MSc. Doctoral Student, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0000-0002-3492-6114>

^VUndergraduate Student, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0009-0000-1741-6377>

^{VI}PhD. Associate Professor, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0000-0001-6544-767X>

^{VII}PhD. Associate Professor, Department of Gerontology, Universidade Federal de São Carlos (UFSCAR), São Carlos (SP), Brazil.

<https://orcid.org/0000-0002-5714-6890>

KEY WORDS (MeSH terms):

Renal insufficiency, chronic.
Renal dialysis.
Social support.
Family support.

AUTHORS' KEY WORDS:

Chronic kidney disease.
Hemodialysis.
Social support.

ABSTRACT

BACKGROUND: End-stage chronic kidney disease (CKD) requires hemodialysis, which carries significant biopsychosocial impact.

OBJECTIVES: This study aimed to investigate the relationship between family functionality and social support among older individuals undergoing hemodialysis.

DESIGN AND SETTING: This cross-sectional and correlational study was conducted at a renal replacement therapy unit located in São Paulo State.

METHODS: A total of 118 participants aged ≥ 60 years undergoing hemodialysis were interviewed individually using the Family adaptability, partnership, growth, affection, and resolve (APGAR), and the Medical Outcomes Study Social Support Scale. All ethical principles were duly respected.

RESULTS: Majority of participants were male (69.5%), aged between 60–69 years (50.8%), white ethnicity (63.6%), with a stable partner (54.2%), and had an income of one to three times the minimum wage (55.9%). Regarding the domains of social support, affective support (AS) had the highest mean score (93.44 ± 12.97), whereas informational support (IS) had the lowest (81.69 ± 24.60). Overall, 83.9% reported good family functioning, 9.3% moderate family dysfunction, and 6.8% severe dysfunction. A statistically significant positive correlation with a moderate magnitude was confirmed between APGAR with all support domains. Material support (MS, $r = 0.327$), AS ($r = 0.425$), emotional support (ES, $r = 0.518$), positive social interaction support (PSIS, $r = 0.466$), and IS ($r = 0.421$).

CONCLUSION: Social support was positively associated with family functioning, indicating that patients with CKD undergoing hemodialysis who reported better family functioning also perceived higher levels of social support.

INTRODUCTION

Noncommunicable diseases (NCDs) are long-lasting conditions arising from a mix of inherited traits, physiological functions, lifestyle choices, and environmental influences. According to the World Health Organization (WHO),¹ NCDs account for a disproportionate burden in low- and middle-income countries, where approximately 75% of the global mortality is attributed to these conditions. NCDs include chronic kidney disease (CKD), which affects approximately 10% of the global population. According to WHO estimates, CKD affects an estimated 6.7% of adults in Brazil, with prevalence approximately tripling among older adults aged 60 years and older.²

According to the kidney disease: Improving Global Outcomes CKD Work Group,³ CKD is characterized by structural or functional abnormalities of the kidneys that persist for a minimum of three months with health implications on the individual. The classification of CKD relies on three key parameters to establish a definitive diagnosis: the cause, glomerular filtration rate, and albuminuria level. Patients with end-stage CKD require treatment, and hemodialysis is one of the main treatment modalities. Although hemodialysis improves health status and prolongs patient survival, it neither halts disease progression nor fully reproduce the normal kidney function.^{4,5}

Individuals receiving hemodialysis treatment must follow strict therapeutic routines that demand significant changes to their daily habits and lifestyle, with important emotional, social, and family repercussions. Given this context, social support is a key determinant to enhance the quality of life among individuals receiving hemodialysis. Sułkowski et al.⁴ highlighted social support as an essential element in reducing the negative impact of CKD in patients undergoing hemodialysis and in promoting better health outcomes, especially among older individuals.⁶

Considering the above, family support is also crucial during hemodialysis as it can influence both health status and quality of life. According to Hulu et al.,⁷ greater family support leads to better the quality of life of the patient. Close relationships with family, friends, and other significant individuals play a central role in shaping the support patients received, and higher family resilience are associated with increased support. In addition, evidence suggests that among older individuals with CKD undergoing hemodialysis, stronger and more consistent family involvement is associated with improved psychological outcomes.^{5,8}

Despite the recognized importance of social and familial support for older adults undergoing hemodialysis, research from Brazil in this area remains limited. Few studies, such as those by Santos et al.⁹ and Wang et al.¹⁰ have simultaneously examined family functioning and social support in this population, and even fewer have employed adjusted analyses or longitudinal designs. This gap highlights the need for a deeper investigation into how these psychosocial factors interact and influence patient outcomes over time.

OBJECTIVE

This study aimed to investigate the relationship between family functionality and social support among older individuals undergoing hemodialysis in the Brazilian context.

METHODS

Study design and population

This cross-sectional and correlational study with a quantitative approach was conducted at a renal care facility in the interior of the state of São Paulo, Brazil.

Clinical data were collected from February to August 2025. Participants were recruited through non-probabilistic consecutive sampling based on the inclusion criteria of being 60 years or older, undergoing hemodialysis treatment for at least 6 months and whose oral communication ability was preserved. The exclusion criteria was a documented diagnosis of dementia in the medical chart that would make the interview unfeasible. For the sample size calculation, a correlation coefficient value of $r \geq 0.40$ was considered, given that $r = 0.40$ is indicative of a moderate correlation between psychosocial variables according to Ajzen & Fishbein,¹¹ and with the fixed values of $\alpha = 0.05$ and $\beta = 0.05$. This estimation yielded a minimum required number of 75 participants.

Data collection and definitions

At the beginning of participants' interactions, the study was explained to eligible participants, who were then invited to take part; those who agreed were provided with informed consent form. Subsequently, an individual interview was conducted using a structured instrument that included a sociodemographic,

socioeconomic, and health conditions questionnaire, the medical outcomes study (MOS) social support assessment instrument; and the Family adaptability, partnership, growth, affection, and resolve (APGAR) to assess family functioning.

The dependent variable was social support, assessed using the MOS Social Support Scale developed by Sherbourne and Stewart¹² and validated and adapted for use in Brazil by Andrade.¹³ The instrument consists of 19 items divided into five categories of support: tangible, affectionate, emotional, positive social interaction, and informational support. Each support type includes four questions, with five response options: 1 = "Never," 2 = "Rarely," 3 = "Sometimes," 4 = "Very frequently," and 5 = "Always." Each response contributes to the individual's perceived level of support. For four of the support domains, categories are summed and divided by 20, and the resulting value is multiplied by 100. Affectionate support is scored differently, as it includes only three questions; the total score is divided by 15 and then multiplied by 100. The final scores for each support category range from 20 to 100, with higher scores indicating greater perceived support.

The principal independent variable was family functionality, evaluated using the Family APGAR scale developed by Smilkstein¹⁴ and validated in Brazil by Duarte.¹⁵ The instrument consists of five questions, each offering five response options: 0 = "Never," 1 = "Rarely," 2 = "Sometimes," 3 = "Very Frequently," and 4 = "Always." The final score, ranging from 1 to 20 (where scores ≥ 12 indicate good functioning and scores < 12 suggest dysfunction), was operationalized as the independent variable in two formats for statistical analysis: the total raw score was treated as a continuous variable and utilized for comparisons of means and medians of social support, whereas the dichotomized score (functional vs. dysfunctional) was used as categorical variable and was employed to calculate Spearman's correlation with social support.

The sociodemographic and socioeconomic questionnaire included items related to age, sex, ethnicity, education level, income, and marital status.

Statistical analysis

The statistical analysis of the data was performed using SPSS software, v.22.0 (IBM Inc., Chicago, Illinois, United States). Descriptive analyses were performed, including central trend data (mean, minimum, and maximum) and dispersion measures (standard deviation). The Kolmogorov–Smirnov test was used to verify the absence of data normality and nonparametric tests were adopted.

To compare the social support scores, according to the levels of family functioning, the Kruskal–Wallis test was adopted. Spearman's correlation coefficients were calculated to verify the relationship between the continuous variables. In the present study, the magnitude of correlations was classified according to Levin and Fox's proposition (2004): weak (< 0.3); moderate (0.3 to 0.59); (0.6 to 0.9), perfect (1.0). The significance level adopted for the statistical tests was 5% ($P \leq 0.05$).¹⁶

Ethical Considerations

All ethical principles were respected, and the research project was approved prior to the start of the study by the Human Research Ethics Committee of the Universidade Federal de São Carlos (UFSCAR; approval no. 7.130.768; CAAE: 80181224.5.0000.5504).

RESULTS

This study included 118 older patients undergoing hemodialysis treatment who met the predefined eligibility criteria. Regarding participants' characterization, 82 were men (69.5%), 60 were aged between 60–69 years (50.8%), 75 were white (63.6%), and 64 reported having a stable partner (54.2%). Regarding income, 66 individuals reported a monthly income between one to three times the minimum wage (55.9%). Regarding education level, categorized as illiterate, 1–4 years, and ≥ 5 years, 65 participants had ≥ 5 years of education (55.1%) (**Table 1**).

Regarding social support domains, the informational support had the lowest mean score 81.69 (± 24.60), followed by positive

social interaction support 82.71 (± 23.08), affective support 85.25 (± 22.62), material support 92.92 (± 14.10), and emotional support 93.44 (± 12.97). Regarding family APGAR scores, 83.9% had good family functioning, 9.3% had moderate family dysfunction, and 6.8% had severe family dysfunction (**Table 2**).

Analysis of social support scores across family functioning levels revealed significant differences for all social support domains ($P < 0.001$). Participants categorized as having severe family dysfunction reported the lowest levels of social support. These levels were evident, especially regarding positive social interaction support. The level was the lowest in the severe family dysfunction group (mean = 55 [± 28.66]), whereas the mean for the good family functioning group was 87.78 (± 19.04). For informational support, the severe family dysfunction group had a mean of 57.5 (± 32.4), whereas the good family functioning group had a mean of 86.11 (± 21.56). The domain of emotional support had a mean score for severe dysfunction of 61.25 (± 28.62), whereas the group of good family function had a mean of 91.26 (± 16.10). Similar trends were observed across the other domains; material support had a mean of 80 (± 22.83) for the severe dysfunction group and 95.15 (± 11.54) for the good functioning group, followed by affective support, with a mean of 83.37 (± 18.11) for the severe dysfunction and a mean of 96.22 (± 10.05) for the good functioning group. These findings collectively underscored that dysfunctional families were strongly associated with a diminished perception of social support (**Table 3**).

In the correlational analysis, a moderate positive correlation was confirmed between family functioning and each domain of social support, with a statistically significant threshold of $P < 0.001$ for all domains of social support (**Table 4**).

Table 1. Percentage distribution of sociodemographic variables among older Brazilian patients undergoing hemodialysis. Ribeirão Preto, SP, Brazil, 2025 (n = 118)

Variables	Categories	n	%
Sex	Male	82	69.5
	Female	36	30.5
Age	60–69 years	60	50.8
	70–79 years	49	41.5
	80+ years	9	7.6
Ethnicity	White	75	63.6
	Non-White	43	36.4
Marital status	With a stable partner	64	54.2
	Without a stable partner	53	44.9
Income	Up to one the minimum wages	31	26.3
	From one to three the minimum wages	66	55.9
	Four or more the minimum wages	19	16.1
Education level	Illiterate	7	5.9
	From 1 to 4 years	46	39
	Equal or more than 5 years	65	55.1

DISCUSSION

Consistent with the study's objectives, the results demonstrated that most hemodialysis participants were men (69.5%), aged 60–69 years (50.8%), of white ethnicity (63.6%), and had a stable partner (54.2%). These findings are consistent with those of both Brazilian and international studies showing a predominance of men and older adults in hemodialysis programs, reflecting the epidemiological profile of CKD in adult and older aged populations. A substantial proportion of participants reported low

Table 2. Descriptive analysis of social support and family functioning. Ribeirão Preto, SP, Brazil, 2025, (n = 118)

Variables	Min	Max	Med	Mean	Standard Deviation
Social support					
Material support	30	100	100	92.92	± 14.10
Affective support	53	100	100	93.44	± 12.97
Emotional support	20	100	100	85.25	± 22.62
Informational support	20	100	95	81.69	± 24.60
Positive social interaction support	20	100	95	82.71	± 23.08

Table 3. Distribution statistics of the mean scores of social support according to the MOS questionnaire and levels of Family Functioning (APGAR). Ribeirão Preto, SP, Brazil, 2025 (n = 118)

Questionnaire	Categories	n	Mean	SD	P value
Material support	Severe family dysfunction	8	80	22.83	< 0.001a,b
	Moderate family dysfunction	11	82.27	18.62	
	Good family function	99	95.15	11.54	
Affective support	Severe family dysfunction	8	83.37	18.11	< 0.001a,b
	Moderate family dysfunction	11	75.73	15.63	
	Good family function	99	96.22	10.05	
Emotional support	Severe family dysfunction	8	61.25	28.62	< 0.001a,b
	Moderate family dysfunction	11	48.64	24.09	
	Good family function	99	91.26	16.10	
Positive social interaction Support	Severe family dysfunction	8	55	28.66	< 0.001a,b
	Moderate family dysfunction	11	57.27	22.84	
	Good family function	99	87.78	19.04	
Informational support	Severe family dysfunction	8	57.5	32.40	< 0.001a,b
	Moderate family dysfunction	11	59.54	24.23	
	Good family function	99	86.11	21.56	

*Severe family dysfunction $\neq$ Good family function; ^bModerate family dysfunction $\neq$ Good family function. SD = standard deviation.

Table 4. Spearman's correlation coefficients between Family Functioning (APGAR) and Social Support (MOS) scores, according to its dimensions (MS, AS, ES, PSIS, IS). Ribeirão Preto, SP, Brazil, 2025 (n = 118)

		MS	AS	ES	PSIS	IS
APGAR	r	0.327	0.425	0.518	0.466	0.421
	P value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

MS = material support; AS = affective support; ES = emotional support; PSIS = positive social interaction support; IS = informational support.

income (55.9%), earning between one and three times the minimum wages, in line with the Instituto Brasileiro de Geografia e Estatística (IBGE)¹⁸ classifications of economically vulnerable populations, reinforcing the association between CKD and unfavorable socioeconomic conditions. Notably, 55.1% of participants had completed ≥ 5 years of schooling, corroborating international findings that suggest a relatively high educational level attainment among hemodialysis patients.^{9,17,19}

Overall mean scores for social support were high, indicating generally positive perceptions. However, the mean scores for informational support (81.69 ± 24.60) and positive social interaction (82.71 ± 23.08) were lower than those for emotional (93.44 ± 12.97) and tangible support (92.92 ± 14.10). This pattern aligns with prior reports and may reflect the constraints of prolonged treatment, restricted social networks, and reliance on specialized information for disease management.^{4,9}

Family functioning, measured by the Family APGAR scale, revealed that 83.9% of participants had functional families, whereas 9.3% and 6.8% reported moderate and severe dysfunction, respectively. Consistent with the findings of Safi et al.,⁵ higher family

functioning was associated with greater perceived social support, highlighting the critical role of the family as a source of emotional and instrumental support during hemodialysis. Analysis by family functioning level revealed statistically significant differences across all support domains ($P < 0.001$). Participants from families with severe dysfunction had the lowest scores in terms of positive social interaction (55 ± 28.66) and informational support (57.5 ± 32.4), whereas those from functional families scored substantially higher (87.78 ± 19.04 and 86.11 ± 21.56 , respectively). These findings suggest that the family environment exerts a direct influence on social support perception, reinforcing its mediating role in psychological and social well-being.

Correlation analyses demonstrated moderate positive associations between family functioning and all social support domains ($P < 0.001$), indicating that better family dynamics are linked to higher perceived social support. This is consistent with the findings of Wang et al.,¹⁰ who reported a longitudinal relationship between family resilience and social support in patients undergoing kidney treatment, emphasizing the interplay between familial and psychosocial factors.

Recent studies have further underscored the potential of targeted interventions to enhance family support and resilience. Moradi et al.²⁰ showed that family-centered virtual programs improved mental health outcomes in patients and caregivers, demonstrating that remote interventions could effectively strengthen family support. Similarly, Supriati et al.²¹ found a significant correlation between peer social support and family psychological resilience, suggesting that interactions with individuals facing similar challenges

reinforce coping capacity. Wang et al.¹⁰ highlighted the mediating roles of family functioning and self-efficacy in the relationship between depression and resilience, while Anita et al.²² demonstrated that perceived family support was associated with lower depressive symptoms and improved health-related quality of life.

These findings provide valuable insights but should be interpreted in light of several limitations. The cross-sectional design precludes causal inferences, and reliance on self-reported measures may have introduced response biases. Additionally, the context-specific sample also limits generalizability to other populations or cultural settings. Future research should employ longitudinal designs, explore the efficacy of interventions across diverse socio-demographic contexts, and evaluate their impact on both psychosocial and clinical outcomes.

Overall, the results indicate that adequate family functioning is strongly related to positive perceptions of social support, particularly emotional and tangible support, among patients undergoing hemodialysis. These findings highlight the importance of multiprofessional care strategies that strengthen support networks, actively involve families, and develop coping mechanisms to promote resilience and overall well-being in individuals with CKD.

CONCLUSION

The findings of this study indicate a moderate correlation between social support and family functioning in older adults undergoing hemodialysis. The limited depth of existing literature on family functionality and social support in this population constrains the corroboration of our findings, underscoring the importance of future research focused on these variables to better inform clinical practice and theoretical development.

REFERENCES

- World Health Organization. Noncommunicable diseases. Available from: <https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases>. Accessed in 2025 (September 15).
- Brazil. Boletim Epidemiológico. Cenário da doença renal crônica no Brasil no período de 2010 a 2023. Ministério da Saúde. 2024;55(12):1-19. Available from: <https://www.gov.br/saude/pt-br/assuntos/noticias/2024/outubro/divulgado-boletim-epidemiologico-sobre-doenca-renal-cronica-no-brasil>. Accessed in 2025 (October 7).
- Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. *Kidney Int.* 2024;105(4S):S117-S314. PMID: 38490803; <https://doi.org/10.1016/j.kint.2023.10.018>.
- Sułkowski L, Matyja A, Matyja M. Social support and quality of life in hemodialysis patients: a comparative study with healthy controls. *Medicina.* 2024;60(11):1732. PMID: 39596917; <https://doi.org/10.3390/medicina60111732>.
- Safi F, Namdar Areshtanab H, Ghafourifard M, Ebrahimi H. The association between self-efficacy, perceived social support, and family resilience in patients undergoing hemodialysis: a cross-sectional study. *BMC Nephrol.* 2024;25(1):207. PMID: 38918709; <https://doi.org/10.1186/s12882-024-03629-4>.
- Kisomi ZS, Taherkhani O, Mollaei M, et al. The moderating role of social support in the relationship between death anxiety and resilience among dialysis patients. *BMC Nephrol.* 2024;25(1):100. PMID: 38493124; <https://doi.org/10.1186/s12882-024-03533-x>.
- Hulu T, Budi NP, Sari RP. Relationship of family support with quality of life of hemodialyzed patients using study literature review method. *J Vocat Nurs.* 2021;2(2):132-41. <https://doi.org/10.20473/jovin.v2i2.30720>.
- Pratiwi WN, Gayatri PR, Astutik WS, Pratama YG, Ayyaroh FA. The role of family support in stress and anxiety conditions in the elderly with chronic kidney disease underwent hemodialysis therapy. *Adi Husada Nurs J.* 2023;9(1):20-9. <https://doi.org/10.37036/ahnj.v9i1.392>.
- Santos DGM, Pallone JM, Manzini CSS, Zazzetta MS, Orlandi FS. Relationship between frailty, social support and family functionality of hemodialysis patients: a cross-sectional study. *Sao Paulo Med J.* 2021;139(6):570-5. PMID: 34706049; <https://doi.org/10.1590/1516-3180.2021.0089.r1.0904221>.
- Wang Y, Qiu Y, Ren L, et al. Social support, family resilience and psychological resilience among maintenance hemodialysis patients: a longitudinal study. *BMC Psychiatry.* 2024;24(1):76. PMID: 38279114; <https://doi.org/10.1186/s12888-024-05526-4>.
- Ajzen I, Fishbein M. Understanding attitudes and predicting social behavior. Englewood Cliffs: Prentice-Hall; 1980.
- Sherbourne CD, Stewart AL. The MOS social support survey. *Soc Sci Med.* 1991;32(6):705-14. PMID: 2035047; [https://doi.org/10.1016/0277-9536\(91\)90150-b](https://doi.org/10.1016/0277-9536(91)90150-b).
- Andrade CR, Chor D, Faerstein E, et al. Social support and breast self-examination in the Pró-Saúde Study. *Rev Saude Publica.* 2005;21(2):379-86. PMID: 15905900; <https://doi.org/10.1590/s0102-311x2005000200004>.
- Smilkstein G. The family APGAR: a proposal for a family function test and its use by physicians. *J Fam Pract.* 1978;6(6):1231-9. PMID: 660126.
- Duarte YAO. Família: rede de suporte ou fator estressor – a ótica de idosos e cuidadores familiares [doctoral thesis]. São Paulo: Escola de Enfermagem, Universidade de São Paulo; 2001.
- Levin J, Fox JA. Elementary statistics in social research. 9th ed. São Paulo: Pearson Prentice Hall; 2004.
- Paula EA, Roth JM, Schwartz E, Lima Spagnolo LM, Lise F. Perfil sociodemográfico e clínico de usuários em hemodiálise no sul do Rio Grande do Sul, Brasil. *Enferm Actual Costa Rica.* 2022;(43):51375. <http://dx.doi.org/10.15517/enferm.actual.cr.v0i43.45296>.
- Instituto Brasileiro de Geografia e Estatística (IBGE). Demographic Census 2022: population and households: preliminary results. Available from: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv102011.pdf>. Accessed in 2025 (October 7).

19. Alatawi AA, Alaamri M, Almutary H. Social support and adherence to treatment regimens among patients undergoing hemodialysis. *Healthcare*. 2024;12(19):1958. PMID: 39408138; <https://doi.org/10.3390/healthcare12191958>.
20. Moradi M, Amiri M, Daneshi S, et al. Investigating the effectiveness of a virtual family-centered support intervention on the mental health of hemodialysis patients and their family caregivers during the COVID-19 pandemic. *Open Public Health J*. 2024;17:e18749445271736. <http://dx.doi.org/10.2174/0118749445271736231123110235>.
21. Supriati L, Sunarto M, Ulya I, et al. The role of peer social support on family psychological resilience in caring for chronic kidney disease patients receiving hemodialysis. *Healthcare Low-Resour S*. 2024;12(s1):1-10. <https://doi.org/10.4081/hls.2024.13048>.
22. Anita DC, Alfrizki ES, Suprayitno E, Fadlilah S, Wantonoro W. Family support, depression, and health-related quality of life of hemodialysis patients. *Turk J Nephrol*. 2025;34(3):195-200. <https://doi.org/10.5152/turkjnephrol.2025.25991>.

Date of first submission: August 8, 2025

Last received: March 9, 2026

Accepted: March 9, 2026

Address for correspondence:

Fabiana de Souza Orlandi
Universidade Federal de São Carlos (UFSCar) - Departamento de Gerontologia
Rod. Washington Luiz, s/n
São Carlos - SP
CEP 13565-905
Tel. (+55 16) 3306-6673
E-mail: forlandi@ufscar.br

Editor responsible for the evaluation process:

Paulo Manuel Pêgo-Fernandes MD, PhD

Authors' contributions: Thomazini ABM; methodology (equal), investigation (equal) formal analysis (equal) writing – original draft (equal), writing - review and editing (equal) funding acquisition (equal); Santos DGM: methodology (equal) formal analysis (equal) writing - original draft (equal), writing - review and editing (equal); Souza YLT: investigation (equal), methodology (equal), writing - review and editing (equal); Apolinário DB: investigation (equal); methodology (equal); Barbosa AL: investigation (equal), methodology (equal), writing - review and editing (equal); Zazzetta MS: project administration (equal), data curation (equal), methodology (equal), writing – review and editing (equal), resources (equal), supervision (equal), funding acquisition (equal); Orlandi FS: supervision (equal), project administration (equal), data curation (equal), methodology (equal), writing – review and editing (equal), resources (equal), funding acquisition (equal). All authors reviewed and approved the final version of the manuscript for publication

Sources of funding: This study was funded by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – 422048/2023-2) and Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – 2024/23513-8)

Conflicts of interest: None

Data availability statement: All data generated or analyzed during this study are included in this published article

Declaration of generative AI in scientific writing: During the preparation of this article, the author used Gemini Pro (Google) to improve readability, grammar and vocabulary accuracy. After using this tool, the authors reviewed and edited the content and take full responsibility for the content of the publication



Multilevel auditory and cognitive processing in post-stroke aphasia: associations with language performance using ABR, LLR, and P300: a cross-sectional observational study

Agit Şimşek^I, Nihal Sümeyye Ulutaş^{II}, Feyza Deniz Saman^{III}

Department of Speech and Language Therapy, Faculty of Health Sciences, İnönü University, Malatya, Türkiye

^IPhD. Assistant Professor, Department of Speech and Language Therapy, Faculty of Health Sciences, İnönü University, Malatya, Türkiye.

 <https://orcid.org/0000-0003-1572-1188>

^{II}PhD. Assistant Professor, Faculty of Medicine, Turgut Özal University, Ankara, Türkiye.

 <https://orcid.org/0000-0001-5768-9863>

^{III}MSc. Independent Researcher, Malatya, Türkiye.

 <https://orcid.org/0000-0001-8481-3348>

KEYWORDS (MeSH terms):

Aphasia.

Stroke.

Evoked Potentials, Auditory.

Evoked Potentials, Auditory, Brain Stem.

Event-Related Potentials, P300.

AUTHOR'S KEYWORDS:

Post-Stroke Aphasia.

Auditory Processing.

Late Latency Responses.

Cognitive Processing Speed.

Language Performance.

Speech and Language Therapy.

ABSTRACT

BACKGROUND: Aphasia following stroke is primarily characterized by language impairment; however, accumulating evidence suggests that deficits in auditory and cognitive processing may also contribute to impaired language function. Auditory evoked potentials provide objective markers of neural processing across subcortical, cortical, and cognitive levels and may help clarify the neurophysiological mechanisms underlying post-stroke aphasia.

OBJECTIVES: To investigate multilevel auditory processing and its relationship with language performance in individuals with post-stroke aphasia using Auditory Brainstem Responses (ABR), Late Latency Responses (LLR), and P300 potentials.

DESIGN AND SETTING: Cross-sectional observational study conducted at a university hospital in Türkiye.

METHODS: Twenty-nine individuals with post-stroke aphasia and 33 age-matched healthy controls were included. Language performance was assessed using the Aphasia Language Assessment Test and a standardized naming task. Auditory processing was evaluated using ABR, LLR (P1, N1, P2, N2), and P300 potentials recorded according to standard electrophysiological protocols. Group comparisons were performed for latency and amplitude measures, and correlation analyses were conducted to examine associations between electrophysiological parameters and language performance.

RESULTS: Compared with healthy controls, individuals with aphasia demonstrated significantly prolonged latencies in ABR waves and interpeak intervals, as well as in all LLR and P300 components ($p < 0.05$). No significant group differences were observed in amplitude measures. Naming performance was significantly lower in the aphasia group. Although correlations between language performance and electrophysiological measures did not reach statistical significance, moderate negative trends were observed between naming scores and N2 and P300 latencies.

CONCLUSION: Post-stroke aphasia is associated with delayed auditory processing across subcortical, cortical, and cognitive levels, reflecting generalized slowing rather than reduced neural recruitment. Latency-based auditory evoked potential measures may complement behavioral language assessments and support a multilevel auditory–cognitive framework for aphasia evaluation and rehabilitation.

INTRODUCTION

Post-stroke aphasia is among the most common and disabling outcomes of cerebrovascular accidents, affecting approximately one-third of stroke survivors and substantially impairing communication and quality of life.¹ Although aphasia has traditionally been conceptualized as a disorder of language production and comprehension, growing evidence indicates that it also involves impairments in auditory processing, attention, and cognitive control mechanisms that support functional language use.^{2,3} From a speech and language therapy perspective, this broader framework highlights the importance of integrating linguistic and neurofunctional approaches in aphasia assessment.

Language performance, particularly naming ability, is a core component of clinical aphasia evaluation and a sensitive indicator of impairment severity and recovery.⁴ Naming tasks require coordinated auditory perception, lexical access, semantic processing, and executive control, making them vulnerable to disruption across multiple neural levels. Although behavioral language assessments provide essential information regarding functional outcomes, they may not fully capture the neural mechanisms underlying impaired language performance, emphasizing the need for complementary objective measures.

Auditory evoked potentials (AEPs) provide a non-invasive and temporally precise method for examining neural processing along the auditory pathway. Auditory Brainstem Responses (ABR) reflect early subcortical auditory transmission, Late Latency Responses (LLR) index cortical auditory detection and discrimination, and the P300 component is associated with attention allocation and working memory updating.⁵ Previous aphasia studies have primarily focused on isolated cortical components and consistently reported prolonged latencies, suggesting slowed auditory and cognitive processing.³ However, amplitude findings remain inconsistent, and subcortical auditory processing has received limited attention. Although ABR is not routinely included in aphasia assessment, early auditory timing may influence the quality of auditory input available for higher-level language processing.

Emerging evidence suggests that early auditory processing may also be impaired in post-stroke aphasia. Mourad et al.⁶ reported delayed auditory brainstem responses in individuals with aphasia, raising the possibility that inefficiencies at early auditory stages contribute to downstream cortical and cognitive processing delays. Despite this theoretical and clinical relevance, few studies have examined brainstem, cortical, and cognitive auditory responses within the same cohort, limiting integrative interpretations of auditory–language interactions.

Accordingly, this study aimed to investigate language performance—particularly naming ability—and auditory evoked potentials across multiple neural levels (ABR, LLR, and P300) in individuals with post-stroke aphasia compared with age- and sex-matched healthy controls. We hypothesized that individuals with aphasia would exhibit prolonged electrophysiological latencies reflecting slowed auditory and cognitive processing and that these delays would be associated with reduced language performance. By integrating behavioral and electrophysiological measures, this study aims to contribute to objective multilevel assessment frameworks in speech and language therapy.

OBJECTIVE

To investigate auditory and cognitive processing across multiple neural levels and examine their association with language performance in individuals with post-stroke aphasia using ABR, LLR, and P300 event-related potentials.

METHODS

Study design and participants

This cross-sectional observational study was conducted at İnönü University. The study included 29 individuals with post-stroke aphasia and 33 age- and sex-matched neurologically healthy controls. Aphasia diagnosis was established by a speech and language therapist based on clinical evaluation and standardized language assessment. Inclusion criteria for the aphasia group were: (i) history

of ischemic or hemorrhagic stroke, (ii) aphasia diagnosis, (iii) age 18–80 years, and (iv) adequate hearing sensitivity for auditory testing. Exclusion criteria for all participants included additional neurological disorders, severe cognitive impairment, psychiatric illness, and uncorrected hearing loss. Pure-tone audiometry confirmed hearing thresholds ≤ 25 dB HL at 500–4,000 Hz in all participants.

Ethical approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the İnönü University Health Sciences Scientific Research Ethics Committee (Approval No 2024/6333; Date: September 24, 2024). Written informed consent was obtained from all participants prior to participation.

Language assessment

Language performance was assessed using the Aphasia Language Assessment Test (ADD) and a naming task. The ADD was administered according to standardized procedures to evaluate overall language ability. In the naming task, participants were asked to name visually or auditorily presented stimuli. Each correct response received one point, whereas incorrect responses received zero points. Total scores were used for statistical analyses.

Electrophysiological recordings

Auditory evoked potentials were recorded in a sound-attenuated room using the Neurosoft Neuro Audio system. Electrode placement followed the international 10–20 system, and electrode impedance was maintained below 5 k Ω throughout testing.

ABR: ABR recordings were obtained using 90 dB nHL rarefaction click stimuli presented at a repetition rate of 21.1/s. Filter settings were 150–3,000 Hz. Absolute latencies of Waves I, III, and V, as well as interpeak intervals (I–III and I–V), were analyzed for both ears.

LLR: LLR recordings were elicited using 65 dB click stimuli presented at a repetition rate of 1.1/s with filter settings of 30–150 Hz. Latencies of the P1, N1, P2, and N2 components and peak-to-peak amplitudes (P1–N1 and P2–N2) were measured. Click stimuli were preferred to ensure temporal precision and comparability across subcortical and cortical auditory responses.

P300 Potentials: P300 responses were recorded using an auditory oddball paradigm with 1,000 Hz frequent and 2,000 Hz rare tones. Participants were instructed to mentally count the rare stimuli. Recordings were obtained from the Cz electrode referenced to M1–M2, with Fpz as ground. P300 latency was measured as the primary outcome. Only participants achieving $\geq 80\%$ accuracy in the counting task were included in the analysis.

Statistical analysis

Statistical analyses were performed using SPSS version 23.0. Data normality was assessed using the Shapiro–Wilk and Kolmogorov–Smirnov

tests. Group comparisons were conducted using independent samples t-tests for normally distributed variables and Mann-Whitney U tests for non-normally distributed variables. Analysis of covariance (ANCOVA) was performed with age as a covariate. False discovery rate (FDR) correction was applied to control for multiple comparisons. Correlations between language performance and electrophysiological measures were examined using partial correlation analyses controlling for age. Statistical significance was set at $p < 0.05$.

RESULTS

Participant characteristics

A total of 62 participants were included, comprising 29 individuals with post-stroke aphasia and 33 healthy controls. Demographic characteristics of both groups are presented in Table 1. No statistically significant differences were observed between groups in age or sex distribution ($p > 0.05$).

Language performance

Language assessment results are summarized in Table 2. The aphasia group demonstrated significantly lower Aphasia Language Assessment Test (ADD) total scores and ADD success percentages than the control group after adjustment for age. These differences remained statistically significant following FDR correction ($p < 0.001$).

ABR

Group comparisons of ABR latency and interpeak interval measures are presented in Table 2. After adjustment for age using ANCOVA

and correction for multiple comparisons, the aphasia group showed significantly prolonged Wave III and Wave V latencies, as well as prolonged Wave I–III interpeak intervals, in both ears compared with controls (FDR-adjusted $p < 0.05$). No significant group differences were observed for Wave V–I interpeak intervals.

LLR

LLR component latencies and amplitudes are shown in Table 2. Latencies of the P1, N1, P2, and N2 components were significantly prolonged in the aphasia group compared with the control group after age adjustment and FDR correction (all FDR-adjusted $p < 0.01$). No statistically significant group differences were observed for P1–N1 or P2–N2 amplitude measures.

Figure 1 illustrates representative LLR waveforms from an individual with aphasia and a healthy control participant. Visual inspection demonstrates prolonged P1, N1, P2, and N2 latencies in the aphasia group relative to the control participant, whereas overall waveform morphology remained preserved.

P300 potentials

P300 latency values are presented in Table 2 and Figure 2. The aphasia group exhibited significantly prolonged P300 latencies compared with the control group after controlling for age ($F = 15.34$, $p < 0.001$), and this difference remained significant following FDR correction (FDR-adjusted $p = 0.001$).

Figure 3 presents group-level comparisons of LLR and P300 latency values across participants. Latency measures for P1, N1, P2, N2, and P300 components were consistently longer in the aphasia group than in controls, reflecting generalized delays in auditory and cognitive processing.

Correlation between language performance and electrophysiological measures

Partial correlation analyses controlling for age were conducted within the aphasia group. As shown in Table 3, moderate negative

Table 1. Demographic characteristics of the study groups

Variable	Aphasia group (n = 29)	Control group (n = 33)	p value
Age (years), mean $\pm$ SD	55.9 $\pm$ 7.8	53.1 $\pm$ 8.3	0.207
Sex (male/female), n	20/9	19/14	0.368

Table 2. Comparison of language and electrophysiological measures between groups (ANCOVA with age as covariate, FDR-corrected)

Measure	Aphasia group (mean $\pm$ SD/median)	Control group (mean $\pm$ SD/median)	ANCOVA F	p value	FDR-adjusted p
ADD Total Score	15.28 $\pm$ 4.96	19.67 $\pm$ 0.82	25.34	< 0.001	< 0.001
ADD Success (%)	79.66 $\pm$ 28.1	98.33 $\pm$ 4.7	18.67	< 0.001	< 0.001
Right ABR Wave III (ms)	3.85 $\pm$ 0.04	3.54 $\pm$ 0.04	12.67	< 0.001	0.002
Right ABR Wave V (ms)	6.00 $\pm$ 0.39	5.65 $\pm$ 0.28	8.92	0.005	0.009
Left ABR Wave III (ms)	3.89 $\pm$ 0.05	3.58 $\pm$ 0.03	13.45	< 0.001	0.002
Left ABR Wave V (ms)	5.93 $\pm$ 0.46	5.62 $\pm$ 0.21	7.89	0.007	0.012
P1 Latency (ms)	80.86 $\pm$ 5.7	66.84 $\pm$ 1.22	9.23	0.004	0.009
N1 Latency (ms)	128.43 $\pm$ 7.15	98.97 $\pm$ 1.88	22.18	< 0.001	< 0.001
P2 Latency (ms)	192.99 $\pm$ 6.74	120.96 $\pm$ 2.59	35.67	< 0.001	< 0.001
N2 Latency (ms)	249.15 $\pm$ 8.54	145.99 $\pm$ 2.45	42.15	< 0.001	< 0.001
P300 Latency (ms)	346.76 $\pm$ 41.3	269.72 $\pm$ 18.7	15.34	< 0.001	< 0.001

correlations were observed between naming performance and N2 latency ($\rho = -0.324$, $p = 0.087$) and P300 latency ($\rho = -0.337$, $p = 0.073$); however, these associations did not reach statistical significance. No significant correlations were observed between language performance and ABR or earlier LLR components ($p > 0.05$).

Overall, latency measures consistently differentiated individuals with aphasia from controls across subcortical, cortical, and cognitive levels, whereas amplitude measures did not.

DISCUSSION

The present study examined language performance and auditory evoked potentials across multiple neural levels in individuals with post-stroke aphasia. By integrating ABR, LLR, and P300 potentials within the same cohort, this study provides a comprehensive

electrophysiological characterization of auditory-cognitive processing in aphasia. The findings demonstrated that individuals with aphasia exhibited significant latency prolongations at both subcortical and cortical levels, whereas amplitude measures remained largely preserved. In parallel, language performance—particularly naming ability—was significantly reduced in the aphasia group. Collectively, these findings suggest that aphasia is associated with generalized slowing in auditory and cognitive processing while neural response magnitude remains preserved.

Auditory brainstem processing

A key finding of this study was the presence of prolonged ABR wave latencies in individuals with aphasia. Although aphasia is traditionally conceptualized as a cortical language disorder,

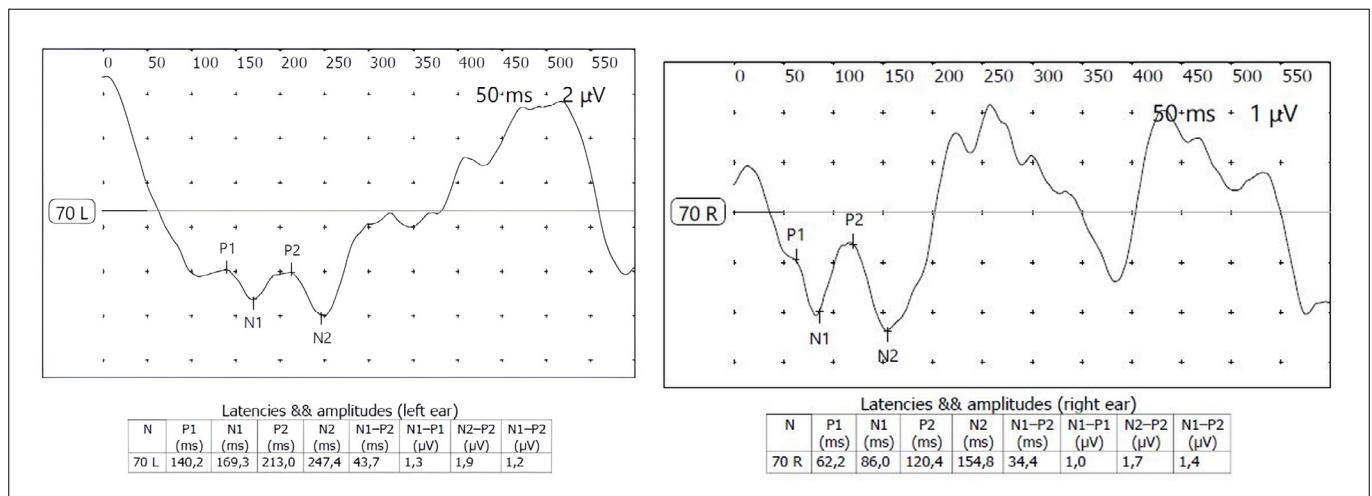


Figure 1. The left panel presents the LLR (Late Latency Response) waveform of a participant from the aphasia group, while the right panel illustrates the LLR waveform of a participant from the control group.

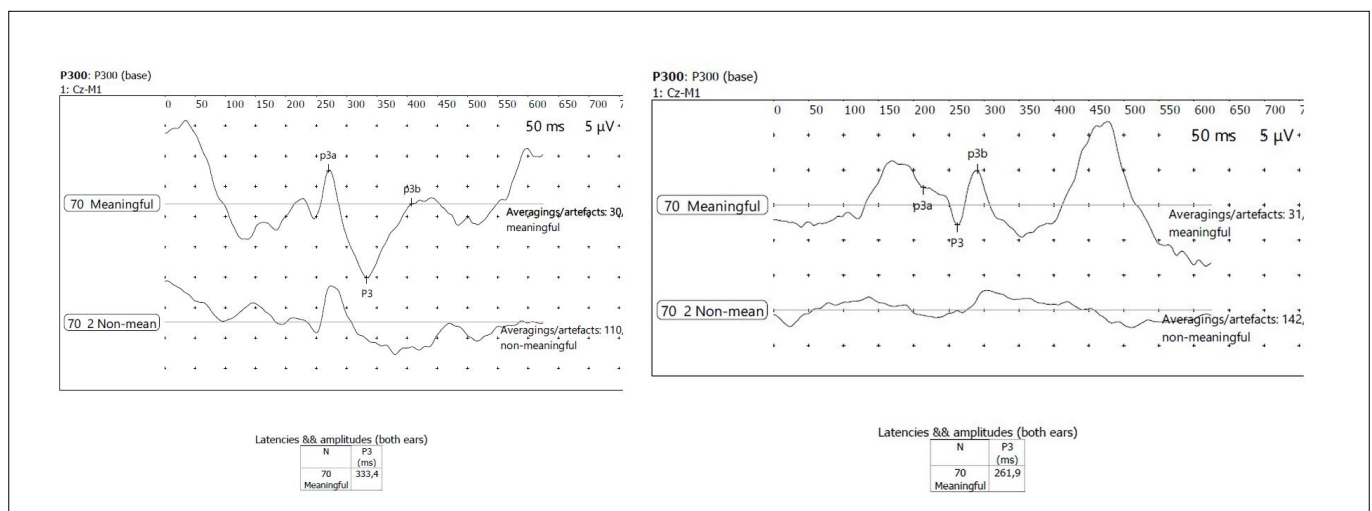


Figure 2. The left panel displays the P300 waveform of a participant from the aphasia group, while the right panel shows the P300 waveform of a participant from the control group.

delayed Waves III and V and prolonged interpeak intervals suggest reduced efficiency in subcortical auditory transmission following stroke. Similar findings were reported by Mourad et al.⁶ in post-stroke aphasia. Nevertheless, ABR findings should be interpreted cautiously because subcortical auditory responses may also be influenced by age-related and vascular factors. Rather than indicating primary peripheral auditory pathology, the observed delays likely reflect altered neural synchrony within the auditory brainstem, which may contribute to degraded auditory input available for higher-level language processing.

Cortical auditory processing and LLR

At the cortical level, individuals with aphasia demonstrated consistent latency prolongations across all LLR components (P1, N1, P2, and N2), whereas amplitude measures did not differ significantly from controls. This pattern suggests preserved cortical responsiveness with reduced temporal efficiency. Previous studies have similarly reported latency-dominant alterations in cortical auditory evoked potentials in aphasia.³ Given the temporal precision required for speech perception and lexical access, such cortical processing delays may interfere with the rapid integration of auditory information necessary for naming and functional communication.

Auditory-cognitive processing and P300

The most robust electrophysiological difference observed in this study was the marked prolongation of P300 latency in the aphasia

group. The P300 component is widely regarded as an index of attention allocation, working memory updating, and stimulus classification.⁵ Prolonged P300 latency therefore reflects slowed cognitive processing rather than diminished attentional capacity. This finding is consistent with previous literature demonstrating delayed P300 responses in aphasia and supports the view that language impairment is embedded within broader auditory-cognitive processing inefficiencies. From a speech and language therapy perspective, these findings underscore the importance of addressing attentional and processing speed components alongside linguistic targets in rehabilitation programs.

Relationship between language performance and electrophysiological measures

Correlation analyses revealed moderate negative trends between language performance and both N2 and P300 latencies, although these associations did not reach statistical significance. The absence of statistically significant correlations may partly reflect limited statistical power due to the modest sample size rather than a lack of functional association. Nevertheless, the observed trends suggest that delayed cortical and cognitive processing may be functionally relevant to naming performance in aphasia, consistent with prior reports linking cortical auditory evoked potentials to language recovery.⁷ Future studies with larger cohorts are warranted to further clarify these relationships.

Clinical and translational implications

From a clinical and translational perspective, the present findings highlight the value of latency-based auditory evoked potential measures as objective indices that may complement behavioral language assessments in aphasia.² The preservation of response amplitudes suggests retained neural responsiveness, supporting the potential for rehabilitation-induced plasticity. Integrating auditory-processing and attention-focused interventions into speech and language therapy protocols may therefore improve treatment outcomes. Moreover, electrophysiological measures may provide sensitive tools for monitoring neurophysiological changes during rehabilitation and tailoring individualized intervention strategies.⁸ Such measures may help clinicians identify processing-level limitations not fully captured by behavioral language tests alone.

Limitations and future directions

Several limitations should be acknowledged. The modest sample size may have limited the detection of statistically significant associations between electrophysiological measures and language performance. Additionally, aphasia subtypes and lesion characteristics were not systematically analyzed, which may have contributed to variability in electrophysiological responses. Future research should incorporate lesion localization, aphasia subtype

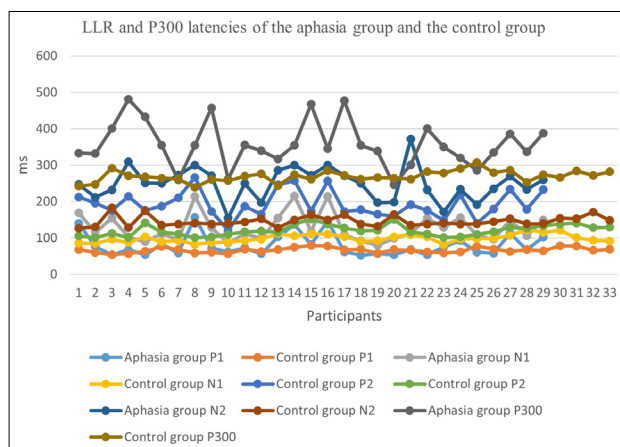


Figure 3. Latency values (ms) of LLR and P300 components in the aphasia and control groups.

Table 3. Partial correlations between language performance and electrophysiological measures in the aphasia group (controlled for age)

Measure	ρ	p value
N2 Latency (ms)	-0.324	0.087
P300 Latency (ms)	-0.337	0.073

classification, and longitudinal designs to better elucidate the role of auditory–cognitive processing in language recovery. The lack of statistically significant correlations likely reflects limited statistical power rather than the absence of a functional relationship.

CONCLUSION

In conclusion, this study demonstrates that post-stroke aphasia is characterized by delayed auditory processing at both subcortical and cortical levels, accompanied by slowed cognitive processing reflected by prolonged P300 latency. These electrophysiological alterations parallel impairments in language performance and support a multilevel auditory–cognitive framework for understanding aphasia. By integrating objective neural measures with behavioral language assessment, this study contributes to the development of more comprehensive and clinically relevant approaches in speech and language therapy.

REFERENCES

- Engelter ST, Gostynski M, Papa S, et al. Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. *Stroke*. 2006 Jun;37(6):1379–84. <https://doi.org/10.1161/01.STR.0000221815.64093.8c>.
- Shah-Basak P, Boukrina O, Li XR, Jebahi F, Kielar A. Targeted neurorehabilitation strategies in post-stroke aphasia. *Restorative Neurology and Neuroscience*. 08 2023 Dec;41(3–4):129–91. <https://doi.org/10.3233/RNN-231344>.
- Meechan RJH, McCann CM, Purdy SC. The electrophysiology of aphasia: A scoping review. *Clinical Neurophysiology*. 2021 Dec;132(12):3025–34. <https://doi.org/10.1016/j.clinph.2021.08.023>.
- Kertesz A. Western Aphasia Battery—Revised (WAB-R) [Database record] [Internet]. APA PsycTests; 2006. [cited 2026 May 14]. <https://doi.org/10.1037/t15168-000>.
- Polich J. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*. 2007 Oct;118(10):2128–48. <https://doi.org/10.1016/j.clinph.2007.04.019>.
- Mourad S, Abd Al-Ghaffar M, Bassiony MAA, Fawzi G. Value of complex evoked auditory brainstem response in patients with post-stroke aphasia (prospective study). *Egyptian Journal of Ear, Nose, Throat and Allied Sciences*. 2017 Nov;18(3):217–21. <https://doi.org/10.1016/j.ejenta.2017.05.006>.
- Fonseca J, Ferreira JJ, Pavão Martins I. Cognitive performance in aphasia due to stroke: a systematic review. *International Journal on Disability and Human Development*. 24 2017 Jan;16(2). <https://doi.org/10.1515/ijdh-2016-0011>.
- Rojas Sosa MC, Fraire Martínez MI, Olvera Gómez JL, Jáuregui-Renaud K. Early auditory middle latency evoked potentials correlates with recovery from aphasia after stroke. *Clinical Neurophysiology*. 2009 Jan;120(1):136–9. <https://doi.org/10.1016/j.clinph.2008.10.011>.

Authors' contributions: Şimşek A: conceptualization, methodology, formal analysis, investigation, data curation, writing – original draft, writing – review and editing, project administration. Ulutaş NS: investigation, data curation, formal analysis, writing – original draft, visualization. Saman FD: investigation, resources, data curation, writing – review and editing. All authors reviewed and approved the final version of the manuscript for publication.

Acknowledgments: The authors would like to thank all participants for their time and cooperation in this study. The authors also thank the clinical staff who supported data collection and assessment procedures.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: Data supporting the findings of this study are available upon request from the corresponding author, Agit Şimşek.

Declaration of generative AI in scientific writing: During the preparation of this study, the authors did not use generative AI or AI-assisted technologies.

Date of first submission: February 14, 2026

Last received: May 13, 2026

Accepted: April 8, 2026

Address for correspondence:

Agit Şimşek
Dil ve Konuşma Terapisi Bölümü, Sağlık Bilimleri Fakültesi, İnönü Üniversitesi
Elazığ Yolu 15. km
Battalgazi — Malatya — Türkiye
Tel. (+90 422) 3410220
E-mail: agit.simsek@inonu.edu.tr

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Prevalence of melasma in pregnant women and its relationship with body perception: a cross-sectional study in Türkiye

Demet Toktaş¹, Rabia Atilla^{II}

Niğde Ömer Halisdemir University, Niğde, Türkiye

¹MSc. Midwife, Niğde Ömer Halisdemir University Training and Research Hospital, Niğde, Türkiye.

 <https://orcid.org/0000-0001-5807-6639>

^{II}PhD. Assistant Professor, Department of Nursing, Zübeyde Hanım Faculty of Health Sciences, Niğde Ömer Halisdemir University, Niğde, Türkiye.

 <https://orcid.org/0000-0003-2854-9508>

KEYWORDS (MeSH terms):

Pregnancy.
Melanosis.
Chloasma.
Body image.
Perception.

AUTHOR'S KEYWORDS:

Prevalence.
Risk factors.
Melasma.
Body perception.
Nursing.

ABSTRACT

BACKGROUND: Melasma is one of the most common physiological hyperpigmentation disorders observed during pregnancy, and its clinical characteristics and risk factors have been extensively investigated. However, evidence regarding the relationship between pregnancy-related melasma and body perception remains limited.

OBJECTIVE: This descriptive cross-sectional study aimed to determine the prevalence of melasma among pregnant women and to examine its association with body perception.

METHODS: This study included 450 pregnant women in the central city center of the Central Anatolia Region of Türkiye who met the inclusion criteria. Data were collected using the Introductory Information Form, the Melasma and Skin Type Characteristics Assessment Form, and the Body Perception Scale (BPS) and analyzed using descriptive statistics, comparative analyses, and logistic regression.

RESULTS: Melasma was identified in 37.8% of the participants, and the centrofacial pattern was the most common presentation (56.5%). A statistically significant association was found between pregnancy-related melasma and total BPS scores ($p < 0.05$). Pregnant women without melasma had significantly higher mean BPS scores than those with melasma (150.83 ± 17.18 vs. 146.15 ± 15.16 ; $p = 0.004$). Logistic regression analysis demonstrated that low educational and income levels, a personal or family history of melasma, and Fitzpatrick skin type IV were significant predictors of pregnancy-related melasma ($p < 0.05$), accounting for 61% of the variance.

CONCLUSION: The findings indicate the need for a holistic assessment of physiological skin changes and body perception during pregnancy, as well as the integration of counseling, education, and nursing care interventions targeting risk factors into routine prenatal care.

INTRODUCTION

Melasma is an acquired, chronic hyperpigmentation disorder characterized by increased melanin deposition in the epidermal and/or dermal layers of the skin.¹ Clinically, it presents as asymptomatic light-to-dark-brown macules with irregular borders and symmetrical distribution, most commonly affecting sun-exposed areas, such as the face and neck.^{1,2} Although the exact etiology remains unclear, genetic predisposition, environmental exposure, and hormonal changes are known to contribute to its development. Pregnancy is a major risk factor, largely because of increased levels of estrogen, progesterone, and melanocyte-stimulating hormone.^{2,3} The prevalence of melasma during pregnancy has been reported to range from 36.4% to 75%, varying by geographic region and skin type.³ Although spontaneous regression is generally expected within one year postpartum, approximately 30% of cases persist.⁴ Although considered a physiological change, pregnancy-related melasma may lead to aesthetic concerns and negatively affect body image perception among pregnant women.^{5,6}

Body perception refers to an individual's perception of their physical appearance, along with associated feelings and thoughts.⁷ Physiological and psychological changes during pregnancy influence body perceptions of the expectant mother in multiple ways⁸ and in some cases, may contribute to poor body image.⁹ During pregnancy, discrepancies between changes in a woman's face and body and her internalized ideal body image may result in either positive or negative body image perceptions.⁶ Visible skin changes, particularly pregnancy mask (melasma), striae gravidarum, and edema, have been associated with increased body dissatisfaction.⁵ Qualitative studies have shown that many pregnant women perceive pregnancy-related melasma as distressing

and upsetting, report feeling unattractive, limit social interactions, and experience anxiety about whether their skin will return to its pre-pregnancy state.^{6,10}

Healthcare professionals play a significant role in identifying aesthetic concerns related to pregnancy-associated melasma and planning preventive and treatment-related care.^{4,11,12} However, given the limited number of studies addressing this topic, there is a need for a comprehensive examination of the prevalence of pregnancy-related melasma and its association with body perception.

OBJECTIVES

This study aimed to determine the prevalence of melasma among pregnant women and to examine its relationship with body perception.

METHODS

Study design and population

This study employed descriptive cross-sectional design. It was conducted between March and May 2024 in a city in the Central Anatolia Region of Türkiye.

The study population comprised pregnant women in their second or third trimester who attended routine examinations at the obstetrics and gynecology outpatient clinic of a training and research hospital in the Central Anatolia Region of Türkiye. The inclusion criteria were as follows: pregnant women aged ≥ 18 years who were in the second or third trimester, able to communicate in Turkish, and voluntarily agreed to participate. Pregnant women in the first trimester; those with multiple pregnancies; those diagnosed during pregnancy with dermatological conditions (e.g., psoriasis, atopic dermatitis, urticaria, connective tissue diseases, skin cancer, and pemphigus); those with pregnancy-related dermatoses, including hemangioma, cholestasis, spider angioma, palmar erythema, and malignant melanoma; those with facial manifestations of these conditions; and those with a diagnosed psychiatric disorder were excluded.

The sample size was calculated using the formula for estimating sample size in populations of unknown size. Based on an assumed melasma prevalence of 23.5%,¹² the minimum required sample size was calculated to be 278.¹³ Participants were recruited between March 1 and May 15, 2024, using convenience sampling. The study was completed with the participation of a total of 450 pregnant women who met the inclusion criteria during the data collection process and consented to participate.

Measures

Introductory information form: This form was developed by the researchers based on a review of relevant literature^{2,4,5,12,14,15} and expert opinions. It comprised 21 items assessing participants' sociodemographic and obstetric characteristics.

Melasma and skin type characteristics assessment form: This form comprised nine items assessing participants' skin type, presence of melasma, personal and family history of melasma, use of preventive measures, and health-care-seeking behaviors, including physician consultation and receipt of treatment.^{2,4,5,12,14} The findings obtained through the physical examination conducted by the researcher—including the pregnant women's natural hair color, eye color, skin tone, presence and number of facial freckles/moles, pregnancy melasma status, melasma type, and skin type characteristics—were evaluated and recorded using this form. Skin type was additionally assessed by the researchers through physical examination according to the Fitzpatrick skin type classification,¹⁶ and the findings were recorded on the form.

Body perception scale (BPS): The BPS is a 40-item instrument designed to assess satisfaction with different body regions and bodily functions. In this five-point Likert-type scale, each item is rated from 1 (strongly dislike) to 5 (strongly like). Total scores range from 40 to 200, with higher scores indicating greater satisfaction. In the Turkish reliability study conducted by Hovardaoğlu,¹⁷ Cronbach's alpha coefficient was 0.91; in the present study, it was 0.896.

Data collection

Data were collected through face-to-face interviews conducted by researchers between March and May 2024, using the 'Introductory information form', 'Melasma and skin type characteristics assessment form', and 'BPS'. Facial regions of the participants were assessed by the researchers during a physical examination. Melasma status and skin type were recorded using relevant assessment forms.

Statistical analyses

Data were analyzed using IBM SPSS Statistics for Windows (version 25; IBM Corp., Armonk, New York). Descriptive statistics, including frequencies, percentages, means, and standard deviations, were used to summarize the data. Normality of the data distribution was assessed using the Kolmogorov-Smirnov test. As the data were normally distributed, independent-samples *t*-tests were used. Logistic regression analysis was performed to evaluate categorical dependent variables. Variables that were significant in the chi-square analyses were entered into the regression model. Statistical significance was set at $p < 0.05$.

Ethical considerations

Ethical approval was obtained from the Ethics Committee of the university in the province where the study was conducted (Decision No: 2024/01-65; January 31, 2024). Institutional permission was also obtained, and all procedures were performed in accordance with the principles of the Declaration of Helsinki.

RESULTS

The distribution of selected descriptive characteristics of the pregnant women is presented in Table 1. Pregnancy-related melasma was identified in 37.8% of the participants, of whom 56.5% exhibited the centrofacial type. Furthermore, 40.2% of all women had Fitzpatrick skin type III. Overall, 31.6% of participants were high school graduates, 79.1% were unemployed, 62.8% resided in urban areas, and 70.9% reported that their income was equal to their expenses. Most of the pregnant women reported no history

of melasma prior to pregnancy (90.9%), during a previous pregnancy (72.3%), or in their family (76.9%). The majority (99.3%) reported not using any specific preventive methods for melasma; 46.2% applied sunscreen to the face, 21.1% reported consistently applying sunscreen, and 88.2% did not reapply it during the day.

A comparison of BPS scores according to pregnancy-related melasma status is presented in Table 2. A significant difference was observed between the total BPS scores and the presence of melasma during pregnancy ($p < 0.05$). The mean total BPS score was higher

Table 1. Distribution of selected descriptive characteristics of the pregnant women (n = 450)

Characteristics		$\bar{X} \pm SD$	Min–Max/Range
Age (years)		27.23 $\pm$ 5.29	18–44 (26)
Gravidity		2.2 $\pm$ 1.33	1–8 (7)
Gestational age (weeks)		33.63 $\pm$ 4.75	16–41 (25)
Pregnancy BMI (kg/m ²)		29.62 $\pm$ 4.55	19.1–44.4 (25.3)
Gestational weight gain (kg)		11.11 $\pm$ 5.16	0–26 (26)
		n	%
Presence of pregnancy-related melasma	Present	170	37.8
	Not present	280	62.2
Type of pregnancy melasma (n = 170)	Centrofacial type	96	56.4
	Malar type	70	41.2
	Mandibular type	4	2.4
Fitzpatrick skin type*	Skin type II	95	21.1
	Skin type III	181	40.2
	Skin type IV	174	38.7
Educational level	Literate/primary school	38	8.4
	Secondary school	105	23.3
	High school	142	31.6
	University	133	29.6
	Postgraduate	32	7.1
Employment status	Employed	94	20.9
	Unemployed	356	79.1
Place of residence	Province	283	62.8
	District	70	15.6
	Village	97	21.6
History of melasma before pregnancy	Present	41	9.1
	Not present	409	90.9
History of melasma in previous pregnancy (n = 274)	Present	76	27.7
	Not present	198	72.3
Family history of melasma	Present	104	23.1
	Not present	346	76.9
Use of preventive methods for melasma	Yes	3	0.7
	No	447	99.3
Use of sunscreen on the face	Yes	208	46.2
	No	242	53.8
Frequency of sunscreen use	Always	95	21.1
	Occasionally	113	25.1
	Never	242	53.8
Reapplication of sunscreen during the day	Yes	53	11.8
	No	397	88.2
Physician consultation for pregnancy melasma	Yes	5	1.1
	No	445	98.9

$\bar{X}$, mean; SD, standard deviation; Min–Max, minimum–maximum; n, calculations were based only on respondents; BMI, body mass index. *There were no participants with skin type I or V in the study.

Table 2. Comparison of body perception scale scores based on pregnancy-related melasma status

Pregnancy-Related Melasma Status	$\bar{X} \pm SD$	Min-Max (Range)	t test	p value
Present	146.15 $\pm$ 15.16	80–195 (115)		
Not present	150.83 $\pm$ 17.18	93–199 (106)	–2.924	0.004*

SD, standard deviation; Min–Max, minimum–maximum. * $p < 0.05$;

t test, independent samples t test.

among pregnant women without melasma (150.83 $\pm$ 17.18) than among those with melasma (146.15 $\pm$ 15.16).

The distribution of variables associated with the presence of pregnancy-related melasma is presented in Table 3. Logistic regression analysis showed that educational level, income status, history of melasma prior to pregnancy, history of melasma in previous pregnancies, family history of melasma, and skin type were significantly

Table 3. Distribution of variables associated with the presence of pregnancy-related melasma

	β	SE	Wald	<i>p</i> value	OR (B)	95% CI OR	
						Lower	Upper
Educational level							
<i>Literate/primary school</i>			10.082	0.039			
<i>Secondary school</i>	−0.861	0.588	2.139	0.144	0.423	0.133	1.34
<i>High school</i>	−1.662	0.627	7.031	0.008*	0.19	0.056	0.648
<i>University</i>	−1.961	0.702	7.811	0.005*	0.141	0.036	0.557
<i>Postgraduate</i>	−1.741	0.961	3.283	0.07	0.175	0.027	1.153
Place of residence							
<i>Village</i>			3.011	0.222			
<i>Province</i>	−0.858	0.507	2.866	0.09	0.424	0.157	1.145
<i>Village</i>	−0.296	0.581	0.26	0.61	0.744	0.238	2.323
Social security (None)							
<i>Yes</i>	0.467	0.458	1.041	0.308	1.595	0.65	3.914
Perceived income status							
<i>Less than expenses</i>			6.164	0.046*			
<i>Equal to expenses</i>	−1.201	0.51	5.536	0.019*	0.301	0.111	0.818
<i>Greater than expenses</i>	−0.519	0.747	0.483	0.487	0.595	0.138	2.572
Number of births							
<i>3 or more</i>			5.443	0.142			
<i>No previous births</i>	0.718	0.818	0.77	0.38	2.049	0.413	10.175
<i>1</i>	0.691	0.651	1.126	0.289	1.995	0.557	7.146
<i>2</i>	−0.282	0.65	0.188	0.665	0.754	0.211	2.697
Hormone therapy for infertility treatment (Not received)							
<i>Received</i>	0.489	0.718	0.464	0.496	1.631	0.399	6.657
History of combined oral contraceptive use (Not used)							
<i>Used</i>	0.094	0.55	0.029	0.864	1.099	0.374	3.227
History of melasma before pregnancy (Not present)							
<i>Present</i>	2.252	0.964	5.461	0.019*	9.503	1.438	62.809
History of melasma in previous pregnancy (Present)							
<i>Not present</i>	2.65	0.552	23.016	0*	14.16	4.795	41.814
Family history of melasma (Present)							
<i>Not Present</i>	1.335	0.475	7.907	0.005*	3.798	1.498	9.629
Frequency of sunscreen application							
<i>Never</i>			3.769	0.152			
<i>Always</i>	−0.579	0.574	1.016	0.313	0.561	0.182	1.727
<i>Occasionally</i>	0.59	0.445	1.764	0.184	1.805	0.755	4.314
Regular medication use (No)							
<i>Yes</i>	−0.364	0.613	0.352	0.553	0.695	0.209	2.311
Fitzpatrick skin type							
<i>Skin type II</i>			25.423	0*			
<i>Skin type III</i>	0.831	0.572	2.11	0.146	2.296	0.748	7.045
<i>Skin type IV</i>	2.515	0.582	18.649	0*	12.365	3.949	38.715
Model statistics							
Hosmer–Lemeshow Test = 11.246, <i>p</i> = 0.188; Cox & Snell R ² =0.457; Nagelkerke R ² = 0.61							

β , Regression coefficient; * $p < 0.05$; SE, standard error; OR, odds ratio; CI, confidence interval.

associated with pregnancy-related melasma ($p < 0.05$). The likelihood of pregnancy-related melasma was 3.32 times lower among women whose income was equal to their expenses than among those whose income was less than their expenses ($1/0.301 = 3.32$; odds ratio [OR] = 0.301; $p = 0.019$). Compared with women with primary school education or below, the likelihood of pregnancy-related melasma was 5.26 times lower among high school graduates ($1/0.19 = 5.26$; OR = 0.19; $p = 0.008$) and 7.09 times lower among university graduates ($1/0.141 = 7.09$; OR = 0.141; $p = 0.005$).

In contrast, the likelihood of pregnancy-related melasma was 12.36 times higher among women with Fitzpatrick skin type IV than among those with skin type II (OR = 12.365; $p < 0.001$). A history of melasma prior to pregnancy was associated with a 9.50-fold increase in the likelihood of pregnancy-related melasma (OR = 9.503; $p = 0.019$), a history of melasma in previous pregnancies with a 14.16-fold increase (OR = 14.16; $p < 0.001$), and a family history of melasma with a 3.80-fold increase (OR = 3.798; $p = 0.005$). The final model accounted for 61% of the variance in the presence of melasma during pregnancy.

DISCUSSION

This study examined the prevalence of pregnancy-related melasma, its association with body perception, and factors associated with its occurrence among pregnant women. The findings indicate that approximately four out of ten pregnant women experienced pregnancy-related melasma, and the presence of melasma was associated with negative body perception. Furthermore, skin type, history of melasma, and educational and income levels were identified as significant factors that correlated with its occurrence.

In this study, pregnancy-related melasma was identified in 37.8% of pregnant women based on physical examinations performed by the researchers. Globally, the prevalence of melasma during pregnancy has been reported to range from 36.4% to 75%.³ Substantial variations in prevalence rates have been reported across countries. Reported prevalence rates are 10.7% in Brazil,¹⁸ 15.8% in Iran,¹⁵ 50.8% in India,¹⁹ and 63.5% in Pakistan.²⁰ In Türkiye, the reported prevalence rates range from 23.5% to 46.9%.^{12,21} Differences in skin type, hormonal changes, and geographic factors may contribute to this variability. Although the prevalence observed in the present study is consistent with the findings from other studies conducted in Türkiye, it can be considered moderate relative to international reports.

Most participants in the present study had a Fitzpatrick skin type of III, and approximately half of the melasma cases exhibited a centrofacial type. These findings are consistent with previous reports indicating that skin type III is common among pregnant women,^{14,22,23} and that the centrofacial pattern is the most frequently observed clinical presentation in pregnant women with melasma.^{23,24} Furthermore, most participants did not use specific preventive measures for melasma. Although nearly half reported applying sunscreen

to the face, only one-fifth reported consistent use, and the majority did not reapply sunscreen during the day. Similar patterns of limited sun-protective behavior and irregular use of preventive measures during pregnancy have been reported in previous studies.^{22,25}

Melasma, although generally considered a physiological skin change during pregnancy, has been reported to adversely affect body perception.⁵ In this study, it was determined that women with pregnancy-related melasma had significantly lower BPS scores compared to women without melasma; in other words, melasma reduced body perception satisfaction in pregnant women. Our findings are consistent with both qualitative and quantitative studies reporting that pregnancy-related melasma is associated with a more negative body/face perception in women.^{5,6} In this context, pregnancy-related melasma should be considered not only a dermatological change but also an important condition that may affect women's body perception, and it should be addressed within a holistic approach in prenatal care.

The regression analysis revealed that education level, perceived income, history of melasma before and during previous pregnancies, family history of melasma, and skin type together explained 61% of the probability of melasma occurring during pregnancy. This demonstrates that melasma is not limited to pregnancy-specific physiological changes but constitutes a multifactorial condition resulting from the interaction of sociodemographic, obstetric, and genetic components, supporting the need for a holistic approach in risk assessment.

The results of this study suggest that high school and university graduation, as well as having an adequate income level, are factors that reduce the risk of melasma. Specifically, the probability of developing melasma was 5.26-fold lower in high school graduates and 7.09-fold lower in university graduates; also, it was 3.32-fold lower in pregnant women whose income equaled their expenses. This can be explained by the determining role of education and economic status in the health behaviors of individuals. The literature also reports that low income and education levels,²⁶ socioeconomic status, and living conditions²⁷ may be associated with the development of melasma. Women with low income and education levels may have a higher risk of melasma because of an increased likelihood of working outdoors and prolonged unprotected sun exposure, as well as limited access to dermatological care and treatment.^{2,26} Every ten hours spent outdoors per week increases the risk of pregnancy-related melasma by 27%.²⁸ However, individuals with higher socioeconomic status have better access to dermatological care services and photoprotective products, which strengthens preventive health behaviors and potentially reduce the development of melasma.²⁷ These results suggest that education and income levels exert protective effects, not only indirectly but also through behavioral mechanisms.

The history of melasma before pregnancy, during previous pregnancies, and familial predisposition significantly increased the risk of melasma in this study. The risk was 9.5-fold higher in those with

a history of melasma before pregnancy, 14.16-fold higher in those who experienced melasma in a previous pregnancy, and 3.8-fold higher in those with a family history clearly, clearly demonstrating the recurrent nature of melasma and the role of individual/genetic predisposition. Another study conducted in Türkiye reported that the risk was approximately 44-fold higher in women with a history of melasma in previous pregnancies, and that family history was also significantly associated with this increase.¹² Similarly, other studies emphasize a strong association between family history and melasma^{23,24,28–30} highlighting that genetic factors are among the important determinants.^{23,29} Based on our findings and the available literature, pregnant women with an individual or family history of melasma should be considered a high-risk group.

In the present study, Fitzpatrick skin type IV (brown/dark skin) was found to be the strongest risk factor for pregnancy-related melasma, with the probability of melasma being 12.36-fold higher in this group compared with those with skin type II (light skin). The literature also indicates that having a dark skin type (Fitzpatrick skin types III, IV, and V) is a significant risk factor for the development of melasma and for the spread of existing lesions.^{2,23,31} Melasma is rarely detected in extreme skin types (Fitzpatrick skin types I and VI) because of the limited pigmentation response, whereas it is frequently observed in intermediate skin types (Fitzpatrick skin types III, IV, and V). Melasma is more common in individuals with moderate skin types because of their more pronounced and variable melanin production responses when exposed to triggering factors, such as ultraviolet radiation. Skin type I has a low tanning capacity and therefore cannot produce sufficient melanin, whereas in skin type VI, the existing pigmentation level is already high, which limits the production of additional pigment.³¹ Previous studies have reported that pregnancy, as well as having dark skin (skin type V) and lifelong sun exposure, are significant risk factors for melasma.^{28,32} The risk of severe melasma is approximately three-fold higher under these circumstances.³² In light of these findings, it is important that pregnant women with dark skin types be more closely evaluated for melasma development during prenatal checkups and provided with counseling on preventive measures.

CONCLUSION

Pregnancy-related melasma is a common condition, affecting approximately 40% of pregnant women, and is associated with poorer body perception. Educational attainment and economic status were identified as protective factors, whereas personal history, genetic predisposition, and skin type emerged as important risk factors for its development. Accordingly, prenatal care should incorporate the assessment of physiological skin changes and body image perception. This should be accompanied by educational interventions aimed at increasing awareness of risk factors and appropriate counseling and supportive care.

REFERENCES

1. Rigopoulos D, Gregoriou S, Katsambas A. Hyperpigmentation and melasma. *J Cosmet Dermatol*. 2007 Sep;6(3):195–202. PMID: 17760699; <https://doi.org/10.1111/j.1473-2165.2007.00321.x>.
2. Handel AC, Miot LDB, Miot HA. Melasma: a clinical and epidemiological review. *An Bras Dermatol*. 2014 Sep;89(5):771–82. PMID: 25184917; <https://doi.org/10.1590/abd1806-4841.20143063>.
3. Zhao L, Fu X, Cheng H. Prevention of Melasma During Pregnancy: Risk Factors and Photoprotection-Focused Strategies. *Clin Cosmet Investig Dermatol*. 2024 Oct 15;17:2301–10. <https://doi.org/10.2147/CCID.S488663>. Erratum in: *Clin Cosmet Investig Dermatol*. 2024 Nov 02;17:2445–6. PMID: 39430643; <https://doi.org/10.2147/CCID.S503955>.
4. Lopes CCC. Prevention of Melasma in Pregnant Women: Aspects of Approach in Aesthetic Treatment. *EJ*. 2022;6(3):1–9. <https://doi.org/10.23880/eij-16000245>.
5. Kartal B, Kızılırmak A, Zedelenmez A, Erdem E. The effect of physical changes in pregnancy on women's body image. *Journal of Contemporary Medicine*. 2018 Dec 26. <https://doi.org/10.16899/gopctd.447159>.
6. Sohrabi Z, Kazemi A, Farajzadegan Z, Janighorban M. Body perception in pregnant women: a qualitative study. *BMC Pregnancy Childbirth*. 2023 Mar 11;23(1):165. <https://doi.org/10.1186/s12884-023-05467-y>.
7. Kale M, Köksal Akyol A, Karagöz S. Beden algısı: tanımı, tarihçesi ve beden algısıyla ilgili kavramlar. In: Kale M, editor. *Beden Algısı*. Ankara: Nobel Akademik Yayıncılık; 2023.
8. Linde K, Lehnig F, Nagl M, Stepan H, Kersting A. Course and prediction of body image dissatisfaction during pregnancy: a prospective study. *BMC Pregnancy Childbirth*. 2022 Sep 20;22(1):719. PMID: 36127633; <https://doi.org/10.1186/s12884-022-05050-x>.
9. Desai K, Poojari DP, Shwetha TS, Upadhyaya R, Ramachandra P. Understanding the body image perception of pregnant women during their third trimester in a tertiary care setting in Southern India. *BMC Pregnancy Childbirth*. 2024 Oct 14;24(1):671. PMID: 39402515; <https://doi.org/10.1186/s12884-024-06864-7>.
10. Kazemi F, Nahidi F, Kariman N. Disorders Affecting Quality of Life During Pregnancy: A Qualitative Study. *J Clin Diagn Res*. 2017 Apr;11(4):QC06–QC10. PMID: 28571212; <https://doi.org/10.7860/JCDR/2017/23703.9560>.
11. Chandra M, Paray AA. Natural Physiological Changes During Pregnancy. *Yale J Biol Med*. 2024 Mar 29;97(1):85–92. PMID: 38559455; <https://doi.org/10.59249/JTIV4138>.
12. Türkmen H, Yörük S. Risk factors of striae gravidarum and chloasma melasma and their effects on quality of life. *J of Cosmetic Dermatology*. 2023 Feb;22(2):603–12. <https://doi.org/10.1111/jocd.14783>.
13. Sümbüloğlu K, Sümbüloğlu V. *Biyoistatistik*. 15. baskı. Ankara: Hatiboglu Yayıncılık; 2012.
14. Kaplan Ö, Bülbül T, ŞahiN MG, Çınar SL, MüderriS İ. Kadınlarda Stria Gravidarum Oluşumu ile İlişkili Faktörler ve Vücut Algısı. *Ordu Üniversitesi Hemşirelik Çalışmaları Dergisi*. 2021 Sep 8;4(2):163–71. <https://doi.org/10.38108/ouhcd.868766>.
15. Moin A, Jabery Z, Fallah N. Prevalence and awareness of melasma during pregnancy. *Int J Dermatology*. 2006 Mar;45(3):285–8. PMID: 16533230; <https://doi.org/10.1111/j.1365-4632.2004.02470.x>.

16. Fitzpatrick TB. The Validity and Practicality of Sun-Reactive Skin Types I Through VI. *Arch Dermatol*. 1988 Jun 1;124(6):869. PMID: 3377516; <https://doi.org/10.1001/archderm.1988.01670060015008>.
17. Hovardaoğlu S. Vücut algısı ölçeği'nin geliştirilmesi. *Psikiyatri, Psikoloji, Psikofarmakoloji (3P) Dergisi*. 1993;1 (Özel sayı 2):21–7.
18. Hexsel D, Rodrigues TC, Dal'Forno T, Zechmeister-Prado D, Lima MM. Melasma and pregnancy in southern Brazil. *J Eur Acad Dermatol Venereol*. 2009 Mar;23(3):367–8. PMID: 18631207; <https://doi.org/10.1111/j.1468-3083.2008.02885.x>.
19. Rathore S, Gupta S, Gupta V. Pattern and prevalence of physiological cutaneous changes in pregnancy: A study of 2000 antenatal women. *Indian J Dermatol Venereol Leprol*. 2011;77(3):402. PMID: 21508591; <https://doi.org/10.4103/0378-6323.79741>.
20. Ikram S, Malik A, Suhail M. Physiological skin changes during pregnancy. *J Pakistan Assoc Dermatologists*. 2018;28(2):219–23. <https://doi.org/https://www.jpap.com.pk/index.php/jpap/article/view/1167>.
21. Özdemir S, Özdemir M. Gebelikte melazma. *Turkderm-Turkish Arch Dermatology Venereol*. 2006;40(3):98–100.
22. Ünlü E, Demirci Saadet E. Evaluation of Epidemiological and Clinical Features of Melasma Patients. *atfm*. 2023 Oct 24;76(3):220–6. <https://doi.org/10.4274/atfm.galenos.2023.24196>.
23. Tamega AA, Miot LDB, Bonfietti C et al. Clinical patterns and epidemiological characteristics of facial melasma in Brazilian women. *Acad Dermatol Venereol*. 2013 Feb;27(2):151–6. PMID: 22212073; <https://doi.org/10.1111/j.1468-3083.2011.04430.x>.
24. Platsidaki E, Markantoni V, Nicolaidou E et al. A Clinical and Epidemiological Single-Group Observational Study in the Greek Population. *Dermatol Ther (Heidelb)*. 2024 Nov;14(11):3183–92. PMID: 39485595; <https://doi.org/10.1007/s13555-024-01297-8>.
25. Aladwan S, Issa R, Al. Safadi W, Alnsour L, Al-Halaseh LK. Perceptions and Management of Pregnancy-Related Skin Changes: A Cross-Sectional Study on Knowledge, Practices, and Use of Skincare Product. *J Cosmetic Dermatology*. 2025 Apr;24(4):e70132. PMID: 40145230; <https://doi.org/10.1111/jocd.70132>.
26. Pollo CF, Miot LDB, Meneguín S, Miot HA. Factors associated with quality of life in facial melasma: a cross-sectional study. *Int J Cosmet Sci*. 2018 May 7. PMID: 29734511; <https://doi.org/10.1111/ics.12464>.
27. Ullman LE, Nasir-Moin M, Hoffman V et al. Sunscreen use and affordability attitudes based on ethnicity, socioeconomic status, and Fitzpatrick skin type. *Arch Dermatol Res*. 2024 May 25;316(6):266. PMID: 38795207; <https://doi.org/10.1007/s00403-024-02997-z>.
28. Ortonne JP, Arellano I, Berneburg M et al. A global survey of the role of ultraviolet radiation and hormonal influences in the development of melasma. *J Eur Acad Dermatol Venereol*. 2009 Nov;23(11):1254–62. PMID: 19486232; <https://doi.org/10.1111/j.1468-3083.2009.03295.x>.
29. Alifa D, Hamzah RA, E. Ladi J. Influencing Factors of Melasma. *SCRJ*. 2021 Oct 25;9(10):26–8. <https://doi.org/10.31364/SCRJ/v9.i10.2021.P1021884>.
30. KrupaShankar DS, Somani VK, Kohli M et al. A cross-sectional, multicentric clinico-epidemiological study of melasma in India. *Dermatol Ther (Heidelb)*. 2014 Jun;4(1):71–81. PMID: 24643868; <https://doi.org/10.1007/s13555-014-0046-1>.
31. Majid I, Aleem S. Melasma: Update on Epidemiology, Clinical Presentation, Assessment, and Scoring. *J Skin Stem Cell*. 2022 Jan 25;8(4). <https://doi.org/10.5812/jssc.120283>.
32. Guinot C, Cheffai S, Latreille J et al. Aggravating factors for melasma: a prospective study in 197 Tunisian patients. *Acad Dermatol Venereol*. 2010 Sep;24(9):1060–9. PMID: 20202051; <https://doi.org/10.1111/j.1468-3083.2010.03592.x>.

Authors' contributions: Toktaş D: conceptualization, investigation, methodology data curation, writing – original draft, formal analysis, writing – review and editing; Atilla R: conceptualization, investigation, methodology, supervision, writing – original draft, formal analysis, writing – review and editing. Both authors reviewed and approved the final version of the manuscript for publication.

Acknowledgments: We thank the women who participated in this study.

Notice: This article is derived from master's dissertation entitled *Gebe Kadınlarda Melazma Görülme Sıklığı ve Beden Algısı İle İlişkisi*, supervised by Rabia Atilla at Niğde Ömer Halisdemir University, Niğde, Türkiye, 2025.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: Data supporting the findings of this study are available upon reasonable request from the corresponding author, Rabia Atilla.

Declaration of generative AI in scientific writing: During the preparation of this work, the authors did not use any generative AI or AI-assisted technologies.

Date of first submission: February 17, 2026

Last received: March 24, 2026

Accepted: March 27, 2026

Address for correspondence:

Rabia Atilla
Niğde Ömer Halisdemir Üniversitesi, Zübeyde Hanım Sağlık Bilimleri
Fakültesi, Hemşirelik Bölümü, Derbent Yerleşkesi
Aşağı Kayabaşı Mahallesi; Atatürk Bulvarı
Merkez — Niğde — Türkiye
Tel. (+90) 0388 211 28 65
E-mail: sert_rabia@hotmail.com; rsert@ohu.edu.tr

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



Treatment patterns and clinical outcomes in chronic myeloid leukemia treated with second-line nilotinib—association of molecular milestones and prognostic scores with deep molecular response: a real-world retrospective cohort study

Irena Čojbašić^I, Ivan Tijanić^{II}, Žarko Čojbašić^{III}

Clinic of Hematology, Allergology and Clinical Immunology, University Clinical Centre Niš, Niš, Serbia

^IPhD; MD. Physician. Associate Professor, Faculty of Medicine, University of Niš, Niš, Serbia.

 <https://orcid.org/0000-0001-6684-1249>

^{II}PhD; MD. Physician. Assistant Professor, Faculty of Medicine, University of Niš, Niš, Serbia.

 <https://orcid.org/0000-0002-6411-4352>

^{III}PhD; Full Professor, Faculty of Mechanical Engineering, University of Niš, Niš, Serbia.

 <https://orcid.org/0000-0002-4581-1048>

KEY WORDS (MeSH terms):

Leukemia, myelogenous, chronic, BCR-ABL positive.

Nilotinib.

Prognosis.

Risk assessment.

Treatment outcome.

AUTHORS' KEY WORDS:

Deep molecular remission.

European LeukemiaNet milestones.

Treatment-free remission.

ABSTRACT

BACKGROUND: Real-world data on the treatment patterns and clinical outcomes of second-line nilotinib use in chronic myeloid leukemia (CML) remain limited, particularly regarding predictors of deep molecular response (DMR).

OBJECTIVES: To characterize treatment patterns and long-term clinical outcomes in patients with CML receiving second-line nilotinib and to evaluate whether early molecular milestones combined with baseline prognostic scores predict DMR.

DESIGN AND SETTING: Retrospective, single-center real-world study conducted at a tertiary academic center between 2011 and 2025.

METHODS: The data of 45 patients with CML receiving second-line nilotinib because of imatinib resistance or intolerance were analyzed. Treatment patterns, including duration and reasons for switching therapy, were assessed. Baseline prognostic scores (Sokal, European Treatment and Outcome Study [EUTOS], EUTOS Long-Term Survival [ELTS], Hammersmith) were calculated, and molecular responses were assessed according to the European LeukemiaNet (ELN) criteria. Predictors of DMR were examined using univariate and multivariate logistic regression, and survival outcomes were estimated using Kaplan–Meier analysis.

RESULTS: The median duration of nilotinib administration was 67.3 months, with 60% of patients achieving a best response of DMR and 46.7% maintaining a sustained DMR. Univariate analysis showed that a favorable Hammersmith score ($p = 0.001$), low ELTS risk ($p = 0.028$), and attainment of ELN-defined milestones at 6 ($p = 0.020$) and 12 months ($p = 0.006$) were significantly associated with DMR. Multivariate analysis demonstrated that the 12-month ELN milestone ($p = 0.008$) and favorable Hammersmith score ($p = 0.023$) were independently associated with a DMR. Patients achieving a DMR showed a nonsignificant trend toward improved overall survival.

CONCLUSION: In this real-world cohort, a DMR and favorable survival outcomes were achieved when nilotinib was administered as second-line treatment. Integrating baseline prognostic scores with molecular milestones enables individualized monitoring and supports a sustained DMR as a marker of durable disease control.

INTRODUCTION

Chronic myeloid leukemia (CML) is a clonal myeloproliferative neoplasm driven by the constitutively active *BCR::ABL1* (breakpoint cluster region::Abelson 1) tyrosine kinase generated by the Philadelphia chromosome, t(9;22)(q34;q11).¹ Quantitative monitoring of *BCR::ABL1* transcripts is central to managing CML, enabling early identification of patients at risk of sub-optimal response and informing therapeutic decision-making.² High rates of cytogenetic and molecular responses are achieved with the use of the first-generation tyrosine kinase inhibitor (TKI) imatinib and second-generation TKIs, including nilotinib, dasatinib, and bosutinib.^{3–5}

Despite these advances, a subset of patients discontinues first-line therapy due to resistance or intolerance and requires second-line treatment.^{3–5} Treatment failure is frequently associated with *BCR::ABL1* kinase domain mutations and other resistance mechanisms, whereas intolerance reflects clinically significant treatment-related toxicities necessitating that therapy be modified or

switched.⁵ Both clinical trials and real-world studies have demonstrated that second-line treatment with nilotinib induces durable molecular responses and favorable survival outcomes.^{6–8}

Although *BCR::ABL1* transcript kinetics are well established for response monitoring, predictors of deep molecular response (DMR) in the second-line setting remain incompletely defined. DMR reflects profound and sustained suppression of *BCR::ABL1* transcripts; it is associated with improved long-term outcomes, such as progression-free survival (PFS) and the potential for treatment-free remission, confirming its clinical relevance as a therapeutic endpoint.^{6–8} European LeukemiaNet (ELN)-defined molecular milestones at 6 and 12 months provide dynamic response assessment, and baseline prognostic scoring systems—including the Sokal, European Treatment and Outcome Study (EUTOS), EUTOS Long-Term Survival (ELTS), and Hammersmith systems—offer complementary information on disease risk.^{9–12} However, most studies have evaluated these parameters separately, and data integrating static baseline risk with molecular dynamics during treatment to predict the DMR in patients receiving second-line nilotinib remain limited.

OBJECTIVE

In this study, we aimed to evaluate whether integrating 6- and 12-month ELN-defined molecular milestones with baseline prognostic scores improved prediction of the DMR in a real-world cohort of patients with chronic-phase CML who were undergoing second-line nilotinib treatment, thereby supporting risk-adapted monitoring and individualized clinical decision-making in routine practice.

METHODS

Study design and population

In this retrospective, real-world study, data from a database of patients with chronic-phase CML were analyzed to evaluate treatment patterns and molecular outcomes at a single center between 2011 and 2025. In total, 45 patients received nilotinib as second-line treatment because they were resistant or intolerant to imatinib. The small sample size reflects the rarity of second-line use of nilotinib in routine practice, highlighting the real-world nature of the cohort. Nilotinib was administered at the standard approved dose of 800 mg/day. Primary resistance was defined as failure to achieve an adequate cytogenetic or molecular response within the expected timeframe. Secondary resistance was defined as the loss of a previously achieved response.

Baseline prognostic assessment

Baseline prognostic scores were calculated at diagnosis to stratify patients according to their expected clinical outcomes and

likelihood of responding to TKI therapy. The Sokal score was used to predict overall survival (OS), with patients stratified into low, intermediate, and high risk at diagnosis, categories. Using the EUTOS score to estimate the probability of a cytogenetic response, patients were classified as being at low or high risk. The ELTS score was used to predicted long-term CML-related mortality in patients treated with a TKI.^{9,10} The Hammersmith score was used to evaluate the likelihood of the response to a second-generation TKI based on the prior response to imatinib, Sokal risk, and recurrence of grade 3 or 4 neutropenia; accordingly, patients were stratified into categories of good, intermediate, or poor risk.¹¹

Cytogenetic and molecular monitoring

Conventional cytogenetic analysis was performed on bone marrow samples, with ≥ 20 metaphases analyzed; the cytogenetic response was defined as the proportion of Philadelphia chromosome-positive metaphases among mitotic cells. Peripheral blood samples were collected longitudinally for molecular monitoring in accordance with ELN recommendations. The *BCR::ABL1* transcript levels were quantified using the GeneXpert® system (Cepheid; Sunnyvale, California) via automated RT-qPCR (reverse transcription quantitative polymerase chain reaction). The results were reported using the International Scale, and internal quality controls were applied. This standardized approach enabled early and DMRs to be assessed throughout therapy.

Response evaluation according to the ELN

Recommendations

Cytogenetic and molecular responses were evaluated according to the current ELN recommendations. A complete cytogenetic response (CCyR) was defined as the absence of Philadelphia chromosome-positive metaphases among mitotic cells. A major molecular response (MMR, MR3) was defined as *BCR::ABL1* $\leq 0.1\%$ IS, MR4 as *BCR::ABL1* $\leq 0.01\%$ IS, DMR as MR4 or deeper, and stable DMR as sustained deep molecular remission confirmed in ≥ 2 consecutive molecular assessments.^{13–15} Early molecular response was assessed at 6 and 12 months using ELN-defined milestones, and the best molecular response represented the deepest response achieved during therapy.

Survival endpoints

PFS was defined as the time from initiation of second-line TKI therapy to progression to the accelerated phase or blast crisis, treatment discontinuation due to disease progression, or death from any cause. OS was defined as the time from initiation of second-line TKI treatment to death from any cause.

Statistical analysis

Patient characteristics and treatment outcomes were summarized using descriptive statistics. Continuous variables are reported as medians with ranges, and categorical variables as frequencies and percentages. Univariate logistic regression was used to identify factors associated with achieving a DMR. Given the limited sample size, a parsimonious multivariate logistic regression model that included the 12-month ELN milestone, and ELTS risk and Hammersmith scores was constructed to avoid overfitting and collinearity; results are reported as odds ratios (ORs) with 95% confidence intervals (95% CIs). All statistical tests were two-sided, and $p < 0.05$ was considered statistically significant.

Kaplan–Meier analysis with the log-rank test was used to estimate survival. Confidence intervals were calculated using Greenwood's formula with log–log transformation; in the small subgroups, the upper limits reached 100%. Landmark analyses at 6 and 12 months were used to estimate the 7-year OS by molecular response, excluding patients who died or were censored before each time point to avoid guarantee-time bias. Molecular milestones at 6 and 12 months were analyzed as fixed-time variables within these landmark analyses. In the additional analyses, the best molecular response achieved was considered to capture the overall depth of the response.

All statistical analyses were performed using MATLAB (version 2025b; The MathWorks Inc., Natick, Massachusetts) with the Statistics and Machine Learning Toolbox (version 25.2).

Approval for the study was granted by the Ethics Committee of the University Clinical Centre Niš, Niš, Serbia (approval number 19486/7, dated April 7th, 2025), and the research was performed according to the principles of the Declaration of Helsinki.

RESULTS

Patient characteristics and baseline risk stratification

The cohort consisted of 45 patients with chronic-phase CML who received nilotinib as second-line treatment. The median duration of prior imatinib therapy before switching to nilotinib was 27.3 months (range: 3–138). The main reasons for patients switching to the second-generation TKI were primary (44.5%) and secondary (40%) resistance, with a smaller proportion discontinuing imatinib treatment due to intolerance (15.5%). The baseline demographic and clinical characteristics of the study population are summarized in **Table 1**.

Risk stratification according to the established prognostic scores showed that, based on the ELTS and Hammersmith scores, approximately half the cohort was classified as low risk (51.1%) and as good risk (53.3%), respectively. However, based on the Sokal score, intermediate-risk disease predominated (48.9%). Based on the EUTOS score, most evaluable patients were categorized as low risk (84.5%). The prognostic scores were calculated using the available baseline data.

Cytogenetic and molecular response dynamics during second-line nilotinib therapy

The median duration of nilotinib therapy was 67.3 months (range: 6–150), enabling longitudinal assessment of cytogenetic and molecular responses. The response rates at the predefined time points are summarized in **Table 2**. The number of evaluable

Table 1. Baseline characteristics and prognostic risk distribution of patients with chronic-phase chronic myeloid leukemia receiving nilotinib as second-line treatment

Characteristic	Variable	Number (%) or Median (range)
Age	Median, years	52.2 (20–85)
Sex, n (%)	Male/female	27 (60)/18 (40)
Imatinib therapy	Median, months (range)	27.3 (3–138)
Sokal score	Low	11 (24.5)
	Intermediate	22 (48.9)
	High	10 (22.2)
	Unknown	2 (4.4)
EUTOS score	Low	38 (84.5)
	High	5 (11.1)
	Unknown	2 (4.4)
ELTS score	Low	23 (51.1)
	Intermediate	14 (31.1)
	High	6 (13.3)
	Unknown	2 (4.4)
Hammersmith score	Good	24 (53.3)
	Intermediate	12 (26.7)
	Poor	7 (15.6)
	Unknown	2 (4.4)
Reason for second-generation TKI	Primary resistance	20 (44.5)
	Secondary resistance	18 (40)
	Intolerance	7 (15.5)

Percentages were calculated based on the total study population (N = 45). Unknown indicates missing baseline data. ELTS, EUTOS long-term survival; EUTOS, TKI, tyrosine kinase inhibitor.

Table 2. Cytogenetic and molecular responses during second-line treatment with nilotinib at predefined time points

Therapeutic response	6 Months, n/N (%)	12 Months, n/N (%)	Best response, n/N (%)
Cytogenetic response			
CCyR	37/45 (82.2)	35/41 (85.4)	40/45 (88.9)
Molecular response			
No MMR	14/32 (43.8)	10/38 (26.3)	10/45 (22.2)
MMR (without MR4)	10/32 (31.3)	13/38 (34.2)	8/45 (17.8)
DMR (MR4 or deeper)	8/32 (25)	15/38 (39.5)	27/45 (60)
Stable DMR	NA	NA	21/45 (46.7)

CCyR, complete cytogenetic response; MMR, major molecular response defined as $BCR::ABL1 \leq 0.1\%$ IS; DMR, deep molecular response defined as MR4 or deeper ($BCR::ABL1 \leq 0.01\%$ IS); stable DMR, stable deep molecular response defined as sustained MR4 or deeper confirmed on ≥ 2 consecutive assessments; NA, not applicable.

patients varied across the time points due to missing molecular assessments and treatment discontinuation.

A CCyR was achieved in 82.2% of patients at 6 months and 85.4% of patients at 12 months, reaching a best response during therapy of 88.9%. An MMR or better ($BCR::ABL1 \leq 0.1\%$ IS) was observed in 56.3% of evaluable patients at 6 months; this increased to 73.7% at 12 months, and 77.8% overall. The DMR (MR4 or deeper; $BCR::ABL1 \leq 0.01\%$ IS) showed a progressive accumulation over time, from 25% at 6 months to 39.5% at 12 months. This DMR reached 60%, forming the best molecular response. A stable DMR was documented in 46.7% of patients. Among patients who did not previously achieve a CCyR while undergoing imatinib treatment, 81.8% achieved a CCyR; among those who did not previously achieve an MMR, 59.3% achieved an MMR while undergoing second-line therapy.

Performance of ELN: defined molecular milestones and prediction of DMR

Molecular responses were assessed according to the ELN recommendations. ELN-defined molecular milestones were achieved by 68.8% of evaluable patients at 6 months and 73.7% of patients at 12 months. Patients meeting these milestones showed higher rates of subsequent DMR than those who did not. Specifically, at last follow-up, a DMR was documented in 68.2% of patients who achieved the 6-month milestone versus 30% of those who did not. Similarly, a DMR was observed in 75% of patients who met the 12-month milestone versus 20% of those who did not.

Univariate predictors of DMR

In the univariate logistic regression analysis, achieving the ELN-defined molecular milestones at 6 and 12 months was associated with an increased likelihood of achieving a DMR. Attaining the 12-month ELN milestone significantly increased the odds of achieving a DMR (OR, 12; 95% CI, 2.04–70.45; $p = 0.006$). A favorable Hammersmith score was also significantly associated with a DMR (OR, 11.08; 95% CI, 2.62–46.84; $p = 0.001$). Patients with low ELTS risk had a significantly higher probability of achieving a DMR compared with those at intermediate or high risk (OR, 0.24; 95% CI, 0.06–0.86; $p = 0.028$). Early cytogenetic and molecular responses at 6 months similarly correlated with subsequent DMRs. The wide confidence intervals observed for several estimates reflect the limited sample size and should be interpreted with caution.

Multivariable analysis of predictors of DMR

A multivariate logistic regression model was constructed to identify independent predictors of DMR, including the 12-month ELN milestone, ELTS risk score, and Hammersmith score. Given the limited number of events, the model was restricted to these

key variables to avoid overfitting. Achievement of the 12-month ELN milestone remained independently associated with a DMR (OR, 34.6; 95% CI, 2.56–467.67; $p = 0.008$), and a favorable Hammersmith score also independently predicted a DMR (OR, 14; 95% CI, 1.43–137.1; $p = 0.023$). By contrast, the ELTS risk category did not retain independent significance in the multivariate analysis, indicating that its predictive value was largely captured by the dynamic molecular response and Hammersmith score.

Molecular responses, prognostic scores, and survival outcomes

The OS and PFS were analyzed as long-term clinical endpoints. At a median follow-up of 67.3 months, the 7-year OS and PFS for the entire cohort were 80.5% and 68.7%, respectively. Sex did not significantly influence survival, with a 7-year OS of 77.1% observed in males versus that of 94.4% observed in females ($p = 0.148$); the respective PFSs observed were 62.9% versus 77.5% ($p = 0.116$). Age at diagnosis was associated with OS: Patients younger than 65 years exhibited superior 7-year OS compared with those 65 years or older (89.8% versus 63%, respectively; $p = 0.002$), whereas PFS did not differ significantly ($p = 0.107$).

Baseline prognostic scores demonstrated differential performance. The Sokal score was associated with PFS (7-year PFS: low, 100%; intermediate, 56%; and high, 50.8%; $p = 0.021$) but not OS ($p = 0.056$), whereas the ELTS score discriminated both OS (7-year OS: low, 95%; intermediate, 78.8%; and high, 50%; $p = 0.005$) and PFS (7-year PFS: low, 78.3%; intermediate, 67.5%; and high, 33.3%; $p = 0.003$) (Figures 1 and 2).

Landmark analyses at 6 and 12 months showed no statistically significant differences in OS according to molecular response at isolated time points. At 6 months, the 7-year OS was 74.5% in patients who did not achieve a MMR, 80% in those who achieved an MMR without DMR, and 100% in those who achieved a DMR ($p = 0.268$). At 12 months, the corresponding OS rates were 75%, 82.5%, and 92.3% ($p = 0.582$) (Figure 3).

The analysis based on the best molecular response achieved during therapy showed numerically different survival outcomes. Patients failing to achieve an MMR had a 7-year OS of 59.1%, those achieving an MMR but not DMR had an OS of 71.4%, and patients achieving a DMR had an OS of 95.2% ($p = 0.068$) (Figure 4). Although differences did not reach statistical significance, consistent numerical trends were observed across the response categories.

DISCUSSION

Nilotinib, a second-generation TKI, selectively inhibits $BCR::ABL1$ and demonstrates greater potency than imatinib, effectively overcoming most imatinib-resistant $BCR::ABL1$ mutations.¹⁶ Clinical studies have shown that nilotinib is efficacious and generally well-tolerated in patients with imatinib-resistant

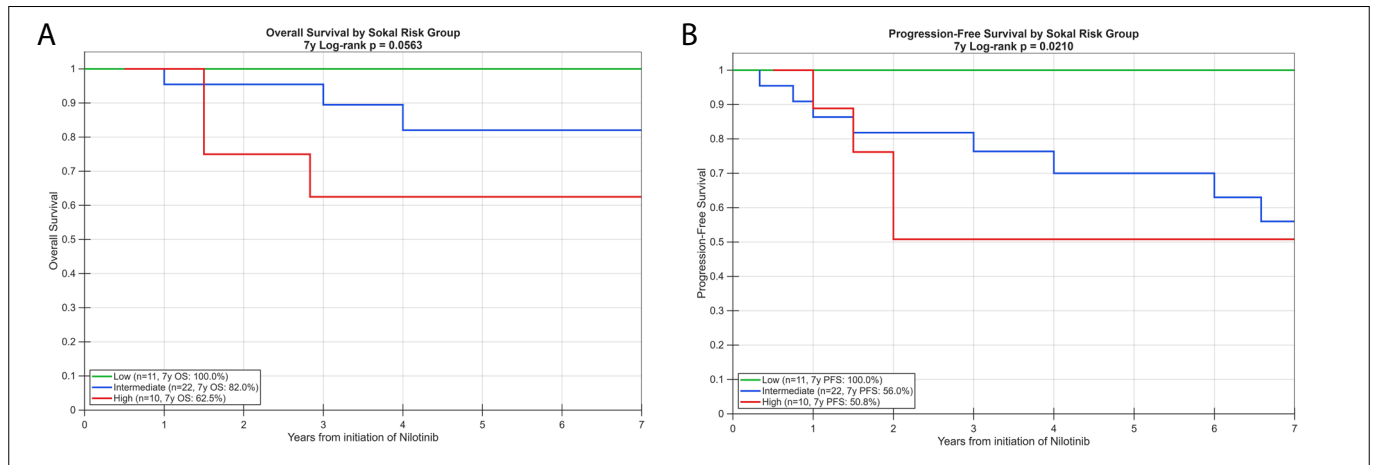


Figure 1. Overall and progression-free survivals according to the Sokal risk category. (a) The Kaplan–Meier curve for overall survival (OS) stratified by the Sokal risk category. (b) The Kaplan–Meier curve for progression-free survival (PFS) stratified by the Sokal risk category.

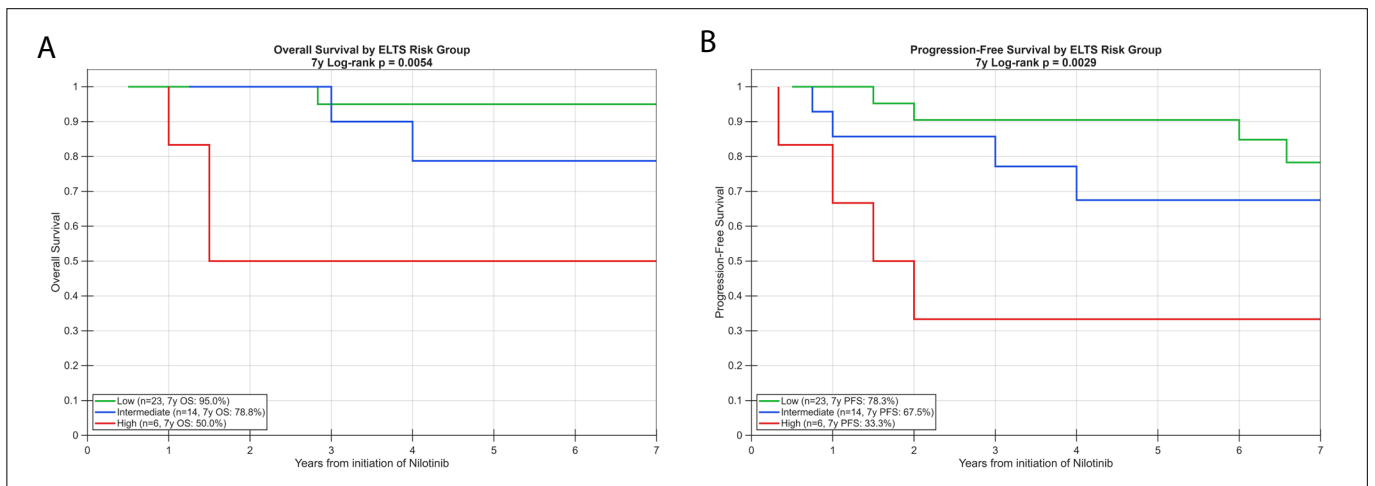


Figure 2. Overall and progression-free survivals according to ELTS risk category. (a) The Kaplan–Meier curve for overall survival (OS) stratified by ELTS risk category. (b) The Kaplan–Meier curve for progression-free survival (PFS) stratified by ELTS risk category.

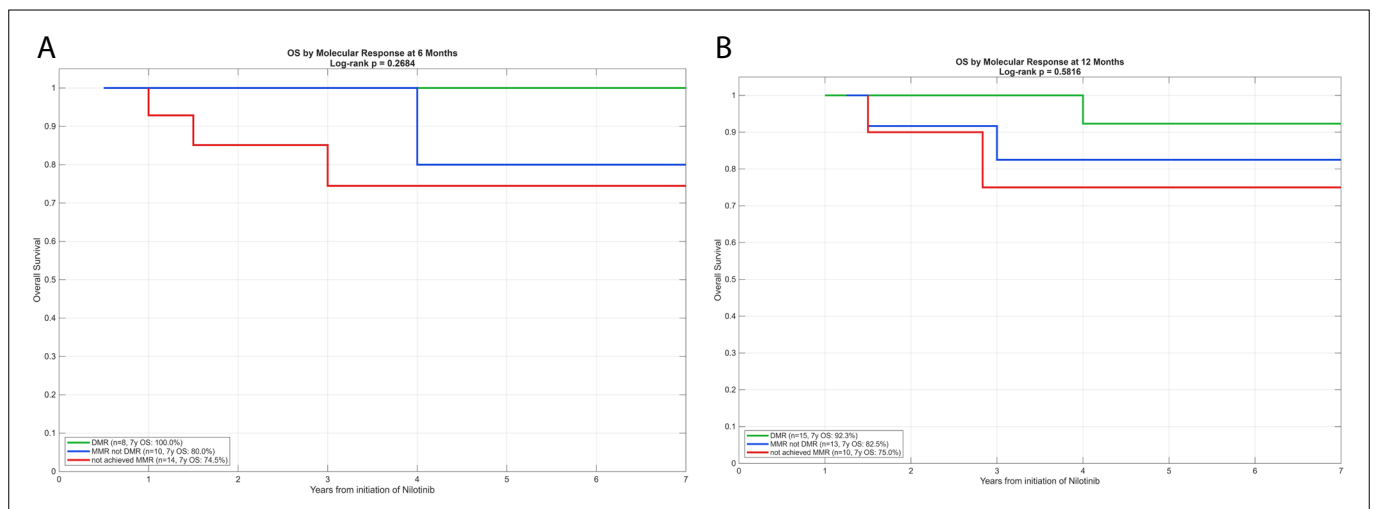


Figure 3. Landmark overall survival according to molecular response category. (a) The Kaplan–Meier curve for the 7-year overall survival (OS) stratified by molecular response (no major molecular response [MMR], MMR without deep molecular response [DMR], and DMR) at 6 months. (b) The Kaplan–Meier curve for the 7-year OS stratified by molecular response (no MMR, MMR without DMR, and DMR) at 12 months.

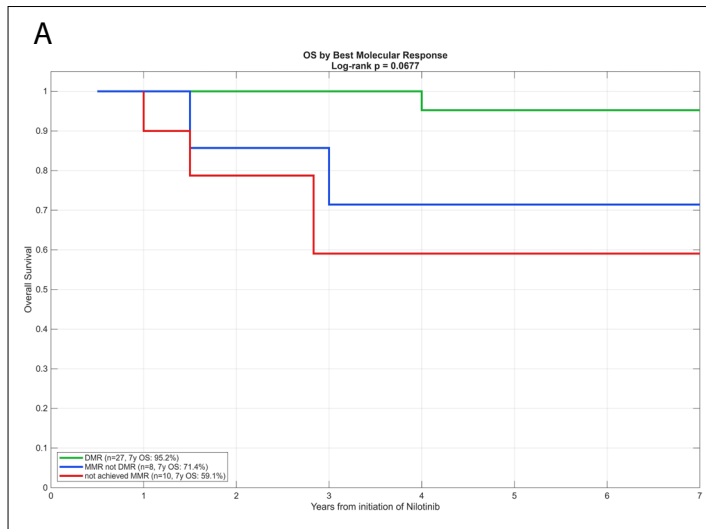


Figure 4. Overall survival (OS) according to best molecular response achieved during therapy. The Kaplan–Meier curve for the 7-year OS stratified by best molecular response category (no major molecular response [MMR], MMR without deep molecular response [DMR], and DMR).

or -intolerant disease, including those with resistance mechanisms independent of *BCR::ABL1* mutations.¹⁶ In newly diagnosed chronic-phase CML, nilotinib has been shown to reduce the emergence of resistant clones, suggesting a potential clinical advantage over imatinib.¹⁷ In this real-world cohort of patients with chronic-phase CML receiving second-line nilotinib treatment after treatment with imatinib failed, early ELN-defined molecular milestones and selected baseline prognostic scores were independently associated with subsequent DMRs. Real-world data integrating early molecular response with baseline prognostic scores to predict DMR remain limited, particularly for second-line use of nilotinib.

In the present study, longitudinal quantitative *BCR::ABL1* monitoring demonstrated that early attainment of ELN-defined molecular milestones at 6 and 12 months was strongly associated with subsequently achieving a DMR. Notably, the 12-month ELN milestone remained independently associated with DMR in the multivariable analysis, emphasizing the prognostic relevance of the dynamic assessment of response during treatment. Patients with prior imatinib resistance, including secondary resistance, achieved clinically meaningful rates of CCyR, MMR, and DMR, confirming the efficacy of second-line nilotinib use. In controlled clinical studies, MMR rates of approximately 73% and DMR rates of up to 34% were achieved within 2 years with the use of nilotinib as a second-line treatment, whereas real-world analyses reported cumulative MMR and DMR rates of approximately 65% and 40%, respectively, albeit with greater heterogeneity in the early response kinetics.^{18–22} The DMR rates observed in our cohort (60% best, 46.7% stable) align with these real-world data, reinforcing the effectiveness

of second-line nilotinib treatment and supporting the use of early molecular milestones as predictive markers for deep remission.

In the modern era of multiple *BCR::ABL1* TKIs, DMRs and treatment-free remission have emerged as key therapeutic end-points.²³ However, most evidence on early *BCR::ABL1* transcript decline as a predictor of DMR is from frontline therapy, with limited data specifically addressing second-line use of nilotinib. Large real-world studies formally evaluating 6- or 12-month ELN milestones in this context remain lacking. Real-world data highlight the prognostic value of early molecular kinetics: Achieving MR4.5 was more likely with second-line use of nilotinib than dasatinib after adjustment for baseline characteristics.²⁴ Furthermore, *BCR::ABL1* levels at 3 and 6 months independently predicted MR4, with higher rates of early molecular response being achieved with nilotinib use than with imatinib use.²⁵ Early transcript dynamics within the first 3 to 6 months predicted subsequent DMR,²⁶ whereas residual disease at 6 months was the strongest early predictor among the molecular parameters for deeper responses of MR4 or MR4.5.²⁷ Real-world DMR rates with nilotinib generally remain lower than those reported in clinical trials, emphasizing the importance of contextualizing efficacy in routine practice.²⁸ Collectively, these studies reinforce that early *BCR::ABL1* transcript kinetics during treatment provide robust predictive information for DMRs in second-line nilotinib therapy.

The results of the landmark analyses at isolated time points did not reach statistical significance for OS; however, the numerical gradients observed across the molecular response categories suggest that the cumulative depth of response may reflect long-term disease control. These observations should be interpreted cautiously due to the limited sample size. This aligns with prior studies highlighting that sustained DMR provides clinically meaningful information regarding durable disease control and potential treatment-free remission.^{21–23} Although ELTS risk stratification predicted DMR in the univariate analyses, its significance was attenuated in the multivariate models once molecular response during treatment and prior treatment sensitivity were included, confirming the dominant role of longitudinal molecular monitoring over baseline risk alone. Importantly, the ability of the ELTS score to discriminate overall and progression-free survivals was retained; however, the lack of independent significance for DMR observed for this score in the multivariable analysis suggests that dynamic molecular response during treatment provides more robust predictive information than baseline risk alone.

The integration of baseline prognostic scores with molecular response enabled additional stratification of patients who were at higher risk for suboptimal outcomes. Favorable Hammersmith scores were independently associated with DMR, consistent with baseline response-related factors, whereas dynamic kinetics during treatment were captured through ELN-defined molecular milestones; together, they provide complementary prognostic information.

These observations are supported by prior studies in which baseline risk scores, particularly those for the ELTS, were shown to predict both survival and molecular outcomes in CML^{29–31}; this highlights that integrating baseline and dynamic variables refines risk stratification and guides individualized monitoring.

From a clinical perspective, these results support the importance of achieving and maintaining DMRs in second-line nilotinib therapy. Our findings confirm the relevance of sustained DMR for long-term disease control and for the identification of patients who may be candidates for treatment-free remission. Integrating longitudinal molecular monitoring with baseline prognostic assessment and prior treatment sensitivity may help clinicians tailor monitoring intensity and identify patients most likely to achieve durable deep remission, thereby supporting individualized management in routine practice.

The limitations of this study include its retrospective, single-center design and modest sample size, which may limit statistical power and generalizability. Data on *BCR::ABL1* kinase domain mutations were not systematically available and thus were not included in the analysis. Future studies incorporating kinase domain mutation data are warranted to further refine predictors of DMR in second-line nilotinib therapy. The confidence intervals were wide for some predictors, reflecting the limited number of events. The survival analyses based on best molecular response may have been influenced by immortal time bias and should be interpreted cautiously. Nevertheless, by integrating baseline and dynamic variables, the findings of this study support a precision-oriented management approach to enhancing prognostic stratification and individualized monitoring even within a modest, single-center cohort.

CONCLUSION

In this real-world cohort of patients with chronic-phase CML treated with second-line nilotinib, early *BCR::ABL1* kinetics combined with baseline prognostic scores provided complementary information for predicting DMR. Achievement of 12-month ELN milestones and a favorable Hammersmith score independently identified those patients most likely to reach a DMR, a key indicator of durable long-term disease control. Integrating longitudinal molecular monitoring with baseline risk enables clinicians to personalize follow-up and guide management in routine practice, reinforcing sustained DMR as a marker of durable disease control.

REFERENCES

- Hochhaus A, Saussele S, Rosti G, et al. Chronic myeloid leukaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017 Jul 1;28(suppl_4):iv41–51. <https://doi.org/10.1093/annonc/mdx219>.
- Hochhaus A, Saglio G, Hughes TP, et al. Long-term benefits and risks of frontline nilotinib vs imatinib for chronic myeloid leukemia in chronic phase: 5-year update of the randomized ENESTnd trial. *Leukemia*. 2016 May;30(5):1044–54. <https://doi.org/10.1038/leu.2016.5>.
- Hochhaus A, Larson RA, Guilhot F, et al. Long-term outcomes of imatinib treatment for chronic myeloid leukemia. *N Engl J Med*. 2017 Mar 9;376(10):917–27. <https://doi.org/10.1056/NEJMoa1609324>.
- Hehlmann R, Lauseker M, Sauße S, et al. Assessment of imatinib as first-line treatment of chronic myeloid leukemia: 10-year survival results of the randomized CML study IV and impact of non-CML determinants. *Leukemia*. 2017 Nov;31(11):2398–406. <https://doi.org/10.1038/leu.2017.253>.
- Soverini S, Hochhaus A, Nicolini FE, et al. BCR-ABL kinase domain mutation analysis in chronic myeloid leukemia patients treated with tyrosine kinase inhibitors: recommendations from an expert panel on behalf of European LeukemiaNet. *Blood*. 2011 Aug 4;118(5):1208–15. <https://doi.org/10.1182/blood-2010-12-326405>.
- Giles FJ, Le Coutre PD, Pinilla-Ibarz J, et al. Nilotinib in imatinib-resistant or imatinib-intolerant patients with chronic myeloid leukemia in chronic phase: 48-month follow-up results of a phase II study. *Leukemia*. 2013 Jan;27(1):107–12. <https://doi.org/10.1038/leu.2012.181>.
- Kantarjian HM, Giles FJ, Bhatta KN, et al. Nilotinib is effective in patients with chronic myeloid leukemia in chronic phase after imatinib resistance or intolerance: 24-month follow-up results. *Blood*. 2011 Jan 27;117(4):1141–5. <https://doi.org/10.1182/blood-2010-03-277152>.
- Kuo CY, Wang PN, Hwang WL, et al. Safety and efficacy of nilotinib in routine clinical practice in patients with chronic myeloid leukemia in chronic or accelerated phase with resistance or intolerance to imatinib: results from the NOVEL study. *Ther Adv Hematol*. 2018 Mar;9(3):65–78. <https://doi.org/10.1177/2040620718756603>.
- Pfirschmann M, Baccarani M, Saussele S, et al. Prognosis of long-term survival considering disease-specific death in patients with chronic myeloid leukemia. *Leukemia*. 2016 Jan;30(1):48–56. <https://doi.org/10.1038/leu.2015.261>.
- Apperley JF, Milojkovic D, Cross NCP, et al. 2025 European LeukemiaNet recommendations for the management of chronic myeloid leukemia. *Leukemia*. 2025 Aug;39(8):1797–813. <https://doi.org/10.1038/s41375-025-02664-w>.
- Milojkovic D, Nicholson E, Apperley JF, et al. Early prediction of success or failure of treatment with second-generation tyrosine kinase inhibitors in patients with chronic myeloid leukemia. *Haematologica*. 2010 Feb 1;95(2):224–31. <https://doi.org/10.3324/haematol.2009.012781>.
- Baccarani M, Deininger MW, Rosti G, et al. European LeukemiaNet recommendations for the management of chronic myeloid leukemia: 2013. *Blood*. 2013 Aug 8;122(6):872–84. <https://doi.org/10.1182/blood-2013-05-501569>.
- Hochhaus A, Baccarani M, Silver RT, et al. European LeukemiaNet 2020 recommendations for treating chronic myeloid leukemia. *Leukemia*. 2020 Apr;34(4):966–84. <https://doi.org/10.1038/s41375-020-0776-2>.
- Cross NCP, Ernst T, Branford S, et al. European LeukemiaNet laboratory recommendations for the diagnosis and management of chronic

- myeloid leukemia. *Leukemia*. 2023 Nov;37(11):2150–67. <https://doi.org/10.1038/s41375-023-02048-y>.
15. Senapati J, Sasaki K, Issa GC, et al. Management of chronic myeloid leukemia in 2023 – common ground and common sense. *Blood Cancer J*. 2023 Apr 24;13(1):58. <https://doi.org/10.1038/s41408-023-00823-9>.
 16. Kantarjian HM, Giles F, Gattermann N, et al. Nilotinib (formerly AMN107), a highly selective BCR-ABL tyrosine kinase inhibitor, is effective in patients with Philadelphia chromosome–positive chronic myelogenous leukemia in chronic phase following imatinib resistance and intolerance. *Blood*. 2007 Nov 15;110(10):3540–6. <https://doi.org/10.1182/blood-2007-03-080689>.
 17. Hochhaus A, Rosti G, Cross NCP, et al. Frontline nilotinib in patients with chronic myeloid leukemia in chronic phase: results from the European ENEST 1st study. *Leukemia*. 2016 Jan;30(1):57–64. <https://doi.org/10.1038/leu.2015.270>.
 18. Yeung DT, Osborn MP, White DL, et al. TIDEL-II: first-line use of imatinib in CML with early switch to nilotinib for failure to achieve time-dependent molecular targets. *Blood*. 2015 Feb 5;125(6):915–23. <https://doi.org/10.1182/blood-2014-07-590315>.
 19. Hughes TP, Leber B, Cervantes F, et al. Sustained deep molecular responses in patients switched to nilotinib due to persistent BCR-ABL1 on imatinib: final ENESTcmr randomized trial results. *Leukemia*. 2017 Nov;31(11):2529–31. <https://doi.org/10.1038/leu.2017.247>.
 20. Tiribelli M, Bonifacio M, Binotto G, et al. Excellent outcomes of 2G-TKI therapy after imatinib failure in chronic phase CML patients. *Oncotarget*. 2018 Mar 6;9(18):14219–27. <https://doi.org/10.18632/oncotarget.24478>.
 21. Scalzulli E, Caocci G, Efficace F, et al. Real-life comparison of nilotinib versus dasatinib as second-line therapy in chronic phase chronic myeloid leukemia patients. *Ann Hematol*. 2021 May;100(5):1213–9. <https://doi.org/10.1007/s00277-021-04477-0>.
 22. Breccia M, Cucci R, Marsili G, et al. Deep molecular response rate in chronic phase chronic myeloid leukemia: eligibility to discontinuation related to time to response and different frontline TKI in the experience of the Gimema Labnet CML National Network. *Clin Lymphoma Myeloma Leuk*. 2025 Jan;25(1):e34–9. <https://doi.org/10.1016/j.clml.2024.08.009>.
 23. Kantarjian H, Branford S, Breccia M, et al. Are there new relevant therapeutic endpoints in the modern era of the BCR::ABL1 tyrosine kinase inhibitors in chronic myeloid leukemia? *Leukemia*. 2024 May;38(5):947–50. <https://doi.org/10.1038/s41375-024-02229-3>.
 24. Cortes J, Huynh L, Mendelson E, et al. Treatment patterns and deep molecular response in chronic phase — chronic myeloid leukemia patients treated with second-line nilotinib or dasatinib: a multi-country retrospective chart review study. *Leuk Lymphoma*. 2020 Jan 2;61(1):98–107. <https://doi.org/10.1080/10428194.2019.1644332>.
 25. Wang R, Cong Y, Li C, Zhang C, Lin H. Predictive value of early molecular response for deep molecular response in chronic phase of chronic myeloid leukemia. *Medicine*. 2019 Apr;98(15):e15222. <https://doi.org/10.1097/MD.00000000000015222>.
 26. Okamoto Y, Hirano M, Morino K, et al. Early dynamics of chronic myeloid leukemia on nilotinib predicts deep molecular response. *NPJ Syst Biol Appl*. 2022 Oct 13;8(1):39. <https://doi.org/10.1038/s41540-022-00248-3>.
 27. Saugues S, Lambert C, Daguene E, et al. The initial molecular response predicts the deep molecular response but not treatment-free remission maintenance in a real-world chronic myeloid leukemia cohort. *Haematologica*. 2024 May 2;109(9):2893–2907. <https://doi.org/10.3324/haematol.2023.284860>.
 28. Collins JB, Muluneh B, Proffitt S. Evaluation of real-world versus clinical trial outcomes of tyrosine kinase inhibitor therapy for chronic myeloid leukemia. *J Oncol Pharm Pract*. 2024 Mar;30(2):385–96. <https://doi.org/10.1177/10781552231217694>.
 29. Specchia G, Pregno P, Breccia M, et al. Prognostic factors for overall survival in chronic myeloid leukemia patients: a multicentric cohort study by the Italian CML GIMEMA Network. *Front Oncol*. 2021 Aug 26;11:739171. <https://doi.org/10.3389/fonc.2021.739171>.
 30. Pffirmann M, Clark RE, Prejzner W, et al. The EUTOS long-term survival (ELTS) score is superior to the Sokal score for predicting survival in chronic myeloid leukemia. *Leukemia*. 2020 Aug;34(8):2138–49. <https://doi.org/10.1038/s41375-020-0931-9>.
 31. Sato E, Iriyama N, Tokuhira M, et al. The EUTOS long-term survival score predicts disease-specific mortality and molecular responses among patients with chronic myeloid leukemia in a practice-based cohort. *Cancer Med*. 2020 Dec;9(23):8931–9. <https://doi.org/10.1002/cam4.3516>.

Authors' contributions: Čojbašić I: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, supervision, validation, writing – original draft, writing – review and editing; Tijanić I: data curation, formal analysis, investigation, methodology, visualization, writing – review and editing; Čojbašić Ž: formal analysis, methodology, validation, writing – review and editing; supervision, statistical analysis. All authors reviewed and approved the final version of the manuscript for publication.

Sources of funding: This work was supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia [grant numbers 451-03-136/2025-03/200113 and 451-03-137/2025-03/200109].

Conflicts of interest: None.

Data availability statement: Due to ethical and privacy restrictions, the raw data is not publicly available. The datasets generated and analyzed during the current study are available from the corresponding author, Irena Čojbašić, on reasonable request.

Declaration of generative AI in scientific writing: During the preparation of this work, the authors did not use any generative AI or AI-assisted technologies.

Date of first submission: February 19, 2026

Last received: May 12, 2026

Accepted: May 6, 2026

Address for correspondence:

Irena Čojbašić
Klinika za hematologiju, alergologiju i kliničku imunologiju, Univerzitetski
klinički centar Niš
Bulevar Zorana Đinđića 48
Medicinski fakultet, Univerzitet u Nišu
Bulevar Zorana Đinđića 81
Niš — Nišavski okrug — Srbija
Telefon: (+381 63) 1045736
E-mail: icojbasic@gmail.com

Editors responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



The effect of BIS-guided sevoflurane anesthesia on postoperative nausea and vomiting in pediatric adenoidectomy: a prospective randomized controlled trial

Engin Cetin^I, Ayse Sencan^{II}, Ahmet Yuksek^{III}, Bedirhan Gunel^{IV}, Fatihhan Zeytun^V, Mehmet Yilmaz^{VI}

Kocaeli City Hospital, Kocaeli, Türkiye

^IMD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0000-0002-3011-3968>

^{II}MD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0000-0003-2225-5269>

^{III}MD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0000-0002-7529-2971>

^{IV}MD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0000-0003-2292-9403>

^VMD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0009-0009-6818-2908>

^{VI}MD. Department of Anesthesiology and Reanimation, University of Health Sciences Kocaeli City Hospital, Izmit, Kocaeli, Türkiye.
 ID <https://orcid.org/0000-0002-5353-9996>

KEYWORDS (MeSH terms):

Monitoring, Intraoperative.
 Postoperative Nausea and Vomiting.
 Pediatric Anesthesia.
 Sevoflurane.
 Adenoidectomy.

AUTHOR'S KEYWORDS:

BIS-Guided Anesthesia.
 Pediatric Adenoidectomy.
 Postoperative Recovery.
 Emergence Time.
 Volatile Anesthetic Consumption.

ABSTRACT

BACKGROUND: Postoperative nausea and vomiting (PONV) is a frequent complication in pediatric patients undergoing adenoidectomy. Bispectral index (BIS) monitoring, an electroencephalography-based method for assessing anesthetic depth, enables individualized sevoflurane titration and may reduce unnecessary exposure to volatile anesthetics.

OBJECTIVES: To determine whether BIS-guided sevoflurane anesthesia reduces PONV in children undergoing adenoidectomy.

DESIGN AND SETTING: Single-center, prospective, randomized, assessor-blinded trial conducted at University of Health Sciences, Kocaeli City Hospital, Kocaeli, Türkiye.

METHODS: Fifty-eight children aged 3–8 years (American Society of Anesthesiologists score I–II) undergoing elective adenoidectomy were randomized to BIS-guided anesthesia (n = 29) or standard care (n = 29). Both groups received sevoflurane anesthesia. The primary outcome was PONV within the first 24 postoperative hours. Secondary outcomes were emergence time, total sevoflurane consumption, remifentanyl dose, and rescue antiemetic use.

RESULTS: PONV within 24 hours occurred less frequently in the BIS-guided group than in controls (p = 0.015), with the greatest between-group differences observed at 0–1 hour and 1–6 hours. Sevoflurane consumption and emergence time were lower in the BIS-guided group (p < 0.05). BIS-guided anesthesia was independently associated with lower odds of PONV (OR = 0.212; 95% CI: 0.067–0.675; p = 0.009).

CONCLUSIONS: BIS monitoring may be a useful strategy for optimizing anesthetic depth and reducing early PONV in pediatric patients.

CLINICAL TRIAL REGISTRATION: ClinicalTrials.gov (NCT07189845). Available at: <https://clinicaltrials.gov/study/NCT07189845>.

INTRODUCTION

Postoperative nausea and vomiting (PONV) is among the most common postoperative complications, with an overall incidence of approximately 30%.¹ In pediatric patients, particularly following otorhinolaryngologic surgeries, this incidence may reach as high as 73%.² Complications associated with PONV include dehydration, electrolyte imbalance, aspiration, and postoperative bleeding.² In addition, PONV is associated with lower patient satisfaction and prolonged length of stay in the post-anesthesia care unit (PACU).^{1,3} Adenotonsillectomy and strabismus surgery have been reported to be the surgical procedures most frequently associated with PONV in children.^{4,5} Other factors associated with the development of PONV include surgical duration exceeding 30 minutes, a positive family history, the use of inhalational anesthetics, nitrous oxide, and opioids.^{4–7}

The pharmacological effects of general anesthetics in children vary according to age and body surface area,⁸ which may lead to excessive or insufficient anesthetic doses.⁸ During inhalational anesthesia, anesthetic depth is commonly monitored using the minimum alveolar concentration (MAC).⁹ However, MAC-based monitoring may not accurately reflect the hypnotic component of anesthesia.⁹ Monitoring anesthetic depth using the bispectral index (BIS) allows more precise and individualized titration of anesthetic agents.¹⁰ BIS-guided anesthesia improves postoperative outcomes and reduces anesthetic drug consumption.^{10,11}

OBJECTIVES

The aim of this study was to evaluate the effect of BIS-guided sevoflurane anesthesia on PONV in pediatric patients undergoing adenoidectomy. We hypothesized that BIS-guided sevoflurane anesthesia would reduce the incidence of PONV within the first 24 postoperative hours.

METHODS

Study design and ethics

This study was a single-center, prospective, randomized, assessor-blinded clinical trial conducted at University of Health Sciences, Kocaeli City Hospital, in accordance with the principles of the Declaration of Helsinki and Good Clinical Practice. The design, conduct, and reporting of the study complied with the CONSORT 2025 (Consolidated Standards of Reporting Trials) guidelines.¹²

The study protocol was approved by the Kocaeli University Clinical Research Ethics Committee (Approval No. KAEK/12.bI.06; date of approval: July 17, 2025). Written informed consent was obtained from the parents or legal guardians of all participants. The trial was prospectively registered in the ClinicalTrials.gov database on September 2, 2025 (ClinicalTrials.gov Identifier: NCT07189845).

Participants

Pediatric patients aged 3–8 years with American Society of Anesthesiologists (ASA) physical status I–II who were scheduled for elective adenoidectomy were included in the study. The exclusion criteria were a history of PONV, preoperative use of antiemetics, steroids, or sedative medications, acute upper respiratory tract infection, history of neurological or psychiatric disorders, body mass index below the 5th percentile or above the 95th percentile for age, and anatomical airway abnormalities.

Patient flow, randomization, and blinding

A total of 62 patients were assessed for eligibility. Four patients were excluded before randomization because intravenous access could not be established during induction or because deviation from the standard anesthesia induction protocol was required. These protocol deviations resulted in exclusion before randomization.

The remaining 58 patients were randomly assigned in a 1:1 ratio to two groups using a computer-generated block randomization method with a block size of four. Accordingly, 29 patients were allocated to the BIS-guided group and 29 patients to the control group. Group allocation was concealed using sequentially numbered, opaque, sealed envelopes prepared by an independent researcher not involved in the study process (Figure 1). Patients and investigators were blinded to group allocation. Blinding of the attending anesthesiologists was not feasible due to the nature of the BIS-guided anesthesia protocol, which may have introduced performance bias.

Anesthesia protocol

Patients were transferred to the operating room after intravenous access had been established, and anesthesia induction was initiated intravenously. Intravenous fentanyl (1 µg/kg) was administered for analgesia, and anesthesia induction was achieved with propofol (2 mg/kg). Rocuronium (0.6 mg/kg) was administered to facilitate neuromuscular blockade and tracheal intubation.

Following endotracheal intubation, anesthesia was maintained with sevoflurane in an air/oxygen mixture. Mechanical ventilation was applied in pressure-controlled ventilation mode, with tidal volumes set at 6–8 mL/kg, respiratory rate adjusted to maintain end-tidal CO₂ between 35 and 40 mmHg, and fresh gas flow set at 2 L/min. Intraoperative analgesia was provided with a remifentanyl infusion titrated according to hemodynamic responses.

In the BIS-guided group, the sevoflurane concentration was titrated during maintenance of anesthesia to maintain a BIS value between 45 and 60. If BIS values exceeded the target range (> 60), the sevoflurane concentration was increased incrementally, whereas values below the target range (< 45) prompted a gradual reduction in sevoflurane concentration.

In the control group, the sevoflurane concentration was adjusted according to standard clinical practice to maintain age-adjusted MAC values between 1 and 1.3 and to keep hemodynamic parameters (heart rate and blood pressure) within $\pm 20\%$ of baseline values. When MAC or hemodynamic targets were exceeded, anesthetic depth was adjusted by modifying the sevoflurane concentration.

Measurement of anesthetic consumption

Total sevoflurane consumption (mL) during surgery was obtained from anesthesia machine records, and total remifentanyl dose

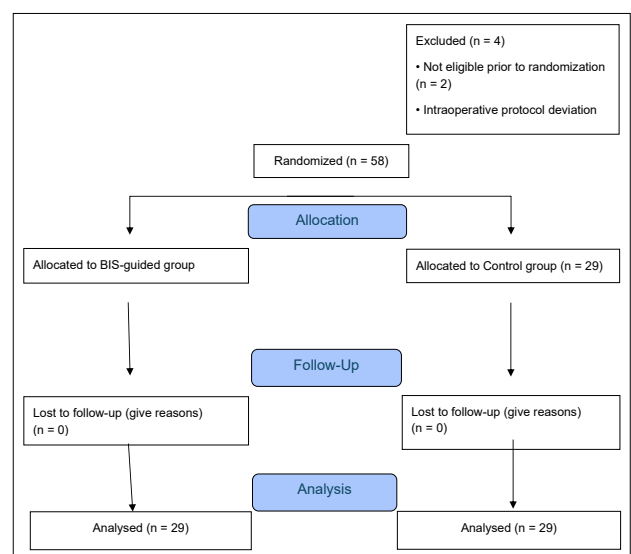


Figure 1. CONSORT flow diagram of patient enrollment and follow-up.

($\mu\text{g/kg}$) was recorded from the infusion pump and anesthesia charts. These variables were analyzed as potential mediators of the effect of BIS-guided anesthesia on PONV.

Standard intraoperative monitoring included continuous assessment of heart rate, noninvasive blood pressure, peripheral oxygen saturation, end-tidal CO_2 , and body temperature throughout the procedure. BIS monitoring was performed using a BIS Vista™ monitor (Medtronic, Minneapolis). Anesthesia was delivered using an anesthesia workstation (Aisys™ Carestation, GE Healthcare, Madison).

Emergence time

Emergence time was defined as the interval between discontinuation of sevoflurane at the end of surgery and eye opening in response to verbal commands. Neuromuscular blockade was reversed at the end of surgery with intravenous sugammadex (2 mg/kg). Emergence time was measured using a stopwatch and recorded in minutes. In cases where eye opening could not be assessed, limb movement or the onset of adequate spontaneous respiration was used as the emergence criterion.

Outcome measures

Primary outcome

The primary outcome was the incidence of PONV within the first 24 postoperative hours.

Secondary outcomes

Secondary outcomes included emergence time, total sevoflurane consumption, total remifentanyl dose, and the requirement for antiemetic therapy within the first 24 postoperative hours.

Assessment of PONV

PONV was assessed at predefined time intervals after extubation. In the PACU, assessments were performed during the first 0–1 hour after extubation. On the ward, PONV assessments were conducted during the 1–6 hour, 6–12 hour, and 12–24 hour postoperative periods.

The severity of PONV was evaluated using a predefined four-point ordinal scoring system: score 0 indicated no nausea or vomiting; score 1 indicated mild nausea; score 2 indicated severe nausea or a single episode of vomiting; and score 3 indicated severe nausea accompanied by multiple episodes of vomiting or refractory PONV. Assessments were based on clinical observation and nursing reports.¹³ Rescue antiemetic treatment was administered when patients experienced vomiting or when the PONV score was ≥ 2 .

The time points were defined as follows: the first postoperative hour in the PACU (T1), 6 hours after extubation (T2), 12 hours after extubation (T3), and 24 hours after extubation (T4). Secondary outcomes included emergence time, total sevoflurane consumption,

total remifentanyl dose, and antiemetic use and timing. In younger children who were unable to report nausea verbally, PONV assessment was based on observable clinical signs such as retching, vomiting, increased salivation, feeding refusal, and behavioral distress, as well as nursing observations and parental reports.

Antiemetic use and timing

Rescue antiemetic treatments administered within the first 24 postoperative hours were recorded. Rescue antiemetic administration followed predefined criteria (vomiting or a PONV score ≥ 2). Ondansetron was used as the first-line rescue antiemetic at an intravenous dose of 0.1 mg/kg (maximum dose 4 mg), in accordance with institutional clinical practice. The type of rescue antiemetic agent, timing of administration, and total number of doses were obtained from anesthesia and nursing records. Outcomes were reported as the proportion of patients requiring rescue antiemetic therapy.

Sample size calculation

Sample size calculation was based on a previous randomized controlled study evaluating the effect of BIS-guided anesthesia on PONV incidence in children, which reported PONV rates of 53% in the control group and 17% in the BIS group. With a 95% confidence level ($\alpha = 0.05$) and 80% power ($1 - \beta = 0.80$), a minimum of 52 patients was required. Considering a potential dropout rate of 10%, the target sample size was set at 58 patients.¹⁴

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp., Armonk, New York). Descriptive statistics were presented as mean $\pm$ standard deviation, median (minimum–maximum), and frequencies and percentages (n, %). Data distribution was assessed using the Kolmogorov–Smirnov test.

For normally distributed independent continuous variables, Student's t-test was used, and results are expressed as mean $\pm$ standard deviation. For non-normally distributed independent continuous variables, the Mann–Whitney U test was applied, and results are reported as median (minimum–maximum). Independent categorical variables were analyzed using Pearson's chi-square test with Yates' continuity correction, and categorical data are presented as frequencies and percentages. A p value < 0.05 was considered statistically significant.

To further evaluate the primary outcome of 24-hour PONV incidence, binary logistic regression analysis was performed to identify independent predictors of PONV. Initially, univariable logistic regression analyses were conducted for potential risk factors including age, height, weight, sex, patient group, total doses of propofol, rocuronium, fentanyl, and remifentanyl, total sevoflurane consumption, mean BIS values, ASA scores, and duration

of surgery. Variables with a p value < 0.20 in univariable analysis were entered into the multiple model. The final model was determined using the backward likelihood ratio method, in which only patient group remained significant; results are presented in Table 5. A p value < 0.05 was considered statistically significant.¹⁵

To reduce the risk of type I error due to multiple comparisons for secondary outcomes, Bonferroni correction was applied, and statistical significance was set at $p < 0.00625$ ($0.05/8$). All randomized patients were analyzed in the groups to which they were originally assigned (intention-to-treat analysis).

RESULTS

A total of 58 patients were included in the study, and analyses compared the BIS-guided group ($n = 29$) with the control group ($n = 29$). Baseline demographic and clinical characteristics, including age, sex, body weight, height, and ASA physical status classification, are presented in Table 1. No significant differences were observed between the groups for these variables.

No statistically significant differences were observed between the groups in terms of duration of surgery, mean heart rate, end-tidal CO_2 , mean arterial pressure, or total doses of propofol, rocuronium, and fentanyl ($p > 0.05$ for all). The mean BIS value was significantly higher in the BIS-guided group than in the control group ($p < 0.001$) (Table 2).

PONV within the first 24 postoperative hours was observed in 13 patients (44.8%) in the BIS-guided group and 23 patients (79.3%) in the control group. The difference between the groups was statistically significant (Relative risk [RR] = 0.565; 95% CI: 0.362–0.881; $p = 0.015$) (Table 3).

Time-interval analyses showed that, during the early postoperative period (0–1 hour and 1–6 hours), PONV incidence was significantly lower in the BIS-guided group than in the control group (for both intervals: RR = 0.294; 95% CI: 0.125–0.691; $p = 0.003$). No significant differences were observed between the groups during the later postoperative periods (6–12 hours and 12–24 hours) ($p > 0.00625$, Bonferroni-adjusted) (Table 3).

Emergence time was significantly shorter in the BIS-guided group than in the control group ($p < 0.001$). Total sevoflurane consumption was also significantly lower in the BIS-guided group ($p < 0.001$). Remifentanyl consumption showed no significant difference between the groups ($p = 0.726$). Although the requirement for antiemetic therapy was lower in the BIS-guided group, the difference did not reach statistical significance after Bonferroni correction ($p = 0.015$) (Table 4).

In the multivariable logistic regression analysis, BIS-guided anesthesia was identified as an independent protective factor against the development of PONV (OR = 0.212; 95% CI: 0.067–0.675; $p = 0.009$) (Table 5).

Harms

No serious adverse events were observed during the study period.

DISCUSSION

This randomized controlled trial demonstrated that, compared with a non-BIS-guided approach, BIS-guided anesthesia

Table 2. Perioperative data

	BIS-guided group	Control group	p value*
Duration of surgery (min)	53.3 $\pm$ 4.7	53.6 $\pm$ 4.6	0.822 [†]
SpO ₂ (%)	99 (97–100)	98 (97–100)	0.048 ^m
Heart rate (bpm)	110 (85–125)	110 (88–124)	0.75 ^m
ETCO ₂	36 (33–41)	38 (33–41)	0.106 ^m
Mean arterial pressure (mmHg)	81 (72–92)	84 (73–92)	0.87 ^m
Propofol dose (mg)	37.8 $\pm$ 8.8	40.5 $\pm$ 7.9	0.228 [†]
Rocuronium dose (mg)	10 (5–15)	10 (9–16)	0.636 ^m
Fentanyl dose (mcg)	20 (10–25)	20 (14–27)	0.132 ^m
Mean BIS value	57 (48–60)	50 (40–54)	< 0.001^{m*}

Data are presented as number (%), mean $\pm$ standard deviation or median (minimum–maximum). [†] Student's t-test. ^m Mann–Whitney U test. * p value < 0.05 was considered statistically significant.

Table 1. Baseline characteristics

	BIS-guided group	Control group
Age (years)	5 (3–8)	6 (3–8)
Weight (kg)	19.3 $\pm$ 4.3	20 $\pm$ 3.9
Height (cm)	109 (92–130)	115 (90–130)
History of PONV	0 (0)	0 (0)
Sex		
Male	13 (44.8)	11 (37.9)
Female	16 (55.2)	18 (62.1)
ASA physical status		
ASA I	26 (89.7)	24 (82.8)
ASA II	3 (10.3)	5 (17.2)

Data are presented as mean $\pm$ standard deviation or median (minimum–maximum) for continuous variables and as number (%) for categorical variables. No statistical tests were performed for baseline characteristics.

Table 3. Incidence of Postoperative Nausea and Vomiting (PONV) according to time intervals

	BIS-guided group	Control group	Relative Risk (95% CI)	p value*
0–1 h PONV	5 (17.2)	17 (58.6)	0.294 (0.125–0.691)	0.003^{y†}
1–6 h PONV	5 (17.2)	17 (58.6)	0.294 (0.125–0.691)	0.003^{y†}
6–12 h PONV	6 (20.7)	11 (37.9)	0.546 (0.233–1.278)	0.249 ^y
12–24 h PONV	4 (13.8)	6 (20.7)	0.667 (0.21–2.118)	0.728 ^y
24 h PONV	13 (44.8)	23 (79.3)	0.565 (0.362–0.881)	0.015^{y*}

Data are presented as number (%). Relative risks (RR) with 95% confidence intervals (CI) are reported. ^y Pearson's chi-squared test with Yates's continuity correction. * A p value < 0.05 was considered statistically significant for the primary outcome (24-hour PONV). [†] For secondary outcomes, Bonferroni correction was applied, and a p value < 0.00625 was considered statistically significant.

management was associated with a reduced incidence of PONV within the first 24 postoperative hours. This effect was particularly pronounced during the early postoperative period (0–1 hour and 0–6 hours). In the multivariable analysis, BIS-guided anesthesia was identified as an independent protective factor against the development of PONV. This finding corresponds to an absolute risk reduction of approximately 34.5%, suggesting a potentially clinically meaningful reduction in PONV.

PONV remains one of the most common postoperative adverse events in pediatric patients and is strongly associated with the use of inhalational anesthesia.¹⁴ Apfel et al.¹³ demonstrated that the association between inhalational anesthesia and PONV is especially prominent during the early postoperative period (0–2 hours).

Previous studies on BIS-guided anesthesia in adult populations have reported a reduction in PONV incidence with BIS monitoring, likely due to more precise and controlled titration of inhalational anesthetics.^{16,17} In contrast, a recent pediatric study reported that BIS-guided inhalational anesthesia did not significantly reduce PONV during the very early postoperative period (0–1 hour) but resulted in a marked reduction during the 1–24 hour period.¹⁴

Our findings are generally consistent with the published literature and further support the beneficial effect of BIS monitoring on PONV. Notably, unlike previous pediatric studies, our results show that BIS-guided anesthesia was also associated with a significant reduction in PONV during the immediate early postoperative period (0–1 hour). This difference may be attributable to variations in surgical type, anesthetic management protocols, and postoperative observation periods.

Liao et al.¹⁸ reported no significant difference in PONV incidence between BIS-guided and non-BIS-guided sevoflurane anesthesia in patients undergoing ambulatory surgery. Although sevoflurane consumption was lower in the BIS-guided group, PONV rates were similar between groups. This discrepancy may be explained by the inclusion of ambulatory urological procedures and the limited postoperative follow-up period, which was limited to the first postoperative hour. In contrast, our study focused on patients undergoing adenoidectomy, a population known to be at higher risk for PONV, and included a longer postoperative observation period. Therefore, differences in surgical type and follow-up duration may account for the divergent findings. Taken together, the available evidence

Table 4. Secondary outcomes

	BIS-guided group	Control group	p value
Emergence time (min)	7 (5–9)	11 (7–14)	< 0.001 ^{m*}
Sevoflurane consumption (mL)	20 (15–25)	26 (20–30)	< 0.001 ^{m*}
Remifentanyl consumption (µg/kg)	5.35 ± 0.6	5.30 ± 0.5	0.726 ^t
Antiemetic use	4 (13.8)	22 (75.9)	0.015 ^y

Data are presented as number (%) or median (minimum–maximum).

^y Pearson's chi-squared test with Yates's continuity correction. ^m Mann–Whitney U test. ^t Student's t-test. * For secondary outcomes, Bonferroni correction was applied, and a p value < 0.00625 (0.05/8) was considered statistically significant.

Table 5. Logistic regression analysis for Postoperative Nausea and Vomiting (PONV)

	PONV		Total	Univariable		Multiple	
	No	Yes		OR (95% CI)	p	OR (95% CI)	p value
Groups							
BIS-guided group	16 (72.7)	13 (36.1)	29 (50)	0.212 (0.067 – 0.675)	0.009*	0.212 (0.067 – 0.675)	0.009*
Control group	6 (27.3)	23 (63.9)	29 (50)	Reference			
Age (years)	5.1 ± 1.7	5.8 ± 1.6	5.5 ± 1.6	1.317 (0.939 – 1.849)	0.111	–	–
Height (cm)	110.4 ± 11.7	113.7 ± 10.5	112.4 ± 11	1.028 (0.978 – 1.08)	0.274	–	–
Weight (kg)	19.3 ± 4.2	19.9 ± 4.1	19.6 ± 4.1	1.036 (0.909 – 1.181)	0.595	–	–
Sex							
Male	7 (31.8)	17 (47.2)	24 (41.4)	1.917 (0.632 – 5.82)	0.251	–	–
Female	15 (68.2)	19 (52.8)	34 (58.6)	Reference			
ASA physical status							
ASA I	18 (81.8)	32 (88.9)	50 (86.2)	0.996 (0.887 – 1.119)	0.946	–	–
ASA II	4 (18.2)	4 (11.1)	8 (13.8)	Reference			
Propofol dose (mg)	38.3 ± 8.7	39.7 ± 8.4	39.2 ± 8.4	1.021 (0.958 – 1.088)	0.523	–	–
Rocuronium dose (mg)	10.9 ± 2.6	11.7 ± 2.4	11.4 ± 2.5	1.14 (0.915 – 1.42)	0.242	–	–
Fentanyl dose (µg)	18.1 ± 4.2	19.2 ± 4	18.8 ± 4.1	1.065 (0.933 – 1.217)	0.352	–	–
Sevoflurane consumption (mL)	21.7 ± 3.3	23.7 ± 3.5	22.9 ± 3.6	1.185 (1.005 – 1.397)	0.044*	–	–
Remifentanyl dose (µg/kg)	5.36 ± 0.6	5.31 ± 0.6	5.33 ± 0.6	0.886 (0.366 – 2.231)	0.766	–	–
Mean BIS value	54.4 ± 4.8	52.4 ± 4	53.1 ± 4.4	0.893 (0.783 – 1.019)	0.094	–	–
Duration of surgery (min)	53.5 ± 5.4	53.4 ± 4.1	53.4 ± 4.6	0.996 (0.887 – 1.119)	0.946	–	–
Constant	–	–	–	–	–	1.765 (0–0)	0.055*

Cox and Snell R Square = 0.122; Nagelkerke R Square = 0.166; Accuracy = 0.672. Data are presented as number (%) and mean ± standard deviation. * p < 0.05 was considered statistically significant.

suggests that the effect of BIS-guided anesthesia on PONV may vary according to surgical population, anesthetic protocol, postoperative observation duration, and the use of prophylactic antiemetics.

The incidence of PONV is proportional to both the duration and total dose of inhalational anesthetics administered.¹⁹ In the present study, surgical durations were comparable between groups; however, total sevoflurane consumption was significantly lower in the BIS-guided group. Another notable finding was the significantly shorter emergence time in the BIS-guided group. These secondary outcomes are consistent with findings from previous meta-analyses and comparable studies, which have demonstrated that BIS-guided anesthesia can reduce anesthetic consumption and accelerate recovery.^{8,9,16}

In a large cohort study, Portnoy et al.⁵ evaluated the incidence of PONV, prophylactic strategies, and adherence to clinical guidelines in pediatric patients. The authors reported lower PONV rates than those commonly cited in the literature, largely attributable to the routine use of antiemetic prophylaxis. In our study, routine antiemetic prophylaxis was intentionally omitted to specifically evaluate the isolated effect of BIS-guided anesthesia on PONV. The observed PONV incidence was 44% in the BIS-guided group and 79% in the non-BIS-guided group. The relatively higher PONV incidence in the BIS-guided group compared with that in the study by Portnoy et al.⁵ may be explained by the absence of routine pharmacological prophylaxis. Although the requirement for rescue antiemetic therapy was numerically lower in the BIS-guided group, this difference did not remain statistically significant after Bonferroni correction.²⁰ Nevertheless, the observed numerical reduction may still be clinically relevant and should be interpreted cautiously given the relatively small sample size.

Taken together, our findings and the published literature suggest that BIS-guided anesthesia may represent a valuable strategy for PONV prevention. The results of this study, obtained without routine antiemetic prophylaxis, show that BIS-guided sevoflurane anesthesia significantly reduces the incidence of PONV in pediatric patients undergoing adenoidectomy. BIS-guided anesthesia may therefore be considered a useful adjunctive strategy for reducing PONV in pediatric surgeries associated with a high risk of PONV. Future multicenter studies with larger sample sizes and standardized antiemetic prophylaxis protocols are warranted to further validate these findings.

Limitations

This study has certain limitations. First, it was conducted at a single center with a relatively small sample size, which may reduce statistical power and limit the generalizability of the findings to broader pediatric surgical populations. In particular, the relatively small sample size may have increased the risk of type II error for certain secondary outcomes, including rescue antiemetic use. Second, the study population was restricted to pediatric patients undergoing adenoidectomy; therefore, the results may not be applicable to other surgical procedures or broader

pediatric age groups. Third, blinding of the attending anesthesiologists was not feasible because of the BIS-guided anesthesia protocol, which may have introduced performance bias. Fourth, the absence of routine antiemetic prophylaxis may limit the applicability of the findings to routine clinical practice, although this approach allowed clearer evaluation of the isolated effect of BIS-guided anesthesia on PONV. Finally, PONV assessment was based partly on clinical observation, nursing assessments, and parental reports, particularly in younger children unable to verbalize nausea reliably; therefore, some degree of subjective assessment bias and interobserver variability cannot be excluded.

CONCLUSIONS

This prospective randomized controlled trial demonstrated that BIS-guided sevoflurane anesthesia was associated with a reduced incidence of PONV within the first 24 postoperative hours in pediatric patients undergoing adenoidectomy. BIS monitoring was also associated with lower sevoflurane consumption and shorter emergence time. These findings suggest that BIS-guided anesthesia may help optimize anesthetic titration and improve early postoperative recovery in pediatric patients at high risk for PONV. However, larger multicenter studies are needed to confirm these findings and evaluate their applicability in routine clinical practice.

REFERENCES

1. Jin Z, Gan TJ, Bergese SD. Prevention and Treatment of Postoperative Nausea and Vomiting (PONV): A Review of Current Recommendations and Emerging Therapies. *TCRM*. 2020 Dec;Volume 16:1305–17. <https://doi.org/10.2147/TCRM.S256234>.
2. Frelich M, Burša F, Jor O, et al. Effect of second-hand smoking on the incidence of postoperative nausea and vomiting in children after adenoidectomy: a single-centre retrospective study. *J Anesth Analg Crit Care*. 2025 Oct 17;5(1):67. <https://doi.org/10.1186/s44158-025-00284-5>.
3. Liang D, Shan Y, Wang L. The effect of prophylactic rewarming on postoperative nausea and vomiting among patients undergoing laparoscopic hysterectomy: a prospective randomized clinical study. *Sao Paulo Med J*. 2020 Oct;138(5):414–21. <https://doi.org/10.1590/1516-3180.2020.0059.r2.06072020>.
4. Kovac AL. Postoperative Nausea and Vomiting in Pediatric Patients. *Pediatr Drugs*. 2021 Jan;23(1):11–37. <https://doi.org/10.1007/s40272-020-00424-0>.
5. Portnoy Y, Glebov M, Orkin D, Katsin M, Berkenstadt H. Postoperative Nausea and Vomiting in Pediatrics: Incidence and Guideline Adherence—a Retrospective Cohort Study. *Anesthesia & Analgesia*. 2025 Jul;141(1):77–85. <https://doi.org/10.1213/ANE.00000000000007291>.
6. Gan TJ, Belani KG, Bergese S, et al. Fourth Consensus Guidelines for the Management of Postoperative Nausea and Vomiting. *Anesthesia & Analgesia*. 2020 Aug;131(2):411–48. <https://doi.org/10.1213/ANE.0000000000004833>.

7. Talih G, Yüksek A, Şahin E. Evaluation of emergence agitation after general anaesthesia in rhinoplasty patients: Inhalation anaesthesia versus total intravenous anaesthesia. *American Journal of Otolaryngology*. 2020 May;41(3):102387. <https://doi.org/10.1016/j.amjoto.2020.102387>.
8. Oliveira DAC, Santos Monteiro J, Macêdo NR, Tavares Costa Santos AH, Rosa Neves SM. Bispectral index-guided anesthesia in children: A systematic review and meta-analysis. *Anaesthesia Critical Care & Pain Medicine*. 2025 Nov;44(6):101607. <https://doi.org/10.1016/j.accpm.2025.101607>.
9. Templeton TW, Alex G, Eloy JD, Stollings L, et al. BIS Guided Titration of Sevoflurane in Pediatric Patients Undergoing Elective Surgery: A Randomized Controlled Trial. *Pediatric Anesthesia*. 2025 Apr;35(4):277–86. <https://doi.org/10.1111/pan.15057>.
10. Gu Y, Hao J, Wang J, et al. Effectiveness Assessment of Bispectral Index Monitoring Compared with Conventional Monitoring in General Anesthesia: A Systematic Review and Meta-Analysis. *Monaco F, editor. Anesthesiology Research and Practice*. 2024 Jan;2024(1):5555481. <https://doi.org/10.1155/2024/5555481>.
11. Sargin M, Uluer MS, Şimşek B. The effect of bispectral index monitoring on cognitive performance following sedation for outpatient colonoscopy: a randomized controlled trial. *Sao Paulo Med J*. 2019 Jul;137(4):305–11. <https://doi.org/10.1590/1516-3180.2018.0383210519>.
12. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. *BMJ*. 2025 Apr 14;389:e081123. <https://doi.org/10.1136/bmj-2024-081123>.
13. Apfel CC, Stoocklein K, Lipfert P. PONV: A problem of inhalational anaesthesia? *Best Practice & Research Clinical Anaesthesiology*. 2005 Sep;19(3):485–500. <https://doi.org/10.1016/j.bpa.2005.03.001>.
14. Frelich M, Sklienka P, Romanová T, et al. The effect of BIS-guided anaesthesia on the incidence of postoperative nausea and vomiting in children: a prospective randomized double-blind study. *BMC Anesthesiol*. 2024 Jul 9;24(1):228. <https://doi.org/10.1186/s12871-024-02610-w>.
15. Gunel B, Sencan A, Cevik T, Islamoglu Z, Yilmaz M. Comparison of high flow nasal cannula versus procedural oxygen mask for hypoxemia prevention during endoscopic retrograde cholangiopancreatography: a randomized parallel-group trial. *Minerva Anesthesiol*. 2026 Feb;92(1–2). <https://doi.org/10.23736/S0375-9393.25.19322-X>.
16. Oliveira CRD, Bernardo WM, Nunes VM. Benefit of general anesthesia monitored by bispectral index compared with monitoring guided only by clinical parameters. *Systematic review and meta-analysis. Brazilian Journal of Anesthesiology (English Edition)*. 2017 Jan;67(1):72–84. <https://doi.org/10.1016/j.bjane.2015.09.001>.
17. Nelskylä KA, Yli-Hankala AM, Puro PH, Korttila KT. Sevoflurane Titration Using Bispectral Index Decreases Postoperative Vomiting in Phase II Recovery After Ambulatory Surgery. *Anesthesia & Analgesia*. 2001 Nov;93(5):1165–9. <https://doi.org/10.1097/0000539-200111000-00021>.
18. Liao WW, Wang JJ, Wu GJ, Kuo CD. The effect of cerebral monitoring on recovery after sevoflurane anesthesia in ambulatory setting in children: A comparison among bispectral index, A-line autoregressive index, and standard practice. *Journal of the Chinese Medical Association*. 2011 Jan;74(1):28–36. <https://doi.org/10.1016/j.jcma.2011.01.004>.
19. Bannister CF, Brosius KK, Sigl JC, Meyer BJ, Sebel PS. The Effect of Bispectral Index Monitoring on Anesthetic Use and Recovery in Children Anesthetized with Sevoflurane in Nitrous Oxide. *Anesthesia & Analgesia*. 2001 Apr;92(4):877–81. <https://doi.org/10.1097/0000539-200104000-00015>.
20. Talih G, Yüksek A. Septorinoplasti Hastalarında Postoperatif Analjezik Ajan Tercihleri: Retrospektif Analiz. *Ankara Eğitim ve Araştırma Hastanesi Tıp Dergisi*. 2021 Apr 27;54(1):7–11. <https://doi.org/10.20492/aeahtd.754077>.

Authors' contributions: Cetin E: conceptualization, methodology, investigation, formal analysis, writing – original draft, writing – review and editing, supervision, project administration; Sencan A: investigation, data curation, writing – review and editing; Yuksek A: methodology, data curation; Gunel B: methodology, data curation, formal analysis, writing – original draft, writing – review and editing; Zeytun F: investigation; Yilmaz M: formal analysis. All authors reviewed and approved the final version of the manuscript for publication.

Sources of funding: None.

Conflicts of interest: None.

Data availability statement: The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request. The data are not publicly available because of privacy and ethical restrictions concerning pediatric patient information.

Declaration of generative AI in scientific writing: ChatGPT (OpenAI) were used exclusively for language editing and translation purposes and did not influence the scientific content.

Date of first submission: February 11, 2026

Last received: April 4, 2026

Accepted: May 13, 2026

Address for correspondence:

Engin Çetin
Sağlık Bilimleri Üniversitesi, Kocaeli Şehir Hastanesi, Anesteziyoloji ve Reanimasyon Kliniği
Tavşantepe Mahallesi, Akif Sokak, 63
İzmit — Kocaeli — Türkiye
PK. 41060
Tel.: (+90) 532 120 6004
E-mail: doccetin52@gmail.com

Editor responsible for the evaluation process:

Marianne Yumi Nakai, MD, PhD (AE)
Paulo Manuel Pêgo-Fernandes, MD, PhD (EIC)



INSTRUCTIONS FOR AUTHORS

Scope and indexing

São Paulo Medical Journal (formerly Revista Paulista de Medicina) was founded in 1932 and is published bimonthly by Associação Paulista de Medicina, a regional medical association in Brazil.

The Journal accepts articles in English in the fields of evidence-based health, including internal medicine, epidemiology and public health, specialized medicine (gynecology & obstetrics, mental health, surgery, pediatrics, urology, neurology and many others), and also physical therapy, speech therapy, psychology, nursing and healthcare management/administration.

São Paulo Medical Journal's articles are indexed in MEDLINE, LILACS, SciELO, Science Citation Index Expanded, Journal Citation Reports/Science Edition (ISI) and EBSCO Publishing.

Editorial policy

Papers with a commercial objective will not be accepted: please review the Journal's conflicts of interest policy below.

São Paulo Medical Journal accepts manuscripts previously deposited in a trusted preprint server.

São Paulo Medical Journal supports Open Science practices. It invites reviewers to join Open Peer Review practices through acceptance that their identities can be revealed to the authors of articles. However, this is purely an invitation: reviewers may also continue to provide their input anonymously.

São Paulo Medical Journal is an open-access publication. This means that it publishes full texts online with free access for readers.

São Paulo Medical Journal applies a publication fee in the form of an article processing charge (APC) for all studies conducted outside of Brazil. This rate will be charged to the corresponding author when the study has been accepted on the grounds of its scientific merit. This fee is US\$ 500.00 and is independent of the length of the text. The corresponding author should wait to receive the journal's invoice before making the payment. The article will only be published after presentation of the proof of payment. Submission is free for all. Associação Paulista de Medicina provides financial support for the Journal.

Articles accepted for publication become the Journal's property for copyright purposes, in accordance with Creative Commons attribution type BY.

Transparency and integrity: guidelines for writing

The Journal recommends that all articles submitted should comply with the editorial quality standards established in the Uniform Requirements for Manuscripts Submitted to Biomedical Journals,¹ as updated in the Recommendations for the Conduct, Reporting, Editing and Publication of Scholarly Work in Medical Journals. These standards were created and published by the International

Committee of Medical Journal Editors (ICMJE) as a step towards integrity and transparency in science reporting and they were updated in December 2018.¹

All studies published in *São Paulo Medical Journal* must be described in accordance with the specific guidelines for papers reporting on clinical trials (CONSORT),² systematic reviews and meta-analyses (PRISMA),^{3,4} observational studies (STROBE),^{5,6} case reports (CARE)⁷ and accuracy studies on diagnostic tests (STARD).^{8,9} These guidelines ensure that all methodological procedures have been described, and that no result has been omitted. If none of the above reporting guidelines are adequate for the study design, authors are encouraged to visit the EQUATOR Network website (<http://www.equator-network.org/>) to search for appropriate tools.

Conflicts of interest

Authors are required to describe any conflicts of interest that may exist regarding the research or the publication of the article. Failure to disclose any conflicts of interest is a form of misconduct.

Conflicts of interest may be financial or non-financial. The Journal recommends that the item "Conflicts of interest" at <http://www.icmje.org> should be read to obtain clarifications regarding what may or may not be considered to be a conflict of interest. The existence and declaration of conflicts of interest is not an impediment to publication at all.

Acknowledgements and funding

Grants, bursaries and any other financial support for studies must be mentioned separately, after the references, in a section named "Acknowledgements." Any financial support should be acknowledged, always with the funding agency name, and with the protocol number whenever possible. Donation of materials used in the research can and should be acknowledged too.

This section should also be used to acknowledge any other contributions from individuals or professionals who have helped in producing or reviewing the study, and whose contributions to the publication do not constitute authorship.

Authorship

The Journal supports the position taken by the ICMJE (<http://www.icmje.org>) regarding authorship. All authors should read ICMJE's recommendations to obtain clarifications regarding the criteria for authorship and to verify whether all of them have made enough contributions to be considered authors.¹⁰

All authors of articles published in *São Paulo Medical Journal* need to have contributed actively to the discussion of the study results and should review and approve the final version that is to be released. If one author has not contributed enough or has not approved the final version of the manuscript, he/she must be transferred to the Acknowledgement section.

The corresponding author is the primary guarantor of all ethical issues relating to the manuscript, before, during and after its publication. However, *São Paulo Medical Journal* and ICMJE consider that all authors are held fully responsible for the study, regarding the accuracy or integrity of data and data interpretation in the text. Contributions such as data collection only do not constitute authorship.

The addition or deletion of authors' names in the manuscript byline is possible only if the corresponding author provides the reason for the rearrangement and a written signed agreement from all authors. Modifications to the order of the authors are possible, but also need to be justified. Authors whose names are removed or inserted must agree with this in writing. Publication of the article cannot proceed without a declaration of authorship contributions signed by all authors.

São Paulo Medical Journal supports the ORCID initiative. All authors should create an ORCID identification (ID) record (in www.orcid.org) before submitting their article and should link the submission to their existing ORCID ID in the electronic submission system. ORCID identifications help to distinguish researchers with similar names, give credit to contributors and link authors to their professional affiliations. In addition, this may increase the ability of search engines to retrieve articles.

São Paulo Medical Journal supports Open Science practices. Authors must therefore complete an open science compliance form, which is available from: https://wp.scielo.org/wp-content/uploads/Open-Science-Compliance-Form_en.docx.

Redundant or duplicate publication

São Paulo Medical Journal will avoid publishing redundant or duplicate articles. The Journal agrees with the ICMJE definition of redundant publication,¹¹ i.e. an attempt to report or publish the same results from a study twice. This includes but is not limited to publication of patient cohort data that has already been published, without clear reference to the previous publication. In situations in which authors are making a secondary analysis on data that has already published elsewhere, they must state this clearly. Moreover, the outcomes assessed in each analysis should be clearly differentiated.

The Journal's peer review policy and procedures

After receipt of the article through the electronic submission system, it will be read by the editorial team, who will check whether the text complies with the Journal's Instructions for Authors regarding format. The Journal has adopted the *CrossRef Similarity Check* system for identifying plagiarism and any text that has been plagiarized, in whole or in part, will be promptly rejected. Self-plagiarism will also be monitored.

When the general format of the manuscript is deemed acceptable and fully compliant with these Instructions for Authors, and only then, the editorial team will submit the article to the Editor-in-Chief, who will firstly evaluate its scope. If the editor finds that the topic is of interest for publication, he will assign at least two reviewers/referees

with expertise in the theme, to evaluate the quality of the study. After a period varying from one to several weeks, the authors will then receive the reviewers' evaluations and will be required to provide all further information requested and the corrections that may be necessary for publication. These reviewers, as well as the Editorial Team and the Editor-in-Chief, may also deem the article to be unsuitable for publication by *São Paulo Medical Journal* at this point.

At the time of manuscript submission, the authors will be asked to indicate the names of three to five referees. All of them should be from outside the institution where the authors work and at least two should preferably be from outside Brazil. The Editor-in-Chief is free to choose them to review the paper or to rely on the *São Paulo Medical Journal's* Editorial Board alone.

Articles will be rejected without peer review if:

- they do not present Ethics Committee approval (or a justification for the absence of this);
- they fail to adhere to the format for text and figures described here.

After peer review

Peer reviewers, associated editors and the Editor-in-Chief may ask for clarifications or changes to be made to the manuscript. The authors should then send their article back to the Journal, with the modifications made as requested. Changes to the text should be highlighted (in a different color or using a text editor tool to track changes). Failure to show the changes clearly might result in the paper being returned to the authors.

The modified article must be accompanied by a letter answering the referees' comments, point by point. The modified article and the response letter are presented to the editorial team and reviewers, who will verify whether the problems have been resolved adequately. The text and the reviewers' final evaluations, along with the response letter, will then be sent to the Editor-in-Chief for a decision.

Manuscripts that are found to be suitable for publication through their scientific merit will be considered "provisionally accepted". However, all articles will subsequently be scrutinized to check for any problems regarding the reporting, i.e. sentence construction, spelling, grammar, numerical/statistical problems, bibliographical references and other matters that may arise, especially in the Methods section. The adherence to reporting guidelines will be checked at this point, and the staff will point out any information regarding methodology or results that the authors should provide. This is done in order to ensure transparency and integrity of publication, and to allow reproducibility.

The editorial team will then provide page proofs for the authors to review and approve. No article is published without this final author approval. All authors should review the proof, although the Journal asks the corresponding author to give final approval.

Submission

Articles should be submitted only after they have been formatted as described below. Texts must be submitted exclusively through the Internet, using the Journal's electronic submission system, which

is available at <http://mc04.manuscriptcentral.com/spmj-scielo>. Submissions sent by e-mail or through the post will not be accepted.

The manuscript should be divided into two files. The first of these, the main document ("blinded"), should contain the article title, article type, keywords and abstract, article text, references and tables, but must omit all information about the authors. The second of these, the "title page", should contain all the information about the authors.

To format these documents, use Times New Roman font, font size 12, line spacing 1.5, justified text and numbered pages.

The corresponding author is responsible for the submission. However, all authors should approve the final version of the manuscript that is to be submitted and should be aware of and approve any changes that might be made after peer review.

Covering letter

All manuscripts must be submitted with a covering letter signed at least by the corresponding author. The letter must contain the following five essential items relating to the manuscript:

1. a declaration that the manuscript is original and that the text is not under consideration by any other journal;
2. a statement that the manuscript has been approved by all authors, who agree to cede the copyrights to the Journal, disclose all sources of funding and declare all potential conflicts of interest;
3. a statement that the study protocol was endorsed by an Internal Review Board (Ethics Committee), including the date and number of the approval (in the case of original articles). This is required for absolutely all studies involving human subjects or patient data (such as medical records), in accordance with the Committee on Publication Ethics (COPE) guidelines, and even for case reports. A copy of the approval document must be submitted to the Journal;
4. each author should indicate a valid, up-to-date email address for contact;
5. a list of a minimum of five potential referees outside of the authors' institutions, who could be invited, at the Editor-in-Chief's discretion, to evaluate the manuscript.

General guidelines for original articles

The following are considered to be full-text original articles: clinical trials; cohort, case-control, prevalence, incidence, accuracy and cost-effectiveness studies; case series (i.e. case reports on more than three patients analyzed together); and systematic reviews with or without meta-analysis. These types of article should be written with a maximum of 3,500 words (from the introduction to the end of the conclusion).

Typical main headings in the text include Introduction, Methods, Results, Discussion and Conclusion. The authors can and should use short subheadings too, especially those concerning the reporting guideline items.

Trial and systematic review registration policy

São Paulo Medical Journal supports the clinical trial registration policies of the World Health Organization (WHO) and the International Committee of Medical Journal Editors (ICMJE) and recognizes the importance of these initiatives for registration and international dissemination of information on randomized clinical trials, with open access. Thus, since 2008, manuscripts on clinical trials are accepted for publication if they have received an identification number from one of the public clinical trial registration database (such as ClinicalTrials.gov and/or REBEC and/or the World Health Organization; the options are stated at <http://www.icmje.org>). The identification number should be declared at the end of the abstract. Articles describing systematic reviews must provide the protocol registration number from a reliable database, such as PROSPERO, Open Science Framework, Cochrane, Joanna Briggs and others. Articles presenting clinical trials or systematic reviews without registration protocols will be promptly rejected without peer review.

Results from cases with DNA sequences must be deposited in appropriate public databases. The protocol number or URL can be requested at any time during the editorial review. Publication of other research data in public repositories is also recommended, since it contributes towards replicability of research, increases article visibility and possibly improves access to health information.

Sample size

All studies published in SPMJ must present a description of how the sample size was arrived at. If it was a convenience or purposive sample, the authors must declare so and explain the characteristics of this sample and recruitment method. For clinical trials, for instance, it is mandatory to inform each of the three main values used to calculate sample size:

- power (usually 80% or more);
- level of significance (usually 0.05 or lower);
- clinically meaningful difference (effect size targeted), according to the main outcome measurement.

Regardless of study results (if "positive" or "negative"), the journal will probably reject articles of trials using underpowered samples, when sample size has not been properly calculated or the calculation has not been fully described as indicated above.

Abbreviations, acronyms and products

Abbreviations and acronyms must not be used, even those in everyday use, unless they are defined when first used in the text. However, authors should avoid them for clarity whenever possible. Drugs or medications must be referred to using their generic names (without capital letters), with avoidance of casual mention of commercial or brand names.

Interventions

All drugs, including anesthetics, should be followed by the dosage and posology used.

Any product cited in the Methods section, such as diagnostic or therapeutic equipment, tests, reagents, instruments, utensils, prostheses, orthoses and intraoperative devices, must be described together with the manufacturer's name and place (city and country) of manufacture in parentheses. The version of the software used should be mentioned.

Any other interventions, such as exercises, psychological assessments or educational sessions, should be described in enough details to allow reproducibility. The Journal recommends that the TIDieR reporting guidelines should be used to describe interventions, both in clinical trials and in observational studies.¹³

Supplementary material

Because supplementary material comprises documents that do not form part of the text of the manuscript, *São Paulo Medical Journal* will not publish it. The authors should cite an access link that allows readers to view the supplementary material.

Short communications

Short communications are reports on the results from ongoing studies or studies that have recently been concluded for which urgent publication is important. They should be structured in the same way as original articles. The authors of this kind of communication should explain, in the covering letter, why they believe that publication is urgent. Short communications and case reports must be limited to 1,000 words (from the introduction to the end of the conclusion).

Case reports, case series, narrative reviews and letters to the editor

Starting in June 2018, only individual case reports dealing with situations of public health emergencies will be accepted by *São Paulo Medical Journal*. Case reports that had already been accepted for publication up to May 2018 will still be published in a timely manner.

After initial evaluation of scope by the editor-in-chief, case reports, case series and narrative reviews will be considered for peer-review evaluation only when accompanied by a systematic search of the literature, in which relevant studies found (based on their level of evidence) are presented and discussed.¹² The search strategy for each database and the number of articles obtained from each database should be shown in a table. This is mandatory for all case reports, case series and narrative reviews submitted for publication. Failure to provide the search description will lead to rejection before peer review.

The access route to the electronic databases used should be stated (for example, PubMed, OVID, Elsevier or Bireme). For the search strategies, MeSH terms must be used for Medline, LILACS, and Cochrane Library. DeCS terms must be used for LILACS.

EMTREE terms must be used for Embase. Also, for LILACS, the search strategy must be conducted using English (MeSH), Spanish (DeCS) and Portuguese (DeCS) terms concomitantly. The search strategies must be presented exactly as they were used during the search, including parentheses, quotation marks and Boolean operators (AND, OR, and NOT). The search dates should be indicated in the text or in the table.

Patients have the right to privacy. Submission of case reports and case series must contain a declaration that all patients gave their consent to have their cases reported (even for patients cared for in public institutions), in text and images (photographs or imaging examination reproductions). The Journal will take care to cover any anatomical part or examination section that might allow patient identification. For deceased patients whose relatives cannot be contacted, the authors should consult the Editor-in-Chief. All case reports and case series must be evaluated and approved by an ethics committee.

Case reports should be reported in accordance with the CARE Statement,⁷ including a timeline of interventions. They should be structured in the same way as original articles.

Case reports must not be submitted as letters. Letters to the editor address articles that have been published in the *São Paulo Medical Journal* or may deal with health issues of interest. In the category of letters to the editor, the text has a free format, but must not exceed 500 words and five references.

FORMAT: FOR ALL TYPES OF ARTICLES

Title page

The title page must contain the following items:

1. Type of paper (original article, review or updating article, short communication or letter to the editor);
2. Title of the paper in English, which should be brief but informative, and should mention the study design.¹⁴ Clinical trial, cohort, cross-sectional or case-control study, and systematic review are the most common study designs. Note: the study design declared in the title should be the same in the methods and in the abstract;
3. Full name of each author. The editorial policy of the *São Paulo Medical Journal* is that abbreviations of authors' names must not be used; therefore, we ask that names be stated in full, without using abbreviations;
4. Place or institution where the work was developed, city and country;
5. Each author should indicate the way his/her name should be used in indexing. For example: for "João Costa Andrade", the indexed name could be "Costa-Andrade J." or "Andrade JC", as preferred;
6. The author's professional background (Physician, Pharmacist, Nurse, Dietitian or another professional description, or Undergraduate Student); and his/her position currently held (for

- example, Master's or Doctoral Student, Assistant Professor, Associate Professor or Professor), in the department and institution where he/she works, and the city and country (affiliations);
7. Each author should present his/her ORCID identification number (as obtained from HYPERLINK "<http://www.orcid.org/>" www.orcid.org/);
 8. Each author must inform his contribution, preferably following the CRediT system (see above in Authorship);
 9. Date and venue of the event at which the paper was presented, if applicable, such as congresses, seminars or dissertation or thesis presentations.
 10. Sources of financial support for the study, bursaries or funding for purchasing or donation of equipment or drugs. The protocol number for the funding must be presented with the name of the issuing institution. For Brazilian authors, all grants that can be considered to be related to production of the study must be declared, such as fellowships for undergraduate, master's and doctoral students; along with possible support for post-graduate programs (such as CAPES) and for the authors individually, such as awards for established investigators (productivity; CNPq), accompanied by the respective grant numbers.
 11. Description of any conflicts of interest held by the authors (see above).
 12. Complete postal address, e-mail address and telephone number of the author to be contacted about the publication process in the Journal (the "corresponding author"). This author should also indicate a postal address, e-mail address and telephone number that can be published together with the article. *São Paulo Medical Journal* recommends that an office address (rather than a residential address) should be informed for publication.

Second page: abstract and keywords

The second page must include the title and a structured abstract in English with a maximum of 250 words. References must not be cited in the abstract.

The following headings must be used in the structured abstract:

- Background – Describe the context and rationale for the study;
- Objectives - Describe the study aims. These aims need to be concordant with the study objectives in the main text of the article, and with the conclusions;
- Design and setting – Declare the study design correctly, and the setting (type of institution or center and geographical location);
- Methods – Describe the methods briefly. It is not necessary to give all the details on statistics in the abstract;
- Results – Report the primary results;
- Conclusions – Make a succinct statement about data interpretation, answering the research question presented previously. Check that this is concordant with the conclusions in the main text of the article;
- Clinical Trial or Systematic Review Registration – Mandatory for clinical trials and systematic reviews; optional for observational

studies. List the URL, as well as the Unique Identifier, on the publicly accessible website on which the trial is registered.

- MeSH Terms - Three to five keywords in English must be chosen from the Medical Subject Headings (MeSH) list of Index Medicus, which is available at <http://www.ncbi.nlm.nih.gov/sites/entrez?db=mesh>. These terms will help librarians to quickly index the article.
- Author keywords - The authors should also add three to six "author keywords" that they think express the main article themes. These keywords should be different from the MeSH terms and preferably different from words already used in the title and abstract, so as to improve the discoverability of the article by readers doing a search in PubMed. They provide an additional chance for the article to be retrieved, read and cited. Combinations of words and variations (different wording or plurals, for example) are encouraged.

References

For any manuscript, all statements in the text that do not result from the study presented for publication in the *São Paulo Medical Journal* but from other studies must be accompanied by a quotation of the source of the data. All statements regarding health statistics and epidemiological data should generally be followed by references to the sources that generated this information, even if the data are only available electronically.

São Paulo Medical Journal uses the reference style known as the "Vancouver style," as recommended by the International Committee of Medical Journal Editors (ICMJE). Follow the instructions and examples at www.icmje.org, item "References", for the format.

In the text, the references must be numbered in the order of citation. The citation numbers must be inserted after periods/full stops or commas in sentences, and in superscript (without parentheses or square brackets). References cited in the legends of tables and figures must maintain sequence with the references mentioned in the text.

In the list of references, all the authors must be listed if there are up to and including five authors; if there are six or more, the first three should be cited, followed by the expression "et al." For books, the city of publication and the name of the publishing house are mandatory. For texts published on the internet, the complete uniform resource locator (URL) or address is necessary (not only the main home page of a website or link), so that by copying the complete address into a computer internet browser, the Journal's readers will be taken to the exact document cited, and not to a general website.

At the end of each reference, please insert the "PMID" number (for papers indexed in PubMed) and the link to the "DOI" number if available.

Authors are responsible for providing a complete and accurate list of references. All references cited in the text must appear in the reference list, and every item in the reference list must be cited in the text. Also, citations must be in the correct sequence.

Manuscripts that do not follow these guidelines for references will be returned to the authors for adjustments.

The reference list should be inserted after the conclusions and before the tables and figures.

Figures and tables

Images must be submitted at a minimum size that is reproducible in the printed edition. Figures should be sent at a resolution of 300 DPI and minimum size of 2,500 pixels (width) and be recorded in ".jpg" or ".tif" format. Images submitted in inadequate formats will not be accepted.

Images must not be embedded inside Microsoft PowerPoint or Microsoft Word documents, because this reduces the image size. Authors must send the images separately, outside of .doc or .ppt documents. Failure to send the original images at appropriate sizes leads to paper rejection before peer review.

Flowcharts are an exception: these must be drawn in an editable document (such as Microsoft Word or PowerPoint), and should not be sent as an image that can't be changed.

Figures such as bars of line graphs should be accompanied by the tables of data from which they have been generated (for example, sending them in the Microsoft Excel spreadsheets, and not as image files). This allows the Journal to correct legends and titles if necessary, and to format the graphs according to the Journal's style. Graphs generated from software such as SPSS or RevMan must be generated at the appropriate size, so that they can be printed (see above). Authors must provide internal legends/captions in correct English.

All the figures and tables should be cited in the text. All figures and tables must contain legends or titles that precisely describe their content and the context or sample from which the information was obtained (i.e. what the results presented are and what the kind of sample or setting was). The reader should be able to understand the content of the figures and tables simply by reading the titles (without the need to consult the text), i.e. titles should be complete. Acronyms or abbreviations in figure and table titles are not acceptable. If it is necessary to use acronyms or abbreviations inside a table or figure (for better formatting), they must be spelled out in a legend below the table or figure.

For figures relating to microscopic findings (i.e. histopathological results), a scale must be embedded in the image to indicate the magnification used (just like in a map scale). The staining agents (in histology or immunohistochemistry evaluations) should be specified in the figure legend.

DOCUMENTS CITED

1. Internal Committee of Medical Journal Editors. Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals. Available from: <http://www.icmje.org/recommendations/>.
2. The CONSORT Statement. Available from: <http://www.consort-statement.org/>. Accessed in 2018 (May 3).
3. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Br J Surg* 2002. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2168.2000.01610.x>. Accessed in 2019 (April 4).
4. PRISMA. Transparent Reporting of Systematic Reviews and Meta-Analyses. Available from: www.prisma-statement.org. Accessed in 2019 (April 4).
5. STROBE Statement. Strengthening the reporting of observational studies in epidemiology. What is strobe? Available from: <http://www.strobe-statement.org/>. Accessed in 2018 (May 3).
6. von Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61(4):344-9. PMID: 18313558. doi: 10.1016/j.jclinepi.2007.11.008.
7. The CARE Guidelines: Consensus-based Clinical Case Reporting Guideline Development. Enhancing the QUALity and Transparency Of health Research. Available from: <https://www.equator-network.org/reporting-guidelines/care/>. Accessed in 2018 (May 3).
8. STARD Statement. STAndards for the Reporting of Diagnostic accuracy studies. Available from: <http://www.equator-network.org/reporting-guidelines/stard/>. Accessed in 2018 (May 3).
9. Rennie D. Improving reports of studies of diagnostic tests: the STARD initiative. *JAMA*. 2003;289(1):89-90. doi:10.1001/jama.289.1.89.
10. International Committee of Medical Journal Editors (ICMJE). Defining the Role of Authors and Contributors. Available from: <http://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html>. Accessed in 2019 (March 11).
11. International Committee of Medical Journal Editors. Overlapping Publications. Available from: <http://www.icmje.org/recommendations/browse/publishing-and-editorial-issues/overlapping-publications.html>. Accessed in 2018 (Feb 18).
12. Phillips B, Ball C, Sackett D, et al. Oxford Centre for Evidence-based Medicine Levels of Evidence (March 2009). Available from: <https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/>. Accessed in 2018 (May 3).
13. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687. PMID: 24609605; doi: 10.1136/bmj.g1687.
14. Non-randomised controlled study (NRS) designs. Available from: <http://childhoodcancer.cochrane.org/non-randomised-controlled-study-nrs-designs>. Accessed in 2018 (May 3).



Portal MBE
MEDICINA BASEADA
EM EVIDÊNCIAS



Aproveite o Portal MBE

Acesso fácil às evidências científicas mais atualizadas!

O Portal MBE - Medicina Baseada em Evidências foi desenvolvido para tornar a sua rotina de pesquisa mais ágil e eficiente. É o momento aproveitar esse poderoso recurso que coloca as mais recentes evidências científicas ao seu alcance!

Comece a utilizar agora!

Acesse: www.apm.org.br/portal_mbe

Realize sua busca de maneira prática. Se precisar de ajuda, agende um atendimento com nossos especialistas!



Busca Rápida e Integrada:

Realize uma pesquisa única e obtenha resultados das principais bases de dados de saúde.



Suporte Personalizado:

Encontre assistência especializada, agendando atendimentos para dúvidas ou pesquisas específicas.



Acesso às Informações Mais Recentes:

Mantenha-se atualizado com as últimas pesquisas para uma prática médica ainda mais assertiva.

Escaneie para
acessar agora



10% de desconto¹ no plano SulAmérica Saúde para médicos associados à APM.

130 ANOS
SulAmérica



Benefícios



Cuidado 360°: oferta promocional com Saúde + Odonto e Vida sem acréscimo na mensalidade²



Opções de planos com reembolso⁴



Faça uma cotação



(11) 3188-4200

¹ Condição válida apenas para novos clientes, com CNPJ de 3 a 29 vidas. O desconto é aplicado automaticamente na contratação do plano de saúde da operadora SulAmérica.

² Condição promocional válida apenas por 12 ou 24 meses de vigência, conforme vigência do contrato do plano de saúde, podendo ser prorrogado a critério da SulAmérica. Apenas novos proponentes serão elegíveis a esta promoção. O proponente poderá optar por aderir apenas ao Seguro Saúde, a seu exclusivo critério. Este material contém informações resumidas que poderão sofrer alterações sem prévio aviso. As coberturas e a rede referenciada variam de acordo com o Plano contratado. O valor do prêmio do Seguro Odontológico será informado no momento da contratação, e este valor será subsidiado pela SulAmérica, durante o período da campanha. Os Seguros de Vida da SulAmérica estão devidamente registrados na SUSEP. As coberturas do Seguro de Vida, que são válidas durante o período de vigência da apólice, possuem os seguintes números de processos na SUSEP: Cobertura de Morte – 15414.003475/2006-67 e Invalidez Permanente Total ou Parcial por Acidente – 15414.003485/2006-01. Os registros na SUSEP não implicam, por parte da autarquia, incentivo ou recomendação à sua comercialização. Os serviços de assistências não constituem coberturas de seguro e, portanto, não garantem indenização ou reembolso de quaisquer despesas efetuadas. Os Seguros de Vida da SulAmérica, farão parte do contrato de Seguro Saúde, como um benefício adicional. O Seguro Saúde, com um benefício adicional de Seguro de Vida, e o Seguro Odontológico, serão oferecidos pela SulAmérica Companhia de Seguro Saúde, CNPJ/MF nº 01.685.053/0001-56. SulAmérica Serviços de Saúde S.A., CNPJ/MF nº 02.866.602/0001-51 - Registro ANS 416428. O Seguro Saúde, com um benefício adicional de Seguro de Vida, obedece às Condições Gerais, que devem ser lidas previamente à sua contratação, conforme link: <https://portal.sulamericaseguros.com.br/para-empresa/saude/>. O Seguro Odontológico obedece às Condições Gerais, que devem ser lidas previamente à sua contratação, conforme link: <https://portal.sulamericaseguros.com.br/para-empresa/odonto/>. A apólice do Seguro de Vida será emitida pela SULAMÉRICA SEGUROS DE PESSOAS E PREVIDÊNCIA S.A., CNPJ: 01.704.513/0001-46 com registro na SUSEP nº 6220, e obedece suas Condições Gerais, que devem ser lidas previamente à sua contratação, conforme link: <https://portal.sulamericaseguros.com.br/para-voce/vida/segundo-menu-servicos-gerais/consulta-de-condicoes-gerais/>. SAC 0800 722 0504 e Ouvidoria 0800-725-3374, horário atendimento: segunda a sexta, das 8h30 às 17h30.

³ Verifique com o consultor as opções de planos com reembolso. Planos de saúde empresarial, conforme as regras da ANS. Informações resumidas.

A comercialização dos planos respeita a área de abrangência dos produtos e das respectivas operadoras. A disponibilidade dos produtos pode variar de acordo com a região e a entidade de classe com a qual os proponentes mantêm o vínculo.

Os planos podem ser coparticipados. Para conhecer a tabela completa de procedimentos e valores de coparticipação, consulte a operadora.

Todas as informações referentes aos planos, incluindo condições, preços, rede de prestadores e sua abrangência geográfica, são de responsabilidade exclusiva das respectivas operadoras de saúde e podem estar sujeitas a alterações por parte delas, mesmo após a contratação do plano, respeitadas as disposições contratuais e legais (Lei nº 9.656/98). Janeiro.