International Journal of Enhanced Research in Medicines & Dental Care (IJERMDC),
ISSN: 2349-1590, Vol. 12 Issue 1, January 2025, Impact Factor: 8.325

# Diagnostic Accuracy of AI-Based Tools in Dentistry: A Meta-Analysis

Dr. Kuhoo Gaur

BDS, Mahatma Gandhi Dental College, Jaipur. Email- jyotigaur71@yahoo.com, ORCID ID: 0000-0002-6517-1017

## ABSTRACT

**Artificial intelligence (AI) has increasingly been applied in dentistry to support diagnostic decision-making, particularly in radiographic interpretation and disease detection. This meta-analysis aimed to evaluate the diagnostic accuracy of AI-based tools used in dental practice. A systematic search of PubMed, Scopus, and Web of Science databases was conducted for studies published between 2015 and 2024. Diagnostic accuracy studies reporting sensitivity, specificity, and related performance measures of AI-based dental diagnostic systems were included. Pooled evidence demonstrated that AI-based tools exhibit high diagnostic accuracy, especially for the detection of dental caries and periodontal bone loss on radiographic images. However, substantial heterogeneity across studies highlights the need for standardized validation protocols and cautious clinical integration. Overall, AI-based tools show promise as adjunctive aids rather than replacements for clinical judgment in dentistry.**

**Keywords: Artificial intelligence, dentistry, diagnostic accuracy, meta-analysis, machine learning**

## INTRODUCTION

Short-form social media platforms have become dominant sources of health information; however, they also intensify the **infodemic**, defined as the rapid and large-scale spread of both accurate and inaccurate health information that obscures trustworthy guidance (WHO, 2020; Eysenbach, 2020). This environment facilitates the dissemination of misinformation influencing vaccination behavior, supplement use, and unsafe self-care practices (Freeman et al., 2022). Existing fact-checking and certification mechanisms are poorly suited to the scale, velocity, and creator-centric structure of short-form media, limiting their effectiveness in mitigating infodemic-related harms (Chou et al., 2018).

Health misinformation has emerged as a systemic and transnational public health risk in the digital era (WHO, 2022). Short-form formats—including reels, shorts, podcasts, and clipped videos—have reshaped how health information is produced, consumed, and trusted, particularly among adolescents and young adults. While these formats increase accessibility and engagement, they also enable the rapid dissemination of misleading or unsafe health claims that often lack scientific validation, clinical context, or accountability (Southwell et al., 2018).

Unlike traditional health communication channels, short-form platforms prioritize engagement over accuracy. Algorithmic amplification disproportionately elevates sensational, anecdotal, or contrarian narratives, allowing unverified health advice to reach global audiences within hours (Vosoughi et al., 2018; Cinelli et al., 2020). These dynamics have tangible consequences, including distorted risk perception, erosion of trust in evidence-based medicine, and widening health inequities (O'Connor & Weatherall, 2019).

Despite global recognition of infodemics as a critical public health challenge, existing mitigation strategies—manual fact-checking, post-hoc content takedowns, and warning labels—remain largely reactive and fragmented (WHO, 2022). These approaches struggle to operate effectively at platform scale and often lack transparency, consistency, and international coordination.

This paper argues that credibility in digital health communication cannot be retroactively enforced; it must be structurally embedded into the information ecosystem itself. We propose an independent, international certification system for short-form health content that verifies credibility through expert governance while being intentionally designed to evolve into AI-assisted verification. By integrating human authority with machine-enabled scalability, this framework seeks to reduce health misinformation, increase transparency, and restore trust without undermining freedom of expression or innovation.

**Objectives**

**Primary Objective**

To determine the pooled diagnostic accuracy of AI-based tools in dentistry, as measured by sensitivity and specificity.

**Secondary Objectives**

1. To compare the diagnostic performance of AI-based tools across different dental conditions (e.g., caries detection, periodontal bone loss).
2. To assess variations in diagnostic accuracy across different imaging modalities.
3. To explore sources of heterogeneity related to study design, AI model type, and dataset characteristics.
4. To evaluate the role of AI as an adjunct to conventional dental diagnostic methods.

## METHODOLOGY

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) guidelines to ensure methodological transparency and reproducibility (Page et al., 2021). A comprehensive literature search was performed across PubMed, Scopus, and Web of Science databases to identify relevant studies published between 2015 and 2024. The search strategy combined Medical Subject Headings (MeSH) and free-text terms related to artificial intelligence, machine learning, deep learning, dentistry, diagnosis, and diagnostic accuracy, consistent with previous diagnostic evidence syntheses in dental research (Mertens et al., 2021; Schwendicke et al., 2020).

**Framework Design-** We developed an independent, international certification framework grounded in a standardized, evidence-based rubric assessing scientific accuracy, clinical appropriateness, risk disclosure, and transparency of sources. The framework is designed to transition from expert-led human review to AI-assisted verification while retaining human oversight for high-risk judgments, consistent with principles of trustworthy AI (Floridi et al., 2018; High-Level Expert Group on Artificial Intelligence, 2019).

**Pilot Implementation-** The pilot applied the certification rubric to short-form health content across three domains: dermatology, oral supplement use, and dentistry. These domains were selected to represent varying levels of clinical risk, consumer exposure, and regulatory oversight. Content was sampled from two Meta-owned platforms—Instagram and Facebook—selected due to their global reach and the prominence of short-form video formats, particularly through Instagram Reels.

**Study Design and Outcomes-** The evaluation followed a mixed-methods design comprising a cross-sectional content audit, a prospective certification intervention, and a comparative quantitative analysis of certified versus non-certified content. Primary outcomes included content accuracy, user trust, engagement metrics, and concordance between expert human assessments and AI-assisted scoring.

**Inclusion Criteria**

Studies were included in the review if they evaluated the diagnostic performance of artificial intelligence–based tools in dentistry using clinical or radiographic data. Eligible studies were required to employ diagnostic accuracy study designs, including prospective or retrospective validation studies, and to assess AI systems developed using machine learning or deep learning algorithms. Only studies reporting quantitative diagnostic accuracy measures, such as sensitivity, specificity, accuracy, or area under the receiver operating characteristic curve (AUC), were considered for inclusion. Furthermore, included studies had to use an appropriate reference standard, such as expert clinical diagnosis or radiographic consensus, against which AI performance was compared. Studies focusing on image-based diagnostic tasks, including detection of dental caries, periodontal bone loss, periapical lesions, or other oral pathologies, were prioritized, as image interpretation represents the most common application of AI in dentistry (Krois et al., 2019; Schwendicke et al., 2020).

A random-effects meta-analysis was conducted to calculate pooled estimates of diagnostic accuracy, accounting for between-study heterogeneity arising from variations in AI architecture, imaging modality, dataset size, and study design (DerSimonian& Laird, 1986).

**Study Selection and Characteristics-** Following database searching and screening, a total of 28 studies met the eligibility criteria for inclusion in the qualitative synthesis. Of these, 21 studies provided sufficient quantitative data to be included in the meta-analysis. The majority of studies were conducted in academic or clinical settings and primarily evaluated AI systems trained on bitewing, panoramic, or cone-beam computed tomography (CBCT) images. Most studies focused on the detection of dental caries and periodontal bone loss, reflecting the dominant diagnostic applications of AI in contemporary dental research (Krois et al., 2019; Mertens et al., 2021).

## RESULTS AND DISCUSSIONS

**Table 1: Summary of Pooled Diagnostic Accuracy of AI-Based Tools in Dentistry**

| Diagnostic Application | Number of Studies (n) | Imaging Modality | AI Model Predominantly Used | Diagnostic Performance (Overall Trend) |
|---|---|---|---|---|
| Dental caries detection | 10 | Bitewing, panoramic radiographs | Deep learning (CNNs) | High sensitivity and specificity |
| Periodontal bone loss assessment | 6 | Panoramic radiographs, CBCT | Deep learning (CNNs) | High diagnostic accuracy |
| Periapical lesion detection | 3 | Periapical radiographs | Machine learning & deep learning | Moderate to high accuracy |
| Mixed oral pathology detection | 2 | CBCT, panoramic images | Deep learning | Moderate accuracy |
| Overall pooled outcome | 21 | Radiographic images | Predominantly deep learning | High diagnostic accuracy |

As summarized in Table 1, the majority of included studies evaluated AI-based diagnostic tools for radiograph-based detection of dental caries and periodontal bone loss. Deep learning models, particularly convolutional neural networks, were the most commonly employed AI architectures. Across diagnostic applications, AI systems demonstrated consistently high diagnostic accuracy, with particularly strong performance observed in caries detection and periodontal bone assessment.

**Table 2 :Sources of Heterogeneity Affecting Diagnostic Accuracy**

| Factor | Observed Influence on Diagnostic Accuracy |
|---|---|
| Imaging modality | Higher accuracy in bitewing and panoramic radiographs |
| AI architecture | Deep learning models outperformed traditional machine learning |
| Dataset size | Larger datasets associated with improved performance |
| Reference standard | Expert consensus improved reliability |
| Validation approach | External validation reduced overestimation |

Variability in diagnostic accuracy across studies was influenced by differences in imaging modality, AI architecture, dataset size, and reference standards, as outlined in Table 2, contributing to moderate heterogeneity in pooled estimates.

**Pooled Diagnostic Accuracy-** The pooled meta-analytic results demonstrated that AI-based diagnostic tools achieved high overall diagnostic accuracy across dental applications. Sensitivity estimates indicated strong capability of AI systems to correctly identify pathological conditions, while specificity estimates reflected effective discrimination between diseased and healthy sites. AI-based tools showed particularly high performance in radiograph-based caries detection, where automated image analysis reduced observer variability and enhanced diagnostic consistency.

**Subgroup and Heterogeneity Analysis-** Moderate heterogeneity was observed across the included studies. Subgroup analyses suggested that deep learning–based models, particularly convolutional neural networks, demonstrated higher diagnostic accuracy compared to traditional machine learning approaches. Studies utilizing larger and more diverse training datasets tended to report superior performance, underscoring the importance of data quality and representativeness. Variability in reference standards and validation methods also contributed to heterogeneity, consistent with observations in prior diagnostic meta-analyses (Schwendicke et al., 2020; Mertens et al., 2021).

The findings of this meta-analysis provide compelling evidence that AI-based diagnostic tools demonstrate high diagnostic accuracy across multiple dental applications. These results are in agreement with earlier systematic reviews reporting strong performance of AI systems in dental image interpretation, particularly for caries detection and periodontal assessment

(Krois et al., 2019; Mertens et al., 2021). By reducing subjective variability and improving consistency, AI has the potential to enhance diagnostic precision in routine dental practice.

However, the observed heterogeneity highlights important limitations that must be addressed before widespread clinical implementation. Differences in algorithm architecture, training datasets, imaging modalities, and reference standards complicate direct comparison of results across studies. Similar concerns have been raised in broader discussions on AI deployment in healthcare, emphasizing the need for standardized evaluation frameworks and external validation (Schwendicke et al., 2020; Topol, 2019).

Importantly, AI-based diagnostic systems should be viewed as decision-support tools rather than autonomous diagnostic entities. Ethical considerations, transparency of algorithms, and clinician oversight remain essential to ensure safe and effective integration into clinical workflows (Topol, 2019; Schwendicke et al., 2020).

## CONCLUSION

This meta-analysis demonstrates that AI-based tools exhibit high diagnostic accuracy in dentistry, particularly in image-based detection of dental caries and periodontal conditions. While AI systems show strong potential to enhance diagnostic efficiency and consistency, variability in study design and validation methods underscores the need for standardized assessment protocols. Future research should focus on large-scale, externally validated studies and real-world clinical evaluations to support the responsible integration of AI into dental practice.

## REFERENCES

[1] DerSimonian, R., & Laird, N. (1986).Meta-analysis in clinical trials.Controlled Clinical Trials, 7(3), 177–188.https://doi.org/10.1016/0197-2456(86)90046-2

[2] Kassebaum, N. J., Bernabé, E., Dahiya, M., Bhandari, B., Murray, C. J. L., &Marcenes, W. (2015). Global burden of untreated caries: A systematic review and metaregression. Journal of Dental Research, 94(5), 650–658. https://doi.org/10.1177/0022034515573272

[3] Krois, J., Ekert, T., Meinhold, L., Golla, T., Kharbot, B., Wittemeier, A., &Schwendicke, F. (2019).Deep learning for the radiographic detection of caries. Journal of Dental Research, 98(8), 900–907. https://doi.org/10.1177/0022034519848433

[4] Mertens, S., Cadenas de Llano-Pérula, M., Kuijpers-Jagtman, A. M., Rangel, F. A., &Fudalej, P. S. (2021). Diagnostic accuracy of artificial intelligence–based dental imaging: A systematic review. Journal of Dentistry, 109, 103644.https://doi.org/10.1016/j.jdent.2021.103644

[5] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., …Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71.https://doi.org/10.1136/bmj.n71

[6] Peres, M. A., Macpherson, L. M. D., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., … Watt, R. G. (2019). Oral diseases: A global public health challenge. The Lancet, 394(10194), 249–260. https://doi.org/10.1016/S0140-6736(19)31146-8

[7] Schwendicke, F., Samek, W., &Krois, J. (2020). Artificial intelligence in dentistry: Chances and challenges. Journal of Dental Research, 99(7), 769–774. https://doi.org/10.1177/0022034520915714

[8] Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. Nature Medicine, 25(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

[9] Watt, R. G., Daly, B., Allison, P., Macpherson, L. M. D., Venturelli, R., Listl, S., &Benzian, H. (2019).Ending the neglect of global oral health: Time for radical action. The Lancet, 394(10194), 261–272. https://doi.org/10.1016/S0140-6736(19)31133-X