Int. J. Oral Maxillofac. Surg. 2026; 55: 756–773
https://doi.org/10.1016/j.ijom.2026.02.020, available online at https://www.sciencedirect.com

International Journal of Oral & Maxillofacial Surgery

Scoping Review
Clinical Pathology

Artificial intelligence in oral and maxillofacial surgery: a scoping review of clinical applications, ethical challenges, and legal considerations

G. Consorti[a,b], L. Catarzi[c], A. Frosolini[c], L.A. Vaira[d], U. Committeri[e], G. Cirignaco[c]

[a]Division of Maxillofacial Surgery, Department of Neurological Sciences, Marche University Hospitals-Umberto I, Ancona, Italy; [b]Department of Clinical Specialistic and Dental Sciences, Marche Polytechnic University, Ancona, Italy; [c]Department of Medicine, Section of Maxillofacial Surgery, University of Siena, Siena, Italy; [d]Maxillofacial Surgery Operative Unit, Department of Medicine, Surgery and Pharmacy, University of Sassari, Sassari, Italy; [e]Department of Maxillofacial Surgery, "S. Maria" Hospital, Terni, Italy

G. Consorti, L. Catarzi, A. Frosolini, L. A. Vaira, U. Committeri, G. Cirignaco: Artificial intelligence in oral and maxillofacial surgery: a scoping review of clinical applications, ethical challenges, and legal considerations. Int. J. Oral Maxillofac. Surg. 2026; 55: 756–773.

Abstract. Artificial intelligence (AI) is increasingly used in oral and maxillofacial surgery (OMS) for imaging, planning, and robotic/navigation support, yet ethical and legal governance remains unsettled. We conducted a PRISMA-ScR scoping review to map clinical applications and associated risk domains. PubMed/MEDLINE, Scopus, and Web of Science were searched for English-language articles published from January 2015 to July 2025. Eligible records reported a clinical AI application in OMS or analyzed ethics/legal issues relevant to OMS; purely technical papers without clinical linkage, non-indexed sources and conference abstracts, non-English items, and records without an abstract were excluded. Two reviewers screened studies independently with consensus on inclusion. Of 2158 records, 99 met eligibility. Ethical discussions most often addressed accuracy/reliability, transparency/explainability, and bias/fairness. External multicenter validation and calibration were uncommon, prospective decision-impact or effectiveness studies were rare, and no large randomized trials were identified. Recurrent ethical–legal themes included informed consent for AI involvement, fairness auditing, privacy and data protection, allocation of responsibility for decision-support versus autonomous functions, and post-deployment monitoring for performance drift. AI shows encouraging technical performance in OMS, but patient benefit will require stepwise, monitored integration, standardized reporting, prospective studies, strengthened consent processes, fairness and privacy safeguards, and clear professional and manufacturer accountability.

Keywords: Artificial intelligence; Ethics; Medical liability; Machine learning; Maxillofacial surgery.

Accepted for publication 24 February 2026
Available online 5 March 2026

Artificial intelligence (AI)—defined as a group of computer systems capable of performing tasks that normally require human intelligence—has emerged as an innovative technology in medicine, surgery, and dentistry with a recent exponential growth of interest in its applications[1–3]. AI systems are now applied in various tasks ranging from diagnostic image analysis to treatment planning and patient management. In the field of oral and maxillofacial surgery (OMS), which encompasses complex diagnostic and surgical care for craniofacial conditions, AI offers tools that can enhance a surgeon's capabilities in unprecedented ways[4–8]. For instance, in dentomaxillofacial imaging tasks, machine-learning systems have matched expert-level accuracy while shortening the interpretation time[9–11]. AI-driven surgical planning software can help simulate orthognathic surgery outcomes or design patient-specific implants, improving precision in reconstructive procedures[12,13]. Robotics and intelligent systems are also being explored to enhance intraoperative guidance and execute repetitive tasks with high precision[14,15]. Collectively, these innovations have demonstrated early, context-dependent performance gains in diagnostic accuracy and operative efficiency in OMS; however, their impact on patient-level outcomes remains provisional pending robust multicentre validation and governance safeguards. Without robust external validation, bias auditing, and clear oversight, AI systems risk propagating diagnostic errors, entrenching inequities, or introducing workflow and medico-legal risks.

Despite the enthusiasm surrounding AI, its implementation in healthcare must be approached cautiously and thoughtfully[16,17]. The deployment of AI in clinical settings raises ethical questions about patient safety and privacy, necessitating transparency in algorithmic decision-making[16,18,19]. For example, if an AI model recommends a treatment based on patterns in data that clinicians find problematic, how should that influence clinical decision-making?[20,21] Is it ethical to act on recommendations that may lack explainability to the care team or the patient? Similarly, biases in training data could result in unequal performance of AI systems across different patient populations, potentially exacerbating health disparities[22].

Legal considerations are equally critical. Regulatory bodies are working to establish guidelines for AI application in healthcare: in the United States, the Food and Drug Administration (FDA) has begun formulating frameworks for AI/machine learning-based medical devices[23–25], and in Europe, lawmakers are advancing the Artificial Intelligence Act (AIA) aimed at ensuring "human-centric and trustworthy AI" in high-risk domains, including medicine[26,27]. Questions of liability loom large: if an AI system error contributes to patient harm, current legal frameworks must determine who is accountable: the clinician using the tool, the hospital, or the software manufacturer. Additionally, compliance with data protection regulations, such as HIPAA (Health Insurance Portability and Accountability Act) in the United States and GDPR (General Data Protection Regulation) in Europe, is mandatory when AI systems handle sensitive health data, adding further complexity to their development and deployment[28–30].

While the application of AI in dentistry has been explored increasingly, no comprehensive mapping has yet addressed both the breadth of AI applications and the associated ethical-legal landscape, specifically in OMS. Hence, the following research question was posed: "How can AI be ethically and legally integrated into oral and maxillofacial surgery practice, and what evidence currently supports its clinical applications?".

This paper presents a scoping review of the literature on AI in OMS, with a focus on: (1) the clinical applications of AI explored or implemented in OMS; (2) the ethical challenges arising from AI use (including informed consent, transparency/explainability, bias/fairness, and data governance); and (3) the legal and regulatory considerations relevant to surgical AI; e.g., FDA/SaMD (software as a medical device) and the European Union (EU) AIA, liability, and cybersecurity.

Specifically, it was sought to: catalogue and classify clinical applications and their reported performance; synthesize ethical issues—with particular attention to patient informed consent, transparency/explainability, bias/fairness, and data governance; summarize legal and regulatory frameworks relevant to OMS (FDA/SaMD landscape, EU AIA, liability allocation, and cybersecurity); and identify evidence gaps and priority areas for future research and policy.

Accordingly, and consistent with the objectives stated above, the dual aims of this scoping review were to map the current state of clinical applications of AI in OMS and to examine, in depth, the ethical, legal, and regulatory implications, explicitly distinguishing surgeon-led decision-support uses from embedded AI within robotics and navigation.

Materials and methods

Protocol

This scoping review was conducted according to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR)[31] to ensure rigor and transparency. The review was registered in the Open Science Framework (OSF) platform on May 4, 2025 (https://doi.org/10.17605/OSF.IO/KG3UH). After registration, a single amendment was made only to extend the time window to include studies published in 2025; there were no changes to the review question, eligibility criteria, data items, or outcomes. The last search date for each database was July 30, 2025.

A systematic search was conducted in January 2025 across three major databases: PubMed/MEDLINE, Scopus, and Web of Science. The search strategy combined keywords and subject headings related to AI, including 'artificial intelligence', 'machine learning', 'deep learning', 'neural network', and 'expert system', with terms relevant to OMS, such as 'oral surgery', 'maxillofacial surgery', 'dentistry', and 'oral and maxillofacial surgery'. To capture ethical and legal discussions, terms like 'ethics', 'ethical', 'legal', 'law', 'regulation', and 'liability' were also incorporated in secondary searches. For instance, a PubMed query was: (Artificial Intelligence [MeSH] OR machine learning OR deep learning OR neural network) AND (oral surgery OR maxillofacial surgery OR dentistry).

Filters were applied to include English-language studies published between January 2015 and July 2025.

In addition, a hand-search of the reference lists from relevant articles was performed to identify additional studies.

As a scoping review, no formal risk-of-bias assessment or certainty-of-evidence grading (e.g., GRADE) was conducted. This approach aligns with PRISMA-ScR and the scoping-review framework, which prioritize mapping

the breadth and characteristics of the evidence over quantitative synthesis[31–33]. The sources included were heterogeneous (algorithm development studies, retrospective diagnostic accuracy reports, narrative and perspective pieces) and rarely reported comparable patient-level outcomes or standardized effect measures.

Data charting and extraction

A data extraction form was developed to systematically chart the key information from each included study. The following data were extracted: bibliographic details (author, year, journal); study type (e.g., clinical study, review, perspective); AI application in OMS (e.g., imaging diagnosis, surgical planning, robotics, etc.); reported outcomes or performance metrics; ethical, legal, or regulatory discussions.

Sections discussing ethical considerations (e.g., bias, informed consent) and legal or regulatory aspects (e.g., FDA approval, policy recommendations) were focused upon.

Eligibility criteria

Titles and abstracts were screened for relevance. This was followed by a full-text assessment based on the following inclusion criteria: studies addressing AI applications in any aspect of OMS (e.g., diagnosis, treatment planning, surgery, and postoperative care); studies discussing ethical or legal issues related to AI in medicine and surgery, specifically with relevance to OMS; articles published in English; articles published between January 2015 and July 2025, as deep-learning frameworks underpinning contemporary medical-imaging AI were only standardized after 2015[34].

Exclusion criteria included studies focusing purely on technical AI development without clinical relevance; articles unrelated to medicine or surgery; studies that did not explicitly mention AI-related technologies; articles without an available abstract online; studies published in languages other than English; conference abstracts and papers published in non-indexed sources (i.e., not included in Web of Science, Scopus, or PubMed).

Methodological clarity

The guiding research question for this review was: "How can AI be ethically and legally integrated into oral and maxillofacial surgery practice, and what evidence currently supports its clinical applications?".

To ensure methodological transparency and reproducibility, all records identified were imported into Mendeley, where duplicate records (n = 250) were automatically removed.

The data extraction was performed independently by two reviewers (G.C., G.C.) employing a pre-piloted extraction form designed to capture key bibliographic details, study types, AI applications, outcomes, and ethical, legal, and regulatory discussions. The extraction form was pilot-tested on a sample of studies to ensure clarity and consistency. Discrepancies between the two reviewers were discussed; if consensus could not be reached, a third reviewer was consulted.

Conference abstracts and other non-indexed sources were excluded to preserve a standard of peer-reviewed methodological transparency. This is because such abstracts are rarely peer-reviewed, often lack sufficient detail for critical appraisal, and their results frequently change when later published as full articles[35,36].

Synthesis of results

Considering the scoping nature of this review, a descriptive synthesis was performed. The findings were grouped into three primary thematic categories, corresponding to the study objectives: clinical applications of AI in OMS; ethical challenges related to AI for OMS; legal and regulatory considerations for AI in OMS.

Within each category, the state-of-the-art developments, notable insights, and gaps in the literature were summarized, thereby providing a structured overview of the role of AI in OMS.

Results

Study selection and characteristics

The search identified 2158 records, of which 250 were duplicates and removed. Consequently, 303 were deemed relevant after screening based on titles, of which 129 remained for full-text assessment after abstract screening. Overall, 99 (4.6%) publications met all of the inclusion criteria and were included in this scoping review (Fig. 1). The included sources were published between January 2015 and July 2025, with a notable increase in publications in the last 5 years reflecting the growing interest in AI in dentistry and surgery. Most clinical studies were exploratory in nature—testing the accuracy of AI algorithms on retrospective datasets—and only a few reported on prospective clinical implementation of AI systems. No large-scale randomized controlled trials (RCTs) of AI in OMS were found. This is attributable to the nascent state of the field. Geographically, the literature included contributions primarily from North America, Europe, and East Asia, paralleling the global interest in medical AI.

For clarity of presentation, the findings are presented under three thematic sub-sections: (1) the clinical applications of AI in OMS, (2) safety considerations including common modes of error and potential clinical consequences, and (3) ethical, legal, and regulatory considerations, which are presented through focused subheadings (bias and fairness, data privacy and confidentiality, informed consent, and the clinician–patient relationship). There is some overlap; for instance, data privacy may be considered under both an ethical and a legal lens. Such cases are presented where they were most emphasized in the literature and their inter-relationships are discussed later. An overview of the articles included in this scoping review is provided in Supplementary material Table S1.

Clinical applications of AI in OMS

AI has been applied to various clinical tasks in OMS, leveraging methods such as machine learning (including deep neural networks) to assist clinicians.

Beyond the broad task categories, 11 discrete maxillofacial clinical situations illustrate how AI maps onto specific decisions, user groups, evidence profiles, and unrealized gaps.

Early jaw lesion triage and characterization

Task: The differentiation and flagging of odontogenic keratocysts, dentigerous cysts, ameloblastomas, and other radiolucent lesions for earlier specialist referral. Deep convolutional and detector architectures applied to panoramic radiographs and cone-beam computed tomography (CBCT) images aim to detect these lesions earlier[37–39]. Evidence: The YOLO v2 convolutional

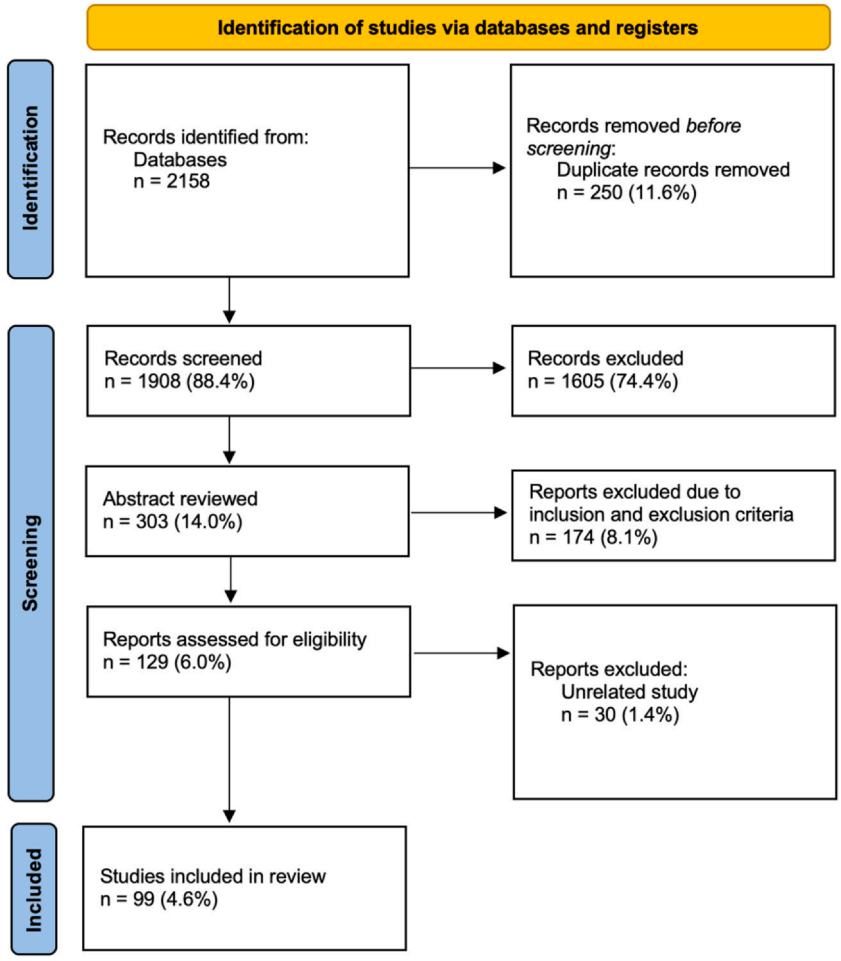

Fig. 1. PRISMA flow chart of the study selection process.

neural network (CNN) introduced by Yang et al.[40] was trained on 1602 annotated lesions (dentigerous cysts, odontogenic keratocysts, ameloblastomas), achieving a precision of 0.707 and recall of 0.680, comparable to OMS and general practitioners. A recent meta-analysis by Shoorgashti et al.[41] reported pooled sensitivity of 83.7% (95% confidence interval (CI) 73.8–93.6%) and specificity of 82.9% (95% CI 70.3–95.5%) for detecting odontogenic keratocysts on panoramic radiographs. Primary users: General dentists or junior clinicians. The immediate value is early differentiation between lesions that can be treated routinely and those requiring a specialist. Gaps and limitations: There has been sparse multicentre external validation; limited calibration reporting (no Brier/net benefit metrics); an absence of prospective 'time-to-referral' or morbidity impact studies; potential false-negatives for small/overlapping root lesions. Boundary: There is currently no prospective evidence that AI triage reduces the delay to intervention or complication rates.

Cephalometric landmark localization and craniofacial segmentation for orthodontic–orthognathic planning

Task: Automated two-dimensional (2D) landmark detection and three-dimensional (3D) segmentation to accelerate virtual surgical planning and reduce inter-observer variability. Evidence: Fully deep learning and region-based CNN systems have achieved mean radial errors (MRE) around or below 2 mm, reducing the planning time and inter-observer variability[42,43]. Kim et al.[44] developed a model using 950 radiographs, achieving a MRE of 1.84 mm—comparable to inter-examiner variability—and outperforming clinicians for landmarks such as A-point and the incisal tips. Jiao et al.[45] applied a Mask R-CNN model to 400 lateral cephalometric radiographs, achieving a detection rate of 98.3%, with 68.3% of landmarks within 2 mm of reference points. Table 1 presents key diagnostic accuracy metrics for representative AI systems in cephalometric workflows. Primary users: Orthodontists and orthognathic teams seeking geometric precision. The outputs influence osteotomy design and the prediction of post-treatment facial proportions. Gaps and limitations: These include heterogeneous reporting (accuracy versus successful detection rate (SDR) versus MRE; variable thresholds), rare calibration/decision curve analysis, and limited evidence that a shorter planning time alters surgical accuracy, relapse, patient-

Table 1. Diagnostic accuracy of recent AI models for cephalometric landmark detection and related tasks (2015–2025).

Study Year	Imaging modality	Task/landmarks	Dataset (images)	Main metric (s)	Performance
Tang et al.[43] 2025	3D CBCT	Mx and Md segmentation + detection of 18 landmarks (orthognathic planning)	200 training 50 test	Dice: Mx: 93.41% Md: 96.89%	Mean landmark error: 1.91 ± 1.17 mm
Han et al.[42] 2025	Lat Ceph	Detection of 22 landmarks + lip-thickness classification	1040 training 260 test	MRE: 0.97 ± 0.52 mm SDR ≤2 mm = 95.4%	AUC for lip-thickness classes ≥0.97
Jiao et al.[45] 2024	Lat Ceph	Detection of 19 landmarks (Mask R-CNN)	320 training 80 test	Detection rate 98.29%;	Average error 2.19 mm SDR ≤2 mm = 68.3%

3D, three-dimensional; AUC, area under the receiver-operating-characteristic curve; CBCT, cone-beam computed tomography; Dice, Dice similarity coefficient; Lat Ceph, lateral cephalogram; Md, mandible; MRE, mean radial error; Mx, maxilla; R-CNN, region-based convolutional neural network; SDR, successful detection rate.

reported outcomes, or costs. Regarding failure modes, these include soft tissue boundary landmarks and metallic artefact distortion. Boundary: There is no multicentre RCT evidence linking automated landmarking to improved functional or patient-reported outcomes.

Classification of temporomandibular joint (TMJ) disc displacement

Task: Magnetic resonance imaging (MRI)-based classification of anterior disc displacement; segmentation of the condyle/fossa structures. Evidence: MRI classifiers distinguish anterior disc displacement with external precision in the low 90% range[46], while segmentation methods delineate condyle–fossa structures[47,48]. Primary users: TMJ specialists and radiologists seeking earlier stratification for conservative management versus invasive intervention. Gaps and limitations: These include inter-MRI scanner variability; minimal longitudinal correlation between the automated findings and the trajectory of symptoms or treatment response; limited analysis of fairness (age, sex). Failure modes include ambiguous intermediate disc positions and motion artefacts. Boundary: There is no prospective evidence that AI classification changes management pathways or outcomes.

Emergency mandibular fracture triage

Task: Automated flagging of suspected fractures in emergency/rural settings before specialist availability. Panoramic fracture detection networks support on-call or rural clinicians in identifying potential fractures before maxillofacial referral[11]. Gaps and limitations: There is a lack of external multicentre validation across detector vendors; few workflow studies (impact on time to fixation or imaging escalation); limited false-positive characterization (normal anatomical lines, ghost images). Failure modes include subtle/hairline fractures and overlap artefacts. Boundary: There is no prospective data showing that AI-driven triage reduces missed fractures or expedites treatment.

Implant site optimization and implant brand/stage classification

Task: Pre-surgical angulation/depth suggestion; automated identification of implant system and stage; support for guided or robotic placement. Planning algorithms propose optimal angulation or depth, and differentiate implant designs or treatment stages on panoramic imaging[12,49]. These systems may improve implant positioning accuracy and surgical consistency, reducing errors and secondary procedures. Evidence: Early studies suggest AI-assisted implant surgery results in fewer misplacements and reduced revision surgeries compared to freehand techniques. AI algorithms integrated into navigation and augmented reality systems help identify critical structures (e.g., neurovascular bundles) intraoperatively, enhancing surgical safety. Primary users: Implant surgeons and prosthodontists. Gaps and limitations: These include the limited cross-scanner generalizability; scant data on 'near miss' rates (e.g., proximity to the inferior alveolar canal), complications, and long-term implant survival; sparse calibration. Failure modes include misclassification of uncommon implant designs; trajectory suggestions ignoring soft tissue/prosthetic constraints. Boundary: There are no controlled prospective trials showing reduced complication or revision rates.

Orthognathic sequencing, stability, and relapse risk prediction

Task: Predict suitability for surgery-first protocols and the magnitude of relapse or soft tissue changes. Models forecast skeletal relapse, the soft tissue response, or suitability for surgery-first protocols[6,50–52]. Primary users: Planning teams and surgeons tailoring sequencing and patient counselling. Gaps and limitations: These include calibration drift over time; an absence of prospective decision-impact studies; lack of long-term functional or sensory outcome data. Failure modes include extrapolation beyond training morphologies and underestimation in severe asymmetry. Boundary: There is no prospective evidence that model-guided sequencing reduces relapse or improves quality of life.

Robot-assisted transoral, implant, and reconstructive procedures

Task: Enhanced precision of execution, ergonomics, and access to confined corridors. Surgical robotics promise precision in execution, ergonomic advantages, and potential quality of life gains[15,53–55]. Primary users: Operating surgeons. Gaps and limitations: There is a paucity of comparative OMS trials measuring operative times, complications, and cost-effectiveness; unresolved medico-legal issues related to partial autonomy. Failure modes include registration/tracking drift, latency, and constrained instrument articulation. Boundary: There is limited controlled outcome evidence for OMS indications[15,20].

High-resolution CBCT/CT multi-structure segmentation for osteotomy planning and anatomical risk mitigation

Task: Automated segmentation of the jaws, condyles, fossae, and pathological lesions to support virtual osteotomy vectors and neurovascular avoidance. Semi-supervised and tracking-based approaches are used to segment the maxilla, mandible, condyles, fossae, and pathological areas, and achieve Dice scores often > 0.90, supporting collision avoidance and precise osteotomy vectoring. Primary users: Virtual planning teams and intraoperative navigation systems[38,39,43]. Failure modes include artefact zones (metal scatter), severe deformity, and under-segmentation of canals near a pathology. Gaps and limitations: Cross-vendor robustness testing is scarce; there is limited reporting on how segmentation accuracy translates to the operative time, nerve injury, or margin status. Boundary: There is no demonstrated downstream reduction in morbidity.

Natural language processing (NLP) for safety surveillance and documentation efficiency

Task: Summarize the clinical notes, extract perioperative risk factors, flag drug interactions, draft consent documents, and generate patient education. NLP systems can reduce the cognitive load by summarizing notes and extracting risk factors[56], while conversational agents support peri- and postoperative clarification[18]. Primary users: Perioperative teams. The value includes reduced transcription error, faster handovers, and potential prevention of adverse events. Gaps and limitations: Samples have been small; there is a risk of hallucination or omission of rare complications; there is sparse systematic error auditing; medico-legal standardization is limited. Failure modes include the omission of

patient-specific contraindications; outdated guidelines; incorrect medication interaction flags. Boundary: There is no robust evidence of reduced adverse events or documentation litigation risk.

Detection of oral potentially malignant disorders (OPMD)

Task: AI-assisted adjunctive imaging or photographic analysis to identify OPMD suspicious lesions. A recent systematic review highlighted the role of AI-assisted clinical imaging in the detection of OPMD and oral cancer, evaluating the variations in accuracy among different imaging tools employed in these diagnostic processes, with a sensitivity of 89.9% (95% CI 86.6–92.5%) and a specificity of 89.2% (95% CI 85.1–92.2%), alongside a high negative predictive value of 89.5% (95% CI 85.1–92.7%)[2]. Primary users: General dental practitioners and multidisciplinary tumour boards. Gaps and limitations: There is heterogeneity of imaging modalities and preprocessing; limited head-to-head comparisons with clinician visual examinations under standardized conditions; external validation cohorts have been small; there is a risk of false-negatives in early, low-contrast lesions. Failure modes include misclassification of keratotic benign lesions and illumination artefacts. Boundary: There is no evidence that AI screening reduces diagnostic delay or stage at presentation.

Outcome prediction and patient monitoring

Task: Machine learning stratification of the postoperative complication risk or likelihood of relapse; monitoring via integrated records and imaging. By training on large datasets of surgical cases, machine learning models can predict the likelihood of complications such as infection, nerve injury, or postoperative relapse in orthognathic cases based on preoperative and intraoperative variables[51,52]. Beyond direct clinical care, AI has been piloted to streamline health records management and improve medication safety. Examples include NLP systems that automatically update patient records or extract relevant clinical summaries, and AI tools that cross-check patient medication lists to flag potential drug–drug interactions or allergies, thereby reducing prescription errors[56]. Gaps and limitations: There is a predominance of retrospective datasets; absence of prospective adaptive intervention trials (e.g., adjusted follow-up schedules); limited reporting of calibration and decision curves; minimal fairness auditing. Failure modes include overestimation of the risk leading to unnecessary imaging or interventions; underestimation delaying timely intervention. Boundary: There are currently no prospective data demonstrating reductions in complication rates or improved resource utilization.

Embedded systems versus predictive analytics

Hardware-embedded intraoperative systems such as robot-assisted surgery have advanced more rapidly from proof-of-concept to routine clinical use than data-driven predictive-analytics algorithms. A national registry analysis of 169,404 general surgery cases reported an increase in robotic utilization climbing from 1.8% in 2012 to 15.1% by 2018 following platform installation, largely supplanting laparoscopy[54]. In OMS specifically, a 2024 scoping review documented year-on-year growth in robot-assisted case reports and reported patient quality-of-life gains, particularly in trans-oral and implant procedures[55]. Several factors underpin this disparity: (1) robots provide an immediately visible procedural advantage and align with existing reimbursement codes, whereas predictive models offer indirect, often opaque benefits; (2) device-level FDA and CE approval pathways for robotics are well established, while regulatory frameworks for algorithmic tools under 'Software as a Medical Device' (SaMD) guidance continue to evolve; (3) surgeons tend to trust technology they can directly observe and override, whereas 'black-box' predictions elicit hesitancy, particularly when liability is unclear[20,27]. Consequently, most predictive-analytics applications, such as the predictive model for guidance on the treatment strategy for mandibular reconstruction reported by Dong et al.[57], remain in retrospective validation phases without prospective clinical deployment.

Performance differences by anatomical region and imaging modality

Subgroup analysis by region and modality revealed notable performance differences. CBCT-driven models targeting bony structures demonstrated strong performance: a semi-supervised 3D CNN achieved Dice coefficients of 93.4% for the maxilla and 96.9% for the mandible, with a mean landmark error of 1.9 mm when assisting orthognathic-surgery planning[43], while a U-Net-based condyle/fossa segmentation on low-dose CT achieved Dice scores of 0.92 ± 0.03 for condyles and 0.90 ± 0.04 for glenoid fossae[47]. These findings indicate that soft tissue MRI excels for TMJ disc pathology, whereas CBCT/CT provides greater geometric fidelity for skeletal tasks, highlighting the importance of aligning the imaging modality with the clinical objective.

Common modes of error and potential clinical consequences

Characteristic error patterns are observed across applications: (1) Lesion detection: false-negatives for small peri-apical or root-superimposed radiolucencies and false-positives on anatomical variants may delay specialist referral or trigger unnecessary imaging. (2) Landmarking/segmentation: drift >2 mm in low-contrast or artefact-affected regions can propagate into inaccurate osteotomy vectors or guide fabrication errors. (3) TMJ classification: ambiguity in intermediate disc positions risks inappropriate escalation or missed initiation of conservative therapy. (4) Fracture triage: missed subtle or hairline fractures on panoramic radiographs can defer fixation, while false-positives increase unnecessary CBCT. (5) Implant planning: misclassification of an implant system or a suggested trajectory ignoring the prosthetic envelope may raise the risk to the inferior alveolar canal or sinus. (6) Relapse/outcome prediction: poorly calibrated probabilities may induce overtreatment (additional fixation) or underplanning retention strategies. (7) Multistructure segmentation: under-segmentation of the mandibular canal or tumour margins may increase nerve injury or the risk of a positive margin. (8) OPMD detection: false-negatives in early mucosal dysplasia may delay biopsy. (9) NLP/large language model (LLM) consent and documentation: omissions ('negative hallucinations') of rare but material complications or the insertion of fabricated references create medico-legal vulnerability. (10) Monitoring dashboards: automation bias may lead clinicians to discount discrepant clinical signs.

Each error type aligns with the mitigation strategies outlined in Table 2, including external validation,

Table 2. Key (non-exhaustive) AI risk domains in oral and maxillofacial surgery and layered mitigation approaches.

Risk domain	Why it matters in OMS	Layered mitigation (mechanism-focused)	References
Data representativeness and sampling bias	Regional/ demographic skew in CBCT, panoramic, cephalometric datasets can degrade performance on under-represented skeletal classes or ethnic groups, risking unequal diagnostic accuracy	Stratified acquisition → transparent cohort reporting (TRIPOD-AI items) → subgroup performance and calibration metrics each release → dataset re-balancing/ domain adaptation if disparity threshold exceeded	Cross et al.[80] Norori et al.[81] Ueda et al.[82]
Model performance and generalization	Internal test results can overstate accuracy; lack of external multicentre validation leads to silent failure on new scanners or populations	External multicentre validation → prospective pilot logging (DECIDE-AI) → continuous drift monitoring with trigger thresholds	Collins et al.[74] Vasey et al.[104]
Calibration and clinical reliability	Well-discriminating models may still misestimate absolute risk (malignancy, complication probability)	Report calibration plots and Brier score → recalibrate on new data → decision-curve analysis for net benefit	Collins et al.[74]
Fairness and equity (beyond dataset diversity)	'Diverse' training alone may leave residual sensitivity gaps by sex, age, ethnicity	Deployment subgroup error tracking → predefined remediation rule (e.g. absolute sensitivity gap > 5%) → bias root-cause analysis → targeted re-training + documentation	Cross et al.[80] Norori et al.[81] Ueda et al.[82]
Explainability and informed consent	Opaque outputs impede disclosure, comprehension, voluntariness in consent	Saliency/ attention maps + textual rationales → readability-adjusted patient summaries → archive artefacts with consent record	Ali et al.[83] Vaira et al.[75] Teasdale et al.[84]
Privacy and data protection	Multimodal imaging + text integration raises re-identification and cross-border transfer risk (GDPR/HIPAA)	Data minimization and de-identification → encryption and access logging → compliant transfer agreements → privacy impact assessments	Yadav et al.[29] Marks et al.[89] McGraw and Mandl[78]
Security and integrity	Adversarial perturbations or tampering with planning/ robotic files could cause procedural harm	Secure version control → cryptographic signing of model and plan files → intrusion detection and integrity checks pre-execution	Ramos et al.[100] López González et al.[101] Elendu et al.[102]
Liability and accountability	Ambiguity (surgeon versus manufacturer) can chill adoption or misallocate risk	Clear labelling (decision support versus autonomous) → inference and version logs → contractual allocation of product versus professional liability	Maliha et al.[20] Terranova et al.[88] Cestonaro et al.[98]
Regulatory alignment and adaptive updates	Continuous-learning models may drift beyond cleared indications.	Algorithm change protocol → classify updates (minor/major) → re-submission or internal review per class → public change log	Smith et al.[23] Gilbert et al.[24] Busch et al.[71] Aboy et al.[72] van Kolfschooten et al.[73]
Post-market surveillance and drift	Performance decay (scanner upgrade, new demographics) may be unnoticed	Continuous dashboards → statistical drift tests → scheduled (e.g. 12-month) re-validation → rollback triggers	Busch et al.[71] Aboy et al.[72] Ueda et al.[82]
Workforce competence and training	Misinterpretation undermines safety and trust	Formal AI literacy (residency + CME) → competency assessment before privileged use → refresher after major updates	Weidener and Fischer[27] Quah et al.[86]

AI, artificial intelligence; CBCT, cone-beam computed tomography; CME, continuing medical education; DECIDE-AI, reporting guideline for early-stage clinical evaluation of AI decision support; GDPR/HIPAA, General Data Protection Regulation/Health Insurance Portability and Accountability Act; OMS, oral and maxillofacial surgery; TRIPOD-AI, Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis—AI extension.

calibration with decision curves, subgroup performance auditing, human verification checkpoints, logging, and drift surveillance. Transparent disclosure of these limitations is essential to prevent over-reliance and guide safe iterative improvements.

Ethical challenges associated with AI for OMS

This framework is intentionally non-exhaustive and mechanism-focused: individual controls (e.g., dataset diversification) reduce but do not eliminate inequity or safety risks and must be combined with ongoing auditing. The key ethical challenges and legal/regulatory considerations are summarized in Table 2, which focuses on current risk domains and layered mitigation approaches. The integration of AI into surgical practice introduces a spectrum of ethical considerations. Patient safety and well-being remain paramount[58]. A major concern is the accuracy and reliability of AI systems[59,60]. An AI model misdiagnosing an imaging finding or providing an erroneous treatment recommendation could directly lead to patient harm. Thus, ensuring that AI tools have been rigorously validated and are capable of performing at least as well as current standards of care is an ethical imperative. A recent review of ethical issues in surgical AI found that accuracy was the most discussed concern, underscoring the need for reliable AI outputs[2].

A primary challenge in this regard is the current lack of validated tools to assess the quality of health information provided by AI systems[61,62].

This ties into the ethical principle of non-maleficence—AI should not introduce new risks to patients. Thus, developers and clinicians should have an ethical obligation to verify AI performance and use it within its validated scope.

Transparency and explainability also remain central challenges[63,64]. Many advanced AI algorithms, particularly deep learning models, lack interpretable decision-making pathways. In healthcare, this opacity can erode trust: patients may be uneasy about AI-driven recommendations, and clinicians may be reluctant to adopt 'black-box' tools. From an ethical standpoint, autonomy and informed consent are implicated—patients have the right to understand how decisions about their care are made. If AI contributes to those decisions, should patients be informed of AI's role? Several publications argue that transparency with patients regarding AI involvement is necessary to respect patient autonomy and maintain trust in the patient–provider relationship[65–68]. Ethical AI in OMS should ideally be 'explainable AI', providing interpretable outputs or highlighting key factors influencing a prediction. This will enable surgeons to integrate AI advice responsibly, thereby safeguarding clinical judgement. The literature emphasizes the importance of balancing AI efficacy with explainability in healthcare innovation[63,64,69].

Beyond accuracy and high-level principles, several practice-facing obligations now define ethical AI use in surgery. First, post-deployment stewardship is an ethical duty: models may drift as scanners, populations, and workflows change, so continuous monitoring of discrimination, calibration, and failure modes—with rollback thresholds and versioned audit trails—should be routine[70–74]. Institutions should treat drift dashboards and periodic re-validation as safety infrastructure, not research extras.

Second, automation bias and human-factors risk require interface-level safeguards: clinician-in-the-loop verification, thresholded 'weak-signal' flags, and explanation artefacts tied to each recommendation reduce over-trust and support accountable override[63,64,67,68]. For LLM-assisted tasks (e.g., consent and patient education), OMS teams should pair human review with domain-specific evaluation instruments to mitigate hallucinations and omissions[60–62,75].

Third, data governance and consent for secondary use must move beyond generic broad consent: privacy-by-design, data minimization, federated/synthetic approaches to reduce raw transfers, and explicit handling of cross-border flows are ethically required in OMS datasets that mix imaging and clinical text[28–30,73,76–78]. Under GDPR, the right not to be subject to automated decision-making reinforces the need to document human oversight for any AI-influenced decision[79].

Fourth, equity demands more than 'diverse' training sets: centres should report subgroup performance (e.g., age, sex, ethnicity, skeletal class) and predefine remediation triggers for unacceptable gaps, then document corrective re-training steps[22,80–82].

Finally, clinician competence is an ethical precondition for safe deployment. OMS professionals should be trained to interpret probabilities and calibration plots, recognize bias or drift, and explain the role of AI to patients using readable materials and visual rationales. In this view, patient-facing explainability is part of the consent record, supported by versioned overlays and structured risk statements[27,69,75,83–85].

Bias and fairness issues in AI

AI systems inherit the biases present in their training data. In OMS, this could manifest as poorer diagnostic accuracy for underrepresented ethnic or age groups, thereby perpetuating disparities in care. Ethically, this conflicts with the principle of justice, which requires equitable care. The literature calls for ensuring diversity in training datasets, auditing algorithmic performance across subpopulations, and scrutinizing automated decisions (e.g., surgical prioritization) for unintended discrimination. Collaboration between AI developers, clinicians, and ethicists is essential to address these risks, and regulatory frameworks should require fairness auditing. Maintaining a human-in-the-loop approach—where clinicians validate AI outputs—remains a crucial safeguard[24,80–82,86].

Data privacy and confidentiality

Training powerful AI models often requires large datasets of patient information. This raises questions about how the data is collected, stored, and used. Patients typically consent to their data being used for care and scientific studies; however, secondary use for AI training may not always be fully disclosed. Ethically, it must be ensured that patient data are used with respect for their privacy and with robust safeguards. This includes de-identifying data to protect patient identity and following best practices in data security. Moreover, if AI systems are deployed in practice, they may handle personal health information (for instance, an AI analysing patient records to flag drug interactions). In such cases, maintaining confidentiality is critical; similar to any clinician, an AI must be bound by privacy principles. Consequently, the concept of data governance has emerged. It recommends that institutions have clear policies on data usage, with oversight committees to ensure ethical use of patient data in AI projects[29,65,66,76].

Informed consent in the context of AI

A new topic raised in several studies is "When a surgeon uses AI for decision support, to what extent should this be disclosed to the patient during the consent process?".

Some argue that if AI materially influences a diagnosis or treatment plan, patients should be informed as part of obtaining consent—similar to how patients are informed about trainees being involved in their care or novel techniques being used. Others suggest that as long as the surgeon remains the decision-maker, explicit consent for the involvement of AI may not be necessary, although transparency is encouraged. This remains an open ethical debate.

Informed consent for AI-enabled care should move beyond generic statements to a concise, task-specific disclosure covering the following: (1) the role of the tool (decision support versus execution), (2) its validation context (population/modality, external testing, and known limits), (3) material failure modes relevant to the indication (e.g., segmentation errors in metal artefact regions; false-negative small periapical lesions), (4) the degree of human oversight/override, and (5) data use and updating (what data are processed; whether/when the model is updated and how that is governed). These elements align with health-AI governance principles and high-risk AI duties discussed

for healthcare under the EU AIA and related analyses[70–73,79,87].

Consent should be proportional to risk. For decision-support algorithms (e.g., CBCT segmentation, OPMD triage, relapse prediction), a brief, standardized disclosure within routine consent is appropriate, provided that the surgeon remains the final decision-maker and verifies outputs. For (semi-) autonomous functions (e.g., robotic osteotomies), consent should be more detailed, including the robot's task boundaries, safety interlocks, override protocols, and allocation of responsibility (manufacturer product responsibility for autonomous execution versus clinician duty of care for indication, planning, and supervision)[15,20,66,88].

Documentation should capture the AI tool name, version/build, date/time of inference, and, where available, a rationale artefact (e.g., saliency map or structured probability statement) archived with the consent or operative note to create an audit trail[23–25].

Regarding data use and privacy, disclosures should state what patient data are analysed, retention/sharing policies, and, where applicable, cross-border transfer safeguards, consistent with GDPR/HIPAA-oriented guidance[73,76–78,89]. For continuous-learning systems, patients should be told that model parameters may be updated under an approved algorithm-change/control plan with post-market monitoring[23,24,28,71,72].

Although high-quality patient-level evidence is limited, early studies suggest that AI-assisted consent materials, for example LLM drafts reviewed by clinicians, can improve readability and patient understanding without sacrificing completeness; clinician oversight remains essential to ensure case-specific accuracy and legal adequacy[69,75,83–85,90]. These tools should augment and not replace the clinician–patient discussion.

For patients with limited capacity or language barriers, consent artefacts should target appropriate reading levels and be interpreter-mediated; for paediatric/young adult cases, assent principles apply alongside parental permission. For foundation or multimodal models incorporated into clinical systems, sites should specify governance, monitoring for drift/bias, and routes for patient questions or objection where feasible[70,71,87].

A pragmatic OMS template (to be adapted locally) would be: "Your images/data were analysed with an AI-based tool that helps your surgical team [task]. The tool has been validated for [population/modality] with typical performance of [metric(s)], but may be less reliable when [known limits/failure modes]. Your surgeon reviews and may override AI outputs. Your data are handled under [policy], and the software is [static/periodically updated] under regulatory oversight. Please ask if you want more detail or prefer not to use AI assistance for elements where that is feasible."

Embedding these structured, proportionate disclosures within routine OMS consents could align ethical duties of disclosure and comprehension with concrete technical design and governance choices, while avoiding both under- and over-disclosure that could erode trust or impede care.

In OMS, these principles can be translated into auditable safeguards: document external validation on local data with model-card details (version, data provenance), report subgroup performance and calibration, maintain AI registries with versioned inference/audit logs, and use proportionate, task-specific consent that clearly states the role of the tool, known limits, human oversight, and update policy. These safeguards align with contemporary reporting and governance frameworks (TRIPOD-AI/TRIPOD-LLM; CHAI Blueprint)[74,87,91] and with EU duties for high-risk health AI[71]. For LLM-assisted consent materials specifically, recent peer-reviewed studies in OMS and allied surgery show improved readability with non-inferior factual completeness when clinician oversight is retained. In addition, oversight responsibilities differ for decision-support tools used by surgeons versus embedded AI within robotics and navigation systems, and should be addressed distinctly.

Impact on the clinician–patient relationship

If AI assumes more diagnostic and planning tasks, there is a risk that clinicians may become over-reliant on these tools, potentially reducing engagement in cognitive and communicative processes central to care. Ethically, concerns have been raised about potential deskilling and erosion of the human touch in medicine[21]. For example, relying on AI for prognostication could diminish opportunities for empathetic, individualized discussions with patients. Preserving the art of medicine—empathy, compassion, and holistic judgement—remains a critical ethical priority. Incorporating AI-related training into surgical education is recommended to ensure practitioners understand the strengths and limitations of AI, thereby upholding professional competence and ethical standards in the era of AI[85,92,93]. In summary, the ethical challenges of AI in OMS revolve around maintaining trust, safety, fairness, autonomy, and data integrity. Proactive measures—rigorous validation and monitoring, transparency and explainability, bias mitigation, robust privacy safeguards, and clinician and patient education—are essential to ensure that AI integration aligns with core principles of medical ethics.

Legal and administrative integration

The rapid emergence of AI in healthcare has outpaced the development of laws and regulations in many respects. Legal and regulatory discussions, while fewer than clinical or ethical discussions, are gaining momentum as the need for clear governance of AI in medicine is increasingly recognized. Several key legal considerations for AI in OMS (and healthcare more broadly) have been identified: regulatory approval and oversight; liability and malpractice; data protection and privacy laws; cybersecurity.

Regarding regulatory approval, before AI-based tools can be widely used in clinical practice, they typically require regulatory clearance. In the United States, the FDA is responsible for this oversight. In 2019, the FDA released a discussion paper proposing a "predetermined change control plan" for AI software, outlining how AI could continue to learn post-approval while maintaining safety[23]. By 2022, documentation indicated that dozens of AI-enabled medical devices had been cleared by the FDA, reflecting an 'updated landscape' of approved AI tools, including dental/OMS-related AI software such as radiographic diagnostic aids. Regulatory science is gradually catching up, with authorities establishing requirements for validation datasets, performance metrics, and post-market surveillance[70]. Similarly, in the EU, medical AI falls under the EU Medical Device Regulation (MDR) and the upcoming AIA. The AIA, approved by the European Parliament in

2024, specifically addresses AI in high-risk domains such as healthcare, and emphasizes requirements for trustworthy AI, including risk assessment, transparency, human oversight, and robustness of AI systems. Moreover, the AIA will impose legal obligations on developers and users of AI in OMS, such as mandatory documentation and performance monitoring. For clinicians and hospitals, maintaining compliance with these evolving regulations will be crucial. Regulatory oversight ultimately aims to ensure that any AI tool used in patient care has demonstrated safety and effectiveness and continues to do so throughout its use[71–73].

A major legal question concerns how liability is determined when an AI system contributes to an adverse outcome. At present, AI in clinical practice is generally considered a decision-support tool, with the clinician remaining the final decision-maker. Thus, if an error occurs (e.g., a misdiagnosis or inappropriate treatment decision), the surgeon would typically bear responsibility in a malpractice claim, even if they relied on AI recommendations. The current review noted legal analyses suggesting that since advanced AI is not yet the standard of care, clinicians should use AI as an aid rather than a sole decision-maker to reduce liability[88,94–98]. In other words, surgeons must exercise independent judgement. However, as AI becomes more prevalent and potentially standard in practice, the question arises: If a surgeon follows a reputable AI recommendation and harm still occurs, does liability fall on the clinician or could it be considered a 'product defect' in the AI? Some have proposed applying product liability to AI systems, placing responsibility on manufacturers or developers when an AI device is proven faulty, although such claims are challenging under current law. Another consideration is the liability of healthcare institutions: hospitals deploying AI could face claims if they fail to properly vet tools or adequately train staff.

In practice, liability should track the AI role. For decision-support uses, surgeons retain primary professional responsibility and must document independent clinical reasoning, validated indications for use, and verification of outputs. For autonomous or semi-autonomous functions, manufacturers bear product-safety obligations within the cleared indications, while institutions assume duties to vet, credential, and monitor deployment. Post-market assurance requires shared mechanisms: versioned model/change-control logs, inference-level audit trails to support causation analysis, and explicit vendor–institution indemnity arrangements. These steps are consistent with evolving EU requirements and with analyses of AI-related malpractice and product liability in clinical care[71–73,88,94,96–98].

For embedded AI in robotics and navigation, governance must reflect device-level obligations alongside clinical oversight. Before deployment, sites should verify indications for use, performance on local cases, fail-safe behaviours, and human-factors checks (credentialing of operators, simulation drills, and stop/override procedures). During use, surgeon verification remains mandatory for critical steps, but the manufacturer and institution share duties for safety monitoring (update/change-control documentation, cybersecurity patches, uptime and fault logs) and for post-market surveillance. Liability generally follows function: decision-support remains clinician-led, whereas device-embedded AI engages product-safety duties for manufacturers and institutional duties for safe deployment and monitoring; all parties should maintain auditable logs that support causation analysis and corrective actions.

Overall, the legal community is actively debating these issues. For now, a practical consensus suggests the need for clear clinical guidelines on integrating AI into care—for example, professional societies might specify that AI should support second opinions or triage but that final diagnoses must be confirmed by a specialist. Proposed OMS-specific guidance could adopt a tiered liability model: First, defining when an AI tool operates as an 'autonomous device' (e.g., a robot executing bone cuts) versus 'decision support' (e.g., a risk-prediction algorithm). In the former case, primary product liability would rest with the manufacturer, similar to existing rules for surgical robots[15,20]. In the latter, liability would remain with the surgeon, provided the software was used within its indications and with reasonable oversight. Second, the guidelines should require algorithmic audit trails (version control, decision logs) to trace causation in the event of harm[88]. Third, they should recommend that hospitals establish shared indemnity agreements with vendors, mirroring practices in radiology AI, to distribute financial risk and encourage adoption[96]. Finally, these guidelines should align with emerging EU AIA provisions that place post-market monitoring obligations on manufacturers while allowing professional bodies to issue specialty-level 'codes of practice' clarifying clinician responsibilities in day-to-day use.

Regarding data protection, using AI in OMS often involves large-scale data, which must comply with strict privacy regulations. In the United States, HIPAA governs personal health data usage, while in Europe, GDPR sets similar stringent standards. When patient data is used to train AI algorithms, developers must ensure proper consent or legal justification and secure data storage. GDPR, for instance, includes provisions on automated decision-making and profiling, potentially applying to AI systems that significantly influence patient care. Under its 'right to an explanation', patients may eventually gain rights to understand the logic behind algorithmic decisions in their care. Authors of articles included in this review emphasized that privacy laws guide the ethical design of AI (e.g., privacy-by-design approaches) and require institutions to conduct privacy impact assessments. Non-compliance can lead to legal penalties and erode public trust. Thus, legal risk management for AI in OMS necessitates strong data governance, anonymization techniques, and, where feasible, the use of synthetic data for training models without exposing real patient information[77–79,89,99].

Another important legal consideration is ensuring robust cybersecurity for AI systems. As healthcare becomes increasingly digitized and AI systems are network-connected, they become potential targets for cyberattacks. A malicious alteration to an AI's functioning—such as hacking a surgical robot or diagnostic tool—could directly endanger patients. Legal standards are evolving to mandate that medical device software (including AI) incorporates adequate safeguards against unauthorized access[100–102]. Manufacturers could be held liable if inadequate cybersecurity results in patient harm. Hospitals also face legal risk if they fail to implement proper IT security for deployed AI tools. This area intersects with broader health IT regulations and remains a significant concern, often addressed alongside privacy in governance frameworks.

To conclude, from the results reported above, it is evident that

alongside the impressive capabilities of AI in OMS, navigating a complex ethical and legal landscape is equally crucial. The following Discussion section interprets these findings, explores their implications, and suggests pathways for future developments in research, practice, and policy.

Discussion

This scoping review provides a comprehensive analysis of the integration of AI into OMS, while also addressing the ethical and legal considerations that must be navigated. Clinically, the included studies reveal a growing interest in the deployment of AI for diagnostic imaging, surgical planning, robotics, and patient monitoring. Although many reports highlight early performance gains, particularly in enhancing diagnostic accuracy and reducing processing time, most remain limited to pilot studies or retrospective analyses[14,70]. This raises concerns about their external validity and real-world applicability. A substantial body of literature underscores the need for prospective clinical trials that account for patient variability, diverse clinical workflows, and potential biases in AI training data[103]. The transition from small-scale validation to large, multi-institutional studies is essential to ensure that AI meets the rigorous standards of precision and reliability required in OMS.

Implications for clinical practice and recommendations

The review findings support a cautious, stepwise approach to integrating AI into OMS. Initially, AI-driven diagnostic tools should function as adjuncts to traditional methods[70,104]. As their reliability is further validated, their role can be expanded progressively to include preoperative planning and intraoperative guidance[43,51,52]. To facilitate safe and effective implementation, the development of standardized clinical guidelines and structured AI training modules for both residents and practising surgeons is crucial[105].

Knowledge gaps and future research needs

Despite early, largely preliminary data, several critical gaps remain: (1) a significant lack of large-scale, prospective RCTs evaluating the clinical efficacy and safety of AI tools in OMS; (2) an absence of standardized protocols for assessing AI performance, hindering reliable comparisons across studies; (3) underdeveloped ethical strategies to ensure transparency, mitigate bias, and secure informed consent; (4) legal uncertainties surrounding liability and the need for robust frameworks governing data security, privacy, and cybersecurity in AI-driven OMS applications.

Ethical and legal considerations

Beyond clinical performance, ethical concerns remain paramount. Data privacy and patient confidentiality are particularly critical when sensitive medical information is aggregated for AI training[29,76,78]. Issues surrounding data anonymization, secondary data use, and governance require transparent protocols to maintain patient trust. The 'black-box' nature of many AI models also raises concerns for both surgeons and patients, underscoring the need for interpretable AI systems that support informed clinical decision-making[63,80]. Preventing biases from non-representative training datasets is equally essential to ensure equitable care and avoid disparities in treatment outcomes[24,80,81].

Parallel to these ethical challenges, legal and regulatory frameworks for AI in healthcare remain dynamic and often incomplete. Regulatory bodies such as the US FDA and European legislative institutions have outlined requirements for AI-based medical devices, but significant gaps persist, particularly concerning liability in adverse events and post-market surveillance[23,25]. To navigate these uncertainties, careful documentation of AI use in clinical practice and adherence to professional standards are prudent strategies for mitigating legal risk.

Implications for policy

AI technologies are moving from proof-of-concept studies into routine OMS workflows at a pace exceeding the specialty's established governance structure. A coherent policy response must rest on two pillars already articulated internationally: transparent reporting, captured by the EQUATOR Network's new TRIPOD-AI checklist[74], and ethical–legal assurance, codified in the Coalition for Health AI (CHAI) Blueprint for Trustworthy AI in Healthcare[87]. Read together, these frameworks offer a near-complete roadmap for safe, accountable, and equitable adoption—adaptable with modest effort to the realities of maxillofacial practice.

The EQUATOR Network has long provided structured guidance for scientific reporting, but its 2024 extension, TRIPOD-AI[74], finally brings machine-learning studies under the same discipline that has governed randomized trials and observational research for decades. Each of its 27 items addresses recurring weaknesses that hamper appraisal of medical AI: vague data provenance, absent calibration statistics, performance claims based on single-centre test sets, and insufficient detail on post-deployment model updates. These gaps are particularly evident in the maxillofacial literature, where many image-analysis papers report only accuracy or Dice scores and omit external validation[106]. By requiring disclosure of data sources, version numbers, and calibration plots, TRIPOD-AI allows editors and reviewers to distinguish robust algorithms from proof-of-concept demonstrations, placing maxillofacial imaging research on a footing comparable with other evidence hierarchies. Importantly, TRIPOD-AI is complemented by sister extensions—TRIPOD-LLM for large language models[91], MINIMAR for medical imaging[107], and CLAIM for radiology AI[106]—covering nearly every tool now in development for cephalometric prediction, tumour segmentation, or automated consent generation. A journal policy mandating the relevant checklist at submission would not be an administrative burden but a catalyst for embedding reproducibility into model design.

However, transparency alone cannot guarantee that a system remains safe and fair once it leaves the laboratory. That responsibility is addressed in the CHAI Blueprint, published after a multistakeholder process involving regulators, academia, industry, and patient groups. The Blueprint distils AI ethics discourse into seven actionable principles: usefulness, safety, accountability, transparency, fairness, security, and privacy, and couples them to a four-level assurance label[87]. Level I prototypes may be explored only in silico; level II systems can inform decisions under explicit human supervision; level III tools integrate into workflows provided that mechanisms for

monitoring drift and bias exist; level IV systems are autonomous and require the most stringent oversight. Mapping typical OMS applications onto this schema is straightforward: a retrospective convolutional network flagging mandibular fractures on panoramic radiographs would be level I or II[11,12,40], whereas an autonomous robotic osteotomy platform executing pre-planned cuts with sub-millimetre precision is level IV and subject to continuous post-market audit. The Blueprint thus provides the procedural detail that turns abstract ethical aspirations into measurable obligations for manufacturers, hospitals, and surgeons.

Together, these frameworks suggest policy priorities specific to maxillofacial surgery. First, editors, grant reviewers, and ethics committees should reject diagnostic or prognostic models without a completed TRIPOD-AI (or TRIPOD-LLM) checklist[74,91]; such gatekeeping will not stifle innovation but enforce clarity on external validation, calibration, and dataset diversity, factors critical to trust at the bedside. Second, hospitals adopting CHAI level III or IV systems should establish AI safety committees to approve devices, monitor performance, and commission root-cause analyses when error rates rise[20,56,92]. These deliberations are most effective when manufacturers provide continuous, de-identified logs, as recommended in the Blueprint, enabling early detection of drift. Third, professional societies should translate medico-legal scholarship into template contracts allocating product liability to vendors for autonomous functions while clarifying that surgeons retain a duty of care for decision-support systems[20]. Without shared liability, fear of litigation will remain a brake on adoption[20,88,94]. Fourth, fairness auditing must be taken as seriously as dosimetry or sterility: models should report performance across age, sex, ethnicity, and skeletal pattern strata, with retraining triggered when subgroup disparities exceed clinically acceptable margins[61,86]. Finally, sustainable trust depends on education. Residency curricula need core modules in AI literacy—how models are trained, how bias arises, what calibration means—and practising surgeons should be recertified when major software updates are introduced, mirroring current expectations for radiation safety training[85].

Even with these safeguards, the evidence base will evolve at software speed. Thus, the specialty should commit to 'living' systematic reviews—updated at fixed intervals rather than every 5 or 10 years—and require revalidation of deployed models on contemporary data. This vigilance has modest costs compared with the risks of silent performance decay in systems guiding high-dose imaging, implant placement, or nerve-sparing resections.

While framed in maxillofacial vocabulary, this approach is exportable. Orthopaedic navigation, cardiac mapping, and neurosurgical targeting face the same tension between innovation and oversight, and the same need to reconcile transparency with privacy[15,29]. Yet OMS provides a particularly stringent test case: procedures hinge on geometry measured in fractions of a millimetre, and emerging tools—LLMs drafting consent documents or robots milling bone—operate at informational and physical extremes[6,8]. The marriage of the reporting discipline of TRIPOD-AI with the continuous-assurance culture of CHAI is not redundant bureaucracy but a prerequisite for ethical progress. Together, they chart a course for embracing algorithmic assistance while upholding commitments to safety, accountability, and equitable care.

Framework for phased integration of AI in OMS

AI systems rarely move directly from algorithmic potential to reliable clinical performance; they mature through stages requiring different evidentiary standards and governance controls. Drawing on the transparent-reporting discipline of TRIPOD-AI[74] and the progressive assurance ladder articulated by the CHAI Blueprint[87], a four-phase pathway tailored to OMS practice is proposed, covering the preclinical development phase, supervised clinical pilots, controlled integration, and routine deployment under adaptive governance.

In the preclinical development phase, models are trained and validated exclusively on retrospective data. Transparency is paramount: investigators must disclose data provenance, calibration statistics, and versioning per TRIPOD-AI, and make code or detailed parameters available for peer evaluation. No patient-facing use is permitted until these requirements are met and external validity is demonstrated.

Completion of these benchmarks justifies supervised clinical pilots. Here, algorithms act as second readers; outputs are logged alongside surgeon judgements but exert no direct influence on operative plans. Prospective comparison of human and AI decisions generates real-world performance data, while institutional AI-safety committees review deviations from pre-specified error thresholds. These deployments align with the 'pre-deployment' testing recommended in the DECIDE-AI framework for clinical evaluation[104].

Once equivalence—or superiority—has been demonstrated under supervision, systems may enter controlled integration. Here, the algorithm informs decisions in limited workflows, with recommendations linked to peri- and postoperative outcomes for audit. Biannual public reports document accuracy, calibration drift, and emerging bias, satisfying CHAI level III monitoring. Examples include automated cephalometric landmarking in virtual planning software or CBCT-based fracture detection in emergency pathways.

The final stage, routine deployment under adaptive governance, applies to mature tools meeting CHAI level IV criteria, such as autonomous robotic osteotomy platforms[108]. At this point, shared-liability agreements are essential: manufacturers bear product liability for autonomous functions, while surgeons retain professional responsibility for decision-support outputs, echoing medico-legal recommendations[20]. Real-time performance logging, drift alarms, and rapid-rollback protocols are mandatory safeguards; annual recertification of both software and clinical teams ensures that technological evolution does not outstrip human competence.

Two cross-cutting obligations span all phases. First, AI literacy. Residency curricula and continuing-education programmes must equip surgeons to interpret model output, identify bias, and balance algorithmic advice with clinical judgement. Second, because algorithmic performance can decay in the face of shifting data, specialty societies should commission living systematic reviews that update the evidence base at fixed intervals[109,110]. By embedding continuous education and dynamic appraisal into the lifecycle of each

system, the OMS community can transform ethical aspiration into operational reality. This will convert the potential of AI into predictable, accountable benefit for patients.

Model of care for AI integration in OMS

The authors propose a phased, human-in-the-loop pathway tailored to surgical services: (1) governance and readiness: register each tool with indications/contraindications, dataset provenance, subgroup performance, calibration, cybersecurity posture, and site-specific consent language; (2) pre-deployment validation: test retrospectively on local cases versus clinician benchmarks, report discrimination and calibration by subgroup, and set go/no-go thresholds; (3) supervised deployment: use as second reader or planning adjunct with mandatory surgeon verification, automatic inference/version logging, and structured override documentation; (4) continuous monitoring: quarterly performance/bias audits with drift tests, near-miss reviews, and documented remediation; (5) change control: classify model updates, re-train/rollback as needed, and re-credential users after major updates. Informed-consent materials may leverage LLM-assisted drafts to improve readability while preserving surgeon authorship and case-specific accuracy, as shown in recent OMS and allied-surgery evaluations; dataset transparency and licensing clarity, particularly for CBCT/CT repositories, are prerequisites for fair and reproducible performance[26,71–73,79,82,84,87,91,100–102].

Oversight and responsibilities by AI modality (decision-support versus embedded systems)

Decision-support AI (diagnostics, surgical planning, monitoring) should operate under human-in-the-loop controls: pre-deployment local validation against clinician benchmarks; documented calibration and subgroup performance; mandatory surgeon verification for outputs that affect diagnosis, resection margins, flap choices, or postoperative surveillance; and inference-level audit logging within an institutional AI registry[71–73,79]. Embedded AI in robotics/navigation requires additional device-centric assurances: operator credentialing and simulation, verification of indications for use, defined override/stop protocols, continuous device-log monitoring (faults, drifts, patches), and formal change-control for software/model updates[23,74,91,104]. Institutions should maintain a governance pathway that maps responsibilities among surgeon, service line, biomedical engineering/IT, and vendor, including incident reporting and indemnity arrangements. This division of responsibilities ensures that ethics, reliability, bias controls, and data governance are implemented proportionately to the AI modality, while clarifying liability across surgeon, manufacturer, and institution[88,94,96,97].

Implications for practice

AI tools are ready to serve as adjuncts that enhance, not replace, clinical judgement. Surgeons should integrate validated systems into diagnostic workflows and preoperative planning while retaining final decision authority and documenting AI use in the record. Institutions must pair deployment with staff training, robust data-security infrastructure, and clear governance policies.

Next steps for research and policy can be summarized as follows: (1) prospective, multi-institutional studies—including randomized and real-world effectiveness trials—to confirm clinical benefit, generalizability, and economic value; (2) standardized reporting and benchmarking of datasets, performance metrics, and validation protocols to enable transparent comparison across studies; (3) explainable and bias-audited models to strengthen trust and ensure equitable performance across patient subgroups; (4) interdisciplinary guidelines that codify best practices for consent, documentation, and liability sharing, harmonized with emerging FDA and EU requirements; (5) continuous post-market surveillance for safety, drift, and cybersecurity threats, ideally embedded in learning-health-system infrastructures.

By addressing these priorities, the OMS community can translate AI innovations from proof-of-concept to routine, ethically grounded, and legally robust clinical practice, ultimately improving diagnostic precision, surgical accuracy, and patient outcomes.

Strengths and limitations of this review

This scoping review has several strengths. First, it is novel in mapping clinical applications, ethical challenges, and legal-regulatory aspects of AI specifically in OMS. By adopting this interdisciplinary lens, it provides a holistic panorama beyond purely technical evaluations. Second, the methodology closely followed the PRISMA-ScR framework, employing a prospectively designed search strategy, predefined eligibility criteria, duplicate screening, and standardized data extraction forms—enhancing transparency, reproducibility, and methodological rigour. Third, the search spanned three multidisciplinary databases (PubMed/MEDLINE, Scopus, and Web of Science) and included backward citation tracking, reducing the risk of omitting relevant studies. Finally, by synthesizing evidence across diverse study designs and highlighting research gaps, the review generates a practical roadmap for clinicians, researchers, and policymakers—facilitating targeted investigations and informed AI governance in OMS.

This review also has limitations. First, as a scoping review, it maps evidence breadth rather than critically appraising study quality or providing a quantitative synthesis; conclusions are therefore bounded by the quality of underlying studies. Second, the literature is dominated by exploratory, retrospective, or pilot investigations with small samples and heterogeneous methodologies, limiting comparability and generalizability. Third, the search was confined to English-language articles indexed in PubMed, Scopus, and Web of Science and published between January 2015 and July 2025. Hence, non-English publications, the grey literature, and studies in other databases may have been excluded, introducing publication and language bias. Fourth, focusing on articles published between January 2015 and July 2025 captures the deep-learning era during which almost all clinically relevant maxillofacial AI systems emerged, while excluding earlier, methodologically obsolete rule-based work. Meta-research shows that non-English studies can yield different effect estimates and often reflect research undertaken in under-represented regions[111]. Moreover, 87% of AI studies screened originated from high-income countries, mirroring broader analyses that highlight a paucity of evidence from low- and middle-income settings; this geographic imbalance limits the direct transferability of the findings to resource-constrained environments[81,82]. However, restricting

the start date to 2015 is appropriate because earlier work predates modern deep-learning methods and contributes little to current clinical practice. Fifth, outcome measures and reporting standards varied widely, impeding direct comparison and precluding meta-analysis. Finally, recognizing that rapid innovation will have continued to produce new models and trials after the July 2025 search cut-off, the review was structured around functional domains (rather than exhaustive topic catalogues), and the authors commit to scheduled literature updates, so that emerging high-quality evidence can be incorporated without re-expanding into breadth-only summaries. As AI research is evolving at an unprecedented rate and more than 3000 new PubMed entries on 'medical AI' were indexed in the first half of 2025 alone, the current synthesis may become outdated rapidly. A living-review methodology, with scheduled literature updates every 12 months, is therefore recommended to maintain relevance.

Consistent with the methodology for scoping reviews and PRISMA-ScR guidance, no formal risk-of-bias or GRADE certainty appraisals were performed; given the heterogeneity and paucity of comparable effect estimates across studies, such grading would be premature. Future domain-specific systematic reviews with homogeneous outcomes should incorporate risk of bias (RoB) and GRADE tools to quantify certainty and support practice recommendations.

While AI may advance OMS meaningfully—improving diagnostic accuracy, surgical precision, and patient management—its integration must be gradual, evidence-based, and rigorously regulated. The establishment of standardized guidelines, continuous quality monitoring, and robust interdisciplinary collaboration among clinicians, data scientists, ethicists, and legal experts will be essential to ensure responsible and safe AI-driven innovations. Addressing the identified knowledge gaps through future research will be critical for advancing the safe and effective clinical adoption of AI in OMS.

The aim of this review was to map the clinical applications of AI in OMS and identify the parallel ethical-legal landscape shaping its safe adoption. The 99 studies included, published between January 2015 and July 2025, revealed a field advancing quickly in diagnostic imaging and steadily branching into surgical planning, robotics, and outcome prediction. Across these domains, AI systems often match or exceed expert-level accuracy and can reduce operative error, yet the evidence base remains dominated by small, retrospective, single-centre investigations; prospective, multi-centre trials and cost-effectiveness analyses are still lacking.

Ethical approval

Not required.

Patient consent

Not applicable.

Funding

Dr Luigi Angelo Vaira received funding from the project e.INS – Ecosystem of Innovation for Next Generation Sardinia (code ECS00000038), funded by the Italian Ministry for University and Research (MUR) under the National Recovery and Resilience Plan (NRRP) – Mission 4, Component 2, "From Research to Business", Investment 1.5 "Creation and Strengthening of Innovation Ecosystems" and the development of "Territorial R&D Leaders", CUP J83C21000320007.

Competing interests

None.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ijom.2026.02.020.

References

1. Bajwa J, Munir U, Nori A, Williams B. Artificial intelligence in healthcare: transforming the practice of medicine. *Future Healthc J* 2021;**8**:e188–94. https://doi.org/10.7861/fhj.2021-0095
2. Mithany RH, Aslam S, Abdallah S, Abdelmaseeh M, Gerges F, Mohamed MS, Manasseh M, Wanees A, Shahid MH, Khalil MS, Daniel N. Advancements and challenges in the application of artificial intelligence in surgical arena: a literature review. *Cureus* 2023;**15**:e47924. https://doi.org/10.7759/cureus.47924
3. Guni A, Varma P, Zhang J, Fehervari M, Ashrafian H. Artificial intelligence in surgery: the future is now. *Eur Surg Res* 2024. https://doi.org/10.1159/000536393
4. Qiu P, Cao R, Li Z, Huang J, Zhang H, Zhang X. Applications of artificial intelligence for surgical extraction in stomatology: a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2024;**138**:346–61. https://doi.org/10.1016/j.oooo.2024.05.002
5. Dong F, Yan J, Zhang X, Zhang Y, Liu D, Pan X, Xue L, Liu Y. Artificial intelligence-based predictive model for guidance on treatment strategy selection in oral and maxillofacial surgery. *Heliyon* 2024;**10**:e35742. https://doi.org/10.1016/j.heliyon.2024.e35742
6. Bouletreau P, Makaremi M, Ibrahim B, Louvrier A, Sigaux N. Artificial intelligence: applications in orthognathic surgery. *J Stomatol Oral Maxillofac Surg* 2019;**120**:347–54. https://doi.org/10.1016/j.jormas.2019.06.001
7. Puladi B, Gsaxner C, Kleesiek J, Hölzle F, Röhrig R, Egger J. The impact and opportunities of large language models like ChatGPT in oral and maxillofacial surgery: a narrative review. *Int J Oral Maxillofac Surg* 2024;**53**:78–88. https://doi.org/10.1016/j.ijom.2023.09.005
8. Sillmann YM, Monteiro JLGC, Eber P, Baggio AMP, Peacock ZS, Guastaldi FPS. Empowering surgeons: will artificial intelligence change oral and maxillofacial surgery? *Int J Oral Maxillofac Surg* 2025;**54**:179–90. https://doi.org/10.1016/j.ijom.2024.09.004
9. Zirek T, Öziç MÜ, Tassoker M. AI-driven localization of all impacted teeth and prediction of winter angulation for third molars on panoramic radiographs: clinical user interface design. *Comput Biol Med* 2024;**178**:108755. https://doi.org/10.1016/j.compbiomed.2024.108755
10. Hartoonian S, Hosseini M, Yousefi I, Mahdian M, Ghazizadeh Ahsaie M. Applications of artificial intelligence in dentomaxillofacial imaging—a systematic review. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2024;**138**:641–55. https://doi.org/10.1016/j.oooo.2023.12.790
11. Warin K, Limprasert W, Suebnukarn S, Inglam S, Jantana P, Vicharueang S. Assessment of deep convolutional neural network models for mandibular fracture detection in panoramic radiographs. *Int J Oral Maxillofac Surg* 2022;**51**:1488–94. https://doi.org/10.1016/j.ijom.2022.03.056
12. Yan Y, Lv C, Wang B, Wang X, Han W, Sun M, Kim BS, Zhang Y, Bao J, Lin L, Chai G. Applying artificial intelligence algorithm in the design of a guide plate for

mandibular angle ostectomy. *J Plast Reconstr Aesthet Surg* 2023;**84**:595–604. https://doi.org/10.1016/j.bjps.2023.05.030
13. Mohaideen K, Negi A, Verma DK, Kumar N, Sennimalai K, Negi A. Applications of artificial intelligence and machine learning in orthognathic surgery: a scoping review. *J Stomatol Oral Maxillofac Surg* 2022;**123**:e962–72. https://doi.org/10.1016/j.jormas.2022.06.027
14. Birkhoff DC, van Dalen ASHM, Schijven MP. A review on the current applications of artificial intelligence in the operating room. *Surg Innov* 2021;**28**:611–9. https://doi.org/10.1177/1553350621996961
15. Hernigou P, Lustig S, Caton J. Artificial intelligence and robots like us (surgeons) for people like you (patients): toward a new human–robot-surgery shared experience. What is the moral and legal status of robots and surgeons in the operating room? *Int Orthop* 2023;**47**:289–94. https://doi.org/10.1007/s00264-023-05690-4
16. Ferreres AR. Ethical and legal issues regarding artificial intelligence (AI) and management of surgical data. *Eur J Surg Oncol* 2025;**51**:108279. https://doi.org/10.1016/j.ejso.2024.108279
17. Karimian G, Petelos E, Evers SMAA. The ethical issues of the application of artificial intelligence in healthcare: a systematic scoping review. *AI Ethics* 2022;**2**:539–51. https://doi.org/10.1007/s43681-021-00131-7
18. McGreevey JD, Hanson CW, Koppel R. Clinical, legal, and ethical aspects of artificial intelligence–assisted conversational agents in health care. *JAMA* 2020;**324**:552. https://doi.org/10.1001/jama.2020.2724
19. Sablone S, Bellino M, Cardinale AN, Esposito M, Sessa F, Salerno M. Artificial intelligence in healthcare: an Italian perspective on ethical and medico-legal implications. *Front Med* 2024;**11**:1343456. https://doi.org/10.3389/fmed.2024.1343456
20. Maliha G, Gerke S, Cohen IG, Parikh RB. Artificial intelligence and liability in medicine: balancing safety and innovation. *Milbank Quarterly* 2021;**99**:629–47. https://doi.org/10.1111/1468-0009.12504
21. Naik N, Hameed BMZ, Shetty DK, Swain D, Shah M, Paul R, Aggarwal K, Ibrahim S, Patil V, Smriti K, Shetty S, Rai BP, Chlosta P, Somani BK. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? *Front Surg* 2022;**9**:862322. https://doi.org/10.3389/fsurg.2022.862322
22. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;**366**:447–53. https://doi.org/10.1126/science.aax2342
23. Smith JA, Abhari RE, Hussain Z, Heneghan C, Collins GS, Carr AJ. Industry ties and evidence in public comments on the FDA framework for modifications to artificial intelligence/machine learning-based medical devices: a cross sectional study. *BMJ Open* 2020;**10**:e039969. https://doi.org/10.1136/bmjopen-2020-039969
24. Gilbert S, Fenech M, Hirsch M, Upadhyay S, Biasiucci A, Starlinger J. Algorithm change protocols in the regulation of adaptive machine learning-based medical devices. *J Med Internet Res* 2021;**23**:e30545. https://doi.org/10.2196/30545
25. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *NPJ Digit Med* 2020;**3**:118. https://doi.org/10.1038/s41746-020-00324-0
26. Cancela-Outeda C. The EU's AI act: a framework for collaborative governance. *Internet Things* 2024;**27**:101291. https://doi.org/10.1016/j.iot.2024.101291
27. Weidener L, Fischer M. Role of ethics in developing AI-based applications in medicine: insights from expert interviews and discussion of implications. *JMIR AI* 2024;**3**:e51204. https://doi.org/10.2196/51204
28. Bradford L, Aboy M, Liddell K. International transfers of health data between the EU and USA: a sector-specific approach for the USA to ensure an 'adequate' level of protection. *J Law Biosci* 2020;**7**:lsaa055. https://doi.org/10.1093/jlb/lsaa055
29. Yadav N, Pandey S, Gupta A, Dudani P, Gupta S, Rangarajan K. Data privacy in healthcare: in the era of artificial intelligence. *Indian Dermatol Online J* 2023;**14**:788–92. https://doi.org/10.4103/idoj.idoj_543_23
30. Tschider C, Compagnucci MC, Minssen T. The new EU–US data protection framework's implications for healthcare. *J Law Biosci* 2024;**11**:lsae022. https://doi.org/10.1093/jlb/lsae022
31. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, Moher D, Peters MDJ, Horsley T, Weeks L, Hempel S, Akl EA, Chang C, McGowan J, Stewart L, Hartling L, Aldcroft A, Wilson MG, Garritty C, Lewin S, Godfrey CM, Macdonald MT, Langlois EV, Soares-Weiser K, Moriarty J, Clifford T, Tunçalp Ö, Straus SE. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med* 2018;**169**:467–73. https://doi.org/10.7326/M18-0850
32. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;**8**:19–32. https://doi.org/10.1080/1364557032000119616
33. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci* 2010;**5**:69. https://doi.org/10.1186/1748-5908-5-69
34. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;**521**:436–44. https://doi.org/10.1038/nature14539
35. Scherer RW, Saldanha IJ. How should systematic reviewers handle conference abstracts? A view from the trenches. *Syst Rev* 2019;**8**:264. https://doi.org/10.1186/s13643-019-1188-0
36. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database Syst Rev* 2007;**18**:MR000005. https://doi.org/10.1002/14651858.MR000005.pub3
37. Orhan K, Bilgir E, Bayrakdar IS, Ezhov M, Gusarev M, Shumilov E. Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans. *J Stomatol Oral Maxillofac Surg* 2021;**122**:333–7. https://doi.org/10.1016/j.jormas.2020.12.006
38. Kirnbauer B, Hadzic A, Jakse N, Bischof H, Stern D. Automatic detection of periapical osteolytic lesions on cone-beam computed tomography using deep convolutional neuronal networks. *J Endod* 2022;**48**:1434–40. https://doi.org/10.1016/j.joen.2022.07.013
39. Preda F, Morgan N, Van Gerven A, Nogueira-Reis F, Smolders A, Wang X, Nomidis S, Shaheen E, Willems H, Jacobs R. Deep convolutional neural network-based automated segmentation of the maxillofacial complex from cone-beam computed tomography: a validation study. *J Dent* 2022;**124**:104238. https://doi.org/10.1016/j.jdent.2022.104238
40. Yang H, Jo E, Kim HJ, Cha IH, Jung YS, Nam W, Kim JY, Kim JK, Kim YH, Oh TG, Han SS, Kim H, Kim D. Deep learning for automated detection of cyst and tumors of the jaw in panoramic radiographs. *J Clin Med* 2020;**9**:1839. https://doi.org/10.3390/jcm9061839
41. Shoorgashti R, Alimohammadi M, Baghizadeh S, Radmard B, Ebrahimi H, Lesan S. Artificial intelligence models accuracy for odontogenic keratocyst detection from panoramic view radiographs: a systematic review and meta-analysis. *Health Sci Rep* 2025;**8**:e70614. https://doi.org/10.1002/hsr2.70614
42. Han M, Huo Z, Ren J, Zhu H, Li H, Li J, Mei L. Automated landmark detection and lip thickness classification using a convolutional neural network in lateral cephalometric radiographs. *Diagnostics*

2025;**15**:1468. https://doi.org/10.3390/diagnostics15121468
43. Tang H, Liu S, Shi Y, Wei J, Peng J, Feng H. Automatic segmentation and landmark detection of 3D CBCT images using semi supervised learning for assisting orthognathic surgery planning. *Sci Rep* 2025;**15**:8814. https://doi.org/10.1038/s41598-025-93317-6
44. Kim YH, Lee C, Ha EG, Choi YJ, Han SS. A fully deep learning model for the automatic identification of cephalometric landmarks. *Imaging Sci Dent* 2021;**51**:299. https://doi.org/10.5624/isd.20210077
45. Jiao Z, Liang Z, Liao Q, Chen S, Yang H, Hong G, Gui H. Deep learning for automatic detection of cephalometric landmarks on lateral cephalometric radiographs using the Mask region-based convolutional neural network: a pilot study. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2024;**137**:554–62. https://doi.org/10.1016/j.oooo.2024.02.003
46. Yu Y, Wu SJ, Zhu YM. Deep learning-based automated diagnosis of temporomandibular joint anterior disc displacement and its clinical application. *Front Physiol* 2024;**15**:1445258. https://doi.org/10.3389/fphys.2024.1445258
47. Liu Y, Lu Y, Fan Y, Mao L. Tracking-based deep learning method for temporomandibular joint segmentation. *Ann Transl Med* 2021;**9**:467. https://doi.org/10.21037/atm-21-319. [–467].
48. Ozsari S, Güzel MS, Yılmaz D, Kamburoğlu K. A comprehensive review of artificial intelligence based algorithms regarding temporomandibular joint related diseases. *Diagnostics* 2023;**13**:2700. https://doi.org/10.3390/diagnostics13162700
49. Sukegawa S, Yoshii K, Hara T, Matsuyama T, Yamashita K, Nakano K, Takabatake K, Kawai H, Nagatsuka H, Furuki Y. Multi-task deep learning model for classification of dental implant brand and treatment stage using dental panoramic radiograph images. *Biomolecules* 2021;**11**:815. https://doi.org/10.3390/biom11060815
50. Chang JS, Ma CY, Ko EWC. Prediction of surgery-first approach orthognathic surgery using deep learning models. *Int J Oral Maxillofac Surg* 2024;**53**:942–9. https://doi.org/10.1016/j.ijom.2024.05.003
51. Ma Q, Kobayashi E, Fan B, Hara K, Nakagawa K, Masamune K, Sakuma I, Suenaga H. Machine-learning-based approach for predicting postoperative skeletal changes for orthognathic surgical planning. *Int J Med Robot* 2022;**18**:e2379. https://doi.org/10.1002/rcs.2379
52. Salazar D, Rossouw PE, Javed F, Michelogiannakis D. Artificial intelligence for treatment planning and soft tissue outcome prediction of orthognathic treatment: a systematic review. *J Orthod* 2023;**51**:107–19. https://doi.org/10.1177/14653125231203743
53. Liu HH, Li LJ, Shi B, Xu CW, Luo E. Robotic surgical systems in maxillofacial surgery: a review. *Int J Oral Sci* 2017;**9**:63–73. https://doi.org/10.1038/ijos.2017.24
54. Sheetz KH, Claflin J, Dimick JB. Trends in the adoption of robotic surgery for common surgical procedures. *JAMA Netw Open* 2020;**3**:e1918911. https://doi.org/10.1001/jamanetworkopen.2019.18911
55. Bandara DL, Kanmodi KK, Salami AA, Aladelusi TO, Chandrasiri A, Amzat J, Jayasinghe RD. Quality of life of patients treated with robotic surgery in the oral and maxillofacial region: a scoping review of empirical evidence. *BMC Oral Health* 2024;**24**:276. https://doi.org/10.1186/s12903-024-04035-w
56. Choudhury A, Asan O. Role of artificial intelligence in patient safety outcomes: systematic literature review. *JMIR Med Inform* 2020;**8**:e18599. https://doi.org/10.2196/18599
57. Dong F, Yan J, Zhang X, Zhang Y, Liu D, Pan X, Xue L, Liu Y. Artificial intelligence-based predictive model for guidance on treatment strategy selection in oral and maxillofacial surgery. *Heliyon* 2024;**10**:e35742. https://doi.org/10.1016/j.heliyon.2024.e35742
58. Drabiak K. Leveraging law and ethics to promote safe and reliable AI/ML in healthcare. *Front Nucl Med* 2022;**2**:983340. https://doi.org/10.3389/fnume.2022.983340
59. Vaira LA, Lechien JR, Abbate V, Allevi F, Audino G, Beltramini GA, Bergonzani M, Bolzoni A, Committeri U, Crimi S, Gabriele G, Lonardi F, Maglitto F, Petrocelli M, Pucci R, Saponaro G, Tel A, Vellone V, Chiesa-Estomba CM, Boscolo-Rizzo P, Salzano G, De Riu G. Accuracy of ChatGPT-generated information on head and neck and oromaxillofacial surgery: a multicenter collaborative analysis. *Otolaryngol Head Neck Surg* 2024;**170**:1492–503. https://doi.org/10.1002/ohn.489
60. Frosolini A, Franz L, Benedetti S, Vaira LA, de Filippis C, Gennaro P, Marioni G, Gabriele G. Assessing the accuracy of ChatGPT references in head and neck and ENT disciplines. *Eur Arch Otorhinolaryngol* 2023;**280**:5129–33. https://doi.org/10.1007/s00405-023-08205-4
61. Lechien JR, Maniaci A, Gengler I, Hans S, Chiesa-Estomba CM, Vaira LA. Validity and reliability of an instrument evaluating the performance of intelligent chatbot: the artificial intelligence performance instrument (AIPI). *Eur Arch Otorhinolaryngol* 2024;**281**:2063–79. https://doi.org/10.1007/s00405-023-08219-y
62. Vaira LA, Lechien JR, Abbate V, Allevi F, Audino G, Beltramini GA, Bergonzani M, Boscolo-Rizzo P, Califano G, Cammaroto G, Chiesa-Estomba CM, Committeri U, Crimi S, Curran NR, di Bello F, di Stadio A, Frosolini A, Gabriele G, Gengler IM, Lonardi F, Maglitto F, Mayo-Yáñez M, Petrocelli M, Pucci R, Saibene AM, Saponaro G, Tel A, Trabalzini F, Trecca EMC, Vellone V, Salzano G, De Riu G. Validation of the quality analysis of medical artificial intelligence (QAMAI) tool: a new tool to assess the quality of health information provided by AI platforms. *Eur Arch Otorhinolaryngol* 2024;**281**:6123–31. https://doi.org/10.1007/s00405-024-08710-0
63. Amann J, Vetter D, Blomberg SN, Christensen HC, Coffee M, Gerke S, Gilbert TK, Hagendorff T, Holm S, Livne M, Spezzatti A, Strümke I, Zicari RV, Madai VI. Z-Inspection initiative. To explain or not to explain? Artificial intelligence explainability in clinical decision support systems. *PLoS Digit Health* 2022;**1**:e0000016. https://doi.org/10.1371/journal.pdig.0000016
64. Amann J, Blasimme A, Vayena E, Frey D, Madai VI. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak* 2020;**20**:310. https://doi.org/10.1186/s12911-020-01332-6
65. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon* 2024;**10**:e26297. https://doi.org/10.1016/j.heliyon.2024.e26297
66. O'Sullivan S, Nevejans N, Allen C, Blyth A, Leonard S, Pagallo U, Holzinger K, Holzinger A, Sajid MI, Ashrafian H. Legal, regulatory, and ethical frameworks for development of standards in artificial intelligence (AI) and autonomous robotic surgery. *Int J Med Robot* 2019;**15**:e1968. https://doi.org/10.1002/rcs.1968
67. Steerling E, Siira E, Nilsen P, Svedberg P, Nygren J. Implementing AI in healthcare—the relevance of trust: a scoping review. *Front Health Serv* 2023;**3**:1211150. https://doi.org/10.3389/frhs.2023.1211150
68. Sauerbrei A, Kerasidou A, Lucivero F, Hallowell N. The impact of artificial intelligence on the person-centred, doctor–patient relationship: some problems and solutions. *BMC Med Inform Decis Mak* 2023;**23**:73. https://doi.org/10.1186/s12911-023-02162-y

69. Abou-Abdallah M, Dar T, Mahmudzade Y, Michaels J, Talwar R, Tornari C. The quality and readability of patient information provided by ChatGPT: can AI reliably explain common ENT operations? *Eur Arch Otorhinolaryngol* 2024;**281**:6147–53. https://doi.org/10.1007/s00405-024-08598-w
70. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaney-Israni S, Goldenberg A. Do no harm: a roadmap for responsible machine learning for health care. *Nat Med* 2019;**25**:1337–40. https://doi.org/10.1038/s41591-019-0548-6
71. Busch F, Kather JN, Johner C, Moser M, Truhn D, Adams LC, Bressem KK. Navigating the European Union artificial intelligence act for healthcare. *NPJ Digit Med* 2024;**7**:210. https://doi.org/10.1038/s41746-024-01213-6
72. Aboy M, Minssen T, Vayena E. Navigating the EU AI act: implications for regulated digital medical products. *NPJ Digit Med* 2024;**7**:237. https://doi.org/10.1038/s41746-024-01232-3
73. van Kolfschooten H, van Oirschot J. The EU artificial intelligence act (2024): implications for healthcare. *Health Policy* 2024;**149**:105152. https://doi.org/10.1016/j.healthpol.2024.105152
74. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L, Logullo P. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024:e078378. https://doi.org/10.1136/bmj-2023-078378
75. Vaira LA, Lechien JR, Maniaci A, Tanda G, Abbate V, Allevi F, Arena A, Beltramini GA, Bergonzani M, Bolzoni AR, Crimi S, Frosolini A, Gabriele G, Maglitto F, Mayo-Yáñez M, Orrù L, Petrocelli M, Pucci R, Saibene AM, Troise S, Tel A, Vellone V, Chiesa-Estomba CM, Boscolo-Rizzo P, Salzano G, De Riu G. Evaluating AI-generated informed consent documents in oral surgery: a comparative study of ChatGPT-4, Bard Gemini advanced, and human-written consents. *J Craniomaxillofac Surg* 2025;**53**:18–23. https://doi.org/10.1016/j.jcms.2024.10.002
76. Price WN, Cohen IG. Privacy in the age of medical big data. *Nat Med* 2019;**25**:37–43. https://doi.org/10.1038/s41591-018-0272-7
77. Forcier MB, Gallois H, Mullan S, Joly Y. Integrating artificial intelligence into health care through data access: can the GDPR act as a beacon for policymakers? *J Law Biosci* 2019;**6**:317–35. https://doi.org/10.1093/jlb/lsz013
78. McGraw D, Mandl KD. Privacy protections to encourage use of health-relevant digital data in a learning health system. *NPJ Digit Med* 2021;**4**:2. https://doi.org/10.1038/s41746-020-00362-8
79. van Kolfschooten HB. A health-conformant reading of the GDPR's right not to be subject to automated decision-making. *Med Law Rev* 2024;**32**:373–91. https://doi.org/10.1093/medlaw/fwae029
80. Cross JL, Choma MA, Onofrey JA. Bias in medical AI: implications for clinical decision-making. *PLoS Digit Health* 2024;**3**:e0000651. https://doi.org/10.1371/journal.pdig.0000651
81. Norori N, Hu Q, Aellen FM, Faraci FD, Tzovara A. Addressing bias in big data and AI for health care: a call for open science. *Patterns* 2021;**2**:100347https://doi.org/10.1016/j.patter.2021.100347
82. Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, Matsui Y, Nozaki T, Nakaura T, Fujima N, Tatsugami F, Yanagawa M, Hirata K, Yamada A, Tsuboyama T, Kawamura M, Fujioka T, Naganawa S. Fairness of artificial intelligence in healthcare: review and recommendations. *Jpn J Radiol* 2024;**42**:3–15. https://doi.org/10.1007/s11604-023-01474-3
83. Ali R, Connolly ID, Tang OY, Mirza FN, Johnston B, Abdulrazeq HF, Lim RK, Galamaga PF, Libby TJ, Sodha NR, Groff MW, Gokaslan ZL, Telfeian AE, Shin JH, Asaad WF, Zou J, Doberstein CE. Bridging the literacy gap for surgical consents: an AI-human expert collaborative approach. *NPJ Digit Med* 2024;**7**:63. https://doi.org/10.1038/s41746-024-01039-2
84. Teasdale A, Mills L, Costello R. Artificial intelligence-powered surgical consent: patient insights. *Cureus* 2024;**16**:e68134. https://doi.org/10.7759/cureus.68134
85. Helvacioglu-Yigit D, Demirturk H, Ali K, Tamimi D, Koenig L, Almashraqi A. Evaluating artificial intelligence chatbots for patient education in oral and maxillofacial radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol* 2025;**139**:750–9. https://doi.org/10.1016/j.oooo.2025.01.001
86. Quah B, Yong CW, Lai CWM, Islam I. Performance of large language models in oral and maxillofacial surgery examinations. *Int J Oral Maxillofac Surg* 2024;**53**:881–6. https://doi.org/10.1016/j.ijom.2024.06.003
87. Dullabh P, Dhopeshwarkar R, Leaphart D, Peterson C, Gauthreaux N, Grigorescu V, Eloso J, Wei S. Trustworthy artificial intelligence (TAI) for patient-centered outcomes research (PCOR). Washington DC: Office of the Assistant Secretary for Planning and Evaluation (ASPE); 2023.
88. Terranova C, Cestonaro C, Fava L, Cinquetti A. AI and professional liability assessment in healthcare. A revolution in legal medicine? *Front Med* 2024;**10**:1337335. https://doi.org/10.3389/fmed.2023.1337335
89. Marks M, Haupt CE. AI chatbots, health privacy, and challenges to HIPAA compliance. *JAMA* 2023;**330**:309. https://doi.org/10.1001/jama.2023.9458
90. Park HJ. Patient perspectives on informed consent for medical AI: a web-based experiment. *Digit Health* 2024;**10**. https://doi.org/10.1177/20552076241247938. [20552076241247938].
91. Gallifant J, Afshar M, Ameen S, Aphinyanaphongs Y, Chen S, Cacciamani G, Demner-Fushman D, Dligach D, Daneshjou R, Fernandes C, Hansen LH, Landman A, Lehmann L, McCoy LG, Miller T, Moreno A, Munch N, Restrepo D, Savova G, Umeton R, Gichoya JW, Collins GS, Moons KGM, Celi LA, Bitterman DS. The TRIPOD-LLM reporting guideline for studies using large language models. *Nat Med* 2025;**31**:60–9. https://doi.org/10.1038/s41591-024-03425-5
92. De Micco F, Di Palma G, Ferorelli D, De Benedictis A, Tomassini L, Tambone V, Cingolani M, Scendoni R. Artificial intelligence in healthcare: transforming patient safety with intelligent systems—a systematic review. *Front Med* 2025;**11**:1522554https://doi.org/10.3389/fmed.2024.1522554
93. Cai Y, Zhao R, Zhao H, Li Y, Gou L. Exploring the use of ChatGPT/GPT-4 for patient follow-up after oral surgeries. *Int J Oral Maxillofac Surg* 2024;**53**:867–72. https://doi.org/10.1016/j.ijom.2024.04.002
94. Mello MM, Frakes MD, Blumenkranz E, Studdert DM. Malpractice liability and health care quality. *JAMA* 2020;**323**:352. https://doi.org/10.1001/jama.2019.21411
95. Sullivan HR, Schweikart SJ. Are current tort liability doctrines adequate for addressing injury caused by AI? *AMA J Ethics* 2019;**21**:E160–6. https://doi.org/10.1001/amajethics.2019.160
96. Jorstad KT. Intersection of artificial intelligence and medicine: tort liability in the technological age. *J Med Artif Intell* 2020;**3**. https://doi.org/10.21037/jmai-20-57
97. Shumway DO, Hartman HJ. Medical malpractice liability in large language

model artificial intelligence: legal review and policy recommendations. *J Osteopath Med* 2024;**124**:287–90. https://doi.org/10.1515/jom-2023-0229

98. Cestonaro C, Delicati A, Marcante B, Caenazzo L, Tozzo P. Defining medical liability when artificial intelligence is applied on diagnostic algorithms: a systematic review. *Front Med* 2023; **10**:1305756. https://doi.org/10.3389/fmed.2023.1305756
99. Pruski M. AI-enhanced healthcare: not a new paradigm for informed consent. *J Bioeth Inq* 2024;**21**:475–89. https://doi.org/10.1007/s11673-023-10320-0
100. Ramos S, Ellul J. Blockchain for artificial intelligence (AI): enhancing compliance with the EU AI Act through distributed ledger technology. A cybersecurity perspective. *Int Cybersecur Law Rev* 2024;**5**:1–20. https://doi.org/10.1365/s43439-023-00107-9
101. López González A, Moreno M, Moreno Román AC, Hadfeg Fernández Y, Cepero Pérez N. Ethics in artificial intelligence: an approach to cybersecurity. *Inteligencia Artificial* 2024;**27**:38–54. https://doi.org/10.4114/intartif.vol27iss73pp38-54
102. Elendu C, Omeludike EK, Oloyede PO, Obidigbo BT, Omeludike JC. Legal implications for clinicians in cybersecurity incidents: a review. *Medicine* 2024; **103**:e39887. https://doi.org/10.1097/MD.0000000000039887
103. Ibrahim H, Liu X, Rivera SC, Moher D, Chan AW, Sydes MR, Calvert MJ, Denniston AK. Reporting guidelines for clinical trials of artificial intelligence interventions: the SPIRIT-AI and CONSORT-AI guidelines. *Trials* 2021; **22**:11. https://doi.org/10.1186/s13063-020-04951-6
104. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P. DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med* 2022; **28**:924–33. https://doi.org/10.1038/s41591-022-01772-9
105. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc* 2016; **23**:899–908. https://doi.org/10.1093/jamia/ocv189
106. Mongan J, Moy L, Kahn CE. Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell* 2020; **2**:e200029https://doi.org/10.1148/ryai.2020200029
107. Hernandez-Boussard T, Bozkurt S, Ioannidis JPA, Shah NH. MINIMAR (MINimum information for medical AI reporting): developing reporting standards for artificial intelligence in health care. *J Am Med Inform Assoc* 2020; **27**:2011–5. https://doi.org/10.1093/jamia/ocaa088
108. Haltaufderheide J, Pfisterer-Heise S, Pieper D, Ranisch R. The ethical landscape of robot-assisted surgery: a systematic review. *J Robot Surg* 2025; **19**:102. https://doi.org/10.1007/s11701-025-02228-1
109. Elliott JH, Synnot A, Turner T, Simmonds M, Akl EA, McDonald S, Salanti G, Meerpohl J, MacLehose H, Hilton J, Tovey D, Shemilt I, Thomas J. Living systematic review network. Living systematic review: 1. Introduction—the why, what, when, and how. *J Clin Epidemiol* 2017; **91**:23–30. https://doi.org/10.1016/j.jclinepi.2017.08.010
110. Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JP, Mavergames C, Gruen RL. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med* 2014; **11**:e1001603https://doi.org/10.1371/journal.pmed.1001603
111. Morrison A, Polisena J, Husereau D, Moulton K, Clark M, Fiander M, Mierzwinski-Urban M, Clifford T, Hutton B, Rabb D. The effect of English-language restriction on systematic review-based meta-analyses: a systematic review of empirical studies. *Int J Technol Assess Health Care* 2012; **28**:138–44. https://doi.org/10.1017/S0266462312000086

Correspondence: Giulio Cirignaco
Department of Medicine
Section of Maxillofacial Surgery
University of Siena
Viale Mario Bracci
53100 Siena
Italy.
E-mail: g.cirignaco@student.unisi.it