Journal of Pioneering Medical Sciences
Received: August 23, 2026 | Accepted: October 19, 2026 | Published: November 15, 2026
Volume 15, Issue S02, Pages 121-128

DOI https://doi.org/10.47310/jpms202615S0213



Artificial Intelligence in Oral Medicine and Radiology: A Narrative Review

Tarang K. Mehta[1*], M.R. Vaishali[2], Sonal Yashwant Patil[3], Avani Soni[4], Mona Janak Shah[5] and Jigar Joshi[6]

[1]Department of Oral and Maxillofacial Pathology, K.M. Shah Dental College and Hospital, Sumandeep Vidyapeeth Deemed to be University, Vadodara, Gujarat, India
[2]Department of Oral Medicine and Radiology, M.R. Ambedkar Dental College and Hospital, Rajiv Gandhi University of Health Sciences, Bengaluru, Karnataka, India
[3]Department of Oral Medicine and Radiology, MGV's KBH Dental College and Hospital, Nashik, Maharashtra, India
[4]Department of Oral Medicine and Radiology, Karnavati School of Dentistry, Karnavati University, Uvarsad, Gandhinagar, Gujarat, India
[5]Department of Oral Medicine and Radiology, Faculty of Dental Sciences, Dharmsinh Desai University, Nadiad, Gujarat, India
[6]Department of Oral Medicine and Radiology, Ahmedabad Dental College and Hospital, Gujarat University, Ahmedabad, Gujarat, India

Author Designation: [1]Senior Lecturer

*Corresponding author: Tarang K. Mehta (e-mail: tarangmehta111@gmail.com).

©2026 the Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0

Abstract This narrative review examines Artificial Intelligence (AI) applications in oral medicine and dentomaxillofacial radiology, including mucosal-lesion triage, oral potentially malignant disorders and cancer, two-dimensional imaging, cone-beam computed tomography and teledentistry. The review differentiates classification, detection, segmentation and prediction tasks; considers internal versus external validation and clinical reference standards; and identifies requirements for prospective human–AI evaluation. Evidence was located by a targeted literature search rather than a fully reproducible systematic review; the search yield and review dates were not fully documented. Strong internal discrimination or segmentation scores should not be equated with improved referrals, safe clinical decisions or patient outcomes. Broad multicentre external validation and implementation studies remain priorities.

Key Words Artificial Intelligence, Oral Medicine, Oral and Maxillofacial Radiology, Deep Learning, Oral Cancer, Cone-Beam Computed Tomography, Teledentistry

INTRODUCTION

AI has transitioned from a more theoretical to a practical research platform throughout dentistry. Patterns can be found in structured clinical data with machine learning algorithms and hierarchical image features can be learned directly from photographs, radiographs and volumetric scans with deep learning systems. These are desirable in oral medicine and radiology, as both rely heavily on pattern recognition, integration of diverse information and rapid differentiation of benign, potentially malignant and malignant disease. Meanwhile, enthusiasm needs to be tempered with the difference between a high-performing curated set of data and real-world clinical value. The primary question is no longer if an algorithm can generate a probability or segmentation but whether it can generate one that is valid across patients, devices, institutions, disease spectra and real clinical workflows [1].

Systematic reviews of dental AI have reported that the number of applications in the field has increased rapidly, ranging from caries, periodontal disease, endodontic lesions, oral cancer, cephalometry, anatomical segmentation, to treatment planning [2]. Broad reviews also have noted that AI is not a monolithic technology-supervised machine learning, convolutional neural networks, transformer-based models, ensemble methods, NLP and multimodal architectures are used to address various problems and need different types of validation [3]. The field of oral and maxillofacial radiology is a very active area, as digital imaging offers a wealth of easily standardized data but initial reviews revealed significant data heterogeneity, reference standards, outcome measures and risk of bias [4]. A scoping review of CNNs for dental image diagnostics revealed many promising studies technically but not designed to show clinical impact [5].

The range of the field is captured by two recent publications that were used as a starting point for this review. One spoke about oral-health surveillance using AI in underserved communities and the possibilities of automated or remotely assisted screening to broaden the reach of specialists [6]. The other systematically reviewed AI for the

diagnosis of mouth ulcers, which is a clinically important and heterogeneous group of lesions, can be challenging to triage due to the visual similarity, variable duration and overlapping etiologies [7]. These papers collectively represent a helpful spectrum; AI can be used to aid population-level surveillance, as well as individual-level diagnostic reasoning. But the link between image classification and improved outcomes is not direct. A surveillance tool is only useful if it is effective in the target population, generates actionable referrals and is part of a service that can handle positive results. A clinically relevant lesion classification model must increase the sensitivity, specificity, time to diagnosis and/or decision consistency without adding new errors that are not acceptable.

This narrative review thus focuses on the current evidence of the use of AI in oral medicine and radiology with particular attention to: (1) Oral mucosal lesions, potentially malignant disorders and oral cancer; (2) Two-dimensional and three-dimensional dentomaxillofacial imaging; (3) AI in surveillance and teledentistry; (4) The potential for multimodal integration of imaging with microbial and molecular information and (5) Methodological, ethical and implementation priorities necessary for safe translation. The review is designed to consider AI as a clinical intervention, not just an accuracy rating and to do so within a human system.

REVIEW APPROACH

A narrative search was performed in PubMed and Scopus databases to identify clinically relevant studies and reviews related to the use of AI, machine learning, deep learning, oral medicine, oral mucosal lesions, oral cancer, dental radiology, panoramic radiography, periapical radiography, cone-beam computed tomography, teledentistry and oral-health surveillance. Studies that are peer-reviewed, have clear clinical tasks, have established reviews, have papers with external or clinically oriented validation and reporting or education frameworks relevant to translation were given priority. The two separate PubMed Central articles mentioned by the authors were included. Furthermore, the Google Scholar profile of Dr Tarang K. Mehta was explored and two papers were chosen from the PubMed-indexed profile that were relevant to the topic of future multimodal AI, with the help of microbial and molecular biomarkers. Due to the nature of this narrative review, the literature was synthesized conceptually rather than by formal meta-analysis and findings of the studies were interpreted in terms of clinical utility, generalizability and implementation.

Conceptual Foundations: Automated detection to Augmented Intelligence

AI has different clinical implications depending on the task. Classification categorizes a lesion, benign or suspicious lesion. Detection determines if there is an abnormality and its location. Segmentation is the process of identifying structures or lesions, one pixel (or voxel) at a time. Prediction: prediction of future outcomes, natural language systems: Summarization or generation of natural language. These tasks are not interchangeable and the metrics should be aligned to the intended use. For instance, a screening system can be evaluated on the basis of sensitivity and safe referral, while a segmentation system can be evaluated on overlap measures which are not directly related to diagnostic correctness. Even though a model is technically sound, it may be clinically unsafe if the operating threshold is not suitable for the population it is used in.

The best conceptualization of dental AI is augmented intelligence, where algorithms complete specific tasks and clinicians interpret and manage the situation. This is particularly important in oral medicine as the diagnosis may rely on lesion history, pain, recurrence, tobacco use, medication use, systemic disease, palpation, lymph-node findings and histopathology. An image-only algorithm might recognize a pattern in the image but not automatically recreate all clinically relevant context. Likewise, a radiographic algorithm can identify a periapical radiolucency but fail to make a diagnosis of pulpal status, symptoms, prognosis for the restoration or if intervention is warranted. So, the comparison isn't AI versus clinician but unaided clinician versus clinician with well-calibrated AI.

Broad dental reviews highlight a wide range of possible applications and common challenges as well: small or nonrepresentative datasets, leakage between training and test data, poor external validation, inconsistent annotations and overinterpretation of accuracy [8]. These restrictions are especially significant in the presence of rare but serious, conditions and common benign conditions. A model can have a high overall accuracy but have a low accuracy for the minority class that is most clinically relevant. Therefore, it is important to assess sensitivity by disease category, calibration, false-negative analysis and subgroup performance as well as a single headline metric.

Artificial Intelligence in Oral Medicine

Oral Mucosal Lesions and Ulcerative Disorders: Oral ulceration is a good test case for AI because it is a common clinical phenotype with a wide range of causes. The clinical appearance of traumatic ulcers, recurrent aphthous disease, immune-mediated disorders, infections, drug reactions and malignancy can be similar. Tiwari *et al.* [7] conducted a systematic review that found evidence of the potential of AI to assist in the diagnosis of ulcers but noted that the available evidence was inconsistent due to small sample sizes, a lack of consistency in lesion types and a lack of validation. This is important because a model that discriminates between two preselected categories in a research data set is not the same as a clinical diagnostic system that will be used to discriminate among all the categories of lesions seen in practice.

Photography-based models are still appealing, however, since the images are easily acquired at low cost and are becoming increasingly possible with the use of a smartphone. In the field of oral cancer diagnosis, deep learning has been applied to oral squamous cell carcinoma

(OSCC) using clinical photographs and discriminative visual features have been learned from the images of the oral cavity [9]. There are other studies that demonstrate automatic classification and detection of oral cancer in photographic data [10] and related models have been developed for oral potentially malignant disorders (OPMDs) [11]. Other machine-learning methods have been investigated for automated categorization of lesions in wider intraoral image datasets [12]. All these studies have confirmed the biological plausibility of image-based AI but their clinical impact would be limited if they do not take into account the confusing mimics that cause diagnostic uncertainty in real practice.

One of the main drawbacks is that it is limited to a specific spectrum. There are numerous research datasets that include good quality images of well known disease and relatively clear control images. The challenging cases in clinics are often small, partially treated, inflamed, poorly lit, blocked by saliva or in difficult anatomical locations. Lesion appearance is also dependent on the pigmentation of the skin, the keratinization of the mucous membranes, the amount of light admitted to the camera, compression and the skill of the operator. To be generalizable, the acquisition conditions must be varied and prospective testing must be conducted. A systematic review of AI in head and neck precancerous and cancerous lesions also highlighted the potential benefits of computational approaches but also noted that there is heterogeneity and a need for improved validation [13].

Another question is what is the goal? In many cases, the clinically valuable result of a photo is not a definitive diagnosis but a risk category that is associated with action. For instance, an AI system might categorize lesions as a routine follow-up, for clinical review or for urgent specialist referral. This triage model could be safer and more realistic than the autonomous diagnosis as it recognises uncertainty and maintains confirmatory examination and biopsy. Future trials should therefore focus on the assessment of referral appropriateness, time to definitive diagnosis, missed malignancies, patient anxiety and clinician workload, in addition to image accuracy.

The oral potentially malignant disorders and oral cancer are discussed. Oral potentially malignant disorders and oral cancer are discussed.

Early detection of oral cancer is a critical need, making OPMDs and OSCC a significant target for AI. Clinical photographs provide information about the surface texture, colour heterogeneity, ulceration, margins and anatomical distribution, while histopathology and molecular assays provide further layers of information. The current image-based models show that deep learning can capture clinically relevant patterns but most of the studies are retrospective and prone to overfitting, particularly when images from the same patient are used for both training and testing. Patient-level data separation is, therefore, a must.

The greater opportunity is in multimodal risk estimation. The two other papers chosen from Dr Tarang K. Mehta's Google Scholar profile are not AI studies but are relevant as they are biological information that could be incorporated into images by future AI systems. In one study, microbial associations in oral lichen planus were assessed and differences in the presence of some microorganisms were found between the cases and controls [14]. The other investigated genetic markers associated with methylation in oral premalignant and malignant lesions and their possible association with progression and patient outcome [15]. These studies are not meant to be interpreted as proof of the validation of AI. Instead, they represent the type of structured biological variables that may be used in future multimodal models, along with pictures, x-rays, clinical history and histopathology.

One of the challenges in oral oncology that can be solved by multimodal AI is the fact that no single modality encompasses the whole disease process. A photograph is a reflection of the phenotype; palpation gives information on texture and induration; histopathology gives information on the architecture of the tissue; molecular markers give information on biological pathways; and imaging gives information on depth, bone involvement or nodal disease. The use of these sources can enhance risk stratification but multimodal models can also pose new challenges. Data is often missing, assays are different, sample sizes decrease with the number of modalities and complex models can be hard to audit. Thus, combinations of variables for multimodal research should be clinically motivated rather than simply adding all variables that are available.

Future models should address questions that will change management for OPMDs, such as: What lesions should be immediately biopsied, which patients should have more intense surveillance and what lesions are most likely to become malignant? These are tasks that take time to follow up and are longitudinal tasks. While classification of an existing cancer is helpful, predicting the progression is another more challenging task. AI will need to be able to support long-term risk-based surveillance but this will not be possible until prospective cohorts are established with standardized outcomes and these cohorts are geographically diverse.

Artificial Intelligence in Dentomaxillofacial Radiology

Two-Dimensional Radiography: Radiology is a relatively well-developed field for AI, as panoramic and intraoral images are regularly digitized. Tooth detection and numbering are basic tasks that can be automated and used to facilitate downstream diagnostic pipelines, charting and data organization. Tuzoff *et al.* [16] illustrated the feasibility of automating a labor-intensive interpretive step using convolutional neural networks for tooth detection and numbering on panoramic radiographs. While these applications may seem straightforward in comparison with disease detection, they demonstrate the potential of AI to enhance workflow uniformity and provide structured inputs for more intricate systems.

There has also been improvement in the detection of pathology. The use of deep learning in the radiographic

detection of apical lesions [17] and periodontal bone loss [18] has been investigated. The tasks are clinically relevant, as subtle findings may be missed, multiple measurements are time consuming and interobserver variation is likely. Automated systems can be used as a second reader, can highlight areas that need to be read or can give quantitative measurements. The reference standard is still a big constraint, however. A single clinician may have used model labels, which can lead to individual interpretation being reproduced by the algorithm. Consensus annotation, clinical correlation or longitudinal confirmation can enhance the quality of the reference but each method has certain limitations.

Another high value use is anatomical localization. The use of deep learning for automated detection of third molars and mandibular nerve has been investigated [19]. In theory, these systems may speed up the preoperative review process and minimise omissions but in order to be clinically meaningful, difficult anatomy, superimposition, low quality images, altered anatomy and various scanner manufacturers should be included. The algorithm should also be able to show uncertainty, instead of a confident output when image quality is bad.

Cone-Beam Computed Tomography

Cone-beam computed tomography (CBCT) provides 3D data but also adds complexity to the data. Each volume has hundreds of slices and several anatomical structures, providing opportunities for detection, segmentation and quantitative analysis. The usefulness of AI has been tested in the detection of periapical lesions in CBCT [20]. These tools could assist radiologists in overcoming the burdens of volumetric search, although they rely on the false-positive rate and sensitivity. Too many reminders can lead to “alert fatigue” and can actually decrease attention.

Recent systematic reviews on the use of AI in dentomaxillofacial CBCT have reported the use of AI for tooth and bone segmentation, localization of the mandibular canal, classification of cystic lesions, cephalometric landmarking and planning of surgical procedures [21]. A systematic review and meta-analysis of detection and segmentation of oral and maxillofacial structures also revealed promising results with methodological heterogeneity being highlighted [22]. Domain shift is the main problem in translation. CBCT images vary from machine to machine, depending on the field of view, voxel size, reconstruction algorithm, artifacts and exposure protocol. Even if the anatomy is similar, a model trained on one scanner may suffer significant degradation when applied to another scanner.

Three-dimensional segmentation can save a lot of clinician time especially in treatment planning and quantitative research. However, geometric overlap measures (e.g., Dice coefficient) should not be used alone. Clinically important errors could be more important in a critical region than in a low-risk region, even though the errors are smaller in the critical region. Task-specific safety margins, boundary errors and downstream impact on decisions should therefore be considered in the evaluation. For instance, segmentation of the mandibular canal should be evaluated based on the surgical risk and not average voxel overlap.

There is also a concern with incidental findings as it relates to CBCT AI. Abnormalities may be present in a comprehensive volume that are not related to the reason for imaging. A negative AI result can sometimes give the illusion of completeness with narrow-task algorithms if the clinician thinks that a negative result means no disease is present. The scope of the system should be clearly communicated in the user interface: an algorithm for teeth segmentation is not a general radiology reader. The safest systems are systems where the intended use, limitations and uncertainty are apparent at the point of care.

AI systems, both commercial and web-based, are increasingly coming into practice, leaving a gap between research evidence and actual use. The feasibility of studying automated web-based AI interpretation of structures and treatments on panoramic radiographs was evaluated [23]. However, software changes, proprietary training data, threshold changes and post-market drift make reproducibility difficult. Independent evaluation should be on-going, not a single approval event.

AI for Oral-Health Surveillance and Teledentistry

AI's potential for public health is greatest when specialists are scarce. The review on oral-health surveillance among underserved communities highlighted the potential of AI, digital imaging and remote care pathways to aid in screening and prioritization [6]. The key is not to replace dentists with AI but to focus limited skills on those who will benefit the most. In rural, remote or high volume areas, simple digital records and smartphone photographs can be sent for automated triage or for asynchronous specialist review.

Studies in teledentistry have demonstrated the ability to compare the use of a smartphone camera with the face-to-face caries examination using photographs [24]. Automated caries detection using machine-learning systems from smartphone colour photographs has also been developed [25] and AI has been investigated for caries detection on intraoral images [26]. These are all ways to get from the image capture to the triage but the quality of the pathway depends on more than the algorithm. All of these are critical components of image acquisition protocols, operator training, infection control, patient consent, secure transmission, referral capacity and follow up completion.

Equity is a special consideration. While AI can help to overcome geographic barriers, it can also perpetuate inequities if it is trained using data that underrepresents certain populations, languages, devices or disease patterns. Lower cost cameras, inconsistent connectivity and more advanced disease than tertiary center data sets may be used in underserved settings. A good model in a university clinic might then not work exactly where it's most needed. Community deployment should start with local validation and monitoring of subgroup performance and have a clear escalation pathway for cases that are uncertain or high-risk.

The other practical problem is prevalence. If sensitivity and specificity do not change, the positive predictive value will change with the prevalence of the disease. A model designed to be used in a high-risk referral clinic can give a lot of false positive results in mass screening. Thresholds might need to be adjusted for each setting and the desired outcome of a positive result should be pre-determined. Screening programs should be assessed throughout the whole process of image capture to completed diagnosis and not just the first algorithmic step.

Clinical Integration and Human-AI Collaboration

The clinical usefulness is a result of the interplay between model, user, patient and workflow. AI can cut down on repetitive tasks, ensure consistency, identify areas that may have been missed and assist junior clinicians. It can also cause algorithm aversion, where users disregard a system after noticing a prominent error or automation bias, where they become overly reliant on a system that they believe is correct. The interface design and training, therefore, play a significant part in safety, in addition to the model architecture.

Multiple conditions should be compared in human-AI studies: clinician alone, AI alone and clinician with AI. The outcomes that are relevant are: sensitivity, specificity, reading time, confidence, referral decisions, error type and disagreement resolution. A system that works well for the experts may be of little value to them but of great value to less experienced users. On the other hand, an algorithm with a relatively low individual accuracy could be beneficial to the workflow if it consistently selects normal exams or presegments structures for review.

Calibration is essential. A model must be correct not only often but also have the right level of confidence that matches the actual probability of correctness. Unsafe decisions are made when there is poor calibration of confidence. Systems should be able to abstain or defer when inputs are outside the training distribution, for high stakes oral lesions. Uncertainty-aware referral might be more clinically useful than a hard classification into a fixed label.

The idea of having a 'second reader' is especially applicable in radiology and photographic triage. The AI can identify potential abnormalities and the clinician can accept, reject or explore the potential abnormalities. This ensures professional accountability and enables contextual reasoning. Second-reader systems, however, should still be monitored as false prompts may be a burden to the system. The best operating point should be determined based on the clinical context instead of maximizing a generic metric.

Methodological Limitations, Bias, Explainability and Governance

The biggest hurdle to credible dental AI is not the lack of sophistication in the algorithms but the lack of evidence. Most studies are retrospective, single center, convenience sample. Leakage frequently occurs when multiple images from the same patient are present in both the training set and the testing set, which is not the case for data that are divided at the patient level.When data are divided at the image level instead of the patient level, leakage can occur if multiple images from the same patient are present in both the training set and the testing set. The optimization of hyperparameters can be repeated on a test set and external validation may be lacking. These practices make it seem like there is more performance than there actually is.

Another type of bias is reference standards. Oral lesion labels can be clinical, histopathology or a combination of clinical and histopathology. One or more radiographers can make radiographic labels. Whereas the ground truth is not certain, the evaluation by AI is not certain. Studies should specify who made the labels, how disagreements were handled, if the annotators were blind to relevant information and whether or not clinical or pathological confirmation was available. It is important to consider false negatives and false positives in clinically relevant categories in the analysis of errors.

The representativeness of the datasets is also crucial. Performance can be affected by age, sex, ethnicity, disease prevalence, comorbidities, scanner type, image quality and clinical setting. Assessment of fairness should be planned, not added after development. For underrepresented populations, it is particularly critical to include them in AI proposals, as the impact of underrepresentation can be most severe.

Explainability is frequently suggested as a remedy to opacity. Heat maps and saliency visualizations are useful for visualizing image regions associated with a prediction but do not imply causal reasoning or correctness. An explanation that is plausible but not correct, can be accompanied by a visual explanation. Therefore, explainability should be used to aid in auditing and communication but not as a measure of validity. Better transparency can be achieved by providing clear documentation of intended use, training data, known failure modes, uncertainty and performance in settings.

The dental researchers have suggested a particular checklist for planning, conducting and reporting AI studies, which includes study goals, data sampling, reference standards, model development, evaluation, uncertainty, explainability and reporting [27]. CONSORT-AI extends the reporting requirements for clinical trials with AI, covering aspects such as acquisition of inputs, integration of outputs and human-AI interaction [28]. DECIDE-AI offers guidance on safety, human factors and real-world workflow assessment for early live clinical evaluation [29]. These frameworks are useful because the method goes beyond the algorithm and is necessary for reproducibility and clinical translation.

Governance needs to also cover privacy, cyber security, version control, accountability and post deployment drift. Dental images are health information that can be identified and cloud-based processing can add to the exposure. Institutions require policies with regards to data transfer, vendor access, model updates and incident response. If a

model changes following deployment, performance should be re-evaluated as the clinically used system could be different from the one validated.

DISCUSSION

The evidence presented here lends itself to a balanced conclusion. AI has proven its technical ability in the field of oral lesion photography, classification of oral cancer and OPMD, panoramic radiography, detection of periapical disease, periodontal bone-loss assessment, anatomical localization, CBCT analysis and smartphone-based screening. The articles provided are especially valuable as they take the discussion beyond the specialist imaging labs: one discusses the use of imaging in diagnostic support for mouth ulcers and the other discusses surveillance in underserved communities [6,7]. These use cases reflect the two main tenets of dental AI: increased diagnostic uniformity and access to expertise.

The clinical evidence is not as developed as the technical literature, however. Most studies respond to the question, 'Can a model determine a pattern in this data set? Much fewer respond to the question, 'Does the model improve care in this workflow? The difference is key. To be clinically useful, it needs to be evaluated prospectively, compared with current practice, assessed for consequences and shown to be beneficial and to have no new harms or burdens. An algorithm can be made more sensitive and lead to too many referrals. It can make reading easier and make them more dependent. It can do well in the overall population and poorly in a subpopulation. Accuracy alone does not represent these trade-offs.

Oral medicine is a challenge due to the integration of history, examination, serial change, biopsy, microbiology and molecular findings for definitive diagnosis. The two other Scholar-profile publications were thus added as conceptual bridges to multimodal AI. The oral lichen planus study is an example of microbial variables that could be used to enhance disease characterization [14] and the genetic-marker study is an example of molecular information that could be associated with the risk of disease progression [15]. The next generation of oral-medicine AI should not just be bigger image classifiers but also determine which combinations of data provide clinically relevant information and can be collected reliably.

Radiology is likely to be the first specialty to routinely adopt digital workflows and tasks can be more clearly defined. Natural candidates for assistive automation include the following: Tooth numbering, lesion highlighting, quantitative bone measurements and segmentation. However, external validation and workflow trials are still crucial in radiology. If there are differences in the way images are acquired, a model trained in one imaging ecosystem may not work in another. Furthermore, AI should not limit the clinicians' search, as it assumes that the area not flagged is normal.

This has significant public health consequences. Teledentistry with AI may help with screening in schools, community camps, rural clinics and in geographically isolated communities. However, to be successful, service design is needed. Positive screen should result in appointment, urgent referral should be to a center that can perform a biopsy, ambiguous should be reviewed and communication should be in appropriate languages. Technologies that do not have downstream capacity can only be a solution to an unmet need and not a solution to the need itself. So, implementation research should be done on completion of care, not detection.

There are three particular evidence gaps that are significant. First, there are still not enough prospective multicenter datasets, especially for oral mucosal disease. Second, the prediction of malignant transformation or disease progression is much less advanced in the longitudinal dimension than in the cross-sectional dimension of classification. Third, few studies assess the human-AI teams in terms of patient-important outcomes. These gaps will need to be filled by the cooperation of oral physicians, radiologists, pathologists, data scientists, primary-care dentists, public-health researchers and patients.

Education is an integral part of implementation. A core curriculum for AI in oral and dental care has been suggested to support undergraduate and postgraduate students to gain knowledge of data, model building, performance, ethics and clinical evaluation [30]. This is essential, as safe use of the AI is not just about clicking an AI button. Clinicians should become familiar with prevalence, uncertainty, false positives, false negatives, domain shift and automated explanations' limitations. Thus, AI literacy must be part of evidence-based dental practice.

The key clinical take-home message from this review is that AI should be considered as an intervention in a care pathway. The most successful systems will probably be those that address a well-defined constraint, acknowledge uncertainty, augment the capabilities of humans and are monitored following deployment. The weakest will be those systems that are touted as universal diagnostic replacements based on internal test accuracy.

Future Directions

Multicenter prospective studies with patient-level separation of data and with prespecified analysis plans should be the focus of future research. Before claims of generalizability are made, external validation should take place across institutions, devices and populations. In rare oral lesions, collaborative data networks may be required to obtain sufficient numbers and diversity of lesions.

Hypotheses should be clinically meaningful and the development of multimodal models should be based on them. For oral cancer and OPMDs, a combination of clinical photographs, risk factors, histopathology, radiology and validated biomarkers may enhance risk stratification. Each additional modality added should show incremental value and feasibility, however. Expensive or unavailable data will not be likely to enhance access in resource limited settings.

Uncertainty-aware systems should be ubiquitous. Models should be sensitive to poor image quality, unfamiliar devices and out-of-distribution presentations. Safe abstention, referral and escalation guidelines should be assessed in a prospective manner. Outputs can be useful at certain clinical thresholds, which can be determined by calibration and decision-curve analysis.

Technical development should be accompanied by human factors research. The design of the interface, the burden of alerts, trust, training, cognitive load and the actions of novice and expert users should be measured. Before using the AI system actively, local performance information can be provided by silent or shadow deployment, where the AI system operates without impacting on care. Future trials should assess patient outcomes, efficiency of workflow and cost-effectiveness.

Implementation needs to be co-designed with local services for underserved communities. A system for smartphone-based systems should be tested with the actual devices and image-acquisition conditions in the field. Programs should have security measures in place for data handling, culturally responsive communication, a specialist review for ambiguous cases and a way to ensure referrals are made. Equity metrics should be reported in a clear manner.

Last but not least, post deployment monitoring is crucial. The prevalence of disease, equipment, clinical behavior and software versions vary over time. Performance drift, subgroup errors, safety events and the effect of model changes should be monitored. The reporting framework and AI education should be a part of this lifecycle, allowing clinicians to evaluate and manage systems instead of simply being handed them.

CONCLUSIONS

AI algorithms have demonstrated potential for defined image-based tasks in oral medicine and radiology but current evidence is insufficient to conclude that their broad deployment improves patient outcomes. Models require external validation across devices and populations, prospectively specified clinical thresholds, comparison with clinician workflows and monitoring for unequal performance, drift and inappropriate automation. Clinical responsibility, confirmatory diagnostics and access to referral pathways remain essential.

REFERENCES

[1] Schwendicke, F. *et al.* "Artificial Intelligence in Dentistry: Chances and Challenges." *Journal of Dental Research*, vol. 99, no. 7, 2020, pp. 769-774. https://doi.org/10.1177/0022034520915714.

[2] Khanagar, S.B. *et al.* "Developments, Application and Performance of Artificial Intelligence in Dentistry: A Systematic Review." *Journal of Dental Sciences*, vol. 16, no. 1, 2021, pp. 508-522. https://doi.org/10.1016/j.jds.2020.06.019.

[3] Shan, T. *et al.* "Application of Artificial Intelligence in Dentistry." *Journal of Dental Research*, vol. 100, no. 3, 2021, pp. 232-244. https://doi.org/10.1177/0022034520969115.

[4] Hung, K. *et al.* "The Use and Performance of Artificial Intelligence Applications in Dental and Maxillofacial Radiology: A Systematic Review." *Dentomaxillofacial Radiology*, vol. 49, no. 1, 2020. https://doi.org/10.1259/dmfr.20190107.

[5] Schwendicke, F. *et al.* "Convolutional Neural Networks for Dental Image Diagnostics: A Scoping Review." *Journal of Dentistry*, vol. 91, 2019. https://doi.org/10.1016/j.jdent.2019.103226.

[6] Tiwari, A. *et al.* "Artificial Intelligence in Oral Health Surveillance Among Under-Served Communities." *Bioinformation*, vol. 19, no. 13, 2023, pp. 1329-1335. https://doi.org/10.6026/973206300191329.

[7] Tiwari, A. *et al.* "Artificial Intelligence's Use in the Diagnosis of Mouth Ulcers: A Systematic Review." *Cureus*, vol. 15, no. 9, 2023. https://doi.org/10.7759/cureus.45187.

[8] Patil, S. *et al.* "Artificial Intelligence in the Diagnosis of Oral Diseases: Applications and Pitfalls." *Diagnostics*, vol. 12, no. 5, 2022. https://doi.org/10.3390/diagnostics12051029.

[9] Fu, Q. *et al.* "A Deep Learning Algorithm for Detection of Oral Cav Squamous Cell Carcinoma from Photographic Images: A Retrospective Study." *EClinicalMedicine*, vol. 27, 2020. https://doi.org/10.1016/j.eclinm.2020.100558.

[10] Warin, K. *et al.* "Automatic Classification and Detection of Oral Cancer in Photographic Images Using Deep Learning Algorithms." *Journal of Oral Pathology & Medicine*, vol. 50, no. 9, 2021, pp. 911-918. https://doi.org/10.1111/jop.13227.

[11] Warin, K. *et al.* "Performance of Deep Convolutional Neural Network for Classification and Detection of Oral Potentially Malignant Disorders in Photographic Images." *International Journal of Oral and Maxillofacial Surgery*, vol. 51, no. 5, 2022, pp. 699-704. https://doi.org/10.1016/j.ijom.2021.09.001.

[12] Dinesh, Y. *et al.* "Machine Learning in the Detection of Oral Lesions with Clinical Intraoral Images." *Cureus*, vol. 15, no. 8, 2023. https://doi.org/10.7759/cureus.44018.

[13] Mahmood, H. *et al.* "Use of Artificial Intelligence in Diagnosis of Head and Neck Precancerous and Cancerous Lesions: A Systematic Review." *Oral Oncology*, vol. 110, 2020. https://doi.org/10.1016/j.oraloncology.2020.

[14] Mehta, T. *et al.* "Evaluation of the Role of Oral Microbes in Pathogenesis of Oral Lichen Planus." *Journal of Pharmacy & Bioallied Sciences*, vol. 16, no. Suppl 3, 2024, pp. S2153-S2155. https://doi.org/10.4103/jpbs.jpbs_90_24.

[15] Wasti, A. *et al.* "Investigation of Genetic Markers for Predicting Oral Cancer Progression and Patient Outcomes." *Journal of Pharmacy & Bioallied Sciences*, vol. 16, no. Suppl 3, 2024, pp. S2129-S2131. https://doi.org/10.4103/jpbs.jpbs_113_24.

[16] Tuzoff, D.V. *et al.* "Tooth Detection and Numbering in Panoramic Radiographs Using Convolutional Neural Networks." *Dentomaxillofacial Radiology*, vol. 48, no. 4, 2019. https://doi.org/10.1259/dmfr.20180051.

[17] Ekert, T. *et al.* "Deep Learning for the Radiographic Detection of Apical Lesions." *Journal of Endodontics*, vol. 45, no. 7, 2019, pp. 917-922. https://doi.org/10.1016/j.joen.2019.03.016.

[18] Krois, J. *et al.* "Deep Learning for the Radiographic Detection of Periodontal Bone Loss." *Scientific Reports*, vol. 9, no. 1, 2019. https://doi.org/10.1038/s41598-019-44839-3.

[19] Vinayahalingam, S. *et al.* "Automated Detection of Third Molars and Mandibular Nerve by Deep Learning." *Scientific Reports*, vol. 9, no. 1, 2019. https://doi.org/10.1038/s41598-019-45487-3.

[20] Setzer, F.C. *et al.* "Artificial Intelligence for the Computer-Aided Detection of Periapical Lesions in Cone-Beam Computed Tomographic Images." *Journal of Endodontics*, vol. 46, no. 7, 2020, pp. 987-993. https://doi.org/10.1016/j.joen.2020.03.025.

[21] Muresanu, S. *et al.* "Artificial Intelligence Models for Clinical Usage in Dentistry with a Focus on Dentomaxillofacial CBCT: A Systematic Review." *Oral Radiology*, vol. 39, no. 1, 2023, pp. 18-40. https://doi.org/10.1007/s11282-022-00660-9.

[22] Abesi, F. *et al.* "Accuracy of Artificial Intelligence in the Detection and Segmentation of Oral and Maxillofacial Structures Using Cone-Beam Computed Tomography Images: A Systematic Review and Meta-Analysis." *Polish Journal of Radiology*, vol. 88, 2023, pp. e256-e263. https://doi.org/10.5114/pjr.2023.127624.

[23] Bonfanti-Gris, M. *et al.* "Evaluation of an AI Web-Based Software to Detect and Classify Dental Structures and Treatments in Panoramic Radiographs." *Journal of Dentistry*, vol. 126, 2022. https://doi.org/10.1016/j.jdent.2022.104301.

[24] Estai, M. *et al.* "Comparison of a Smartphone-Based Photographic Method with Face-to-Face Caries Assessment: A Mobile Teledentistry Model." *Telemedicine and e-Health*, vol. 23, no. 5, 2017, pp. 435-440. https://doi.org/10.1089/tmj.2016.0122.

[25] Duong, D.L. *et al.* "Automated Caries Detection with Smartphone Color Photography Using Machine Learning." *Health Informatics Journal*, vol. 27, no. 2, 2021. https://doi.org/10.1177/14604582211007530.

[26] Kuhnisch, J. *et al.* "Caries Detection on Intraoral Images Using Artificial Intelligence." *Journal of Dental Research*, vol. 101, no. 2, 2022, pp. 158-165. https://doi.org/10.1177/00220345211032524.

[27] Schwendicke, F. *et al.* "Artificial Intelligence in Dental Research: Checklist for Authors, Reviewers, Readers." *Journal of Dentistry*, vol. 107, 2021. https://doi.org/10.1016/j.jdent.2021.103610.

[28] Liu, X. *et al.* "Reporting Guidelines for Clinical Trial Reports for Interventions Involving Artificial Intelligence: The CONSORT-AI Extension." *Nature Medicine*, vol. 26, no. 9, 2020, pp. 1364-1374. https://doi.org/10.1038/s41591-0201034-x.

[29] Vasey, B. *et al.* "Reporting Guideline for the Early-Stage Clinical Evaluation of Decision Support Systems Driven by Artificial Intelligence: DECIDE-AI." *Nature Medicine*, vol. 28, no. 5, 2022, pp. 924-933. https://doi.org/10.1038/s41591-022-01772-9.

[30] Schwendicke, F. *et al.* "Artificial Intelligence for Oral and Dental Healthcare: Core Education Curriculum." *Journal of Dentistry*, vol. 128, 2023. https://doi.org/10.1016/j.jdent.2022.104363.