

COMMENT OPEN



Implementing quality management systems to close the AI translation gap and facilitate safe, ethical, and effective health AI solutions

Shauna M. Overgaard¹✉, Megan G. Graham¹, Tracey Brereton¹, Michael J. Pencina², John D. Halamka¹, David E. Vidal¹ and Nicoleta J. Economou-Zavlanos¹

The integration of Quality Management System (QMS) principles into the life cycle of development, deployment, and utilization of machine learning (ML) and artificial intelligence (AI) technologies within healthcare settings holds the potential to close the AI translation gap by establishing a robust framework that accelerates the safe, ethical, and effective delivery of AI/ML in day-to-day patient care. Healthcare organizations (HCOs) can implement these principles effectively by embracing an enterprise QMS analogous to those in regulated industries. By establishing a QMS explicitly tailored to health AI technologies, HCOs can comply with evolving regulations and minimize redundancy and rework while aligning their internal governance practices with their steadfast commitment to scientific rigor and medical excellence.

npj Digital Medicine (2023)6:218; <https://doi.org/10.1038/s41746-023-00968-8>

QMS AS A FRAMEWORK FOR HEALTH AI

The advancements in healthcare software, encompassing artificial intelligence, machine learning (AI/ML), and Software as a Medical Device (SaMD), have brought about opportunities for transformative changes in clinical workflows and patient care to effectively meet patient and clinician needs. However, healthcare software exists within a complex regulatory and technical landscape¹. The need for more readiness among healthcare organizations (HCOs) magnifies the disparity in translating research into effective predictive clinical decision support interventions. Without a collaborative enterprise approach, the intricate nature of this system delays the translation of AI solutions into clinical practice. Characterized by the continuous evolution and maturation of AI/ML capabilities, such as large language models (LLMs), this ecosystem escalates the demand for software-driven clinical solutions and a regulatory framework that must effectively adapt to govern the distinctive nature of in-house-built and procured software². The growing engagement of HCOs in AI calls for alignment among diverse stakeholders, encompassing industry, academic institutions, and the medical community. This alignment should focus on harmonizing assurance standards for health AI technologies, but also practices and infrastructure to enable HCOs to develop and deploy AI solutions meeting rigorous medical-grade standards while ensuring accountability across all involved parties. While regulatory authorities, AI coalitions, medical device manufacturers, and the medical informatics community have acknowledged the current gap not only in common standards but also in the maturity of HCOs to develop and/or deploy health AI, a primary concern for HCOs remains unresolved: "How might our enterprise establish a coordinated, robust strategy that ensures the safe, effective, and ethically sound delivery of AI/ML in day-to-day patient care?"^{3–7}.

We propose using the Quality Management System (QMS) framework to offer HCOs a consistent and adaptable structure to translate research-based health AI technology into clinical practice systematically and transparently. QMS is a structured framework

that documents processes, procedures, and responsibilities to achieve quality policies and objectives. The QMS framework effectively manages evolving regulatory requirements, promotes continuous improvement, and ensures adherence to cutting-edge standards over the life cycle of the design, development, deployment, and maintenance of regulated healthcare software⁸. QMS's are often certified to external standards (e.g., ISO 13485), thus demonstrating organizational commitment to quality, continuous improvement, and regulatory compliance. Aligning standards with risk-based approaches facilitates the least burdensome path for an HCO to meet regulatory requirements and maintain compliance⁹. Thus, the streamlined incorporation of these regulatory requirements into business processes via the QMS assures enduring safety, effectiveness, ethicality, regulatory compliance, and alignment with organizational and user needs as AI-enabled methodologies, such as LLMs, evolve¹⁰.

We aim to elucidate the primary components of a QMS (Fig. 1)^{8,9,11}, namely People & Culture, Process & Data, and Validated Technology, as the impetus for HCO's strategic efforts to integrate research rigor and clinical excellence into a cohesive system and close the AI translation gap.

ESTABLISHING A PROACTIVE CULTURE OF QUALITY

In HCOs, AI/ML technologies are often initiated as siloed research or quality improvement initiatives. However, when these AI technologies demonstrate potential for implementation in patient care, development teams may encounter substantial challenges and backtracking to meet the rigorous quality and regulatory requirements^{12,13}. Similarly, HCO governance and leadership may possess a strong foundation in scientific rigor and clinical studies; however, without targeted qualifications and training, they may find themselves unprepared to offer institutional support, regulatory oversight, or mobilize teams toward interdisciplinary scientific validation of AI/ML-enabled technologies required for regulatory submissions and deployment of SaMD. Consequently,

¹Mayo Clinic, Rochester, MN, USA. ²Duke University, Durham, NC, USA. ✉email: Overgaard.Shauna@mayo.edu

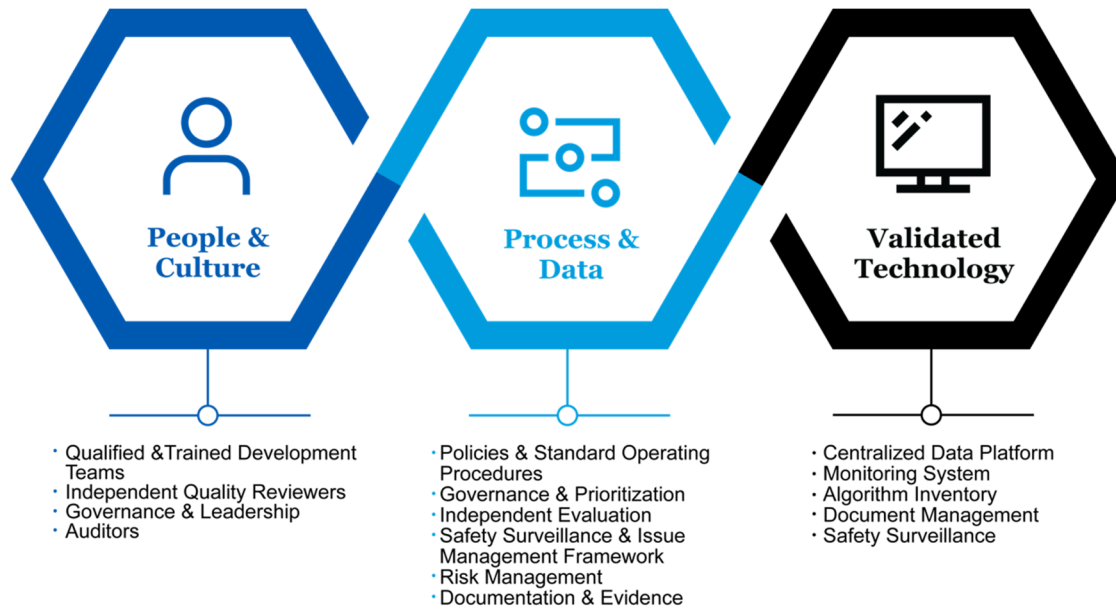


Fig. 1 Primary components of a Quality Management System (QMS).

Term	Definition
End-User Acceptance Testing	Testing that demonstrates that the product satisfies end-user defined requirements ¹² .
Corrective and Preventive Action (CAPA)	Changes to products and/or the quality management system to eliminate the cause of or prevent the occurrence of issues that affect the quality of products or processes ⁹ .
Design Controls	Interrelated set of practices and procedures that are incorporated into the design and development process, i.e., a system of checks and balances ^{8,12} .
Design Verification and Validation	Testing or evaluation activities that provide objective evidence that the product satisfies its intended purpose (or use) and specifications ⁹ .
Installation Qualification	Testing or evaluation activities that demonstrate the software is installed correctly in its intended environment ¹² .
Intended Purpose (Intended Use)	The objective intent and general purpose of a product ²⁹ .
Operational Qualification	Testing or evaluation activities that demonstrate the software functions as intended ¹² .
Performance Qualification	Testing or evaluation activities that demonstrate the software satisfies user and business requirements ¹² .
Records	Documents or other artifacts that demonstrate that a process was completed ⁸ .
Software Development Life Cycle (SDLC)	A framework of processes, activities, and tasks phases to design, develop, and maintain software ³⁰ .
Monitoring and Surveillance	Tracking and analyzing the performance of and changes to products after deployment to identify quality or safety issues that may necessitate corrective or preventive action ⁹ .
Version Control	Activity related to software configuration management intended to track and manage changes to software. Version control is used during all stages of the software development life cycle to prevent mix-ups ^{8,30} .

the unpreparedness of HCOs exacerbates the translation gap between research activities and the practical implementation of clinical solutions¹⁴. The absence of a systematic approach to ensuring the effectiveness of practices and perpetuating them throughout the organization can lead to operational inefficiencies or harm. Thus, HCOs must first contend with a culture shift when faced with quality control rigor inherent to industry-aligned software development and deployment, specifically design controls, version control, installation qualification, operational qualification, performance qualification, that primarily focuses on end-user acceptance testing and the product meeting its intended purpose (improving clinical outcomes or processes compared to the standard of care or the current state), and the traceability and auditability of proof records (Table 1).

Consider that even in cases where a regulatory submission is not within the scope, it remains imperative to adhere to practices

encompassing ethical and quality principles. Examples of such principles identified by the Coalition for Health AI and the National Institute for Standards and Technology (NIST) include effectiveness, safety, fairness, equity, accountability, transparency, privacy, and security^{3,7,15–20}. It is also feasible that the AI/ML technology could transition from a non-regulated state to a regulated one due to updated regulations or an expanded scope. In that case, a proactive approach to streamlining the conversion from a non-regulatory to a regulatory standard should address the delicate balance of meeting baseline requirements while maintaining a least-burdensome transition to regulatory compliance.

As utilized by the FDA for regulating SaMD, a proactive culture of quality recognizes the same practices familiar to research scientists well-versed in informatics, translational science, and AI/ML framework development. For example, the FDA has published good machine learning practices (GMLP)²¹ that enumerate its

expectations across the entire AI/ML life cycle grounded in emerging AI/ML science. The FDA's regulatory framework allows for a stepwise product realization approach that HCOs can follow to augment this culture shift. This stepwise approach implements ethical and quality principles by design into the AI product lifecycle, fostering downstream compliance while allowing development teams to innovate and continuously improve and refine their products. Using this approach allows for freedom to iterate at early research stages. As the product evolves, the team is prepared for the next stage, where prospectively planned development, risk management, and industry-standard design controls are initiated. At this stage, the model becomes a product, incorporating all the software and functionality needed for the model to work as intended in its clinical setting. QMS procedures outline practices, and the records generated during this stage create the level of evidence expected by industry and regulators^{22,23}. HCOs may either maintain dedicated quality teams responsible for conducting testing or employ alternative structures designed to carry out independent reviews and audits.

Upon deployment, QMS rigor increases again to account for standardized post-deployment monitoring and change management practices embedded in QMS procedures (Fig. 2). By increasing formal QMS consistency as the AI/ML gets closer to clinical deployment, the QMS can minimize disruption to current research practices and embolden HCO scientists with a clear pathway as they continue to prove their software safe, effective, and ethical for clinical deployment.

ESTABLISHING RISK-BASED DESIGN, DEVELOPMENT, AND MONITORING

The medical device industry has utilized a risk-based infrastructure for years to support a least burdensome approach to designing, developing, and deploying healthcare technologies^{9,24}. This approach systematically enables HCOs to proactively focus resources on key areas of concern, such as safety, equity, and data privacy, to prevent errors and malfunctions and promote a culture of accountability and continuous improvement.

Risk-based practices have been extended to healthcare AI/ML in not only the medical device domain, such as with AAMI's Technical Information Report 34971²⁵, but more broadly in emerging frameworks such as the NIST AI Risk Management Framework³, the Whitehouse Blueprint for an AI Bill of Rights⁵, the Coalition for Health AI Blueprint for Trustworthy AI Implementation Guidance and Assurance for Healthcare²⁶, and the Health AI Partnership Key Decision Points^{27,28}. Risk management is grounded in the intended use and informed by a prospective risk management plan. It follows the process of identification, enumeration, mitigation, and monitoring (Fig. 3) to analyze and classify potential sources of harm (known as hazards) caused by the healthcare software or its impact on the clinical workflow. As

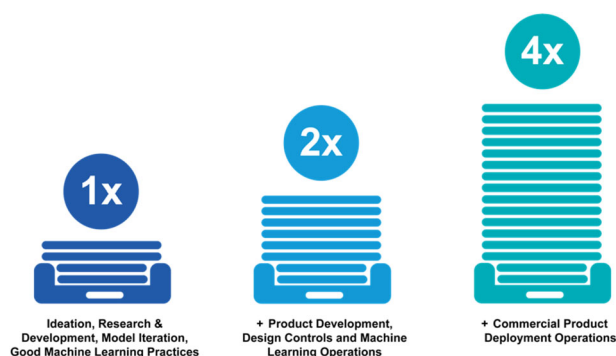


Fig. 2 Staged process for applying increasing regulatory rigor throughout product realization.

the healthcare software is designed and developed, features or attributes that reduce or minimize the risk (known as mitigations) are included in the product design; for example, incorporating features that improve the user experience or providing user training or documentation to clarify how the software should or should not be used. As risks and potential issues are anticipated for the health software's implementation, a risk management plan is put in place, a document articulating how safety, bias, and other anticipated risks will be identified and resolved. Risks continue to be monitored, reported, and reviewed after the software is deployed to ensure the software remains safe for use. Systematic feedback, monitoring, and corrective & preventive action (CAPA) frameworks are key to identifying and triaging issues, escalating issues to relevant accountable departments of the organization depending on their severity, performing root-cause analysis, and continuously controlling risks and improving the AI technology.

Risk-based practices formalized and implemented within a QMS will systematically identify risks associated with an AI solution, document mitigation strategies, and offer a framework for objective testing and auditing of individual technology components. Further, such technologies can be informed by AI/ML and software life cycle best practices to address common issues within phases of the AI lifecycle. This allows for capturing performance metrics across various levels of rigor and data transparency in requirements, version, and design controls. These insights from initial testing can then support the calibration and maintenance of AI solutions during deployment, guided by a multidisciplinary governance system to proactively mitigate future risks²⁶. Moreover, establishing a change management plan and access controls can eliminate business continuity risks, providing transparency into responsible parties and outlining the risks of any given change. Back-up (downtime) processes are in place in the event that risk cannot be managed, and the technology needs to be turned off. Effectively, a risk-based approach ensures the proper rigor and controls are in place at the right time throughout the product life cycle.

ESTABLISHING A COMPLIANCE-FACILITATING INFRASTRUCTURE

The regulations for healthcare software are evolving. Software may or may not be regulated based on its intended use or by changes to regulatory agency enforcement. A QMS that facilitates compliance with applicable legal and regulatory requirements enables HCOs to design, implement, and deploy healthcare software to clinical practice while minimizing overall operational risk. A QMS fosters compliance to internal (e.g., institutional review board) and external (e.g., federal, and local regulatory) bodies by standardizing multi-faceted stakeholder responsibilities with its governance, allowing auditability and traceability through the appropriate evidence and documentation, maintaining an inventory of AI technologies developed and deployed, and hosting



Fig. 3 Example QMS risk management plan and risk assessment phases. Risks are identified, assessed and analyzed, mitigated and controlled, and continuously monitored. Reporting is performed at pre-defined intervals.

infrastructure that will allow document management and monitoring within the deployment platform.

A QMS involves establishing policies and standard operating procedures that outline the process for governance and prioritization, development, independent evaluation, maintenance and monitoring, issue reporting and safety surveillance. Procedures outline the roles and responsibilities of stakeholders such as design and testing responsibilities of the champion stakeholder representing the end-users in the product development process. Procedures should also articulate training and/or qualification requirements for the stakeholders participating in AI technology development teams as safety and other risks can be eliminated with stakeholder education. Procedures also outline the systems and communication channels available to the community impacted by the deployed algorithmic tools ensuring their compliance. Communication in a regulated QMS is bidirectional, where issues, safety surveillance and outcome data are gathered via real-time monitoring and tightly integrated with the risk management and patient safety operations of a given healthcare system to determine the behavior and impact on patients and their healthcare delivery.

Establishing an innovation infrastructure that facilitates compliance requires governance and leadership support to create a communicated mandate that all algorithmic tool-related activities impacting patient health comply with quality and ethical standards. For example, the governing body may have direct integration with existing IRB processes to ensure ethical conduct. With proper governance, algorithm inventory, and transparency, HCOs can begin to implement tools, testing, and monitoring capabilities into their QMS to reduce the burden and achieve safe, effective, ethical ML/AI at scale. Implementing QMS involves formal documentation encompassing quality, ethical principles, and processes, ensuring transparency and traceability to regulatory requirements.

CONCLUSION

HCOs can utilize a QMS framework to accelerate the translation of AI from research to clinical practice. A proactive quality culture, risk-based framework for design, development, monitoring, and compliance-oriented infrastructure enables continuous ethical review, ensuring the effectiveness, safety, and equity of AI/ML technologies and meet regulatory requirements. Implementing a QMS requires adaptability, customization, and interdisciplinary collaboration, fostering awareness, education, and organizational growth. Drawing on regulatory precedents and incorporating insights from expert stakeholders, the QMS framework enables HCOs to prioritize patient needs and foster trust in adopting innovative AI technologies, including those enabled by LLMs.

Received: 14 July 2023; Accepted: 15 November 2023;
Published online: 25 November 2023

REFERENCES

- Yaghi, M. & Jacobo, N. *California AG Sends Letter to Hospital CEOs on Use of Artificial Intelligence*. <https://www.regulatoryoversight.com/2022/10/california-ag-sends-letter-to-hospital-ceos-on-use-of-artificial-intelligence/> (2022).
- U.S. Food and Drug Administration. *Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML) Enabled Device Software Functions*. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial-utm_medium=email&utm_source=govdelivery (2023).
- Tabassi, E. *Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology* <https://doi.org/10.6028/NIST.AI.100-1> (2023).
- U.S. Food and Drug Administration. *Artificial Intelligence and Machine Learning in Software as a Medical Device*. <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device> (2021).
- The White House. *Blueprint for an AI Bill of Rights*. <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (2022).
- International Telecommunication Union (ITU). *Focus Group on "Artificial Intelligence for Health"*. <https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/default.aspx> (2023).
- The MITRE Corporation. *Coalition for Health AI*. <https://coalitionforhealthai.org/> (2022).
- U.S. Food and Drug Administration. *21 CFR PART 820*. <https://www.govinfo.gov/content/pkg/CFR-2012-title21-vol8/pdf/CFR-2012-title21-vol8-part820.pdf> (2023).
- International Standards Organization. *ISO 13485: Medical Devices—Quality Management Systems—Requirements for Regulatory Purposes*. <https://www.iso.org/standard/59752.html> (2016).
- Jiang, L. Y. et al. Health system-scale language models are all-purpose prediction engines. *Nature* **619**, 357–362 (2023).
- International Medical Device Regulators Forum. *Software as a Medical Device (SaMD): Application of Quality Management System* <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-151002-samd-qms.pdf> (2015).
- U.S. Food and Drug Administration. *General Principles of Software Validation*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/general-principles-software-validation> (2022).
- Schott, D. H., Collins, R. N., Bretscher, A. & Hernandez-Boussard, T. The AI life cycle: a holistic approach to creating ethical AI for health decisions. *Nat. Med.* **28**, 2247–2249 (2022).
- Aristidou, A., Jena, R. & Topol, E. J. Bridging the chasm between AI and clinical implementation. *Lancet Lond. Engl.* **399**, 620 (2022).
- European Commission. *Ethics By Design and Ethics of Use Approaches for Artificial Intelligence*. https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/guidance/ethics-by-design-and-ethics-of-use-approaches-for-artificial-intelligence_he_en.pdf (2021).
- Matheny, M., Israni, S. T., Ahmed, M. & Whicher, D. Editors. *Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril*. <https://nam.edu/wp-content/uploads/2019/12/AI-in-Health-Care-PRE-PUB-FINAL.pdf> (2022).
- Partnership on Health AI. *PAI's Responsible Practices for Synthetic Media: A Framework for Collective Action*. https://syntheticmedia.partnershiponai.org/#read_the_framework (2023).
- Kak, A. & West, S. M. *AI Now 2023 Landscape Confronting Tech Power*. <https://www.ainowinstitute.org/2023-landscape> (2023).
- Vasey, B. et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat. Med.* **28**, 924–933 (2022).
- Collins, G. S. et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. *BMJ Open* **11**, e048008 (2021).
- U.S. Food and Drug Administration. *Good Machine Learning Practice for Medical Device Development Guiding Principles*. <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles> (2021).
- U.S. Food and Drug Administration. *Content of Premarket Submissions for Device Software Functions*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/content-premarket-submissions-device-software-functions> (2021).
- U.S. Food and Drug Administration. *Cybersecurity in Medical Devices Quality System Considerations and Content of Premarket Submissions*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/cybersecurity-medical-devices-quality-system-considerations-and-content-premarket-submissions> (2022).
- International Standards Organization. *ISO 14971: Medical Devices—Application of Risk Management to Medical Devices*. <https://www.iso.org/standard/72704.html> (2019).
- Association for the Advancement of Medical Instrumentation. *CR 34971 Guidance on the Application of ISO 14971 to Artificial Intelligence and Machine Learning*. <https://array.aami.org/content/news/new-aami-consensus-report-guidance-risk-management-ai-ml> (2022).
- Coalition for Health AI. *Blueprint for Trustworthy AI Implementation Guidance and Assurance for Healthcare*. https://www.coalitionforhealthai.org/papers/blueprint-for-trustworthy-ai_V1.0.pdf (2023).
- Health AI Partnership (HAIP). *Key Decision Points*. <https://healthaipartnership.org/guiding-question/identify-and-mitigate-risks> (2023).
- U.S. Food and Drug Administration. *Design Control Guidance for Medical Device Manufacturers*. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/design-control-guidance-medical-device-manufacturers> (1997).

29. U.S. Food and Drug Administration. 21 CFR Ch. I (4–1–22 Edition) § 801.4. <https://www.govinfo.gov/content/pkg/CFR-2022-title21-vol8/pdf/CFR-2022-title21-vol8-sec801-4.pdf> (2023).
30. International Standards Organization. *IEC 62304: Medical Device Software—Software Life Cycle Processes*. <https://www.iso.org/standard/38421.html> (2006).

ACKNOWLEDGEMENTS

We acknowledge Stephanie Bernthal, M.Ed., of Mayo Clinic, for her work creating the visualizations included in the manuscript.

AUTHOR CONTRIBUTIONS

S.M.O. and N.J.E.-Z. jointly conceived the manuscript, designed the approach, and prepared the manuscript with substantial contributions from M.G.G., D.E.V., T.B. M.J.P., and J.D.H. made substantial contributions to the design and strategic positioning of the work and supervised the development from its inception. S.M.O., N.J.E.-Z., M.G.G., D.E.V., T.B., M.J.P., J.D.H. critically revised the manuscript for important intellectual content, approved the completed version, and are accountable for all aspects of the work.

COMPETING INTERESTS

Mayo Clinic and Duke University are founders of the Coalition for Health AI, which received funding from the Gordon and Betty Moore Foundation.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Shauna M. Overgaard.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023