Editorial

Validation and reporting standards for dental artificial intelligence: Lessons from materials biosafety

Marcin Mikulewicz[A–F]

Division of Facial Developmental Defects, Department of Maxillofacial Orthopaedics and Orthodontics, Faculty of Dentistry, Wroclaw Medical University, Poland

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

Dental and Medical Problems, ISSN 1644-387X (print), ISSN 2300-9020 (online) *Dent Med Probl*. 2026;63(5):1091–1094

Address for correspondence
Marcin Mikulewicz
E-mail: marcin.mikulewicz@umw.edu.pl

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
None declared

Received on July 10, 2026
Reviewed on July 13, 2026
Accepted on July 16, 2026

Published online on September 24, 2026

Keywords: dentistry, artificial intelligence, biocompatible materials

Extending dentistry's rigorous biomaterial-biosafety validation principles to artificial-intelligence (AI) treatment-planning tools requires prospective and external validation, clearly defined clinical indications, and post-deployment performance monitoring.

An orthodontic archwire cannot enter a patient's mouth until it has cleared a defined evidentiary gate. Its corrosion behavior and ion-release profile are quantified under standardized immersion conditions, benchmarked against the International Organization for Standardization (ISO) biocompatibility expectations, and reported in the literature, with the element-specific concentrations measured using techniques such as inductively coupled plasma optical emission spectrometry (ICP-OES).[1,2] A clinician selecting that wire can, in principle, trace its safety profile to reproducible measurements. The same cannot be said of the algorithm that increasingly helps determine how that wire will be used.

Dentistry has spent decades building a mature, quantitative paradigm for the biosafety of what it places in the patient – alloys, resins and ceramics – yet it is adopting artificial intelligence (AI) into what it decides for the patient with comparatively little validation, transparency or regulatory scaffolding. A discipline that would never accept an uncharacterized biomaterial is, in parallel, accepting largely uncharacterized decision tools. *Dental and Medical Problems* is a natural venue to identify this inconsistency and propose a corrective framework.

Two standards of evidence, one discipline

The biomaterials side of the ledger is well developed. A systematic synthesis of in vitro ion-release studies established, more than a decade ago, that the doses released from fixed appliances generally fell well below toxic thresholds, while also exposing the methodological limits of static immersion models that ignore biofilm, salivary dynamics and pH variation.[2] Comparative ICP-OES work continues to refine this picture – for example, showing that nickel-titanium and chromium-cobalt wires release significantly more nickel and chromium than stainless steel, with direct relevance to allergy risk.[1] The same measurement culture extends beyond wrought alloys – surface-engineered implant coatings are screened by combined physicochemical characterization, corrosion and cytocompatibility testing before any clinical claim is made,[3] and where static

Cite as
Mikulewicz M. Validation and reporting standards for dental artificial intelligence: Lessons from materials biosafety. *Dent Med Probl.* 2026;63(5):1091–1094. doi:10.17219/dmp/226044

DOI
10.17219/dmp/226044

immersion falls short, pulp-capping materials are taken into in vivo tissue-reaction testing by subcutaneous implantation, with the biological response assessed directly rather than inferred.[4] Whatever its imperfections, this is a culture of measurement: defined media, defined exposure, element-level quantification, and explicit acknowledgment of what the model does not capture.

The AI side looks different. The critical appraisals of AI across dental medicine repeatedly identify the same weaknesses – inconsistent data quality, the risk of bias, opacity, and limited clinical validation – which together undermine study quality and impede safe integration into routine care.[5] In teleorthodontics specifically, a scoping review found real-world deployment narrowly concentrated in commercial monitoring platforms, with variable repeatability of clinical decisions and limited independent validation; its conclusion was that AI should function as a complementary tool and never as a substitute for clinical judgement.[6] A recent *Dental and Medical Problems* editorial reaches a similar conclusion from a broader vantage point – AI as an adjunct rather than a replacement, with its promise contingent on validation, standardization and clinician training.[7] That is the same "assistant, not replacement" position the orthodontic community has reached elsewhere – sound as a principle, but not yet operationalized as an evidentiary requirement.

Adoption is outrunning preparation

The disconnect is not theoretical. A survey of postgraduate orthodontic programs in North America with an accreditation process found that more than half had introduced or planned to introduce AI training, yet most had no organized AI training, and about three quarters did not encourage residents to use AI for clinical patient care, research or teaching.[8] In other words, the technology is entering educational programs faster than curricula, oversight, and competency statements are being developed to govern its use. Adoption is preceding instruction – the inverse of how the profession introduced, for example, cone-beam imaging or new bonding systems.

Other branches of medicine have already articulated what a serious answer looks like. In cardiovascular device innovation, expert consensus frames the legitimate lifecycle as an early feasibility study, followed by randomized comparison with the accepted standard of care where one exists, and then post-market real-world evidence – while explicitly noting that regulatory standards and pathways for machine-learning applications remain under development.[9] Dentistry rarely subjects an AI planning aid to anything resembling that sequence. We would not adopt a new luting cement on the basis of a retrospective case series and a favorable impression; we should be equally reluctant to adopt a treatment-planning model on less.

Toward a biosafety frame for algorithms

The constructive move is to borrow the logic the profession already trusts. "Biosafety" in dental materials is not a single test, but a layered expectation: characterize the agent, define exposure conditions, quantify outputs, declare limitations, and report reproducibly. An analogous minimum set for an AI-driven dental tool is not exotic. It would characterize the tool along the same axes we already apply to a biomaterial – what it is, how it performs under realistic conditions, how it fails, and where it may be used – which the checklist below sets out concretely (Fig. 1).

A static alloy is characterized once, and thereafter behaves as characterized; a learning algorithm is a moving target – retrained, updated, and liable to drift as it encounters populations unlike those on which it was developed. Yet this disanalogy sharpens rather than defeats the frame, for materials science has never assumed that characterization at time zero is sufficient: it fatigue-tests materials over their projected service life and mandates post-market surveillance precisely, as performance can change under use. The algorithmic counterpart is neither exotic nor unattainable – versioned revalidation tied to each model update, coupled with continuous monitoring of real-world performance, simply extends to software the post-market-surveillance logic that cardiovascular consensus already applies to algorithmic change.[9]

Framed this way, the recurring "assistant, not replacement" formula stops being a reassurance and becomes a specification: an assistant whose reliability is characterized to the same standard we demand of the materials it helps deploy. This is also where a journal spanning dentistry and general medicine has a distinctive contribution to make. The validation problem is shared across radiology, cardiology and dentistry; in cardiology, for instance, AI for heart-failure diagnosis and risk stratification remains constrained by overfitting, limited external validation and limited explainability – the very barriers that impede routine deployment.[10] The materials-science culture of quantified biosafety, by contrast, is comparatively specific to dentistry. Translating the second into a template for the first is a genuinely cross-disciplinary task, and it is the kind of contribution the journal's broadened article types can accommodate, with perspectives, comments and guidelines now included alongside editorials.[11]

Concretely, this journal could ask any submission describing an AI-driven dental tool to report 5 things (Fig. 1):

1. Specify the composition and representativeness of the training population, including the settings and demographics from which it was drawn (the analog of declaring alloy composition).
2. Report prospective and, where feasible, external validation against outcomes in deployment-relevant populations, rather than retrospective performance on held-out data alone (the analog of in vivo rather than bench testing).

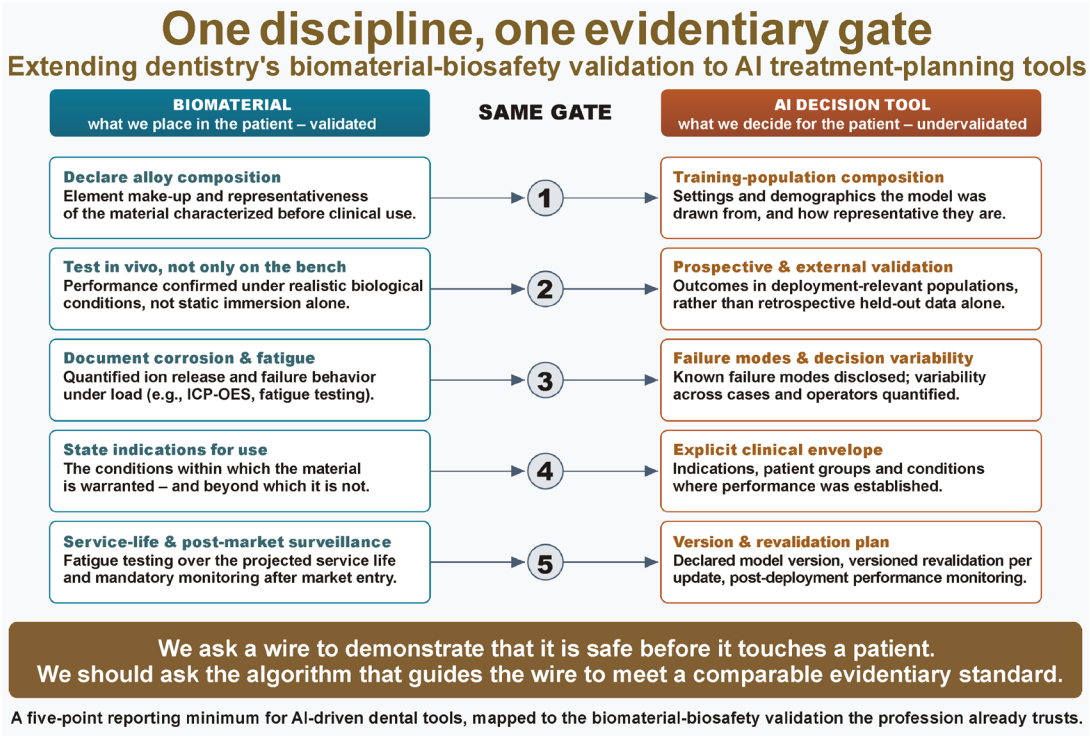

Fig. 1. Artificial intelligence (AI) biosafety

One discipline, one evidentiary gate. The five-point reporting minimum proposed for AI-driven dental tools (right), with each item mapped to the corresponding element of the biomaterial-biosafety validation the profession already applies (left).
ICP-OES – inductively coupled plasma optical emission spectrometry.

3. Disclose known failure modes and quantify decision variability across cases and operators (the analog of documented corrosion and fatigue behavior).
4. State the clinical envelope explicitly – the indications, patient groups and conditions within which performance was established, and beyond which the tool is not warranted (the analog of indications for use).
5. Declare a model version and a plan for versioned revalidation and post-deployment performance monitoring (the analog of service-life fatigue testing and post-market surveillance).

None of this is an argument against AI in dentistry; the case for its diagnostic and workflow value is strong and growing. It is an argument against a double standard. We ask a wire to demonstrate that it is safe before it touches a patient. We should ask the algorithm that guides the wire to meet a comparable evidentiary standard. Closing that gap – conceptually, and eventually through reporting and validation expectations that editors and reviewers can enforce – is a concrete, near-term agenda that this journal is well placed to lead.

Ethics approval and consent to participate

Not applicable.

Data availability

Not applicable.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

The author declares that no generative artificial intelligence (AI) or AI-assisted technologies were used in the preparation of this manuscript.

ORCID iD

Marcin Mikulewicz https://orcid.org/0000-0001-5754-0284

References

1. Mikulewicz M, Suski P, Tokarczuk O, Warzyńska-Maciejewska M, Pohl P, Tokarczuk B. Metal ion release from orthodontic archwires: A comparative study of biocompatibility and corrosion resistance. *Molecules*. 2024;29(23):5685. doi:10.3390/molecules29235685
2. Mikulewicz M, Chojnacka K. Release of metal ions from orthodontic appliances by in vitro studies: A systematic literature review. *Biol Trace Elem Res*. 2011;139(3):241–256. doi:10.1007/s12011-010-8670-9
3. Biju D, Arumugam P, Kannan S, et al. Development, characterization, and biocompatibility and corrosion analyses of a silver-decorated graphene oxide and chitosan surface coating for titanium dental implants: A preliminary report. *Dent Med Probl*. 2024;61(4):627–632. doi:10.17219/dmp/187107
4. Rady D, El Moshy S, Yousry YM, Ramadan M, Radwan IA. Biocompatibility of NeoPUTTY™ versus TheraCal PT® and TheraCal LC® pulp-capping materials, and tissue reaction following subcutaneous implantation in rats: Histological and molecular investigation. *Dent Med Probl*. 2026;63(1):131–143. doi:10.17219/dmp/168627

5. Sitaras S, Tsolakis IA, Gelsini M, et al. Applications of artificial intelligence in dental medicine: A critical review. *Int Dent J*. 2025;75(2):474–486. doi:10.1016/j.identj.2024.11.009
6. Polizzi A, Serra S, Leonardi R, Isola G. Clinical applications of artificial intelligence in teleorthodontics: A scoping review. *Medicina (Kaunas)*. 2025;61(7):1141. doi:10.3390/medicina61071141
7. Paradowska-Stolarz AM. AI in modern dentistry: Hype, hope, or real transformation? *Dent Med Probl*. 2025;62(6):1013–1016. doi:10.17219/dmp/211356
8. Hanenkrath J, Park JH, Bay C. Training, use, and modifications related to artificial intelligence in postgraduate orthodontic programs in North America. *Am J Orthod Dentofacial Orthop*. 2025;167(1):89–94.e2. doi:10.1016/j.ajodo.2024.09.008
9. Windecker S, Gilard M, Achenbach S, et al. Device innovation in cardiovascular medicine: A report from the European Society of Cardiology Cardiovascular Round Table. *Eur Heart J*. 2024;45(13):1104–1115. doi:10.1093/eurheartj/ehae069
10. Resch J, Lelonek M. The role of artificial intelligence in the diagnosis, risk stratification, and treatment of heart failure: A narrative review [published online as ahead of print July 3, 2026]. *Adv Clin Exp Med*. doi:10.17219/acem/218578
11. Więckiewicz M, Martynowicz H. Journal Impact Factor and highly cited papers: The beginning of a new era in *Dental and Medical Problems*. *Dent Med Probl*. 2023;60(4):541–542. doi:10.17219/dmp/176039