BJAN

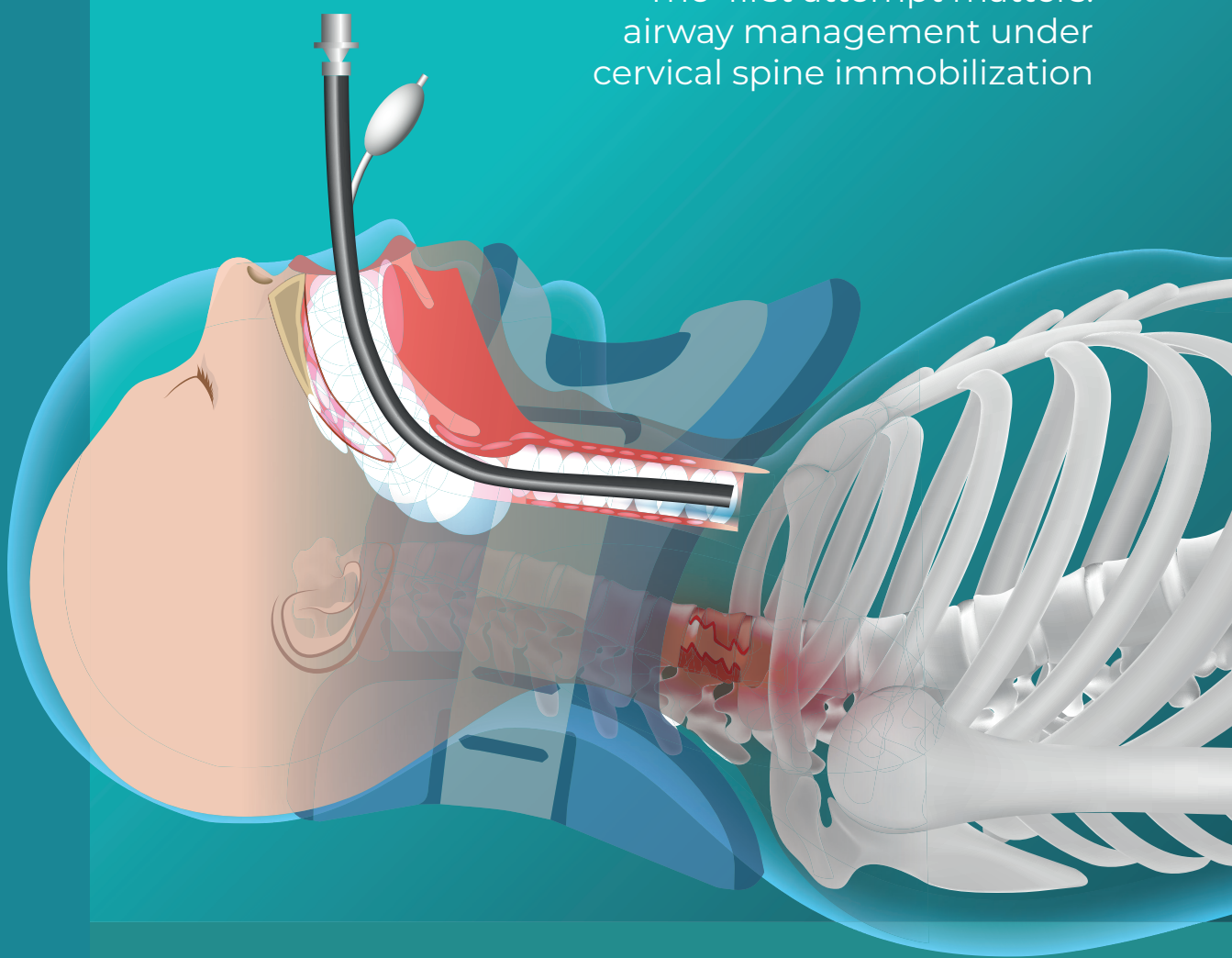
Vol. 76, N. 6 | Nov-Dec 2026 | ISSN 0104-0014

bjan-sba.org

# BJAN
Brazilian Journal of Anesthesiology

# Brazilian Journal of ANESTHESIOLOGY

The first attempt matters: airway management under cervical spine immobilization



Brazilian Journal of Anesthesiology | Vol. 76 N. 6 | November/December 2026

Sociedade Brasileira de Anestesiologia

# Brazilian Journal of ANESTHESIOLOGY

Sociedade Brasileira de Anestesiologia



**Editor-in-Chief**
Liana Maria Torres de Araújo Azi - Departamento de Anestesiologia e Cirurgia da Faculdade de Medicina da Bahia - Universidade Federal da Bahia - Salvador, BA, Brazil

**Co-Editor**
Rodrigo Leal Alves - Hospital São Rafael, Salvador, BA, Brazil

**Associate Editors**
Ana Maria Menezes Caetano - Universidade Federal de Pernambuco, Recife, PE, Brazil
André Prato Schmidt - Hospital de Clínicas de Porto Alegre, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil
Célio Gomes de Amorim - Universidade Federal de Uberlândia, Uberlândia, MG, Brazil
Claudia Marquez Simões - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Durval Campos Kraychete - Universidade Federal da Bahia, Salvador, BA, Brazil
Edmundo Pereira de Souza Neto - Clinique du Pont de Chaume - ELSAN, Montauban, Tarn-et-Garonne, France
Eric Benedet Lineburger - Hospital São José, Criciúma, SC, Brazil
Florentino Fernandes Mendes - Universidade Federal de Ciências da Saúde de Porto Alegre, Porto Alegre, RS, Brazil
Gabriel Magalhães Nunes Guimarães - Universidade de Brasília, Brasília, DF, Brazil
João Manoel da Silva Júnior - Hospital do Servidor Público, São Paulo, SP, Brazil
Lais Helena Navarro e Lima - University of Manitoba, Manitoba, Winnipeg, Canada
Lígia Andrade da S. Telles Mathias - Irmandade da Santa Casa de Misericórdia de São Paulo, São Paulo, SP, Brazil
Lorena Ibiapina Mendes de Carvalho - Hospital Getúlio Vargas, Teresina, PI, Brazil
Luis Vicente Garcia - Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, Ribeirão Preto, SP, Brazil
Luiz Guilherme Villares da Costa - Hospital Israelita Albert Einstein, São Paulo, SP, Brazil
Luiz Marcelo Sá Malbouisson - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Luiz Marciano Cangiani - Hospital da Fundação Centro Médico Campinas, Campinas, SP, Brazil
Marcello Fonseca Salgado-Filho - Universidade Federal Fluminense, Niterói, RJ, Brazil
Maria José Carvalho Carmona - Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Mariana Fontes Lima Neville - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Neuber Martins Fonseca - Faculdade de Medicina da Universidade Federal de Uberlândia, Uberlândia, MG, Brazil
Norma Sueli Pinheiro Módolo - Faculdade de Medicina de Botucatu da Universidade Estadual Paulista, Botucatu, SP, Brazil
Oscar César Pires - Universidade de Taubaté, Taubaté, SP, Brazil
Paulo Corrêa da Silva Neto - Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
Paulo do Nascimento Junior - Faculdade de Medicina de Botucatu da Universidade Estadual Paulista, Botucatu, SP, Brazil
Priscilla Ferreira Neto Cardoso - Centre Hospitalier du Haut Anjou (CHHA), Pays de Loire, France
Ricardo Vieira Carlos - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Rodrigo Moreira e Lima - University of Manitoba, Manitoba, Winnipeg, Canada
Vanessa Henriques Carvalho - Universidade Estadual de Campinas, Campinas, SP, Brazil
Vinicius Caldeira Quintão - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil

**Editorial Committee**
Achal Dhir - London Health Sciences Centre, Western University, London, ON, Canada
Adrian Alvarez - Hospital Italiano de Buenos Aires, Buenos Aires, BA, Argentina
Adrian Gelb - University of California, San Francisco, CA, USA
Alexandra Rezende Assad - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil
Alexis F. Turgeon - Université Laval, Québec, QC, Canada
Antônio Carlos Aguiar Brandão - Universidade do Vale do Sapucaí, Pouso Alegre, MG, Brazil
Ary Serpa Neto - Monash University, Melbourne, Australia
Bernd W. Böttiger - University Hospital of Cologne, Klinikum Köln, NW, Germany
Bobbie Jean Sweitzer - Northwestern Medicine, Chicago, IL, USA
Carlos Galhardo Júnior - Instituto Nacional de Cardiologia (INC/MS), Rio de Janeiro, RJ, Brasil
Carlos Manuel Correia Rodrigues de Almeida - Hospital CUF Viseu, Viseu, Beira Alta, Portugal
Catia Sousa Govêia - Universidade de Brasília, Brasília, DF, Brazil
Clarita Bandeira Margarido - Sunnybrook Health Sciences Care, Toronto, ON, Canada
Claudia Regina Fernandes - Universidade Federal do Ceará, Fortaleza, CE, Brazil
Clovis Tadeu Bevilacqua Filho - Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil
Clyde Matava - The Hospital for Sick Children, Toronto, ON, Canada
Cyril David Mazer - St. Michael's Hospital, Toronto, ON, Canada
dade de São Paulo, São Paulo, SP, Brazil
Daniel Cordovani - McMaster University, Hamilton, Canada
David Ferez - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Deborah Culley - Harvard University, Boston, MA, USA
Deepak K. Tempe - GB Pant Institute of Postgraduate Medical Education and Research, New Delhi, India
Edoardo Quarenghi - Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, MI, Italy
Eduarda Schütz Martinelli - Toronto General Hospital, Toronto, ON, Canada
Eduardo Giroud Joaquim - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Eliane Cristina de Souza Soares - Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
Emery Brown - Massachusetts Institute of Technology, Cambridge, MA, USA
Enrique A. Goytizolo - Hospital for Special Surgery (HSS), New York, NY, USA
Fabiana A. Penachi Bosco Ferreira - Universidade Federal de Goiás, Goiânia, GO, Brazil
Fábio Papa - University of Toronto, Toronto, ON, Canada
Federico Bilotta - Sapienza Università Di Roma, Roma, RM, Italy
Felipe Chiodini - Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Fernanda Bono Fukushima - Universidade Estadual Paulista, Botucatu, SP, Brazil
Fernando Abelha - Hospital de São João, Porto, Portugal
Filipe Nadir Caparica Santos - University of Toronto, Toronto, ON, Canada
Francisco A. Lobo - Cleveland Clinic Abu Dhabi, Abu Dhabi, United Arab Emirates
Frederic Michard - MiCo, Consulting and Research, Denens, Switzerland
Getúlio Rodrigues de Oliveira Filho - Universidade Federal de Santa Catarina, Florianópolis, SC, Brazil
Gildásio de Oliveira Júnior - Alpert Medical School, Brown University, Providence, RI, USA
Giovanni Landoni - Vita-Salute San Raffaele University, Milano, LOM, Italy
Glauber Gouvêa - Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil
Glenio Bitencourt Mizubuti - Queen's University, Kingston, Canada
Gregory Hare - University of Toronto, Toronto, ON, Canada
Guilherme Antonio Moreira de Barros - Faculdade de Medicina de Botucatu da Universidade Estadual Paulista, Botucatu, SP, Brazil
Hazem Adel Ashmawi - Universidade de São Paulo, São Paulo, SP, Brazil
Ismar Lima Cavalcanti - Universidade Federal Fluminense, Niterói, RJ, Brazil
Jean-Jacques Rouby - Pierreand Marie Curie University, Paris, France
Jean-Louis Teboul - Paris-Sud University, Paris, France
Jean-Louis Vincent - Université Libre De Bruxelles, Bruxelles, Belgium
Jean-Michael Constantin - La Pitié Salpetriere Hospital, University Paris-Sorbonne, Paris, France
Joana Berger-Estilita - University of Bern, Bern, Switzerland
João Batista Santos Garcia - Universidade Federal do Maranhão, São Luís, MA, Brazil
João Paulo Jordão Pontes - Universidade Federal de Uberlândia, Uberlândia, MG, Brazil
José Carlos Rodrigues Nascimento - Hospital Geral de Fortaleza, Fortaleza, Ceará, Brazil
José Otavio Costa Auler Junior - Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Judymara Lauzi Gozzani - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Kurt Ruetzler - Cleveland Clinic, Cleveland, OH, USA
Laszlo Vutskits - Geneva University Hospitals, Geneva, GE, Switzerland
Leandro Gobbo Braz - Faculdade de Medicina de Botucatu da Universidade Estadual Paulista, Botucatu, SP, Brazil
Leonardo Henrique Cunha Ferraro - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Leopoldo Muniz da Silva - Faculdade de Medicina de Botucatu da Universidade Estadual Paulista, Botucatu, SP, Brazil
Luciana Paula Cadore Stefani - Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil
Luciano Gattinoni - University of Göttingen, Göttingen, Germany
Luis Antonio dos Santos Diego - Universidade Federal Fluminense, Niterói, RJ, Brazil
Luiz Fernando dos Reis Falcão - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Marcelo Gama de Abreu - Cleveland Clinic, Cleveland, OH, USA
Márcio Matsumoto - Hospital Sírio Libanês, São Paulo, SP, Brazil
Marcos Antônio Costa de Albuquerque - Universidade Federal de Sergipe, São Cristóvão, SE, Brazil
Marcos Francisco Vidal Melo - Columbia University, New York, NY, USA
Maria Angela Tardelli - Universidade Federal de São Paulo, São Paulo, SP, Brazil
Mário J. da Conceição - Fundação Universidade Regional de Blumenau, Blumenau, SC, Brazil
Massimiliano Sorbello - AOU Policlinico Vittorio Emanuele, Catania, SIC, Italy
Matheus Fachini Vane - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Mônica Maria Siaulys - Hospital e Maternidade Santa Joana, São Paulo, SP, Brazil
Nádia Maria da Conceição Duarte - Universidade Federal de Pernambuco, Recife, PE, Brazil
Nicola Disma - Istituto Giannina Gaslini, Génova, GE, Italy
Oscar César Pires - Universidade de Taubaté, Taubaté, SP, Brazil
Paulo Adilson Herrera - Hospital Evangélico de Londrina, Londrina, PR, Brazil
Paulo Alípio - Universidade Federal Fluminense, Niterói, RJ, Brazil
Pedro Amorim - Centro Hospitalar e Universitário do Porto, Porto, Portugal
Pedro Francisco Brandão - Universidade Federal do Espírito Santo, Vitória, ES, Brazil
Peter D. Slinger - University of Toronto, Toronto, ON, Canada
Philip Peng - University of Toronto, Toronto, ON, Canada
Plínio da Cunha Leal - Hospital São Domingos, São Luís, MA, Brazil
Raffael Pereira Cezar Zamper - London Health Science Center, London, UK
Rajinder K. Mirakhur - Royal Hospital, Belfast, Northern Ireland, UK
Ricardo Antônio Guimarães Barbosa - Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, SP, Brazil
Ricardo Esper Treml - Department of Anesthesiology, Perioperative and Pain Medicine, Stanford Health Care, CA, USA
Roberto Fumagalli - Università degli studi Milano Bicocca, Milano, MI, Italy
Rodrigo Daly Guris - Children's Hospital of Philadelphia, Philadelphia, PA, USA
Rogean Rodrigues Nunes - Centro de Ensino e Treinamento/SBA (Residência Médica) em Anestesiologia do Hospital Geral de Fortaleza, Fortaleza, CE, Brazil
Sara Lúcia Ferreira Cavalcante - Hospital Geral do Inamps de Fortaleza, Fortaleza, CE, Brazil
Suzana Margareth Lobo - Faculdade de Medicina de São José do Rio Preto (FAMERP), São José do Rio Preto, SP, Brazil
Thais Cançado - Serviço de Anestesiologia de Campo Grande, Campo Grande, MS, Brazil
Thomas Engelhardt - Montreal Children's Hospital, McGill University, Montreal, Canada
Tom G. Hansen - Akershus University Hospital, Oslo, Norway
Vicente Faraon Fonseca - Unisinos, Porto Alegre, RS, Brazil
Waynice Paula-Garcia - Universidade de São Paulo, São Paulo, SP, Brazil
Wesla Pfeifer - University of Toronto, Toronto, ON, Canada
Wolnei Caumo - Universidade do Rio Grande do Sul, Porto Alegre, RS, Brazil

**Previous Editors-in-Chief**
Oscar Vasconcellos Ribeiro (1951-1957)
Zairo Eira Garcia Vieira (1958-1964)
Bento Mário Villamil Gonçalves (1965-1979)
Masami Katayama (1980-1988)
Antonio Leite Oliva Filho (1989-1994)
Luiz Marciano Cangiani (1995-2003)
Judymara Lauzi Gozzani (2004-2009)
Mario José da Conceição (2010-2015)
Maria Ângela Tardelli (2016-2018)
Maria José Carvalho Carmona (2019-2021)
André Prato Schmidt (2022-2024)

**Editorial Office**
Editorial Coordinator: Gregory Neres
Marketing Analyst: Thadeu de Moraes Carvalho
Librarian: Isabelle Nunes


The Brazilian Journal of Anesthesiology (BJAN) is the official journal of Sociedade Brasileira de Anestesiologia (SBA). The BJAN only accepts original articles for publication submitted and published in English. Before submitting a manuscript, authors must read carefully the Instructions to Authors. It can be found at: <https://bjan-sba.org/instructions>. Manuscripts must be submitted electronically via the Journal's online submission system <http://www.editorialmanager.com/bjan>.

The BJAN publishes original work in all areas of anesthesia, surgical critical care, perioperative medicine and pain medicine, including basic, translational and clinical research, as well as education and technological innovation. In addition, the Journal publishes review articles, relevant case reports, pictorial essays or contextualized images, special articles, correspondence, and letters to the editor. Special articles such as guidelines and historical manuscripts are published upon invitation only, and authors should seek subject approval by the Editorial Office before submission.

The BJAN accepts only original articles that are not under consideration by any other journal and that have not been published before, except as academic theses or abstracts presented at conferences or meetings. A cloud-based intuitive platform is used to compare submitted manuscripts to previous publications, and submissions must not contain any instances of plagiarism. Authors must obtain and send the Editorial Office all required permissions for any overlapping material and properly identify them in the manuscript to avoid plagiarism.

All articles submitted for publication are assessed by two or more members of the Editorial Board or external peer reviewers, assigned at the discretion of the Editor-in-chief or the Associate Editors. Published articles are a property of the Brazilian Society of Anesthesiology (SBA), and their total ou partial reproduction can be made with previous authorization. The BJAN assumes no responsibility for the opinions expressed in the signed works.


**Edited by | Editada por**
Sociedade Brasileira de Anestesiologia (SBA)
Rua Prof. Alfredo Gomes, 36, Botafogo, Rio de Janeiro/RJ, Brazil - CEP 22251-080
Telefone: +55 21 3528-1050
E-mail: contato@sbahq.org
www.sbahq.org

Published by | Publicada por
Elsevier Editora Ltda.
Telefone RJ: +55 21 3970-9300
Telefone SP: +55 11 5105-8555
www.elsevier.com
ISSN: 0104-0014

© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier Editora Ltda. All rights reserved.

Brazilian Journal of ANESTHESIOLOGY

Sociedade Brasileira de Anestesiologia

ISSN 0104-0014 • Volume 76 • Number 6 • November—December 2026

The Brazilian Journal of Anesthesiology is indexed by **Literatura Latino-Americana e do Caribe em Ciências da Saúde (LILACS)** since 1989, **Excerpta Médica Database (EMBASE)** since 1994, **Scientific Electronic Library Online (SciELO - Brasil)** since 2002, **MEDLINE** since 2008, **Scopus since** 2010 and **Web of Science (SCIE - Science Citation Index Expanded)** since 2011.

## Editorial

**844843** **Integrity and trust in anesthesiology research: recognizing and combating scientific misconduct**
*André P. Schmidt, Federico Bilotta, Maria José C. Carmona, and Liana M.T.A. Azi*

**844845** **Nitrous oxide and the environmental case in anesthesia: scale, priority, and Brazilian evidence**
*Ricardo Vieira Carlos, Britta Regli von Ungern-Sternberg, Vinícius Caldeira Quintão, Paulo Sergio Mateus Marcelino Serzedo, and Marcelo Luis Abramides Torres*

**844844** **Videolaryngoscopy in trauma and cervical immobilization: what should guide our first attempt?**
*Liana M.T.A. Azi*

## Original Investigations

**844847** **Cervical spine motion during tracheal intubation with videolaryngoscopy versus Macintosh laryngoscopy in adults undergoing cervical spine surgery: a randomized controlled trial**
*Hieu Trung Doan, Thang Toan Nguyen, Anh Thi Hong Nguyen, and Du Gia Hoang*

**844836** **GlideRite® rigid stylet vs Flex-it® articulated stylet for intubation with videolaryngoscopy in patients with semi-rigid cervical collar: a randomized trial**
*John Viet Nguyen, Eric Zhu, Said Dababneh, Pierre Drolet, Marie-Ève Bélanger, Rami Issa, Valerie Zaphiratos, Nadia Godin, Jean S. Bussières, Mihai Georgescu, and Issam Tanoubi*

**844824** **Efficacy and safety of ciprofol versus remimazolam combined with alfentanil for sedation during outpatient colonoscopy: a randomized double-blind noninferiority trial**
*Xuan Liu, Qianqian Tang, Ning Zou, Changming Wang, and Xuejun Guo*

**844839** **Implementing patient blood management in a Brazilian public women's hospital: a pre—post implementation study**
*Lorena Costa Corrêa, Nataly Andrade Alves de Paiva, Ana Carolina Dias Almeida, Ulisses Montarroyos, and Cesar de Almeida-Neto*

**844837** **Remifentanil for modified rapid sequence induction in children: a proof-of-concept study in a selected pediatric population**
*Julien Pico, Chrystelle Sola, Laurent Hertz, Julien Riera, Christopher Scott, Kévin Chapron, Philippe Pirat, Sophie Bringuier, and Christophe Dadure*

## Review Articles


**844840** **Awake versus deep airway device removal in pediatric tonsillectomy: a systematic review and meta-analysis**
*Henrique Baggio Pivetta, Alice O. Almeida, Rafael L. Miranda, Rafaela P. Silva, Raphaela G. Mendes, João M.M. Borges, Raquel E. Saldanha, Matheus Arraes, Pablo P. Sanín, and Walid Alrayashi*

**844848** **Intraoperative methadone in adult laparoscopic surgery: a systematic review and meta-analysis of randomized trials**
*Julio S. Curi, Matheus Arraes, Jasmeen Kaur, Carlos Martins, Rafael A. Neves, Eduardo M. Pereira, Shaun E. Gruenbaum, and Rafael A. Lombardi*


## Short Communication


**844834** **PRISMA 2020: strengthening transparency and usefulness of systematic reviews**
*Leonardo Roever, Matthew J. Page, and David Moher*


## Letters to the Editor


**844821** **Post-dural puncture headache in patients ≥ 50 years: a multicenter cohort study**
*Scott Seadon, Andrew Walker, Abraham O. Fifen, and Joanna J. Fifen*

**844835** **Beyond forest plots: trial sequential analysis of magnesium sulfate as an adjuvant in supraclavicular brachial plexus block**
*Willyam Barros Saraiva, Rafael Wanderley Persiano Malta, Roberta Ribeiro Brandão Caldas, Olavo Barbosa de Oliveira-Neto, Fabiano Timbó Barbosa, and Célio Fernando de Sousa-Rodrigues*

**844846** **The premature abandonment of thoracic epidural analgesia in the Era of video-assisted thoracoscopic surgery thoracic epidural analgesia in the VATS Era**
*Gabriel Trindade Avelar, André Versiani Caldeira Rocha, Regina Bicalho Gomes de Faria, and Marina Ayres Delgado*

**844825** **Smartphone-based immersive virtual reality for pediatric preoperative preparation: development and initial feasibility report**
*Andreia Cardoso, and Durval Campos Kraychete*

**844816** **Opioid-free versus opioid-containing labor epidural analgesia: interpreting non-inferiority claims in a retrospective cohort study**
*Pourya Baharlooei*

**844849** **ESPB versus thoracic paravertebral block in pediatric nephrolithotomy: analytic population and inference after a negative superiority trial**
*Swarupanjali Padhi, Prashant Ramdas Kokiwar, Janvi Patel, and Ankita Kalra*

**844818** **Authors' reply to comments on erector spinae plane block versus thoracic paravertebral block in pediatric percutaneous nephrolithotomy**
*Fatma Nabil, Ahmed M. Mandour, Amr M. Abdelgawad, Deiaaeldin M. Tamer, Ahmed A. Shahat, Mohamed Anwar, and Hany M. Osman*

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



EDITORIAL

# Integrity and trust in anesthesiology research: recognizing and combating scientific misconduct



Biomedical research is a cornerstone of clinical advancement, providing the evidence that underpins modern medical practice. That enterprise rests on trust. Readers accept a paper because they expect that the participants existed, the data were collected, and the analysis was reported as planned. When an article is retracted for fabrication, falsification, or plagiarism, the damage is not confined to one citation. The retraction undermines confidence in the surrounding literature and, if the record is left uncorrected, in the credibility of science itself. The consequences are concrete: misleading clinical practice, jeopardized patient safety, wasted resources, and eroded public trust.[1,2] The medical community must remain vigilant in an era of heightened scrutiny and "publish or perish" pressures. Recognizing, preventing, and addressing scientific misconduct is a collective responsibility to protect scientific credibility and patient welfare.

Research misconduct and related failures are not of equal gravity, and they should not be handled as if they were. Research misconduct, in the narrow sense of fabrication, falsification, and plagiarism, is intentional deception. It is grounds for retraction, institutional sanction, and, in some jurisdictions, professional or legal penalty. Questionable research practices sit below that threshold but still distort the record: p-hacking, selective reporting, inadequate blinding, outcome switching, and hypothesizing after the results are known. They are not always prosecutable as fraud, yet they inflate false-positive findings and should be treated as integrity problems, not as style. Methodological weakness is a third category. Inadequate blinding, an underpowered comparison, or a poorly defined endpoint can make a study unusable without anyone having invented a number. Honest error is a fourth. A mislabeled table or a coding mistake calls for a correction, not a misconduct finding, as long as it is disclosed and fixed. All of them can waste funding, distort guidelines, and feed irreproducible findings. Retracted papers, whether withdrawn for misconduct or left uncorrected after a fatal flaw, can linger as "zombie" citations and continue to influence subsequent research and meta-analyses.[3,4] In anesthesiology and perioperative medicine, the stake is clinical. Tainted evidence has supported protocols later judged unsafe, which is why appraisal of data integrity belongs in specialty training alongside appraisal of design.

Scientific misconduct is not confined to any single specialty but has left a notable imprint on clinical anesthesiology and perioperative medicine. Several high-profile cases have underscored the devastating effects of misconduct in our field. For instance, anesthesiologist Scott Reuben fabricated data in at least 21 papers between 1996 and 2008, inventing patient outcomes to support pain management protocols involving drugs like Rofecoxib and Celecoxib.[3] This led to multiple retractions, a six-month prison sentence for healthcare fraud, and a reevaluation of multimodal analgesia guidelines that had influenced clinical practice for years.[3] Similarly, Joachim Boldt's fabrication across studies of hydroxyethyl starch for fluid management remains one of the largest retraction scandals in medical history.[4] By mid-2026, his retraction count had reached more than 200, and journals were still removing papers more than a decade after the original institutional findings, including articles previously left standing.[5] Some of those papers continued to be cited years after the scandal became public, illustrating how slowly the scientific record self-corrects even when fraud is no longer in dispute.[5] Notably, some of the forged papers on hydroxyethyl starches supported clinical practices that were subsequently shown to threaten patient safety.[6,7] These incidents underscore the direct harm to patient outcomes, as falsified data led to the adoption of potentially dangerous protocols in perioperative settings, necessitating urgent corrections to clinical guidelines and practices.

The case of Yoshitaka Fujii, an anesthesiologist who fabricated data in more than 160 papers on postoperative nausea and vomiting over two decades, has also exposed flaws in peer review systems. Statistical inconsistencies in his results went undetected for years, leading to revisions of antiemetic guidelines worldwide and underscoring the need for enhanced detection tools to support post-publication

https://doi.org/10.1016/j.bjane.2026.844843
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

scrutiny.[8] A more recent specialty-specific scandal involves Hironobu Ueshima, formerly of Showa University Hospital in Tokyo. Institutional and Japanese Society of Anesthesiologists investigations concluded that 142 of his articles, including original papers, case reports, and a large volume of letters on regional anesthesia and plane blocks, contained fabricated or falsified data and/or improper authorship.[9] More than 100 of those papers were subsequently retracted, placing Ueshima among the most-retracted authors in the biomedical literature and contaminating a body of work that had been rapidly absorbed into regional-anesthesia practice.[9] The scale of these cases has led the specialty in Japan to treat prevention as a society-level responsibility, including mandatory ethics training, centralized data stewardship, and confidential reporting pathways.[10] They also show that fraud is not confined to large multicenter trials: high-volume letters and "technique" papers, once absorbed into everyday practice, can contaminate the clinical literature just as efficiently.

Another egregious example involves cardiologist Don Poldermans, whose falsified data in the DECREASE trials on beta-blockers for non-cardiac surgery patients infiltrated European clinical guidelines.[11] Dismissed from his institution in 2011, Poldermans' work resulted in over 10 retractions and forced revisions to the European Society of Cardiology's perioperative care recommendations, demonstrating how fraudulent research can misdirect treatment protocols on a continental scale and endanger patient safety. Investigations uncovered not only data fabrication but also ethical lapses, such as unreported adverse events and conflicts of interest with pharmaceutical sponsors. The fallout extended beyond retractions: a meta-analysis[12] excluding his work showed an increased mortality rate associated with beta-blockers, prompting revisions to guidelines by the European Society of Cardiology and the American College of Cardiology.

A comprehensive analysis of retraction processes for papers by Boldt, Reuben, and Fujii highlights ongoing deficiencies in journal and publisher responses.[13] By mid-2018, 94% of the 313 eligible papers had been retracted. Still, median retraction times varied significantly (1 month for Reuben, 4 months for Boldt, and 13 months for Fujii), with delays attributed to the larger number of affected journals and a lack of coordinated notification. Only 85% of retraction notices complied with the Committee on Publication Ethics (COPE) guidelines, and 21% of retracted articles lacked transparent watermarks to alert readers.[13] The study's inquiries prompted 16 additional retractions, revealing that many journals were unaware of the need to act or cited legal concerns, emphasizing the necessity for standardized, proactive mechanisms to swiftly remove fraudulent work from the scientific record and minimize its lingering impact on research and clinical practice.

These cases not only undermined specific therapeutic recommendations but also exposed systematic vulnerabilities in editorial oversight, peer review, and institutional controls. A seminal statistical analysis by J. B. Carlisle in 2012[8] scrutinized 168 randomized controlled trials published by Yoshitaka Fujii from 1991 to July 2011, evaluating data integrity by comparing reported distributions of categorical and continuous variables with those expected from random sampling. Employing chi-squared goodness-of-fit tests and Monte Carlo simulations, the study found that 85% of variable distributions were inconsistent with chance expectations, with probabilities ranging from 1 in 25 to less than 1 in $10^{33}$. In 141 human studies, all 13 continuous variables and 7 of 11 categorical variables deviated significantly; individually, 96 of 134 randomized studies showed inconsistencies, compared to only 12 of 139 in trials by other authors. These extreme deviations indicate non-random data patterns, likely due to fabrication, underscoring the need to exclude Fujii's work from scientific evidence and highlighting the role of statistical tools in uncovering fraud.

Fraud is not confined to clinical research. Experimental sciences, particularly in biology, are equally susceptible. In a recent perspective, Dubiner and Arbel-Groissman[14] emphasized the alarming ease with which raw data from common laboratory devices, such as microscopes, gel blots, imaging systems, microplate readers, and spectrophotometers, can be fabricated or manipulated without detection, often through editable output files like spreadsheets or images. Their study noted undoubtedly false data (duplicated or impossible values) during peer review, underscoring that the effort to generate and publish fraudulent results is minimal compared to detection. Positive result bias further compounds the problem, as negative findings are rarely published, creating fertile ground for selective reporting, misconduct, and a "secret third option" to the publish-or-perish dilemma: "simply cheat". Conducting clinical and experimental research presents formidable challenges, including securing funding, recruiting suitable participants, navigating ethical approvals, and dealing with the inherent variability in biological responses that often yields inconclusive or negative results for specific interventions. These difficulties can tempt researchers into misconduct, such as data fabrication or selective reporting, to produce "positive" outcomes that satisfy publication demands and career-advancement pressures in a competitive academic environment. Reasons for producing fabricated data encompass academic pressures, where career progression hinges on publication metrics; financial incentives from grants or industry ties that favor positive results; personal ambitions for recognition; institutional demands for productivity; and a lack of robust oversight, all of which can drive researchers to invent outcomes rather than report genuine, often negative, findings. Additionally, funding agencies bear significant responsibility for mitigating fraud, as their emphasis on high-productivity metrics, such as publication counts and grant deliverables, can inadvertently exacerbate pressures that lead to misconduct. To counter this, agencies should revise evaluation criteria to prioritize research quality, reproducibility, and ethical compliance over sheer output, while incorporating integrity audits into grant reviews and providing resources for ethics training.

A recent investigation into a specific open-access journal reveals stark anomalies in editorial handling, where a small subset of editors, comprising just 0.25% of the total, oversaw the publication of 30.2% of retracted articles, despite editing only 1.3% of the journal's overall output.[15] This disproportionate involvement, coupled with networks of reciprocal authorship and editing among flagged individuals from multiple countries, underscores the coordinated nature of fraud facilitation. Such patterns not only evade random editor assignment but also highlight how interconnected groups

exploit journal processes, leading to retractions that often cite concerns over authorship, competing interests, and peer-review integrity.

Efforts to detect and prevent research misconduct have evolved but remain insufficient. Journals have adopted stricter protocols, including mandatory clinical trial registration, professional editorial offices with comprehensive statistical review, comprehensive guidelines for editors and peer reviewers, and the systematic use of plagiarism detection software.[16] However, as shown by Fiore et al.[17] in their 2021 scoping review analyzing 423 retracted publications in anesthesiology, the risk of retraction for fraud does not differ significantly across journals of different impact factors or quartiles. These findings reveal that high-profile journals are not immune, underscoring the need for universal and continuous vigilance across all tiers to address systemic vulnerabilities. Despite the increasing number of retractions, it is crucial to emphasize that continuing medical education and evidence-based clinical practice still rely heavily on regular literature reading. Clinicians must be trained to appraise sources critically, cross-reference with retraction databases, and integrate updated guidelines to maintain high standards of care, ensuring that retractions do not deter engagement with the evolving body of evidence but instead foster a more discerning approach to its application.

A recent review by Nato and Bilotta[18] further illuminates these challenges, reporting over 10,000 paper retractions in 2023, with perioperative medicine heavily impacted. The review highlights high-profile cases in which forged data infiltrated guidelines, leading to suboptimal care and patient harm. The authors propose leveraging machine learning and artificial intelligence (AI) for pre-publication screening, alongside tools like the REAPPRAISED checklist[19] and Benford's Law,[20] to detect anomalies and enhance peer review. Integrating databases such as Retraction Watch with Scopus and Zotero is also suggested to exclude retracted papers from reviews, reinforcing the need for robust, technology-driven safeguards to protect research integrity. Importantly, excluding retracted papers from a review is not a software preference, and it is not a task to leave to the reader after publication. It is an author's duty, enforced by the journal at revision and editorial analysis.

Platforms like Retraction Watch play a pivotal role in exposing fraudulent researchers by maintaining a comprehensive, publicly accessible database of retractions, which tracks misconduct, plagiarism, and errors across disciplines, including anesthesiology.[21] By highlighting patterns, such as disproportionate retractions from paper mills or individual authors, and advocating for transparency, Retraction Watch not only accelerates the removal of flawed studies but also deters future fraud through public accountability, often prompting investigations that lead to additional retractions and institutional actions. Similar strategies, including whistleblower protections and collaborative publisher alerts, amplify this impact by fostering a culture of vigilance; for example, by revealing how publication pressures fuel retractions and ensuring "zombie papers" are flagged in citations, these tools help restore trust in science while underscoring the need for proactive, coordinated responses to misconduct. Independent organizations like Retraction Watch and PubPeer, supported by regional legislation, are essential for monitoring and exposing misconduct. Advocacy is needed for actual punishments, including criminal charges, license revocations, and institutional sanctions, to create deterrents beyond mere retractions and ensure accountability. Whistleblowing plays a crucial role in uncovering misconduct, yet many researchers hesitate because of fears of retaliation. Institutions must implement robust protection mechanisms, including anonymous reporting channels and legal safeguards, to encourage disclosures without professional repercussions. One practical suggestion is to add a standardized question at the end of each downloadable article, or in a dedicated channel on the journal's site, asking readers whether they have any concerns or suspicions about integrity issues in any publication, to support proactive community vigilance and streamline investigations.

The rise of paper mills (clandestine operations that fabricate manuscripts *en masse* and sell authorship) has become a profitable industry, infiltrating even reputable journals and undermining scientific credibility. As detailed by Candal-Pedreira et al.,[22] an analysis of 416 retracted mill-originated papers from 2022 uncovered dense networks where these fraudulent works cite and are cited by other retracted papers, forming clusters of mutually interconnected retractions and a broader web of papers, highlighting citation manipulation as a detection signal and evasion tactic. Parker et al.[23] further elucidate the operational evolution, from template-based fabrication and fake reviewer infiltration to editorial board compromises, estimating mills now comprise over 1.5% of the literature, with AI exacerbating production and obfuscation. Recent escalations underscore the urgency: in 2023-2025, publishers like Hindawi retracted over 11,000 mill papers, Sage withdrew hundreds, and Wiley paused special issues after detecting mill activity in one-in-seven submissions, while investigations revealed mills bribing editors and hijacking journals.[23] The industrial scale of this activity is now measurable. A 2026 machine-learning screen of 2.65 million cancer papers flagged 9.87% as textually similar to known paper-mill products, with the signal rising over time and appearing even in journals in the top decile by impact factor.[24] Flagged papers also attracted more citations than apparently genuine work in the same field, which can inflate journal metrics and recycle fraudulent claims through systematic reviews. These findings do not prove that every flagged article is fabricated, but they show that mills are no longer confined to obscure outlets and that textual and citation forensics must become routine editorial tools, including in perioperative journals. Recent examples in Brazil highlight the sale of authorship positions and paper mill operations, which facilitated fraudulent publications and led to retractions in SciELO-indexed journals.[25] Investigations reveal networks selling slots in fabricated manuscripts, often infiltrating reputable outlets.[26] The lack of effective measures, such as nationwide AI-driven detection systems or mandatory authorship audits, impedes control, allowing these practices to persist amid limited resources and enforcement.

To combat the rising threat of organized fraud via paper mills, which has shown exponential growth in retractions and fabricated products, institutions must prioritize prevention through ethics training and strict data-integrity policies. The COPE guidelines[27] serve as a critical tool in this effort, providing standardized frameworks for publishers, editors, and researchers to address misconduct transparently

and consistently. These guidelines outline best practices for handling allegations of fraud, ensuring timely retractions, and promoting ethical publishing standards, thereby acting as a vital ally in safeguarding research integrity. Complementary tools like plagiarism detectors (e.g., iThenticate, Turnitin), statistical anomaly checks, and open science initiatives for data sharing further fortify defenses, ensuring that integrity is valued over publication metrics.[28] While similarity detection software effectively identifies textual plagiarism through similarity checks, it falls short in detecting more sophisticated forms of fraud, including data fabrication, image manipulation, or statistical anomalies. This limitation underscores the need for complementary tools, such as statistical audits and image forensics, to provide a more comprehensive safeguard against misconduct. In this context, editors and reviewers serve as frontline defenders of research integrity, responsible for detecting anomalies in methodology, data, and ethics during peer review. Their actions include requiring conflict-of-interest disclosures, using plagiarism checkers, requesting raw data for verification, and applying statistical assessments to identify irregularities. When misconduct is suspected, editors must initiate investigations per COPE protocols, issue expressions of concern, and execute timely retractions, while reviewers should flag image manipulations or implausible results, fostering a collaborative effort to uphold standards.[27] The 2025 revision of the COPE retraction guidelines makes that duty more explicit. Editors may retract when they no longer have confidence in the findings, including for paper-mill activity, identity theft or fictitious authorship, and undisclosed involvement of artificial intelligence; the guidance also endorses batch retractions when systematic manipulation of the publication process is identified, rather than waiting for article-by-article institutional verdicts.[27] COPE's 2026 relaunch of its paper-mills working group, together with the multi-stakeholder United2Act initiative, further recognizes that coordinated, AI-amplified fraud now outpaces case-by-case investigation.[29] Journals in anesthesiology and perioperative medicine should align their instructions to authors and their retraction workflows with this updated standard.

The model of access, gold versus diamond open access, may play a critical role in influencing integrity. In gold open access, authors pay an article processing charge. Many gold journals keep a rigorous review process and contribute to dissemination. The same fee, however, can reward volume. Some profit-driven outlets have weakened review, and retraction rates are higher in parts of that market, especially where predatory titles exploit authors in low-resource settings.[30] Diamond open access removes the fee for authors and readers and is usually supported by a society, an institution, or a public fund. That has clear advantages for access and for authors who cannot pay. It does not, by itself, lower the risk of misconduct. Integrity in either model depends on editorial governance: independent editors, a transparent review process, the ability to reject weak or suspicious work, and a retraction policy that is actually used. A diamond journal with a captured editorial board is not safer than a gold journal with a functioning one. The useful question is not which label the journal carries, but who controls acceptance and who can be required to correct the record. Nevertheless, while neither model is immune to abuse and both require ongoing vigilance to prevent it, transitioning to broader adoption of diamond open access could enhance research integrity by democratizing access without the pitfalls of pay-to-publish systems.

The struggle to ensure controlled, high-quality literature continues, with fraudulent publishing activities posing significant challenges despite advances in detection. A recent publication highlights innovative approaches to combat this, such as red-flagging methods that identify the mass production of fake publications in biomedical science using indicators such as unnatural citation patterns and textual anomalies.[31] Open science initiatives, which emphasize transparency through data sharing, preregistration, and accessible repositories, can mitigate research misconduct by enabling independent verification and reducing selective reporting.[32] Beyond access models, open science encompasses key pillars such as the FAIR (Findable, Accessible, Interoperable, Reusable) principles for data sharing, which ensure that datasets are easily discoverable and usable by others, thereby enhancing verification and collaboration.[33] Additionally, open peer review, where reviewer identities and comments are disclosed, and the sharing of analysis codes create a comprehensive ecosystem that promotes reproducibility and transparency, potentially inhibiting misconduct by allowing broader scrutiny and reducing opportunities for hidden manipulations. In this framework, readers play a vital role as post-publication reviewers, contributing through platforms like PubPeer to identify errors or inconsistencies after publication, thereby extending the peer review process and strengthening community-driven accountability.

Traditional strategies for enhancing research integrity, such as prospective trial registration on platforms like ClinicalTrials.gov, adherence to CONSORT guidelines for transparent reporting, and compliance with PRISMA for systematic reviews, suffer from inconsistent implementation. Many trials remain unregistered, allowing outcome switching and publication bias to thrive. At the same time, non-adherence to guidelines obscures methodological flaws, enabling misconduct to evade detection and underscoring the gap between policy and practice. Registered reports represent a particularly potent preventive tool against publication bias and p-hacking, as this format involves peer review and in-principle acceptance of the study methodology and analysis plan before data collection, ensuring that results are published regardless of outcome and reducing the incentive for *post hoc* manipulations. By shifting the emphasis from results to rigorous design, prospective registration of studies promotes greater transparency and reproducibility in research, with growing adoption demonstrating its efficacy in curbing questionable practices.

Research misconduct that has already entered the perioperative literature will not leave unless the workflow that admitted it changes. Moving forward, several additional strategies may help protect research integrity in anesthesiology and beyond:[1,28,32,34] i. Retraction check — Authors of reviews and other evidence syntheses must search the cited record, state that no retractions were found or that a specific number were removed from the analysis; ii. Plagiarism screening — Every submission should be run through similarity software, such as iThenticate, before review; iii. Bibliographic authentication — Each cited title, year, and identifier should

resolve in PubMed or Crossref before acceptance, because fabricated references may pass conventional review; iv. Pre-publication integrity screening — Statistical anomaly checks, image forensics, and data forensics may be considered routine in peer review, alongside tools such as the REAPPRAISED checklist; v. Raw data on request — Authors should deposit or supply data in a form that preserves the original export and records later changes, so reviewers can separate honest error from fabrication; vi. Authorship that can be defended — Gift, guest, and sold authorship should be grounds for rejection; every listed author must meet a stated contribution standard; vii. Limited use of artificial intelligence — Models may assist screening for anomalies. They may not invent text, data, images, or references, and any use must be disclosed; viii. Negative results — Well-conducted negative studies should be publishable, so a null finding is not an incentive to switch outcomes or fabricate a positive one; ix. Institutional accountability — Departments and funders should audit high-volume output, and an independent body should be able to investigate when the host institution is conflicted; x. Education and culture change — Widespread and ongoing training in research ethics, transparency, and reproducibility is essential. Cultivating a culture that prioritizes research quality over quantity remains a critical long-term goal. Ethics and data-handling courses should be required for residents and graduate students, and professional societies, such as the Brazilian Society of Anesthesiology, could lead training in high-risk areas like statistical analysis and ethical data handling to deter misconduct.

While AI holds promise for detecting fraud through anomaly screening and enhanced peer review,[34] it also exacerbates misconduct by enabling the rapid fabrication of deceptive content. Generative AI models, such as large language models, can produce plausible but entirely invented research papers, datasets, images, and even statistical analyses, lowering the barrier for fraudsters to create "paper mill" outputs at scale.[35] A quieter but rapidly growing marker of the same problem is the fabricated citation. An audit of 2.5 million biomedical papers in PubMed Central found 4,046 references whose titles matched no existing publication, across 2,810 articles. The rate rose more than twelve-fold between 2023 and early 2026, reaching one in 277 papers in the first weeks of 2026, a trajectory that tracks the publication lag after widespread adoption of large language models.[36] Because hallucinated references are often correctly formatted and attributed to real authors, conventional peer review rarely catches them. For anesthesiology editors, automated reference verification should therefore sit alongside plagiarism and image-forensics checks at submission, and authors should be required to attest that they confirmed every citation against a bibliographic database. For instance, AI-driven tools have been implicated in generating fraudulent medical articles that mimic authentic research, leading to undetected plagiarism, misinformation, and manipulated results that impair scientific credibility.[35] The same tools that can screen for anomalies can also produce the material being screened. The risk that matters here is not generic misuse of AI. It is the fabricated paper, the invented dataset, the manipulated image, and the citation that points to no article. Anesthesiology journals need AI-specific forensics for those products, and instructions to authors that say what use of a model must be disclosed.

A clear research agenda for anesthesiology may include systematic studies on fraud detection methods, such as the application of statistical tools like Benford's Law for identifying anomalies in reported data from anesthesia trials, as demonstrated in analyses of falsified papers where deviations from expected digit distributions have proven effective in uncovering fabrication.[18] Additionally, it should explore the impact of retracted studies on clinical practice, examining how "zombie papers" persist in citations within systematic reviews and guidelines, leading to misguided perioperative protocols and potential patient harm, with evidence showing that retracted randomized controlled trials continue to be inappropriately referenced despite clear retraction notices. Furthermore, the agenda ought to prioritize the development of tools to track research integrity indicators prospectively, including automated platforms for evaluating adherence to good clinical practices in randomized controlled trials, or integrated databases that monitor retraction patterns and citation networks to preempt the dissemination of flawed evidence in anesthesiology literature. To broaden its scope, this agenda could incorporate evaluations of AI's dual role in fraud detection and perpetuation, investigating how AI-driven anomaly screening enhances peer review while also enabling sophisticated data fabrication in paper mills, necessitating the creation of AI-specific ethical forensics tailored to high-stakes fields like anesthesia and perioperative medicine. It should also assess the effects of open science practices on reducing misconduct, such as through preregistration and data sharing that promote transparency and hinder selective reporting, with studies indicating that these initiatives foster a culture of accountability and decrease the incidence of questionable research behaviors in biomedical fields. Moreover, research on the prevalence and consequences of predatory publishing in anesthesiology is essential, given exploitative models that prioritize volume over rigor, leading to higher retraction rates and the infiltration of low-quality studies into clinical decision-making. Finally, the agenda must evaluate the efficacy of ethical training programs in preventing scientific fraud, including mandatory responsible conduct of research workshops that emphasize transparency and institutional accountability, as evidence suggests these interventions can significantly mitigate misconduct when integrated early in researchers' careers.[37] Ethical training should be integrated into the curricula of medical residents and graduate students, emphasizing responsible conduct of research, data integrity, and the consequences of misconduct to raise a generation of clinicians and researchers committed to upholding scientific standards from the outset of their careers.

National postgraduate program evaluations have traditionally been heavily reliant on publication metrics, such as the number of papers and their impact factors, which inadvertently fuel publication pressures and contribute to scientific misconduct. To address these vulnerabilities, adaptations are essential to incorporate broader, more holistic criteria that value quality, societal relevance, and ethical practices over volume of publications. In Brazil, the Coordination for the Improvement of Higher Education Personnel (CAPES), a federal agency that coordinates efforts to improve the quality of Brazilian faculty and postgraduate students, has implemented a transformative evaluation system

for postgraduate programs *stricto sensu* during the 2025–2028 cycle assessment.[38] This system adopts a multidimensional approach that integrates bibliometric indicators with qualitative evaluations of societal impact, open science practices, diversity, inclusion, and alignment with United Nations Sustainable Development Goals, guided by principles of peer review, comparability, retrospective analysis, classification, collaboration, and transparency to ensure community involvement and holistic assessment. A further change is that, for this cycle, individual articles, not only the journals that publish them, are classified, combining venue metrics with article-level quality, societal impact, and open-science practices.[38] This shift is intended to weaken the incentive to chase journal strata regardless of the integrity of the underlying work. Future assessments in Brazil will be crucial to determine the effectiveness of these new strategies in reducing misconduct and promoting sustainable research cultures, offering valuable insights that could inspire similar regulatory changes in other countries grappling with comparable challenges in postgraduate education.

In conclusion, biomedical research is fundamental to improving patient care, but its foundation must be built on integrity. In anesthesiology and perioperative medicine, where clinical decisions often carry high stakes, the need to recognize, prevent, and address scientific misconduct is especially urgent. Safeguarding the scientific record is not merely a technical task, it is a moral imperative shared by researchers, reviewers, editors, and institutions alike. Reaffirming our commitment to transparency, honesty, and rigorous science is essential to ensuring the sustainable advancement of medical knowledge and the continued trust of the society we serve.

## Data availability statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

## Conflicts of interest

The authors declare no conflicts of interest.

## Editor

Liana Azi

## References


1. Li W, van Wely M, Gurrin L, Mol BW. Integrity of randomized controlled trials: challenges and solutions. Fertil Steril. 2020;113:1113–9.
2. Flier JS. Misconduct in Bioscience Research: a 40-year perspective. Perspect Biol Med. 2021;64:437–56.
3. Bornemann-Cimenti H, Szilagyi IS, Sandner-Kiesling A. Perpetuation of Retracted Publications Using the Example of the Scott S. Reuben Case: Incidences, Reasons and Possible Improvements. Sci Eng Ethics. 2016;22:1063–72.
4. Wise J. Boldt: the great pretender. BMJ. 2013;346:f1738.
5. Marcus A, Oransky I. Now, 240: more than a decade later, journals are still retracting Joachim Boldt's papers. Retraction Watch. 15 July 2026. Available from: https://retractionwatch.com/2026/07/15/joachim-boldt-retractions-journals-more-than-decade-later/.
6. Nato CG, Tabacco L, Bilotta F. Fraud and retraction in perioperative medicine publications: what we learned and what can be implemented to prevent future recurrence. J Med Ethics. 2022;48:479–84.
7. Schmidt AP, Bilotta F. Hydroxyethyl starch for perioperative fluid management: a critical appraisal. Braz J Anesthesiol. 2023;73:529–31.
8. Carlisle JB. The analysis of 168 randomised controlled trials to test data integrity. Anaesthesia. 2012;67:521–37.
9. Retraction Watch. More than 100 of an anesthesiologist's papers retracted. 10 February 2022. Available from: https://retractionwatch.com/2022/02/10/more-than-100-of-an-anesthesiologists-papers-retracted/.
10. Hirota K, Saito J, Mazda Y, et al. Anesthesia research misconduct in Japan: understanding the status is critical to prevention. J Anesth. 2026;40:106–10.
11. Chopra V, Eagle KA. Perioperative mischief: the price of academic misconduct. Am J Med. 2012;125:953–5.
12. Bouri S, Shun-Shin MJ, GD Cole, Mayet J, Francis DP. Meta-analysis of secure randomised controlled trials of $\beta$-blockade to prevent perioperative death in non-cardiac surgery. Heart. 2014;100:456–64.
13. McHugh UM, Yentis SM. An analysis of retractions of papers authored by Scott Reuben, Joachim Boldt and Yoshitaka Fujii. Anaesthesia. 2019;74:17–21.
14. Dubiner S, Arbel-Groissman M. Advancing data honesty in experimental biology. J Exp Biol. 2024;227:jeb247383.
15. Richardson RAK, Hong SS, Byrne JA, Stoeger T, Amaral LAN. The entities enabling scientific fraud at scale are large, resilient, and growing rapidly. Proc Natl Acad Sci U S A. 2025;122:e2420092122.
16. Schmidt AP, Carmona MJC. Registration of clinical trials in anesthesiology: promoting transparency in clinical research. Braz J Anesthesiol. 2022;72:685–7.
17. Fiore M, Alfieri A, Pace MC, et al. A scoping review of retracted publications in anesthesiology. Saudi J Anaesth. 2021;15:179–88.
18. Nato CG, Bilotta F. Fraud in Medical Publications. Anesthesiol Clin. 2024;42:607–16.
19. Grey A, Bolland MJ, Avenell A, Klein AA. Gunsalus CK. Check for publication integrity before misconduct. Nature. 2020;577:167–9.
20. Eckhartt GM, Ruxton GD. Investigating and preventing scientific misconduct using Benford's Law. Res Integr Peer Rev. 2023;8:1.
21. Retraction Watch [Internet]. Available from: https://retractionwatch.com.
22. Candal-Pedreira C, Guerra-Tort C, Ruano-Ravina A, et al. Retracted papers originating from paper mills: a cross-sectional analysis of references and citations. J Clin Epidemiol. 2024;172:111397.
23. Parker L, Boughton S, Bero L, Byrne JA. Paper mill challenges: past, present, and future. J Clin Epidemiol. 2024;176:111549.
24. Basu M. Paper mill cancer studies get double the number of citations as genuine papers. Nature. 2026;655:552–3.
25. Santos-d'Amorim K, Wang T, Lund B, Macedo Dos Santos RN. From plagiarism to scientific paper mills: a profile of retracted articles within the SciELO Brazil collection. Ethics Behavior. 2024;34:40–57.
26. Tang G, Cai H. Citation Contamination by Paper Mill Articles in Systematic Reviews of the Life Sciences. JAMA Netw Open. 2025;8:e2515160.

27. Committee on Publication Ethics. Retraction guidelines. Version 3. 4 September 2025. doi: 10.24318/cope.2019.1.4. Available from: https://publicationethics.org/guidance/guideline/retraction-guidelines.
28. Marusic A, Wager E, Utrobicic A, Rothstein HR, Sambunjak D. Interventions to prevent misconduct and promote integrity in research and publication. Cochrane Database Syst Rev. 2016;4: MR000038.
29. Committee on Publication Ethics. Rekindling the flame: back to roots on paper mills. 23 June 2026. Available from: https://publicationethics.org/news-opinion/rekindling-flame-back-roots- paper-mills.
30. Wang T, Xing QR, Wang H, Chen W. Retracted Publications in the Biomedical Literature from Open Access Journals. Sci Eng Ethics. 2019;25:855—68.
31. Sabel BA, Knaack E, Gigerenzer G, Bilc MI. Fake publications in biomedical science: red-flagging method indicates mass production. Naunyn Schmiedebergs Arch Pharmacol. 2026;399:2943—55.
32. Lumbard H, Routledge D. Open science and transparency are our strongest tools in the fight against fraudulent publishing activities. PLoS Med. 2025;22:e1004774.
33. UNESCO Recommendation on Open Science. Paris: United Nations Educational, Scientific and Cultural Organization; 2021 Nov 23. Available from: https://en.unesco.org/science-sustainable-future/open-science/recommendation. doi: 10.54677/MNMH8546.
34. Gomase VS, Ghatule AP, Sharma R, Sardana S. Leveraging Artificial Intelligence for Data Integrity, Transparency, and Security in Technology-enabled Improvements to Clinical Trial Data Management in Healthcare. Rev Recent Clin Trials. 2025. https://doi.org/10.2174/0115748871371119250818102753.
35. Chen Z, Chen C, Yang G, et al. Research integrity in the era of artificial intelligence: Challenges and responses. Medicine (Baltimore). 2024;103:e38811.
36. Topaz M, Roguin N, Gupta P, Zhang Z, Peltonen LM. Fabricated citations: an audit across 2.5 million biomedical papers. Lancet. 2026;407:1779—81.
37. Zhaksylyk A, Zimba O, Yessirkepov M, Kocyigit BF. Research Integrity: Where We Are and Where We Are Heading. J Korean Med Sci. 2023;38:e405.
38. Coordenação de Aperfeiçoamento de Pessoal de Nível Superior. Portaria n° 109, de 2025. Disciplina o processo de Avaliação de Permanência dos Programas de Pós-Graduação stricto sensu no país. Brasília: CAPES; 2025. Available from: https://cad.capes.gov.br/ato-administrativo-detalhar?idAtoAdmElastic=17995.

André P. Schmidt [a,b,c,d,e,f,*], Federico Bilotta [g], Maria José C. Carmona [h], Liana M.T.A. Azi [i]

[a] *Universidade Federal do Rio Grande do Sul (UFRGS), Hospital de Clínicas de Porto Alegre (HCPA), Department of Anesthesia and Perioperative Medicine, Porto Alegre, RS, Brazil*

[b] *Santa Casa de Porto Alegre, Department of Anesthesia, Porto Alegre, RS, Brazil*

[c] *Hospital Nossa Senhora da Conceição (GHC), Department of Anesthesia, Porto Alegre, RS, Brazil*

[d] *Universidade Federal do Rio Grande do Sul (UFRGS), Programa de Pós-graduação em Ciências Cirúrgicas, Porto Alegre, RS, Brazil*

[e] *Universidade Federal do Rio Grande do Sul (UFRGS), Programa de Pós-graduação em Ciências Pneumológicas, Porto Alegre, RS, Brazil*

[f] *Faculdade de Medicina da Universidade de São Paulo (FMUSP), Programa de Pós-Graduação em Anestesiologia, Ciências Cirúrgicas e Medicina Perioperatória, São Paulo, SP, Brazil*

[g] *"Tor Vergata" University of Rome, Department of Anesthesiology and Intensive Care, Rome, Italy*

[h] *Faculdade de Medicina da Universidade de São Paulo (FMUSP), Discipline of Anesthesiology, São Paulo, SP, Brazil*

[i] *Universidade Federal da Bahia (UFBA), Hospital Universitário Professor Edgard Santos, Department of Anesthesiology, Salvador, BA, Brazil*

[*] Corresponding author.
*E-mail:* apschmidt@hcpa.edu.br (A.P. Schmidt).

Available online 11 October 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



EDITORIAL

# Nitrous oxide and the environmental case in anesthesia: scale, priority, and Brazilian evidence

Our responsibility as anesthesiologists does not end at the edge of the operating table. The gases we deliver leave the operating room, persist in the atmosphere for a century or more, and affect the health of populations we will never meet; the environment is therefore not an adjunct to our clinical duty but part of it. Seen in that light, the momentum to remove nitrous oxide ($N_2O$) from operating rooms is a step in the right direction. It has placed the environmental footprint of anesthesia on the agenda of departments that had never previously considered it and has prompted concrete, measurable action where there had long been inertia. Yet the conversation has also become polarized, drifting towards a binary in which a clinician is expected either to defend $N_2O$ or to repudiate it. But in medicine, the right path is rarely found in black and white. A more useful question, surely, is not whether $N_2O$ should be used at all, but where its elimination fits among the many interventions available to a profession that genuinely wants to reduce its emissions. Our argument is one of proportionality: environmental effort, like any finite clinical resource, should be directed first to where it yields the greatest return, and a campaign organized around the most visible target is not the same as one organized around the largest.

It is worth stating clearly what this editorial does not address. There are legitimate, long-standing clinical concerns about $N_2O$, including inactivation of vitamin $B_{12}$ and methionine synthase with prolonged or repeated exposure, the hazards of chronic occupational exposure, and a contribution to postoperative nausea and vomiting. These questions have been examined in dedicated clinical investigations, including the large ENIGMA-II randomized trial, which found no significant increase in death or cardiovascular complications when $N_2O$ was added to general anesthesia in at-risk patients undergoing major non-cardiac surgery.[1] The clinical risk-benefit calculus of $N_2O$ is a real and ongoing discussion that deserves its own forum. It is, however, separate from the environmental case, and conflating the two has muddied both: a clinical reservation about the drug is repeatedly invoked as evidence for an environmental policy, and an environmental anxiety is dressed up as a patient-safety imperative. Here we confine ourselves to the environmental argument, which rests on a single empirical question: where does clinical $N_2O$ actually sit in the global emissions ledger, and where would an equivalent quantum of effort yield greater environmental benefit?

The numbers are not ambiguous. Volatile (halogenated) anesthetics alone are estimated to account for approximately 0.1% of global warming, equivalent to roughly one million automobiles on the road.[2] Worldwide, anesthetic $N_2O$ is estimated to contribute 1% to 3% of global $N_2O$ emissions.[2] The healthcare sector as a whole is responsible for about 4.4% of global greenhouse gas emissions, and within an individual acute hospital, anesthetic gases account for roughly 5% of the institutional carbon footprint.[2,3] These are not trivial figures, nor are they an argument for complacency. But they must be read against the dominant source of atmospheric $N_2O$: the most recent global $N_2O$ budget attributes about 74% of the anthropogenic fraction of $N_2O$ emissions to agriculture, driven by synthetic nitrogen fertilizers and animal manure applied to soils.[4] This is not a marginal contribution that clinical practice might meaningfully offset; it is the overwhelming human source. Against that backdrop, clinical $N_2O$ occupies a fraction of a fraction of the global $N_2O$ footprint (Fig. 1).

The primary leverage for atmospheric $N_2O$ lies in agriculture. Meta-analyses of agricultural soils show that emission factors vary substantially with the form, rate, and timing of nitrogen application, and field studies across the world's principal cereal crops confirm that inorganic fertilizer can substantially increase soil $N_2O$ emissions, leaving optimized nitrogen management as a large and largely unrealized mitigation opportunity.[5,6] The arithmetic of the global budget therefore points unambiguously towards nitrogen-use efficiency in food production as the intervention with the greatest planetary return. This does not absolve anesthesiologists of responsibility for their own emissions, but it sets the scale against which clinical $N_2O$ must be judged.

https://doi.org/10.1016/j.bjane.2026.844845
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

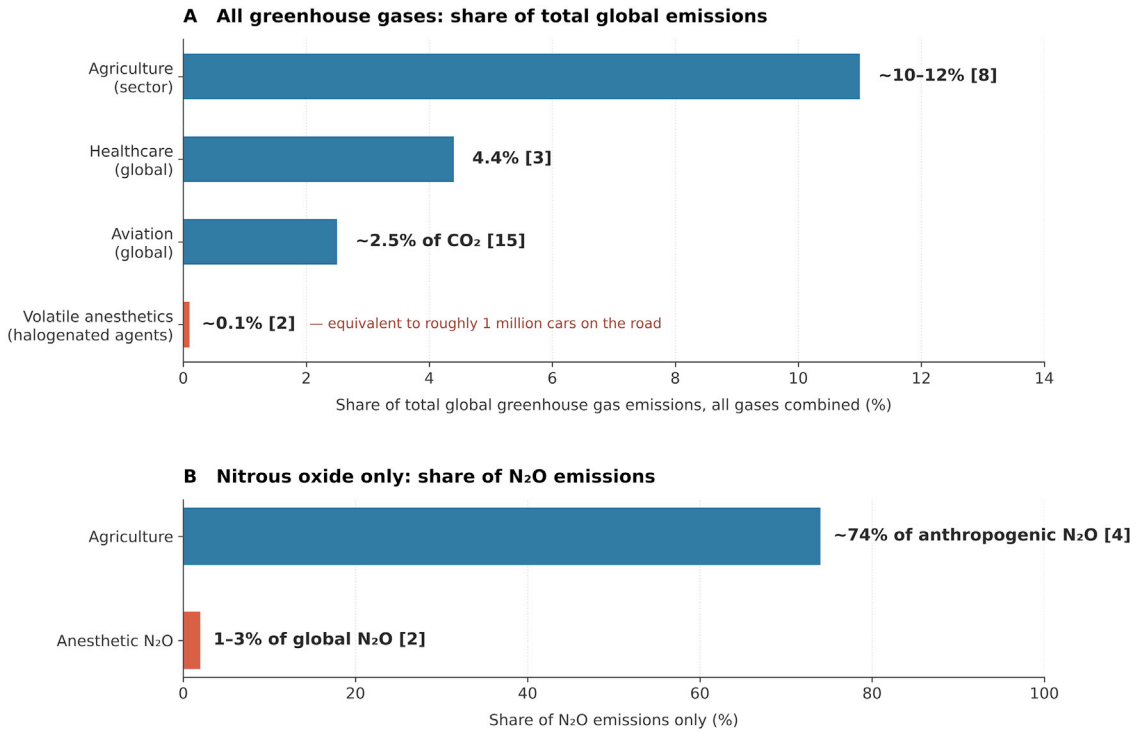



**Figure 1** Anesthetic gases in the global emissions context. The two panels use different denominators. (A) Share of total global greenhouse gas emissions, all gases combined: agriculture accounts for ~10–12%, reflecting its combined methane and $N_2O$ footprint;[8] healthcare for about 4.4%;[3] commercial aviation for approximately 2.5% of global $CO_2$ emissions (more of total warming once non-$CO_2$ effects are included);[15] and volatile (halogenated) anesthetics for approximately 0.1% of global warming, equivalent to roughly one million cars on the road.[2] (B) Share of $N_2O$ emissions only: agriculture accounts for ~74% of anthropogenic $N_2O$ emissions,[4] whereas anesthetic $N_2O$ contributes only 1–3% of total global $N_2O$ emissions.[2] Clinical use therefore represents a fraction of a fraction of the global $N_2O$ burden. Numbers in brackets within the figure correspond to the reference list.

The chemistry underscores the point. $N_2O$ has a 100-year global warming potential of approximately 273 times that of carbon dioxide and an atmospheric lifetime of roughly 109 years. It is now the single most important ozone-depleting substance still emitted at scale.[7,8] None of this is a reason to treat clinical use as the appropriate target. A gas can be both individually potent and collectively minor as a clinical emission, and $N_2O$ is exactly that: its environmental significance is the product of mass emitted and warming potential, and the mass emitted from operating rooms is small relative to the rest of healthcare and, overwhelmingly, to the rest of the economy.

The most important feature of $N_2O$'s hospital footprint, however, is that it does not arise primarily from clinical use. Across multiple international health systems, between 47.2% and 99.8% of the $N_2O$ a hospital procures is lost to the atmosphere through leaks in central pipeline (manifold) systems, without ever reaching a patient.[9] A United Kingdom multicenter study quantified this gap directly: across the participating hospitals, clinical use accounted for only about 18% of the $N_2O$ procured, implying that more than four-fifths was lost before use; comparable work from Australia estimated roughly 77% pre-utilization loss.[7,10] The waste is structural — embedded in reticulated infrastructure that leaks continuously at joints and fittings and compounded by the venting of residual gas when partially full cylinders are exchanged — and it has nothing to do with how judiciously any individual anesthesiologist prescribes the gas. This relocates the problem from the bedside to the building, and from clinical behavior to procurement and engineering: a department could reduce its clinical $N_2O$ use to zero and still leak the great majority of what the hospital buys.

The implication is that the greatest environmental gains come not from restricting clinical use but from decommissioning central pipelines and replacing them with point-of-care portable cylinders. This is now the explicit recommendation of the World Federation of Societies of Anaesthesiologists, whose 2026 statement frames the change as a shift from high-volume waste to low-volume, precision point-of-care use rather than as elimination from clinical care; the American Society of Anesthesiologists has issued a parallel statement. Crucially, evidence that this works has been generated in Brazil: a quality improvement initiative across two quaternary hospitals in São Paulo achieved an 82.5% reduction in anesthesia-related $N_2O$ $CO_2e$ emissions through an educational program combined with workstation-level disconnection of the central pipeline, with $N_2O$ preserved selectively for pediatric inhalational induction and no adverse patient outcomes attributed to the change in delivery.[11] Notably, residual post-intervention emissions in that study, at about 23 kg $CO_2e$ per case, remained roughly threefold higher than benchmarks from settings without pipeline-related leakage — a finding that reinforces rather than weakens the central conclusion: the supply pipeline is

the dominant source, and even an aggressive workstation-level intervention leaves residual upstream losses that full decommissioning can capture. The one element of that intervention we would now qualify is the retention of $N_2O$ for inhalational induction, for the reasons set out below.

This argument extends well beyond the operating room, which matters for any policy that aspires to be comprehensive. Clinical $N_2O$ is used not only in surgical anesthesia but also in obstetric analgesia, pediatric and dental sedation, and prehospital and emergency settings, all of which rely on the same upstream supply infrastructure. Reducing pipeline-related losses therefore captures benefits across the entire spectrum of medical $N_2O$ use, and it does so without removing a tool that retains genuine value in specific contexts: brief pediatric interventions such as intravenous cannulation, dressing changes, and minor wound care, and labor analgesia, where the alternative is often physical restraint or a general anesthetic and where the volumes involved are small. A pipeline-focused strategy is, in this sense, both more effective and less disruptive than prohibition: it eliminates waste while preserving care.

One pediatric indication, however, should be conceded rather than defended. $N_2O$ does not facilitate tranquil mask placement in an unwilling child and does not meaningfully accelerate inhalational induction. Its presence in the fresh gas precludes adequate preoxygenation — a narrow margin to surrender in patients whose smaller functional residual capacity, higher oxygen consumption, and greater risk of laryngospasm already shorten the time to desaturation.[12] Practice has moved accordingly: a quality improvement program spanning 33,285 inhalational inductions reduced $N_2O$ use from 80% to below 20% and cut maximum fresh gas flows by 28%, with induction times and behavior unchanged.[13] That is what one would expect if $N_2O$'s contribution to mask induction were largely a matter of tradition rather than pharmacology. Abandoning it in this setting costs nothing clinically and should be uncontroversial — and it was achieved through education and changed defaults rather than prohibition, which is precisely the point.

The case for prioritizing infrastructure over prohibition is strengthened by what we know about how Brazilian anesthesiologists already use the gas. A recent nationwide cross-sectional survey conducted within a large private hospital network and published in this journal captured responses from 941 of 1,238 eligible anesthesiologists (a 76% response rate).[14] Clinical $N_2O$ use in adults is already low and declining: 59.3% reported never using it in adult patients, and only 14.5% reported frequent or constant use.[14] Pediatric practice differed — 38.9% reported frequent or constant use in children, compared with 27.5% who never used it — reflecting the historical role of $N_2O$ in inhalational induction, a practice that the pediatric evidence no longer supports.[14] Eighty percent reported that their use had fallen over their careers, indicating a substantial behavioral shift already underway without any categorical ban, while 34.5% were unaware of $N_2O$'s contribution to climate change.[14]

Two implications follow. First, because the clinical trend is already moving in the right direction, the marginal return on a formal ban on adult use would be modest. Second, the persistence of pediatric use calls for a more discriminating response than a ban. When that use is inhalational induction, it should be discontinued on clinical as well as environmental grounds, and the international trend is already sharply downward.[12,13] When it is brief procedural analgesia — cannulation, dressing changes, minor wound care — or labor analgesia, it reflects genuine clinical value: precisely the low-volume, high-justification use that a portable-cylinder model is designed to preserve. A policy that decommissions the pipeline while retaining cylinder-based $N_2O$ for those indications aligns environmental and clinical interests rather than setting them against one another and is far more likely to be accepted by the clinicians who must implement it.

If clinical $N_2O$ emissions are small, where should comparable effort go? Comparisons with other sources are useful for prioritization, even if they are not, in themselves, a license for inaction. A single intercontinental flight emits on the order of one to three tonnes of $CO_2e$ per passenger, and commercial aviation as a whole contributes roughly 2.5% of global $CO_2$ emissions and more of total warming once non-$CO_2$ effects are included.[15] The point is not that anesthetic emissions are acceptable because others are larger — that would be the very error of false exemption we are warning against — but that a profession serious about its footprint should sequence its actions by tractable magnitude rather than by symbolic visibility. That sequencing points inward as well as outward. Within our own operating rooms, single-use plastics, over-specified procedure packs, unrecycled packaging and the waste stream generated by every case are substantial, largely unmeasured and entirely within our control, and reducing them requires no trade-off against clinical care at all.[16]

On that logic, several interventions deliver substantially more benefit than withdrawing $N_2O$ from the adult formulary. The first is decommissioning central $N_2O$ pipelines and replacing them with closed portable cylinders, which addresses the dominant source of loss. The second is eliminating desflurane, whose global warming potential of approximately 2,540 is roughly an order of magnitude greater than that of $N_2O$;[8,17] many departments removed it from the formulary years ago with no clinical detriment, and doing so eliminates a large per-case emission at almost no cost, given the availability of sevoflurane and total intravenous anesthesia. The third is optimizing fresh gas flow and expanding low-flow and intravenous techniques, since volatile-agent consumption scales directly with fresh gas flow, halving the flow halves the emission, and reductions of one-half to three-quarters are achievable without changing agent.[17] The fourth is routine integrity checks of portable cylinder valves and connections, paired with institutional clinician-education programs that the Brazilian quality-improvement literature has shown to be effective.[11] None of these requires removing a clinically useful agent from the formulary.

The desflurane phase-out is instructive because it illustrates what proportionate regulation looks like. Scotland removed desflurane from routine use in 2023, the first health system to do so; NHS England completed its decommissioning in 2024; and the European Union's revised regulation on fluorinated greenhouse gases prohibits routine desflurane use from January 2026, with an explicit exemption when no other agent is medically suitable.[18] Regulators acted because the magnitude of a single agent's emissions justified it: desflurane's per-case footprint is large, its clinical advantages are narrow, and substitutes are readily available — the

conditions under which prohibition is the proportionate instrument. The modest clinical contribution of $N_2O$ does not approach that threshold, and the structure of its waste — overwhelmingly pre-utilization pipeline loss rather than clinical administration - points to a different remedy. To treat $N_2O$ as though it were desflurane is to misapply a sound precedent.

A thoughtful case can still be made for elimination on grounds other than tonnage: that a clean break is administratively simpler than maintaining a cylinder supply for a few indications; that the precautionary principle favors removing a potent, ozone-depleting gas wherever feasible; and that the symbolic value of the specialty renouncing $N_2O$ might catalyze wider change. These are not trivial points. But simplicity that removes a valuable option for children and laboring women is a false economy; precaution is better served by eliminating the large fraction of procured gas now vented unused rather than the small fraction delivered to patients; and symbolism that diverts attention from larger sources is, in the end, a cost. The most credible signal the specialty can send is not the renunciation of a single agent but a demonstrated willingness to follow the evidence wherever the emissions actually are.

The equity dimension reinforces the same conclusion. In resource-constrained systems, including much of the public sector in low- and middle-income countries, managerial attention, capital, and clinical goodwill for environmental work are scarce, and the opportunity cost of misdirected effort is correspondingly high. Brazil, with one of the world's largest public health systems and a sophisticated private sector, is well positioned to model a proportionate approach that delivers large, verifiable reductions without rationing care and to export it to other middle-income systems facing the same constraints.

The next step for the anesthesia community is to align local interventions with evidence on where emissions originate. For $N_2O$, that evidence is unusually clear and actionable: the dominant hospital source is a leaking supply infrastructure; the dominant global source is agriculture; and clinical administration — already declining in adult practice and clinically valuable in selected pediatric and obstetric use — is a small contributor by every available measure. Environmental leadership in our specialty will not come from the categorical removal of an agent that retains specific clinical value, but from priority-ranked action that addresses the largest and most tractable sources first: decommissioning pipelines, eliminating desflurane, lowering fresh gas flows, and educating clinicians. Proportionality is not an excuse for inaction; it is the discipline that makes our action count.

## Data availability statement

No new data were generated or analyzed for this Editorial. All data discussed are derived from the cited publications.

## Funding

This editorial received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author(s) used Claude (Anthropic) to generate Figure 1. The author(s) reviewed and edited the output as needed and take full responsibility for the content of the published article.

## Declaration of competing interest

BSvUS is part-funded by the Stan Perron Charitable Foundation and through a National Health and Medical Research Council Investigator Grant (2009322). VCQ is supported by a Research Productivity Fellowship (Bolsa de Produtividade em Pesquisa, Level C) from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (Brazilian National Council for Scientific and Technological Development, CNPq). VCQ and RVC are associate editors of the Brazilian Journal of Anesthesiology.

## Editor

Liana Azi

## References

1. Myles PS, Leslie K, Chan MTV, et al. The safety of addition of nitrous oxide to general anaesthesia in at-risk patients having major non-cardiac surgery (ENIGMA-II): a randomised, single-blind trial. Lancet. 2014;384:1446—54.
2. McGain F, Muret J, Lawson C, Sherman JD. Environmental sustainability in anaesthesia and critical care. Br J Anaesth. 2020;125:680—92.
3. Lenzen M, Malik A, Li M, et al. The environmental footprint of health care: a global assessment. Lancet Planet Health. 2020;4:e271—9.
4. Tian H, Pan N, Thompson RL, et al. Global nitrous oxide budget (1980—2020). Earth Syst Sci Data. 2024;16:2543—604.
5. Charles A, Rochette P, Whalen JK, Angers DA, Chantigny MH, Bertrand N. Global nitrous oxide emission factors from agricultural soils after addition of organic amendments: A meta-analysis. Agric Ecosyst Environ. 2017;236:88—98.
6. Hou D, Meng X, Qin M, et al. Nitrous oxide (N2O) emission characteristics of farmland (rice, wheat, and maize) based on different fertilization strategies. PLoS One. 2024;19:e0305385.
7. Thomas MAF, Ward CJ, Sinnott ME, et al. Nitrous Oxide Manifold and Other Reduction of Emissions (NoMoreGas): a multicentre observational study evaluating pre-utilisation loss of nitrous oxide. Br J Anaesth. 2024;133:1427—34.
8. Intergovernmental Panel on Climate Change (IPCC). Climate Change 2022 - Mitigation of Climate Change. Intergovernmental Panel on Climate Change (IPCC), editor. Cambridge University Press; 2023.
9. Chesebro BB, Gandhi S. Mitigating the systemic loss of nitrous oxide: a narrative review and data-driven practice analysis. Br J Anaesth. 2024;133:1413—8.
10. Liu Y, Lee-Archer P, Sheridan NM, Seglenieks R, McGain F, Eley VA. Nitrous Oxide Use in Australian Health Care: Strategies to Reduce the Climate Impact. Anesth Analg. 2023;137:819—29.

11. da Silva LM, de Araújo ACLF, Defácio L, et al. Reduction in anesthesia-related nitrous oxide consumption and environmental footprint via a quality improvement initiative at two quaternary hospitals in São Paulo, Brazil. Can J Anaesth. 2025;72:1249—59.
12. Gordon DW, Chatterjee D, McGain F. It's time to stop using nitrous oxide for pediatric mask induction. Paediatr Anaesth. 2024;34:104—7.
13. Isserman RS, Yuan I, Elliott EM, et al. Reducing the environmental impact of mask inductions in children: A quality improvement report. Paediatr Anaesth. 2023; 33:728—35.
14. da Silva LM, Nersessian RSF, Araújo ACLF, et al. Changing patterns of nitrous oxide use and environmental awareness among Brazilian anesthesiologists: a nationwide cross-sectional survey. Braz J Anesthesiol. 2026;76:844739.
15. Lee DS, Fahey DW, Skowron A, et al. The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. Atmos Environ. 2021;244:117834.
16. Sampaio JBP, Quintão VC, Teixeira FP, de Souza GD, Simões CM, Carmona MJC. Recycling and sustainability in anaesthesia: a Brazilian national survey. Br J Anaesth. 2024; 133:1380—2.
17. Yamakage M. Sustainable anesthesiology: evidence-based strategies for environmental stewardship in perioperative care. J Anesth. 2026;40:607—14.
18. European Parliament and Council of the European Union. Regulation (EU) 2024/573 of the European Parliament and of the Council of 7 February 2024 on fluorinated greenhouse gases. Off J Eur Union. 2024;L 2024/573.



Ricardo Vieira Carlos [a,b,*], Britta Regli von Ungern-Sternberg [c,d,e,f], Vinícius Caldeira Quintão [a,g,h], Paulo Sergio Mateus Marcelino Serzedo [b], Marcelo Luis Abramides Torres [a]

[a] *Universidade de São Paulo (USP), Faculdade de Medicina (FM), Discipline of Anesthesiology, São Paulo, SP, Brazil*
[b] *Santa Casa de Misericórdia de Ribeirão Preto, Ribeirão Preto, SP, Brazil*
[c] *Child and Adolescent Health Service, Perth Children's Hospital, Department of Anaesthesia and Pain Medicine, Nedlands, Western Australia, Australia*
[d] *The University of Western Australia, Institute for Paediatric Perioperative Excellence, Anaesthesia and Pain Management Program, Perth, Western Australia, Australia*
[e] *Perioperative Medicine Team, The Kids Research Institute Australia, Nedlands, Western Australia, Australia*
[f] *Division of Emergency Medicine, Anaesthesia and Pain Medicine, The University of Western Australia, Perth, Western Australia, Australia*
[g] *Universidade de São Paulo (USP), Faculdade de Medicina (FM), Department of Pediatrics, São Paulo, SP, Brazil*
[h] *Universidade de São Paulo (USP), Faculdade de Medicina (FM), Hospital das Clínicas (HC), Instituto da Criança e do Adolescente, São Paulo, SP, Brazil*

[*] Corresponding author.
*E-mail:* ricardo.vieira@hc.fm.usp.br (R.V. Carlos).


Available online 13 October 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



EDITORIAL

# Videolaryngoscopy in trauma and cervical immobilization: what should guide our first attempt?



Two recent editorials in the Brazilian Journal of Anesthesiology have addressed some practical concerns regarding intubation: how to describe the view obtained with videolaryngoscopy, and why a better image does not replace training and clinical judgement.[1,2] These concerns are particularly relevant when the cervical movement is restricted. The anesthesiologist must obtain a useful view, advance the tracheal tube, and particularly avoid unnecessary movement without prolonging an unsuccessful attempt. These are related tasks, but the success in one does not establish the success in the others.

These outcomes should not be conflated. Videolaryngoscopy may improve intubation success, but whether it also reduces cervical movement and, more importantly, neurological deterioration remains less certain. Current evidence supports its use for the first question more clearly than for the other two. The 2024 multisociety guideline recommends videolaryngoscopy, where possible, in suspected or confirmed cervical spine injury, but cannot recommend a particular device or blade. Importantly, the guideline also states that there are no data showing a greater risk of secondary spinal cord injury with direct laryngoscopy.[3] So, the preference for videolaryngoscopy should not be presented as a proof of spinal protection.

There is now directly relevant randomized evidence in trauma. In the preplanned trauma analysis of the DEVICE trial, published in 2026, first-attempt intubation succeeded in 151/171 patients allocated to videolaryngoscopy and 114/167 allocated to direct laryngoscopy: 88% versus 68%, an absolute difference of 20 percentage points (95% confidence interval, 11-29). Even so, the severe complications during intubation and in-hospital outcomes did not differ significantly.[4] This finding supports videolaryngoscopy for emergency intubation after trauma, but it does not establish a mortality benefit or a neurological protection in unstable cervical injury.

This result is also consistent with the National Emergency Airway Registry analysis of 4,449 trauma intubations, which reported first-pass success of 90% with video and 79% with direct laryngoscopy. The association persisted after propensity matching, although residual confounding could be possible.[5] Earlier results were less favorable. Yeatts and colleagues found no improvement in first-pass success or overall survival with GlideScope, and intubation took longer. A retrospectively identified severe head injury subgroup had more hypoxemia and higher mortality with GlideScope; this remains a cautionary, post hoc finding, not an established contraindication.[6] Taken together, these findings favor videolaryngoscopy, but they also remind us that the device alone does not determine the success of an intubation.

The presence of a cervical collar in those patients, however, is not a diagnosis of cervical instability. Nor does applying a collar to an elective surgical patient reproduce an emergency trauma airway. In the network meta-analysis of 80 trials involving 8,039 participants, several devices improved first-pass success relative to Macintosh laryngoscopy. The authors nevertheless called for more studies using manual in-line stabilization rather than collar-based models.[7] Results obtained with a complete collar, a removed anterior piece, or manual stabilization should not be treated as if they were interchangeable.

In this issue, the two new BJAN studies help clarify different parts of this problem. Nguyen and colleagues compared GlideRite rigid and Flex-it articulated stylets with a hyperangulated GlideScope blade in selected elective patients wearing a semi-rigid collar, with the head fixed to the operating table. Among 75 analyzed patients, median intubation duration was 49 seconds with GlideRite and 88 seconds with Flex-it. Protocol-defined failure occurred in 2/37 and 15/38 patients, respectively.[8] The comparison was between stylets, not between video and direct laryngoscopy. Also, they measured neither cervical movement nor neurological outcomes. Here, the practical lesson is that the stylet is part of the intubation technique, not a minor accessory. In this trial, an articulated design did not improve tube delivery. The magnitude of the failure difference deserves attention, but the trial was not powered specifically for that outcome, and unequal familiarity with the stylets may have influenced

https://doi.org/10.1016/j.bjane.2026.844844
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

performance.[8] These failure rates should not be transported to trauma patients or interpreted as a universal ranking of stylets. But they do justify training with the precise blade, tube and stylet combination that will be used clinically. In the second manuscript in this issue, Doan and colleagues studied 56 adults undergoing elective cervical spine surgery, comparing hyperangulated ClearVue videolaryngoscopy with Macintosh laryngoscopy under manual in-line stabilization, without a collar. Mean measured C0-C5 angular change was 19.71° versus 26.32°, a difference of 6.61° (95% confidence interval, 2.91-10.31). The C0-C1 difference was also significant, but differences at the other individual segments were not. First-attempt success did not differ significantly. Some participants were classified as having cervical trauma, but this was not an emergency trial in patients with documented unstable injuries; acute trauma with established spinal cord injury was excluded. The study provides useful biomechanical evidence, but its clinical boundaries are important. Images were obtained in the neutral position and at the best glottic view, so they need not capture peak movement during tube advancement. Also, neurological outcomes were not measured.[9]

The movement reduction is not a uniform property of videolaryngoscopy, as Robitaille and colleagues found better glottic visualization with GlideScope but no significant reduction in segmental movement under manual stabilization in 20 patients without cervical pathology.[10] Different blades, measurement methods and patient populations matter. A difference in degrees cannot yet be translated into an established reduction in spinal cord injury risk. The collar management also deserves more attention than a simple instruction to "keep the neck still". The 2024 guideline recommends removing a rigid or a semi-rigid collar during intubation attempts, most easily by removing its anterior part; this is a weak recommendation based on low-grade evidence.[3] In 102 elective patients without cervical spine disease, Jo and colleagues compared Macintosh-style videolaryngoscopy with the posterior collar piece alone versus both pieces. Median intubation time was shorter without the anterior piece, 23 versus 33 seconds, while mean between-group motion differences were approximately one degree and not statistically significant.[11] The study therefore updates intubating conditions in this model, but it cannot establish the neurological safety of collar removal in an unstable cervical spine.

Flexible bronchoscopic intubation should not be dismissed based on collar studies. Choi and colleagues found higher first-attempt success with Macintosh-style videolaryngoscopy than with flexible fibrescopy in 330 elective cervical surgery patients wearing a collar to simulate difficulty.[12] That comparison does not answer which awake technique is preferable for a cooperative patient with known instability. In a separate randomized trial of 46 patients with upper cervical instability, Dutta and colleagues found greater C1-C2 movement with awake McGrath intubation than with awake fibreoptic intubation. Neither group developed a new motor deficit, but the small sample cannot establish equivalent neurological safety.[13] The decision about awake intubation must remain distinct from the choice of laryngoscope after induction.

For Brazilian anesthesiologists, the immediate implication is very practical: a videolaryngoscope intended for trauma should not be a device used only after direct laryngoscopy has failed! Regular use and training should include restricted mouth opening, the chosen stylet, suction, tube advancement and an agreed rescue plan. The guideline explicitly recommends regular videolaryngoscopy training with cervical immobilization and consideration of a stylet or bougie.[3]

Further studies should move beyond reporting only the best glottic view. They should distinguish acute trauma from elective cervical disease, describe injury stability and the exact immobilization technique, and record blade geometry, operator experience, first-attempt success, and physiological complications. Where cervical motion is measured, its timing and segmental distribution should be explicit. Neurological assessments are needed, although uncommon adverse events will require larger multicenter studies. The useful clinical objective is a successful, well-prepared first attempt with no unnecessary movement, and not the best image at any cost.

## Data availability statement

No new data were created or analyzed in this study. Data sharing is not applicable to this article.

## IA disclosure

During the preparation of this manuscript, the author used Grammarly solely for English-language editing, including grammar, spelling, style, and readability. The author reviewed and approved all suggested changes and take full responsibility for the final content of the manuscript.

## Conflicts of interest

The author declares no conflicts of interest.

## Editor

Liana Azi

## References

1. Perin D, Malito ML, Neto MA. Imaging classification in videolaryngoscopy: are we on the right track? Braz J Anesthesiol. 2025;75:844671.
2. Perin D, Alves RL. Airway Management: why training, judgement, and competence remain the most important technologies. Braz J Anesthesiol. 2026;76:844781.
3. Wiles MD, Iliff HA, Brooks K, et al. Airway management in patients with suspected or confirmed cervical spine injury. Anaesthesia. 2024;79:856–68.
4. Trent SA, Schauer SG, Prekker ME, et al. Video versus direct laryngoscopy for tracheal intubation in trauma: A secondary analysis of the DEVICE trial. J Trauma Acute Care Surg. 2026;101:359–65.
5. Trent SA, Kaji AH, Carlson JN, McCormick T, Haukoos JS, Brown CA. Video Laryngoscopy Is Associated With First-Pass Success in

Emergency Department Intubations for Trauma Patients: A Propensity Score Matched Analysis of the National Emergency Airway Registry. Ann Emerg Med. 2021;78:708–19.
6. Yeatts DJ, Dutton RP, Hu PF, et al. Effect of video laryngoscopy on trauma patient survival. J Trauma Acute Care Surg. 2013;75:212–9.
7. Singleton BN, Morris FK, Yet B, Buggy DJ, Perkins ZB. Effectiveness of intubation devices in patients with cervical spine immobilisation: a systematic review and network meta-analysis. Br J Anaesth. 2021;126:1055–66.
8. Nguyen JV, Zhu E, Dababneh S, et al. GlideRite® rigid stylet vs Flex-it® articulated stylet for intubation with videolaryngoscopy in patients with semi-rigid cervical collar: a randomized trial. Braz J Anesthesiol. 2026:844836. https://doi.org/10.1016/j.bjane.2026.844836.
9. Doan HT, Nguyen TT, Nguyen ATH, Hoang DG. Cervical spine motion during tracheal intubation with videolaryngoscopy versus Macintosh laryngoscopy in adults undergoing cervical spine surgery: a randomized controlled trial. Braz J Anesthesiol. 2026;76:844847.
10. Robitaille A, Williams SR, Tremblay MH, Guilbert F, Thériault M, Drolet P. Cervical Spine Motion During Tracheal Intubation with Manual In-Line Stabilization: Direct Laryngoscopy versus GlideScope® Videolaryngoscopy. Anesth Analg. 2008;106:935–41.
11. Jo WY, Choi JH, Kim J, Shin KW, Choi S, Park HP, et al. Cervical spine motion during videolaryngoscopic intubation using a Macintosh-style blade with and without the anterior piece of a cervical collar: a randomized controlled trial. Can J Anesth/Journal Canadien D'anesthésie. 2025;72:142–51.
12. Choi S, Yoo HK, Shin KW, et al. Videolaryngoscopy vs. flexible fibrescopy for tracheal intubation in patients with cervical spine immobilisation: a randomised controlled trial. Anaesthesia. 2023;78:970–8.
13. Dutta K, Sriganesh K, Chakrabarti D, Pruthi N, Reddy M. Cervical Spine Movement During Awake Orotracheal Intubation With Fiberoptic Scope and McGrath Videolaryngoscope in Patients Undergoing Surgery for Cervical Spine Instability: A Randomized Control Trial. J Neurosurg Anesthesiol. 2020;32:249–55.

Liana M.T.A. Azi *
*Universidade Federal da Bahia, Hospital Universitário Professor Edgard Santos, Department of Anesthesiology, Salvador, BA, Brazil*

* Corresponding author.
*E-mail:* liana.araujo@ufba.br

Available online 13 October 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



ORIGINAL INVESTIGATION

# Cervical spine motion during tracheal intubation with videolaryngoscopy versus Macintosh laryngoscopy in adults undergoing cervical spine surgery: a randomized controlled trial

**Hieu Trung Doan** [a], **Thang Toan Nguyen** [a,b,*], **Anh Thi Hong Nguyen** [a], **Du Gia Hoang** [c]

[a] *Bach Mai Hospital, Center for Anesthesia and Surgical Intensive Care, Hanoi, Vietnam*
[b] *Hanoi Medical University, Department of Anesthesia and Critical Care, Hanoi, Vietnam*
[c] *Bach Mai Hospital, Department of Orthopedic and Spine Surgery, Hanoi, Vietnam*

Received 10 April 2026; accepted 5 October 2026
Available online 15 October 2026

**KEYWORDS**
Airway management;
Cervical vertebrae;
Fluoroscopy;
Randomized controlled trial;
Tracheal intubation;
Videolaryngoscopes

## Abstract

*Background:* Excessive cervical spine motion during tracheal intubation may increase neurological risk in individuals with cervical spine pathology. Videolaryngoscopy may reduce cervical motion compared with Macintosh direct laryngoscopy, but any biomechanical reduction should be distinguished from neurological protection.

*Methods:* Adults scheduled for elective cervical spine surgery were randomized 1:1 to intubation with a ClearVue videolaryngoscope (group V) or a Macintosh laryngoscope (group M) under manual in-line stabilization. Lateral fluoroscopy was acquired immediately before laryngoscopy (G0, neutral position) and at best glottic view (G1). The prespecified primary endpoint was directly measured cumulative C0–C5 cervical motion. Cervical spine motion (ΔG) was also calculated at C0–C1, C1–C2, C2–C3, C3–C4 and C4–C5. ΔG was expressed in degrees (°).

*Results:* Fifty-six patients were randomized and analyzed (28 per group). Mean (SD) ΔG at C0–C5 was 19.71 (4.75) in group V and 26.32 (8.46) in group M (mean difference [M–V], 6.61; 95% CI 2.91–10.31; $p < 0.001$). At C0–C1, ΔG was 2.68 (2.53) vs. 4.68 (2.42) (difference = 2.00; 95% CI 0.67–3.33; $p = 0.004$). First-attempt success was 89.3% (25/28) vs. 78.6% (22/28) ($p = 0.47$). Among patients intubated successfully on the first attempt, intubation time was shorter in group V (21.75 ± 3.64 s) than group M (31.71 ± 21.25 s).

*Conclusion:* ClearVue videolaryngoscopy reduced fluoroscopically measured upper cervical and directly measured cumulative C0–C5 motion compared with Macintosh direct laryngoscopy.

* Corresponding author.
*E-mail:* nguyentoanthang@hmu.edu.vn (T.T. Nguyen).

https://doi.org/10.1016/j.bjane.2026.844847
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

These findings should be interpreted as biomechanical surrogate outcomes, not as evidence of neurological protection.

© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

Tracheal intubation in individuals with cervical spine pathology is challenging because laryngoscopy can produce cervical spine extension and translation, particularly at the upper cervical segments.[1-3] Although the motion threshold associated with neurological deterioration is unknown, minimizing cervical spine motion is widely advocated in both traumatic and non-traumatic cervical spine disease.[1] In a cadaver model with a destabilized cervical segment, common airway maneuvers produced measurable motion across the injured level, supporting the rationale for motion-minimizing techniques.[4] Awake flexible bronchoscopic intubation is often recommended when feasible; however, direct laryngoscopy under manual in-line stabilization (MILS) remains common in many operating room and emergency settings.[1]

Videolaryngoscopes may provide an improved glottic view with less lifting force and potentially less upper cervical motion compared with Macintosh direct laryngoscopy, although results vary by device design and immobilization strategy.[5-11] Evidence specific to individual videolaryngoscope designs remains clinically relevant because blade geometry, optical axis, tube-shaping strategy, and operator familiarity can influence both glottic exposure and cervical spine kinematics. We designed this randomized trial to compare cervical spine motion measured by lateral fluoroscopy and intubation performance using a ClearVue videolaryngoscope versus a Macintosh laryngoscope in adults undergoing elective cervical spine surgery with MILS.

We hypothesized that videolaryngoscopy would reduce directly measured cumulative cervical spine motion (ΔG at C0–C5) during laryngoscopy compared with Macintosh direct laryngoscopy.

## Methods

### Study design, ethics, and registration

This prospective, single-center, parallel-group randomized controlled trial was conducted at Bach Mai Hospital. Patient recruitment and enrollment occurred between February 2024 and September 2024. The study protocol had been approved before patient enrollment by the Ethics Committee at Bach Mai Hospital (Decision n° 4665/QĐ-BM, dated August 14, 2023; project code: BM_2023_23). All patients provided written informed consent. The trial was conducted in accordance with the Declaration of Helsinki.

The ethics-approved protocol prespecified the directly measured cumulative C0–C5 cervical spine motion as the primary outcome, the segmental cervical spine motion and intubation performance outcomes as secondary outcomes, and included the sample-size assumptions and statistical analysis plan before enrollment began. No changes were made to the trial design, outcomes, sample-size assumptions, or statistical analysis plan after patient enrollment started. Recruitment ended when the planned sample size was achieved.

The trial was not prospectively registered in a public registry. Retrospective registration was completed in the Australian New Zealand Clinical Trials Registry (ANZCTR; ACTRN12626000389303; registered on March 27, 2026). The full protocol is available from the corresponding author upon reasonable request.

### Patients

Adults aged 18 years or older scheduled for elective cervical spine surgery under general anesthesia were screened. Inclusion criteria were ASA physical status I–II and consent to participate. Exclusion criteria were rapid-sequence induction requirement, intubation without neuromuscular blockade, limited mouth opening ($< 2$ cm), body mass index $> 35$ kg.m$^{-2}$, history of cervical spine or maxillofacial surgery, ankylosing spondylitis, or acute cervical spine trauma with established spinal cord injury.

Cervical pathology was classified preoperatively as traumatic cervical spine injury or non-traumatic pathology based on the surgical diagnosis. Non-traumatic indications included cervical disc herniation, cervical spinal stenosis, and cervical spinal cord tumor.

### Randomization and blinding

Patients were randomized 1:1 to videolaryngoscopy (group V) or Macintosh direct laryngoscopy (group M) using a computer-generated block randomization table created before recruitment by an investigator not involved in patient enrollment. Allocation assignments were placed in sequentially numbered, opaque, sealed envelopes, which were opened immediately before induction after eligibility confirmation and consent. Because the devices differ, the intubating anesthesiologist could not be blinded.

To minimize performance variability, all tracheal intubations were performed by the same study anesthesiologist, who had more than 5 years of anesthesia and intensive care experience and was familiar with both Macintosh direct laryngoscopy and the ClearVue videolaryngoscope. The fluoroscopy operator and the assessor who measured cervical angles were blinded to group assignment. The intubator was not allowed to view fluoroscopy images during intubation.

### Anesthesia, airway devices, and MILS protocol

Standard monitoring included electrocardiography, pulse oximetry, and noninvasive arterial blood pressure. Neuromuscular transmission was monitored with TOFscan (Dräger). Patients

were positioned supine on a firm horizontal table with head and neck in neutral position. After preoxygenation (6 $L.min^{-1}$ for 3–5 minutes), anesthesia was induced with propofol 2 $mg.kg^{-1}$ and fentanyl 3 $\mu g.kg^{-1}$, followed by rocuronium 0.6 $mg.kg^{-1}$. Mask ventilation was performed after apnea; an oral airway was permitted if needed. Laryngoscopy began after confirmation of complete neuromuscular blockade (TOF count 0).

Tracheal intubation was performed using either a ClearVue videolaryngoscope (Infinium Medical, USA) or a Macintosh laryngoscope according to allocation. The ClearVue device used a non-channeled, hyperangulated D-blade (size 3 or 4) with a distal camera near the blade tip. In group V, a malleable stylet was inserted into the tracheal tube and precurved to approximately 60 degrees to match the hyperangulated blade curvature and optical axis. In group M, a standard Macintosh blade was used, and the tracheal tube was shaped with a malleable stylet into a hockey-stick configuration with an approximately 35-degree distal bend. These device-specific features were considered when interpreting generalizability to other videolaryngoscopes, particularly Macintosh-style or differently angulated devices.

Two dedicated anesthesia assistants from the spine operating room team were trained before study implementation and were responsible for applying MILS. During laryngoscopy, the assistant stood at the head of the patient beside the intubating anesthesiologist, held the mastoid processes bilaterally with the fingertips, and supported the occiput with the palms without applying traction. The same standardized MILS technique was maintained throughout induction and laryngoscopy. External laryngeal manipulation was permitted when the glottic opening was < 50% visible, but this maneuver was not systematically recorded as a separate outcome.

Intubation was limited to 120 seconds per attempt. If oxygen saturation decreased to < 90% or the time limit was reached, mask ventilation was resumed for 1-minute before reattempting intubation. If intubation failed, the assigned device was abandoned, and an alternative rescue technique was used according to local difficult-airway practice. Successful tracheal intubation was confirmed by capnography and bilateral chest auscultation.

## Fluoroscopy acquisition and cervical angle measurement

Lateral fluoroscopic images of the skull base and cervical spine were obtained at two time points: immediately before laryngoscopy with the head-neck in neutral position (G0), and at the point of best glottic view (G1). Fluoroscopy adhered to ALARA principles (tight collimation, shortest feasible exposure, and standard shielding) to minimize any incremental radiation. After confirming alignment, the operating table and fluoroscopy equipment were not moved during image acquisition.

Bone reference lines were defined using two anatomic landmarks per structure (Fig. 1), following previously described fluoroscopy-based techniques.[3,6,8,11] Reference lines were drawn for C0 (McGregor line, between the posterior margin of the hard palate and the opisthion), for C1 (a line through the anterior and posterior arches of the atlas),

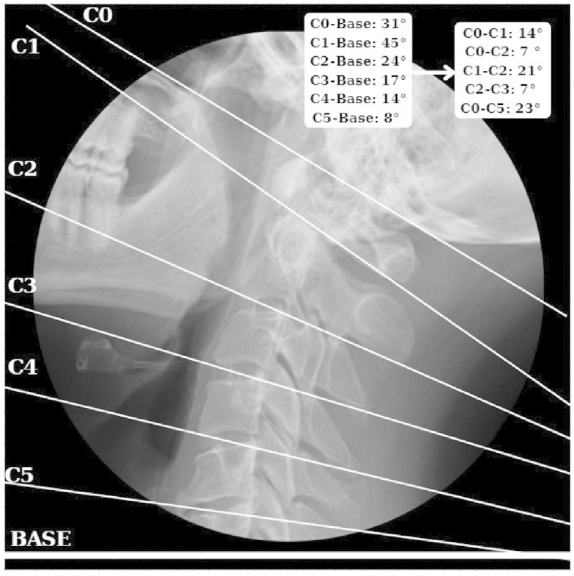



**Figure 1** Reference lines and angle measurements on lateral fluoroscopy used to calculate intervertebral and directly measured cumulative cervical motion (ΔG).

and for C2–C5 (tangential lines through the anterior and posterior endplates of each vertebral body). For each image, the absolute angle between each reference line and a horizontal baseline was measured.

Intervertebral segmental angles (C0–C1, C1–C2, C2–C3, C3–C4, and C4–C5) were calculated from adjacent reference lines. Segmental motion was reported as the absolute angular change between G0 and G1 for each segment. The prespecified cumulative C0–C5 outcome was calculated directly at the patient level as the absolute change in the global C0–C5 angle between G0 and G1, using the C0 and C5 reference lines. Therefore, the C0–C5 value was not derived by summing group-level mean segmental values; the latter can differ from the direct composite measurement because adjacent segments may move in different directions and because absolute segmental magnitudes do not preserve directionality.

All measurements were performed by a blinded assessor. Each angle was measured twice and averaged to reduce measurement error. Formal intra-observer reliability statistics (e.g., intraclass correlation coefficient) were not prespecified and were not calculated; this is acknowledged as a limitation.

## Outcomes

The prespecified primary outcome was directly measured cumulative cervical spine motion (ΔG) at C0–C5. Secondary outcomes included segmental ΔG at C0–C1, C1–C2, C2–C3, C3–C4, and C4–C5; first-attempt intubation success; intubation time; glottic view assessed using the Cormack-Lehane grade; and perioperative airway-related adverse events.

First-attempt success was defined as successful tracheal tube placement on the first laryngoscopy attempt, confirmed by continuous capnography. An intubation attempt was defined as insertion of the laryngoscope into the mouth;

removal of the device and/or resumption of mask ventilation marked the end of an attempt. Intubation time was measured in seconds from mouth opening/laryngoscope insertion to confirmation of successful tracheal intubation by capnography. Glottic view was recorded at the point of best glottic view. Prespecified safety outcomes included oxygen desaturation < 90%, dental injury, mucosal bleeding, and perioperative airway-related complications.

## Sample size

Sample size was informed by prior cinefluoroscopic studies comparing airway devices,[3,6] and by the sample-size approach used in recent cervical motion trials.[11] Although the prespecified primary endpoint was the directly measured cumulative C0–C5 motion, the planning assumptions were based on C0–C1 motion because the occipito-atlanto-axial complex is clinically important, highly mobile during laryngoscopy, and better characterized in prior radiographic studies than composite C0–C5 motion. A between-group C0–C1 difference of 2.0 degrees was used for sample-size planning based on prior radiographic studies, with an assumed SD of 2.5 degrees. With a two-sided alpha = 0.05 and 80% power, 25 patients per group were required. We planned to recruit 28 patients per group (total 56), allowing for approximately 10% attrition or missing fluoroscopy images. The primary C0–C5 result is interpreted using the observed mean difference and 95% confidence interval rather than post-hoc power.

## Statistical analysis

Analyses were performed using SPSS (version 20.0) and were conducted on an intention-to-treat basis. Continuous variables are presented as mean ± SD and categorical variables as counts (%). Between-group comparisons of continuous outcomes used Welch's *t*-test (or Mann-Whitney *U*-test if non-normal). Categorical outcomes were compared using Fisher's exact test. Effect sizes are reported as mean differences with 95% confidence intervals for continuous outcomes and odds ratios with 95% confidence intervals for categorical outcomes. A two-sided p-value < 0.05 was considered statistically significant.

Given multiple cervical segments were evaluated, the directly measured cumulative C0–C5 motion was treated as the prespecified primary endpoint. Segmental analyses (C0–C1, C1–C2, C2–C3, C3–C4, and C4–C5) were secondary, exploratory and interpreted accordingly. No formal adjustment for multiple comparisons was applied. The statistical tables and SPSS output were independently checked by a second author.

# Results

Between February 2024 and September 2024, 62 patients were assessed for eligibility. Six were excluded (ankylosing spondylitis, n = 2; acute cervical spine trauma with established spinal cord injury, n = 2; obesity, n = 1; declined participation, n = 1). Fifty-six patients were randomized, allocated, and analyzed (28 per group). No patients were lost to follow-up or excluded from the primary analysis (Fig. 2). Recruitment ended when the target sample size was achieved.

Baseline demographic and airway characteristics were similar between groups (Table 1). Traumatic cervical pathology was present in 9 patients (32.1%) in each group. Non-traumatic pathology was present in 19 patients (67.9%) in each group and included cervical disc herniation, cervical spinal stenosis, and cervical spinal cord tumor.

## Cervical spine motion

Videolaryngoscopy was associated with lower directly measured cumulative C0–C5 motion than Macintosh direct laryngoscopy (Table 2). Mean difference (M–V) at C0–C5 was 6.61 (95% CI 2.91–10.31; p < 0.001). At C0–C1, the mean difference was 2.00 (95% CI 0.67–3.33; p = 0.004). Segmental differences at C1–C2, C2–C3, C3–C4, and C4–C5 were not statistically significant; C4–C5 motion was 3.21 ± 2.45 in group V and 3.66 ± 2.96 in group M (mean difference [M–V], 0.45; 95% CI -1.00 to 1.91; p = 0.54) (Table 2).

C1–C2 showed the largest segmental motion in both groups. This pattern was verified during data review and is consistent with the biomechanical predominance of the upper cervical spine during laryngoscopy, especially around the occipito-atlanto-axial complex, and with prior fluoroscopic studies showing that airway instrumentation concentrates motion in upper cervical segments.[3,4,6,8,13]

## Intubation performance and safety

First-attempt success and glottic view favored videolaryngoscopy numerically but did not reach statistical significance (Table 3). Among patients intubated successfully on the first attempt, intubation time was shorter in group V than group M (21.75 ± 3.64 seconds [range 17–29] vs. 31.71 ± 21.25 seconds [range 14–106]; mean difference [M–V], 9.96 seconds; 95% CI 0.45–19.47; p = 0.041). No rescue airway technique, protocol deviation, or study-related adverse event was observed (including desaturation < 90%, dental injury, or mucosal bleeding).

# Discussion

In this randomized trial in adults undergoing elective cervical spine surgery with MILS, ClearVue videolaryngoscopy reduced both directly measured cumulative cervical motion (C0–C5) and upper cervical motion (C0–C1) at the time of best glottic view compared with Macintosh direct laryngoscopy. The between-group difference in the primary outcome was 6.61 (95% CI 2.91–10.31), and the C0–C1 difference was 2.00 (95% CI 0.67–3.33). These results support the biomechanical premise that videolaryngoscopy can reduce cervical extension required to obtain an adequate glottic view.

The finding should be interpreted in the context of prior fluoroscopic and radiographic studies, in which indirect or optical laryngoscopic devices generally reduced upper cervical extension or composite cervical motion compared with Macintosh laryngoscopy under immobilization or MILS.[6,8,10,12,13-15] The present study adds device-specific evidence for the ClearVue hyperangulated videolaryngoscope, but its clinical novelty is incremental rather than definitive;

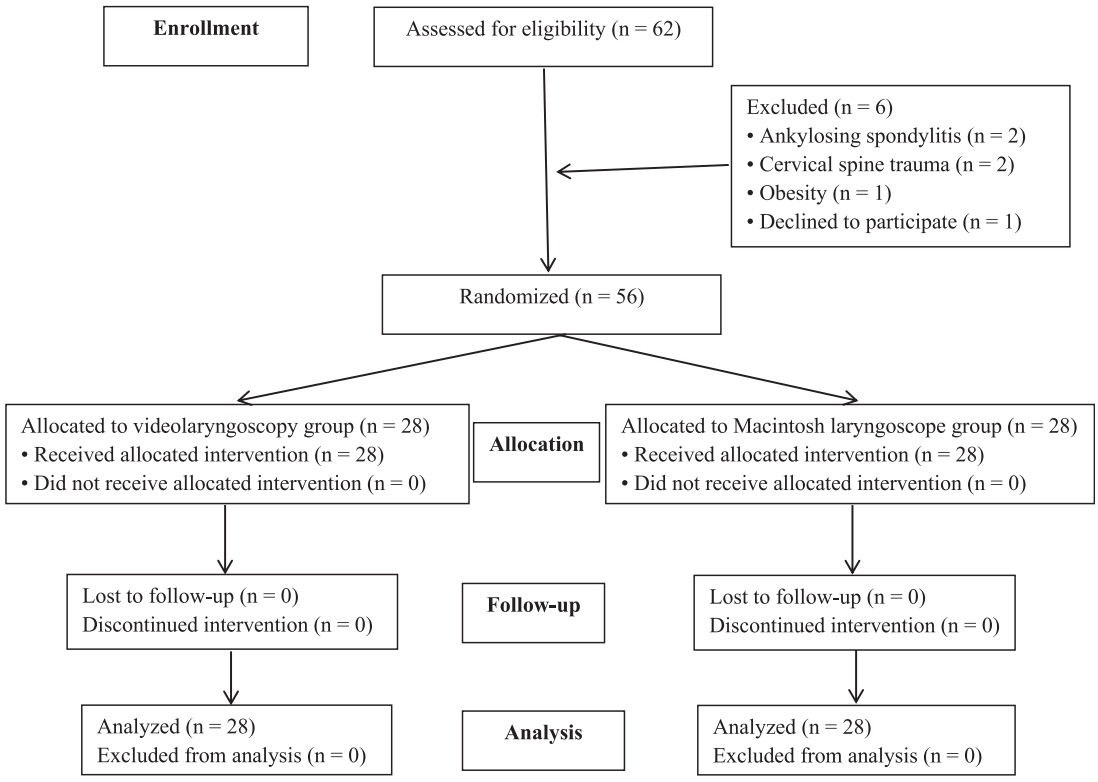



**Figure 2** CONSORT flow diagram of patient recruitment, allocation, follow-up, and analysis.

similar biomechanical patterns have been reported with other videolaryngoscopes and optical devices. Differences among studies likely reflect device geometry, blade curvature, tube-shaping strategy, immobilization method, and operator experience.

The C1–C2 values were larger than adjacent lower cervical segmental values in both groups. This distribution is plausible because the occipito-atlanto-axial complex is the most mobile part of the cervical spine during airway instrumentation, and because laryngoscopic force directed at the tongue base and vallecula can be transmitted primarily to upper cervical segments.[3,4,6,8,13] The addition of C2–C3, C3–C4, C4–C5 data confirms that lower cervical motion was smaller and did not differ significantly between groups, whereas the largest inter-device difference was concentrated at C0–C1 and the directly measured composite C0–C5 outcome. Our findings are consistent with prior fluoroscopic and radiographic studies comparing videolaryngoscopes or optical laryngoscopes with Macintosh laryngoscopy under immobilization or MILS, which generally report reduced upper cervical extension and/or reduced composite cervical motion with indirect laryngoscopy.[6,8,10,12] In contrast, flexible bronchoscopy techniques may minimize cervical motion even further, but can be less practical in time-critical scenarios and require specific expertise and preparation.[1,5,13] A plausible mechanism for reduced motion with videolaryngoscopy is the lower lifting

**Table 1** Baseline characteristics and cervical pathology.

| Variable | Videolaryngoscope (V) (n = 28) | Macintosh (M) (n = 28) |
|---|---|---|
| Age, years | 53.32 ± 15.67 (22–78) | 54.89 ± 13.23 (22–75) |
| Male sex, n (%) | 17 (60.7) | 18 (64.3) |
| Weight, kg | 53.89 ± 7.98 (40–70) | 55.96 ± 8.27 (36–71) |
| Height, cm | 158.00 ± 9.81 (145–177) | 157.03 ± 8.91 (143–175) |
| BMI, $kg.m^{-2}$ | 21.74 ± 3.60 (16.7–30.1) | 23.75 ± 4.65 (16.3–33.2) |
| ASA I/II, n | 10/18 | 9/19 |
| Traumatic cervical pathology, n (%) | 9 (32.1) | 9 (32.1) |
| Non-traumatic cervical pathology, n (%) | 19 (67.9) | 19 (67.9) |
| Mouth opening, cm | 3.66 ± 0.52 (3.0–4.5) | 3.71 ± 0.51 (3.0–4.5) |
| Thyromental distance, cm | 6.58 ± 0.85 (5.0–8.0) | 6.55 ± 0.80 (5.0–8.0) |
| Mallampati I/II/III/IV, n | 4/17/5/2 | 5/18/4/1 |

Non-traumatic cervical pathology included cervical disc herniation, cervical spinal stenosis, and cervical spinal cord tumor. Sex, ASA physical status, and Mallampati class are reported as categorical distributions. Continuous variables are presented as mean ± SD (min-max) and categorical variables as n (%).

**Table 2** Cervical spine motion during tracheal intubation.

| Segment | ΔG, V (mean ± SD) ° | ΔG, M (mean ± SD) ° | Mean difference (M−V), ° (95% CI) | p-value |
|---|---|---|---|---|
| C0–C1 | 2.68 ± 2.53 | 4.68 ± 2.42 | 2.00 (0.67–3.33) | 0.004 |
| C1–C2 | 15.86 ± 7.37 | 16.75 ± 7.89 | 0.89 (-3.20 to 4.98) | 0.66 |
| C2–C3 | 2.14 ± 2.49 | 2.18 ± 2.52 | 0.04 (-1.30 to 1.38) | 0.95 |
| C3–C4 | 1.43 ± 1.71 | 1.68 ± 1.94 | 0.25 (-0.73 to 1.23) | 0.61 |
| C4–C5 | 3.21 ± 2.45 | 3.66 ± 2.96 | 0.45 (-1.00 to 1.91) | 0.54 |
| C0–C5 | 19.71 ± 4.75 | 26.32 ± 8.46 | 6.61 (2.91-10.31) | < 0.001 |

Delta G (ΔG) denotes the change in intervertebral angle between the neutral position immediately before laryngoscopy (G0) and the point of best glottic view (G1) on lateral fluoroscopy. Positive values indicate extension. Between-group comparisons were performed using Welch's *t*-test. The C0–C5 outcome was measured directly at the patient level from the C0 and C5 reference lines and was not calculated by summing group-level mean segmental values.

force required to obtain an adequate glottic view, because less force may translate into less cervical extension and less bending moment at the cervical spine.[7] This biomechanical relationship has been demonstrated experimentally and provides a physiologic rationale for preferential use of videolaryngoscopy when minimizing cervical motion is prioritized.

Contemporary difficult-airway guidance increasingly incorporates videolaryngoscopy into structured airway strategies rather than treating it as a rescue-only device.[16,17] Recent expert guidelines on universal implementation of videolaryngoscopy similarly emphasize equipment standardization, training, availability, and team-level implementation.[18] Therefore, the present results should not be interpreted as a device-only solution. The potential benefit of videolaryngoscopy in cervical spine pathology depends on standardized MILS, adequate neuromuscular blockade, gentle technique, operator expertise, ergonomics, team communication, and early rescue planning.[19]

Although first-attempt success and Cormack-Lehane grade favored videolaryngoscopy numerically, the trial was not powered for intubation performance outcomes. The shorter first-attempt intubation time observed with videolaryngoscopy should also be interpreted cautiously because the analysis included only patients successfully intubated on the first attempt. Nonetheless, the direction of effect is consistent with evidence that videolaryngoscopy can improve glottic visualization in immobilized or anticipated difficult airway conditions.[9,12]

Most importantly, reduced angular cervical motion is a biomechanical surrogate outcome and should not be interpreted as evidence of neurological protection. The threshold of cervical motion associated with neurological deterioration remains unknown, and this study was not powered to detect neurological complications. Device choice should therefore be integrated into a broader movement-minimizing airway strategy rather than used as the sole determinant of safety.

## Limitations

This study has limitations. First, the intubator could not be blinded to the device, introducing potential performance bias, although all intubations were performed by the same experienced anesthesiologist using a standardized protocol. Second, cervical spine motion was derived from two static fluoroscopy time points (neutral position and best glottic view) rather than continuous imaging; therefore, peak motion during tube advancement may not have been captured. Third, formal intra-observer reliability statistics were not prespecified, although each angle was measured twice by a blinded assessor and averaged. Fourth, the trial was not prospectively registered, which may raise concerns about selective reporting despite prior ethics approval, prespecified outcomes, and subsequent retrospective registration in ANZCTR. Fifth, the study enrolled ASA I–II adults undergoing elective cervical spine surgery and excluded severe obesity,

**Table 3** Intubation performance and safety.

| Outcome | V (n = 28) | M (n = 28) | Effect estimate | p-value |
|---|---|---|---|---|
| First-attempt intubation success | 25/28 (89.3%) | 22/28 (78.6%) | OR = 2.27 (95% CI 0.51–10.18); RD 10.7% | 0.47 |
| Cormack-Lehane grade III–IV | 4/28 (14.3%) | 9/28 (32.1%) | OR = 0.35 (95% CI 0.09–1.32) | 0.21 |
| Intubation time among first-attempt successes, seconds (s) | 21.75 ± 3.64 (17–29); n = 25 | 31.71 ± 21.25 (14–106); n = 22 | MD = 9.96s (95% CI 0.45–19.47) | 0.041 |
| Rescue airway technique | 0/28 | 0/28 | Not estimable | – |
| Protocol deviation | 0/28 | 0/28 | Not estimable | – |
| Study-related adverse events | 0/28 | 0/28 | Not estimable | – |

OR, Indicates odds ratio comparing group V vs. group M; RD, Indicates risk difference (V–M); MD, Indicates mean difference (M–V). Fisher's exact test was used for categorical p-values; Welch's *t*-test was used for intubation time.

ankylosing spondylitis, acute cervical spine trauma with established spinal cord injury, and anticipated highly difficult airway situations; therefore, findings may not be generalizable to emergency airway management, acute unstable cervical trauma, established spinal cord injury, severe obesity, high ASA physical status, or physiologically unstable patients. Finally, we did not directly measure laryngoscopic lifting force, hemodynamic responses, sore throat, hoarseness in a structured manner, radiation dose, or patient-centered neurological outcomes.

## Conclusion

In adults undergoing elective cervical spine surgery with MILS, tracheal intubation using the ClearVue videolaryngoscope was associated with reduced fluoroscopically measured upper cervical (C0–C1) and directly measured cumulative C0–C5 motion compared with Macintosh direct laryngoscopy. These findings represent biomechanical surrogate outcomes and should not be interpreted as proof of neurological protection. Future studies should evaluate continuous motion profiles, broader airway populations, and patient-centered neurological outcomes.

## Data availability statement

The de-identified numerical data underlying the reported tables and the statistical output files are available from the corresponding author upon reasonable request after publication, subject to institutional policies, approval of a scientifically sound proposal, and completion of an appropriate data-sharing agreement. Because the dataset comes from a single-center clinical trial with a relatively small sample size, unrestricted public deposition was not undertaken at this stage to minimize the risk of indirect participant identification. The authors will comply with any additional data-sharing requirements requested by the journal.

## Trial registration

Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12626000389303; registered on March 27, 2026. The trial was retrospectively registered.

## Ethics approval

Bach Mai Hospital Ethics Committee approval (Decision n° 4665/QĐ-BM, dated August 14, 2023; project code: BM_2023_23). Written informed consent was obtained from all participants.

## Use of AI and AI-assisted technologies

During the preparation of this revised work, the authors used ChatGPT (OpenAI) solely to improve language clarity, readability, and organization of the response to reviewers. The tool was not used to generate data, perform statistical analyses, or create or alter figures or results. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the integrity of the work.

## Authors' contributions

All authors contributed to the study conception and design. Data collection was performed by the study team. Statistical analysis was performed by an author with responsibility for the analysis plan. The first draft of the manuscript was written by the authors, and all authors critically revised the manuscript and approved the final version.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

None.

## Funding

This research received no external funding.

## Associate Editor

Ana Maria Menezes Caetano

## References

1. Crosby ET, Warltier DC. Airway management in adults after cervical spine trauma. Anesthesiology. 2006;104:1293—318.
2. White 3rd AA, Johnson RM, Panjabi MM, Southwick WO. Biomechanical analysis of clinical stability in the cervical spine. Clin Orthop Relat Res. 1975(109):85—96.
3. Sawin PD, Todd MM, Traynelis VC, et al. Cervical spine motion with direct laryngoscopy and orotracheal intubation: an in vivo cinefluoroscopic study of subjects without cervical abnormality. Anesthesiology. 1996;85:26—36.
4. Brimacombe J, Keller C, Künzel KH, Gaber O, Boehler M, Pühringer F. Cervical spine motion during airway management: a cinefluoroscopic study of the posteriorly destabilized third cervical vertebrae in human cadavers. Anesth Analg. 2000;91:1274—8.
5. Wong DM, Prabhu A, Chakraborty S, Tan G, Massicotte EM, Cooper R. Cervical spine motion during flexible bronchoscopy compared with the Lo-Pro GlideScope®. Br J Anaesth. 2009;102:424—30.
6. Robitaille A, Williams SR, Tremblay MH, Guilbert F, Thériault M, Drolet P. Cervical spine motion during tracheal intubation with manual in-line stabilization: direct laryngoscopy versus GlideScope® videolaryngoscopy. Anesth Analg. 2008;106:935—41.
7. Hindman BJ, Santoni BG, Puttlitz CM, From RP, Todd MM. Intubation biomechanics: laryngoscope force and cervical spine motion during intubation with Macintosh and Airtraq laryngoscopes. Anesthesiology. 2014;121:260—71.

8. Turkstra TP, Pelz DM, Jones PM. Cervical spine motion: a fluoroscopic comparison of the AirTraq laryngoscope versus the Macintosh laryngoscope. Anesthesiology. 2009;111:97—101.
9. Kleine-Brueggeney M, Greif R, Schoettker P, Savoldelli GL, Nabecker S, Theiler LG. Evaluation of six videolaryngoscopes in 720 patients with a simulated difficult airway: a multicentre randomized controlled trial. Br J Anaesth. 2016;116:670—9.
10. Kaur K, Raja R, Kumar P, Singh R, Vashishth S, Singh HD, et al. A comparative study to evaluate the cervical spine movements during laryngoscopy using Macintosh and Airtraq laryngoscopes. J Anaesthesiol Clin Pharmacol. 2024;40:101—7.
11. Jo WY, Choi JH, Kim J, et al. Cervical spine motion during videolaryngoscopic intubation using a Macintosh-style blade with and without the anterior piece of a cervical collar: a randomized controlled trial. Can J Anaesth. 2025;72:142—51.
12. Akbar SHM, Ooi JSM. Comparison between C-MAC video-laryngoscope and Macintosh direct laryngoscope during cervical spine immobilization. Middle East J Anaesthesiol. 2015;23:43—50.
13. Rudolph C, Schneider JP, Wallenborn J, Schaffranietz L. Movement of the upper cervical spine during laryngoscopy: a comparison of the Bonfils intubation fibrescope and the Macintosh laryngoscope. Anaesthesia. 2005;60:668—72.
14. Stier GR, Gabriel CL, Cole DJ. Neurosurgical diseases and trauma of the spine and spinal cord: anesthetic considerations. In: Cottrell JE, Young WL, eds. Cottrell and Young's Neuroanesthesia, 5th ed., Philadelphia (PA): Mosby Elsevier; 2010:343—89.
15. Şahin T, Arslan ZI, Akansel G, et al. Fluoroscopic comparison of cervical spine motion using LMA CTrach, C-MAC videolaryngoscope and Macintosh laryngoscope. Turk J Anaesthesiol Reanim. 2018;46:44—50.
16. Gomez-Rios MA, Sastre JA, Onrubia-Fuertes X, et al. Spanish Society of Anesthesiology, Reanimation and Pain Therapy (SEDAR), Spanish Society of Emergency and Emergency Medicine (SEMES) and Spanish Society of Otolaryngology, Head and Neck Surgery (SEORL-CCC) Guideline for difficult airway management. Part I. Rev Esp Anestesiol Reanim (Engl Ed). 2024;71:171—206.
17. Gomez-Rios MA, Sastre JA, Onrubia-Fuertes X, et al. Spanish Society of Anesthesiology, Reanimation and Pain Therapy (SEDAR), Spanish Society of Emergency and Emergency Medicine (SEMES) and Spanish Society of Otolaryngology, Head and Neck Surgery (SEORL-CCC) Guideline for difficult airway management. Part II. Rev Esp Anestesiol Reanim (Engl Ed). 2024;71:207—47.
18. Gomez-Rios MA, Van Zundert AAJ, McNarry AF, et al. Guidelines on strategies for the universal implementation of videolaryngoscopy. Eur J Anaesthesiol. 2025;42:872—88.
19. Gomez-Rios MA, Michalek P, Gaszynski T, Van Zundert AAJ. Human Factors in Airway Management: Designing Systems for Safer, Team-Based Care. J Clin Med. 2025;14:8850.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



ORIGINAL INVESTIGATION

# GlideRite® rigid stylet vs Flex-it® articulated stylet for intubation with videolaryngoscopy in patients with semi-rigid cervical collar: a randomized trial

John Viet Nguyen [a], Eric Zhu [b,*], Said Dababneh [c], Pierre Drolet [a,c], Marie-Ève Bélanger [a], Rami Issa [a], Valerie Zaphiratos [a], Nadia Godin [a], Jean S. Bussières [d], Mihai Georgescu [a,c], Issam Tanoubi [a,c,e]

[a] *Université de Montréal, CIUSSS de l'Est-de-l'Île-de-Montréal, Maisonneuve-Rosemont Hospital, Department of Anesthesiology and Pain Medicine, Montréal, QC, Canada*
[b] *Université de Montréal, Faculty of Medicine, Montréal, QC, Canada*
[c] *Université de Montréal, Medical Simulation Centre, Centre d'Apprentissage des Attitudes et Habiletés Cliniques (CAAHC), Montréal, QC, Canada*
[d] *Institut Universitaire de Cardiologie et de Pneumologie de Québec (IUCPQ), Department of Anesthesiology and Intensive Care, Québec City, QC, Canada*
[e] *Université de Montréal, Centre de Recherche en Pédagogie de la Santé (CRPS), Montréal, QC, Canada*

Received 10 January 2026; accepted 23 August 2026
Available online 24 September 2026

**KEYWORDS**
Airway management;
Immobilization;
Intubation, intratracheal;
Laryngoscopes;
Laryngoscopy;
Randomized controlled trial

## Abstract

*Background:* Compared with the GlideRite® rigid stylet used with the GlideScope® hyperangulated blade, the Flex-it® stylet permits dynamic orientation of the tracheal tube toward the glottis. This randomized trial compared GlideRite and Flex-it stylets regarding GlideScope intubation duration in patients wearing a semi-rigid cervical collar, simulating conditions of cervical instability.

*Methods:* Eighty adult patients undergoing elective surgery were randomized to tracheal intubation using the GlideScope with either the GlideRite or the Flex-it stylet. A semi-rigid cervical collar was applied following anesthesia induction. The primary endpoint was intubation duration in seconds. For our modified Intention-To-Treat (mITT) analysis, failed intubations were assigned the maximum duration of 180 seconds. Secondary outcomes included tracheal tube handling duration, intubation failure, Intubation Difficulty Scale, and ease of stylet removal.

*Results:* Seventy-five patients were included in the mITT analysis (n = 38 for Flex-it, 37 for GlideRite). Intubation duration was longer when using Flex-it compared with GlideRite (median [IQR]: 88 [65–180] vs. 49 [39–61] s) with a median difference of 37 s (95% CI: 22 to 64; $p < 0.0001$). Tube handling duration was longer with Flex-it (median [IQR]: 48 [33–66] vs. 34 [25–39] s; $p = 0.007$). Failure rate was significantly higher with Flex-it (15/38, 39.5%) than with GlideRite

* Corresponding author.
*E-mail:* eric.zhu@umontreal.ca (E. Zhu).

https://doi.org/10.1016/j.bjane.2026.844836
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

(2/37, 5.4%; $p < 0.001$). Intubation Difficulty Scale score was higher for Flex-it (median [IQR]: 6 [5–7] vs. 2 [2–5]; $p < 0.001$).
*Conclusion:* In elective patients with simulated cervical instability, videolaryngoscopy with the Flex-it stylet resulted in longer intubation, greater difficulty, and higher failure rate compared with the GlideRite stylet.
© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

The GlideScope® video laryngoscope (Verathon® Medical Inc., Bothell, WA, USA) has a hyperangulated blade that improves glottic view over direct laryngoscopy. However, this angulation can cause difficulties when positioning the Tracheal Tube (TT), requiring routine use of an intubation stylet. A stylet mimics the curvature of the GlideScope blade, directing the TT toward the glottic aperture.[1] The GlideRite® (Verathon® Medical Inc., Bothell, WA, USA) is a reusable rigid metal stylet with a 90° curvature and a radius origin of 6 cm, designed for GlideScope intubation. Nonetheless, GlideRite was not superior to a malleable stylet in elective patients without signs of difficult intubation, when used by either experienced[2] or novice[3] anesthesiologists. The Flex-it® (Parker Medical, Highlands Ranch, CO, USA) is a disposable plastic articulated intubation stylet, enabling operators to vary the angulation (Fig. 1) to orient the TT tip between the vocal cords, facilitating advancement down the trachea. It is also more pliable than GlideRite, potentially reducing tissue trauma during laryngoscopy. The Flex-it and GlideRite stylets have rarely been compared during videolaryngoscopy with a hyperangulated blade, especially in non-manikin-based difficult airway situations.

In this randomized trial, we hypothesized that intubation duration would differ between the GlideRite and Flex-it stylets when used with the GlideScope hyperangulated blade in patients wearing a semi-rigid cervical collar. Secondary outcomes included TT handling duration, intubation failure rate, ease of intubation, hemodynamic response, and side effects.

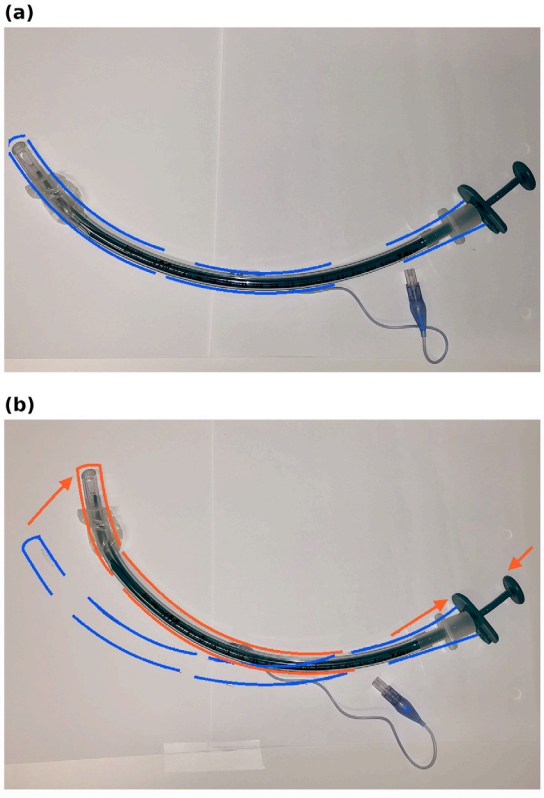



**Figure 1** Impact of the Flex-it stylet tensioning on the curve of the tracheal tube. Flex-it released (a); Flex-it with tension (b).

## Methods

This was a single-center, parallel-group, randomized controlled trial conducted at Maisonneuve-Rosemont Hospital from 14 June 2014 to 17 November 2016. The study is reported in accordance with the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) CONsolidated Standards Of Reporting Trials (CONSORT) guidelines.[4] Our institutional research ethics board approved the study protocol for this trial on 12 May 2014 (IRB approval number: 14013). The trial was registered before patient enrollment on 29 May 2014 (ClinicalTrials.gov registration: NCT02177240). Written informed consent was obtained from all participants.

### Study population

Inclusion criteria were American Society of Anesthesiologists (ASA) physical status I or II, age 18 to 65 years, body mass index of $\leq$ 35 kg.m$^{-2}$, and elective surgery under general anesthesia requiring a TT. Exclusion criteria were Mallampati score of $\geq$ 3, mouth opening $<$ 2.5 cm, thyromental distance $<$ 6 cm, limited cervical mobility, upper respiratory tract abnormalities (e.g., trauma, radiotherapy), difficult intubation history, rapid sequence induction, or contraindication to GlideScope or cervical collar.

### Procedure

Simple randomization with 1:1 allocation was performed using an electronic randomization list (randomizer.org) generated by an independent research assistant not involved in patient enrollment, with codes kept in opaque, sequentially numbered envelopes. Each envelope was opened only after the patient had undergone anesthesia induction, immediately prior to intubation. Given the nature of the intervention, stylet type was revealed once the TT was handed to the operator. Therefore, while participants and outcome assessors were blinded, operators were not blinded to the intervention.

For each patient, the respiratory therapist prepared two TTs with the studied stylets, concealing them under a towel. Upon patient arrival, the respiratory therapist established

intravenous access and started an infusion of lactated Ringer's solution. Standard monitoring included pulse oximetry, non-invasive blood pressure, 5-lead ECG, Train-Of-Four (TOF) monitoring of neuromuscular blockade, and Bispectral Index (BIS) (Medtronic, Kirkland, QC, Canada). The attending anesthesiologist decided on additional monitoring and administered premedication with 1 mg intravenous midazolam unless contraindicated. Patients were preoxygenated to an end-tidal oxygen concentration of 85% before induction. The anesthesiologist induced the patient using propofol 1 to 3 $mg.kg^{-1}$, sufentanil 0.1 to 0.3 $\mu g.kg^{-1}$, and rocuronium 0.4 to 0.6 $mg.kg^{-1}$. The respiratory therapist performed mask ventilation until the response to ulnar neurostimulation disappeared, determined by a TOF count of 0/4 (no muscle response to any of the four stimuli). The research assistant then applied a Philadelphia semi-rigid cervical collar (Össur Americas, Paulsboro, NJ, USA) and secured the patient's head to the operating table with tape.

One of three participating anesthesiologists inserted the GlideScope to achieve a satisfactory glottic view. All intubations were performed using the same blade model in both study arms. No standardized instructions were provided to operators regarding technique. The respiratory therapist then provided the TT (7 mm for female patients, 8 mm for male patients) with either GlideRite or Flex-it stylet according to allocation. Only then did the operator learn the allocated stylet. They confirmed TT positioning using a satisfactory End-Tidal $CO_2$ ($ETCO_2$) waveform, then ventilated the patient according to standard practice. No changes to the study protocol or outcomes occurred after trial initiation.

The anesthesiologists had eight to ten years of practice and were proficient in both direct laryngoscopy and videolaryngoscopy. They completed a minimum of 10 GlideScope intubations with each stylet before participating, either on patients or on manikins.[5]

Removing the TT or GlideScope blade from the patient's mouth before a successful intubation was considered a failed attempt. Between attempts, the patient was ventilated at the anesthesiologist's discretion. Intubation failure was defined as a procedure requiring more than three attempts or exceeding 180 seconds in duration. Failed intubations were assigned the maximum time (180 seconds) for our modified Intention-To-Treat (mITT) analysis. In such cases, the attending anesthesiologist managed the airway and restored standard care.

During the procedure, the principal investigator filmed the patient's monitors with a video camera (Handycam Camcorder, Sony Corporation, Tokyo, Japan) while the assistant verbally signaled events. The assistant said "blade" to signal the insertion of the blade into the patient's mouth and "tube" to signal the operator's handling of the tube. Capnograph waveforms were visible on the recordings. We did not capture the actual intubation procedure. This method allowed blinded assessment of timing outcomes.

## Primary outcome

The primary outcome was the intubation duration in seconds, defined as the interval from placing the GlideScope blade into the patient's mouth until detection of $ETCO_2 \geq 30$ mmHg.

## Secondary outcomes

Secondary outcomes included TT handling duration in seconds (interval from TT insertion to detection of $ETCO_2 \geq 30$ mmHg), intubation failure (more than three attempts or longer than 180 seconds), number of attempts (any removal of the TT or the GlideScope blade from the patient's mouth), Intubation Difficulty Scale (IDS) score, difficulty in removing the stylet (assessed on a 5-point Likert scale), postoperative hoarseness and sore throat (assessed immediately after the surgery and after 24 hours on a 5-point Likert scale), and glottic trauma incidence (blood stains on the TT after extubation). The IDS has seven parameters, providing a progressive, quantitative assessment of the complexity involved in both direct and indirect laryngoscopy.[6] Although the IDS was developed for direct laryngoscopy and does not include parameters unique to videolaryngoscopy, it remains a clinically useful scale for evaluating intubation difficulty.[7,8] In this study, the IDS was applied using the standard scoring method. An IDS score $\geq 4$ indicates a moderately difficult intubation. We also compared blood pressure, heart rate, and oxygen saturation measurements before and 1 minute after intubation to evaluate the response to the procedure.

## Statistical analysis

### Primary outcome

We performed statistical analysis using Prism 7 for macOS (GraphPad Software, La Jolla, CA). We analyzed time to successful intubation under an mITT framework. Failed intubations were assigned the maximum allowed duration of 180 seconds. Given the non-normal distribution of intubation duration (assessed using the D'Agostino-Pearson omnibus test), we summarized the data as medians with interquartile ranges and compared them between groups using the Mann-Whitney test. We used the Hodges-Lehmann estimator to quantify the median difference between groups, with corresponding 95% Confidence Intervals.

In addition, we performed time-to-event analysis using Kaplan-Meier methods to describe the temporal dynamics of successful intubation. Because failed intubations were assigned a fixed duration, this analysis is meant to be descriptive and hypothesis-generating. Survival curves were compared using the Gehan-Breslow-Wilcoxon test, which places greater weight on early events and was selected given the clinical relevance of early intubation success. Hazard ratios with 95% confidence intervals were estimated using the Mantel-Haenszel method and log-rank test.

### Secondary outcomes

Because the TT handling duration was non-normally distributed, we chose to use the D'Agostino-Pearson test. It was then compared using the Mann-Whitney test and presented as median [IQR]. We compared the incidence of intubation failure and glottic trauma using the Pearson Chi-square test and presented them as percentages. We reported the Relative Risk (RR) and Absolute Risk Difference (ARD) with 95% confidence intervals. We used a Mann-Whitney test to analyze data on the number of attempts, IDS scores, stylet removal difficulty, and hoarseness or sore throat intensity and presented this data as median [IQR]. The blood pressure,

heart rate and oxygen saturation measurements were presented as mean ± SD. As these parameters were compared at a single time point (1 minute after intubation), we used an unpaired *t*-test after confirming normality of the data using the Shapiro-Wilk test. The significance level was 0.05 for each hypothesis test.

### Sample size calculation

Based on a study comparing the Flex-it stylet with a malleable guide during videolaryngoscopy,[9] median [IQR] intubation duration was 15 s [8–28]. We estimated the corresponding mean (SD) value using the method described by Hozo et al.[10] at 16 seconds (5.7), which we used to calculate the sample size needed to demonstrate a 33% difference (5 s) between interventions in intubation duration, assuming the same SD for both. We estimated sample size using Bernard Rosner's method[11] and concluded that we needed 36 patients per intervention, with $\alpha = 0.05$ and $\beta = 0.20$. We aimed to recruit 80 patients to compensate for exclusions or losses.

## Results

Seventy-five participants were included in the mITT analysis, with 38 patients in the Flex-it group and 37 patients in the GlideRite group (Fig. 2). Two patients in the Flex-it group and three in the GlideRite group were excluded from analysis because monitoring data could not be extracted. Patient characteristics are presented in Table 1. Demographic parameters and categorical variables are compared using the Absolute Standardized Difference (ASD) (Table 1).

### Primary outcome

Time distributions were markedly non-normal in both groups (D'Agostino-Pearson $p < 0.0001$). Median intubation duration was significantly shorter in the GlideRite group compared with the Flex-it group (median [IQR]: 49 [39–61] vs. 88 [65–180] s), with a median difference of 37 s (Hodges-Lehmann estimate, 95% CI: 22 to 64; $p < 0.0001$).

We examined time to successful intubation using Kaplan-Meier methods as an additional analysis (Fig. 3). Survival curves differed significantly using the Gehan-Breslow-Wilcoxon test ($p = 0.0418$), suggesting an early advantage with GlideRite. The hazard ratio did not differ significantly between groups (log-rank HR = 1.03, 95% CI: 0.65 to 1.62). This analysis was exploratory and should be interpreted as supportive rather than confirmatory of the primary analysis.

### Secondary outcomes

TT handling duration was longer with Flex-it than with GlideRite (median [IQR]: 48 [33–66] vs. 34 [25–39] s) with median difference of 14 s (Hodges-Lehmann estimate, 95% CI: 3 to 25; $p = 0.007$). Failure rate was significantly higher with Flex-it (39.5%) than with GlideRite (5.4%) (RR = 7.30;

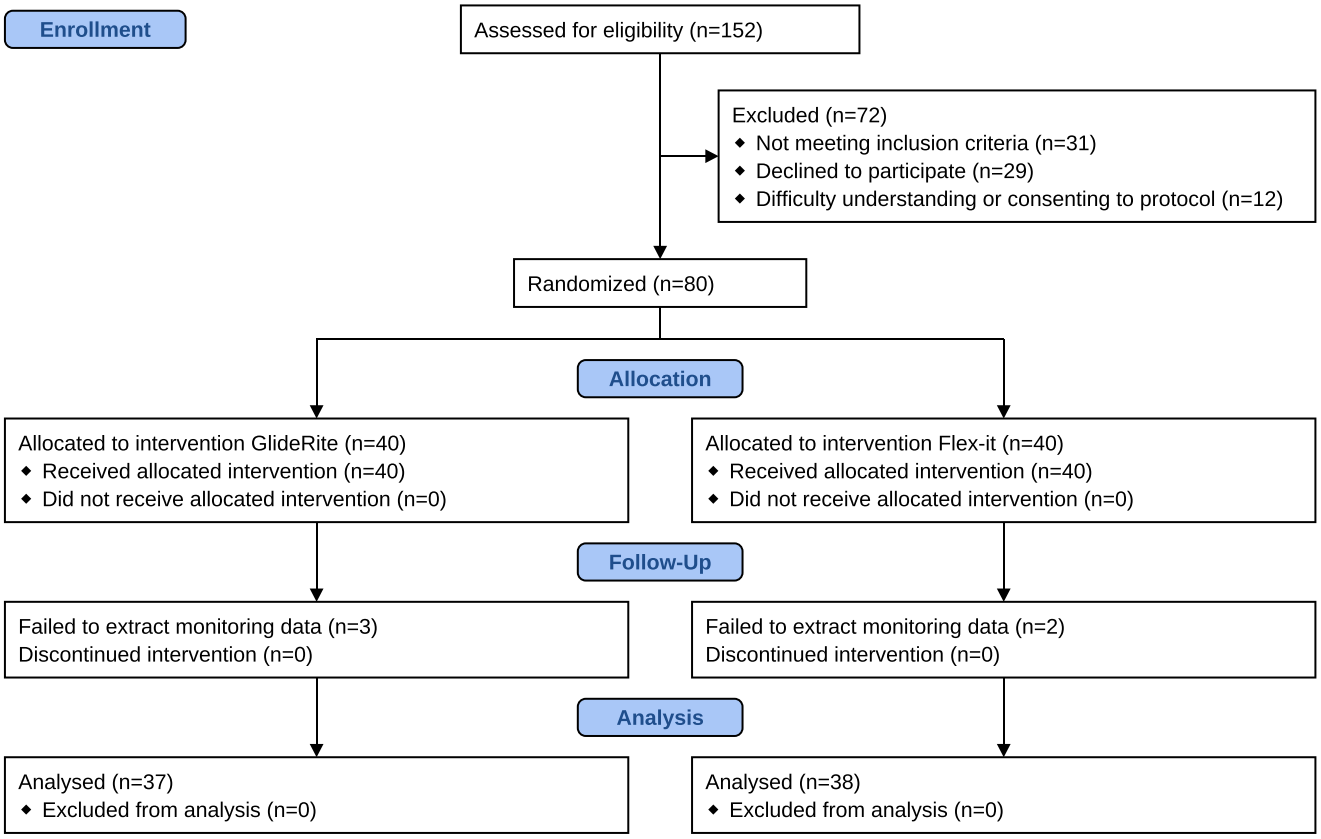



**Figure 2** CONSORT flow diagram.

**Table 1** Characteristics of patients intubated with GlideScope + GlideRite stylet (GlideRite) and with GlideScope + Flex-it stylet (Flex-it).

| | GlideRite | Flex-it | Absolute standardized difference (ASD) |
|---|---|---|---|
| **Patients, n** | 37 | 38 | |
| **Age (mean ± SD)** | 45 ± 12 | 50 ± 12 | 0.4 |
| **Males (%)** | 47 | 42 | 0.10 |
| **Body Mass Index (kg.m$^{-2}$, mean ± SD)** | 25 ± 4.2 | 27 ± 4.7 | 0.44 |
| **Edentulous patients (%)** | 23 | 18 | 0.12 |
| **American Society of Anesthesiologists (ASA) physical status (I/II)** | 11/26 | 10/28 | 0.08 |
| **BIS index before intubation (mean ± SD)** | 32 ± 10 | 36 ± 13 | 0.38 |

95% CI: 1.79 to 29.74, ARD = 34.07%; 95% CI: 16.90% to 51.23%, Number Needed to Harm [NNH] = 3; $p < 0.001$). Fewer attempts were required with GlideRite than with Flex-it (median [IQR]: 1 [1–2] vs. 3 [1–3]), Hodges-Lehmann median difference of 1, 95% CI: [0 to 2] ($p < 0.001$). IDS score was significantly lower with GlideRite (median [IQR]: 2 [2–5] vs. 6 [5–7]) with median difference of 3 (Hodges-Lehmann estimate, 95% CI: 1 to 4; $p < 0.001$). No serious adverse events occurred. Ease of stylet removal, hoarseness, sore throat, incidence of glottic trauma, blood pressure, and heart rate were not different between interventions. Oxygen saturation 1 minute after intubation was slightly higher with GlideRite (98.8 ± 0.9% vs. 98.3 ± 0.9%; $p = 0.027$) (Table 2).

## Discussion

In this randomized trial, tracheal intubation with a hyperangulated GlideScope blade was significantly inferior when the Flex-it stylet was used compared with the GlideRite rigid stylet across the primary and principal secondary outcomes. Intubation duration was significantly longer with the Flex-it stylet. However, the most clinically meaningful finding was intubation failure, with the Flex-it group experiencing a failure rate of 39.5% versus 5.4% in the GlideRite group. The NNH was 3, indicating one additional failure for approximately every three patients intubated with the Flex-it. However, the wide confidence interval (RR: 95% CI 1.79 to 29.74) reflects

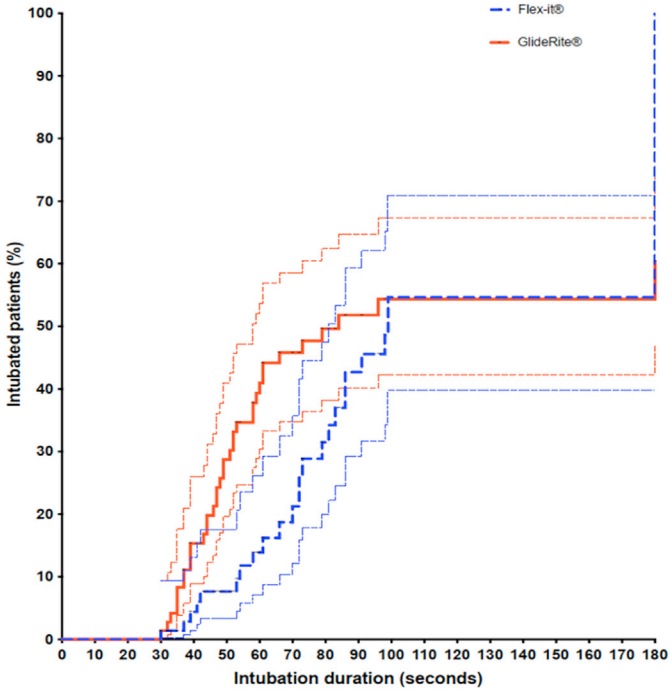



**Figure 3** Kaplan-Meier curves show time to successful intubation for Flex-it (blue) and GlideRite (red). Failed intubations were assigned a duration of 180 seconds. The GlideRite group demonstrated faster intubation, particularly at earlier time points (Gehan-Breslow-Wilcoxon p = 0.0418).

**Table 2** Primary and secondary outcomes for GlideRite and Flex-it interventions.

| | | GlideRite | Flex-it | p-value | Treatment effect: Difference between medians [95% CI] or difference between means ± SEM | Relative Risk (RR) [95% CI] | Absolute Risk Difference [95% CI] |
|---|---|---|---|---|---|---|---|
| **Primary** | Intubation duration (s, median [IQR]) | 49 [39–61] | 88 [65–180] | < 0.0001 | 37 [22 to 64] | | |
| **Secondary** | Tracheal tube handling duration (s, median [IQR]) | 34 [25–39] | 48 [33–66] | 0.007 | 14 [3 to 25] | | |
| | Failed intubation (n, incidence %) | 2, 5.4% | 15, 39.5% | < 0.001 | | 7.30 [1.79 to 29.74] | 34.07% [16.90 to 51.23] |
| | Number of attempts for intubated patients (median [IQR]) | 1 [1–2] | 3 [1–3] | < 0.001 | 1 [0 to 2] | | |
| | Intubation Difficulty Scale score (median [IQR]) | 2 [2–5] | 6 [5–7] | < 0.001 | 3 [1 to 4] | | |
| | Stylet removal difficulty (median [IQR]) | 2 [1–3] | 2 [1–3] | 0.527 | 0 [0 to 1] | | |
| | Hoarseness intensity (median [IQR]) | 1 [0–1.25] | 0 [0–1] | 0.465 | 0 [0 to 1] | | |
| | Sore throat intensity (median [IQR]) | 0 [0–1] | 0.5 [0–1] | 0.730 | 0 [0 to 0] | | |
| | Incidence of glottic trauma (n, incidence %) | 1, 2.7% | 3, 7.9% | 0.317 | | 2.92 [0.32 to 26.83] | 5.19% [-4.85 to 15.23] |
| | Systolic blood pressure 1 min after intubation (mmHg, mean ± SD) | 125 ± 30 | 141 ± 37 | 0.086 | 16.4 ± 9.4 | | |
| | Heart rate 1 min after intubation (bpm, mean ± SD) | 85 ± 10 | 89 ± 12 | 0.164 | 4.4 ± 3.1 | | |
| | Oxygen saturation 1 min after intubation (%, mean ± SD) | 98.8 ± 0.9 | 98.3 ± 0.9 | 0.027 | 0.59 ± 0.26 | | |

Primary and secondary outcomes for patients randomized to GlideScope intubation with the GlideRite rigid stylet or the Flex-it articulating stylet. Data are presented as median [interquartile range] or mean ± standard deviation unless otherwise indicated. Treatment effects for continuous outcomes are expressed as the Hodges-Lehmann median difference [95% Confidence Interval] or as the difference between means ± Standard Error of the Mean (SEM). Treatment effects are expressed as absolute differences between groups. Binary outcomes are reported as number of events and incidence (%), with Relative Risk (RR) and Absolute Risk Difference (ARD) with 95% confidence intervals. Intubation failure was defined as a procedure exceeding 180 seconds or requiring more than three attempts; a failed attempt was defined as any removal of the tracheal tube from the patient's mouth or repositioning of the GlideScope blade. The Intubation Difficulty Scale score ≥ 4 indicates at least moderate intubation difficulty. Stylet removal difficulty was rated by the performing anesthesiologist on a 5-point scale (1 = very easy, 5 = very difficult). Hoarseness and sore throat were assessed by a blinded recovery room nurse on a 5-point Likert scale.

the limited number of failures and warrants cautious interpretation of the precise magnitude of this effect.

Turkstra et al.[9] compared the Flex-it to a malleable stylet and found that intubation was longer and more difficult with the Flex-it, consistent with our results. However, they did not report a comparable failure rate disparity.

Our findings contrast with those of Bak et al.[12], who reported no difference in intubation duration between the GlideRite and Flex-it stylets in a simulated difficult airway, despite a subjective higher difficulty rating for the Flex-it stylet. Similarly, Sheta et al.[13] found comparable intubation durations and success rates for both stylets in patients with normal airways using a GlideScope hyperangulated blade. Methodological differences may explain the divergence between these results and our own. Neither study reported intubation failure as a primary or secondary outcome. Thus, the high failure rate we observed using the Flex-it stylet may have gone undetected in smaller samples or less challenging airways. Meanwhile, Reus et al.[14] reported a higher first-attempt success rate with the Flex-it when used with the McGrath® video laryngoscope. This was done across multiple simulated airway scenarios in manikins and may not be generalizable to our findings. Their findings may also reflect a device-specific interaction, as the McGrath employs a different blade geometry and may be more forgiving of variable stylet characteristics than the GlideScope blade.

Our results may be explained by mechanical differences between the stylets. The articulated design of the Flex-it was intended to provide dynamic modulation of the TT tip,[15] theoretically offering an advantage over rigid stylets for indirect visualization of the glottis.[16,17] Instead, we observed that the tension required to deflect the TT tip anteriorly caused irreversible deformation of the stylet. This caused it to effectively function as an unpredictable malleable stylet rather than a dynamic articulated device. Because the GlideScope hyperangulated blade requires reliable upward redirection of the TT tip to navigate the acute angle between the blade and the glottis, the excessive malleability of the Flex-it and its inability to maintain or restore its original curvature may explain both the prolonged intubation duration and high failure rate observed in our study. Consistent with this interpretation, studies performed with normal airways[13] or blades with reduced angulation[18] have not demonstrated the same inferiority, suggesting that the choice of stylet becomes clinically consequential under specific conditions that require reliable TT tip redirection. Our findings might vary if a less angulated blade was used, which requires fewer adjustments at the distal end.[18]

The combination of both prolonged intubation duration and higher failure rates carries meaningful clinical implications in situations where speed and reliability are paramount, such as in patients with suspected cervical instability. Based on our primary mITT analysis, the GlideRite stylet should be considered the preferred adjunct with the GlideScope hyperangulated blade in this patient population.

This study was conducted between 2014 and 2016. The protocol, endpoints, and analysis plan remain unchanged since trial registration. We acknowledge that videolaryngoscopy practice, technology, and familiarity have substantially evolved in the past decade. Hyperangulated videolaryngoscopy, including the GlideScope, has become widely adopted as a primary intubation technique.[19] Contemporary anesthesiologists may have greater baseline familiarity with hyperangulated videolaryngoscopy than the participants in our trial. This could attenuate the magnitude of the effect observed, though it is equally plausible that the intrinsic mechanical limitations of the Flex-it stylet would persist.

## Limitations

This study has several limitations. First, since rigid stylets are more commonly used than malleable ones like Flex-it, operator familiarity with the GlideRite may be greater. A study comparing intubation duration with two different video laryngoscopes suggests that time to intubation by an operator converges by the sixth procedure.[5] Although we standardized training to a minimum of 10 intubations with each stylet, performance bias favoring the GlideRite cannot be completely excluded. In particular, the markedly higher failure rate observed with the Flex-it stylet may partly reflect differential operator familiarity rather than an intrinsic limitation of the device itself. In addition, because no instructions were provided concerning technique, operator variability represents an additional consideration. The contribution of operator technique to the observed results cannot be determined.

Second, this trial excluded patients with predictors of anatomically difficult airway, including Mallampati score $\geq 3$, mouth opening $< 2.5$ cm, thyromental distance $< 6$ cm, prior history of difficult intubation, and limited cervical mobility. Difficulty was therefore imposed extrinsically via cervical collar and head fixation, which does not replicate actual cervical instability encountered in clinical practice, nor does it reflect intrinsic anatomical challenges. This may underestimate the difference between stylets in true difficult airway scenarios involving distorted anatomy, trauma, hemorrhage, or other emergency scenarios. Consequently, our study design may have limited the ability to detect a potential advantage of articulated stylets, which theoretically permit greater adaptability in patients with anatomically unusual airways.

Third, all intubations were performed by experienced anesthesiologists. While this supports internal consistency, it limits generalizability to less experienced operators, in whom performance differences may be even more pronounced.

Fourth, the initial sample size calculation was based on parametric assumptions, whereas the primary analysis was ultimately conducted using non-parametric methods. As such, the estimated sample size should be interpreted as approximate. Moreover, the trial was not explicitly designed to detect the observed difference in failure rate. Intubation duration was chosen as the primary outcome because it is a commonly used primary endpoint for stylet comparison studies. Additional studies with intubation failure as a primary endpoint may be warranted.

## Conclusion

In elective adult patients without predictors of anatomically difficult airway, using a cervical collar to simulate conditions of cervical instability, intubation with the GlideScope hyperangulated blade using the Flex-it stylet was significantly slower and more difficult, and had a markedly higher failure rate than with the GlideRite stylet.

## Data availability statement

The de-identified dataset supporting the findings of this study has been deposited in Zenodo and is publicly available at 10.5281/zenodo.22004806.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

No artificial intelligence tools were used during the preparation of this work. The authors take full responsibility for the content of the published article.

## Authors' contributions

John Viet Nguyen: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; validation; writing-original draft.

Eric Zhu: Formal analysis; validation; writing-review and editing.

Said Dababneh: Writing-review and editing.

Pierre Drolet: Writing-review and editing.

Marie-Ève Bélanger: Data curation; validation; visualization; writing-review and editing.

Rami Issa: Data curation; validation; visualization.

Valerie Zaphiratos: Validation; visualization; writing-review and editing.

Nadia Godin: Investigation; project administration; resources; supervision; validation.

Jean S. Bussières: Conceptualization; methodology; writing-original draft.

Mihai Georgescu: Writing-review and editing.

Issam Tanoubi: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; resources; supervision; validation; writing-original draft; writing-review and editing.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgements

The authors would like to thank Mr. Kyle Roerick for his professional, comprehensive, and insightful language editing of the manuscript.

## Associate Editor

Ana Maria Menezes Caetano

## References

1. van Zundert A, Maassen R, Lee R, et al. A Macintosh laryngoscope blade for videolaryngoscopy reduces stylet use in patients with normal airways. Anesth Analg. 2009;109:825—31.
2. Turkstra TP, Harle CC, Armstrong KP, et al. The GlideScope-specific rigid stylet and standard malleable stylet are equally effective for GlideScope use. Can J Anaesth. 2007;54:891—6.
3. Xue FS, Liao X, Liu JH, Yuan YJ, Wang Q. Performance of the GlideRite® Rigid Stylet and malleable stylet for tracheal intubation by novices using the GlideScope® videolaryngoscope. Can J Anaesth. 2011;58:660—1.
4. Schulz KF, Altman DG, Moher D. CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
5. Pius J, Noppens RR. Learning curve and performance in simulated difficult airway for the novel C-MAC video-stylet and C-MAC Macintosh video laryngoscope: a prospective randomized manikin trial. PLoS One. 2020;15:e0242154.
6. Adnet F, Borron SW, Racine SX, et al. The intubation difficulty scale (IDS): proposal and evaluation of a new score characterizing the complexity of tracheal intubation. Anesthesiology. 1997;87:1290—7.
7. Vijayakumar V. Intubation Difficulty Scale Score and Videolaryngoscopes. J Neurosurg Anesthesiol. 2017;29:70—1.
8. McElwain J, Simpkin A, Newell J, Laffey JG. Determination of the utility of the Intubation Difficulty Scale for use with indirect laryngoscopes. Anaesthesia. 2011;66:1127—33.
9. Turkstra TP, Jones PM, Ower KM, Gros ML. The Flex-It stylet is less effective than a malleable stylet for orotracheal intubation using the GlideScope. Anesth Analg. 2009;109:1856—9.
10. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol. 2005;5:13.
11. Rosner B. Fundamentals of biostatistics. 7th ed. Boston: Brooks/Cole; 2011.
12. Bak JW, Noh Y, Kim J, et al. Comparison between GlideRite rigid stylet and Parker Flex-It stylet to facilitate GlideScope intubation in simulated difficult intubation: a randomized controlled study. Anesth Pain Med (Seoul). 2022;17:104—11.
13. Sheta SA, Abdelhalim AA, ElZoughari IA, AlZahrani TA, Al-Saeed AH. Parker Flex-It stylet is as effective as GlideRite rigid stylet for orotracheal intubation by Glidescope. Saudi Med J. 2015;36:1446—52.
14. Reus E, Grundmann U, Liening K, Wrobel M. Parker Flex-It intubation stylet versus a 90-degree curved stylet during intubation with the McGrath videolaryngoscope performed by novices: a manikin study with 5 airway scenarios. J Clin Anesth. 2013;25:624—8.
15. Ueshima H, Terao K, Kitamura A. Successful intubation in a patient with limited oral opening by using a combination of the Parker Flex-IT stylet and McGRATH MAC laryngoscopy. Masui. 2015;64:1056—8. Japanese.
16. Arai T, Yamamoto Y, Kikuchi A, Okuda Y. Use of a new videolaryngoscope (McGRATH X-blade) in patients with difficult airways. Masui. 2015;64:1048—51. Japanese.

17. McElwain J, Malik MA, Harte BH, Flynn NH, Laffey JG. Determination of the optimal stylet strategy for the C-MAC videolaryngoscope. Anaesthesia. 2010;65:369—78.
18. Zhang J, Tan LZ, Toh H, et al. Comparing the first-attempt tracheal intubation success of the hyperangulated McGrath X-blade vs the Macintosh-type CMAC videolaryngoscope in patients with cervical immobilization: a two-centre randomized controlled trial. J Clin Monit Comput. 2022;36:1139—45.
19. Griesdale DE, Liu D, McKinney J, Choi PT. GlideScope video-laryngoscopy versus direct laryngoscopy for endotracheal intubation: a systematic review and meta-analysis. Can J Anaesth. 2012;59:41—52.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



ORIGINAL INVESTIGATION

# Efficacy and safety of ciprofol versus remimazolam combined with alfentanil for sedation during outpatient colonoscopy: a randomized double-blind noninferiority trial

Xuan Liu [a,1], Qianqian Tang [b,1], Ning Zou [b], Changming Wang [b,*], Xuejun Guo [c]

[a] *The Dalian Medical University, Department of Anesthesiology, Dalian, China*
[b] *People's Hospital of Liaoning Province (Peoples' Hospital of China Medical University), Department of Anesthesiology, Shenyang, China*
[c] *People's Hospital of Liaoning Province (Peoples' Hospital of China Medical University), Department of Neurosurgery, Shenyang, China*

Received 25 April 2026; accepted 5 September 2026
Available online 12 September 2026

**KEYWORDS**
Alfentanil;
Ciprofol;
Colonoscopy;
Conscious Sedation;
Gastrointestinal Endoscopy;
Remimazolam

**Abstract**
*Introduction:* Colonoscopy often causes significant discomfort to patients. To improve patient experience, sedation-assisted colonoscopy is being increasingly adopted. This study was designed to assess the efficacy and safety of ciprofol versus remimazolam combined with alfentanil during outpatient colonoscopy.
*Methods:* A randomized, double-blind controlled clinical trial was performed among patients scheduled for elective outpatient colonoscopy. Patients were randomly allocated at a 1:1 ratio to two groups: the ciprofol group (Group C, 0.4 $mg.kg^{-1}$ intravenously, n = 88) and the remimazolam group (Group R, 0.3 $mg.kg^{-1}$ intravenously, n = 88). Intravenous alfentanil (3 $\mu g.kg^{-1}$) was administered to all patients in conjunction with the study drug. The primary outcome was the success rate of colonoscopy. Secondary outcomes included induction time, discharge time, satisfaction levels (patients and endoscopists), and adverse events.
*Results:* In the ITT population, colonoscopy success rates were 100.0% (88/88) in Group C and 98.9% (87/88) in Group R (treatment difference [95% CI]: 0.011 [-0.024 to 0.048]). The lower bound of the two-sided 95% CI was above the prespecified noninferiority margin of -0.08, confirming the noninferiority of ciprofol relative to remimazolam. After Bonferroni correction (adjusted threshold $p < 0.0028$), discharge time was longer in Group C (8.5 [6.0–12.0] vs. 7.0 [5.0–9.0] minutes; p = 0.002). Group R showed numerically higher MAP and HR at several time points.

* Corresponding author. C. Wang.
*E-mail:* teamwangcm@163.com (C. Wang).
[1] Xuan Liu and Qianqian Tang share first authorship equally.

https://doi.org/10.1016/j.bjane.2026.844824
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

*Conclusions:* This study demonstrated that intravenous ciprofol (0.4 mg.kg$^{-1}$) or remimazolam (0.3 mg.kg$^{-1}$), combined with alfentanil (3 $\mu$g.kg$^{-1}$), provided effective sedation for outpatient colonoscopy.
© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

Colonoscopy is considered an effective primary screening tool for colorectal cancer, but the procedure often causes discomfort in patients.[1] Therefore, sedation and analgesia are typically required to alleviate these symptoms.[2,3] Propofol has been widely used for sedation during colonoscopy. However, its use in outpatient settings is limited by several drawbacks, including injection pain, hypoxemia, and potential cardiorespiratory depression.[4,5]

Ciprofol is a cyclopropyl-modified propofol analog with high GABA-A receptor affinity, providing rapid onset, reliable sedation, and a safety profile similar to propofol.[6] Remimazolam is an ultra-short-acting benzodiazepine characterized by fast onset, quick recovery, favorable hemodynamic stability, and reversibility with flumazenil.[7,8] Both agents are promising alternatives to propofol for outpatient endoscopic sedation.[9] However, their clinical advantages and disadvantages differ in daily practice. In recent years, a number of clinical studies have compared ciprofol and remimazolam for endoscopic sedation. These studies consistently demonstrate that both agents provide effective sedation with favorable safety profiles and low rates of adverse events.[10,11] Remimazolam is associated with more rapid recovery and better hemodynamic and respiratory stability,[12] while ciprofol exhibits reliable sedative effects and a lower incidence of injection pain.[10] Additionally, comparable overall sedation efficacy has been reported, although the two drugs differ in their effects on intraoperative dreaming.[13] Our previous study demonstrated that intravenous administration of propofol (1.5 mg.kg$^{-1}$) combined with alfentanil (3 $\mu$g.kg$^{-1}$) provides effective sedation for sedation-assisted colonoscopy.[14] Nevertheless, direct head-to-head comparisons of ciprofol versus remimazolam, both combined with alfentanil for outpatient colonoscopy, remain scarce.

To address this issue, we conducted a randomized, double-blind clinical trial to evaluate the efficacy and safety of ciprofol versus remimazolam combined with alfentanil for sedation during outpatient colonoscopy.

## Methods

This double-blind, randomized, controlled study enrolled 176 patients undergoing colonoscopy at our hospital between December 2023 and June 2024. The study protocol was approved by the hospital's Ethics Committee in September 2022 (approval number: 2022 H002). All participating patients provided written informed consent. The study was prospectively registered at the Chinese Clinical Trial Registry on April 28, 2023 (Registration number: ChiCTR2300070996; URL: https://www.chictr.org.cn). This manuscript was prepared in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines.[15]

Patients aged 18–65 years with American Society of Anesthesiologists (ASA) physical status I to stable III and a body mass index (BMI) of 18.0–30.0 kg.m$^{-2}$ scheduled for elective colonoscopy were included. Exclusion criteria were as follows: BMI > 30 kg.m$^{-2}$; expected colonoscopy duration > 30-min; uncontrolled or decompensated severe cardiac, cerebral, pulmonary, hepatic, or renal disease; severe uncontrolled hypertension; preoperative Heart Rate (HR) < 50 beats per minute (bpm) or > 100 bpm; anemia; history of narcotic or sedative abuse; known or anticipated difficult airway; allergy to study medications; pregnancy or lactation; inability to cooperate or communicate; and refusal to participate.

Eligible patients were randomly assigned in a 1:1 ratio to the ciprofol group (Group C, 0.4 mg.kg$^{-1}$ intravenously) and the remimazolam group (Group R, 0.3 mg.kg$^{-1}$ intravenously). A randomization sequence was generated using Statistical Package for Social Sciences (SPSS) software version 27.0 (Inc., Chicago, Illinois, USA) (fixed value: 2000000) by an independent investigator not involved in patient enrollment or clinical care. Patients were enrolled consecutively and allocated to groups in strict chronological order according to this pre-generated random list. The randomization numbers were encoded, printed on opaque paper, and sealed in sequentially numbered, identical opaque envelopes to ensure allocation concealment.

On the day of the procedure, the envelopes were opened in order by the attending anesthesiologist, who administered the study drugs according to the allocated group. Both study medications were prepared in identical syringes with standardized total injection volume, and syringes were covered with opaque paper during administration. Patients were positioned in the lateral decubitus position with venous access concealed, so they remained blinded to group assignment. The examination area was separated by an opaque curtain. A dedicated anesthesiology assistant, fully blinded to group allocation, monitored vital signs, assessed sedation depth, documented all study outcomes and adverse events, and completed postoperative follow-up assessments. Therefore, the subjects, outcome assessors, endoscopists, and data analysts were blinded to group assignment.

Following premedication with intravenous alfentanil (3 $\mu$g.kg$^{-1}$, Yichang Renfu Pharmaceutical Co., Ltd.) given over 30 seconds, sedation was induced after 1-minute with either ciprofol (0.4 mg.kg$^{-1}$, Liaoning Haisho Pharmaceutical Co., Ltd.) in Group C or remimazolam (0.3 mg.kg$^{-1}$, Yichang Renfu Pharmaceutical Co., Ltd.) in Group R.[14,16] Sedation depth was assessed with the Modified Observer's Assessment of Alertness/Sedation (MOAA/S) scale, a validated 6-point scale (0–5): score 5, responds readily to name spoken in a normal tone; score 4, lethargic response to name spoken in a normal tone; score 3, responds only after name is called loudly and/or repeatedly; score 2, responds only to mild prodding or shaking; score 1, responds only to painful

trapezius squeeze; score 0, no response to painful trapezius squeeze. Colonoscopy was initiated when the MOAA/S score was $\leq 2$.[17] All procedures were conducted by experienced endoscopists using an Olympus CV-290 system and CF-H290I colonoscopes. In the event that the MOAA/S score remained $> 3$ following the initial dose, a supplementary dose (0.1 $mg.kg^{-1}$ ciprofol for Group C or 0.05 $mg.kg^{-1}$ remimazolam for Group R) was permitted after a 1-minute interval. Should this prove insufficient, 0.5 $mg.kg^{-1}$ propofol was to be used as the rescue medication, and study participation was terminated thereafter. Sedation was maintained by assessing the MOAA/S score every 5 minutes and administering supplemental doses as required. Intraoperative adverse events were recorded. Intraoperative hypotension was defined as a MAP $< 60$ mmHg or an SBP reduction $> 30\%$ from baseline and was treated with an intravenous bolus of 20 $\mu$g phenylephrine per episode. Bradycardia was defined as a HR $< 40$ bpm during sedation, representing clinically significant severe bradycardia that requires prompt intervention in gastrointestinal endoscopy settings.[18] Our management approach is to immediately pause the procedure and administer 0.5 mg atropine intravenously. Furthermore, a sustained $SpO_2$ between 75% and 90% for over 60 seconds prompted interventions such as jaw thrust, airway repositioning, or increased oxygen supplementation. If $SpO_2$ fell below 75% or apnea persisted beyond 60 seconds, positive-pressure ventilation via face mask or endotracheal intubation was instituted. Upon procedure completion, patients were moved to the Post-Anesthesia Care Unit (PACU). The Post-Anesthetic Discharge Scoring System (PADSS) was used to assess discharge readiness after colonoscopy sedation. The scale includes five items: 1) Stable vital signs with systolic blood pressure and heart rate within 20% of baseline values; 2) Stable ambulatory ability with independent standing and walking; 3) Fully recovered mental status with clear consciousness and oriented response; 4) Well-controlled pain with tolerable discomfort; 5) No obvious gastrointestinal bleeding or procedural bleeding. Discharge required a PADSS score $\geq 9$ along with stable vital signs.[19] The quality of recovery within seven days of surgery was assessed by the Quality of Recovery-15 (QoR-15) questionnaire.[20] The QoR-15 is a validated 15-item questionnaire assessing postoperative recovery, with higher scores indicating better recovery.

The primary outcome was to evaluate the colonoscopy success rate, which must meet the following criteria: 1) The completion of colonoscopy procedure (colonoscopy can sequentially and clearly observe the status of anal canal, rectum, sigmoid colon, colon and ileocecal mucosa); 2) No need for any alternative/rescue sedative medications; 3) The maximum number of subsequent additions should not exceed 5 doses and should be administered within 15 minutes.[21] Secondary outcomes included: 1) Induction time (the time from administration to the moment of MOAA/S $\leq$ 2); 2) Operation time (the time from the insertion of colonoscopy to the withdrawal of colonoscopy); 3) Alert time (the time from the last administration of sedative drugs to the moment of MOAA/S $> 4$); 4) Discharge time (the time from the end of procedure to the moment of the PADSS score $\geq 9$); 5) Patient and endoscopist satisfaction: Patient satisfaction with sedation experience and procedural tolerance, and endoscopist satisfaction with overall sedation quality and procedural conditions, were evaluated using a structured questionnaire. Each assessment was scored on a 10-point scale, with higher scores indicating greater satisfaction; the total score was derived from the sum of individual items;[22] 6) Use of an improved Brice interview questionnaire to assess intraoperative awareness;[2] 7) Systolic Blood Pressure (SBP), Diastolic Blood Pressure (DBP), Mean Arterial Pressure (MAP), HR, and Pulse Oxygen Saturation ($SpO_2$) were monitored and recorded at the following time points: T1, upon entering the unit; T2, 1 minute after receiving an intravenous injection of ciprofol or remimazolam; T3, colonoscope reaching the ileocecal valve; T4, at the end of the procedure; T5, when leaving the room.

Two analytical populations were defined: the Intention-To-Treat (ITT) population included all randomized patients, regardless of protocol deviations; the Per-Protocol (PP) population was defined as patients who completed the study according to the study protocol, which required completion of the prespecified study procedures and follow-up assessments. Patients who were lost to follow-up, withdrew informed consent, or did not complete the prespecified study procedures and follow-up were excluded from the PP population. In the event of sedation failure, rescue propofol was administered according to the prespecified rescue strategy, and the study was terminated for that patient. Because the primary outcome (colonoscopy success or failure) was definitively determined at the time of rescue, this patient was included in the PP analysis of the primary outcome as a treatment failure, in accordance with the principle that excluding treatment failures from the PP population in noninferiority trials would artificially inflate the success rate in the comparator group and bias the analysis in favor of noninferiority. In contrast, this patient was excluded from the PP analysis of secondary outcomes because study termination precluded the collection of complete follow-up data. The ITT population was primary for noninferiority testing, while secondary outcomes were analyzed in the PP population comprising patients who completed all prespecified study procedures and follow-up assessments.

PASS 15.0.5 (NCSS, LLC, Kesville, USA) was used for sample size calculations. Based on previous studies, the success rate of ciprofol was 99%,[3] and the success rate of remimazolam was 97.5%,[23] type I error rate $\alpha = 0.05$ (one-sided), and the target power was 90%. The required sample size was 70 subjects per group, which was increased to 88 subjects per group for a 20% dropout rate for a total sample size of 176.

Noninferiority testing was adopted to compare the sedation efficacy of ciprofol and remimazolam for colonoscopy, with a prespecified absolute non-inferiority margin of 0.08.[24,25] Clinically, an absolute difference $\leq 8\%$ in colonoscopy completion success is considered negligible, without material influence on endoscopic workflow, patient safety, or routine clinical judgment. A 7%–10% absolute non-inferiority margin for procedural success is commonly accepted in analogous gastrointestinal endoscopy sedation trials. The treatment difference was defined as the success rate in Group C minus that in Group R. A two-sided 95% Confidence Interval (95% CI) for the treatment difference was calculated using the Miettinen-Nurminen score method. Noninferiority was concluded if the lower bound of the two-sided 95% CI was above the prespecified noninferiority margin of -0.08.

SPSS 27.0 (Inc., Chicago, Illinois, USA) was used for statistical analysis. Normality was assessed via Shapiro-Wilk test

and histograms; homogeneity of variance was tested by Levene's test. Continuous data were compared by Student's *t*-test or Mann-Whitney *U*-test and presented as mean ± Standard Deviation (SD) or median (Interquartile Range, IQR). Categorical data were analyzed using Pearson's $\chi^2$, corrected $\chi^2$ or Fisher's exact test. Bonferroni correction was applied to adjust for multiple comparisons among secondary outcomes, with a two-sided p-value $< 0.0028$ (0.05/18 secondary comparisons) considered statistically significant. Hemodynamic parameters (MAP, HR, and $SpO_2$) were analyzed separately using repeated-measures analysis of variance, with group, time, and group-by-time interaction effects evaluated. A two-sided p-value $< 0.05$ was considered statistically significant for repeated-measures analyses.

## Results

From December 2023 to June 2024, a total of 350 patients were screened for outpatient colonoscopy at the Digestive Endoscopy Center of our hospital. A total of 174 patients were excluded due to non-compliance with the inclusion criteria. Ultimately, a total of 176 patients were included for randomization. In the ITT population (n = 176), 171 patients were included in the PP analysis of the primary outcome, and 170 patients were included in the PP analysis of secondary outcomes. In Group C, 2 patients were excluded because of loss to follow-up. In Group R, 2 patients were excluded because of loss to follow-up, and one because of withdrawal of informed consent. One patient who experienced sedation failure and received rescue propofol was included in the PP analysis of the primary outcome as a treatment failure, because the colonoscopy success/failure status was unequivocally assessable at the time of rescue; however, this patient was excluded from the PP analysis of secondary outcomes because study termination after rescue precluded the collection of complete follow-up data (Fig. 1). There were no significant differences in the baseline and demographic characteristics of patients between the two groups (Table 1).

The primary noninferiority analysis was conducted in the ITT population (n = 176). The success rate was 100% (88/88) in Group C and 98.9% (87/88) in Group R, with a treatment difference (ciprofol minus remimazolam) of 0.011 (95% CI: -0.024 to 0.048) (Table 2). The lower bound of the two-sided 95% CI was -0.024, which was above the prespecified noninferiority margin of -0.08, confirming the noninferiority of ciprofol relative to remimazolam. In the PP analysis of the primary outcome (n = 171), the success rate was 100% (86/86) in Group C and 98.8% (84/85) in Group R. The treatment difference (ciprofol minus remimazolam) was 0.012 (95% CI:

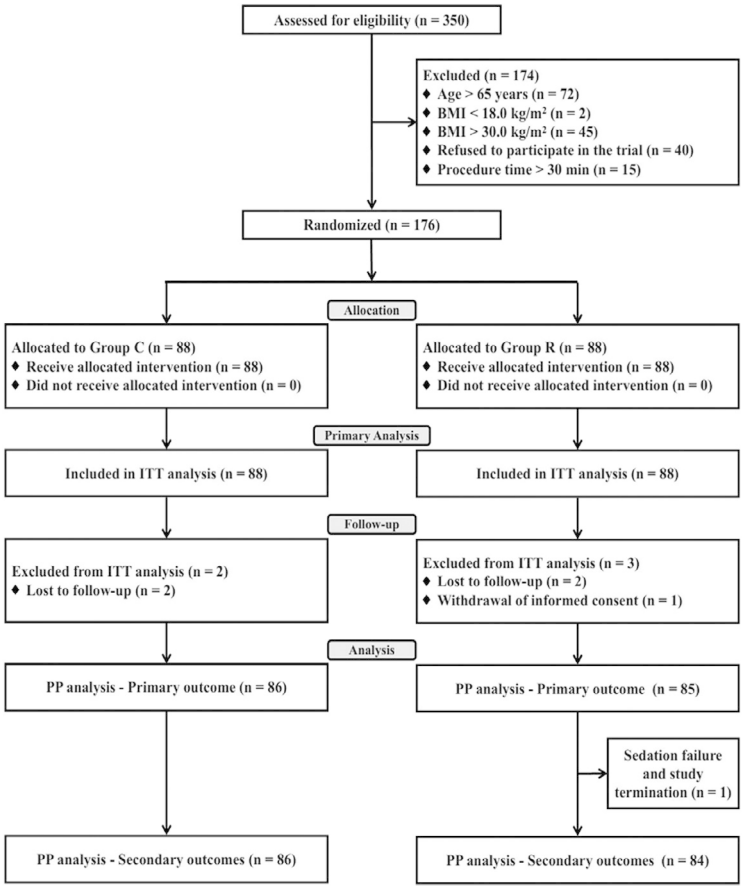



**Figure 1** CONSORT flow diagram. CONSORT indicates Consolidated Standards of Reporting Trials. Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group; BMI, Body Mass Index; ITT, Intention-To-Treat; PP, Per-Protocol. One patient who received rescue propofol was included in the PP analysis of the primary outcome as a treatment failure, but was excluded from the PP analysis of secondary outcomes because study termination after rescue precluded the collection of complete follow-up data.

**Table 1** Characteristics of the patients.

| | Group C (n = 88) | Group R (n = 88) | p-value |
|---|---|---|---|
| **Age[a], median (IQR); years** | 44.0 (34.0–54.0) | 48.0 (40.3–54.0) | 0.103 |
| **Sex[b], Male/Female, n (%)** | 47 (53.4)/41 (46.6) | 53 (60.2)/35 (39.8) | 0.361 |
| **Height[a], median (IQR); m** | 1.7 (1.6–1.8) | 1.7 (1.6–1.7) | 0.490 |
| **Weight[a], median (IQR); kg** | 70 (60–80) | 70 (60–75) | 0.435 |
| **BMI[a], median (IQR); kg.m$^{-2}$** | 24.6 (22.0–27.0) | 23.8 (22.0–26.0) | 0.277 |
| **ASA status[b], n (%)** | | | |
| I | 25/88 (28.4) | 15/88 (17.0) | 0.072 |
| II | 63/88 (71.6) | 73/88 (83.0) | |

A p-value < 0.05 was considered statistically significant. IQR, Interquartile Range (25$^{th}$–75$^{th}$ percentiles); BMI, Body Mass Index; ASA, American Society of Anaesthesiologists; Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group.
[a] Mann-Whitney *U*-test.
[b] Pearson Chi-Square.

**Table 2** Colonoscopy success rates based on Intention-To-Treat (ITT) analysis.

| | Group C (n = 88) | Group R (n = 88) | Estimated difference (95% CI) |
|---|---|---|---|
| **Success rate of colonoscopy, n (%)** | 88 (100.0) | 87 (98.9) | 0.011 (-0.024–0.048) |

Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group; ITT, Intention-To-Treat; CI, Confidence Interval. The treatment difference was calculated as the colonoscopy success rate in Group C minus that in Group R. The 95% CI was calculated using the Miettinen-Nurminen score method. Noninferiority was concluded because the lower bound of the two-sided 95% CI (-0.024) was above the prespecified noninferiority margin of -0.08. In the Per-Protocol (PP) population (n = 171), the colonoscopy success rate was 100% (86/86) in Group C and 98.8% (84/85) in Group R; the estimated treatment difference was 0.012 (95% CI: -0.026 to 0.054, Miettinen-Nurminen score method), with the lower bound also exceeding -0.08. The patient who received rescue propofol was included as a treatment failure in this PP analysis of the primary outcome.

-0.026 to 0.054). The lower bound of the two-sided 95% CI (-0.026) remained above the prespecified noninferiority margin of -0.08, further confirming the noninferiority of ciprofol relative to remimazolam.

For secondary outcomes, the PP population comprised 170 patients (86 in Group C and 84 in Group R), as the sedation failure patient lacked complete follow-up data. After Bonferroni correction for multiple comparisons (adjusted p = 0.05/18 = 0.0028), only two differences remained statistically significant. Discharge time was shorter in Group R (7.0 [5.0–9.0] vs. 8.5 [6.0–12.0] minutes, p = 0.002). The total sedation cost per patient was also significantly lower in Group R (Group C vs. Group R: 8.01 [6.97–9.12] vs. 5.10 [4.54–5.95] US Dollars, p < 0.001). While induction time was numerically shorter in Group R (140 [134–147] vs. 144 [136–153] seconds, p = 0.040), maintenance dose proportion was lower in Group R (Group C vs. Group R: 0.09/0.49 mg [18.4%] vs. 0.02/0.32 mg [6.2%], p = 0.009), and hiccups occurred more frequently in Group R (0/86 [0.0%] vs. 5/84 [6.0%], p = 0.025), these differences were no longer statistically significant after Bonferroni correction. The two groups showed no significant differences in other adverse events (Table 3), postoperative QoR-15 scores at days 1, 3, 5, or 7 (Table 4), alert time, or satisfaction levels reported by endoscopists and patients.

Comparisons of MAP, HR and $SpO_2$ at all predefined intraoperative time points between the two groups are shown in Figure 2. A significant main effect of time was observed for MAP ($F$ = 179.772, p < 0.001). MAP decreased markedly at T2 and gradually recovered from T2 to T5. The main group effect did not reach statistical significance ($F$ = 3.274, p = 0.072); the estimated marginal mean MAP was 86.59 mmHg in Group C and 88.98 mmHg in Group R (mean difference = -2.397 mmHg, p = 0.072). A significant group × time interaction was detected ($F$ = 3.038, p = 0.017), indicating different MAP trajectories between groups. Simple-effect comparisons revealed no significant intergroup differences at T1 (p = 0.931), T2 (p = 0.507) and T3 (p = 0.229). Nominally higher MAP values were observed in Group R at T4 (p = 0.013) and T5 (p = 0.004). A significant main effect of time was detected for HR ($F$ = 9.480, p < 0.001). HR slightly increased from T1 to T2, declined to the nadir at T4, and rebounded at T5, with significant within-time pairwise differences at multiple time points. A significant between-group effect was observed for HR ($F$ = 18.212, p < 0.001); the estimated marginal mean HR was 70.58 bpm in Group C and 76.93 bpm in Group R (mean difference = -6.347 bpm, p < 0.001). A significant group × time interaction was identified ($F$ = 11.187, p < 0.001), showing divergent HR dynamic trends over time. No significant intergroup difference was found at T1 (p = 0.061). Nominally higher HR values were observed in Group R at T2 (p < 0.001), T3 (p < 0.001), T4 (p = 0.003), and T5 (p = 0.002).

## Discussion

This study compared the sedative effects of ciprofol versus remimazolam for colonoscopy sedation, accounting for potential opioid-related confounding. The results showed that both drugs offer favorable sedation, with safety profiles that appeared comparable within the limits of this study. The noninferiority of ciprofol relative to remimazolam was confirmed in both ITT and PP analyses of the primary outcome,

**Table 3** Secondary outcomes during the study period.

| | Group C (n = 86) | Group R (n = 84) | p-value |
|---|---|---|---|
| **Induction time[b], median (IQR); seconds** | 144 (136-153) | 140 (134-147) | 0.040 |
| **Operation time[b], median (IQR); min** | 15.5 (12.0-19.0) | 16.0 (12.0-22.5) | 0.702 |
| **Alert time[b], median (IQR); min** | 10.0 (8.0-13.0) | 11.0 (8.0-14.8) | 0.358 |
| **Discharge time[b], median (IQR); min** | 8.5 (6.0-12.0) | 7.0 (5.0-9.0) | 0.002 |
| **Endoscopist satisfaction[a], mean ± SD; score** | 9.94 ± 0.28 | 9.87 ± 0.43 | 0.211 |
| **Subject satisfaction[a], mean ± SD; score** | 10.00 ± 0.00 | 9.99 ± 0.11 | 0.494 |
| **Induction consumption dose[c], n/total n (%); mg.kg$^{-1}$** | 0.40/0.49 (81.6) | 0.30/0.32 (93.8) | 0.009 |
| **Maintenance dose[c], n/total n (%); mg.kg$^{-1}$** | 0.09/0.49 (18.4) | 0.02/0.32 (6.2) | |
| **Average cost of sedative per case[b], median (IQR); US Dollars** | 8.01 (6.97-9.12) | 5.10 (4.54-5.95) | < 0.001 |
| **Intraoperative adverse reactions[c], n/total n (%)** | | | |
| Hypotension | 55/86 (64.0) | 44/84 (52.4) | 0.126 |
| Hypotension require intervention | 4/86 (4.7) | 1/84 (1.2) | 0.182 |
| Injection pain | 1/86 (1.2) | 7/84 (8.3) | 0.065 |
| Bradycardia | 1/86 (1.2) | 0/84 (0.0) | 1.000 |
| Tachycardia | 0/86 (0.0) | 2/84 (2.4) | 0.476 |
| Respiratory depression | 1/86 (1.2) | 6/84 (7.1) | 0.115 |
| Apnea | 0/86 (0.0) | 1/84 (1.2) | 0.991 |
| Hiccup | 0/86 (0.0) | 5/84 (6.0) | 0.025 |
| **Postoperative adverse reactions[c], n/total n (%)** | | | |
| PONV | 1/86 (1.2) | 1/84 (1.2) | 0.987 |
| Abdominal pain | 11/86 (12.8) | 5/84 (6.0) | 0.125 |
| Intraoperative awareness | 0/86 (0.0) | 0/84 (0.0) | – |

A two-sided p-value < 0.0028 was considered statistically significant after Bonferroni correction for multiple comparisons (0.05/18 secondary comparisons). Data are presented for the PP population for secondary outcomes (n = 170), which excluded the patient who received rescue propofol due to lack of complete follow-up data. IQR, Interquartile Range (25$^{th}$–75$^{th}$ percentiles); NRS, Numerical rating scale (range 0 to 10); PONV, Postoperative Nausea and Vomiting; SD, Standard Deviation; Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group.

[a] *t* test.

[b] Mann-Whitney *U* test.

[c] Pearson Chi-Square.

with consistent conclusions from the two analytical datasets as recommended by non-inferiority trial reporting guidelines.

As a novel intravenous anesthetic, ciprofol has been gradually applied in colonoscopy sedation. A previous study investigated the optimal dose of ciprofol for colonoscopy procedures, and the results demonstrated that ciprofol at doses of 0.4 mg.kg$^{-1}$ and 0.5 mg.kg$^{-1}$ exhibited the most favorable benefit-risk profile.[9] A study showed that the sedative effect of 0.4 mg.kg$^{-1}$ ciprofol and 1.5 mg.kg$^{-1}$ propofol in gastrointestinal endoscopy was equivalent.[3] Consequently, our study selected ciprofol at a dose of 0.4 mg.kg$^{-1}$ for inducing sedation.[9] Remimazolam is often combined with opioids to meet the specific sedation requirements of endoscopic procedures.[11,22,26] Previous studies have shown that a 0.3 mg.kg$^{-1}$ dose of remimazolam provides a safe and effective anesthetic method for patients undergoing upper gastrointestinal endoscopy.[15,27] Our study selected this dose for sedation induction. In our study, colonoscopy procedures using ciprofol or remimazolam, both combined with alfentanil, resulted in high rates of procedural success (100% and 98.9%, respectively). The impressive sedation success observed in both groups aligns with several earlier

**Table 4** Comparison of QoR-15 scores at different time points.

| Time Points | Group C (n = 86) | Group R (n = 84) | p-value |
|---|---|---|---|
| QoR-15 scores on POD1[a] | 147 (140–150) | 147 (138–150) | 0.485 |
| QoR-15 scores on POD3[a] | 150 (144–150) | 149 (140–150) | 0.332 |
| QoR-15 scores on POD5[a] | 150 (146–150) | 150 (143–150) | 0.297 |
| QoR-15 scores on POD7[a] | 150 (147–150) | 150 (146–150) | 0.994 |

Data presented as median (IQR).

A two-sided p-value < 0.0028 was considered statistically significant after Bonferroni correction for multiple comparisons (0.05/18 secondary comparisons). Data are presented for the PP population for secondary outcomes (n = 170). IQR, Interquartile Range (25$^{th}$–75$^{th}$ percentiles); QoR-15, the Quality of Recovery-15; POD1, Postoperative day-1; POD3, Postoperative day-3; POD5, Postoperative day-5; POD7, Postoperative day-7; SD, Standard deviation; Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group.

[a] Mann-Whitney *U* test.

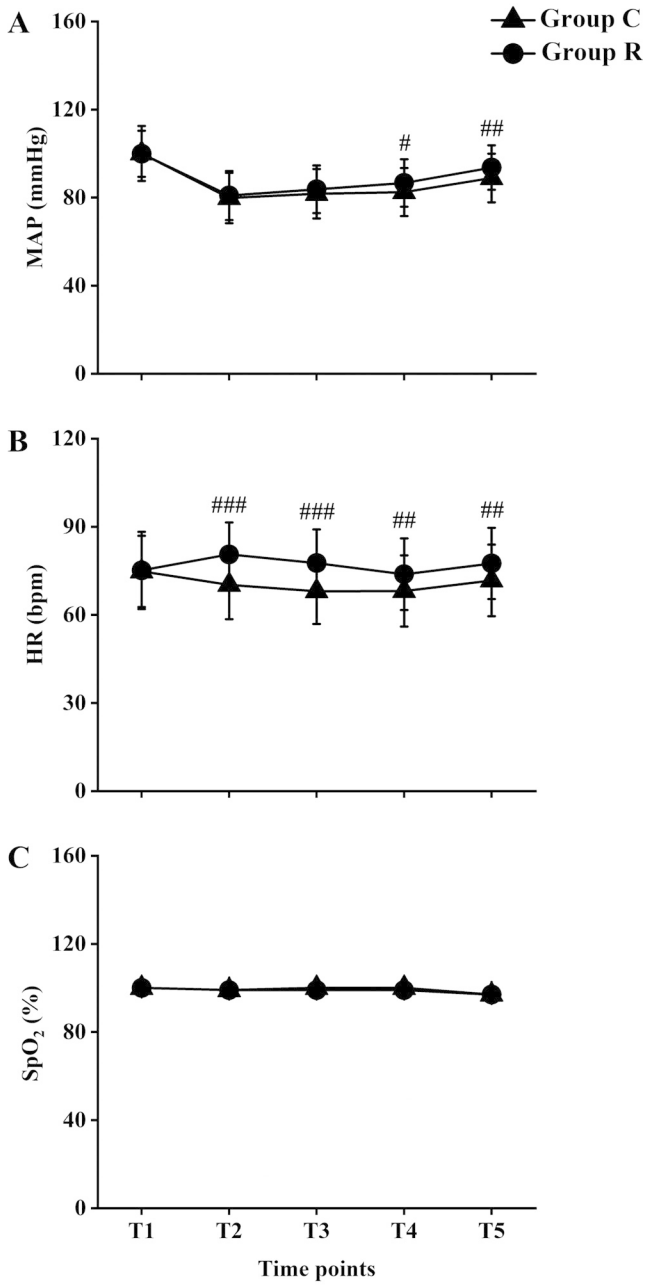



**Figure 2** Comparison of MAP (A), HR (B) and $SPO_2$ (C) between two groups. Group C, Ciprofol-Alfentanil group; Group R, Remimazolam-Alfentanil group; T1, upon entering the unit; T2, 1-minute after receiving an intravenous injection of ciprofol or remimazolam; T3, colonoscope reaching the ileocecal valve; T4, at the end of the procedure; T5, when leaving the room; MAP, Mean Arterial Pressure; HR, Heart Rate; $SPO_2$, Pulse Oxygen Saturation. $^{\#}$ $p < 0.05$, $^{\#\#}$ $p < 0.01$, $^{\#\#\#}$ $p < 0.001$.

reports.[3,12,13] It is noteworthy that the treatment difference in colonoscopy success rates between the ciprofol and remimazolam groups was 0.011 (1.1%) in the ITT population. The lower bound of the two-sided 95% CI for the treatment difference (ciprofol minus remimazolam) was -0.024, which was above the prespecified non-inferiority margin of -0.08. This statistical result definitively confirms that ciprofol met the predefined non-inferiority criterion compared to remimazolam for the colonoscopy success rate. Notably, one case of inadequate sedation occurred in the remimazolam group; this isolated finding is hypothesis-generating and does not allow general conclusions about remimazolam's overall reliability.

Additionally, both groups demonstrated high success rates, which may be associated with the enhanced analgesic effect achieved by the combined use of alfentanil. Compared to other opioids, alfentanil has a faster onset time and shorter duration of action, making it more suitable for anesthesia in outpatient surgeries.[22] Moreover, alfentanil effectively alleviated patient movements induced by intestinal traction during colonoscopy, thereby minimizing procedural interruptions caused by inadequate sedation.[28] Ultimately, this, in combination with sedative drugs of different mechanisms of action, achieved a high success rate.

In our study, although remimazolam had a statistically shorter induction time than ciprofol, the absolute difference of only 4 seconds is too small to exert any practical impact on clinical practice. Meanwhile, the 1.5-minute reduction in discharge time, while not prominent for individual patients, could contribute to higher patient turnover and operational efficiency in high-volume endoscopy departments with a large daily caseload. This not only maximizes resource efficiency but also increases the procedural turnover rate per unit time.[29] This study did not employ any antagonists, thus the recovery time of patients in both groups primarily depended on the metabolic process of the drugs themselves. To further shorten recovery time, consideration could be given to using the specific antagonist flumazenil in the remimazolam group.[30] Meanwhile, in our study, remimazolam showed a lower proportion of maintenance dose relative to total consumption, indicating fewer intraoperative supplemental boluses and more stable sedation. This favorable profile may be attributed to its high receptor affinity. According to the sedative pricing at our center (ciprofol: 11.5 US dollars; remimazolam: 5.6 US dollars), remimazolam was associated with significantly lower drug acquisition cost. This finding should be interpreted cautiously, as the study was not powered for formal pharmacoeconomic analysis.

The two sedation regimens showed different hemodynamic trends and adverse event profiles during the procedure. The incidence of intraoperative hypotension was 64.0% in the ciprofol group and 52.4% in the remimazolam group. Notably, most hypotensive episodes were transient and mild, with spontaneous recovery shortly after occurrence. Only a small proportion of patients required vasopressor intervention, and no repeated or continuous vasopressor administration was needed in any patient. No clinical symptoms associated with hypotension were observed. The relatively high incidence of hypotension may be attributed to the synergistic effect on hemodynamics between sedative drugs and alfentanil, which is a common pharmacological interaction in endoscopic sedation regimens.[7] Close intraoperative monitoring and timely symptomatic management ensured that these hemodynamic changes did not lead to adverse clinical outcomes, suggesting that both sedation regimens were generally well tolerated in this setting. Notably, five patients (5.95%) receiving remimazolam developed hiccups, consistent with prior reports.[10] These episodes were mild and self-limited, with no examination interruption or severe adverse events.

This study has several limitations. Firstly, this single-center trial excluded elderly, obese and ASA III—IV high-risk patients, which reduces the generalizability of our results. Secondly, the sample size was calculated only for the primary endpoint, lacking adequate power to detect rare adverse events. Additionally, safety outcomes were analyzed in the PP population rather than the full treated population, as complete safety data were unavailable for 6 excluded patients. These factors should be considered when interpreting the safety findings. Thirdly, objective sedation depth monitoring was not performed. Sedation depth was assessed merely by clinical observation, which may ignore subtle intergroup differences in sedation stability. Finally, no propofol control group was set in this trial, making direct comparison with traditional propofol sedation unavailable. Further multi-arm trials containing propofol are required for more complete clinical evidence.

## Conclusion

In conclusion, this study demonstrated that ciprofol (0.4 mg.kg$^{-1}$) and remimazolam (0.3 mg.kg$^{-1}$), each combined with alfentanil (3 $\mu$g.kg$^{-1}$), achieved adequate sedation for outpatient colonoscopy. Ciprofol was noninferior to remimazolam with respect to the colonoscopy success rate. Distinct hemodynamic profiles were seen across the two sedation regimens. Since this trial lacked power to capture rare adverse events, safety outcomes ought to be interpreted cautiously. The single sedation failure with remimazolam warrants further assessment within larger patient cohorts.

## Data availability statement

The fully de-identified individual participant raw data supporting the findings of this study are available in the Zenodo repository at https://doi.org/10.5281/zenodo.21625752. All personal identifiable information was completely removed to protect patient confidentiality.

## Ethics statement

The study was conducted following the guidelines of the Declaration of Helsinki. The research project for this study was submitted and approved by the Ethics Committee of the People's Hospital of Liaoning Province on September 28, 2022 (approval number: 2022 H002) and was registered in the clinical registration center under number ChiCTR2300070996 on April 28, 2023 (https://www.chictr.org.cn/showproj.html?proj=195336). Written informed consent was obtained from all subjects participating in the trial.

## Declaration of generative AI and AI-assisted technologies

This study did not utilize generative AI or AI-assisted technologies.

## Authorship contribution statement

Xuan Liu: Conceptualization; data curation; formal analysis; investigation; writing-original draft.

Qianqian Tang: Data curation; methodology; writing-original draft.

Ning Zou: Data curation; investigation.

Changming Wang: Conceptualization; methodology; project administration; resources, supervision, validation; visualization; writing-review & editing.

Xuejun Guo: Project administration; resources; supervision.

## Funding

This work was supported by the Xingliao Talent Program Project [grant number XLYC2412057] and the Research and Development Fund Project for Medical and Health Undertakings [grant number S231], and the China International Science Exchange Foundation [grant number SWYY2025001007]. The funders had no role in study design, data collection, and analysis, preparation of the manuscript, or decision to publish.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors thank all contributors of the study group and all patients participating in this study.

## Associate Editor

Edmundo Pereira Souza Neto

## References

1. Karsenti D, Tharsis G, Burtin P, et al. Adenoma and advanced neoplasia detection rates increase from 45-years of age. WJG. 2018;25:447—56.
2. Gao S-H, Tang Q-Q, Wang C-M, et al. The efficacy and safety of ciprofol and propofol in patients undergoing colonoscopy: A double-blind, randomized, controlled trial. J Clin Anesth. 2024;95:111474.
3. Li J, Wang X, Liu J, et al. Comparison of ciprofol (HSK3486) versus propofol for the induction of deep sedation during gastroscopy and colonoscopy procedures: A multi-centre, non-inferiority, randomized, controlled phase 3 clinical trial. Basic Clin Pharma Tox. 2022;131:138—48.
4. Tang L, Ye C, Wang N, et al. The median effective doses of propofol combined with two different doses of nalbuphine for adult patients during painless gastroscopy. Front Pharmacol. 2022;13:1014486.
5. Kumar S, Khuba S, Agarwal A, Gautam S, Yadav M, Dixit A. Evaluation of efficacy of Valsalva maneuver for attenuating propofol injection pain: a prospective, randomized, single blind, placebo controlled study. Korean J Anesthesiol. 2018;71(6):453—8.

6. Luo Z, Tu H, Zhang X, et al. Efficacy and Safety of HSK3486 for Anesthesia/Sedation in Patients Undergoing Fiberoptic Bronchoscopy: A Multicenter, Double-Blind, Propofol-Controlled, Randomized, Phase 3 Study. CNS Drugs. 2022;36(3):301—13.
7. Zhao N, Zeng J, Fan L, et al. Moderate sedation by total intravenous remimazolam-alfentanil vs. propofol-alfentanil for third molar extraction: A prospective randomized controlled trial. Front Med. 2022;9:950564.
8. Zhou Y-Y, Yang S-T, Duan K-M, et al. Efficacy and safety of remimazolam besylate in bronchoscopy for adults: A multicenter, randomized, double-blind, positive-controlled clinical study. Front Pharmacol. 2022;13:1005367.
9. Teng Y, Ou M, Wang X, et al. Efficacy and safety of ciprofol for the sedation/anesthesia in patients undergoing colonoscopy: Phase IIa and IIb multi-center clinical trials. Eur J Pharm Sci. 2021;164:105904.
10. Li D, Wang Y, Xing Y, Zhao Z, Chang L, Leng Y. Effectiveness and safety of remimazolam tosilate versus propofol for sedation in patients undergoing gastrointestinal endoscopy: a randomized controlled trial. Int J Clin Pharm. 2024;46:1371—80.
11. Wang Z, Zhan R, Mo L, et al. Comparison of effect of remimazolam and propofol on respiration of patients under deep sedation for colonoscopy: a prospective multicenter randomized controlled trial. Eur J Med Res. 2025;30:250.
12. Zhu H, Su Z, Zhou H, et al. Remimazolam Dosing for Gastroscopy: A Randomized Noninferiority Trial. Anesthesiology. 2024;140:409—16.
13. Zhou R, Fu L, Liu S, et al. Influences of Propofol, Ciprofol and Remimazolam on Dreaming During Anesthesia for Gastrointestinal Endoscopy: A Randomized Double-Blind Parallel-Design Trial. DDDT. 2024;18:1907—15.
14. Wang L-L, Guan Z-Y, Wang C-M, Zhang Y-W, Zhang J, Zhao P. A comparative study on the efficacy and safety of propofol combined with different doses of alfentanil in gastroscopy: a randomized controlled trial. J Anesth. 2023;37:201—9.
15. Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. Obstet Gynecol. 2010;115:1063—70.
16. Cui X, Cheng Z, Li H, et al. Efficacy and Safety of Different Doses of Remimazolam Tosilate Applied in Upper Gastrointestinal Endoscopy: A Prospective Randomized Controlled Double-Blind Trial. DDDT. 2023;17:2889—96.
17. Pastis NJ, Hill NT, Yarmus LB, et al. Correlation of Vital Signs and Depth of Sedation by Modified Observer's Assessment of Alertness and Sedation (MOAA/S) Scale in Bronchoscopy. J Bronchology Interv Pulmonol. 2022;29:54—61.
18. Riphaus A, Wehrmann T, Weber B, et al. S3 Guideline: Sedation for gastrointestinal endoscopy 2008. Endoscopy. 2009;41:787—815.
19. Trevisani L. Post-Anaesthetic Discharge Scoring System to assess patient recovery and discharge after colonoscopy. WJGE. 2013;5:502.
20. Stark PA, Myles PS, Burke JA. Development and Psychometric Evaluation of a Postoperative Quality of Recovery Score: The QoR-15. Anesthesiology. 2013;118:1332—40.
21. Zhong J, Zhang J, Fan Y, et al. Efficacy and safety of Ciprofol for procedural sedation and anesthesia in non-operating room settings. J Clin Anesth. 2023;85:111047.
22. Shi W, Cheng Y, He H, et al. Efficacy and Safety of the Remimazolam-Alfentanil Combination for Sedation During Gastroscopy: A Randomized, Double-blind, Single-center Controlled Trial. Clin Ther. 2022;44:1506—18.
23. Pambianco DJ, Borkett KM, Riff DS, et al. A phase IIb study comparing the safety and efficacy of remimazolam and midazolam in patients undergoing colonoscopy. Gastrointest Endosc. 2016;83:984—92.
24. Chen S, Wang J, Xu X, et al. The efficacy and safety of remimazolam tosylate versus propofol in patients undergoing colonoscopy: a multicentered, randomized, positive-controlled, phase III clinical trial. Am J Transl Res. 2020;12(8):4594—603.
25. Liang P, Dai M, Wang X, et al. Efficacy and safety of ciprofol vs. propofol for the induction and maintenance of general anaesthesia: A multicentre, single-blind, randomised, parallel-group, phase 3 clinical trial. Eur J Anaesthesiol. 2023;40 (6):399—406.
26. Wiltshire HR, Kilpatrick GJ, Tilbrook GS, Borkett KM. A Placebo- and Midazolam-Controlled Phase I Single Ascending-Dose Study Evaluating the Safety, Pharmacokinetics, and Pharmacodynamics of Remimazolam (CNS 7056): Part II. Population Pharmacokinetic and Pharmacodynamic Modeling and Simulation. Anesth Analg. 2012;115:284—96.
27. Sessler DI, Bloomstone JA, Aronson S, et al. Perioperative Quality Initiative consensus statement on intraoperative blood pressure, risk and outcomes for elective surgery. Br J Anaesth. 2019;122:563—74.
28. Doğanay G, Ekmekçi P, Kazbek BK, Yılmaz H, Erkan G, Tüzüner F. Effects of alfentanil or fentanyl added to propofol for sedation in colonoscopy on cognitive functions: Randomized controlled trial. Turk J Gastroenterol. 2017;28:453—9.
29. Yao Y, Guan J, Liu L, Fu B, Chen L, Zheng X. Discharge readiness after remimazolam versus propofol for colonoscopy: A randomised, double-blind trial. Eur J Anaesthesiol. 2022;39:911—7.
30. Doi M, Morita K, Takeda J, Sakamoto A, Yamakage M, Suzuki T. Efficacy and safety of remimazolam versus propofol for general anesthesia: a multicenter, single-blind, randomized, parallel-group, phase IIb/III trial. J Anesth. 2020;34:543—53.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



ORIGINAL INVESTIGATION

# Implementing patient blood management in a Brazilian public women's hospital: a pre—post implementation study

Lorena Costa Corrêa [a,b,*], Nataly Andrade Alves de Paiva [a,b], Ana Carolina Dias Almeida [a], Ulisses Montarroyos [c], Cesar de Almeida-Neto [d,e]

[a] *Universidade de Pernambuco, Centro Universitário Integrado Amaury de Medeiros, Recife, PE, Brazil*
[b] *Universidade Federal de Pernambuco, Hospital das Clínicas Romero Marques, Recife, PE, Brazil*
[c] *Universidade de Pernambuco, Programa de Pós-Graduação da Faculdade de Ciências Médicas, Recife, PE, Brazil*
[d] *Faculdade de Medicina da Universidade de São Paulo, Disciplina de Pós-Graduação em Ciências Médicas, São Paulo, SP, Brazil*
[e] *Fundação Pró-Sangue, Hemocentro de São Paulo, São Paulo, SP, Brazil*

Received 24 October 2025; accepted 23 September 2026
Available online 1 October 2026

**KEYWORDS**
Anemia, iron-deficiency;
Blood transfusion;
Bloodless medical and surgical procedures;
Hysterectomy;
Postpartum hemorrhage;
Preoperative care

## Abstract

*Background:* Patient Blood Management (PBM) is a personalized, evidence-based, and cost-effective approach. In surgery, the objective is to detect and treat anemia, minimal perioperative blood loss, and promote restrictive transfusion practices. It improves clinical outcomes, patient engagement, and reduces costs. We evaluated the effect of PBM on preoperative management of hysterectomies for uterine fibroids and Postpartum Hemorrhage (PPH) in a Brazilian public hospital.

*Methods:* In this pre-post intervention study, a PBM outpatient clinic was established, a Hemorrhage Management Protocol (HMP) was implemented, and low-cost educational initiatives were used. The study period was from 2016 to 2023.

*Results:* Overall, 36 patients were included; twenty-two returned to a 30-day appointment and had a mean preoperative hemoglobin increase of 2.48 $g.dL^{-1}$ (95% Confidence Interval: 1.8 −3.16; $p < 0.001$). Twenty-five underwent surgery. No preoperative prophylactic transfusions were administered. Transfusions of Packed Red Blood Cells (PRBC) had a mean reduction of -77.4 ($p = 0.253$) before January 2020 and maintained its behavior until 2023 (1.9; $p = 0.686$). The HMP was activated in five cases for PPH and 23 units of PRBCs and 11 units of Fresh Frozen Plasma (FFP) were transfused (mean of 4.6 PRBC units and 2.2 FFP units per patient). Additionally, 546 healthcare professionals accessed the PBM guidelines via instant messaging during the study period.

* Corresponding author.
*E-mails:* lorena.correa@upe.br (L.C. Corrêa), cesarnt.almeida@gmail.com (C. de Almeida-Neto).

https://doi.org/10.1016/j.bjane.2026.844839
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

*Conclusion:* The implementation of PBM by leveraging existing resources led to appropriate treatment of iron deficiency anemia before hysterectomies; systematic use of hemocomponents for PPH and sustained reduction on PRBC transfusions.
© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

Patient Blood Management (PBM) is a personalized, evidence-based, and cost-effective approach with three well-defined pillars: early detection and treatment of anemia and minimal blood loss in the preoperative period with restricted transfusion practices. These strategies have improved clinical outcomes, increased patient engagement, and reduced costs.[1,2] This care model emerged in 2005 and was globally recommended in 2010 by the World Health Organization (WHO), which called for the urgent global implementation of PBM by 2021.[1,3,4]

PBM may be applied for the management of preoperative anemia in surgeries for uterine myomatosis, Postpartum Hemorrhage (PPH), and postoperative or postpartum scenarios. Preoperative correction of anemia with oral or intravenous iron supplementation is recommended for patients undergoing hysterectomy or myomectomy. Several measures can be adopted to prevent and manage PPH. In the postoperative or postpartum period, for stable patients with anemia, tolerance to anemia is achieved by using antihemorrhagic drugs and micronutrients (e.g., iron) and restricting allogeneic blood transfusions to only required cases.[5-8]

Fibroids are benign myometrial tumors with an incidence of 5%—65%; surgical treatment is indicated in 80%—90% of cases. Women with fibroids develop Iron-Deficiency Anemia (IDA), which causes loss of quality of life, absence from work, and job losses.[5,8] Iron replacement is indicated to normalize Hemoglobin (Hb) levels, correcting anemia and stopping or reducing abnormal menstrual bleeding. There is no clinical indication for the prophylactic use of blood components in the preoperative setting. However, in our country, it is common for hospitalized patients with anemia who are scheduled to undergo myomectomy or hysterectomy to receive Packed Red Blood Cell (PRBC) transfusions with the aim of achieving preoperative hemoglobin levels of 10—12 g. $dL^{-1}$. Nevertheless, this approach is associated with poor clinical outcomes.[6,9-12]

PPH is the loss of $\geq$ 500 mL of blood in the third stage of labor for vaginal delivery or $\geq$ 1,000 mL for cesarean delivery, occurring up to 48h postpartum.[13] It is an important cause of maternal death worldwide and the second one in Brazil. In Pernambuco, PPH caused 12% of maternal deaths in 2018.[14] Therefore, it is essential to establish well-defined, evidence-based, and readily applicable Hemorrhage Management Protocols (HMPs).[13-15] Additionally, the administration of tranexamic acid within three hours of the diagnosis of Postpartum Hemorrhage (PPH), along with the implementation of an HMP, is recommended and may improve patient outcome.[16,17]

Until 2016, the preoperative routine at our institution included the transfusion of PRBCs before hysterectomy, with the aim of increasing hemoglobin levels prior to surgery. Although Brazilian Postpartum Hemorrhage (PPH) protocols were followed, no standardized Massive Transfusion Protocol (MTP) was in place, and Fresh Frozen Plasma (FFP) was administered routinely without an evidence-based indication. Therefore, to assess the impact of a Patient Blood Management (PBM) program on transfusion practices, we reviewed strategies for managing preoperative anemia in patients with uterine myomatosis, implemented an MTP, and promoted a more restrictive, tolerance-based approach to anemia. In addition, we encouraged restrictive transfusion practices and the use of specific therapeutic algorithms for anemia management. To support the management of preoperative anemia and optimize transfusion practices, we implemented low-cost, rapid, and widely accessible educational measures. These interventions addressed all three pillars of PBM using the resources available at our institution.

## Methods

### Type of study

This was an intervention study with pre- and post- evaluations, including the implementation of PBM. As part of this program, three procedures were introduced concurrently at Centro Universitário Integrado de Saúde Amaury de Medeiros (CISAM) based on the pillars of PBM (Fig. 1).

### Study setting

CISAM, linked to the University of Pernambuco (UPE), is a leading institution for women's health located in the northeastern region of Brazil. It has 108 beds and serves the Unified Health System. This unit is part of the UPE Hospital Complex (CH/UPE), which consists of 839 beds. CISAM provides outpatient care to approximately 80 women scheduled for fibroid correction. It conducts averagely 280 gynecological surgeries monthly, of which 25 are uterine myomatosis corrections. CISAM conducts 350 deliveries per month, with a PPH rate of 3%—5% and 1% requiring blood transfusions. As a teaching center, it receives students and residents from medical, nursing, and multidisciplinary areas.[18]

Our institution uses semi-electronic medical records and has no access to cell salvage systems or point-of-care testing.

### Study period

The study was conducted from December 2016 to December 2023. The PBM program was implemented on 31$^{st}$ December 2019, when an outpatient clinic, a hemorrhage management protocol, and educational measures were implemented concurrently.

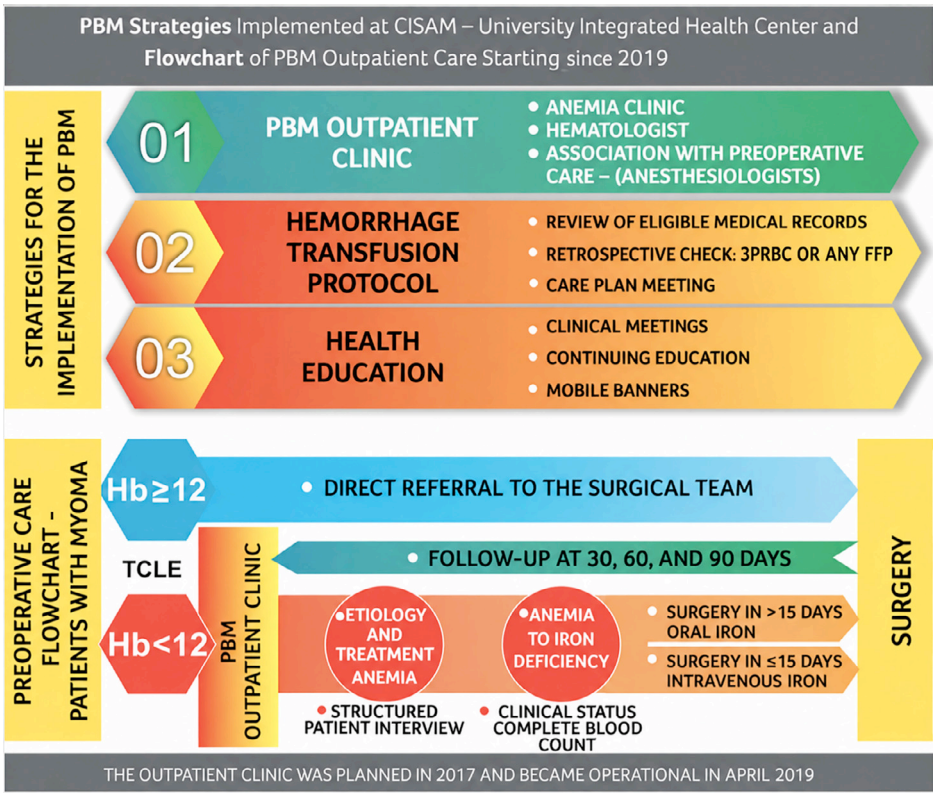



**Figure 1** PBM strategies implemented at the Integrated University Health Center (CISAM) and the flow chart of care at the CISAM PBM outpatient clinic as of January 2020.

### Procedure 1 – Patient blood management outpatient clinic

We identified and treated anemia for patients scheduled for hysterectomy or myomectomy. Patients with Hemoglobin (Hb) levels < 12 $g.dL^{-1}$ who were scheduled for elective hysterectomy or myomectomy were evaluated and followed at the PBM outpatient clinic of CISAM (Fig. 1). Before the initial consultation, all patients were invited to participate in the study and were given the opportunity to ask questions and withdraw from the study at any time. Written informed consent was obtained from all participants.

At the initial consultation, a structured medical history was obtained, including anemia-related clinical data, menstrual parameters, and information suggestive of other potential causes of anemia. Patients with Abnormal Uterine Bleeding (AUB) or ultrasound-confirmed uterine fibroids, Hb < 12 $g.dL^{-1}$, low Mean Corpuscular Volume (MCV), and/or elevated Red cell Distribution Width (RDW) were considered to have iron-deficiency anemia, even when a complete blood count was the only available laboratory test. This diagnostic approach was adopted because of limited resources and the lack of routine iron studies, in conjunction with a well-documented clinical history and the absence of other conditions associated with bleeding.

Patients diagnosed with anemia were treated according to the interval between the consultation and the scheduled surgery. When surgery was scheduled more than 15 days after the consultation, patients received oral iron therapy prescribed by their attending physicians, at a dose of 100 −200 mg of elemental iron per day. When surgery was scheduled within 15 days, intravenous iron therapy with iron sucrose was administered, with the dose calculated according to the Ganzoni formula (Fig. 1).[19] The number of infusions depended on the total iron requirement to correct iron deficiency and anemia. Patients were followed up on day 30, 60 and 90 from the first appointment. A hemoglobin increase of ≥ 1 $g.dL^{-1}$ within 30 days was considered a satisfactory response to iron replacement therapy. Patients who showed an inadequate response to oral iron or were intolerant to it were subsequently treated with intravenous iron. Patients presenting symptoms or signs suggestive of neoplasia were referred for specialist evaluation. No cases of malignancy or hereditary anemia were diagnosed. To assess perioperative PRBC utilization, we reviewed the medical records of patients undergoing surgery at CISAM from 48 hours prior to surgery until hospital discharge.

Transfusion was defined as each episode of blood component administration. PRBC transfusion was quantified in units, whereas FFP and platelet transfusions were recorded as transfusion episodes. The total number of blood components administered was recorded in units. For bleeding control, all patients received tranexamic acid and hormonal therapy when indicated by the attending gynecologists. No intraoperative interventions were performed during elective procedures.

### Procedure 2 – Implementation of the hemorrhage management protocol

The Hemorrhage Management Protocol used at hospital for PPH is based on Brazilian's Ministry of Health, which has recommendations on use of oxytocin for prevention and tranexamic acid for treatment after diagnosis of PPH. Despite

this guide, the massive transfusion protocol (MTP) was not in place. So, we designed it based on the Australian hemorrhage and transfusion protocols and included it in the clinical guidelines in 2019. At that time, Australian protocol for major hemorrhage control recommended use 1:1:1 ratio of PRBC:FFP: Platelet, or 2:1:1, considered the use of tranexamic acid for PPH and use of cryoprecipitate if bleeding continued, or if fibrinogen < 1 lt; 1 $g.L^{-1}$ (100 $mg.dL^{-1}$).[20] So, at CISAM, MTP was designed in 2019 to be activated in cases of administration of four or more PRBC transfusions in < 4h by the direct patient care team (Fig. 2). If the Transfusion Agency (TA) team identifies a situation that fits its parameters, the responsibility for conducting the procedure falls on the TA.[21,22] To analyze retrospectively MTP activation for PPH, we considered an initial searching of records from patients who have received three or more units of PRBC and at least one unit of FFP infusion (Supplementary Materials Fig. 1). After that, all patients who have received four or more PRBC transfusions (met MPT activation criteria) and had the diagnosis of PPH were included for this study (Supplementary Materials Figs. 1 and 2).[13]

### Procedure 3 – Educational measures

Educational measures were implemented through an instant messaging application in addition to presentations of the program in regular clinical meetings at the hospital. CISAM's hemotherapy routines were disseminated through written documentation made available and accessible on an institutional platform (CISAM Wiki) for the hospital's clinical staff. Additionally, information was shared via "WhatsApp" and through periodic meetings with medical staff and residents. The training for the medical team was conducted using digital banners prepared publicly and specifically produced by the researcher. Mastery of the educational material was measured using an icon corresponding to the reading. This project was presented in its entirety or in part at several meetings, including those with the TA team, clinical meetings, conferences with specialists, and senior management staff of the hospital and hemotherapy center. It was also presented at the Symposium of the CH/UPE, in the discipline of hematology and at the rectory.[23]

## Primary outcomes

We analyzed the increase in hemoglobin levels by at least 1 $g.dL^{-1}$ following treatment for iron-deficiency anemia, from the initial consultation to the 30-day follow-up and, when necessary, to the 90-day follow-up.

## Secondary and/or exploratory outcomes

The following outcomes were assessed: 1) The number of Packed Red Blood Cell (PRBC) transfusions received by patients evaluated at the PBM outpatient clinic who underwent hysterectomy or myomectomy during the study period; 2) The number of PRBC transfusions before and after implementation of the PBM program; 3) The total number of activations of the hemorrhage management protocol and the number of blood components administered; and 4) The icon indicating the participants reading of the educational materials distributed through electronic channels.

## Statistical analysis

PBM outpatient clinic: Considering that the service conducts approximately 250 gynecological surgeries monthly, and that 30% of these patients will present with anemia, according to the literature, we included all women that met criteria of AUB and anemia from myomas in the preoperative period of fibroid repair. A convenience sample of 36 patients was used for this study.[24] Data were collected during consultations at the PBM outpatient clinic using an Excel spreadsheet.

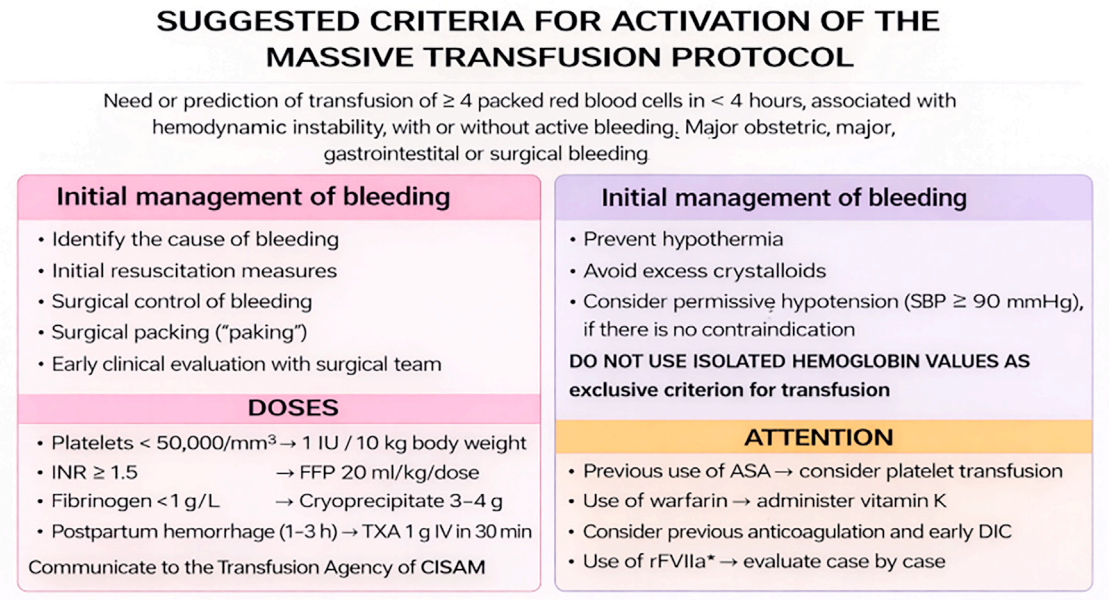



**Figure 2** CISAM's suggested criteria for activation of the massive transfusion protocol 2019.

The data were periodically inputed and later analyzed using STATA 14.[25] The evaluation conducted among the treated patients was unmatched, as a PBM outpatient clinic did not exist before 2019. We examined each patient included in the study on a case-by-case basis and described the results. Categorical variables are expressed as absolute and relative frequencies, whereas quantitative variables are expressed as means and standard deviations.[25]

Patients with gynecological conditions scheduled for elective surgery for the treatment of uterine fibroids were admitted to the gynecology ward, underwent surgery in the operating room, and were subsequently transferred to the post-anesthesia care unit and then to the gynecology ward until discharge. After mapping this patient pathway, we analyzed the total number of PRBC transfusions administered to adults throughout the hospital and stratified the data by an adult-care unit. A total of 4,688 PRBC transfusions were recorded during the study period, of which 1,989 were administered to adult patients between January 1, 2016, and December 31, 2023. We also compared the number of transfusions by the hospital unit before and after the implementation of the PBM program, using December 31, 2019, as the cutoff date.

To assess the temporal trend in the mean annual number of PRBC transfusions among adult patients at CISAM from 2016 to 2023, with December 31, 2019, as the intervention milestone, we performed a simple linear regression model using calendar year as the explanatory variable. Given the limited number of annual observations, we conducted an exploratory segmented trend analysis, dividing the time series into two periods (2016–2019 and 2020–2023) and estimating the regression coefficient for each period.[25]

### Implementation of the hemorrhage management protocol

We described numbers of MPT activations, PRBC, and FFP used and presented them as media and standard deviation.

Anonymized data are deposited in Mendeley Dataset with doi: 10.17632/4whnt9x9dh.2 and doi: 10.17632/hd6kg52v9t.2, which has laboratory data.

### Ethical considerations

This project was approved by the Research Ethics Committee of CISAM (CAAE number: 24402719.00000.5191) in 2019. Each patient's identity was preserved. All the patients signed an informed consent form.

## Results

### Procedure 1 — PBM outpatient clinic

Overall, 57 patients attended the outpatient clinic, after approval of this study by the ethical committee. Of these, 40 were eligible and provided written informed consent; however, four patients were excluded after consent was obtained – two because they were not anemic at presentation and two due to the absence of a diagnosis of uterine myomatosis. Consequently, 36 patients were included in the study (Fig. 3). From those 36 patients, 35 had clinical characteristics documented; 33 had previously been treated with iron or transfusions or both, as described in Table 1. Hemoglobin (Hb) values were available for 31 patients at

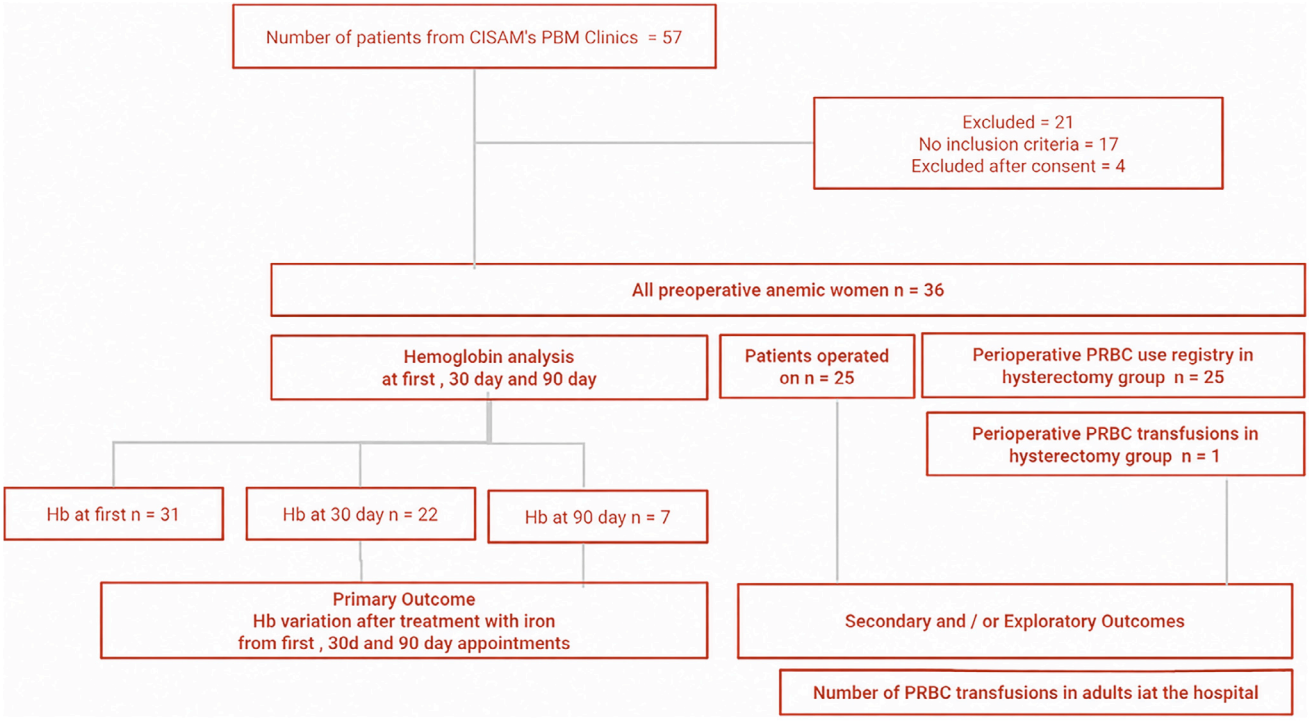



**Figure 3** Flowchart of patients from PBM Clinics of Centro Universitário Integrado de Saúde Amaury de Medeiros.

**Table 1** Clinical characteristics of patients included in the study at the patient blood management outpatient clinic of the Amaury de Medeiros Integrated University Health Center from January 2019 to 2022[a].

| Features | Number (%) |
|---|---|
| Number of respondents | 36 patients[c] |
| Age (years – Mean ± SD) | 42.9 ± 6.8 |
| Conducted surgery | 25 (69.4%) |
| Median time between 1st consultation and surgery (days)[b] | 90 |
| Postoperative bleeding | 1 (4.0%) |
| Clinically indicated PRBC transfusion within 48h before surgery[d] | 1 (4.0%) |
| Clinical characteristics (35 patients) | |
| Irregular cycles | 22 (62.9%) |
| Proiomenorreia | 8 (22.9%) |
| Days per cycle > 7 days | 22 (66.7%) |
| Use of night pad | 31 (91.2%) |
| Duration of anemia, months (3 patients)[b] | 24 (11—90) |
| Previous anemia | 22 (66.7%) |
| Family member with anemia member | 4 (11.8%) |
| Malaise | 21 (61.8%) |
| Daytime cold | 12 (36.4%) |
| Reduced labor | 18 (52.9%) |
| Pica | 11 (33.3%) |
| Does not climb stairs | 20 (60.6%) |
| Restless leg syndrome | 9 (27.3%) |
| Allergies | 2 (6.1%) |
| Weight loss | 7 (20.6%) |
| Shudder | 3 (8.8%) |
| Pyrosis | 6 (17,6%) |
| Bowel habit | 1 (2.9%) |
| Type of previous treatment (33 patients)[e] | |
| Transfusion | 11 (33.3%) |
| Oral iron | 19 (57.6%) |
| Intravenous iron | 12 (36.4%) |
| Referred for intravenous iron therapy[f] | 11 (30.6%) |

[# #]17 were referred by an anesthesiologist (48.6%) and 18 by a gynecologist (51.4%).

[# # #]None of the patients who underwent surgery had transfusions until discharge, and no patient reported fever or pruritus.

[a] Outpatient clinic was conducted by another hematologist during the year of 2023, who did not included patients in this study.

[b] Median (P25—P75).

[c] Clinical characteristics of 35 patients – One of the 36 patients had no clinical data but had treatment and surgery information.

[d] This transfusion was administered for symptomatic anemia and was not intended as prophylactic preoperative transfusion to facilitate surgery.

[e] Type (s) of treatment(s) before the first appointment in outpatients' anemia clinic was available for 33 patients.

[f] Among the patients who were referred for iron therapy, two underwent two sessions, four underwent six sessions, and three underwent 10 sessions; two patients had unavailable data.

the first appointment to PBM Outpatient Clinics. Twenty — two returned after 30 days to the second appointment and just 7 needed a third appointment, 90 days after.

**Table 2** Descriptive measures of hemoglobin levels available for the first, 30-day and 90-day appointments of patients seen at the PBM outpatient clinic of CISAM/UPE between 2019 and 2023.

| Hemoglobin levels | Mean ± SD |
|---|---|
| First consultation — 31 patients[a] | 8.4 ± 1.5 |
| Second visit (30 days) — 22 patients[b] | 10.7 ± 1.3 |
| Third visit (90 days) — 7 patients[c] | 11.3 ± 1.3 |
| Comparison of 1st and 2nd visit (22 patients) | |
| p-value | < 0.001 |
| Mean increase (95% CI) | 2.48 (1.8 to 3.16) |
| Patients with an increase of ≥ 1 g.dL$^{-1}$, n (%) | 19 (86.4%) |
| Comparison of 1st and 3rd visit (7 patients) | |
| p-value | 0.002 |
| Mean increase (95% CI) | 2.53 (1.39 to 3.67) |
| Patients with an increase of ≥ 1 g.dL$^{-1}$, n (%) | 7 (100%) |

CI, Confidence Interval; PBM, Patient Blood Management; CISAM, Centro Universitário Integrado de Saúde Amaury de Medeiros; UPE, Universidade de Pernambuco.

[a] Five patients had on hemoglobina value avaiable for ther first appointment.

[b] Twenty — two patients returned to 30-day appointment.

[c] Seven patients returned to the 90-day appointment.

Primary outcome — There was a mean increase in Hb levels of 2.48 (95% Confidence Interval: 1.8—3.16, p < 0.001) between the first (Hb1) and second (Hb2) visits to the PBM outpatient clinic for the 22 patients who returned to the 30-day appointment. And a mean increase in Hb levels of 2.53 (95% Confidence Interval:1.39—3.67, p = 0.002) between the first (Hb1) and third (Hb3) for 7 patients who needed an appointment on the 90th day (Table 2).

Secondary outcome: Overall, 69.4% of the treated patients underwent the surgical procedure at CISAM. From that 36% who weren't operated on, some stayed waiting for surgery and other lost follow up. Only one patient experienced postoperative bleeding (4%), and one patient required a PRBC transfusion (4%) due to symptomatic anemia. Therefore, the PRBC transfusion was indicated based on the patient's symptoms and was not intended to facilitate or enable surgical intervention. No other transfusions were required before, during, or after the surgery (Fig. 4).

## Analysis of the number of PRBC transfusions in adults

There was a numerical decrease in the number of PRBC transfusions in adults at CISAM. It was 257 (221—375) before January 2020 and 199 (193—204) after that. This trend had a mean reduction of -77.4 (p = 0.253) before January 2020 and maintained its behavior from 2020 to 2023 (1.9; p = 0.686). An expressive reduction occurred at rooming-in unit (-17.8; p = 0.005) and in delivery room (-46.8; p = 0.088). In the emergency room, where we have more severely bleeding

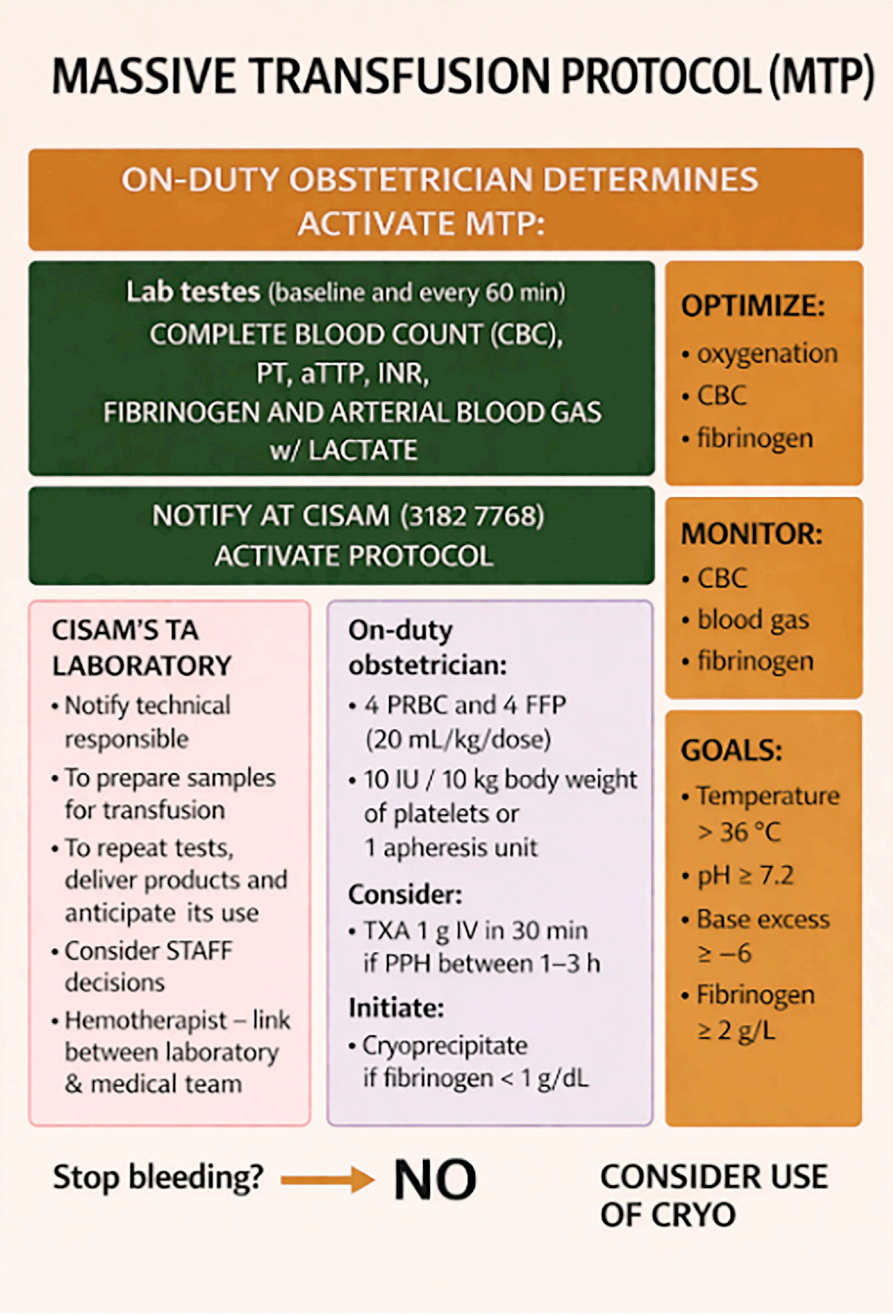



**Figure 4** CISAM's massive transfusion protocol flow chart 2019.

patients, we observed higher numbers (p = 0.094) (Table 3; Supplementary Fig. S1).

## Procedure 2 – Implementation of the hemorrhagic management protocol

Overall, 4688 transfusions were performed at CISAM (adults and pediatrics). Among these, 197 FFP units were transfused in adults, with the transfusion of 89 units occurring before the implementation of the hemorrhage care protocol and 65 occurring after this date. The mean number of FFP unit transfusions was 22.5 (20—29.5) before the protocol was implemented and 23 (17.5—31.5) after (p = 1.00) (Supplementary Fig. S2). To sensitize retrospective search of data for cases of Massive Transfusion Protocol (MTP) activation, we started searching data from all records of transfused patients with ≥ 3 PRBC units or any amount of FFP between 2019 and December 2023 (Supplementary Table S1). This initial search resulted in 69 patients who have received ≥ 3 PRBC units

**Table 3** Regress coefficient for trend of number of PRBC transfusions in adults during two periods at the Amaury de Medeiros Integrated University Health Center from 2016 to 2019 and 2020 to 2023.

| Setting (s) | From 2016 to 2019 $\beta$ (p-value) | From 2020 to 2023 $\beta$ (p-value) |
|---|---|---|
| CISAM | −77.4 (0.253) | 1.9 (0.686) |
| Gynecology clinic | −19.2 (0.391) | −0.7 (0.749) |
| Delivery Room | −46.8 (0.088) | 4.0 (0.404) |
| **Emergency Room**[a] | 21 (0.094) | 1.8 (0.782) |
| Surgery Room | −9.4 (0.446) | 1.8 (0.211) |
| **Post-anesthetic care unit** | −6.9 (0.512) | −2.5 (0.305) |
| **High risk patients' room**[b] | 2.0 (0.555) | −0.4 (0.742) |
| **Rooming-in**[c] | **−17.8 (0.005)** | −1.2 (0.106) |
| **Newborn skin-to-skin contact with mother**[c] | −0.3 (0.845) | −0.9 (0.712) |

CISAM, Centro Universitário Integrado Amaury de Medeiros.
**** Newborn skin-to-skin contact is a care practice in which the newborn is placed in direct skin contact with the mother.
[a] Emergency Room is called or Triage at CISAM.
[b] High risk patients' room is the yard for high-risk inpatient pregnant women.
[c] Rooming-in is the yard where mother and babies stay together.

or any amount of FFP. Six of them (6/69) received 4 or more PRBC units and, at least a FFP unit infusion, which met activation criteria according to MTP designed for CISAM at that time, and 5 (5/6) of MTP were to treat PPH (Supplementary Tables S2–S3). Activation of MTP for PPH (5 cases) used a total of 23 PRBC and 11 FFP units (mean of 4.6 PRBC units and 2.2 FFP units per patient). The ratio of PRBC to FFP units was 2.09:1. There was one transfusion of platelets concentrate, and no records of cryoprecipitate utilization in any of the situations. There was an increase in thawed and unused FFP units from 6 units by 2019 to 11 units by 2020. Tranexamic acid is used in all cases of PPH at CISAM. There was no registry of death among those five patients.

### Procedure 3 — Educational measures

A total of 546 healthcare professionals received and read messages sent through WhatsApp to service groups, including managers, medical managers, working groups of PBM at CISAM and CH/UPE, and resident doctors. PBM and CISAM protocols were presented in 24 meetings with around 1548 people in the audience.

## Discussion

We observed a significant change in care practices in this study. Reducing the number of PRBC transfusions improves clinical practice. The educational strategy using available and low-cost resources, such as "WhatsApp", was important for continuous education of health professionals and in the training of the team local multidisciplinary approach of the PBM. The HMP turned MPT policy into a systematic one and modified transfusional routines of staff during PPH. We used the available oral and intravenous iron compounds, resulting in an adequate increase in the Hb levels of patients treated at the PBM outpatient clinic. Preoperative anemia impairs the outcomes of women scheduled for gynecological surgeries, and transfusions do not correct this risk; however, they are associated with 61% and 16% increase in mortality and compound morbidity, respectively.[5] The impact of anemia on the management of patients hospitalized for leiomyoma was evaluated by Antunes et al., who concluded that the early identification and treatment of anemia improved care. Fajardo et al. (2023) evaluated 26,299 patients with fibroids who underwent myomectomy by laparotomy or video laparoscopy and concluded that an increased hematocrit was associated with a lower risk of blood transfusions. Therefore, optimization of Hb levels using iron replacement without transfusions in the pre-, intra-, and postoperative periods is consistent with the most recent literature and first PBM pillar, as recommended by the WHO.[1,9,26]

The implementation of written protocols, establishment of a clinic specialized in the treatment of anemia, and the promotion of frequent continuous education were fundamental for modifying local prescriptive habits, resulting in a reduction in PRBC transfusions which was sustained throughout the subsequent observation period. Education may help to systemically reduce the number of transfusions. These data reflect changes in the local culture linked to the pillars of PBM and its accelerators. These changes were made possible through accessible technologies in our service. Therefore, the implementation of PBM is feasible in scenarios with limited resources, such as ours.[2,23] Hemorrhage management protocols are recommended by current health regulations in Brazil. The implementation of an easy-to-understand protocol, disseminated through instant messaging, and widely known by the institution's professionals, is fundamental for improving the care of patients with PPH. We found that multiple transfusions at CISAM were unrelated to PPH and instead they were associated to other diagnosis. The mean transfusion ratio (PRBC:FFP) was 2.09:1 per patient with PPH with just one platelet transfusion, different from international protocols and possibly reflecting resource limitations or protocol maturity.[21,27]

This study has some limitations. The relatively short study period and limited sample size may have reduced the statistical power of the study, potentially limiting the ability to detect statistically significant differences between the study periods. However, the absence of transfusions in the elective preoperative setting, overall reduction in PRBC transfusions, and their maintenance for all the years after implementation may demonstrate program consistency. The activities of the PBM outpatient clinic were suspended in March 2020 and resumed in August 2021 because of the COVID-19 pandemic, which may be an important temporal confounder. There was no established practice of referring patients to a hematologist preoperatively. In fact, this interruption limited the growth in the number of outpatients evaluated during the study period. Concerning the number of PRBC transfusions,

no obvious sustained disruption was observed in the available descriptive data as we reviewed the annual number of surgeries performed at our hospital from 2016 to 2023. The mean number of surgeries was 251 per year from 2016 to 2019 and 266 per year from 2020 to 2023. After excluding 2020, the mean increased to 291 surgeries per year. Therefore, despite the disruption caused by the pandemic period, surgical activity and PRBC transfusion requirements remained stable throughout the study period. It is important to note that this study was not designed to isolate the effect of the pandemic period from the PBM implementation.[28] Initially, we did not have kits for ferritin measurement, which made it difficult to accurately diagnose iron-deficiency anemia. However, using data from the clinical history of chronic blood loss, we ruled out other diseases. In addition to the complete blood count, it was possible to make a presumptive diagnosis of IDA and treat patients who had increased Hb levels and did not need perioperative transfusions. Despite these drawbacks, we evaluated the ferrokinetics in all patients.[29] Finally, quantitative blood loss and other perioperative outcomes that could serve as proxy measures of bleeding severity (such as postoperative hemoglobin, hemoglobin decrease, reoperation for bleeding, ICU admission, use of hemostatic agents, or procedure-specific transfusion data) were not routinely collected in our database. Consequently, these outcomes could not be evaluated. Specifically, for the PPH massive transfusion protocol, we analyzed the number of activations and use of blood components.

## Conclusion

Improvements in health care have provided patients with adequate treatment for IDA, which improved results in hemoglobin levels among patients with paired measures. It allowed patients with myomas and with appointments in PBM clinics to undergo surgical procedures without receiving inappropriate prophylactic packed red blood cells PRBC transfusions before hysterectomy, which reduces immediate and late risks associated with the use of blood components. For the institution, PBM outpatient clinic was an innovation because it was the implementation of a new service line, which modified the workflow and the current medical assistance. Additionally, the sustained reduction in the number of PRBC transfusions may support the feasibility of a PBM program in a setting of constraints and may reflect a consistent implementation of the PBM outpatient clinic. The adequate treatment of IDA was possible using resources available in our center. The design of a hemorrhage management protocol was also an advancement. The implementation of these changes, integrating education and clinical practice in a resource-limited setting and leveraging free, widely accessible digital tools, underscores the role of educational activities and it is aligned with the pillars of PBM, which improves care through ethical and evidence-based practices.[2,30]

In summary, our findings support the feasibility of implementing this model, demonstrate increased hemoglobin levels among patients with paired measurements, show the absence of prophylactic preoperative transfusions in the outpatient PBM cohort undergoing surgery, and describe institutional transfusion trends. Further research is needed to evaluate the cost-effectiveness of such models, particularly given their potential to extend high-impact care to remote regions facing financial and technological constraints.

## Data availability statement

The data that support the findings of this study are available through Correa, Lorena (2026), "Implementing Patient Blood Management in a Brazilian Public Women's Hospital: a Pre—Post Implementation Study", Mendeley Data, V2, doi: 10.17632/4whnt9x9dh.2 and through Correa, Lorena (2026), "Implementing Patient Blood Management in a Brazilian Public Women's Hospital: a Pre—Post Implementation Study - with lab", Mendeley Data, V2, doi: 10.17632/hd6kg52v9t.2.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT for translation and for layout of some figures. The text has authors full review, edition and responsibility about it. No AI tool was used for study design, data collection, data analysis, interpretation, or scientific conclusions.

## Institutional research board approval

This project was approved by the Research Ethics Committee of CISAM (CAAE number: 24402719.00000.5191) in 2019 and remains in progress. Each patient's identity was preserved. All the patients signed an informed consent form.

## Authors' contributions

Conceptualization, A.N.C, C.C.L; Acquisition and execution, C.C.L, P.A.A.N, A.D.C.A; Analysis, M.U; Interpretation, A.N.C, C.C.L, M.U; Drafting, A.N.C, C.C.L; Revising, all authors.

## Funding

No financial support was received for the conduct of this study.

## Conflicts of interest

The authors disclose that CSL provided support for the translation and submission of this manuscript.

## Acknowledgements

We would like to thank nurse Ana Cláudia Ramos Claudino, biomedicals Raphael Ferreira Pimentel, Giselly Juliane Barbosa de Albertim and all nursery and laboratory teams from Transfusional Agency of Centro Universitário Integrado de Saúde Amaury de Medeiros, University of Pernambuco, for all partnership and for training our staff during this implementation program. The authors also acknowledge CSL for providing assistance with the translation and submission of the manuscript.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjane.2026.844839.

## Associate Editor

Luís Vicente Garcia

## References

1. World Health Organization. The urgent need to implement patient blood management: policy brief [internet]. Geneva: World Health Organization; 2021. [cited 2025 Oct 21] Available from https://apps.who.int/iris/handle/10665/346655.
2. Shander A, Hardy JF, Ozawa S, et al. A Global Definition of Patient Blood Management. Anesth Analg. 2022;135:476—88.
3. Spence RK, Erhard J. History of patient blood management. Best Pract Res Clin Anaesthesiol. 2013;27:11—5.
4. World Health Organization. World Health Assembly and Executive Board Resolutions on Blood Safety and Availability. [Internet]. Geneva: World Health Organization; 2011. [cited 2025 Oct 21]. Available from http://www.who.int/medicines/publications/essentialmedicines/en/index.html.
5. Richards T, Musallam KM, Nassif J, et al. Impact of Preoperative Anaemia and Blood Transfusion on Postoperative Outcomes in Gynaecological Surgery. PLoS One. 2015;10:e0130861.
6. Sabry M, Al-Hendy A. Medical treatment of uterine leiomyoma. Reprod Sci. 2012;19:339—53.
7. Beris P, Muñoz M, García-Erce JA, Thomas D, Maniatis A, Van der Linden P. Perioperative anaemia management: consensus statement on the role of intravenous iron. Br J Anaesth. 2008;100:599—604.
8. Muñoz M, Stensballe J, Ducloy-Bouthors AS, et al. Patient blood management in obstetrics: prevention and treatment of postpartum haemorrhage. A NATA consensus statement. Blood Transfus. 2019;17:112—36.
9. Fajardo OM, Chaves KF, Zhao Z, Curlin HL, Harvey LFB, Anderson TL. Preoperative Hematocrit Level and Associated Risk of Transfusion for Myomectomy Based on Myoma Burden and Surgical Route. J Minim Invasive Gynecol. 2023;30:115—21.
10. Choi UE, Nicholson RC, Thomas AJ, et al. A Propensity-Matched Cohort Study of Intravenous Iron versus Red Cell Transfusions for Preoperative Iron-Deficiency Anemia. Anesth Analg. 2024;139:969—77.
11. Murji A, Lam M, Allen B, et al. Risks of preoperative anemia in women undergoing elective hysterectomy and myomectomy. Am J Obstet Gynecol. 2019;221:629.e1—629.e18.
12. Breymann C, Römer T, Dudenhausen JW. Treatment of Iron Deficiency in Women. Geburtshilfe Frauenheilkd. 2013;73:256—61.
13. Baird EJ. Identification and Management of Obstetric Hemorrhage. Anesthesiol Clin. 2017;35:15—34.
14. Postpartum hemorrhage [Internet]. [Rio de Janeiro]: Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira; [2019?] [cited 2024 Apr 19]. Available from: https://www.arca.fiocruz.br/bitstream/handle/icict/29895/HEMORRAGIA%20P%C3%93S-PARTO.pdf?sequence=2&isAllowed=y.
15. Hu J, Yu ZP, Wang P, Shi CY, Yang HX. Clinical Analysis of Postpartum Hemorrhage Requiring Massive Transfusions at a Tertiary Center. Chin Med J (Engl). 2017;130:581—5.
16. Zdanowicz JA, Surbek D. Patient blood management in obstetrics - Review. Transfus Apher Sci. 2019;58:412—5.
17. WOMAN Trial Collaborators. Effect of early tranexamic acid administration on mortality, hysterectomy, and other morbidities in women with post-partum haemorrhage (WOMAN): an international, randomised, double-blind, placebo-controlled trial. Lancet. 2017;389:2105—16.
18. Universidade de Pernambuco. Specialties of the CISAM HUOC PROCAPE total beds [Internet]. Recife: Universidade de Pernambuco; 2020. [cited 2024 Apr 19]. Available from https://upe.br/anexos/complexo_hospitalar/ch_em_numeros/CH_EM_NUMEROS-2020.pdf.
19. German Society for Hematology and Medical Oncology. Iron deficit according to Ganzoni's equation [Internet]. Berlin: German Society for Hematology and Medical Oncology; 2025. [cited 2025 Oct 21]. Available from https://www.onkopedia.com/de/onkopedia/guidelines/eisenmangel-und-eisenma.
20. Fundação Oswaldo Cruz. Instituto Nacional de Saúde da Mulher, da Criança e do Adolescente Fernandes Figueira. Portal de Boas Práticas em Saúde da Mulher, da Criança e do Adolescente. Diretrizes Consolidadas da OMS para Hemorragia Pós-parto. 2025. [Internet]. Rio de Janeiro: Fiocruz; 2025 [cited 2 Jan 2026]. Available from https://portaldeboaspraticas.iff.fiocruz.br/atencao-mulher/novas-diretrizes-consolidadas-da-oms-para-hpp/.
21. National Blood Authority (AU). Patient Blood Management Guidelines: Module 5 — Obstetrics and Maternity. Canberra: NBA; 2015.
22. Bormanis J. Development of a massive transfusion protocol. Transfus Apher Sci. 2008;38:57—63.
23. Coleman E, O'Connor E. The role of WhatsApp® in medical education; a scoping review and instructional design model. BMC Med Educ. 2019;19:279.
24. Gurusamy KS, Vaughan J, Fraser IS, Best LM, Richards T. Medical Therapies for Uterine Fibroids – A Systematic Review and Network Meta-Analysis of Randomised Controlled Trials. PLoS One. 2016;11:e0149631.
25. Wagner AK, Soumerai SB, Zhang F, Ross-Degnan D. Segmented regression analysis of interrupted time series studies in medication use research. J Clin Pharm Ther. 2002;27:299—309.
26. Antunes D, Gante I, Carvalho MJ, Medeiros-Borges C, Águas F. The impact of anemia on treatment management and clinical outcomes of women hospitalized for uterine leiomyomas. Ginekol Pol. 2022;93:799—805.
27. Murphy CH, Hess JR. Massive transfusion: red blood cell to plasma and platelet unit ratios for resuscitation of massive hemorrhage. Curr Opin Hematol. 2015;22:533—9.

28. Moynihan R, Sanders S, Michaleff ZA, et al. Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. BMJ Open. 2021;11:e045343.
29. Auerbach M, Adamson JW. How we diagnose and treat iron deficiency anemia. Am J Hematol. 2016;91:31—8.
30. Leahy MF, Hofmann A, Towler S, et al. Improved outcomes and reduced costs associated with a health-system-wide patient blood management program: a retrospective observational study in four major adult tertiary-care hospitals. Transfusion. 2017;57:1347—58.

Brazilian Journal of Anesthesiology 2026;76(6): 844837



Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY

ORIGINAL INVESTIGATION

# Remifentanil for modified rapid sequence induction in children: a proof-of-concept study in a selected pediatric population

Julien Pico [a,b,*], Chrystelle Sola [a,c], Laurent Hertz [a], Julien Riera [a], Christopher Scott [a], Kévin Chapron [a], Philippe Pirat [a], Sophie Bringuier [b,d], Christophe Dadure [a,c]

[a] *Montpellier University Hospital, Department of Maternal, Child and Women's Anaesthesiology and Intensive Care Medicine, Paediatric Anaesthesia Unit, Montpellier, France*
[b] *University of Montpellier, Debrest Institute of Epidemiology and Public Health IDESP, UMR INSERM, Montpellier, France*
[c] *University of Montpellier, CNRS, INSERM, Institute of Functional Genomics, Montpellier, France*
[d] *Montpellier University Hospital, Department of Medical Statistics and Epidemiology, Montpellier, France*

Received 14 March 2026; accepted 20 September 2026
Available online 24 September 2026

**KEYWORDS**
Airway management;
Atropine;
Intubation, intratracheal;
Pediatric anesthesia;
Rapid sequence induction and intubation;
Remifentanil

## Abstract

*Introduction:* Rapid Sequence Induction (RSI) is widely used in pediatric anesthesia for children at risk of pulmonary aspiration. Modified RSI has been proposed to reduce the risk of hypoxemia during induction while maintaining airway protection. This proof-of-concept study aimed to describe peri-intubation outcomes associated with the use of intravenous bolus remifentanil during modified RSI in a selected pediatric population at risk of pulmonary aspiration.

*Methods:* This proof-of-concept study analyzed data from pediatric patients, with at least one risk factor for pulmonary aspiration, undergoing modified RSI with remifentanil between December 2021 and August 2023. Exclusion criteria included the planned use of a neuromuscular blocking agent for RSI, preoperative hypoxemia, hemodynamic shock or difficult intubation risk factors. Remifentanil was administered by direct intravenous injection with optional prophylactic administration of atropine. Data on intubation success rates, major and minor complications and rescue treatment were collected and analyzed.

*Results:* A total of 267 children were included (median age 8.0 years [IQR 5.0–12.0]). Successful first-attempt intubation without major complications was achieved in 92.9% (248/267; 95% CI 88.9–95.6). No major complications were reported. Nineteen patients (7.1%) required more than one intubation attempt. In multivariable analysis, age < 1-year (adjusted OR = 11.4, 95% CI 3.5–34.3; $p < 0.001$) and Cormack–Lehane grade ≥ 2 (adjusted OR = 10.2, 95% CI 1.8–57.2; $p = 0.008$) were associated with failed first-attempt intubation. Minor complications occurred in 15.7% of patients, mainly transient hypotension.

* Corresponding author.
*E-mail:* j-pico@chu-montpellier.fr (J. Pico).

https://doi.org/10.1016/j.bjane.2026.844837
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

*Conclusion:* In this proof-of-concept study of a selected pediatric population, modified RSI with remifentanil was associated with a first-attempt intubation success rate of 92.9%, with no major complications observed. These descriptive findings should be considered hypothesis-generating and support further evaluation of remifentanil-based modified RSI in prospective randomized studies.
© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

As pulmonary aspiration remains a potentially devastating complication of General Anesthesia (GA) in all age groups,[1] Rapid Sequence Induction (RSI) is a standard procedure for tracheal intubation in children at risk of aspiration or regurgitation of gastric contents. Current data from the APRICOT study[2] show that almost 11% of pediatric general anesthesia is performed by RSI, making this a regular practice of pediatric anesthesiologists.

As in adults, the aim of RSI in the pediatric population is to achieve adequate intubation conditions and to minimize the delay between loss of consciousness and airway protection by tracheal intubation. Reduced apnea tolerance combined with reduced functional residual capacity[3] in children leads to an increased risk of hypoxemia: the younger the child, the greater the risk of desaturation.

The French guidelines, entitled "*Management of the child's airway under anesthesia*",[4] recommend using a rapid intravenous neuromuscular blocking agent during a conventional RSI (Grade 1+), with succinylcholine as the first choice. Experts introduce the concept of a "modified" or "controlled" RSI with a gentle mask ventilation not exceeding 15 cm $H_2O$ to maximize oxygen reserves and to reduce the risk of hypoxemia before intubation.

Several authors suggested that modified RSI may better balance the risks of pulmonary aspiration and hypoxemia in pediatric anesthesia.[5,6] Modified RSI therefore represents an adaptation of conventional RSI intended to reduce oxygen desaturation during induction while maintaining airway protection strategies. The choice of neuromuscular blockade during RSI remains a separate consideration, and current guidelines continue to recommend the use of a rapid-acting neuromuscular blocking agent.[4,7,8]

Although rapid-acting neuromuscular blocking agents remain recommended during RSI, remifentanil has been used in our institution as part of a modified RSI strategy in selected pediatric patients without routine neuromuscular blockade.

Remifentanil is an ultra-short-acting opioid with rapid elimination through plasma esterases and favorable pharmacokinetic properties in children, particularly in neonates.[9]

Remifentanil used as an IV injection of 3 to 4 $\mu g.kg^{-1}$ appears to offer optimal intubation conditions based on a scoring system[10] in a pediatric population without causing severe cardiovascular depression or signs of muscular rigidity.[11] However, previous pediatric studies excluded emergency procedures and children at risk of pulmonary aspiration.

The aim of this study was to describe peri-intubation outcomes associated with intravenous bolus remifentanil during modified RSI in a selected pediatric population, assessed by first-attempt intubation success and the occurrence of peri-intubation complications.

## Methods

### Study design

We conducted a retrospective analysis of prospectively recorded data, in the department of pediatric anesthesia. Anonymized anesthesia data from children who received remifentanil for modified RSI between December 2021 and August 2023 were collected.

### Population

Children aged up to 18 years old requiring RSI were eligible for enrolment. For the present retrospective analysis, we included consecutive children who received IV remifentanil bolus for a modified RSI procedure, with at least one of the following risk factors for pulmonary aspiration: bowel obstruction, recent vomiting or nausea, post traumatic ileus, severe chronic gastric dysfunction, insufficient preoperative fasting period (less than 6 hours). Exclusion criteria included the planned use of neuromuscular blocking agent for RSI, preoperative hypoxemia, hemodynamic shock, or difficult intubation risk factors.

### Usual per operative procedures

On arrival to the operating room, an IV catheter was inserted, if not already present. Baseline blood pressure, heart rate and pulse oximetry were routinely recorded. Preoxygenation was performed with 100% oxygen using a tightly fitting face mask until an adequate end-expiratory oxygen concentration was achieved, according to institutional practice and current pediatric airway management guidelines. Remifentanil was administered as a rapid intravenous injection following hypnotic administration. No predefined duration or specific administration rate was standardized in this observational setting. Laryngoscopy was generally performed approximately 60 seconds after induction, once loss of consciousness and adequate intubating conditions had been achieved. Gentle mask ventilation is part of the modified RSI approach and may be performed at the discretion of the anesthesiologist during induction. When applied, inspiratory pressure was limited to 15 cm $H_2O$. The use and frequency of mask oxygenation were not specifically recorded in the database. Intubation was performed by senior anesthesiologists, anesthesia residents with more than one year of training, or nurse anesthetists. Residents performed laryngoscopy under the direct supervision of a senior pediatric

anesthesiologist. The choice to use prophylactic IV atropine (20 $\mu$g.kg$^{-1}$) during induction to prevent bradycardia was made by the anesthesiologist in charge.

In the event of an early complication (within ten minutes of RSI), the use of rescue drugs or emergency ventilation techniques was observed.

## Collected data

All data included in the study come from the "Vigilance Anesthetic in the PeriOperative period" (VAPO) questionnaire. The VAPO is a prospective collection of perioperative anesthetic data, integrated directly into the computerized patient chart. The VAPO has been used in our pediatric center for several years for all anesthesia procedures to document the perioperative anesthetic management, monitor patients at risk of complications and track post-operative outcomes. A special section was added to obtain data on the use of IV bolus remifentanil in the RSI procedure.

The study was based on prospectively collected data routinely recorded in the institutional VAPO database, an institutional quality assurance registry used for all pediatric anesthesia procedures as part of standard clinical practice and quality monitoring. Approval for the retrospective scientific analysis of these anonymized prospectively collected data was obtained in April 2022 (IRB Montpellier University Hospital – IRB-MTP_2022_04_202201075). No consent collection was required for this retrospective observational study.

Patient characteristics included: age, sex, weight, American Society of Anesthesiologists classification (ASA). Surgery data included the timing of surgery (urgent or scheduled) and the type of surgery (urologic, digestive, orthopedic, neurosurgery or other). Intubation data included: risk factors for pulmonary aspiration, induction drugs and their doses, use of prophylactic atropine during induction, use of emergency drugs after intubation. The use of direct or indirect laryngoscopy, Cormack-Lehane grade, number of laryngoscopies attempts prior to intubation, vocal cord position during the intubation procedure, cough response during intubation and operator satisfaction (very satisfied, satisfied, moderately satisfied or unsatisfied) were also reported.

## Outcomes

The primary outcome was successful tracheal intubation at the first attempt without major complications following modified RSI with IV bolus remifentanil used in combination with a hypnotic agent. A successful intubation at the first attempt was defined as a successful endotracheal tube placement during the first laryngoscope insertion. Complications were collected up to ten minutes after the administration of induction drugs. Major complications were defined as: hypotension requiring more than one injection of ephedrine and IV crystalloid administration (20 mL.kg$^{-1}$), bradycardia requiring more than one dose of IV atropine (20 $\mu$g.kg$^{-1}$), hypoxemia below 90% requiring face mask ventilation prior to intubation or manual ventilation following intubation, pulmonary aspiration. Minor complications were defined as: an episode of hypotension reversible after just one injection of ephedrine or IV crystalloid fluid administration (10 mL.kg$^{-1}$), an episode of reversible bradycardia after a single dose of atropine or oxygen desaturation below 90%, quickly reversible after intubation and initiation of ventilation.

The secondary outcomes were to identify factors associated with failed first-attempt intubation and to describe peri-intubation complications.

## Statistical analysis

### Primary outcome

The proportion of successful tracheal intubations at the first attempt without major complications was presented as percentage and its 95% Confidence Interval (95% CI).

### Secondary outcomes

Failed intubation at first attempt: To evaluate factors associated with failed intubation at first attempt, a univariate analysis was performed on patient characteristics, the timing and the type of the surgery, remifentanil and propofol doses and use of prophylactic atropine. Count data, such as the number of laryngoscopy attempts, were compared using Poisson regression. The comparison between groups (success *vs* failed at first attempt) on categorical data were performed using the $\chi^2$ test or the Fisher exact test, as appropriate. Continuous data were compared with the Wilcoxon test. Because of the limited number of first-attempt failures, multivariable logistic regression was restricted to a parsimonious model including three clinically selected covariates: age ($<$ 1-year vs. $\geq$ 1-year), Cormack-Lehane (grade $\geq$ 2 vs. grade 1) and remifentanil (dose $<$ 3.1 vs. $\geq$ 3.1). Associations with failed intubation at first attempt were presented as adjusted Odds Ratios (OR) with 95% confidence intervals.

Descriptive continuous variables were presented as means $\pm$ SD or medians (1Q, 3Q) and categorical data as numbers (%). p-values below 0.05 were considered statistically significant. The statistical analysis was carried out using SAS Enterprise version 8.3 (SAS Institute, Cary, NC).

## Results

We included 267 children at risk for pulmonary aspiration who underwent general anesthesia using a modified RSI procedure.

The median age was 8.0 years (IQR 5.0–12.0). Age ranged from one day to sixteen years, including 21 (7.9%) children younger than 1-year. Patient characteristics are presented in Table 1. The main risk factors for pulmonary aspiration were bowel obstruction, recent vomiting or post traumatic ileus, reported in 81.3% of cases (217 patients). Digestive or urologic surgery was the most frequent procedure, reported in 56.9% of cases (152 patients). The hypnotic agent used in combination with remifentanil was propofol in all patients. The median (IQR) dose of bolus remifentanil was 3.1 (2.9–3.5) $\mu$g.kg$^{-1}$. Direct laryngoscopy was used in 208 cases (78.2%).

Successful intubation at the first attempt without major complications was accomplished in 92.9% (95% CI 88.9–95.6) of children (248/267 patients). No major complications were observed during the study period.

**Table 1** Patient characteristics.

| Characteristics | Remifentanil n = 267 |
|---|---|
| Age, median (Q1, Q3), years | 8.0 (5.0—12.0) |
| < One year n (%) | 21 (7.9) |
| ≥ One year n (%) | 246 (92.1) |
| Male n (%) | 172 (64.4) |
| Weight, median (Q1, Q3), kg | 26.5 (18.0—41.3) |
| American Society of Anesthesiologists Physical Status Classification System score (ASA) n (%) | |
| I | 179 (67) |
| II | 53 (19.9) |
| III | 33 (12.4) |
| IV | 2 (0.8) |
| Timing of surgery n (%) | |
| Urgent or semi-urgent | 249 (93.3) |
| Scheduled | 18 (6.7) |
| Type of surgery n (%) | |
| Digestive, urologic | 152 (56.9) |
| Orthopedic | 73 (27.3) |
| Others: Neurosurgery, endoscopic procedures | 42 (15.7) |
| Risk factors for pulmonary aspiration n (%) | |
| Bowel obstruction, recent vomiting | 154 (57.7) |
| Post-traumatic ileus | 63 (23.6) |
| Others: severe chronic gastric dysfunction, insufficient preoperative fasting period | 50 (18.3) |
| Remifentanil dose, median (Q1, Q3), $\mu g.kg^{-1}$ | 3.1 (2.9—3.5) |
| Propofol dose, median (Q1, Q3), $mg.kg^{-1}$ | 6.0 (5.1—7.5) |
| Prophylactic atropine n (%) | 114 (42.7) |
| Laryngoscopy n (%) (n = 266) | |
| Direct Laryngoscopy | 208 (78.2) |
| Video Laryngoscopy | 58 (21.8) |
| Cormack-Lehane grade n (%) (n = 266) | |
| 1: Full glottis view | 258 (97) |
| 2: Partial view of the glottis | 7 (2.6) |
| 3: Only epiglottis seen, no view of the glottis seen | 1 (0.4) |

Values are presented as mean ± SD, median (Q1, Q3), or number (%).

Patient and procedural characteristics according to first-attempt tracheal intubation success are presented in Table 2. Intubation failed at the first attempt in 19 children (7.1%). Successful intubation required two and three attempts in 16 (6.0%) and 3 (1.1%) patients, respectively. In the multivariable analysis, infant age < 1-year (adjusted OR = 11.4; 95% CI 3.5—34.3; p < 0.001) and Cormack—Lehane grade ≥ 2 (adjusted OR = 10.2; 95% CI 1.8—57.2; p = 0.008) were associated with failed first-attempt intubation, whereas remifentanil dose ≥ 3.1 $\mu g.kg^{-1}$ was not (adjusted OR = 1.3; 95% CI 0.45—3.7; p = 0.63).

Vocal cords were in the abducted position in 245/264 (92.8%) children, and airway reactivity during tube passage was noted in 67/260 (25.8%) children. Operators rated intubation conditions as very satisfactory in 210/257 (81.7%) children and satisfactory in 39/257 (15.2%) children.

Forty-two (15.7%) children had at least one minor complication during the first 10 minutes following induction. Thirty-eight (14.2%) patients presented an episode of minor hypotension, which was corrected by a single injection of ephedrine without IV crystalloid fluid administration. Minor hypotension occurred in 1 of 21 children aged less than one year (4.8%) and in 37 of 246 children aged one year or older (15.0%). Three episodes (1.1%) of bradycardia were recorded in older children: two were 13 years old and the third was 16 years old. Two episodes of bradycardia were associated with minor hypotension. No episode of bradycardia was observed in children aged less than one year. Three patients (1.1%) presented an episode of desaturation below 90%. All were aged less than one year. For two children, the VAPO reported an episode of bronchospasm linked to selective intubation, quickly corrected by adjusting the tracheal tube position. The third episode of desaturation was linked to difficulties with tracheal intubation, requiring a change of operator.

## Discussion

To our knowledge, this study represents one of the first observational cohorts evaluating remifentanil in combination with propofol for modified RSI in pediatric patients at risk of pulmonary aspiration. The results of this study showed a high success rate of 92.9% for tracheal intubation at the first attempt without major complications. These findings provide preliminary descriptive data regarding peri-intubation outcomes associated with this modified RSI approach in a selected pediatric population.

**Table 2** Patient and procedural characteristics according to first-attempt tracheal intubation success.

| Characteristics | Successful intubation at the first attempt (n = 248) | Failure intubation at the first attempt (n = 19) | p-value |
|---|---|---|---|
| Age, median (1Q, 3Q), years | 8.00 (5.00; 12.00) | 1.00 (0.08; 8.00) | < 0.01 |
| < One year n (%) | 13 (5.24) | 8 (42.11) | < 0.01 |
| Weight, median (1Q, 3Q), kg | 27.10 (18.80; 41.70) | 10.61 (3.83; 20.75) | < 0.01 |
| Male n (%) | 157 (63.31) | 15 (78.95) | 0.17 |
| ASA score n (%) | | | < 0.01 |
| I | 172 (69.35) | 7 (36.84) | |
| II | 48 (19.35) | 5 (26.32) | |
| III | 27 (10.89) | 6 (31.58) | |
| IV | 1 (0.40) | 1 (5.26) | |
| Timing of surgery n (%) | | | 0.13 |
| Scheduled | 15 (6.05) | 3 (15.79) | |
| Urgent or semi-urgent | 233 (93.95) | 16 (84.21) | |
| Type of surgery n (%) | | | 0.04 |
| *Orthopedic surgery* | 72 (29.03) | 1 (5.26) | |
| *Digestive, urologic surgery* | 139 (56.05) | 13 (68.42) | |
| *Other surgeries* | 37 (14.92) | 5 (26.32) | |
| Risk factors for pulmonary aspiration n (%) | | | |
| *Bowel obstruction, recent vomiting* | 140 (56.45) | 14 (73.68) | 0.14 |
| *Post-traumatic ileus* | 62 (25.00) | 1 (5.26) | 0.05 |
| *Other risk factors* | 46 (18.55) | 4 (21.05) | 0.76 |
| Direct laryngoscopy n (%) | 194 (78.54) | 14 (73.68) | 0.57 |
| Cormack-Lehane grade n (%) | | | < 0.01 |
| 1 | 243 (98.38) | 15 (78.95) | |
| 2 | 4 (1.62) | 3 (15.79) | |
| 3 | 0 (0.00) | 1 (5.26) | |
| Prophylactic Atropine n (%) | 104 (41.94) | 10 (52.63) | 0.36 |
| Remifentanil dose, median (1Q, 3Q), $\mu g.kg^{-1}$ | 3.10 (2.90; 3.40) | 3.40 (2.80; 3.90) | 0.36 |
| < 3.1 $\mu g.kg^{-1}$ n (%) | 136 (54.84) | 12 (63.16) | 0.48 |
| Propofol dose, median (1Q, 3Q), $mg.kg^{-1}$ | 6.00 (5.10; 7.50) | 6.00 (4.20; 7.00) | 0.33 |

Values are presented as median (Q1, Q3), or number (%).

Few studies have evaluated the use of IV bolus remifentanil for modified RSI in pediatric patients, particularly in children at risk of pulmonary aspiration. Previous studies,[10,11] which reported excellent intubation conditions with remifentanil used in pediatrics, systematically excluded emergency procedures, often associated with a potential risk of aspiration. In addition, isolated pediatric case reports have previously described the use of propofol-remifentanil combinations for RSI in children at risk of aspiration.[12] Blair et al.[10] showed that remifentanil at a dose of 3 $\mu g{\cdot}kg^{-1}$ provides similar intubation conditions to mivacurium. In the same way, Klemola et al.[11] showed that excellent intubation conditions were more frequent in the group receiving IV remifentanil at a dose of 4 $\mu g.kg^{-1}$ than in the rocuronium group. In accordance, the average dose of bolus remifentanil used in the present study was 3.19 $\pm$ 0.66 $\mu g.kg^{-1}$. In addition, in the current study, the selected population was children at risk of aspiration, for whom practitioners were free to choose a classical or modified RSI protocol, including, or not, the planned use of a neuromuscular blocking agent. In the modified RSI current cohort, this study reported a successful tracheal intubation at the first attempt in 92.9% of patients. Previous pediatric emergency RSI cohorts have reported variable first-attempt success rates, including 78% in a multicenter registry and 52% in a single-center study using systematic video review.[13,14] However, no direct comparison can be made given the important differences in study design, patient populations, clinical settings, and airway management strategies across studies.

Regarding intubation conditions, while the vocal cords were in the abducted position in 92.8% of children, airway reactivity during tube passage through the vocal cords was observed in 25.8% of cases. The VAPO questionnaire did not distinguish between mild airway reactivity, coughing during tracheal tube passage, or reactivity related to cuff inflation. These results appear to be consistent with the study by Hanna et al.[15] who reported no difference in the occurrence of a cough reaction between intubations with suxamethonium 1.2 $mg.kg^{-1}$ and remifentanil 4 $\mu g.kg^{-1}$. Acceptable intubation conditions have been shown to increase with increasing doses of remifentanil.[10,16] Considering the pharmacological properties of remifentanil, both the administered dose and timing between drug administration and laryngoscopy may influence airway reactivity. The time interval between remifentanil administration and laryngoscopy was not recorded in this study, which limits interpretation. Variability in this interval may have contributed to the observed airway reactivity. These findings should be taken into account when interpreting intubating conditions in this cohort and suggest that airway

reflex suppression may be inconsistent under the conditions observed in this study.

In a recent adult study by Grillot et al.[17] remifentanil was compared to suxamethonium for RSI. The remifentanil group did not meet the criterion for non-inferiority regarding successful intubation without major complications (66% vs. 71% in the remifentanil group vs the neuromuscular blocking agent group, respectively). RSI failure was attributed to the inability to intubate on the first attempt, while episodes of major hemodynamic instability, prolonged arrhythmia > 30s or cardiac arrest were reported in 20% of patients with similar rates in the remifentanil and neuromuscular blocking agent groups. Although a high first-attempt intubation success rate was observed in the present cohort, the uncontrolled design of this study does not permit any comparison with alternative RSI strategies. As previously indicated, the pediatric population is significantly distinct from the adult population. Children differ from adults in terms of airway anatomy, oxygen reserve, apnea tolerance, and hemodynamic responses during induction of anesthesia, which may influence both intubation conditions and complication profiles during modified RSI. In the present study, episodes of mild and transient hypotension were the most frequent minor complications, observed in 14.2% of cases. All episodes were corrected by a single injection of ephedrine without IV crystalloid fluid administration. Previous pediatric studies by Blair et al.[10] and Klemola et al.[11] systematically administered prophylactic IV atropine (at doses of 10 $\mu g.kg^{-1}$ or 15 $\mu g.kg^{-1}$) when using remifentanil for induction, did not report hypotension or bradycardia requiring treatment. It should be noted, however, that there is currently no consensus on what constitutes hypotension in the literature on pediatric anesthesia. Wani et al.[18] studied variations in blood pressure in children aged 2 to 8 years in the 12 minutes following the induction of anesthesia to investigate the incidence of post-induction hypotension. The incidence varied considerably based on the definition used, 27.5% if the $5^{th}$ percentile of the systolic blood pressure for age was used versus 57.2% (107 of 189 patients) when a 20% decrease of the systolic blood pressure from baseline was used. In adult classical RSI, a high rate of cardiorespiratory complications was reported, with 18% experiencing hypotension after GA induction.

Finally, infant age < 1-year and Cormack-Lehane grade ≥ 2 were associated with failed first-attempt intubation in the multivariable analysis. These findings are consistent with the pediatric airway literature, which indicates that infants and children with more difficult laryngoscopic views are at increased risk of intubation failure.[19]

Therefore, these results provide additional descriptive data regarding the use of remifentanil for modified RSI in selected pediatric patients managed by teams experienced with this technique.

Limitations of the present study include the absence of a control group, the retrospective analysis of prospectively collected data, and a potential selection bias. Patients with known difficult airway risk factors were excluded, which may explain the high proportion of favorable Cormack-Lehane grades observed in this cohort. This selection process may also have contributed to the high first-attempt intubation success rate observed in the present study and may limit the generalizability of the findings to broader pediatric populations. Furthermore, the absence of a control group precludes any comparison with standard RSI techniques using neuromuscular blocking agents and does not allow conclusions regarding comparative efficacy, safety, or intubating conditions. Although no major complications were observed, the sample size was not sufficient to reliably assess rare adverse events such as pulmonary aspiration. Therefore, the absence of such events should not be interpreted as evidence of safety. However, the study was not designed to compare remifentanil with neuromuscular blocking agents or to establish comparative efficacy, but rather to describe the peri-intubation outcomes associated with a modified RSI approach using remifentanil. Similarly, the outcomes of hypotension and bradycardia were subjective and depended on the decision of treatment by the anesthesiologist, with no predefined threshold. However, as previously indicated, the definition of thresholds is a recurrent debate in pediatrics.[18]

## Conclusion

This study provides additional data describing the peri-intubation outcomes associated with intravenous bolus remifentanil during modified RSI in a selected pediatric population ranging from newborns to adolescents. A first-attempt intubation success rate of 92.9% was observed, and no major complications were reported. However, the observational design, absence of a control group, and limited sample size preclude conclusions regarding comparative safety or efficacy. Larger prospective and controlled studies are required to further evaluate this approach.

## Data availability statement

The individual-level dataset generated and analyzed during the current study originates from our institutional quality assurance database (VAPO) and cannot be made publicly available because of institutional and ethical restrictions. A de-identified dataset, data dictionary, and SAS code used for the analyses may be made available by the corresponding author upon reasonable request, subject to approval by the Institutional Review Board and the data governance policies of Montpellier University Hospital.

## Declaration of generative AI in scientific writing

During the preparation of this work the authors used ChatGPT (OpenAI) in order to improve readability and language. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

## Preprint statement

A previous version of this manuscript was posted as a preprint on medRxiv: Pico J, Sola C, Hertz L, Riera J, Scott C, Chapron K, Pirat P, Bringuier S, Dadure C. Remifentanil as

an alternative in modified rapid sequence induction: A proof-of-concept study in a selected paediatric population. medRxiv. 2024. 10.1101/2024.12.11.24318829.

## Conflicts of interest

The authors declare no conflicts of interest.

## Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

## Associate Editor

Vinicius Caldeira Quintão

## References

1. Kluger MT, Short TG. Aspiration during anaesthesia: a review of 133 cases from the Australian Anaesthetic Incident Monitoring Study (AIMS). Anaesthesia. 1999;54:19—26.
2. Habre W, Disma N, Virag K, et al. Incidence of severe critical events in paediatric anaesthesia (APRICOT): a prospective multicentre observational study in 261 hospitals in Europe. Lancet Respir Med. 2017;5:412—25.
3. Patel R, Lenczyk M, Hannallah RS, McGill WA. Age and the onset of desaturation in apnoeic children. Can J Anaesth. 1994;41:771—4.
4. Dadure C, Sabourdin N, Veyckemans F, et al. Management of the child's airway under anaesthesia: The French guidelines. Anaesth Crit Care Pain Med. 2019;38:681—93.
5. Weiss M, Gerber AC. Rapid sequence induction in children – it's not a matter of time!. Paediatr Anaesth. 2008;18:97—9.
6. Russotto V, Myatra SN, Laffey JG, et al. Intubation Practices and Adverse Peri-intubation Events in Critically Ill Patients From 29 Countries. JAMA. 2021;325:1164—72.
7. Klučka J, Štourač P, Štoudek R, Ťoukálková M, Harazim H, Kosinová M. Controversies in Pediatric Perioperative Airways. Biomed Res Int. 2015;2015:368761.
8. Neuhaus D, Schmitz A, Gerber A, Weiss M. Controlled rapid sequence induction and intubation — an analysis of 1001 children. Paediatr Anaesth. 2013;23:734—40.
9. Ross AK, Davis PJ, Dear Gd GL, et al. Pharmacokinetics of remifentanil in anesthetized pediatric patients undergoing elective surgery or diagnostic procedures. Anesth Analg. 2001;93:1393—401.
10. Blair JM, Hill DA, Wilson CM, Fee JPH. Assessment of tracheal intubation in children after induction with propofol and different doses of remifentanil. Anaesthesia. 2004;59:27—33.
11. Klemola UM, Hiller A. Tracheal intubation after induction of anesthesia in children with propofol – remifentanil or propofol-rocuronium. Can J Anaesth. 2000;47:854—9.
12. Dewhirst E, Tobias JD, Martin DP. Propofol and remifentanil for rapid sequence intubation in a pediatric patient at risk for aspiration with elevated intracranial pressure. Pediatr Emerg Care. 2013;29:1201—3.
13. Sagarin MJ, Chiang V, Sakles JC, et al. National Emergency Airway Registry (NEAR) investigators. Rapid sequence intubation for pediatric emergency airway management. Pediatr Emerg Care. 2002;18:417—23.
14. Kerrey BT, Rinderknecht AS, Geis GL, Nigrovic LE, Mittiga MR. Rapid Sequence Intubation for Pediatric Emergency Patients: Higher Frequency of Failed Attempts and Adverse Effects Found by Video Review. Ann Emerg Med. 2012;60:251—9.
15. Hanna SF, Ahmad F, Pappas ALS, et al. The effect of propofol/remifentanil rapid-induction technique without muscle relaxants on intraocular pressure. J Clin Anesth. 2010;22:437—42.
16. Crawford MW, Hayes J, Tan JM. Dose-response of remifentanil for tracheal intubation in infants. Anesth Analg. 2005;100:1599—604.
17. Grillot N, Lebuffe G, Huet O, et al. Effect of Remifentanil vs Neuromuscular Blockers During Rapid Sequence Intubation on Successful Intubation Without Major Complications Among Patients at Risk of Aspiration: A Randomized Clinical Trial. JAMA. 2023;329:28—38.
18. Wani TM, Hakim M, Ramesh A, et al. Risk factors for post-induction hypotension in children presenting for surgery. Pediatr Surg Int. 2018;34:1333—8.
19. Fiadjoe JE, Nishisaki A, Jagannathan N, et al. Airway management complications in children with difficult tracheal intubation from the Pediatric Difficult Intubation (PeDI) registry: a prospective cohort analysis. Lancet Respir Med. 2016;4:37—48.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



REVIEW ARTICLE

# Awake versus deep airway device removal in pediatric tonsillectomy: a systematic review and meta-analysis

**Henrique Baggio Pivetta [a,*], Alice O. Almeida [b], Rafael L. Miranda [c], Rafaela P. Silva [d], Raphaela G. Mendes [c], João M.M. Borges [e], Raquel E. Saldanha [f], Matheus Arraes [g], Pablo P. Sanín [h], Walid Alrayashi [i]**

[a] *Universidade do Sul de Santa Catarina, Department of Medicine, Palhoça, SC, Brazil*
[b] *Faculdade de Medicina Zarns, Department of Medicine, Salvador, BA, Brazil*
[c] *Universidade Federal de Minas Gerais, Department of Medicine, Belo Horizonte, MG, Brazil*
[d] *Universidade Federal do Maranhão, Department of Medicine, São Luís, MA, Brazil*
[e] *The University of Texas Health Science Center at Houston, Houston, TX, USA*
[f] *Hospital Instituto José Frota, Anesthesiology Department, Fortaleza, CE, Brazil*
[g] *University of Miami Miller School of Medicine, Miami, FL, USA*
[h] *Colsubsidio Cardiovascular Clinic, Bogotá, Colombia*
[i] *Boston Children's Hospital and Harvard Medical School, Anesthesiology Department, Critical Care, and Pain Medicine, Boston, Massachusetts, USA*

Received 7 April 2026; accepted 29 September 2026
Available online 9 October 2026

**KEYWORDS**
Adenoidectomy;
Child;
Extubation;
Intubation, intratracheal;
Laryngeal masks;
Tonsillectomy

## Abstract

*Background and objectives:* Tonsillectomy, with or without adenoidectomy, is among the most common pediatric surgeries, associated with a high incidence of perioperative respiratory adverse events. The timing of airway device removal, during deep anesthesia or after awakening, may influence these outcomes.

*Methods:* PubMed, Embase and CENTRAL were systematically searched for studies comparing awake and deep airway device removal in pediatric otolaryngology surgery. The primary analysis was restricted to Randomized Controlled Trials (RCTs). The primary outcome was laryngospasm; secondary outcomes included cough, oxygen desaturation and airway obstruction. Risk Ratios (RR) with 95% Confidence Intervals (95% CI) were pooled using random-effects models.

*Results:* Seven studies (six RCTs and one observational study, the latter used only in sensitivity analyses), including 1,571 children, were analyzed. Deep airway device removal was associated

*Abbreviations:* ASA, American Society of Anesthesiologists; CI, Confidence Interval; ETT, Endotracheal tube; GRADE, Grading of Recommendations Assessment, Development and Evaluation; LMA, Laryngeal Mask Airway; MH, Mantel-Haenszel; PRAE, Perioperative Respiratory Adverse Events; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROSPERO, International Prospective Register of Systematic Reviews; REML, Restricted Maximum Likelihood; RR, Risk Ratio; RCTs, Randomized Controlled Trials; ROBINS-I, Risk of Bias in Non-randomized Studies of Interventions; RoB2, Revised Cochrane Risk-of-Bias Tool for Randomized Trials; OSA, Obstructive Sleep Apnea; TIVA, Total Intravenous Anesthesia..

* Corresponding author.
*E-mail:* henriquebpivetta@hotmail.com (H.B. Pivetta).

https://doi.org/10.1016/j.bjane.2026.844840
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

with a higher risk of airway obstruction (RR = 1.82; 95% CI 1.10—3.03; p = 0.020) and a lower risk of cough (RR = 0.57; 95% CI 0.43—0.75; p < 0.001), with no evidence of a difference between the Laryngeal Mask Airway (LMA) and Endotracheal Tube (ETT) subgroups (p for subgroup differences = 0.992). Laryngospasm did not differ significantly (RR = 0.99; 95% CI 0.43—2.30; p = 0.984). Desaturation was lower with deep removal in the RCT-only analysis (RR = 0.62; 95% CI 0.44 —0.88; p = 0.007).

*Conclusions:* Deep airway device removal was associated with less coughing, with no evidence that this effect differed between LMA and ETT. Deep removal may increase airway obstruction, though sensitive to individual studies. Desaturation findings were sensitive to study design and are not definitive.

© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

Adenotonsillectomy and tonsillectomy are among the most frequently performed surgical procedures in the pediatric population, primarily indicated for the treatment of upper airway obstruction and recurrent infections of the tonsils and adenoids.[1—6] At the end of the procedure, the removal of airway devices, including LMA and ETT, can be performed either while the patient remains in a deep plane of anesthesia, when spontaneous ventilation is established but airway reflexes have not yet returned, or after full emergence and awakening.[7,8] Both techniques have critical respiratory safety implications, particularly in children, who are more susceptible to laryngospasm, coughing, desaturation, airway obstruction, and stridor due to heightened airway reactivity.[9,10] These risks are further amplified during surgeries involving the upper airway, such as adenotonsillectomy or other oropharyngeal procedures, in which the presence of blood and secretions act as potent airway irritants.[9,11,12] These complications highlight the importance of careful anesthetic management during adenotonsillectomy and tonsillectomy procedures in children.[6]

The transition to Total Intravenous Anesthesia (TIVA) and sevoflurane largely replaced older volatile agents like halothane and enflurane. This shift prioritized superior safety, faster and smoother recoveries, and lower toxicity, addressing specific clinical challenges associated with the older medications. Understanding whether modern anesthetic practice mitigates or exacerbates the classic risks associated with each extubation technique is essential for evidence-based perioperative management, particularly in high-risk patients.

However, evidence comparing the available methods of airway device removal (deep vs awake) in tonsillectomy and adenotonsillectomy patients remains conflicting.[8,13,14] In contrast to a 2018 meta-analysis, which evaluated airway device removal timing in a general pediatric population without surgical stratification, the present study focuses specifically on otolaryngology procedures and incorporates recent evidence to provide an updated synthesis.[8] This meta-analysis aims to clarify whether removal of the LMA or ETT under deep anesthesia or after full emergence results in superior respiratory safety and recovery outcomes in pediatric patients.

Importantly, much of the foundational literature addressing extubation depth in pediatric tonsillectomy was conducted during earlier anesthetic eras, prior to widespread use of sevoflurane and TIVA. As a result, existing data may offer mechanistic insight rather than direct guidance for contemporary practice.

## Methods

This systematic review and meta-analysis was pre-registered on the International Prospective Register of Systematic Reviews (PROSPERO) with protocol number CRD420251084368. Institutional ethics approval was waived. Data registration, conduct of the meta-analysis, and reporting adhered to the guidelines outlined in the Cochrane Handbook for Systematic Reviews of Interventions and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[15,16] The complete PRISMA checklist is available in Supplementary Methods 1.[17]

### Search strategy

A comprehensive search was conducted in MEDLINE, Embase, and the Cochrane Central Register of Controlled Trials (CENTRAL) from inception to September 2026. The complete search strategy is provided in Supplementary Methods 2. Two reviewers independently screened studies for eligibility and extracted data using a predefined extraction form. Discrepancies were resolved by consensus.

### Eligibility criteria

Studies were included if they met all of the following: 1) RCTs were eligible for the primary meta-analysis; 2) Comparing awake and deep removal of airway devices; 3) Conducted in pediatric patients undergoing tonsillectomy and adenotonsillectomy; and 4) Reporting at least one predefined outcome. Observational studies meeting criteria (2) through (4) were considered separately and included only in sensitivity analyses.

Articles were excluded if they (1) lacked a control group; (2) were conference abstracts; or (3) involved overlapping study populations. No date restrictions were applied.

### Outcome assessments

The primary outcome was laryngospasm. Secondary outcomes included cough, oxygen desaturation, and airway obstruction. Laryngospasm was defined as complete upper airway obstruction with absence of airflow, associated with inspiratory stridor

and paradoxical chest wall movement. Cough was considered when persistent or severe (> 10s) coughing occurred during emergence or in the post-anesthesia care unit. Oxygen desaturation was recorded continuously by pulse oximetry, typically defined as $SpO_2$ below 92%–95% for more than 30 seconds or requiring supplemental oxygen. Airway obstruction was identified clinically as partial obstruction accompanied by snoring, increased respiratory effort, or the need for airway maneuvers (chin lift or jaw thrust). Outcomes were assessed during the immediate postoperative period. Datawere pooled in meta-analyses when at least three studies reported a given outcome.

## Study selection and data extraction

Two reviewers (H.B.P. and A.O.A.) independently screened the studies in two phases: initial title and abstract screening, followed by full-text review. Discrepancies were resolved through discussion or consultation with a third reviewer (R.P.S.). Two authors (R.G.M. and J.M.M.B.) independently extracted the data for each study using a standardized study form to determine: authors, enrollment period, study publication year, inclusion and exclusion criteria, sample size, follow-up period, baseline patient characteristics, endpoint data and endpoint definition. Disagreements were resolved in a panel discussion with the first author (H.B.P.).

Additionally, all included studies and the references used were screened against the Retraction Watch Database (retractionwatch.com), PubMed/MEDLINE, publisher websites and CrossMark were consulted in September 2026. No retractions or expressions of concern were identified among the studies included.

## Statistical analysis

Risk Ratio (RR) was used for dichotomous outcomes, estimated via the Mantel-Haenszel (MH) method. Pooled estimates were calculated using random-effects models with REML estimators for $\tau^2$ and 95% Confidence Intervals (95% CIs), considering potential variations in clinical and methodological factors across studies. Statistical significance was set at $p < 0.05$, with emphasis on effect size magnitude and precision. Analyses were conducted in RStudio (v4.5.0; RStudio, PBC) using the meta and metafor packages.

The primary meta-analysis was restricted to randomized controlled trials to minimize confounding and intervention-classification bias. The observational study was not included in the primary pooled estimates but was incorporated in a sensitivity analysis when outcome data were available. Additional exploratory subgroup analyses were performed according to airway device type and sevoflurane-based maintenance anesthesia.

Although ETT and LMA differ in their anatomical positioning, sealing mechanisms, and degree of airway stimulation, both devices share the same fundamental intervention under investigation, the timing of removal relative to depth of anesthesia. Furthermore, both devices are routinely used in the same surgical population (pediatric tonsillectomy), and the respiratory outcomes assessed are clinically relevant regardless of device type. To address the clinical heterogeneity introduced by device differences, subgroup analyses stratified by airway device type were performed for all outcomes, and results are interpreted in the context of these subgroups rather than the pooled estimate alone. The statistical significance of differences between subgroups was assessed with the test for subgroup differences (Cochran's *Q*), rather than by comparing the statistical significance of the individual subgroup estimates.

## Deviations from the registered protocol

Three minor deviations from the registered protocol are acknowledged. First, airway obstruction was added as a secondary outcome after protocol registration, based on its clinical relevance to the surgical population studied. Second, although the registered protocol specified inclusion of studies published in English, Spanish, Portuguese and Chinese, no language restrictions were applied during the search to maximize sensitivity and minimize publication bias; this decision was made prior to study selection. Third, the subgroup analyses by airway device type and anesthetic maintenance agent (sevoflurane) were not prespecified in the registration and should therefore be considered exploratory.

This subgroup included studies in which sevoflurane was used for anesthetic maintenance, regardless of induction agent or concomitant use of nitrous oxide. For most outcomes, only randomized controlled trials contributed to this subgroup analysis. For the desaturation outcome, Baijal et al. was the single observational study in this review, also included in the sevoflurane subgroup, as sevoflurane-based maintenance was used in that study.

## Publication bias assessment

Two independent reviewers assessed the risk of bias in the individual trials and the methodological quality of the included studies (H.B.P. and M.A.). Discrepancies were resolved by consensus. To evaluate the risk of bias in RCTs, we used the Cochrane Risk of Bias 2 (RoB2) tool (Suppl. Table 1).[18] For each domain, we assigned a judgement of "high risk", "some concerns" or "low risk" of bias. For the observational study, the Risk of Bias in Non-Randomized Studies of Interventions (ROBINS-I) tool was applied,[19] assessing domains such as selection bias, performance bias, detection bias, attrition bias, and reporting bias (Suppl. Table 2).

The overall certainty of the evidence was evaluated using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.[20] Studies were rated as providing very low-, low-, moderate-, or high-quality evidence based on the presence of risk of bias, inconsistency of results, imprecision, publication bias, and the magnitude of treatment effects.

## Assessment of heterogeneity

Heterogeneity was assessed by visual inspection, the $I^2$ statistic, and Cochran's *Q*-test. Given the small number of included studies, $I^2$ values should be interpreted with caution, as they may not reliably distinguish true homogeneity from insufficient statistical power to detect heterogeneity. Leave-one-out analyses were conducted with all studies. Sensitivity analyses incorporating the observational study were performed to explore the robustness of the findings and potential sources of heterogeneity.

## Results

### Study selection and characteristics

The initial literature review identified 739 records. After removal of duplicates and exclusion based on title and abstract screening, 13 articles were selected for full-text review. Of these, 3 were available only as conference abstracts and 3 were trial registries without published results. Consequently, 7 studies met the inclusion criteria and were included in the meta-analysis, six RCTs[10,11,21–24] contributed to the primary analysis and one observational study was used only in sensitivity analyses.[25]

A total of 1,571 patients with a mean age of 4.88 years (ranging from 1 to 19 years) underwent tonsillectomy or adenotonsillectomy, of whom 647 (41.2%) were extubated awake. The male proportion varied between 41% and 72%, and the weight range across studies spanned from approximately 12 kg to 26 kg. Four studies included patients undergoing adenotonsillectomy,[10,23–25] whereas three studies included tonsillectomy with or without adenoidectomy.[11,21,22] Five studies employed an ETT for airway management,[10,22–25] while two studies utilized LMA.[11,21] Baseline physical status was generally low risk. Four studies included children classified as ASA I–II, and one study was limited to ASA I patients. One study enrolled patients with ASA physical status I–III. ASA classification was not reported in the observational study. Sevoflurane was the most commonly used induction and maintenance anesthetic agent, reported in five studies.[10,11,21,24,25] One study[22] employed halothane for both induction and maintenance of anesthesia while another[23] used enflurane for both purposes. Four studies reported the presence of OSA,[10,11,24,25] with a total of 385 children with obstructive sleep apnea in the awake extubation group and 264 who underwent deep extubation. Study characteristics are summarized in Table 1, and the PRISMA flow diagram is shown in Figure 1.

### Primary analysis: randomized controlled trials

No significant differences were identified between groups in the incidence of laryngospasm (RR = 0.99; 95% CI 0.43–2.30; p = 0.984; $I^2$ = 0%; Fig. 2A). The incidence of cough (RR = 0.57; 95% CI 0.43–0.75; p < 0.001; $I^2$ = 28.3%; Fig. 2C) and desaturation (RR = 0.62; 95% CI 0.44–0.88; p = 0.007; $I^2$ = 32.6%; Fig. 2B) were significantly lower in the deep extubation group compared with the awake group. When stratified by airway device, no significant difference was observed among patients extubated with an ETT (RR = 0.54; 95% CI 0.27–1.06; p = 0.072; $I^2$ = 44.6%; Fig. 2C), whereas the reduction in cough incidence remained statistically significant in the LMA subgroup (RR = 0.53; 95% CI 0.30–0.96; p = 0.035; $I^2$ = 35.3%; Fig. 2C), and the test for subgroup differences provided no evidence of effect modification by airway device (p = 0.992). In contrast, airway obstruction occurred more frequently following deep extubation (RR = 1.82; 95% CI 1.10–3.03; p = 0.020; $I^2$ = 0%; Fig. 2D).

Subgroup analyses by type of airway device yielded similar findings: among studies using an ETT, neither laryngospasm (RR = 2.22; 95% CI 0.41–12.16; p = 0.356; $I^2$ = 0%; Fig. 2A) nor desaturation (RR = 0.77; 95% CI 0.49–1.21; p = 0.254; $I^2$ = 0%; Fig. 2B) differed significantly between deep and awake extubation.

**Table 1** Baseline characteristics of included studies.

| Study, year | Study design | Surgery | ASA-PS | Airway device | Awake/ Deep, n | Age, y Awake/Deep[a] | Weight, kg Awake/Deep[a] | Male, n (%) | Induction | Maintenance | Diagnosis of OSA, n Awake/Deep |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Patel, 1991 | RCT | T/A | I–II | ETT | 18/20 | 4.04 (3.07–5.01)<br>4.65 (3.54–5.76) | 17.0 (7.6–26.4)<br>18.4 (10.8–26.0) | NI | $N_2O$, $O_2$, and halothane or IV thiamylal | Halothane; $N_2O$ and $O_2$ | NI |
| Hong, 1997 | RCT | T&A | I | ETT | 24/25 | 5.8 (4.30–15.95)<br>6.4 (4.80–8.00) | 20.8 (17.09–24.52)<br>22.4 (18.96–25.85) | 36 (72) | Enflurane | Enflurane, $O_2$ and air | NI |
| von Ungern-Sternberg, 2013 | RCT | T&A | I–III | ETT | 50/50 | 5 (1–14)<br>4 (1–14) | 20 (12.7–84.9)<br>18 (12.9–64.0) | 41 (41) | Sevoflurane or IV propofol | Sevoflurane and $N_2O$ | 34/36 |
| Baijal, 2015 | Observational | T&A | NI | ETT | 303/577 | 5.0 (1.00–17.00)<br>4.0 (1.00–19.00) | 19.4 (6.00–157.00)<br>18.56 (6.7–99.5) | NI | Sevoflurane | Sevoflurane | 225/94 |
| Ramgolam, 2018 | RCT | T w/ wo A | I–II | LMA | 140/143 | 6.8 (1.0–16.3)<br>6.4 (1.1–16.0) | 25.3 (11.0–84.6)<br>23.4 (9.9–99.0) | 133 (47) | Sevoflurane or IV propofol | Sevoflurane | 76/87 |
| Kapoor, 2025 | RCT | T w/ wo A | I–II | LMA | 62/62 | 4.54 (1.33–7.00)<br>4.63 (2.00–6.75) | 19.60 (3.00–36.70)<br>18.65 (11.30–36.30) | 72 (58) | Sevoflurane w/ wo $N_2O$ | Sevoflurane | NI |
| Wan, 2026 | RCT | T&A | I–II | ETT | 50/47 | 5.18 ± 1.61<br>5.13 ± 1.53 | 23.2 ± 9.29<br>21.1 ± 5.76 | 63 (65) | Sevoflurane with IV propofol | Sevoflurane | 50/47 |

A, Adenoidectomy; ASA-PS, American Society of Anesthesiologists Physical Status Classification; ETT, Endotracheal Tube; IV, Intravenous; LMA, Laryngeal Mask; NI, Not Informed; OSA: Obstructive Sleep Apnea; RCT, Randomized Controlled Trial; SD, Standard Deviation; T&A, Adenotonsillectomy; T/A, Tonsillectomy and/or Adenoidectomy; T, Tonsillectomy; w/wo, with/without.

[a] Mean ± SD or Median (range).

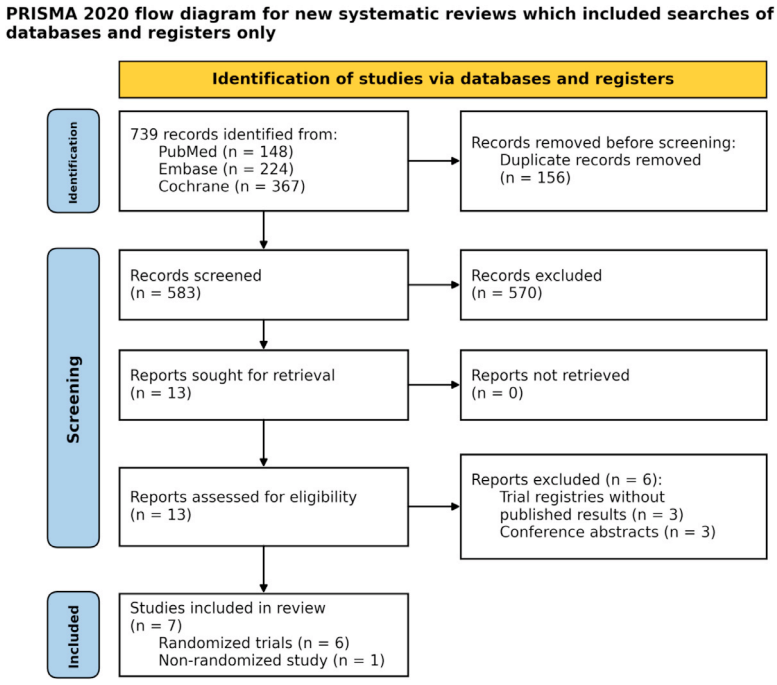



**Figure 1** PRISMA 2020 flowchart.

## Sensitivity analysis including the observational study

When the observational study by Baijal et al. was included in the sensitivity analysis, the pooled estimate for desaturation was attenuated and became non-significant (RR = 0.72; 95% CI 0.51—1.02; p = 0.063; $I^2$ = 44.5%; Fig. 3), suggesting that study design may have influenced this outcome. Laryngospasm, cough, or airway obstruction were not influenced by Baijal et al.

## Exploratory subgroup analysis with sevoflurane-based maintenance anesthesia

In the subgroup analysis of patients with sevoflurane-based maintenance anesthesia,[10,11,21,24,25] the risk of laryngospasm showed no significant difference between the awake and deep extubation groups (RR = 0.79; 95% CI 0.32—1.96; p = 0.60; $I^2$ = 0%; Fig. 4A). Regarding airway obstruction, the results showed no significant difference between the awake and deep group (RR = 1.92; 95% CI 0.96—3.84; p = 0.06; $I^2$ = 38.8%; Fig. 4D). In contrast, the incidence of cough was significantly higher in the awake group (RR = 0.55; 95% CI 0.42—0.73; p < 0.001; $I^2$ = 19%; Fig. 4C), while for desaturation, the only outcome with the observational study, rates showed no statistically significant difference between the two techniques (RR = 0.67; 95% CI 0.44—1.04; p = 0.07; $I^2$ = 61%; Fig. 4B).

## Leave-one-out analysis

Sensitivity analyses using the leave-one-out approach demonstrated that the overall findings were generally robust for most outcomes, with a notable exception for airway obstruction. The leave-one-out analysis for laryngospasm yielded no significant new findings (Suppl. Fig. 1). In the analysis of desaturation, results were consistent with the main findings, except when Baijal et al. was excluded, which rendered the difference statistically significant in favor of deep extubation; both Baijal et al. and Ramgolam et al. were flagged as outliers for this outcome (Suppl. Fig. 2). For cough, statistical significance persisted regardless of the study omitted, although Patel et al. was identified as a potential outlier (Suppl. Fig. 3). For airway obstruction, the pooled effect lost significance when either Hong et al. or von Ungern-Sternberg et al. was omitted, indicating that the significant pooled effect for this outcome is driven by these two studies and should be interpreted with caution (Suppl. Fig. 4).

## Risk of Bias

The assessment of publication bias in this review was constrained by the small number of eligible studies (n = 7). Because statistical methods such as funnel plot inspection and Egger's regression are unreliable with fewer than 10 studies, a formal quantitative evaluation was not conducted. Instead, we relied on a qualitative appraisal of risk across the included studies.

Among the RCTs, three raised "some concerns" regarding the risk of bias; these were primarily related to the randomization process (Hong, 1997), deviations from intended intervention (Wan, 2026), selection of the reported result (Wan, 2026) and measurement of the outcome (Kapoor, 2025; Wan, 2026). The remaining three RCTs (Ramgolam, 2018; von Ungern-Sternberg, 2013; and Patel, 1991) were considered to have a low risk of bias. The observational study (Baijal, 2015) was judged to carry a higher risk, particularly due to potential confounding and issues in the classification of interventions, since the choice between awake and deep extubation was left entirely to the attending anesthesiologist without standardized criteria, leading to possible misclassification. In addition, outcome measurement relied on

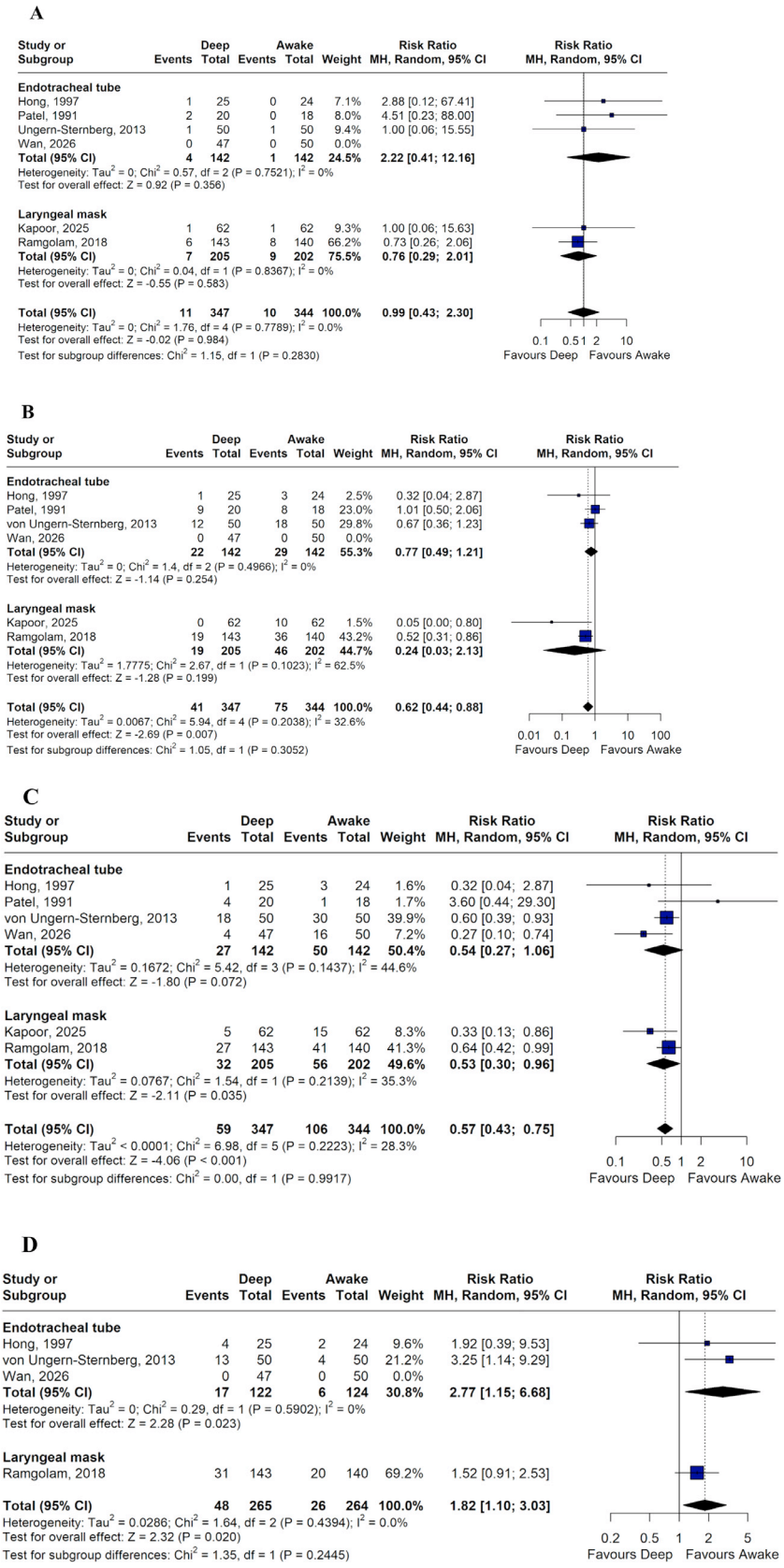



**Figure 2** **Respiratory complications in RCTs.** (A) Incidence of laryngospasm. (B) Incidence of desaturation. (C) Incidence of cough. (D) Incidence of airway obstruction.

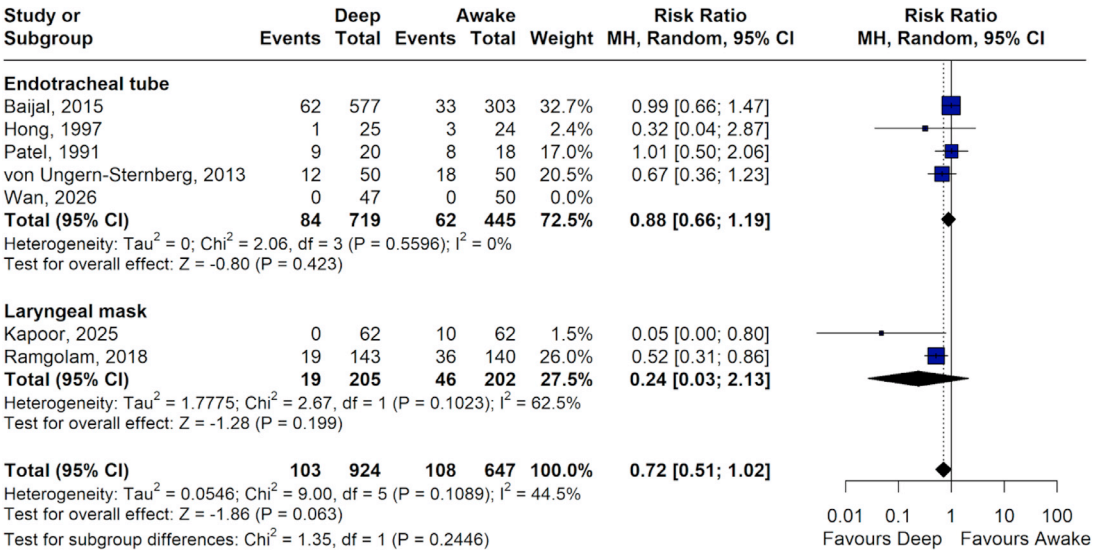



**Figure 3** **Sensitivity analysis including the observational study.** Incidence of desaturation.

voluntary self-reporting by anesthesiologists and PACU nurses, which may have introduced detection bias and underreporting of respiratory complications.

## Certainty of evidence

Using the GRADE approach, the overall certainty of evidence was rated as high for cough. For laryngospasm and airway obstruction, certainty was downgraded to moderate due to serious imprecision: the confidence interval for laryngospasm was wide, crossing the line of no effect with few events across studies; for airway obstruction, only four trials with 529 participants contributed to this outcome, and the pooled estimate was sensitive to the exclusion of individual studies in leave-one-out analyses. For desaturation assessed in the RCT-only primary analysis, certainty was also downgraded to moderate due to serious inconsistency, as heterogeneity was observed ($I^2$ = 32.6%) and the pooled estimate was sensitive to study design and to individual studies in leave-one-out analyses. For desaturation assessed in the non-randomized study, certainty was rated as very low, as non-randomized evidence starts at low certainty by GRADE default and was further downgraded for serious risk of bias related to confounding, classification of the intervention, and outcome measurement (Suppl. Table 3).

## Discussion

We conducted a systematic review and meta-analysis of six RCTs and one observational study, including 1,571 pediatric patients undergoing tonsillectomy or adenotonsillectomy, to compare awake versus deep removal of airway devices on adverse respiratory outcomes. To reduce confounding and intervention-classification bias, the primary meta-analysis was restricted to randomized controlled trials, while the observational study was incorporated as sensitivity analysis when outcome data were available.

Deep airway device removal was associated with a higher incidence of airway obstruction and no difference in laryngospasm. Although desaturation was lower with deep removal in the RCT-only analysis, this finding was attenuated after inclusion of the observational study and should be interpreted cautiously. The overall pooled analysis suggested a significant reduction in cough consistent with earlier trials.[8–10,21,23,26] Although this reduction reached statistical significance only in the LMA subgroup, the point estimates were nearly identical in the two subgroups and the test for subgroup differences did not suggest effect modification. Statistical significance in one subgroup and its absence in another does not, by itself, establish a differential effect; the wide confidence interval in the ETT subgroup most likely reflects imprecision. Although the absence of LMA contact with the vocal cords and trachea offers a plausible mechanism, thereby avoiding direct stimulation of perilaryngeal structures capable of provoking coughing. This is particularly relevant in otolaryngologic procedures, where coughing may lead to complications such as suture dehiscence, bleeding, or airway trauma.

Deep airway device removal is more frequently associated with upper airway obstruction than awake airway device removal.[25] This association was also observed in our population of adenotonsillectomy and tonsillectomy patients. Because pharyngeal muscles recover more slowly from anesthesia than the diaphragm and intercostals, the upper airway is prone to collapse after extubation. When pharyngeal tone cannot counter negative inspiratory pressure, obstruction may occur; therefore, awake extubation is preferred in patients susceptible to residual anesthetic effects.[27] To prevent or manage airway obstruction, supportive measures include placing shoulder rolls, positioning the patient laterally to reduce the impact on oral structures (soft palate, tongue, and tonsils), or turning the head to the side. In neonates and infants, moderate CPAP delivered via mask is particularly effective in maintaining airway patency and preventing collapse.[27]

Among airway complications, laryngospasm and desaturation pose life-threatening risks,[28] accounting for 27% of perioperative cardiac arrests in children, with 6% directly

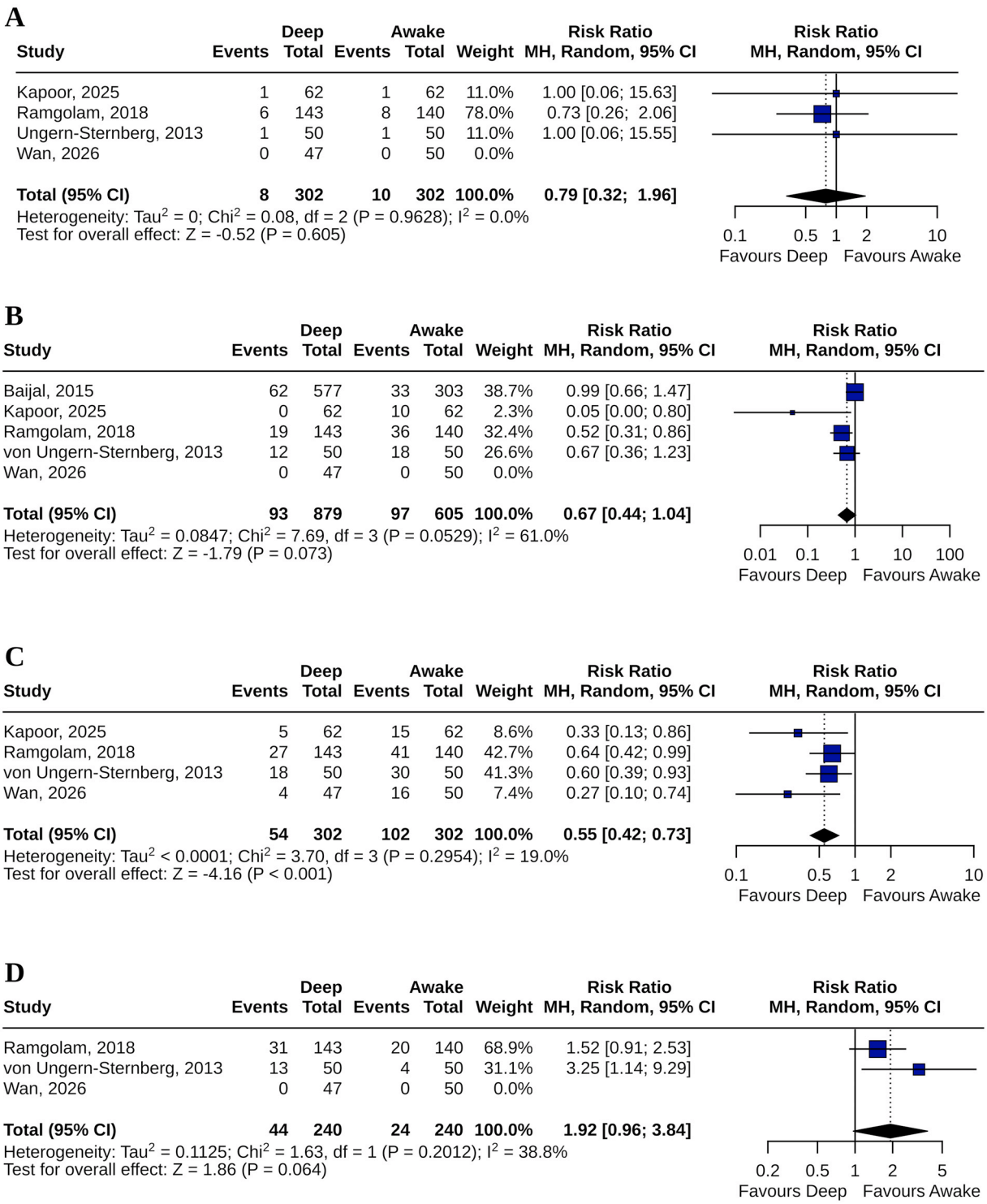



**Figure 4** **Exploratory subgroup analysis with sevoflurane-based maintenance anesthesia.** (A) Incidence of laryngospasm. In the subgroup restricted to studies using sevoflurane for maintenance, deep airway device removal was not associated with a significantly different risk of laryngospasm compared with awake removal (RR = 0.79; 95% CI 0.32−1.96; p = 0.605; $I^2$ = 0%). (B) Incidence of desaturation. In the subgroup restricted to studies using sevoflurane-based maintenance anesthesia, deep airway device removal was not associated with a significantly different risk of oxygen desaturation compared with awake removal (RR = 0.67; 95% CI 0.44−1.04; p = 0.073; $I^2$ = 61.0%). (C) Incidence of cough. In the subgroup restricted to studies using sevoflurane-based maintenance anesthesia, deep airway device removal was associated with a significantly lower risk of cough compared with awake removal (RR = 0.55; 95% CI 0.42−0.73; p < 0.001; $I^2$ = 19%). (D) Incidence of airway obstruction. In the subgroup restricted to studies using sevoflurane-based maintenance anesthesia, deep airway device removal was not associated with a significantly different risk of airway obstruction compared with awake removal (RR = 1.92; 95% CI 0.96−3.84; p = 0.064; $I^2$ = 38.8%). CI, Confidence Interval; df, Degrees of freedom; $I^2$, Inconsistency statistic; MH, Mantel-Haenszel; RR, Risk Ratio; $\tau^2$, Between-study variance.

attributable to laryngospasm.[29] In our analysis, no significant difference was observed among groups regarding the incidence of laryngospasm. Baijal et al.[25] attributed neutral findings to the presence of subspecialty-trained pediatric anesthesiologists and dedicated PACU nurses, whereas Ramgolam et al. suggested that one-to-one nursing ratios enabled immediate intervention and prevented escalation.[11]

Although the RCT-only analysis favored deep airway device removal, this benefit lost significance in the combined analysis with the observational study and with the sevoflurane subgroup due to residual confounding and clinical heterogeneity. Desaturation is a complex endpoint influenced by variable threshold definitions, perioperative oxygen supplementation, rescue maneuvers, staff expertise, anesthetic era, and study design. Critically, none of the included studies were designed to systematically capture post-removal rescue interventions, rendering causal inferences speculative. Given this interplay of confounders, desaturation findings should not be interpreted as definitive evidence favoring either extubation strategy.

Interpretation must account for the anesthetic context of the underlying evidence. Much of the available literature originates from earlier anesthetic eras, including the use of agents no longer employed in contemporary pediatric practice.[30] Desaturation thresholds varied widely, often defined as $SpO_2$ values between 92% and 95% sustained for short durations, levels that may not reflect clinically meaningful hypoxemia in otherwise healthy children. Moreover, oxygen desaturation is not an independent adverse event but rather a downstream physiological consequence of other airway phenomena, such as laryngospasm, upper airway obstruction, coughing, or transient disruption of oxygen delivery during emergence.[31]

Notably, results were consistent in the sevoflurane subgroup. This agent's low blood-gas solubility causes rapid emergence, often associated with increased airway irritability and emergence delirium in children.[32] Specifically, our findings reinforce that while the awake technique shows a strong trend toward reducing airway obstruction ($p = 0.064$), it significantly increases the incidence of coughing, regardless of sevoflurane's known irritant potential during emergence.[10,32] In the overall pooled analysis for airway obstruction, the significant effect was primarily driven by the studies of Hong et al. and von Ungern-Sternberg et al. These studies utilized endotracheal tubes and involved older anesthetic agents (enflurane and sevoflurane/nitrous oxide, respectively), which likely amplified the risk of upper airway collapse during deep extubation compared to the more homogeneous sevoflurane subgroup. These findings across a more homogeneous subgroup strengthen the evidence that the choice between awake and deep extubation should be tailored to the patient's specific risk for obstruction versus irritability.

Clinical heterogeneity was also introduced by variation in anesthetic maintenance. Five studies used sevoflurane, while one used halothane and another enflurane, agents with substantially different pharmacokinetic profiles and airway irritability compared with contemporary practice. As anesthetic maintenance directly influences emergence physiology and upper airway reactivity, pooling across anesthetic eras may introduce heterogeneity beyond what statistical methods can fully capture, reinforcing the exploratory nature of the overall pooled analysis.

LMA removal and ETT extubation were pooled because the intervention under investigation was the depth of anesthesia at the time of device removal, not the type of device itself. Nevertheless, subgroup analyses by device type were performed to explore differential effects, and clinically meaningful differences were observed for cough and airway obstruction, reinforcing the importance of interpreting results at the subgroup level. Although this reduction reached statistical significance only in the LMA subgroup, the point estimates were nearly identical in the two subgroups and the test for subgroup differences did not suggest effect modification ($p = 0.992$).

The pharmacological profile of sevoflurane further underscores the clinical importance of this decision, particularly in high-risk subgroups. Patients with OSA or active respiratory symptoms frequently exhibit airway hyperreactivity; when combined with sevoflurane anesthesia, this state may exacerbate inflammatory responses and consequently elevate the incidence of PRAEs.[33] Children with OSA represent an important high-risk population in whom the balance between airway obstruction and airway irritability deserves particular consideration.

Awake extubation is generally recommended in children with difficult airways, OSA, prematurity, or an increased risk of regurgitation and aspiration.[27] Conversely, deep extubation may be advantageous in patients with asthma, nocturnal cough, or recent upper respiratory infections, where minimizing airway reflexes is desirable.[27] Regardless of the chosen approach, maintaining an adequate depth of anesthesia is the most critical factor for safe airway device removal. Extubation performed in an intermediate state carries the greatest risk, as it predisposes patients to airway obstruction and other serious complications.[27]

A clinically important consideration is that the outcomes assessed differ substantially in severity. Cough, while undesirable in otolaryngologic surgery due to concerns about wound disruption or bleeding, is generally self-limited and rarely threatens patient safety.[33,34] Airway obstruction, laryngospasm and clinically significant desaturation represent true airway emergencies. Therefore, the respiratory outcomes in this meta-analysis cannot be treated as clinically equivalent. While deep extubation reduces the incidence of a minor reflex (cough), its association with an increased risk of a severe complication (airway obstruction) represents a shift in the risk profile from minor physiological disruption to potential patient harm. This hierarchy of severity must be weighed when choosing the extubation timing, as avoiding a minor reflex does not justify increasing the risk of a life-threatening event.

## Limitations

This review has several limitations. First, several included studies were conducted during earlier anesthetic eras, incorporating agents no longer used in contemporary pediatric anesthesia, which limits external validity. Second, the inclusion of one observational study alongside randomized controlled trials may have introduced confounding. Third, the relatively small number of eligible studies precluded a formal assessment of publication bias. Fourth, most studies were conducted in tertiary centers with highly trained pediatric anesthesiologists and recovery staff, which may limit the generalizability of our findings to other settings. Fifth, one trial administered dexmedetomidine only in the deep extubation arm, which may have contributed to the reduction in cough observed in that study.[24]

Taken together, these considerations demonstrate that the observed differences between awake and deep airway device removal in this review should be interpreted as manifestations of well-described physiological trade-offs inherent to anesthetic depth and airway reflex recovery. They do not constitute definitive evidence to guide contemporary clinical practice, particularly in the current sevoflurane and TIVA dominant era.[33] Future studies should prioritize standardized, clinically meaningful outcome definitions and explicitly account for modern anesthetic maintenance strategies to generate practice-relevant evidence.

## Conclusion

In pediatric patients undergoing tonsillectomy and adenotonsillectomy, deep airway device removal was associated with less coughing, with no evidence that this effect differed between LMA and ETT. Deep airway device removal was also associated with a higher incidence of airway obstruction compared with awake removal. No significant differences were observed in laryngospasm. Desaturation findings were influenced by study design and became non-significant after inclusion of observational evidence; therefore, they should not be interpreted as conclusive evidence favoring either strategy. Future clinical trials should be adequately powered, restricted to clearly defined surgical populations, incorporate standardized and clinically meaningful outcome measures and explicitly account for contemporary anesthetic maintenance strategies.

## Data availability statement

All data supporting the findings of this study are available within the article itself and its Supplementary Material. The study-level dataset extracted for this meta-analysis, as well as the statistical code and analysis scripts used, are available from the corresponding author upon reasonable request.

## Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

During the preparation of this work, the author(s) used ChatGPT (OpenAI) in order to review and improve the language and readability of the manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

## Authors' contributions

Henrique B. Pivetta: Conceptualization; methodology; literature search; study selection; statistical analysis; writing; editing; risk of bias assessment. Alice O. Almeida: Study selection; data extraction; writing. Rafael L. Miranda: Statistical analysis; writing. Rafaela P. Silva: Writing; data extraction. Raphaela G. Mendes: Writing; data extraction; review; study selection. João M.M. Borges: Writing; data extraction. Raquel E. Saldanha: Writing; data extraction. Matheus Arraes: Risk of bias assessment; review. Pablo P. Sanín: Supervision; Conceptualization; writing; review. Walid Alrayashi: Supervision; Conceptualization; writing; review.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

## Conflicts of interest

The authors declare no conflicts of interest.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjane.2026.844840.

## Associate Editor

Paulo do Nascimento Junior

## References

1. Baugh RF, Archer SM, Mitchell RB, et al. Clinical Practice Guideline: Tonsillectomy in Children. Otolaryngol Head Neck Surg. 2011;144(1 Suppl):S1—S30.
2. Mitchell RB, Archer SM, Ishman SL, et al. Clinical Practice Guideline: Tonsillectomy in Children (Update). Otolaryngol Neck Surg. 2019 Feb;160(1_suppl):S1—S42.
3. Douglas-Jones P, Fagan JJ. Tonsillectomy rates in the South African private healthcare sector. S Afr Med J. 2016;106:1134.
4. Nguyen BK, Quraishi HA. Tonsillectomy and Adenoidectomy - Pediatric Clinics of North America. Pediatr Clin North Am. 2022;69:247—59.
5. Derkay CS, Darrow DH, LeFebvre SM. Pediatric Tonsillectomy and Adenoidectomy Procedures. AORN J. 1995;62:885—904.
6. McGuire SR, Doyle NM. Update on the safety of anesthesia in young children presenting for adenotonsillectomy. World J Otorhinolaryngol Head Neck Surg. 2021;7:179—85.
7. Lee J, Kim J, Kim S, Kim C, Yoon T, Kim H. Removal of the laryngeal tube in children: anaesthetized compared with awake. Br J Anaesth. 2007;98:802—5.
8. Koo CH, Lee S, Chung S, Ryu JH. Deep vs. Awake Extubation and LMA Removal in Terms of Airway Complications in Pediatric Patients Undergoing Anesthesia: A Systemic Review and Meta-Analysis. J Clin Med. 2018;7:353.
9. Von Ungern-Sternberg BS, Boda K, Chambers NA, et al. Risk assessment for respiratory complications in paediatric anaesthesia: a prospective cohort study. Lancet. 2010;376:773—83.
10. Von Ungern-Sternberg BS, Davies K, Hegarty M, Erb TO, Habre W. The effect of deep vs. awake extubation on respiratory complications in high-risk children undergoing adenotonsillectomy: A randomised controlled trial. Eur J Anaesthesiol. 2013;30:529—36.
11. Ramgolam A, Hall GL, Zhang G, Hegarty M, Von Ungern-Sternberg BS. Deep or awake removal of laryngeal mask airway in children at risk of respiratory adverse events undergoing tonsillectomy – a randomised controlled trial. Br J Anaesth. 2018;120:571—80.

12. Benham-Hermetz J, Mitchell V. Safe tracheal extubation after general anaesthesia. BJA Educ. 2021;21:446—54.
13. Mathew PJ, Mathew JL. Early versus late removal of the laryngeal mask airway (LMA) for general anaesthesia. Cochrane Database Syst Rev. 2015(8):CD007082.
14. Dolling S, Anders NRK, Rolfe SE. A comparison of deep vs. awake removal of the laryngeal mask airway in paediatric dental day-case surgery. A randomised controlled trial. Anaesthesia. 2003;58:1224—8.
15. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ. 2009;339. b2700-b2700.
16. Shamseer L, Moher D, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ. 2015;349. g7647-g7647.
17. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
18. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
19. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.
20. Balshem H, Helfand M, Schünemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol. 2011;64:401—6.
21. Kapoor D, Tweddle EA, Baitch L. The effect of deep versus awake removal of the laryngeal mask airway on the incidence of emergence delirium in paediatric tonsillectomy: A randomised controlled trial. Anaesth Intensive Care. 2025;53:55—62.
22. Patel RI, Hannallah RS, Verghese ST, Casey WF, Norden JM. Emergence airway complications in children: a comparison of tracheal extubation in awake and deeply anesthetized patients. Anesth Analg. 1991;73:266—70.
23. Hong JY, Han SJ, Kil HK, Kim WO. Airway - Related Complications and SpO2 Changes of Deeply Anesthetized Versus Awake Extubation in Children. Korean J Anesthesiol. 1997;32:384.
24. Wan Y, Shi Y, Wang X, Guo J. The effect of deep extubation on the emergence agitation in pediatric patients undergoing obstructive sleep apnea syndrome surgery: a randomized controlled study. Sci Rep. 2026;16:16073.
25. Baijal RG, Bidani SA, Minard CG, Watcha MF. Perioperative respiratory complications following awake and deep extubation in children undergoing adenotonsillectomy. Pediatr Anesth. 2015;25:392—9.
26. De Carvalho ALR, Vital RB, De Lira CCS, et al. Laryngeal Mask Airway Versus Other Airway Devices for Anesthesia in Children with an Upper Respiratory Tract Infection: A Systematic Review and Meta-analysis of Respiratory Complications. Anesth Analg. 2018;127:941—50.
27. Igarashi A. Extubation and removal of supraglottic airway devices in pediatric anesthesia. Anesth Pain Med. 2024;19(Suppl 1):S49—60.
28. Visvanathan T. Crisis management during anaesthesia: laryngospasm. Qual Saf Health Care. 2005;14. e3-e3.
29. Bhananker SM, Ramamoorthy C, Geiduschek JM, et al. Anesthesia-Related Cardiac Arrest in Children: Update from the Pediatric Perioperative Cardiac Arrest Registry. Anesth Analg. 2007;105:344—50.
30. Doi M, Ikeda K. Airway irritation produced by volatile anaesthetics during brief inhalation: comparison of halothane, enflurane, isoflurane and sevoflurane. Can J Anaesth. 1993;40:122—6.
31. Trachsel D, Erb TO, Hammer J, Von Ungern-Sternberg BS. Developmental respiratory physiology. Pediatr Anesth. 2022;32:108—17.
32. Dahmani S, Stany I, Brasher C, et al. Pharmacological prevention of sevoflurane- and desflurane-related emergence agitation in children: a meta-analysis of published studies. Br J Anaesth. 2010;104:216—23.
33. Shen F, Zhang L, Wang X, et al. Effect of Intravenous, Inhalational, or Combined Anesthesia Maintenance on Postoperative Respiratory Adverse Events in Children Undergoing Adenotonsillectomy (AmPRAEC): A Multicenter Randomized Clinical Trial. Anesthesiology. 2025;143:1484—96.
34. Turner RD, Birring SS. Measuring cough: what really matters? J Thorac Dis. 2023;15:2288—99.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



REVIEW ARTICLE

# Intraoperative methadone in adult laparoscopic surgery: a systematic review and meta-analysis of randomized trials

**Julio S. Curi** [a,*], **Matheus Arraes** [b], **Jasmeen Kaur** [c], **Carlos Martins** [d], **Rafael A. Neves** [e], **Eduardo M. Pereira** [f], **Shaun E. Gruenbaum** [g], **Rafael A. Lombardi** [h]

[a] *Universidade do Vale do Taquari, Department of Medicine, Lajeado, RS, Brazil*
[b] *CMA Anesthesia Group, Division of Anesthesiology, São Paulo, SP, Brazil*
[c] *Augusta University Medical Center, Department of Anesthesiology, Georgia, USA*
[d] *Grupo de Anestesia Anestesvel, Department of Anesthesiology, Cascavel, PR, Brazil*
[e] *Centro Universitário Fundação Assis Gurgacz, Department of Medicine, Cascavel, PR, Brazil*
[f] *Universidade Federal de Minas Gerais, Department of Medicine, Belo Horizonte, MG, Brazil*
[g] *Mayo Clinic Florida, Department of Anesthesiology, Florida, USA*
[h] *University of Nebraska Medical Center, Department of Anesthesiology, Nebraska, USA*

Received 20 December 2025; accepted 7 October 2026
Available online 15 October 2026

**KEYWORDS**
Analgesia;
Analgesics, opioid;
Anesthesia recovery period;
Laparoscopy;
Methadone;
Pain, postoperative

## Abstract

*Background:* Opioids are routinely used for pain control in laparoscopic procedures. Methadone, a long-acting opioid, has unclear efficacy in this context. Therefore, we aim to compare intraoperative methadone *vs.* short-acting opioids or placebo for laparoscopic surgeries.

*Methods:* We searched PubMed, Scopus, Embase, and Cochrane for eligible Randomized Controlled Trials (RCTs) up to June 2025. Outcomes were cumulative opioid use at 24 hours (primary endpoint), the incidence of respiratory depression, postoperative nausea/vomiting, and hypoxia; intraoperative propofol use; emergence time; Post-Anesthesia Care Unit (PACU) length of stay and opioid use; pain scores at 30 and 60 minutes and at 24 hours (h). We used a random-effects model, Cochrane risk-of-bias analyses, certainty-of-evidence assessment, trial sequential analyses, and minimal clinically important differences to assess clinical impact.

*Results:* 7 RCTs (591 patients) were included. Compared to short-acting opioids, methadone was significantly associated with a reduction in postoperative hypoxia (RR = 0.24; 95% CI 0.06—0.91; p = 0.03), PACU opioid use (RR = 0.69; 95% CI 0.53—0.90; p < 0.01), pain scores at 30 (MD = -1.21; 95% CI -1.91 to -0.51; p < 0.01) and at 60 minutes (MD = -0.41; 95% CI -0.80 to -0.03; p = 0.03),

PROSPERO registration: PROSPERO 2025 CRD420251075236. Available from https://www.crd.york.ac.uk/PROSPERO/view/CRD420251075236.

* Corresponding author.
*E-mail:* julio.curi@icloud.com (J.S. Curi).

https://doi.org/10.1016/j.bjane.2026.844848
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

and cumulative opioid use at 24h (MD = -1.62; 95% CI -2.75 to -0.50; $p < 0.01$). There were no differences between groups in the remaining outcomes.
*Conclusion:* Intraoperative methadone in laparoscopic procedures was associated with a statistically significant reduction in pain scores, hypoxia, and opioid requirements. However, the reduction in opioid requirements did not reach clinical significance, and most effects were small, heterogeneous, and supported by limited-certainty evidence. Evidence regarding respiratory adverse events remains limited because of the small number of trials and participants.
© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Introduction

Postoperative pain remains a major challenge in laparoscopic procedures. This issue is especially significant in outpatient surgeries, which are increasingly common in modern clinical practice, where quick recovery is crucial for early hospital discharge. Poor pain management in surgical patients is linked to higher morbidity, slower recovery and discharge, reduced quality of life, prolonged opioid use, and increased health care costs.[1-3]

Methadone is a long-acting $\mu$-opioid receptor agonist;[4] nevertheless, it also acts as a non-competitive antagonist of N-Methyl-D-Aspartate (NMDA) receptors and inhibits serotonin and norepinephrine reuptake in the central nervous system.[5] It has a half-life ranging from 13 to 50 hours.[6,7] Because of its long-acting nature, it provides a stable plasma concentration after a single intraoperative dose when compared to fluctuating levels seen with short-acting opioids like fentanyl, morphine, or hydromorphone.[6,8,9]

Some studies have shown positive results when using intraoperative methadone in spine, cardiac, pediatric, and general surgery for managing acute postoperative pain and decreasing opioid requirement.[10-14] Similarly, several Randomized Controlled Trials (RCTs) and observational studies have investigated the use of methadone in laparoscopic surgeries and have shown reduced opioid requirement in bariatric, gynecological, ambulatory, and cholecystectomy surgeries.[15-19] However, these studies are limited by small sample sizes and have yielded inconsistent outcomes. Additionally, few studies have assessed metrics such as quality of recovery or functional outcomes in the laparoscopic setting.

We, therefore, aim to conduct a systematic review, meta-analysis, and Trial Sequential Analysis (TSA) of RCTs to study the effectiveness of methadone *vs*. standard opioid therapy in the management of acute postoperative pain.

## Methods

This study follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[20,21] The protocol was preregistered using the International Prospective Register of Systematic Reviews (PROSPERO, CRD420251075236).

### Eligibility criteria

The inclusion criteria were defined as 1) RCTs; 2) Comparing methadone *vs*. no methadone or placebo intraoperatively; 3) Including patients ≥ 18 years old undergoing laparoscopic surgeries; and 4) Reporting at least one outcome of interest. We excluded 1) RCTs with overlapping populations; 2) Conference abstracts; and 3) Studies using regional anesthesia for all patients. We did not apply restrictions on language or publication year. We contacted the authors of individual RCTs to obtain missing data. We excluded studies using any regional anesthesia (i.e., neuraxial anesthesia or peripheral blocks, excluding port-site infiltration) for all patients.

### Search strategy and data extraction

Two independent authors (R.L. and C.E.) systematically searched MEDLINE, Embase, Scopus, and Cochrane databases for RCTs meeting the inclusion criteria from inception until June 2025. The search was updated through October 1, 2026, and no additional eligible randomized trials were identified. The search strategy is listed in Supplementary Table 1. Additionally, we reviewed references from all included studies to identify additional relevant studies (backward snowballing). Nonetheless, we did not search the grey literature for additional studies. Three independent authors (R.L., C.E., and J.C.) collected the data and resolved disagreements through consensus. As of August 2026, each reference was checked to ensure that no cited study underwent retraction, withdrawal, expression of concern, or correction in the source database or publisher's website.

### Endpoints

The included outcomes of interest were 1) Cumulative postoperative opioid consumption within the first 24 hours after surgery, expressed as Oral Morphine Equivalents (OME); 2) Emergence time; 3) Incidence of hypoxia; 4) Incidence of PACU opioid consumption; 5) Incidence of Postoperative Nausea and Vomiting (PONV), defined as any episode of nausea, vomiting, or both during the postoperative observation period; 6) Incidence of respiratory depression, defined as a respiratory rate < 10 breaths per minute sustained for at least 10 minutes; 7) Intraoperative opioid consumption expressed in OME; 8) Intraoperative propofol consumption; 9) PACU length of stay; 10) PACU (0-min) pain scores; 11) 30 minutes pain scores; 12) 60 minutes pain scores; 13) 24 hours pain scores; and 14) Patient satisfaction. The primary outcome was cumulative 24h postoperative opioid use; all the other outcomes listed above were our secondary outcomes. We assessed pain outcomes using a numeric rating scale (NRS 0 to 10). When studies reported pain using the Visual Analog Scale (VAS 0–100 mm), we converted to NRS by dividing by 10, as NRS and VAS scores are considered sufficiently comparable to permit pooling on the

same scale while preserving clinical interpretability.[22] Opioid doses were converted to OME using the ClinCalc opioid conversion calculator.[23] We pooled results for additional outcomes whenever they were reported in at least three studies. We extracted and analyzed hypoxia according to each trial's original definition. Freire et al.[16] defined hypoxia as sustained desaturation or hypoventilation associated with peripheral oxygen saturation below 90%, whereas Pontes et al.[15] defined it as peripheral oxygen saturation of 90% or lower. Moro et al.[19] reported hypoxemia without providing an operational definition. No standardized oxygen-saturation threshold was applied retrospectively.

### Quality assessment

Risk of bias was assessed with version 2 of the Cochrane Risk of Bias assessment tool (RoB2) by two independent authors (R.L. and J.C.).[24] Furthermore, two authors (R.L. and M.A.) independently applied the Grading Recommendations, Assessment, Development, and Evaluation (GRADE) tool to assess the certainty of evidence for each outcome. Similarly, we resolved any disagreements through discussion with the remaining authors. We also performed a leave-one-out sensitivity analysis for each outcome.

### Statistical analysis

We pooled dichotomous outcomes using the Mantel-Haenszel method and reported Relative Risks (RRs) with 95% Confidence Intervals (95% CIs). We pooled continuous outcomes using the inverse-variance method and reported Mean Differences (MDs) with 95% CIs. We set statistical significance at $p < 0.05$. Given the anticipated clinical and methodological heterogeneity across RCTs, we conducted all analyses using a random-effects model with restricted maximum likelihood estimation. We assessed heterogeneity using the Cochran *Q*-test and the $I^2$ statistic, classifying it as mild ($I^2 <$ 25%), moderate ($I^2$ 25%—49%), or substantial ($I^2 \geq 50\%$).[25]

In accordance with previous evidence, the Minimal Clinically Important Difference (MCID) was set at 10 mm for pain scores and at 40 mg of OME in 24 hours for opioid consumption.[26,27] All analyses were performed in the "meta" package of R 4.1.2 (R Foundation for Statistical Computing, Austria). Lastly, we performed TSA for every included outcome (version 0.9.5.10, Copenhagen Trial Unit, Denmark), with a type I error of 0.05 and a type II error of 0.20 (80% power), and applied the O'Brien-Fleming alpha-spending function for the *Z*-score.[28]

## Results

### Study selection and characteristics

The search identified 345 records. After removing 171 duplicates, 174 records underwent title and abstract screening; 150 were excluded. Full-text assessment was performed for 24 studies (Supplementary Table 3), and 7 were included in the meta-analysis (Fig. 1), comprising 591 adults, of whom 295 (49.9%) were randomized to receive methadone. The baseline characteristics of these RCTs are shown in Table 1. Given the variability in anesthetic regimens, including methadone dosing and timing, as well as endpoint definitions, we provided trial-level protocol details for each study (Supplementary Table 2). The reasons for full-text exclusions are itemized in Figure 1. None of the included trials and cited references underwent retraction, withdrawal, expression of concern, or correction in the source database or on the publisher's website, as of October, 2026.

### Pooled analysis

Cumulative opioid consumption within 24 hours, measured in OME, was significantly lower in the methadone group (MD = -1.62; 95% CI -2.75 to -0.50; p = 0.004; $I^2$ = 55%; three trials; 239 patients, Fig. 2). The incidence of PACU opioid use was also reduced in the methadone group (RR = 0.69; 95% CI 0.53—0.90; p = 0.006; $I^2$ = 52.6%; four trials; 365 patients, Fig. 3A). The risk of postoperative hypoxia decreased in patients receiving methadone (RR = 0.24; 95% CI 0.06—0.91; p = 0.036; $I^2$ = 0%; three trials; 233 patients, Fig. 3B). Pain scores were notably lower in the methadone group at 30 minutes (MD = -1.21; 95% CI -1.91 to -0.51; p = 0.0007; $I^2$ = 47.7%; three trials; 274 patients, Fig. 3C) and at 60 minutes (MD = -0.41; 95% CI -0.80 to -0.03; p = 0.0361; $I^2$ = 0%; four trials; 336 patients, Fig. 3D).

No significant differences were observed between groups in intraoperative opioid consumption (MD = -8.32; 95% CI -26.84 to 10.21; p = 0.37; $I^2$ = 72.1%; five trials; 411 patients, Fig. 3E), intraoperative propofol consumption (MD = -2.11; 95% CI -26.12 to 21.9; p = 0.863; $I^2$ = 0%; four trials; 376 patients, Fig. 3F), emergence time (MD —0.30; 95% CI -3.13 to 2.53; p = 0.835; $I^2$ = 91.6%; three trials; 341 patients, Fig. 4A), PACU length of stay (MD = -11.22; 95% CI -25.69 to 3.25; p = 0.128; $I^2$ = 89.1%; three trials; 234 patients, Fig. 4B), and pain at PACU arrival (MD = -2.47; 95% CI -5.49 to 0.55; p = 0.109; $I^2$ = 97.2%; four trials; 243 patients, Fig. 4C). Pain scores at 24 hours (MD = -0.44; 95% CI -1.68 to 0.8; p = 0.486; $I^2$ = 86.6%; four trials; 301 patients, Fig. 4D) and patient satisfaction (MD = -0.00; 95% CI -0.24 to 0.24; p = 0.987; $I^2$ = 0.6%; three trials; 326 patients, Fig. 4E) were also not significantly different. Similarly, the incidence of postoperative nausea and vomiting (RR = 0.67; 95% CI 0.41—1.1; p = 0.114; $I^2$ = 0%; three trials; 238 patients, Fig. 5A) and respiratory depression (RR = 1.54; 95% CI 0.83—2.85; p = 0.173; $I^2$ not applicable; two trials, 69 patients, Fig. 5B) showed no significant differences between groups.

### Sensitivity analyses

Leave-one-out sensitivity analyses revealed that the statistical significance for 24-hour opioid consumption was lost when Arriaza et al.[1] or Freire et al.[16] were excluded, while removing Friesgaard et al.[18] reduced heterogeneity to 0% and maintained statistical significance. For the emergence time, excluding Simoni et al.[2] resulted in a statistically significant favoring of methadone. Additionally, removing Pontes et al.[15] rendered hypoxia non-significant. For PACU opioid use, excluding Arriaza et al.,[1] caused a loss of significance. Intraoperative opioid and propofol consumption remained unaffected by omitting any individual study. Excluding Freire et al.[16] made pain at PACU arrival statistically significant (MD = -0.89; 95% CI -1.67 to -0.11). Pain at 30 and 60 minutes and at 24 hours, PACU length of stay, PONV, and patient satisfaction remained consistently significant and in the same direction across all sensitivity

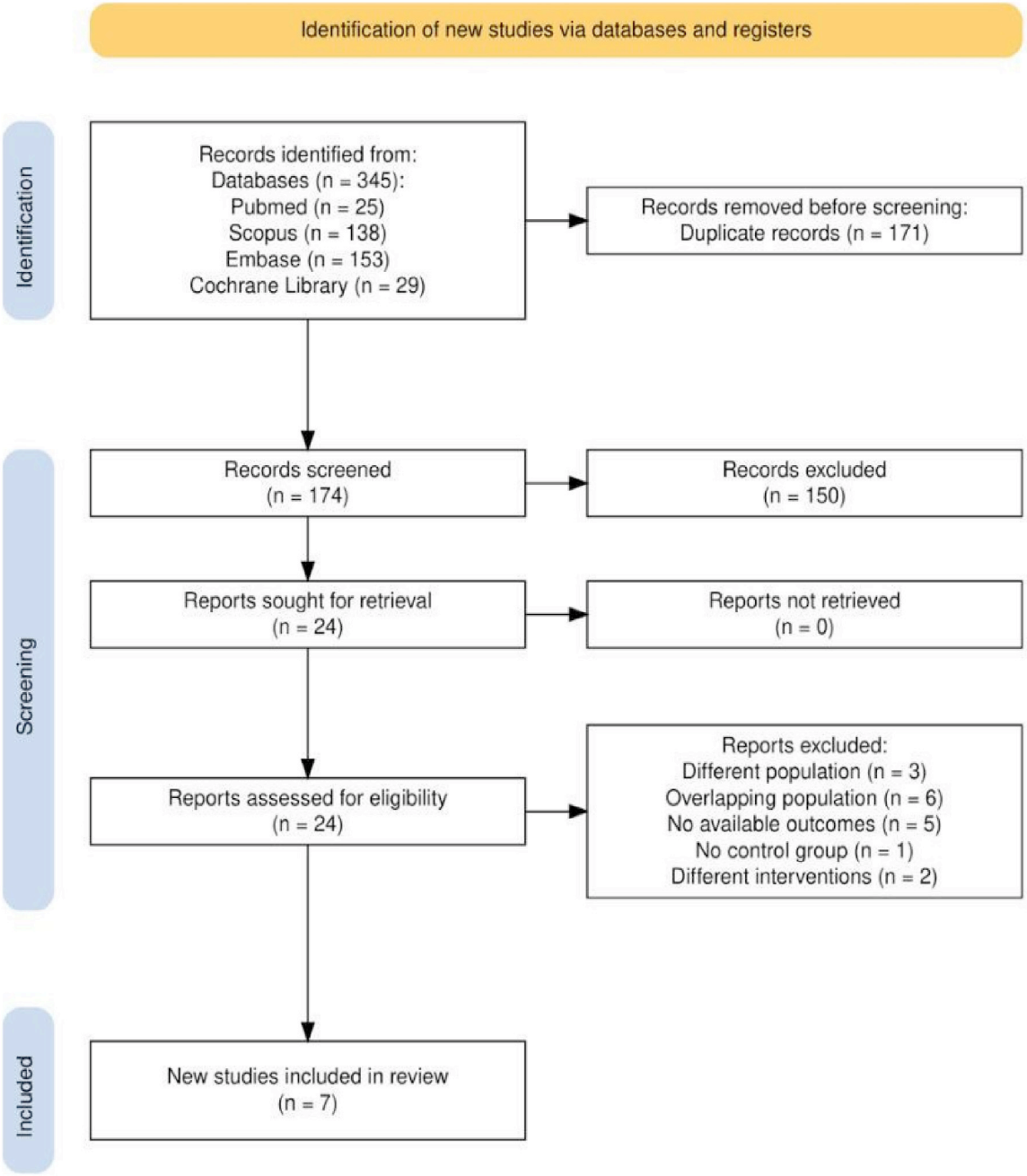



**Figure 1** Preferred reporting items for systematic reviews and meta-analyses flow diagram of study selection.

analyses. Detailed leave-one-out results are available in the Supplementary Figure 1.

## Trial sequential analysis

As shown in Figure S2A, the Trial Sequential Analysis (TSA) indicated that the cumulative *Z*-curve for 24-hour opioid consumption crossed the conventional significance boundary favoring methadone, but did not reach the monitoring boundary or the required information size. Similarly, for hypoxia, as shown in Figure S2B, the *Z*-curve crossed the conventional boundary in favor of methadone but did not meet either the monitoring boundary or the required information size. Regarding pain outcomes, the *Z*-curve at 30 minutes crossed the required information size (Fig. S2C). In contrast, at 60 minutes, it crossed only the conventional boundary favoring methadone, without reaching the monitoring boundary or the required information size (Fig. S2D). Conversely, for the incidence of patients requiring postoperative opioids, the *Z*-curve crossed the conventional boundary favoring methadone and reached the required information size (Fig. S2E).

## Risk of bias

Five of our included studies were deemed to have a low risk of bias.[1,15,16,18,19] Two studies had some concerns: Green et al.[17]

**Table 1** Qualitative characteristics of the included studies.

| Author, year | Country | Type of Surgery | Patients, n Intervention Control | Male, n (%) Intervention Control | Age, years[a] Intervention Control | BMI[a] Intervention Control ($kg.m^{-2}$) | Duration of surgery, min[a] Intervention Control | Intervention Control |
|---|---|---|---|---|---|---|---|---|
| Arriaza, 2021[1] | Chile | Laparoscopic Cholecystectomy | 39 | 15 (35.7%) | 49.4 ± 14.2 | 28 ± 3.4 | 60.2 ± 19.4 | Methadone 0.08 $mg.kg^{-1}$ |
| | | | 39 | 14 (32.6%) | 54.4 ± 12.8 | 27.7 ± 3.4 | 59.4 ± 19.9 | Morphine 0.08 $mg.kg^{-1}$ |
| Simoni, 2009[2] | Brazil | Laparoscopic Cholecystectomy and Hiatal Hernia Repair | 42 | N/R | 45 ± 12 | 79 ± 16 | 65 ± 20 | Methadone 0.1 $mg.kg^{-1}$ |
| | | | 42 | | 40 ± 10 | 76 ± 17 | 57 ± 15 | Placebo (Normal Saline) |
| Pontes, 2021[15] | Brazil | Laparoscopic Bariatric | 69 | 15 (22%) | 36.2 ± 10.5 | 40.0 ± 4.2 | 53 (48–60) | Methadone 0.1 $mg.kg^{-1}$ |
| | | | 68 | 19 (28%) | 34.4 ± 8 | 39.6 ± 3.7 | 53 (47–61)[b] | Morphine 0.1 $mg.kg^{-1}$ |
| Freire, 2025[16] | Brazil | Laparoscopic Bariatric | 16 | 1 (6.2%) | 59.01 ± 12.50 | 42.1 ± 7.1 | 264 ± 78 | Methadone 10 mg diluted in 100 mL saline |
| | | | 18 | 5 (17.8%) | 57.78 ± 9.43 | 44.8 ± 5.3 | 258 ± 66 | Placebo 100 mL Normal saline |
| Green, 2024[17] | Australia | Gynecological Laparoscopy | 35 | 0 | 33.17 ± 10.49 | 25.60 ± 4.20 | N/R | Methadone 10 mg at induction (or 0.15 mg. $kg^{-1}$ if < 10 lt; 10 mg) |
| | | | 34 | 0 | 32.74 ± 9.51 | 25.81 ± 4.09 | | Fentanyl, morphine, or oxycodone (mean intraoperative dose ≈ 47 mg OME). |
| Friesgaard, 2023[18] | Denmark | Laparoscopic Hysterectomy | 64 | 0 | 47.4 (45.7–49.1) | 26.9 (25.7–28.0) | 108 (90–130) | Methadone 0.2 $mg.kg^{-1}$ |
| | | | 63 | 0 | 46.6 (44.6–48.6) | 27.1 (26.3–28.7)[a] | 107 (91–120)[b] | Morphine 0.2 $mg.kg^{-1}$ |
| Moro, 2019[19] | Brazil | Laparoscopic Cholecystectomy | 31 | 3 (9.7%) | 48.0 ± 11.3 | 28 (26–33) | 60 (60–75) | Methadone 0.1 $mg.kg^{-1}$ |
| | | | 31 | 7 (22.6%) | 42.3 ± 10.2 | 28 (24–32)[b] | 60 (50–70)[b] | Morphine 0.1 $mg.kg^{-1}$ |

N, Number; BMI, Body Mass Index; IQR: interquartile range; kg, kilograms; lt: liters; m, meters; min, minutes; mg, milligrams; mL, milliliters; N/R, Not Reported; OME, Oral Morphine Equivalents; SD: standard deviation.

[a] Mean ± SD.

[b] Median (IQR).

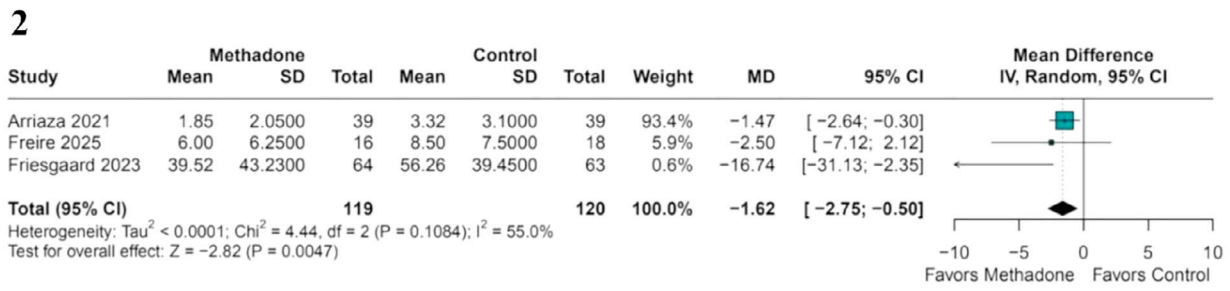



**Figure 2** Forest plot for cumulative opioid use in the first 24 hours. MD, Mean Difference; CI, Confidence Interval.

due to deviations from the intended intervention, and Simoni et al.[2] because of potential bias in the reported results (Fig. S3).

We did not evaluate publication bias because none of the pooled outcomes included at least ten studies. Funnel plots and statistical tests for publication bias are unreliable when few studies are available, and the Cochrane Handbook recommends using these methods only when at least 10 studies are included to ensure adequate power.[29,30]

## Certainty of evidence

The risk-of-bias assessment indicates that the quality of the included RCTs was reasonable. However, the GRADE assessment also demonstrated that the overall certainty of the evidence in this meta-analysis was moderate to low, mainly due to heterogeneity and imprecision (Supplementary Table 4).

## Discussion

In this systematic review, meta-analysis, and TSA including 591 patients from 7 RCTs, we found that methadone was associated with decreased 1) Cumulative opioid consumption in the first 24 hours, 2) Incidence of hypoxia, 3) Incidence of opioid use in PACU, and 4) Pain scores at PACU at

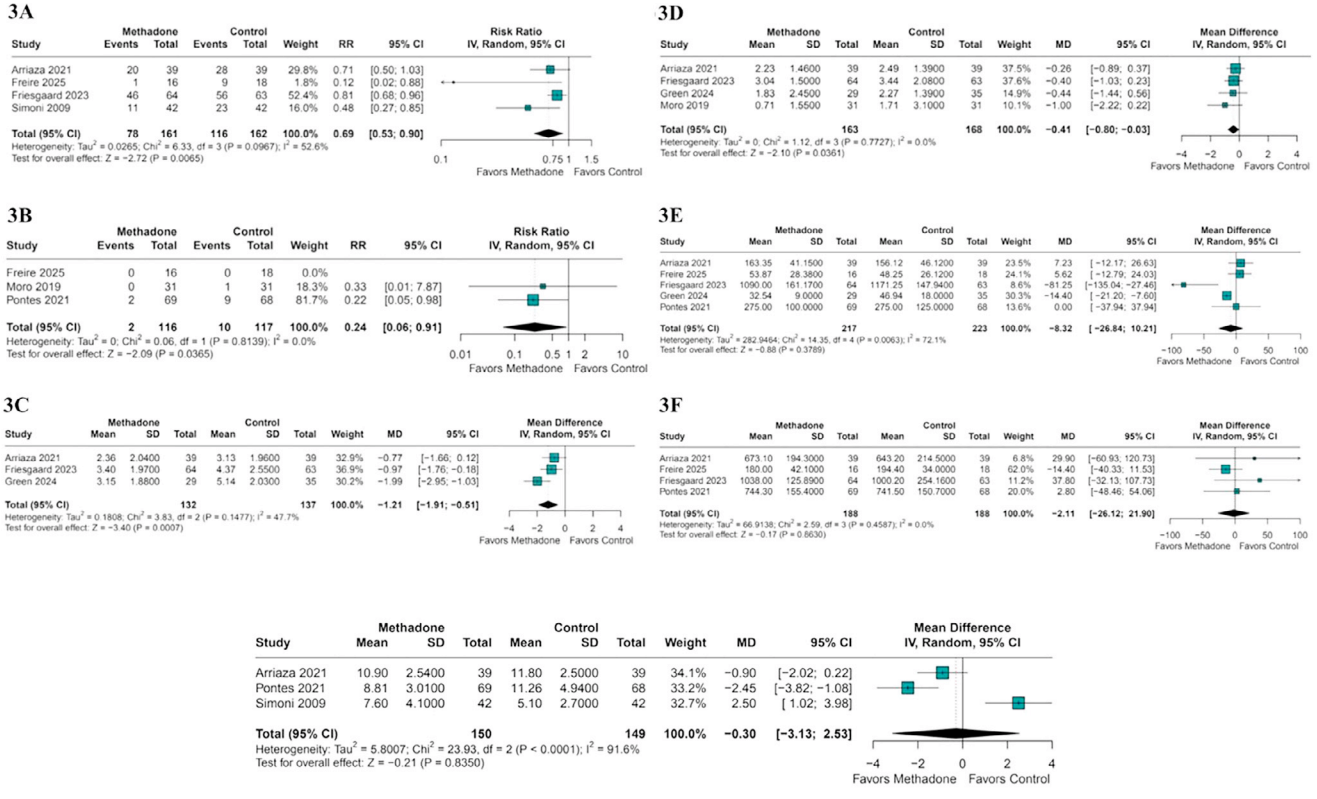



**Figure 3** (A) Forest plot for the incidence of Post-Anesthesia Care Unit (PACU) opioid use. (B) Forest plot for the risk of postoperative hypoxia. (C) Forest plot for pain scores at 30 minutes. (D) Forest plot for pain scores at 60 minutes. (E) Forest plot for intraoperative opioid consumption. (F) Forest plot for intraoperative propofol consumption. RR, Risk Ratio; MD, Mean Difference; CI, Confidence Interval; SD: Standard Deviation.

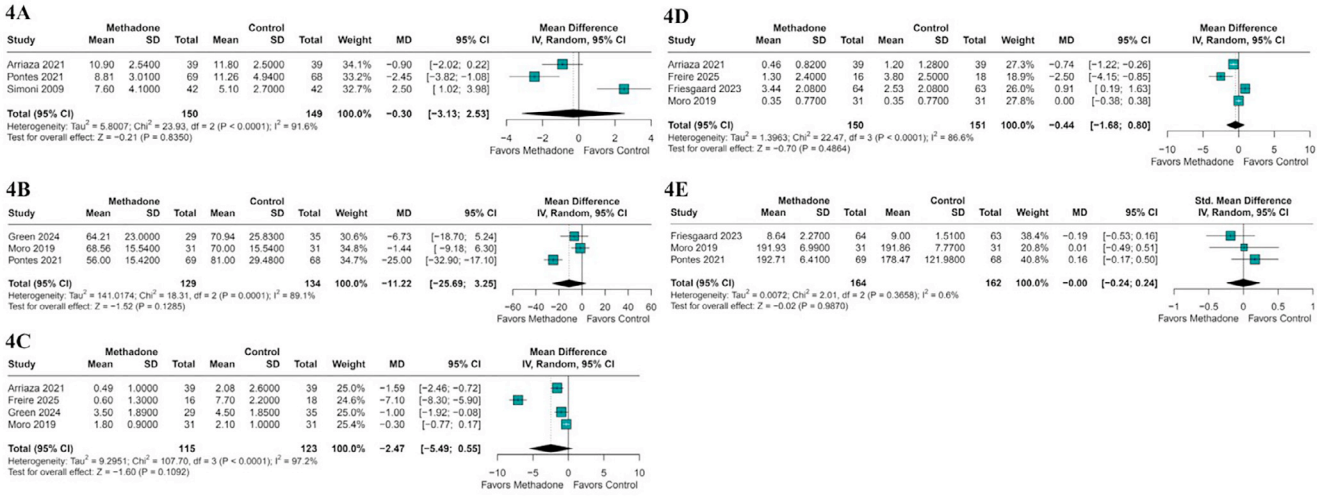



**Figure 4** (A) Forest plot for emergence time, (B) Forest plot for PACU stay, (C) Forest plot for pain at PACU arrival, (D) Forest plot for pain scores at 24 hours, (E) Forest plot for patient satisfaction. MD, Mean Difference; CI, Confidence Interval; SD: Standard Deviation.

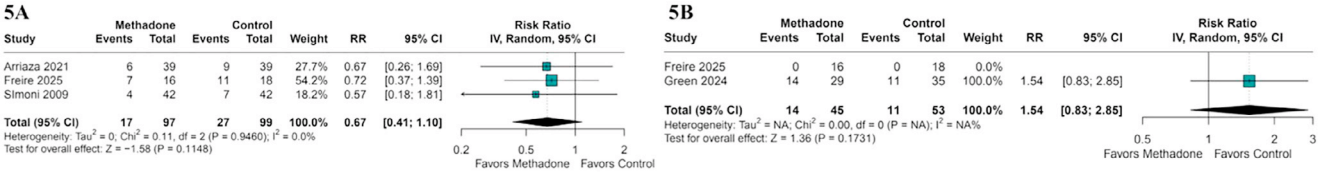



**Figure 5** (A) Forest plot for the incidence of postoperative nausea and vomiting. (B) Forest plot for the incidence of respiratory depression. RR, Risk Ratio; CI, Confidence Interval.

30-min and 60-min, although some of these results did not meet clinical benefit and showed low certainty of evidence. We observed no differences between groups regarding emergence time, incidence of PONV, incidence of respiratory depression, intraoperative opioid use, intraoperative propofol use, PACU length of stay, patient satisfaction, and pain scores on PACU arrival and at 24h.

Our analysis showed that methadone significantly decreased the total opioid use 24 hours after laparoscopic surgery. However, the reduction was much smaller than the MCID suggested by Versyck (30 mg OME) and Laigaard (10 mg IV morphine equivalents). While MCID thresholds vary by context and no universal standard exists, the small effect observed here likely does not provide a meaningful clinical benefit.[27,31] The TSA confirmed that the result was statistically significant but did not cross the monitoring boundaries or achieve the required information size, indicating a potential risk of a type I error.[32]

Sensitivity analysis for cumulative opioid use in 24 hours revealed that removing Friesgaard's[18] study reduced heterogeneity to 0% while maintaining statistical significance. This reduction likely reflects methodological and clinical differences between this study and the other two trials. Friesgaard et al.[18] focused on laparoscopic hysterectomy, used an equal-potency comparator (morphine vs. methadone), and administered the study drug 60 minutes before extubation, unlike the other RCTs that administered it at induction. Additionally, opioid use was recorded as total 24-hour OME from both PACU and ward, following a standardized rescue protocol. In contrast, Arriaza et al.[1] only measured PACU morphine use and reported it as 24-hour opioid consumption, since institutional policy restricted opioid use outside PACU, and ward regimens, including nalbuphine, influenced Freire et al.'s[16] 24-hour totals. Removing Friesgaard et al.[18] aligns the trials regarding design and outcome measurement, thus removing these sources of heterogeneity. This result aligns with Machado et al., who reported a mean difference of -8.42 mg IV morphine equivalents when comparing methadone with other opioids across surgical populations.[8] Although this effect was also below MCID thresholds, it was larger than ours, likely reflecting the inclusion of more painful procedures such as spine surgery, which have higher postoperative opioid requirements.[33] In contrast, laparoscopic procedures like those in our review are less invasive and generally involve lower baseline pain and opioid use, limiting the measurable opioid-sparing potential of methadone.[34,35] Nonetheless, we acknowledge that some key effects lose statistical significance when individual studies are removed, which might indicate fragility of the results, especially given that our results were below the MCID and that the number of studies was insufficient to conduct further subgroup analyses by procedure type.

Our analysis also showed that intraoperative methadone significantly reduced the risk of postoperative hypoxia, with no heterogeneity across studies (76% relative risk reduction). The TSA confirmed that this finding reached conventional statistical significance but did not cross the monitoring boundaries or meet the required information size, also indicating a potential risk of type I error and thus reducing the certainty of the evidence, likely due to the low pooled number of events across included studies. A plausible explanation for the possible reduction in hypoxia might be the lower cumulative opioid consumption observed in the methadone group, since opioids are known to

cause respiratory complications. This interpretation is supported by RCTs where morphine recipients exhibited higher postoperative sedation and required more supplemental opioids compared to those given methadone.[15,19] Although a cohort study by Carlé et al. found no difference in respiratory depression between methadone and morphine, our pooled findings contradict that, despite a limited number of included studies.[36] Given that methadone and morphine have similar potency according to equianalgesic estimates, the most likely explanation is that reduced opioid consumption in the methadone group might indirectly translate into fewer hypoxic episodes.[37,38]

Furthermore, while the incidence of respiratory depression did not differ between groups (despite including only 2 RCTs), the higher frequency of hypoxia in the morphine group may indicate that episodes of respiratory depression were more severe in these patients, even if they were not more common. Overall, these findings suggest that intraoperative methadone may provide a respiratory safety benefit through opioid-sparing effects. However, the certainty of this evidence remains limited due to the TSA results, the low number of included trials and pooled events, and the broad definitions of hypoxia and respiratory depression across the included trials (Supplementary Table 2). We understand that methadone dosing and timing varied widely (i.e., four studies at induction and two studies before extubation; fixed *vs.* weight-based dosing; Supplementary Table 2), which may have influenced early analgesic or respiratory outcomes.

We observed no statistically significant differences in pain scores at PACU arrival. In contrast, pain was significantly reduced in the methadone group at 30 and 60 minutes, probably due to the time to methadone's peak analgesic effect (nearly 30 minutes after intravenous infusion).[6] The MCID for postoperative pain is 10-mm on a 100-mm VAS (equivalent to 1-point on a 0—10 NRS), and the Patient Acceptable Symptom State (PASS) is 33 mm.[26] Across all included trials, mean pain scores at PACU arrival, 30-min, 60-min, and 24 hours were consistently well below the PASS threshold, typically ranging between 0.5 and 3 on the NRS. Therefore, while the reduction at 30 minutes exceeded the MCID and was both statistically and clinically significant, the difference at 60 minutes, although statistically significant, was below the MCID threshold and unlikely to be clinically relevant. This finding highlights the importance of differentiating statistically significant from clinically significant results using MCID, a growing trend in anesthesiology research that aids the interpretability and applicability of findings.[26]

The TSA confirmed that the 30-minute effect was statistically robust and exceeded the required information size, reducing the risk of a type I error. In contrast, the TSA for 60 minutes confirmed statistical significance but did not cross monitoring boundaries or reach the information size, leaving uncertainty about the effect's reliability. The absence of significance at PACU arrival may be due to residual effects of intraoperative opioids or early multimodal analgesia obscuring differences between groups.[39-41] Similarly, the lack of significance at 24 hours probably results from the short, minimally painful nature of laparoscopic procedures, routine use of non-opioid multimodal regimens, and the possibility that methadone's long half-life offered no additional benefit once background pain scores had already dropped below clinically significant levels in both groups.[35]

Finally, although previous reviews suggest that the inclusion of methadone within Enhanced Recovery After Surgery (ERAS) protocols appears to reduce opioid consumption and lead to a significantly increased analgesic effect,[42] some evidence suggests that the use of multimodal anesthesia may also dilute or increase methadone's apparent effect.[43] A previous meta-analysis of 929 patients assessing intraoperative methadone found that its analgesic benefits may extend up to 72 hours postoperatively;[8] however, this study included an even higher variability of procedures and patients (i.e., pediatric and cardiac surgeries, ambulatory procedures), and the authors did not assess the MCID for most endpoints. In our study, we limited procedures to laparoscopic surgeries to reduce expected heterogeneity in baseline and postoperative pain burdens; however, variability in anesthetic regimens, methadone timing, and institutional PACU protocols may have contributed to variability in the results. In addition, although most studies included gynecological and biliary surgeries, we could not perform subgroup analyses because there were insufficient studies to pool by procedure. Nonetheless, despite this dispersion, the MCID findings suggest that methadone's role in pain management seems to be limited to the early postoperative periods, and further studies are warranted to better address its impacts on adverse events, such as QTc prolongation or, most importantly, respiratory outcomes (i.e., delayed respiratory depression), given the limited number of reporting trials.

Nonetheless, this study has some limitations. First, we included a small sample size of 591 patients across seven RCTs, and no single outcome was reported in all trials; the small number of studies per outcome also prevented subgroup analyses and meta-regression to examine the effect of methadone dosing. Furthermore, the GRADE assessment was mostly low to moderate, with several outcomes showing limited reporting of adverse events and heterogeneity due to differences in surgical procedures, populations, and methodologies; we further increased variability by including different laparoscopic surgeries, which might restrict generalizability. In addition, although we employed the "backward snowballing" technique to identify additional studies, we did not search the gray literature, which might have contributed to publication bias. Lastly, TSA indicated that most outcomes did not reach the required information size, suggesting a risk of type I error and highlighting the need for larger studies with more participants.

## Conclusion

Intraoperative methadone in laparoscopic surgery was associated with reduced early postoperative pain scores, reduced opioid use, and less hypoxia without evidence of increased adverse events. However, most effects were small, the decrease in opioid consumption did not reach the minimal clinically important difference, and the certainty of evidence was limited by heterogeneity, imprecision, and TSA findings. Future well-powered trials with standardized methadone dosing and outcome definitions are needed to determine if these early benefits translate into consistent clinical improvements.

## Data availability statement

The extracted study-level dataset, opioid conversion table, GRADE evidence profile, Trial Sequential Analysis settings,

and statistical analysis code used to reproduce the analyses are provided in the Supplementary Material.

## Declaration of artificial intelligence (AI) use

The authors did not use AI or AI-assisted technologies in writing this manuscript.

## Authors' contributions

All authors contributed to the study's conception and design. Rafael A. Lombardi, Julio S. Curi, Carlos Martins, and Matheus Arraes performed material preparation, data collection, and analysis. Rafael A. Lombardi and Julio S. Curi wrote the first draft of the manuscript. All authors critically reviewed previous versions of the manuscript. All authors read and approved the final manuscript. All authors agree to be accountable for all aspects of the work.

## Funding

The authors did not receive support from any organization for the submitted work.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgments

The authors thank Rhanderson Cardoso, MD, for supporting this research.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjane.2026.844848.

## Associate Editor

Maria José Carmona

## References

1. Arriaza N, Papuzinski C, Kirmayr M, et al. Efficacy of methadone for the management of postoperative pain in laparoscopic cholecystectomy: a randomized clinical trial. Medwave. 2021;21:e8135.
2. Simoni RF, Cangiani LM, Pereira AMSA, Abreu MP, Cangiani LH, Zemi G. Efficacy of intraoperative methadone and clonidine in pain control in the immediate postoperative period after the use of remifentanil. Rev Bras Anestesiol. 2009;59:421–30.
3. Gan TJ. Poorly controlled postoperative pain: prevalence, consequences, and prevention. J Pain Res. 2017;10:2287–98.
4. Murphy GS, Szokol JW. Intraoperative Methadone in Surgical Patients: A Review of Clinical Investigations. Anesthesiology. 2019;131:678–92.
5. Mercadante S, Bruera E. Methadone as a first-line opioid in cancer pain management: a systematic review. J Pain Symptom Manage. 2018;55:998–1003.
6. Kharasch ED. Intraoperative methadone: rediscovery, reappraisal, and reinvigoration? Anesth Analg. 2011;112:13–6.
7. Chary S, Abdul-Razzak A, Galloway L. Ultralow-dose adjunctive methadone with slow titration, considering long half-life, for outpatients with cancer-related pain. Palliat Med Rep. 2020;1:119–23.
8. Machado FC, Vieira JE, de Orange FA, Ashmawi HA. Intraoperative methadone reduces pain and opioid consumption in acute postoperative pain: a systematic review and meta-analysis. Anesth Analg. 2019;129:1723–32.
9. Blasiole B, Lavage DR, Lin HHS, et al. Intraoperative methadone versus non-methadone analgesia in pediatric cardiac surgery: A Retrospective Cohort Study. Child Basel Switz. 2025;12:567.
10. Rajkovic C, Vazquez S, Thomas Z, et al. Intraoperative methadone in spine surgery eras protocols: a systematic review of the literature. Clin Spine Surg. 2025;38:280–93.
11. Nunn KP, Velazquez AA, Bebawy JF, et al. Perioperative methadone for spine surgery: a scoping review. J Neurosurg Anesthesiol. 2025;37:31–9.
12. Lobova VA, Roll JM, Roll MLC. Intraoperative methadone use in cardiac surgery: a systematic review. Pain Med Malden Mass. 2021;22:2827–34.
13. Wang DJ, Song P, Nault KM. Impact of intraoperative methadone use on postoperative opioid requirements after cardiac surgery. Am J Health-Syst Pharm. 2022;79:636–42.
14. Iguidbashian JP, Chang PH, Iguidbashian J, Lines J, Maxwell BG. Enhanced recovery and early extubation after pediatric cardiac surgery using single-dose intravenous methadone. Ann Card Anaesth. 2020;23:70–4.
15. Pontes JPJ, Braz FR, Módolo NSP, et al. Intra-operative methadone effect on quality of recovery compared with morphine following laparoscopic gastroplasty: a randomized controlled trial. Anaesthesia. 2021;76:199–208.
16. Freire JM, Dos Santos SAG, Confessor de Sousa RK, et al. Intraoperative use of methadone for postoperative pain control in bariatric surgery: a randomized, double-blind, controlled clinical trial. Cureus. 2025;17:e77127.
17. Green KW, Popovic G, Baitch L. Intraoperative methadone for day-case gynaecological laparoscopy: a double-blind, randomised controlled trial. Anaesth Intensive Care. 2024;52:168–79.
18. Friesgaard KD, Brix LD, Kristensen CB, Rian O, Nikolajsen L. Clinical effectiveness and safety of intraoperative methadone in patients undergoing laparoscopic hysterectomy: a randomised, blinded clinical trial. BJA Open. 2023;7.
19. Moro ET, Lambert MF, Pereira AL, et al. The effect of methadone on postoperative quality of recovery in patients undergoing laparoscopic cholecystectomy: A prospective, randomized, double blinded, controlled clinical trial. J Clin Anesth. 2019;53:64–9.
20. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane; 2023. Available from www.training.cochrane.org/handbook.
21. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
22. Rosas S, Paço M, Lemos C, Pinho T. Comparison between the visual analog scale and the numerical rating scale in the perception of esthetics and pain. Int Orthod. 2017;15:543–60.
23. Kane SP. ClinCalc [Internet]. [cited 2025 May 26]. Opioid (Opiate) Equianalgesia Conversion Calculator. Available from: https://clincalc.com/Opioids/.

24. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
25. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21:1539—58.
26. Myles PS, Myles DB, Galagher W, et al. Measuring acute postoperative pain using the visual analog scale: the minimal clinically important difference and patient acceptable symptom state. Br J Anaesth. 2017;118:424—9.
27. Laigaard J, Pedersen C, Rønsbo TN, Mathiesen O, Karlsen APH. Minimal clinically important differences in randomised clinical trials on pain management after total hip and knee arthroplasty: a systematic review. Br J Anaesth. 2021;126:1029—37.
28. TSA software (Copenhagen Trial Unit, Centre for Clinical Intervention Research, Copenhagen) [Internet]. Available from: https://ctu.dk/.
29. Page MJ, Higgins JPT, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, eds. Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022), Cochrane; 2022. Available from www.training.cochrane.org/handbook.
30. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629—34.
31. Versyck B, Pawa A, Chin KJ. Minimal clinically important difference: a context-specific metric. Reg Anesth Pain Med. 2021;46:933—4.
32. De Cassai A, Tassone M, Geraldini F, et al. Explanation of trial sequential analysis: using a post-hoc analysis of meta-analyses published in Korean Journal of Anesthesiology. Korean J Anesthesiol. 2021;74:383—93.
33. Sanford Z, Broda A, Taylor H, Turcotte J, Patton CM. Predictive risk factors associated with increased opioid use among patients undergoing electivewithne surgery. Int J Spine Surg. 2020;14:189—94.
34. Ammari Smail, Nait Slimane N, Benhocine Y, Taieb M. The impact of laparoscopic surgery on postoperative pain: A Prospective Study. World J Adv Res Rev. 2024;22:1727—32.
35. Basunbul LI, Alhazmi LSS, Almughamisi SA, Aljuaid NM, Rizk H, Moshref R. Recent technical developments in the field of laparoscopic surgery: a literature review. Cureus. 2022;14:e22246.
36. Carlé N, Nikolajsen L, Uhrbrand CG. Respiratory depression following intraoperative methadone: A Retrospective Cohort Study. Anesth Analg. 2025;140:516—23.
37. Beaver WT, Wallenstein SL, Houde RW, Rogers A. A clinical comparison of the analgesic effects of methadone and morphine administered intramuscularly, and of orally and parenterally administered methadone. Clin Pharmacol Ther. 1967;8:415—26.
38. Hanna V, Senderovich H. Methadone in pain management: a systematic review. J Pain. 2021;22:233—45.
39. Santa Cruz Mercado LA, Liu R, Bharadwaj KM, et al. Association of intraoperative opioid administration with postoperative pain and opioid use. JAMA Surg. 2023;158:854—64.
40. Bamgbade OA, Oluwole O, Khaw RR. Perioperative analgesia for fast-track laparoscopic bariatric surgery. Obes Surg. 2017;27:1828—34.
41. Geng Z, Bi H, Zhang D, et al. The impact of multimodal analgesia based enhanced recovery protocol on quality of recovery after laparoscopic gynecological surgery: a randomized controlled trial. BMC Anesthesiol. 2021;21:179.
42. D'Souza RS, Esfahani K, Dunn LK. Pro-con debate: role of methadone in enhanced recovery after surgery protocols – superior analgesic or harmful drug? Anesth Analg. 2023;137:76—82.
43. Silva LMD, Ho AMH, Oliveira DRD, et al. Comparison of three intraoperative analgesic strategies in laparoscopic bariatric surgery: a retrospective study of immediate postoperative outcomes. Braz J Anesthesiol. 2022;72:560—6.

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



SHORT COMMUNICATION

# PRISMA 2020: strengthening transparency and usefulness of systematic reviews

**Leonardo Roever** [a,*], **Matthew J. Page** [b], **David Moher** [c,d,e]

[a] *Brazilian Evidence Based Health Network, Department of Clinical Research, Uberlândia, MG, Brazil*
[b] *Monash University, School of Public Health and Preventive Medicine, Methods in Evidence Synthesis Unit, Melbourne, Australia*
[c] *Ottawa Hospital Research Institute, Centre for Journalology, Methodological and Implementation Research Program, Ottawa, Canada*
[d] *University of Ottawa, Faculty of Medicine, School of Epidemiology and Public Health, Ottawa, Canada*
[e] *University of Toronto, Dalla Lana School of Public Health, Institute of Health Policy, Management & Evaluation, Toronto, Canada*

Received 26 May 2026; accepted 23 August 2026
Available online 20 September 2026

We are currently living in an era of unprecedented expansion of scientific literature. Millions of articles are published annually, making it virtually impossible for clinicians, policymakers, researchers, patients, and other stakeholders to keep up with all the available evidence. Systematic reviews are essential tools for organizing and interpreting this growing volume of information.[1,2]

However, the increasing popularity of systematic reviews has also brought significant challenges. Unnecessary duplicate reviews, opaque methods, incomplete search strategies, and inconsistent reporting can compromise the reliability of the synthesized evidence. In many cases, the absence of methodological details prevents the reproducibility of reviews, limiting their scientific value. Robust reporting guidelines are essential to improve the completeness, transparency, and reproducibility of systematic review reporting. However, transparent reporting alone does not ensure methodological rigor or validity of the review findings. PRISMA should be used in conjunction with tools that address the conduct of the review, risk of bias, certainty of evidence, and methodological assessment.[1-4]

In this context, the PRISMA 2020 guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) represent a fundamental milestone in strengthening the reliability of systematic reviews in contemporary healthcare.[1-5] Figure 1 shows the functional structure of PRISMA.

## Core components of PRISMA 2020 and their contribution to the transparency of systematic reviews

PRISMA 2020 represents the most recent update to the international PRISMA guidelines for reporting systematic reviews. Developed to reflect methodological advances of the preceding decade, PRISMA 2020 offers a structured set of items aimed at improving the transparency, completeness, and reproducibility of systematic reviews.[1-5] There are also many PRISMA extensions providing reporting guidance for different types or aspects of systematic reviews [e.g., for reporting a protocol of a systematic review (PRISMA-P), a scoping review (PRISMA-ScR), and a systematic review with network meta-analysis (PRISMA-NMA)]. All PRISMA documents (e.g., checklist, explanation and elaboration paper for PRISMA 2020 and PRISMA extensions) can be accessed on the PRISMA website (https://www.prisma-statement.org/), which receives more than 200,000 site visits per month. PRISMA 2020 has been translated into 17 languages [including in Brazilian Portuguese and Spanish (see all translated documents at https://www.prisma-statement.org/translations)].

The PRISMA 2020 guideline consists of four main components:

1. The main checklist with 27 items and abstracts checklist with 12 items.

* Corresponding author.
*E-mail:* leonardoroever@hotmail.com (L. Roever).

https://doi.org/10.1016/j.bjane.2026.844834
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).




**Figure 1** Shows the functional structure of PRISMA. The PRISMA 2020 framework (checklist, flow diagram, and explanation & elaboration document) provides a standardized structure that supports authors, editors, and reviewers across various institutions. Its primary goal is to enhance the clarity, completeness, and transparency of systematic review reports. By using this framework, stakeholder's can improves the usability and assessability of evidence syntheses, enabling more rigorous and informed evaluation by guideline developers, clinicians, and policymakers, ultimately contributing to a more robust global health evidence base thorough complete and transparent reporting.

2. An expanded checklist (which includes detailed reporting recommendations for each item).
3. Flow diagram templates for authors to document the flow of documents through the study selection process.
4. The Explanation and Elaboration document (which provides exemplars of good reporting and explanations to enhance the use and understanding of PRISMA 2020).

These components were developed to guide authors, peer reviewers, and editors through all stages of communicating systematic reviews.

## The 27-item checklist

The core of PRISMA 2020 is the 27-item checklist, which defines the critical aspects that should be reported in a systematic review. These items cover all sections of a systematic review manuscript, from the title to the declarations (e.g., funding and competing interests) statements. These steps ensure that critical information about methods and results is presented clearly and completely.[1]

The items are organized into seven major sections of the scientific article:

1. Title: The title should explicitly identify the study as a systematic review. This identification facilitates indexing in databases and improves literature retrievability.
2. Abstract: The abstract should provide a structured description of the review's objectives, methods, results, and conclusions, allowing readers to quickly understand the nature and relevance of the study.
3. Introduction: In this section, authors should explain the rationale for the review in the context of existing knowledge and clearly define the review objectives or questions.
4. Methods: The methods section is one of the most detailed components of the checklist. It includes items related to:
   - Eligibility criteria for studies;
   - Information sources;
   - Search strategies;
   - Study selection process;
   - Data collection process;
   - Data items;
   - Study risk of bias assessment process;
   - Effect measures;
   - Synthesis methods;
   - Reporting bias assessment;
   - Certainty assessment.
5. Results
   Authors should report: The number of studies identified, evaluated, and included; Characteristics of the selected studies; Study risk of bias assessments; Results of syntheses; Reporting bias assessments; Certainty assessments.
6. Discussion: The discussion should interpret the results in light of the existing literature, discuss limitations of the evidence included and the review processes used, and consider implications for clinical practice, health policies, and future research.
7. Other Information
   The checklist also includes items related to: Registration and protocol for the review; Support (e.g., Funding) for the review; Competing interests of review authors; Availability of data, code and other materials.

These elements reinforce transparency and scientific integrity.

## The expanded checklist

In addition to the main list of 27 items, PRISMA 2020 provides an expanded checklist, which offers more detailed guidance on how each item should be reported.[1] The goal of the expanded checklist is to reduce ambiguities in the interpretation of checklist items, helping authors to fully understand the expected level of detail.

## The PRISMA Flow diagram

Another central component of PRISMA 2020 is the study selection flow diagram, which provides a visual representation of the process of identifying, screening, and inclusion of studies. This visual representation significantly improves the transparency of the study selection process, allowing readers to clearly understand how evidence was identified and filtered. Different flow diagram templates are available depending on whether the review is new or updated, and on the type of sources searched.

## Explanation and elaboration paper

The Explanation and Elaboration paper provides an in-depth description of each item on the checklist.[1] This document includes: The scientific rationale for each item; Exemplars of reporting each checklist item drawn from published systematic reviews; Detailed guidance on good reporting practices.

The aim of this document is to help authors understand not only what should be reported, but also why each piece of information is important for the interpretation of the systematic reviews.

## Importance of PRISMA 2020

The importance of robust reporting guidelines has become even more evident in an era characterized by the rapid dissemination of scientific information and misinformation. Spurious associations between clinical interventions and health outcomes can spread rapidly when based on poorly conducted or poorly reported evidence. The increasing integration of Artificial Intelligence (AI) into evidence synthesis offers important opportunities to improve the efficiency of selected tasks, including literature screening, citation management, and data extraction. However, AI-assisted technologies should be regarded as supportive tools rather than replacements for scientific judgment. Critical stages of the systematic reviews process – including study selection, critical appraisal, interpretation of findings, assessment of the certainty of the evidence, and formulation of conclusions – require human expertise, methodological competence, and accountability. Accordingly, the use of AI should always be accompanied by appropriate human oversight and independent verification of outputs. To promote transparency and reproducibility, authors should explicitly report the AI tools used, their specific purpose within the review workflow, the extent of human supervision, the procedures adopted to verify AI-generated outputs, and the measures taken to preserve an auditable record of the research process. Although PRISMA 2020 was developed before the widespread adoption of generative AI, its core principles of transparency, completeness, and reproducibility provide a strong foundation upon which future reporting guidance for AI-assisted evidence synthesis can be built.[4-6]

Although systematic reviews are often seen as methodological tools, their impact extends far beyond academia. They inform clinical guidelines, regulatory decisions, intervention funding, patient decision-making, and global health strategies. When properly reported, systematic reviews can help identify evidence gaps, underrepresented populations, and inequalities in scientific output. In this way, guidelines such as PRISMA 2020 also contribute to promoting greater equity in global health research. By improving the transparency of evidence syntheses, PRISMA 2020 allows researchers, funders, and policymakers to better understand where the evidence is strong and where it is still insufficient.

PRISMA 2020 represents an evolution in the reporting of systematic reviews and meta-analyses. At a historical moment when trust in science faces challenges and the volume of scientific information grows exponentially, robust standards of transparency and reporting are essential to preserving the credibility of health evidence. The true impact of PRISMA 2020, however, will depend on the collective commitment of researchers, reviewers, editors, and funders to fully adopt it. Only in this way will it be possible to ensure that systematic reviews continue to play their central role in promoting evidence-based decisions and, ultimately, in improving the health of populations worldwide.

To maximize the impact of PRISMA 2020 authors, peer reviewers, and journal editors should work collaboratively to strengthen the transparency and reproducibility of systematic reviews. Authors should prospectively register their systematic review protocols and transparently report and justify any deviations from the original protocol. Complete and reproducible search strategies for all information sources should be provided, together with transparent descriptions of duplicate screening, study selection, and data extraction procedures. Systematic reviews should comprehensively report risk-of-bias assessments and the certainty of the evidence using established frameworks, such as GRADE, whenever appropriate. In keeping with policy frameworks in several parts of the world, there is an increasing expectation for data sharing to occur responsibly, where ethical, legal, and commercial obligations allow, in line with FAIR principles: as open as possible, as closed as necessary. Authors of systematic reviews should make extracted datasets, analytic code, statistical scripts, and supplementary materials publicly available to facilitate verification and reproducibility. Finally, as Artificial Intelligence (AI) becomes increasingly integrated into evidence synthesis workflows, any use of AI-assisted technologies should be explicitly disclosed, together with appropriate human oversight. Collectively, these practices reinforce the principles of transparency, reproducibility, and research integrity that underpin PRISMA 2020 and enhance the reliability of systematic reviews for clinical decision-making.

## Data availability statement

No new data were created or analyzed in this article. Data sharing is not applicable to this article.

## Use of generative AI tools

During the preparation of this manuscript, the authors used ChatGPT 5 (OpenAI, San Francisco, CA, USA) solely to

improve language clarity and readability. No AI tool was used to generate scientific conclusions, select references, assess evidence, analyze data, or interpret findings. The authors reviewed and edited all AI-assisted output and take full responsibility for the final content.

## Authors' contributions

Leonardo Roever: Writing-original draft; writing-revision and editing.

Matthew J. Page: Writing-revision and editing; writing-original draft.

David Moher: Writing-original draft; writing-revision and editing; conceptualization.

## Funding

The authors declare that they did not receive any financial support for this work and/or its publication.

## Conflicts of interest

Matthew J. Page and David Moher are authors of the PRISMA 2020 Statement. This role is disclosed for transparency. The authors declare that this work was carried out in the absence of any commercial or financial relationships that could be interpreted as a potential conflict of interest.

## Editor

Liana Azi

## References

1. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
2. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160.
3. Shaheen N, Shaheen A, Ramadan A, et al. Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. Front Res Metr Anal. 2023;8:1268045.
4. Harasgama S, Pearce H, Appel C, et al. Artificial Intelligence Tools for Automating Evidence Synthesis: Scoping Review. J Med Internet Res. 2026;28:e81597.
5. De Cassai A, Dost B, Tulgar S, Boscolo A. Methodological Standards for Conducting High-Quality Systematic Reviews. Biology. 2025;14:973.
6. Lunny C, Kanji S, Thabet P, Haidich A, Bougioukas KI, Pieper D. Assessing the methodological quality and risk of bias of systematic reviews: primer for authors of overviews of systematic reviews. BMJ Medicine. 2024;3:e000604.

Brazilian Journal of Anesthesiology 2026;76(6): 844821



Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY

LETTER TO THE EDITOR

# Post-dural puncture headache in patients ≥ 50 years: a multicenter cohort study



*Dear Editor,*

Post-Dural Puncture Headache (PDPH) remains a recognized complication of spinal anesthesia, with needle design traditionally emphasized as a key modifiable risk factor.[1] Atraumatic (pencil-point) needles are widely recommended due to their lower associated PDPH rates.[2-4] However, PDPH appears to be uncommon in older adults, consistent with the well-established inverse relationship between age and PDPH risk.[1,5] Consequently, two clinically relevant questions arise in this population: how frequently does clinically confirmed PDPH occur, and does needle type remain an important determinant of risk when the underlying event rate is already very low? Given the anticipated rarity of PDPH in older adults, this analysis was principally intended to describe event frequency and explore, rather than definitively test, any association between needle type and PDPH.

We conducted a retrospective multicenter cohort study of patients ≥ 50 years undergoing orthopedic surgery under spinal anesthesia. Our primary objective was to characterize the incidence of clinically confirmed PDPH in this older population, while secondarily exploring whether PDPH occurrence differed according to spinal needle type (22-gauge Quincke vs. 25-gauge Whitacre). Potentially eligible patients were identified retrospectively through institutional perioperative electronic medical records. Patients receiving spinal anesthesia with either a 22-gauge Quincke (cutting) or 25-gauge Whitacre (pencil-point) needle between May 2022 and February 2025 across four acute care hospitals were included. The primary outcome was clinically confirmed PDPH within three weeks of surgery, defined as a clinician-documented diagnosis of PDPH attributed to dural puncture rather than a non-specific postoperative headache. PDPH typically presented as a severe, positional headache accompanied by nausea, vomiting, tinnitus, photophobia, and neck stiffness, consistent with routine clinical documentation and assessment. The three-week follow-up window was selected to capture delayed presentations occurring after hospital discharge, including emergency department visits, anesthesiology consultation assessments, or epidural blood patch procedures. Secondary outcomes included documentation of any postoperative headache, recognizing that non-specific headache is common in the postoperative setting. Inclusion criteria included 22-gauge Quincke, or 25-gauge Whitacre needles used for spinal anesthesia. Exclusion criteria included use of other needle types or gauges, multiple needle attempts with different sizes, failed spinal anesthesia requiring conversion to general anesthesia and repeat surgeries as only the first procedure was considered for each patient. Because needle choice was not randomized and appeared to reflect anticipated technical difficulty, comparisons between needle types should be interpreted as exploratory and hypothesis-generating.

Incident rates with Clopper-Pearson 95% Confidence Intervals (95% CI) and absolute risk differences using 10,000 bootstrap replicates were calculated for documented headaches and PDPH. Relative effects of needle size on headaches and PDPH were assessed using binomial logistic regression and Firth penalized regression respectively and presented with Odds Ratios (OR) and 95% CI. We used a three-moment entropy balance weighting framework to balance demographic variable means, variances, and skewness between needle types. Variables included in the entropy model were age, sex, weight, height, Body Mass Index (BMI), American Society of Anesthesiologists' (ASA) physical status, and procedure regressed on needle type using an Average Treatment effect in the Treated (ATT) estimand. The number of spinal needle insertion attempts was omitted from the weighting model as it represents a post-treatment variable that may lie on the causal pathway between needle type and our outcomes. We considered acceptable balance to be an absolute standardized mean difference $< 0.1$. Equivalent weighted binomial and Firth penalized regression analyses were completed that incorporated entropy weights. Robust sandwich standard errors were applied to the weighted binomial model. We considered $p < 0.05$ to be significant.

A total of 6,306 patients met inclusion criteria, including 418 in the Quincke group and 5,888 in the Whitacre group. There were marked baseline differences between groups. Patients receiving Quincke needles were older, had higher ASA physical status, and more frequently underwent procedures associated with technical difficulty (e.g., hip fracture and hemiarthroplasty) (Table 1). They also required a greater number of spinal needle insertion attempts. Together, these findings strongly suggest that needle choice was driven by anticipated anatomical and procedural challenges rather than random allocation, introducing substantial confounding. Consequently, this study should not be interpreted as a comparative effectiveness analysis and cannot establish equivalence between Quincke and Whitacre needles with respect to PDPH risk.

https://doi.org/10.1016/j.bjane.2026.844821
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

**Table 1** Baseline and procedural characteristics by spinal needle type.

| | 25-gauge Whitacre needle patients (n = 5888) | 22-gauge Quincke needle patients (n = 418) | p-value | ASD | ASD after entropy balancing |
|---|---|---|---|---|---|
| **Age, years, median [IQR]** | 70.0 [63.0–77.0] | 77.0 [68.0–84.3] | < 0.001 | 0.635 | 0.077 |
| **Sex, female, n/total, n (%)** | 3469/5888 (58.9) | 252/418 (60.3) | 0.582 | 0.028 | 0.012 |
| **Weight, kg, median [IQR]** | 80.7 [67.9–94.3] | 76.6 [62.1–94.4] | 0.009 | 0.085 | 0.040 |
| **Height, m, median [IQR]** | 1.65 [1.57–1.73] | 1.65 [1.57–1.73] | 0.850 | 0.023 | 0.004 |
| **BMI, kg·m$^{-2}$, median [IQR]** | 29.1 [25.3–33.5] | 28.8 [24.7–33.6] | 0.266 | 0.029 | 0.012 |
| **ASA physical status, n/total, n (%)** | | | < 0.001 | | |
| 1 | 331/5861 (5.6) | 6/414 (1.4) | | 0.186 | 0.001 |
| 2 | 3571/5861 (60.9) | 156/414 (37.7) | | 0.470 | 0.022 |
| 3 | 1859/5861 (31.7) | 230/414 (55.6) | | 0.525 | 0.016 |
| ≥ 4 | 100/5861 (1.7) | 22/414 (5.3) | | 0.280 | 0.007 |
| **Procedure** | | | <0.001 | | |
| Total knee arthroplasty | 2783/5888 (47.3) | 157/418 (37.6) | | 0.194 | 0.022 |
| Total hip arthroplasty | 1281/5888 (21.8) | 53/418 (12.7) | | 0.242 | 0.008 |
| Hip fracture | 327/5888 (5.6) | 85/418 (20.4) | | 0.451 | 0.033 |
| Hemiarthroplasty | 316/5888 (5.4) | 68/418 (16.3) | | 0.356 | 0.005 |
| Ankle fusion | 190/5888 (3.2) | 13/418 (3.1) | | 0.007 | 0.002 |
| Partial knee replacement | 125/5888 (2.1) | 3/418 (0.7) | | 0.119 | 0.000 |
| Other | 866/5888 (14.7) | 39/418 (9.3) | | 0.166 | 0.005 |
| **Number of spinal needle insertion attempts, n/total, n (%)** | | | < 0.001 | | |
| 1 | 4148/5674 (73.1) | 165/411 (40.1) | | | |
| 2 | 1026/5674 (18.1) | 123/411 (29.9) | | | |
| 3 | 367/5674 (6.5) | 77/411 (18.7) | | | |
| 4 | 81/5674 (1.4) | 23/411 (5.6) | | | |
| ≥ 5 | 52/5674 (0.9) | 23/411 (5.6) | | | |

ASD, Absolute Standardized Difference; IQR, Interquartile Range; BMI, Body Mass Index; ASA, American Society of Anesthesiologists. Entropy balancing completed on pre-exposure demographic variables of age, sex, weight, height, BMI, ASA physical status, and procedure. Acceptable balance after entropy balancing was an ASD < 0.10.

*Maximum missing data was associated with patient height (5.4% in the 25-gauge group and 13.3% in the 22-gauge group). A complete-case analysis was used for all outcomes given the rarity of our post-dural puncture headache (n = 2). Such rarity can lead to separation issues and unstable estimates within a multiple imputation framework. Both patients experiencing post-dural puncture headache had complete baseline data and as such, use of a complete-case approach did not reduce statistical power.

Four-hundred and fifty-four total headaches were documented in 6,306 patients 7.2% (95% CI 6.6% to 7.9%). Postoperative headache of any type was documented in 10.0% (95% CI 7.3% to 13.3%) of patients in the Quincke group and 7.0% (95% CI 6.4% to 7.7%) in the Whitacre group. While the odds of postoperative headache were significantly higher in the Quincke group (Odds Ratio [OR = 1.48], 95% CI 1.05 to 2.05) on unadjusted analysis, this significant difference was no longer observed after adjustment for baseline imbalance using entropy balance weighting (OR = 0.98, 95% CI 0.71 to 1.37). This attenuation supports the interpretation that observed differences were attributable to patient and procedural factors rather than needle design per se. It also underscores the limitations of unadjusted comparisons in large observational datasets, particularly where treatment allocation is non-random.

Only two cases of PDPH were identified across the entire cohort of 6,306 patients 0.03% (95% CI 0.004% to 0.11%) with both occurring in the Whitacre group 0.03% (95% CI 0.004% to 0.12%). No cases were observed in patients receiving a Quincke needle 0.0% (95% CI 0.0% to 0.88%). Although the overall cohort was large, the very low number of PDPH events precluded meaningful comparative modeling or reliable estimation of between-group differences in PDPH risk. The absolute risk difference was -0.03% (95% CI -0.08 to 0.01%). Due to the zero-event count and large denominator imbalance between groups, Firth penalized regression resulted in an unstable sparse data estimated PDPH OR of 2.81 (95% CI 0.02 to 34.6, p = 0.553) in the Quincke group. The corresponding OR after entropy balancing was 3.50 (95% CI 0.02 to 52.6, p = 0.490). The extreme width of the 95% CI reflects large statistical imprecision rather than a true clinical increase in PDPH odds with a Quincke needle. Given the extremely low number of events, all comparative analyses were underpowered and should be interpreted as exploratory. Although the overall cohort was large, only two PDPH events were identified, precluding reliable estimation of between-group differences in PDPH risk. As such, no reliable inference can be made regarding the relative risk of PDPH between needle types. However, it is notable that even the upper bounds of estimated incidence remained very low in both groups, highlighting the rarity of clinically confirmed PDPH in this population.

These findings emphasize a key practical message: clinically confirmed PDPH was very rare in this patient cohort, with very few events observed in either needle group. This observation is consistent with existing literature demonstrating a strong inverse relationship between age and PDPH risk.[1,5] Several mechanisms have been proposed, including decreased dural

compliance, reduced cerebrospinal fluid leakage through puncture sites, and age-related changes in intracranial dynamics.[5] In addition, altered pain perception and reporting in older adults may influence the clinical expression and recognition of PDPH.

In this context, the relative importance of needle design may be less pronounced than in younger populations. While atraumatic needles clearly reduce PDPH incidence in high-risk groups,[2-4] the very low observed incidence of PDPH in this cohort suggests that any absolute difference attributable to needle type would likely be small, although this study was not powered to estimate that difference reliably. Consequently, other factors may reasonably guide spinal needle selection. Technical success and procedural feasibility are important considerations in elderly orthopedic patients, who often present with degenerative spinal changes, limited mobility, and reduced interspinous space. In technically challenging cases, clinicians may choose needle type based on perceived ease of use and procedural feasibility; however, the present study cannot determine whether one needle type improves procedural success, reduces insertion attempts, or minimizes patient discomfort.

From a pragmatic standpoint, our findings support a more individualized approach to needle selection in older adults. Rather than strict adherence to needle type based solely on PDPH prevention, clinicians may consider anatomical factors, anticipated difficulty, and operator experience when choosing equipment. This aligns with real-world practice patterns, as reflected in our data, where Quincke needles were preferentially used in more complex cases.

Several limitations should be acknowledged. First, the extremely low incidence of PDPH limits statistical power and precludes robust comparative conclusions. Although the overall cohort was large, only two PDPH events occurred, preventing reliable estimation of between-group differences in PDPH risk. Second, significant baseline imbalance reflects non-randomized needle selection and introduces substantial potential for confounding by indication despite statistical adjustment. Consequently, any comparison between needle types should be regarded as exploratory and hypothesis-generating rather than causal. Third, the retrospective design depends on clinical documentation, which may underestimate both headache frequency and diagnostic specificity. Misclassification between nonspecific postoperative headache and PDPH is possible. Finally, our findings may not be generalizable beyond older orthopedic populations, where baseline PDPH risk is already low.

In this multicenter retrospective cohort of orthopedic patients $\geq$ 50 years, clinically confirmed PDPH was exceedingly rare. Because only two PDPH events occurred, the study cannot support reliable comparative conclusions regarding needle type. These findings suggest that the absolute risk of PDPH in older orthopedic patients is low and that needle selection may reasonably incorporate technical considerations, patient anatomy, and operator judgment. Prospective studies with standardized PDPH ascertainment would be valuable to confirm these observations.

## Data availability statement

In accordance with the Alberta Health Services Data Disclosure Agreement, we are unable to provide or make the data available to a third party without prior written consent from Alberta Health Services, Calgary, Alberta, Canada. Custom-generated R code used for data analysis is available from the authors upon reasonable request, subject to institutional approval and applicable data-sharing agreements.

## Ethics approval

This study was approved by the University of Calgary Conjoint Health Research Ethics Board. Approval number: REB24-0612.

## Trial registration

Not applicable. Retrospective observational study.

## Declaration of AI and AI-assisted technologies

The authors declare that no Artificial Intelligence (AI) or AI-assisted technologies were used in the conception, data analysis, or writing of this manuscript.

## Authors' contributions

Scott Seadon: Study conception; Data acquisition; Manuscript drafting.

Andrew Walker: Statistical analysis; Data interpretation; Manuscript drafting and revision.

Abraham O. Fifen: Data interpretation; Critical manuscript revision.

Joanna J. Fifen: Study conception and design; Ethics oversight; Data interpretation; Manuscript drafting and revision; Final approval.

All authors meet ICMJE authorship criteria and approved the final manuscript.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for profit sectors.

## Conflicts of interest

The authors declare no conflicts of interest.

## Editor

Liana Azi

## References

1. Turnbull DK, Shepherd DB. Post-dural puncture headache: pathogenesis, prevention and treatment. Br J Anaesth. 2003;91:718—29.
2. Zorrilla-Vaca A, Mathur V, Wu CL, Grant MC. The impact of spinal needle selection on postdural puncture headache: a meta-analysis and

metaregression of randomized studies. Reg Anesth Pain Med. 2018;43:502—8.
3. Uppal V, Russell R, Sondekoppam R, et al. Consensus practice guidelines on postdural puncture headache from a multisociety, international working group: a summary report. JAMA Netw Open. 2023;6:e2325387.
4. Nath S, Koziarz A, Badhiwala JH, et al. Atraumatic versus conventional lumbar puncture needles: a systematic review and meta-analysis. Lancet. 2018;391:1197—204.
5. Kim JE, Kim SH, Han RJW, Kang MH, Kim JH. Postdural puncture headache related to procedure: incidence and risk factors after neuraxial anesthesia and spinal procedures. Pain Med. 2021;22:1420—5.

Scott Seadon [a,b], Andrew Walker [a], Abraham O. Fifen [c], Joanna J. Fifen [a,*]

[a] *University of Calgary, Cumming School of Medicine, Department of Anesthesia, Perioperative and Pain Medicine, Calgary, Alberta, Canada*
[b] *University of Calgary, Cumming School of Medicine, Department of Emergency Medicine, Calgary, Alberta, Canada*
[c] *University of Saskatchewan, Department of Public Health and Preventive Medicine, Saskatoon, Saskatchewan, Canada*

[*] Corresponding author.
*E-mail:* Joanna.Fifen@albertahealthservices.ca (J.J. Fifen).
Received 29 May 2026; accepted 16 August 2026
Available online 25 August 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# Beyond forest plots: trial sequential analysis of magnesium sulfate as an adjuvant in supraclavicular brachial plexus block

*Dear Editor,*

Peripheral nerve block adjuvants are frequently investigated in regional anesthesia to prolong sensory and motor block duration and improve postoperative analgesia, particularly when the duration of local anesthetics alone may be insufficient. In this context, magnesium sulfate has been evaluated as a potential adjuvant to local anesthetics.[1] Although conventional meta-analyses[2] have suggested clinically relevant benefits, these conclusions are commonly based on cumulative forest plots, which are vulnerable to random error and inflation of type I error, especially when the number of trials is limited or when repeated significance testing is performed.

Trial Sequential Analysis (TSA) extends conventional meta-analysis by applying principles analogous to interim monitoring in randomized trials. TSA adjusts significance thresholds according to the accrued information size and constructs sequential monitoring boundaries, thereby distinguishing robust evidence from potentially spurious findings.[3] Therefore, applying TSA to previously reported pooled estimates may help clarify which conclusions are statistically robust after adjustment for random error and which should remain hypothesis-generating.

## Methods

This TSA was based on the original meta-analytic database used in the previously published BJAN meta-analysis on magnesium sulfate as an adjuvant in supraclavicular brachial plexus block.[2] No new literature search, study selection, or data extraction from primary randomized trials was performed. Outcome-level data, including effect estimates and corresponding measures of variance, were retrieved from the original meta-analytic dataset and checked for consistency with the pooled estimates reported in the published meta-analysis. This approach was chosen because the objective of the present study was to reassess whether the original pooled estimates crossed prespecified TSA monitoring boundaries after adjustment for random error and accrued information size. Thus, the TSA was designed as a methodological reassessment of the previously reported meta-analytic estimates rather than as a new effectiveness analysis based on independently defined clinical thresholds. The original meta-analysis included randomized controlled trials comparing local anesthetics with or without magnesium sulfate as an adjuvant for supraclavicular brachial plexus block in adult patients. Additionally, the authors followed a structured checklist[4] based on the EQUATOR Network Reporting Guidelines (https://www.equator-network.org/reporting-guidelines/).

The post-publication status of the randomized trials included in the original dataset and all references cited in this letter to the editor was checked using the Retraction Watch Database, PubMed/MEDLINE, and publishers' websites, including CrossMark when available. No retraction notices or expression of concern were identified.

The following outcomes were evaluated: onset of sensory block, onset of motor block, duration of sensory block, duration of motor block, and rescue analgesic consumption within 24 hours.

TSA was conducted using TSA software.[5,6] For each continuous outcome, the Required Information Size (RIS) was calculated using a random-effects model. The type I error risk was set at 5% using a two-sided alpha, and the type II error risk was set at 20%, corresponding to 80% statistical power. The anticipated intervention effect was defined according to the pooled effect estimate obtained in the previously published conventional meta-analysis. Between-study heterogeneity was incorporated into the RIS using the diversity/heterogeneity adjustment provided by the TSA software. Sequential monitoring boundaries were constructed using the O'Brien-Fleming approach, because this method applies conservative thresholds during early cumulative evidence accrual and reduces the risk of false-positive conclusions due to repeated significance testing. A cumulative *Z*-curve crossing the trial sequential monitoring boundary for benefit was interpreted as evidence that the previously reported pooled estimate was robust after adjustment for random error and accrued information size. Failure to cross the monitoring boundary was interpreted as inconclusive sequential evidence.

## Results

Outcome-specific TSA assumptions are presented in Supplementary Table 1.

https://doi.org/10.1016/j.bjane.2026.844835
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

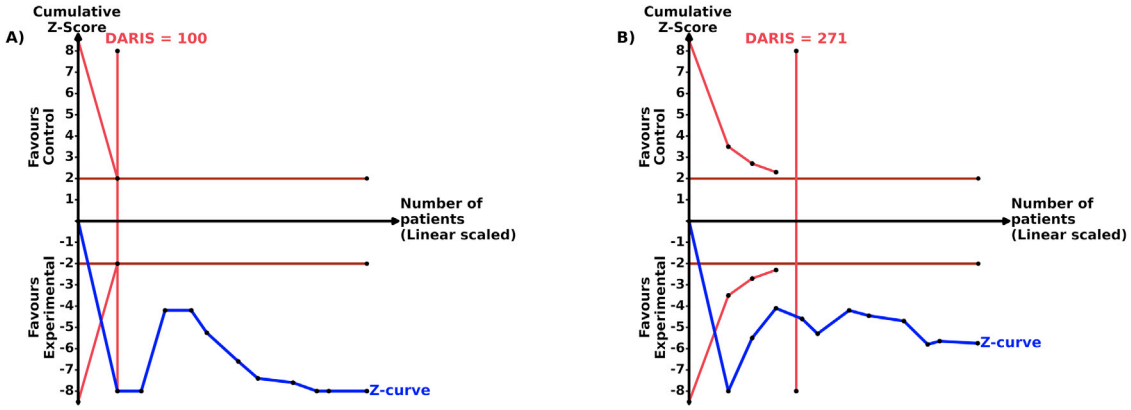



**Figure 1** TSA for duration outcomes: (A) Sensory block duration and (B) Motor block duration. The x-axis shows the cumulative information size, and the y-axis shows the cumulative *Z*-score. Blue lines represent cumulative *Z*-curves; red lines represent conventional significance boundaries and O'Brien-Fleming trial sequential monitoring boundaries; the vertical red line indicates the Diversity-Adjusted Required Information Size (DARIS). Favours experimental: local anesthetic plus magnesium sulfate; and favours control: local anesthetic alone.

## Block duration

For sensory block duration, 10 studies contributing 11 comparison units and 734 participants were included in the final TSA. The Diversity-Adjusted Required Information Size (DARIS) was reached, and the final cumulative Z-curve crossed the O'Brien-Fleming monitoring boundary, favoring magnesium sulfate (Fig. 1A).

For motor block duration, 10 studies contributing 11 comparison units and 734 participants were included in the final TSA. The DARIS was reached, and the final cumulative *Z*-curve crossed the O'Brien-Fleming monitoring boundary, favoring magnesium sulfate (Fig. 1B).

## Block onset

For sensory block onset, 10 studies contributing 11 comparison units and 734 participants were included in the final TSA. The DARIS was reached; however, the final cumulative *Z*-curve did not clearly cross the O'Brien-Fleming monitoring boundary, although it crossed the conventional significance boundary (Fig. 2A). Therefore, this outcome was considered inconclusive after sequential adjustment.

For motor block onset, 10 studies contributing 11 comparison units and 734 participants were included in the final TSA. The DARIS was reached; however, the final cumulative *Z*-curve did not clearly cross the O'Brien-Fleming monitoring boundary, despite crossing the conventional significance boundary (Fig. 2B). Therefore, this outcome was also considered inconclusive after sequential adjustment.

## 24-hour rescue analgesic consumption

For 24-hour rescue analgesic consumption, 4 studies contributing 4 comparison units and 266 participants were included in the final TSA. The DARIS was reached, and the final cumulative *Z*-curve crossed the O'Brien-Fleming monitoring boundary, favoring magnesium sulfate (Fig. 2C).

## Discussion

The present TSA provides a methodological reassessment of previously published pooled estimates on magnesium sulfate as an adjuvant in supraclavicular brachial plexus block, rather than new clinical evidence regarding the intervention itself.

Sequential monitoring is particularly relevant in anesthesiology research, where meta-analyses often rely on a limited number of relatively small, randomized trials. In such settings, repeated significance testing increases the risk of type I error, potentially leading to premature or spurious conclusions. TSA addresses this limitation by applying adjusted significance thresholds and by estimating the DARIS, thereby offering a framework analogous to interim analyses in single randomized controlled trials. Our findings illustrate this principle: for sensory and motor block duration, the cumulative *Z*-curves crossed the prespecified TSA monitoring boundaries under the stated assumptions, whereas outcomes related to block onset did not cross these boundaries despite nominal statistical significance in conventional analyses.

For 24-hour rescue analgesic consumption, the TSA indicated that the cumulative information size exceeded the DARIS, and the final cumulative Z-curve crossed the prespecified O'Brien-Fleming monitoring boundary, favoring magnesium sulfate under the stated assumptions. This suggests that, within the TSA framework, the observed reduction in rescue analgesic consumption was unlikely to be explained by random error alone. However, TSA addresses the risk of random errors due to sparse and repeated testing; it does not by itself resolve other sources of uncertainty, such as clinical heterogeneity, risk of bias, or outcome-level differences across trials.

Taken together, these results emphasize that TSA should be viewed not as a replacement for conventional meta-analysis, but as a complementary tool that distinguishes robust evidence from findings that remain hypothesis-generating. For magnesium sulfate as an adjuvant in supraclavicular block, TSA suggests that the previously reported prolongation of block duration and reduction in 24-hour rescue analgesic consumption are statistically robust, whereas the pooled estimates for

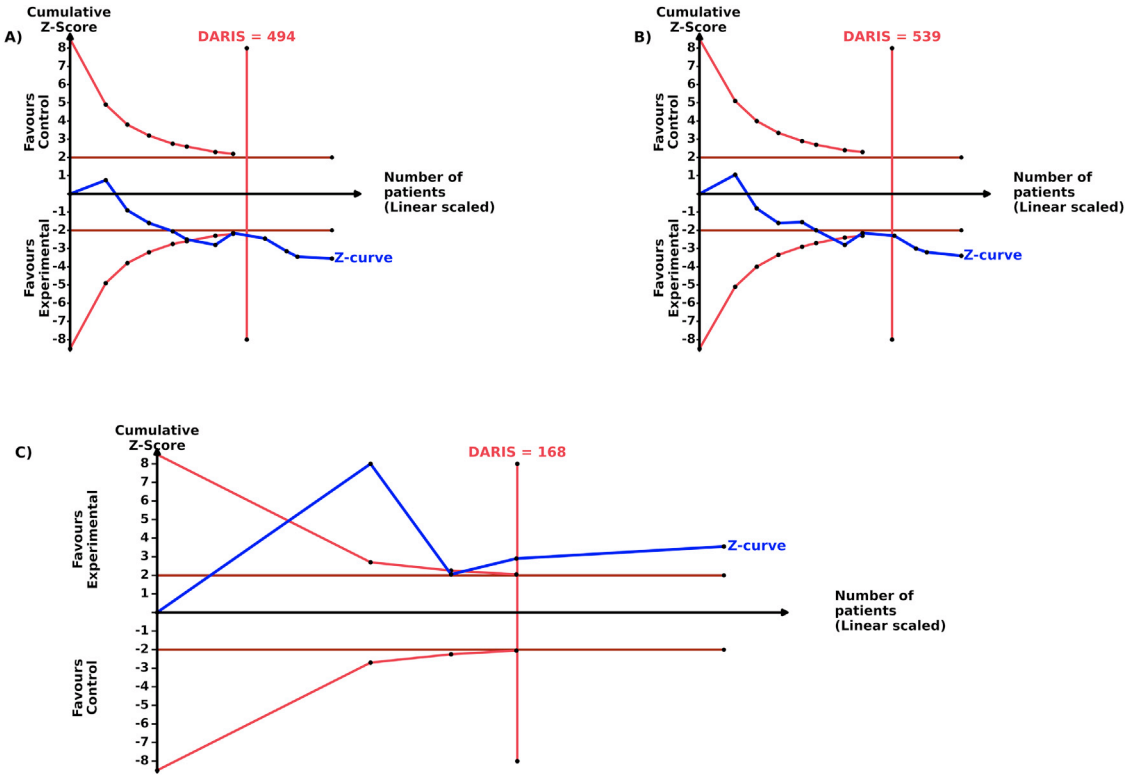



**Figure 2** TSA for onset and postoperative analgesia outcomes: (A) Sensory block onset, (B) Motor block onset, and (C) Rescue analgesic consumption within 24 hours. The x-axis shows the cumulative information size, and the y-axis shows the cumulative Z-score. Blue lines represent cumulative Z-curves; red lines represent conventional significance boundaries and O'Brien-Fleming trial sequential monitoring boundaries; the vertical red line indicates the Diversity-Adjusted Required Information Size (DARIS). Favours experimental: local anesthetic plus magnesium sulfate; and favours control: local anesthetic alone.

block onset remain uncertain and should be interpreted as hypothesis-generating.

## Conclusion

This TSA supports the robustness against random error of the previously reported benefits of magnesium sulfate on sensory and motor block duration and 24-hour rescue analgesic consumption in supraclavicular brachial plexus block. For these outcomes, the cumulative information size exceeded the DARIS and the cumulative Z-curves crossed the prespecified O'Brien-Fleming monitoring boundaries under the stated TSA assumptions. In contrast, sensory and motor block onset did not clearly cross the monitoring boundaries and therefore remained inconclusive after sequential adjustment. These findings clarify which pooled estimates are statistically robust within the TSA framework and which remain uncertain after adjustment for random error.

These findings should not be interpreted as a recommendation for routine perineural magnesium sulfate use, which remains off-label. Interpretation should remain cautious given trial heterogeneity, the secondary nature of this methodological reanalysis, and certainty limitations in the original evidence base. Further adequately powered randomized trials with standardized dosing, outcome reporting, and safety assessment are still needed.

## Data availability statement

The complete dataset and materials supporting this Trial Sequential Analysis have been deposited and are publicly available at https://doi.org/10.5281/zenodo.20891978. The deposited files include the original study-level extraction dataset used in the prior meta-analysis, TSA input files/settings, and outputs. No statistical code was used for TSA; analyses were performed using TSA software.

## AI assistance disclosure

The authors used AI-assisted technology (ChatGPT, OpenAI) exclusively for linguistic support, including grammar revision and translation. All content was critically reviewed and approved by the authors.

## Funding

The authors declare that no financial support or funding was received for this study.

## Conflicts of interest

The authors declare no conflicts of interest.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjane.2026.844835.

## Editor

Liana Azi

## References


1. Li M, Jin S, Zhao X, et al. Does magnesium sulfate as an adjuvant of local anesthetics facilitate better effect of perineural nerve blocks? A meta-analysis of randomized controlled trials. Clin J Pain. 2016;32:1053—61.
2. Saraiva WB, Candido IE, Caldas RRB, Barbosa FT. Efficacy of magnesium sulfate as an adjuvant to local anesthetics in supraclavicular brachial plexus block: a meta-analysis of randomized trials. Braz J Anesthesiol. 2026;76:844689.
3. Kang H. Trial sequential analysis: novel approach for meta-analysis. Anesth Pain Med (Seoul). 2021;16:138—50.
4. Coskun Benlidayi I, Gupta L. Letter to the Editor Standards (LETTERS): considerations for authors, reviewers, and editors. J Korean Med Sci. 2024;39(37):e296.
5. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). Copenhagen: Copenhagen Trial Unit, Centre for Clinical Intervention Research; 2011.
6. Claire R, Gluud C, Berlin I, Coleman T, Leonardi-Bee J. Using trial sequential analysis for estimating the sample sizes of further trials: example using smoking cessation intervention. BMC Med Res Methodol. 2020;20:284.



Willyam Barros Saraiva [a,c,*], Rafael Wanderley Persiano Malta [a,c], Roberta Ribeiro Brandão Caldas [a,b], Olavo Barbosa de Oliveira-Neto [c], Fabiano Timbó Barbosa [b,c], Célio Fernando de Sousa-Rodrigues [a,c]

[a] *Universidade Federal de Alagoas, Hospital Universitário Professor Alberto Antunes, Maceió, AL, Brazil*
[b] *Sociedade Brasileira de Anestesiologia, Higher Degree in Anesthesiology (TSA), Rio de Janeiro, RJ, Brazil*
[c] *Universidade Federal de Alagoas, Maceió, AL, Brazil*

[*] Corresponding author.
*E-mail:* will.bsaraiva@gmail.com (W.B. Saraiva).


Received 12 April 2026; accepted 16 August 2026
Available online 20 September 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# The premature abandonment of thoracic epidural analgesia in the Era of video-assisted thoracoscopic surgery thoracic epidural analgesia in the VATS Era

*Dear Editor,*

The evolution from open thoracotomy to Video-Assisted Thoracic Surgery (VATS) has significantly modified perioperative analgesia strategies. Contemporary procedure-specific recommendations, most notably the PROSPECT guidelines for VATS, now advocate the thoracic Paravertebral Block (PVB) or Erector Spinae Plane Block (ESPB) as first-choice regional techniques rather than Thoracic Epidural Analgesia (TEA).[1] These approaches have gained ground because they are less invasive and better aligned with enhanced-recovery pathways. Contemporary evidence therefore supports moving away from routine TEA for most VATS patients; however, whether TEA has no role in any clinical circumstance remains uncertain.[1]

The randomized trials comparing TEA and PVB illustrate this trade-off rather than supporting a uniform conclusion. Lai et al. randomized 86 patients undergoing VATS lobectomy and found lower pain scores at rest and during coughing at 24 and 48 hours with TEA. However, TEA was associated with more vomiting (32.6% vs. 11.6%), dizziness (55.8% vs. 12.9%), pruritus (27.9% vs. 2.3%), and hypotension (32.6% vs. 4.7%).[2] In contrast, Wu et al., in a randomized trial of 176 patients, found no significant difference in pain scores between PVB and TEA. PVB required less time for catheter placement (median 5 vs. 14 minutes), had a lower catheter-placement failure rate, and was associated with lower incidences of hypotension and urinary retention.[3] These studies highlight the central trade-off: TEA may provide greater analgesic intensity in some comparisons, whereas PVB may provide similar analgesia with practical and safety advantages in others.

Nevertheless, the safety argument favoring peripheral techniques also deserves appropriate context. Less invasive does not mean risk-free; rather, TEA and peripheral blocks have different technical and adverse-effect profiles that should be weighed alongside analgesic efficacy and recovery outcomes.

The recent multicenter randomized non-inferiority trial provides important contemporary context. Among 450 randomized patients, 389 were included in the intention-to-treat analysis. The proportion of pain assessments with NRS scores $\geq 4$ was 20.7% with TEA, 35.5% with PVB, and 29.5% with Intercostal Nerve Block (ICNB). Using a prespecified non-inferiority margin of 17.5%, ICNB met the criterion for non-inferiority to TEA, whereas PVB was inferior to TEA for the primary pain outcome in both intention-to-treat and per-protocol analyses.[4] Both PVB and ICNB nevertheless reduced opioid consumption and improved mobility compared with TEA, while quality-of-recovery scores were similar across groups.[4] These findings reinforce that non-inferiority should not be interpreted as equivalence and that analgesic efficacy must be weighed against recovery-related outcomes. Importantly, the trial excluded patients receiving long-term opioid therapy.[4] This exclusion identifies an evidence gap rather than demonstrating that such patients would benefit more from TEA. Accordingly, patients with chronic opioid use, difficult-to-control pain, or severely compromised respiratory reserve should not yet be considered validated indications for preferential TEA. Instead, they represent clinically relevant populations in whom the comparative effectiveness and net benefit of TEA versus peripheral techniques remain least certain and warrant prospective investigation.

The magnitude of TEA's analgesic advantage should also be interpreted cautiously. A recent systematic review and meta-analysis including four randomized trials and four observational studies (946 patients) found that epidural analgesia reduced pain compared with systemic opioids by a pooled mean difference of 0.8 NRS/VAS points at rest and 1.1 points during activity over postoperative days 1 to 3.[5] These differences were below or at the lower bound of the commonly cited 1- to 2-point minimal clinically important difference, and heterogeneity was substantial.[5] Thus, statistically stronger analgesia does not necessarily translate into a clinically meaningful advantage for the average VATS patient.

Taken together, contemporary evidence supports moving away from routine TEA for most patients undergoing VATS. Procedure-specific recommendations favor PVB or ESPB, and peripheral techniques generally offer a favorable balance between analgesia, recovery, and avoidance of neuraxial adverse effects.[1,4] At the same time, randomized evidence shows that TEA can provide stronger analgesia in some comparisons, and current trials do not adequately define whether particular higher-risk subgroups derive a clinically meaningful net benefit from TEA.[2,4]

Therefore, the clinical paradigm must not simply swing from uncritical routine use to indiscriminate complete abandonment.[1] TEA should no longer serve as the default technique for VATS; instead, its potential role should be considered

https://doi.org/10.1016/j.bjane.2026.844846
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

selectively and on an individual basis when the anticipated analgesic benefit may outweigh its disadvantages.

The appropriate conclusion is therefore not that TEA should remain preferred for a predefined high-risk group, but that its routine use should be abandoned without assuming that its role has been eliminated altogether. Future comparative studies should specifically address patients with chronic opioid use, difficult-to-control pain, and severely compromised respiratory reserve. Until such evidence is available, TEA is best regarded as a non-routine option whose potential value in carefully selected patients remains to be established.

## Data availability statement

No datasets were generated or analyzed for this Letter to the Editor; data sharing is therefore not applicable.

## Declaration of the use of AI and AI-assisted technologies

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) for language editing and improvement of clarity and readability. The authors reviewed and edited all AI-assisted content and take full responsibility for the accuracy, originality, and integrity of the final manuscript. AI and AI-assisted technologies were not used to generate, analyze, or interpret scientific data and select or interpret scientific evidence, nor to make scientific or clinical decisions, and were not used to formulate the clinical conclusions, or generate or verify references.

## Authors' contributions

Gabriel Trindade Avelar: Conceptualization; investigation; writing-original draft preparation.

André Versiani Caldeira Rocha: Conceptualization; investigation; writing-review & editing.

Regina Bicalho Gomes de Faria: Conceptualization; investigation; writing-review & editing.

Marina Ayres Delgado: Conceptualization; supervision; writing-review & editing; project administration.

## Funding

None.

## Conflicts of interest

The authors declare no conflicts of interest.

## Acknowledgments

None.

## Editor

Liana Azi

## References

1. Feray S, Lubach J, Joshi GP, Bonnet F, Van de Velde M. PROSPECT Working Group of the European Society of Regional Anaesthesia and Pain Therapy. PROSPECT guidelines for video-assisted thoracoscopic surgery: a systematic review and procedure-specific postoperative pain management recommendations. Anaesthesia. 2022;77:311–25.
2. Lai J, Situ D, Xie M, et al. Continuous paravertebral analgesia versus continuous epidural analgesia after video-assisted thoracoscopic lobectomy for lung cancer: a randomized controlled trial. Ann Thorac Cardiovasc Surg. 2021;27:297–303.
3. Wu Z, Wang Q, Wu C, et al. Paravertebral vs epidural anesthesia for video-assisted thoracoscopic surgery: a randomized trial. Ann Thorac Surg. 2023;116:1006–12.
4. Spaans LN, Dijkgraaf MGW, Susa D, et al. Intercostal or paravertebral block vs thoracic epidural in lung surgery: a randomized noninferiority trial. JAMA Surg. 2025;160:855–64.
5. Holm JH, Bak M, Brøchner AC. Epidural analgesia versus systemic opioids for postoperative pain management after VATS: a systematic review and meta-analysis. Acta Anaesthesiol Scand. 2025;69:e70122.

Gabriel Trindade Avelar,
André Versiani Caldeira Rocha,
Regina Bicalho Gomes de Faria,
Marina Ayres Delgado *

*Universidade Federal de Minas Gerais (UFMG), Hospital das Clínicas, Department of Anesthesiology, Belo Horizonte, MG, Brazil*

* Corresponding author.
*E-mail*: marina.ayres.delgado@gmail.com (M.A. Delgado).
Received 16 September 2026; accepted 6 October 2026
Available online 10 October 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# Smartphone-based immersive virtual reality for pediatric preoperative preparation: development and initial feasibility report

*Dear Editor,*

Preoperative anxiety affects 40% to 60% of children undergoing anesthesia,[1] representing an important challenge, as it may affect anesthetic management and clinical outcomes.[2] Unfamiliar surroundings, combined with perceived loss of control and uncertainty regarding medical procedures, may contribute to negative expectations and anticipatory distress.[2] Preoperative preparation and other non-pharmacological approaches reduce anxiety and improve the experience of patients and caregivers.[1] Among these, immersive Virtual Reality (VR) has emerged as a promising tool, with benefits attributed to two mechanisms: cognitive distraction and familiarization.[3]

In Brazil, accessible VR materials designed for pediatric preoperative preparation remain scarce. To bridge this gap, a multidisciplinary team developed a tailored immersive tool. This Letter reports the development and initial operational feasibility of the intervention within a prospective, single-arm exploratory pilot study conducted as part of the first author's master's research. The present report does not assess the effect of the intervention on preoperative anxiety, anesthetic outcomes, or postoperative recovery. The remaining outcomes of the pilot study will be reported in the master's dissertation and its derived publication, which will cite this Letter and will not duplicate the data presented here. During content planning, priority was given to determining the type and amount of information conveyed to the child, to promote comprehension without cognitive overload. The video reproduces the child's actual path through the facility, providing realistic and sequential exposure to surgical environments in age-appropriate, playful and reassuring tone. Safety was also considered throughout the design process. Although literature suggests limited evidence of harm from short, supervised exposure,[4] potential discomfort and adverse effects remain relevant. Therefore, length and pacing were optimized to balance clarity and tolerability. Excessively long sequences could lead to boredom and disrupt clinical workflow, whereas an accelerated pace could limit immersion and predispose users to cybersickness symptoms, such as dizziness, nausea, visual disturbances, and headache. A structured description of the intervention following the TIDieR framework[5] is provided as Supplementary Material 1.

Entitled *Viagem ao Centro Cirúrgico* [Journey to the Operating Room], the video combines real hospital footage with animated elements, structured as a spaceship voyage. A first-person corridor segment recorded with a drone simulates dynamic movement through the surgical center. Illustrations, such as fish swimming in scrub sinks and floating paper airplanes, were then incorporated to create a child-friendly scenario. A digital spaceship overlay was added, giving the illusion that the participant is a passenger traveling inside a cockpit, without compromising the realism of the original footage. The flight ends on the operating table, where audio narration invites passengers to disembark, and the perspective then shifts to a 360-degree live-action segment of the operating room. The final version is a passive, non-interactive video in Portuguese, 2 minutes and 51 seconds in duration, comprising a conventional first-person drone sequence followed by a monoscopic 360° equirectangular live-action segment, finalized in July 2024. It is permanently deposited in an open-access repository (doi: 10.5281/zenodo.21312318) and freely available on YouTube (www.youtube.com/watch?v=o3S77cTsly4), playable on regular smartphones and simple VR viewers (Fig. 1).

Between February and August 2025, the intervention was offered to children aged 5 to 12 years, ASA physical status I or II, in the admission area across two units of a private institution in Brazil. Exclusion criteria were neurodevelopmental disorders, history of seizures, visual or auditory impairment, regular use of psychoactive medication, and need for preoperative anxiolytics. Legal guardians provided written informed consent and children provided assent. The study was conducted in accordance with institutional ethical standards and approved by the local Research Ethics Committee (CAAE 59980922.9.0000.5128).

For each session, the child remained seated in an armchair while the video was displayed on a smartphone (Samsung Galaxy S21) inserted into a VR Shinecon headset and played through the YouTube application in VR mode. Audio was delivered through Bluetooth headphones integrated into the headset. All sessions occurred under direct anesthesiologist supervision. Adverse effects were monitored through continuous direct observation, with a prespecified criterion to interrupt playback upon any sign or spontaneous report of discomfort. No standardized symptom

https://doi.org/10.1016/j.bjane.2026.844825
0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).



**Figure 1** QR code providing access to *Viagem ao Centro Cirúrgico* [Journey to the Operating Room], the smartphone-based immersive virtual reality video developed for pediatric preoperative preparation. Scanning the code opens the video on YouTube. (https://www.youtube.com/watch?v=o3S77cTsly4).

instrument was applied. Application time was measured from initial headset positioning to the end of the first complete viewing. The setup used an existing smartphone and a VR headset purchased for R$ 180 (approximately US $ 33). The intervention was integrated into the routine preanesthetic contact in the admission area, requiring no additional dedicated staff.

Of 23 children assessed for eligibility, three were not enrolled: one had received premedication, one was receiving regular psychoactive medication and one declined participation. Therefore, 20 were enrolled, corresponding to a recruitment rate of 87% (20/23). Mean age was 8.2 ± 2.0 years (range, 5.2–12.4), 12 (60%) were boys, and 15 (75%) were classified as ASAI. Scheduled procedures were urological (30%), general (30%), otorhinolaryngological (25%), and orthopedic (15%). All 20 children completed the full viewing (100% intervention completion). There were no technical failures, interruptions, requests to remove the headset, and no observed or spontaneously reported adverse effects. Mean application time was 4 minutes 16 seconds (SD = 13s; range, 4 min 5s to 4 min 51s). No acceptability claims are made in this Letter, as no validated standardized acceptability questionnaire was administered.

The main contribution of this initiative is the materialization of a functional, freely accessible VR tool specifically developed for pediatric preoperative preparation in the Brazilian context. In this small single-arm sample, all enrolled children completed the intervention, and the mean application time was 4 minutes 16 seconds. These descriptive observations should not be interpreted as evidence of safety, efficacy, or broad acceptability. Although the per-unit equipment cost was low and the video is freely available, a relevant feature for access in the Brazilian context, no formal cost-effectiveness or implementation analysis was conducted, and these questions, along with anxiety-related outcomes, require evaluation in further studies. While evidence supports VR-based preoperative preparation in general,[3] the evidence generated by this specific intervention is currently limited to the development process and initial feasibility experience reported here.

## Data availability statement

No new data were generated or analyzed in this article. Data sharing is not applicable to this article.

## Ethics approval

The study was approved by the local Ethics Committee (CAAE: 59980922.9.0000.5128).

## Study registration

Not applicable.

## Permissions

Hospital footage was recorded with written institutional authorization covering academic use and publication. All professionals involved in the production (filming, illustration, editing, and narration) provided written assignment of use rights covering academic publication and open-access deposit. The soundtrack is royalty-free.

## Authors' contributions

Andreia Cardoso contributed to study conception and design, data collection, interpretation of findings, and drafting of the manuscript.

Durval Campos Kraychete contributed to project supervision, intellectual development of the study, critical revision of the manuscript, and final approval of the submitted version.

## Conflicts of interest

The first author and the audiovisual director hold a pending trademark application for *Viagem ao Centro Cirúrgico* at the Brazilian National Institute of Industrial Property (INPI). The intervention is distributed free of charge, and no commercial exploitation exists or is planned. The authors declare no other conflicts of interest.

## Funding

Video production was funded by the first author with personal resources. No external or commercial funding was received. No equipment, software, professional services, or other resources were provided as in-kind support.

## Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjane.2026.844825.

## Editor

Liana Azi

## References


1. Capurso M, Ragni B. Psycho-educational preparation of children for anaesthesia: A review of intervention methods. Patient Educ Couns. 2016;99:173—85.
2. Adler AC, Leung S, Lee BH, Dubow SR. Preparing Your Pediatric Patients and Their Families for the Operating Room: Reducing Fear of the Unknown. Pediatr Rev. 2018;39:13—24.
3. Koo CH, Park JW, Ryu JH, Han SH. The Effect of Virtual Reality on Preoperative Anxiety: A Meta-Analysis of Randomized Controlled Trials. J Clin Med. 2020;9:3151.
4. Bexson C, Oldham G, Wray J. Safety of virtual reality use in children: a systematic review. Eur J Pediatr. 2024;183:2071—90.
5. 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.



Andreia Cardoso [a,b,*], Durval Campos Kraychete [c]

[a] *Anesthesiologist Mater Dei Hospital, Salvador, BA, Brazil*
[b] *Universidade Federal da Bahia, Salvador, BA, Brazil*
[c] *Universidade Federal da Bahia (UFBA), Department of Anesthesiology and Surgery, Salvador, BA, Brazil*

[*] Corresponding author.
*E-mail:* santoscardoso.andreia@gmail.com (A. Cardoso).


Received 29 May 2026; accepted 10 August 2026
Available online 12 September 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# Opioid-free versus opioid-containing labor epidural analgesia: interpreting non-inferiority claims in a retrospective cohort study

*Dear Editor,*

I read with interest the 2026 article titled "The impact of opioidfree labor epidural analgesia maternal and infant outcomes: a retrospective cohort study" by Brahmbhatt et al., examining the impact of Opioid-Free Labor Epidural Analgesia (OFLEA) versus conventional Opioid Epidural Labor Analgesia (OLEA) on maternal and neonatal outcomes.[1] The authors address a timely and clinically important question. While the findings are promising, there appear to be methodological concerns relevant to interpretation of the results.

The methodology does not adequately support a formal non-inferiority claim. No non-inferiority margin was prespecified, and the post-hoc analysis across three candidate thresholds (0.75, 1.0, and 1.5 pain units) is more appropriately characterized as exploratory rather than confirmatory. At the midpoint margin of 1.0 pain units, power is reported at 19.7%, limiting the strength of a non-inferiority conclusion. In the absence of an a priori margin, the results are more appropriately understood as exploratory comparative effectiveness evidence.

Future prospective comparative studies should define a non-inferiority margin before data collection begins. Where such studies take the form of randomized controlled trials, adherence to the CONSORT extension for non-inferiority and equivalence designs would further support transparent and interpretable reporting.[2]

The original report describes 1:1 matching by maternal age; the enrolled groups nonetheless differed substantially in size (OLEA: n = 1,093 vs. OFLEA: n = 330). While unequal group sizes alone do not establish baseline imbalance or non-random allocation, limiting the matching strategy to a single covariate is a meaningful constraint. Group-specific baseline characteristics and post-matching balance diagnostics were not clearly reported, leaving it uncertain whether important obstetric and clinical factors were adequately balanced between groups. Future studies should employ propensity score matching or multivariable adjustment for a broader set of clinical covariates and provide transparent reporting of the matching strategy and the balance achieved.

A further consideration concerns the pharmacological composition of the two epidural regimens. The opioid-free group received 0.2% ropivacaine, whereas the opioid-containing group received 0.1% ropivacaine plus fentanyl. The two strategies therefore differed in both opioid content and local anesthetic concentration. Observed differences in pain outcomes cannot be attributed specifically to opioid avoidance alone; the comparison is more accurately understood as one between two pharmacologically distinct epidural treatment strategies, rather than a direct opioid versus no-opioid comparison.

Additionally, the maximum pain score showed a statistically significant difference in the unadjusted Wilcoxon rank-sum test (p = 0.035); however, the corresponding linear regression model adjusted for age, ethnicity, and fever yielded a non-significant result (p = 0.088, 95% Confidence Interval crossing zero). The adjusted estimate is imprecise and compatible with no meaningful between-group difference on this measure. This outcome should be clearly distinguished from those endpoints that retained statistical significance after covariate adjustment. Authors are encouraged to differentiate outcomes that survive adjustment from those that do not, and to avoid characterizing analgesic benefit as established across all pain metrics when adjusted estimates for key endpoints do not support that conclusion.[3]

These limitations do not lessen the importance of the research question, but they do affect the strength of the conclusions. The observed outcomes are encouraging and warrant prospective validation. However, the field would be best served by prospective, adequately powered, multicenter randomized controlled trials with a pre-registered non-inferiority margin, robust covariate adjustment, and standardized outcome reporting. Such studies would provide the evidentiary foundation necessary to advocate for OFLEA as a standard-of-care alternative.

## Data availability statement

No new data were created or analyzed in this article. Data sharing is not applicable to this article.

## Conflicts of interest

The authors declare no conflicts of interest.

https://doi.org/10.1016/j.bjane.2026.844816

0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## AI assistance disclosure

During the preparation of this manuscript, the author used Claude (Anthropic) solely for language editing and improving clarity. The author reviewed and edited the output as needed and takes full responsibility for the content of the publication.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Editor

Liana Azi

## References

1. Brahmbhatt KS, Reddy AP, Bonilla HAA, et al. The impact of opioid-free labor epidural analgesia maternal and infant outcomes: a retrospective cohort study. Braz J Anesthesiol. 2026;76:844750.
2. Piaggio G, Elbourne DR, Pocock SJ, Evans SJW, Altman DG, Group Consort. Reporting of noninferiority and equivalence randomized trials: extension of the CONSORT 2010 statement. JAMA. 2012;308:2594—604.
3. Schuster NA, Rijnhart JJM, Bosman LC, Twisk JWR, Klausch T, Heymans MW. Misspecification of confounder-exposure and confounder-outcome associations leads to bias in effect estimates. BMC Med Res Methodol. 2023;23:11.

Pourya Baharlooei *

*Tehran University of Medical Sciences, School of Medicine, Tehran, Iran*

* Corresponding author.
*E-mail:* Pourbah2000@gmail.com
Received 31 May 2026; accepted 18 July 2026
Available online 14 August 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# ESPB versus thoracic paravertebral block in pediatric nephrolithotomy: analytic population and inference after a negative superiority trial

*Dear Editor,*

We read with interest the randomized, assessor-blind trial by Nabil et al. comparing erector spinae plane block (ESPB) with thoracic paravertebral block (PVB) in children undergoing percutaneous nephrolithotomy (PCNL). The study found no statistically significant between-group differences in time to first rescue nalbuphine or 24-hour rescue analgesic requirements, while ESPB took less time to perform.[1] This is a clinically relevant question in pediatric urologic surgery.

The interpretation should remain aligned with the superiority design. Failure to demonstrate statistically significant superiority does not establish equivalence or non-inferiority, which require a prespecified clinical margin and an appropriate design and analysis. The confidence intervals reported for the primary outcome describe each group separately; reporting a between-group effect estimate with its confidence interval would better convey the precision of the comparison. Thus, the shorter performance time may inform procedural feasibility, but the absence of a significant analgesic difference should not be interpreted as proof of therapeutic comparability.

Clarification of the analytic population would further aid interpretation. One child allocated to ESPB was excluded after an intraoperative change in the surgical procedure, leaving 27 and 28 children in the respective analyses.[1] If relevant outcome data are available, an intention-to-treat sensitivity analysis retaining the original allocation would help assess the robustness of the findings. If this is not feasible, the authors could explain what data were available and why the changed procedure precluded that analysis. Such clarification would allow readers to judge the implications of the exclusion without assuming that it materially altered the result.

The study defined block failure as requiring two doses of nalbuphine in the post-anesthesia care unit (PACU), and one child in each group received those doses.[1] It would be helpful to state explicitly whether these children remained in the analyses, whether their first PACU rescue doses were counted as primary-outcome events, and whether both doses were included in the 24-hour rescue totals. Because block failure was defined using medication that also contributes to the efficacy endpoints, its handling can affect estimates of analgesic duration and rescue requirements. Reporting these details would distinguish outcomes following assignment to a block from outcomes restricted to successful blocks.

We also request clarification of the fentanyl dose cited as 1 $mg.kg^{-1}$ in the version reviewed, which may represent a unit or typesetting discrepancy. Confirming the intended dose and ensuring consistent units across article formats would improve reproducibility.[1]

The wider evidence should also be interpreted according to its comparators and settings. A pediatric meta-analysis found no statistically significant differences between ESPB and caudal epidural block for selected analgesic outcomes, but this does not establish equivalence to thoracic PVB.[2] A network meta-analysis in PCNL similarly found no significant differences among ESPB, PVB, and quadratus lumborum block for postoperative analgesic efficacy.[3] Such findings cannot replace a pediatric equivalence or non-inferiority trial. In adults undergoing major oncological breast surgery, ESPB did not meet the non-inferiority criterion against PVB and provided less reliable dermatomal coverage.[4] Although that result is not directly transferable to pediatric PCNL, it reinforces the need for procedure-specific inference. Transparent reporting of analysis populations, protocol deviations, and effect estimates would support interpretation, consistent with CONSORT 2025.[5]

We congratulate the investigators on addressing an important pediatric analgesia question. Clarifying the analysis population and rescue events, while interpreting the findings within the superiority design, would strengthen the clinical value of this trial without implying equivalence or non-inferiority.

## Data availability statement

No new data were generated or analyzed in this article. Data sharing is not applicable to this article.

## Funding

No specific funding was received for this work.

https://doi.org/10.1016/j.bjane.2026.844849
0104-0014/© 2026 Sociedade Brasileira de Anestesiologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Authors' contributions

All authors contributed to the conception, drafting, critical revision, and final approval of this letter.

## Declaration of generative AI and AI-assisted technologies in the writing process

ChatGPT (OpenAI) was used to assist with language editing, manuscript polishing, and preparation of revisions responding to the editorial comments. The authors are responsible for the scientific interpretation, reference verification, and final approval of the manuscript.

## Conflicts of interest

The authors declare no conflicts of interest.

## Editor

Liana Azi

## References

1. Nabil F, Mandour AM, Abdelgawad AM, et al. Erector spinae plane block versus thoracic paravertebral block in pediatric patients undergoing percutaneous nephrolithotomy: a prospective randomized clinical trial. Braz J Anesthesiol. 2026;76:844735.
2. Masiero BB, Cavalcante D, Akbarpoor F, et al. Erector spinae plane block versus caudal epidural block in pediatric surgery: a systematic review and meta-analysis of randomized clinical trials. Braz J Anesthesiol. 2025;75:844640.
3. Long K, Zhou C, Liang J, Tang X, Li Z, Chen Q. Effect of different regional anaesthesia techniques on postoperative analgesia following percutaneous nephrolithotomy: a systematic review and network meta-analysis. Indian J Anaesth. 2025;69:12—22.
4. Raft J, Dureau S, Fuzier R, et al. Erector spinae plane block versus paravertebral block for major oncological breast surgery: a multicentre randomised controlled trial. Br J Anaesth. 2025;135:772—8.
5. Hopewell S, Chan AW, Collins GS, et al. CONSORT 2025 statement: updated guideline for reporting randomised trials. BMJ. 2025;389:e081123.

Swarupanjali Padhi [a], Prashant Ramdas Kokiwar [b,*], Janvi Patel [c], Ankita Kalra [d,e]

[a] *Noida Institute of Engineering & Technology (Pharmacy Institute), Department of Pharmaceutics, Greater Noida, India*
[b] *Malla Reddy Institute of Medical Sciences, Malla Reddy Vishwavidyapeeth, Department of Community Medicine, Telangana, India*
[c] *Dr. D. Y. Patil Vidyapeeth (Deemed-to-be-University), Hospital and Research Centre, Dr. D. Y. Patil Medical College, Maharashtra, India*
[d] *Graphic Era Hill University, Faculty of Pharmaceutical Sciences, Dehradun, India*
[e] *Graphic Era Deemed University, Centre for Promotion of Research, Dehradun, India*

[*] Corresponding author.
*E-mail:* kokiwar@mrvv.edu.in (P.R. Kokiwar).
Received 3 June 2026; accepted 11 October 2026
Available online 15 October 2026

Sociedade Brasileira de Anestesiologia

Brazilian Journal of ANESTHESIOLOGY



LETTER TO THE EDITOR

# Authors' reply to comments on erector spinae plane block versus thoracic paravertebral block in pediatric percutaneous nephrolithotomy

*Dear Editor,*

We thank Dr. Peng and Dr. Li,[1] and Dr. Padhi and colleagues[2] for their thoughtful Letters to the Editor regarding our article on erector spinae plane block versus thoracic paravertebral block in pediatric patients undergoing percutaneous nephrolithotomy.[3] Their comments provide an important opportunity to clarify the interpretation of our findings, particularly regarding trial design, analytic population, procedural time, and safety considerations.

First, our study was designed and powered for the primary outcome as a superiority trial, not as an equivalence or non-inferiority trial. Therefore, we agree that the absence of statistically significant difference in the primary outcome should not be interpreted as evidence of equivalence or non-inferiority between ESPB and thoracic PVB. With regard to the conclusion statement "could be considered as an easy alternative", we intended the practical and clinical perspective related to the comparable analgesic effects rather than claiming therapeutic equivalence or superiority.

With regard to the methodological concerns raised by Padhi et al., because the excluded patient did not undergo the planned PCNL procedure after randomization, an intention-to-treat sensitivity analysis including this participant was not considered clinically meaningful. All randomized patients who received the allocated intervention and completed the planned PCNL procedure were included in the final analysis. The intended and actually administered fentanyl dose was 1$\mu$g.kg$^{-1}$. The previous notation as 1 mg.kg$^{-1}$ was a unit error.

We appreciate the important point of Peng and Li that a 40-second difference in the time to conduct the blocks is of negligible clinical benefit despite being statistically significant and we had already stated in the discussion that this result, as a secondary outcome, should be interpreted cautiously. However, it is worth re-emphasizing that all the blocks were conducted by a single anesthesiologist who has a substantial experience in regional anesthesia for pediatric patients. Although this was not specifically evaluated in the current study, the shorter performance time of ESPB may be more relevant in settings where clinicians are less experienced with thoracic PVB or where obtaining the paravertebral sono-anatomic view is more challenging as greater caution is required to avoid plueral puncture.

With regard to the concern raised by Padhi et al. we agree that the current evidence comparing erector spinae plane block and paravertebral block remains inconsistent across different surgical settings. Therefore, we intended to design our methodology to be limited to one surgical procedure in the pediatric population. Hence, our conclusion is confined to this context.

Regarding safety, we agree with Peng and Li that the safety profile in the pediatric population has more robust evidence for PVB since it is an older technique, and we also agree that the current study was not powered to detect the complications as we already mentioned that in the manuscript as a limitation. As a relatively new regional analgesic technique, promising data have started to emerge supporting the safety of ESPB in the pediatric population.[4] However, larger studies are still required to evaluate rare adverse events.

In conclusion, we appreciate the observations raised by the correspondents. We agree that our findings should be interpreted in the context of a superiority trial, and we recommend larger powered studies to evaluate the safety of ESPB in pediatric patients.

## Data availability statement

No new data were generated or analyzed in this article. Data sharing is not applicable to this article.

## Conflicts of interest

The authors declare no conflicts of interest.

## Declaration of the use of AI

During the preparation of this work the authors used ChatGPT (OpenAI; GPT-5.5 version) for editing. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

https://doi.org/10.1016/j.bjane.2026.844818

0104-0014/© 2026 Published by Elsevier España, S.L.U. on behalf of Sociedade Brasileira de Anestesiologia. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Authors' contributions

Fatma Nabil: Conceptualization; Methodology; Writing the original draft.

Ahmed M. Mandour: Methodology; Data curation.

Amr M. Abdelgawad: Methodology; Investigation.

Deiaaeldin M. Tamer: Methodology; Data Curation.

Ahmed A. Shahat: Investigation; Software; Writing the original draft.

Mohamed Anwar: Investigation; Writing-review & editing; Supervision.

Hany M. Osman: Formal analysis; Writing-review & editing; Project administration.

## Funding

No specific funding was received.

## Editor

Liana Azi

## References

1. Peng Y, Li Y. Comment on: Erector spinae plane block versus thoracic paravertebral block in pediatric patients undergoing percutaneous nephrolithotomy: a prospective randomized clinical trial. Braz J Anesthesiol. 2026:844788.
2. Padhi S, Kokiwar PR, Patel J, Kalra A. ESPB versus thoracic paravertebral block in pediatric nephrolithotomy: analytic population and inference after a negative superiority trial. Braz J Anesthesiol. 2026;76:844849.
3. Nabil F, Mandour AM, Abdelgawad AM, et al. Erector spinae plane block versus thoracic paravertebral block in pediatric patients undergoing percutaneous nephrolithotomy: a prospective randomized clinical trial. Braz J Anesthesiol. 2026;76:844735.
4. Lucente M, Ragonesi G, Sanguigni M, et al. Erector spinae plane block in children: a narrative review. Korean J Anesthesiol. 2022;75:473–86.

Fatma Nabil [a], Ahmed M. Mandour [a], Amr M. Abdelgawad [b], Deiaaeldin M. Tamer [c], Ahmed A. Shahat [b], Mohamed Anwar [a], Hany M. Osman [a,*]

[a] *Assiut University, Faculty of Medicine, Department of Anesthesia, Intensive Care, and Pain Management, Assiut, Egypt*
[b] *Assiut University, Faculty of Medicine, Department of Urology, Assiut, Egypt*
[c] *Assiut University, Faculty of Medicine, Pediatric Department, Assiut, Egypt*

[*] Corresponding author.
*E-mail:* hanymeo@aun.edu.eg (H.M. Osman).
Received 29 May 2026; accepted 31 July 2026
Available online 13 August 2026

# BJAN

## Brazilian Journal of Anesthesiology


Contact us

editor.bjan@sbahq.org www.bjan-sba.org +55 21 979 770 024