



Accuracy of Artificial Intelligence-Assisted CBCT Interpretation for the Detection of Endodontic Pathologies and Findings: A Systematic Review and Meta-analysis

1. Faezeh Niroumandfar[✉]: Faculty of Dentistry, Isf.C., Islamic Azad University, Isfahan, Iran
2. Sara Karami Amidabadi[✉]: Faculty of Dentistry, Qa.C., Islamic Azad University, Qazvin, Iran
3. Mahshid Sohofi[✉]: Faculty of Dentistry, Isf.C., Islamic Azad University, Isfahan, Iran
4. Sina Shahabi[✉]: Bogomolets National Medical University, Ukraine
5. Elmira Fardi[✉]: Faculty of Dentistry, TeMS.C., Islamic Azad University, Tehran, Iran
6. Sanaz Keyvanfar^{✉*}: Department of Periodontics, School of Dentistry, Birjand University of Medical Sciences, Birjand, Iran

Corresponding author email address: sanazkeyvanfar@bums.ac.ir

ABSTRACT

Artificial intelligence (AI) has increasingly been applied to cone-beam computed tomography (CBCT) interpretation in endodontics. However, reported diagnostic performance varies across endodontic findings and AI-based approaches. This systematic review and meta-analysis evaluated the diagnostic accuracy of AI-assisted CBCT interpretation for detecting endodontic pathologies and findings. We searched PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Library for studies published between January 2016 and May 2026. We included studies evaluating AI-assisted detection, classification, or segmentation of endodontic pathologies on dental CBCT in human subjects. We extracted data on sensitivity, specificity, and accuracy, and assessed methodological quality using QUADAS-2. We used random-effects models to calculate pooled estimates, and assessed heterogeneity using Cochran's Q and I² statistics. We performed subgroup analyses by diagnostic finding. We included 12 studies, contributing 31 diagnostic evaluations for sensitivity and 25 for specificity and accuracy. The pooled sensitivity, specificity, and accuracy were 79.82% (95% CI: 73.86–85.77), 86.80% (95% CI: 82.69–90.91), and 84.74% (95% CI: 80.53–88.96), respectively. Sensitivity was highest for periapical osteolytic lesions (97.10%) and apical periodontitis (88.27%), while specificity was highest for apical periodontitis (96.27%). For periapical radiolucency, pooled sensitivity, specificity, and accuracy were 87.89%, 71.46%, and 78.89%, respectively. Sensitivity, specificity, and accuracy differed significantly across diagnostic subgroups (all P < 0.001). Substantial heterogeneity was observed across the analyses. AI-assisted CBCT interpretation demonstrated clinically relevant diagnostic accuracy for endodontic findings, with performance varying by pathology.

Keywords: Artificial Intelligence, cone-beam computed tomography, Endodontic.

How to cite this article:

Niroumandfar, F., Karami Amidabadi, S., Sohofi, M., Shahabi, S., Fardi, E., & Keyvanfar, S. (2026). Accuracy of Artificial Intelligence-Assisted CBCT Interpretation for the Detection of Endodontic Pathologies and Findings: A Systematic Review and Meta-analysis. *Journal of Oral and Dental Health Nexus*, 3(3), 1-13. <https://doi.org/10.61838/jodhn.31>

Introduction

Accurate interpretation of dental imaging is central to endodontic diagnosis because many pathologies are defined by subtle changes in tooth structure, root morphology, and periapical bone (1). Conventional periapical radiography remains widely used, but its two-dimensional nature can obscure anatomical structures and lesions through superimposition (2). These limitations are particularly relevant when abnormalities extend buccolingually or involve complex root anatomy. A systematic review has shown that periapical

radiographs have limited ability to reliably characterize periapical disease compared with histopathological findings (3).

Cone-beam computed tomography (CBCT) provides three-dimensional visualization of the dentoalveolar region and can reveal findings that are difficult to identify on conventional radiographs (4). In endodontics, CBCT is used for selected diagnostic problems, including assessment of periapical lesions, root fractures, root resorption, complex canal anatomy, treatment complications, and the relationship of lesions to adjacent structures (5). The 2025 joint position statement of the

American Association of Endodontists and the American Academy of Oral and Maxillofacial Radiology recommends limited-field-of-view CBCT when conventional imaging is inconclusive or insufficient for specific clinical questions (6). However, CBCT interpretation remains affected by image noise, scatter, beam-hardening, and other artefacts, as well as differences in spatial resolution and observer experience (7).

Artificial intelligence (AI), particularly deep learning, has been increasingly investigated for automated dental image analysis (8). AI systems can detect, segment, and classify radiographic findings by learning image features from labeled datasets (9). In endodontics, CBCT-based AI applications include detecting periapical lesions, segmenting teeth and root canals, identifying complex canal anatomy, and detecting vertical root fractures and other treatment-related findings (10). Setzer et al. reported that a deep learning model for detecting and segmenting periapical lesions on CBCT achieved high diagnostic performance, demonstrating the potential of automated image analysis for endodontic diagnosis (11). Subsequent systematic reviews have reported similarly high sensitivity and specificity for AI-based detection of periapical radiolucencies and other endodontic findings (12).

Despite these findings, the diagnostic performance of AI-assisted CBCT interpretation has not been uniform across applications. Differences in study populations, CBCT acquisition parameters, image quality, reference standards, AI architectures, and validation strategies can substantially affect reported accuracy (13). A systematic review and meta-analysis of AI for periapical radiolucency detection reported pooled sensitivity of 0.94 and specificity of 0.96, but also identified methodological limitations, including risk of bias, applicability concerns, and limited external validation (14). Recent evidence continues to show that many studies use relatively small or single-centre datasets, while validation across different imaging devices and clinical populations remains limited (15). Therefore, high performance reported by an individual model does not necessarily reflect consistent diagnostic accuracy in routine clinical settings. The rapid expansion of AI research in endodontic CBCT interpretation has produced evidence across several distinct diagnostic tasks, but the findings remain fragmented. Existing

reviews have primarily focused on individual applications, such as periapical radiolucency detection, pulp-space segmentation, or vertical root fracture detection (14, 16, 17). A broader quantitative assessment is needed to determine whether the diagnostic performance of AI-assisted CBCT interpretation is consistent across different endodontic pathologies and findings.

Different endodontic abnormalities vary in imaging characteristics and diagnostic difficulty, and AI performance may therefore differ across clinical findings. Variations in AI models, reference standards, imaging protocols, and patient populations may also contribute to differences in reported diagnostic accuracy. This systematic review and meta-analysis therefore evaluated the accuracy of AI-assisted CBCT interpretation for endodontic pathologies and findings and explored factors associated with variation in diagnostic performance.

Materials and Methods

Study design

This systematic review and diagnostic accuracy meta-analysis evaluated the performance of artificial intelligence (AI)-assisted CBCT interpretation for detecting endodontic pathologies and imaging findings. The review was designed and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Diagnostic Test Accuracy Studies (PRISMA-DTA) guidelines (18). We defined all methodological procedures, including eligibility criteria, outcomes, risk-of-bias assessment, and the statistical analysis plan, a priori to reduce the potential for post hoc modifications during the review process.

PICO framework and research question

The research question was formulated a priori according to a diagnostic adaptation of the PICO framework:

Population (P): The population comprised patients undergoing dental CBCT examination in whom endodontic pathologies or findings were assessed. Eligible studies included patients with suspected or established endodontic disease, previous or ongoing

root canal treatment, or clinical indications requiring CBCT assessment of teeth and periapical structures.

Index test (I): The index test was AI-assisted interpretation of CBCT images to detect, classify, segment, or identify endodontic pathologies and findings. Eligible AI approaches included machine learning and deep learning systems, whether fully automated or used as computer-assisted diagnostic tools.

Comparison (C): The reference standard comprised the method each primary study used to establish the presence or absence of the target endodontic pathology or finding.

Outcomes (O): The primary outcomes were the diagnostic sensitivity and specificity of AI-assisted CBCT interpretation for detecting endodontic pathologies and findings. Secondary diagnostic outcomes included diagnostic accuracy, positive likelihood ratio (LR+), negative likelihood ratio (LR-), diagnostic odds ratio (DOR), area under the receiver operating characteristic curve (AUROC), and positive and negative predictive values when reported or calculable from the available data. We extracted true-positive (TP), false-positive (FP), false-negative (FN), and true-negative (TN) values when available to reconstruct 2×2 diagnostic contingency tables.

Accordingly, the research question addressed the diagnostic accuracy of AI-assisted CBCT interpretation for identifying endodontic pathologies and findings.

Eligibility criteria

Inclusion criteria: Studies were eligible if they (1) evaluated human dental CBCT images; (2) assessed an AI-based or AI-assisted method for detecting endodontic pathologies or findings; (3) used an appropriate reference standard; (4) reported sufficient data to determine diagnostic accuracy; and (5) were original studies published in English. We included both retrospective and prospective diagnostic accuracy studies.

Exclusion criteria: Case reports, case series, animal or in vitro studies, reviews, meta-analyses, editorials, letters, conference abstracts, and studies without an appropriate reference standard or sufficient diagnostic data were excluded. We also excluded studies that evaluated AI exclusively on non-CBCT images or without an endodontic diagnostic outcome.

Information sources and search strategy

We conducted a comprehensive systematic literature search in PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Library for eligible studies published between January 2016 and May 2026. We selected this search period to capture contemporary applications of AI in dental CBCT interpretation while maintaining a defined, reproducible search window. In addition, we manually screened the reference lists of all included studies and relevant systematic reviews to identify potentially eligible studies not retrieved through the electronic database searches.

The search strategy combined controlled vocabulary and free-text terms related to three principal concepts: artificial intelligence, CBCT, and endodontic diagnosis. We adapted search terms to each database's indexing system. The core search strategy was:

("artificial intelligence" OR "machine learning" OR "deep learning" OR "neural network" OR "convolutional neural network" OR "computer-aided diagnosis") AND ("cone-beam computed tomography" OR CBCT OR "cone beam CT") AND (endodont OR "periapical lesion" OR "periapical patholog" OR "root fracture" OR "vertical root fracture" OR "root resorption" OR "root canal" OR "endodontic finding").

We imposed no restrictions based on AI architecture, CBCT device, voxel size, field of view, or specific endodontic pathology during the initial search.

Study selection process

All records retrieved from the electronic databases were imported into EndNote X8, and duplicate records were identified and removed before screening. Two reviewers independently screened the titles and abstracts of all retrieved records according to the predefined eligibility criteria. We then assessed potentially relevant studies by reviewing full texts. We excluded studies that did not meet the inclusion criteria and recorded the primary reason for exclusion at the full-text stage. The two reviewers resolved disagreements through discussion and consensus. When consensus could not be reached, a third reviewer independently assessed the study and made the final decision. We documented the study identification and selection process using the PRISMA flow diagram.

Data extraction

Two reviewers independently extracted data using a standardized form. Extracted variables included author, publication year, country, study design, sample size, patient characteristics, tooth type, target endodontic pathology or finding, and CBCT characteristics. AI-related variables included the AI model or algorithm, learning approach, validation strategy, and diagnostic task. We also recorded reference standard characteristics and diagnostic criteria; we extracted diagnostic data, including TP, FP, FN, and TN values.

Assessment of risk of bias

Two reviewers independently assessed the methodological quality and risk of bias of the included studies using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool. The assessment covered patient selection, index test, reference standard, and flow and timing, with each domain classified as having low, high, or unclear risk of bias. We assessed applicability concerns for patient selection, the index test, and the reference standard. We resolved disagreements through discussion or consultation with a third reviewer.

Statistical analysis

Statistical analyses were performed using Stata version 17. Diagnostic accuracy was assessed using sensitivity, specificity, and accuracy. When available, TP, FP, FN, and TN values were used to construct 2×2 contingency tables. Pooled estimates were calculated using a random-effects model with restricted maximum likelihood (REML) and reported with 95% confidence

intervals (CIs). Separate meta-analyses were conducted for sensitivity, specificity, and accuracy. Heterogeneity was assessed using Cochran's Q test, I^2 , and τ^2 . Prespecified subgroup analyses were performed according to the type of endodontic pathology or finding, including apical periodontitis, periapical lesion, periapical osteolytic lesion, and periapical radiolucency. Differences between subgroups were assessed using the Qb statistic.

Forest plots were generated to present individual and pooled estimates. Statistical significance was defined as $P < 0.05$.

Results

PRISMA Flow Description

The study selection process followed PRISMA guidelines and is illustrated in Figure 1. A total of 248 records were initially identified. After removal of duplicates, 172 unique records remained for title and abstract screening. During this phase, 47 studies were excluded because they were irrelevant to the research question, used inappropriate study designs (e.g., reviews, case reports, case series, or conference abstracts), evaluated non-CBCT imaging, or did not assess AI-assisted detection of endodontic pathologies or findings. In the eligibility assessment stage, we reviewed 25 full-text articles in detail. Of these, 13 studies were excluded because they lacked an appropriate reference standard, did not provide sufficient diagnostic accuracy data, did not evaluate an endodontic outcome, or did not meet the predefined inclusion criteria. Ultimately, 12 studies met the predefined eligibility criteria and were included in the present systematic review and diagnostic accuracy meta-analysis.

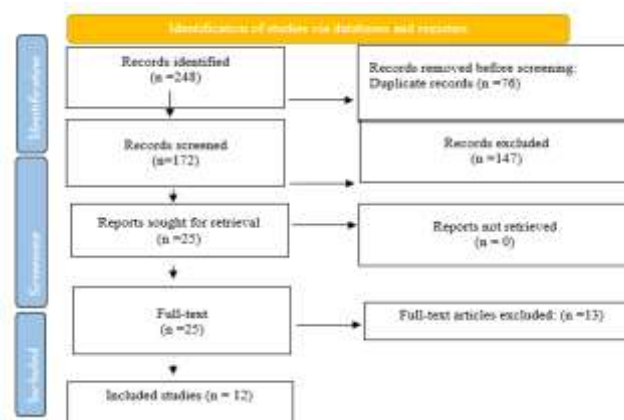


Figure 1. PRISMA 2020 Flow Diagram.

Characteristics of Included Studies

Table 1 summarizes the characteristics of the 12 included studies. All studies used CBCT-based datasets and evaluated AI-assisted detection, classification, or

segmentation of endodontic pathologies, predominantly periapical lesions. Most studies were retrospective and used CNN- or U-Net-based models, with validation performed using independent test sets or cross-validation.

Table 1. Characteristics of Included Studies.

Study, year	Design	Sample	Population/teeth	Target finding	CBCT dataset	AI model	Diagnostic task	Reference standard	Validation
Al Fartousi et al, 2026 (19)	Retrospective diagnostic accuracy	358 teeth/167 scans	Root-filled teeth; 72.1% with apical lesions	Apical lesion and endodontic findings	Private practice, Germany; voxel 125–160 µm	Diagnocat AIS, CNN	Tooth-level detection	Two blinded endodontists; senior expert for disagreements	Independent evaluation; no external cohort
Öksüzöğlu et al, 2026 (20)	Retrospective development/validation	140 scans; 347 lesions	Adults with ≥1 periapical lesion	Periapical lesion	Türkiye; voxel 0.3 mm	nnU-Net v2	3D lesion segmentation/detection	Manual segmentation; senior endodontist review	80/10/10% train/validation/test
Allihaibi et al, 2025 (21)	Retrospective diagnostic accuracy	134 molars; 327 roots	Molars requiring root canal treatment	Periapical radiolucency	London; 4×4 cm FOV	Diagnocat, deep CNN	Tooth- and root-level detection	Two blinded experienced endodontists	Independent evaluation; no external cohort
Chadgal, 2025 (21)	Retrospective comparative	180 patient; 250 scans	Patients with suspected periapical lesions	Early periapical pathology	Tertiary dental centre; parameters NR	3D voxel-based CNN	Lesion detection	Three endodontists; consensus	Separate test set; no external validation
Chau et al, 2025 (22)	Model development/comparison	185 scans	NR	Periapical lesion	CBCT parameters NR	CBCT-SAM; U-Net, PAL-Net	Detection and segmentation	Manual segmentation; radiologist validation	Training/validation
Hadzic et al, 2024 (23)	Retrospective validation	195 CBCT images; 2,947 teeth	Adults with routine CBCT scans	Periapical lesion	Austria; voxel 0.2 mm	Spatial Configuration-Net + U-Net	Tooth-level detection	Consensus of 3 investigators	Independent test set; same centre
Kazimierzak et al, 2024 (24)	Retrospective diagnostic accuracy	49 patients; 1,223 teeth	Mean age 41 y; suspected periapical lesions	Periapical lesion	Poland; voxel 0.3 mm	Diagnocat, CNN	Tooth-level detection	Consensus of orthodontist and radiologist	Independent evaluation; no external cohort
Fu et al, 2024 (25)	Retrospective multicentre	279 patients; 4,951 teeth	Permanent dentition; AP prevalence 6–13%	Apical periodontitis	4 institutions, China	PAL-Net, 3D U-Net CNN	Detection and segmentation	Endodontists + pathology/radiology records	5-fold CV; independent and 3-centre external validation
Calazans et al, 2022 (26)	Retrospective diagnostic accuracy	1,000 treated molars	Endodontically treated maxillary molars	Periapical lesion	Brazil; voxel 0.13–0.40 mm	Siamese CNN; DenseNet-121/VGG-16	Tooth-level classification	Oral radiologist	60/20/20% train/validation/test
Kimbauer et al, 2022 (27)	Retrospective development/validation	144 volumes; 2,128 teeth	Adult teeth with completed roots	Periapical lesion	Austria; voxel 0.2 mm	Spatial Configuration-Net + 3D U-Net	Detection and segmentation	Oral radiologist	4-fold cross-validation; no external validation
Orhan et al, 2020 (28)	Retrospective diagnostic accuracy	153 lesions/109 patients	Teeth with periapical lesions	Periapical lesion	Türkiye; voxel 0.075–0.2 mm	Diagnocat, U-Net-like CNN	Detection, localization, volume	Oral radiologist	Hold-out test set; no external validation
Setzer et al, 2020 (11)	Retrospective algorithm study	20 volumes; 61 roots	Patients referred for endodontic treatment	Periapical lesion	USA; voxel 0.125 mm	Multilabel U-Net	Root-level detection and segmentation	Three clinical reviewers	5-fold cross-validation; no external validation

Risk of bias

In the patient selection domain, 6 studies had a low risk of bias, while 6 had an unclear risk. For the index test, 10 studies had a low risk of bias, and 2 had an unclear risk. All 12 studies were judged to have a low risk

of bias for the reference standard and flow and timing domains. Regarding applicability, 6 studies showed low concern and 6 studies showed unclear concern in the patient selection domain. Applicability concerns for the index test were low in 10 studies and unclear in 2 studies. All studies showed low concern regarding the applicability of the reference standard (Table 2).

Table 2. QUADAS-2 assessment of the included studies.

Study	Patient selection	Index test	Reference standard	F & timing	Applicability: Patient selection	Applicability: Index test	Applicability: Reference standard
Al Fartousi et al., 2026 (19)	●	●	●	●	●	●	●
Öksüzoglu et al., 2026 (20)	●	●	●	●	●	●	●
Allihaibi et al., 2025 (29)	●	●	●	●	●	●	●
Chadgal, 2025 (21)	●	●	●	●	●	●	●
Chau et al., 2025 (22)	●	●	●	●	●	●	●
Hadzic et al., 2024 (23)	●	●	●	●	●	●	●
Kazimierczak et al., 2024 (24)	●	●	●	●	●	●	●
Fu et al., 2024 (25)	●	●	●	●	●	●	●
Calazans et al., 2022 (26)	●	●	●	●	●	●	●
Kirnbauer et al., 2022 (27)	●	●	●	●	●	●	●
Orhan et al., 2020 (28)	●	●	●	●	●	●	●
Setzer et al., 2020 (11)	●	●	●	●	●	●	●

Low risk ● High risk ● Unclear risk ●

Diagnostic sensitivity

Subgroup analysis showed that sensitivity of AI-assisted CBCT interpretation differed by the type of endodontic finding. Across the 31 diagnostic evaluations, the overall pooled sensitivity was 79.82% (95% CI, 73.86–85.77%). Pooled sensitivity was highest for periapical osteolytic lesions (97.10%; 95% CI, 94.81–99.39%), followed by apical periodontitis (88.27%; 95% CI, 76.98–99.56%) and periapical radiolucency (87.89%; 95% CI, 82.42–93.36%). The lowest pooled sensitivity was observed for periapical lesions, at 74.38% (95% CI, 66.38–82.37%). The estimate for periapical osteolytic lesions was based on a single study (Figure 2).

Substantial between-study variation was observed in the apical periodontitis and periapical lesion subgroups. For apical periodontitis, sensitivity estimates ranged from 61.2% to 99.0%, with very high heterogeneity ($Q = 160.08$, $P < 0.001$; $I^2 = 96.76\%$; $\tau^2 = 187.463$). In the periapical lesion subgroup, individual sensitivity

estimates ranged from 35.71% to 100.0%, and heterogeneity was also very high ($Q = 714.73$, $P < 0.001$; $I^2 = 97.98\%$; $\tau^2 = 297.860$). The periapical radiolucency subgroup showed lower heterogeneity, with a pooled sensitivity of 87.89% and I^2 of 54.99% ($Q = 6.70$, $P = 0.082$) (Figure 2).

The between-subgroup comparison was statistically significant ($Qb = 35.96$, $\chi^2(3)$, $P < 0.001$), indicating that sensitivity differed across the four endodontic finding categories. Within the periapical radiolucency subgroup, sensitivity was higher for tooth-level detection than for root-level detection in both preoperative and postoperative assessments. The sensitivity estimates were 93.9% for preoperative tooth-level detection, 86.2% for preoperative root-level detection, 88.6% for postoperative tooth-level detection, and 80.3% for postoperative root-level detection. Overall heterogeneity across the diagnostic evaluations was extremely high ($Q = 952.89$, $P < 0.001$; $I^2 = 99.15\%$; $\tau^2 = 259.870$) (Figure 2).

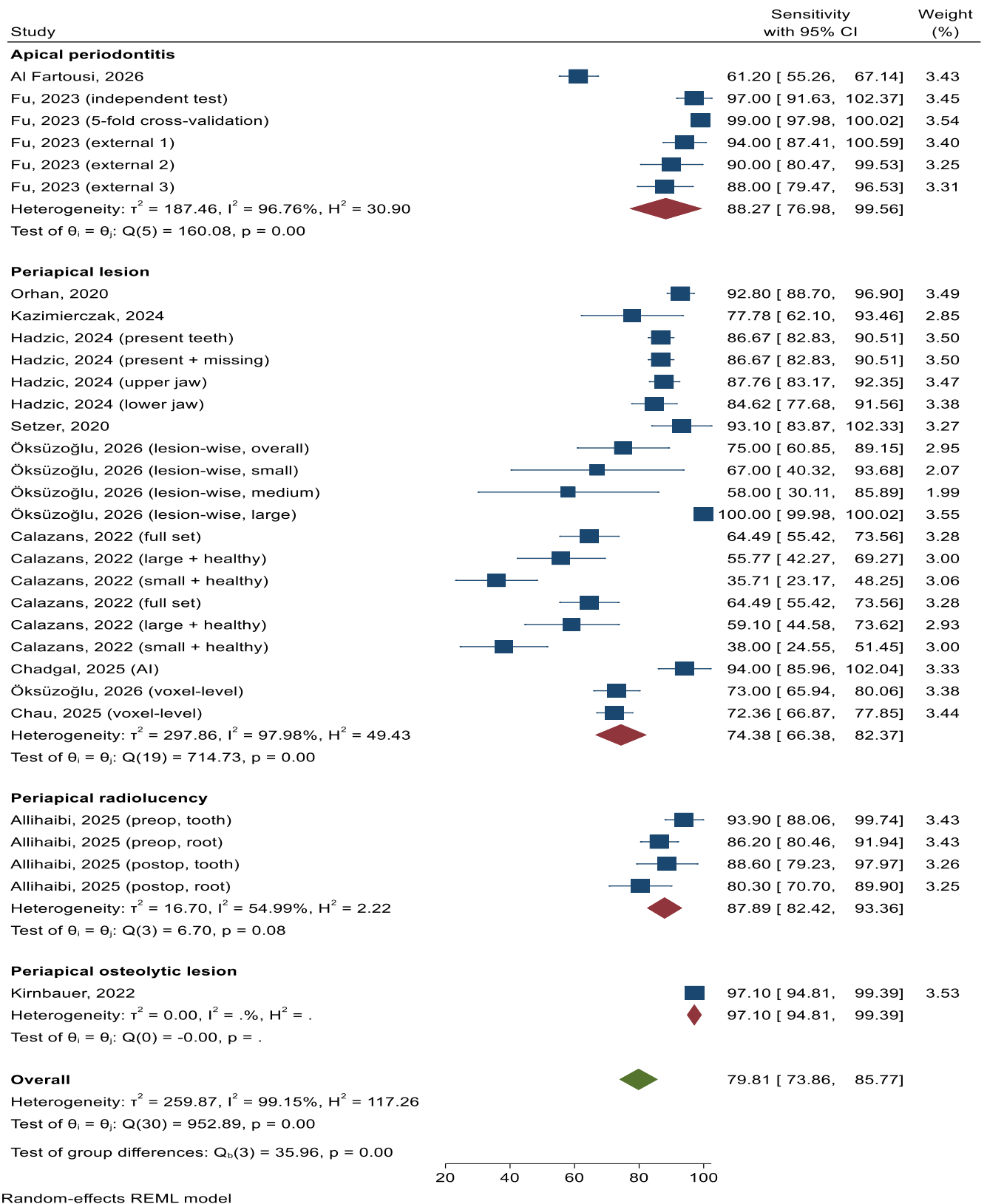


Figure 2. Forest plot of pooled sensitivity of artificial intelligence-assisted CBCT interpretation for the detection of endodontic pathologies and findings.

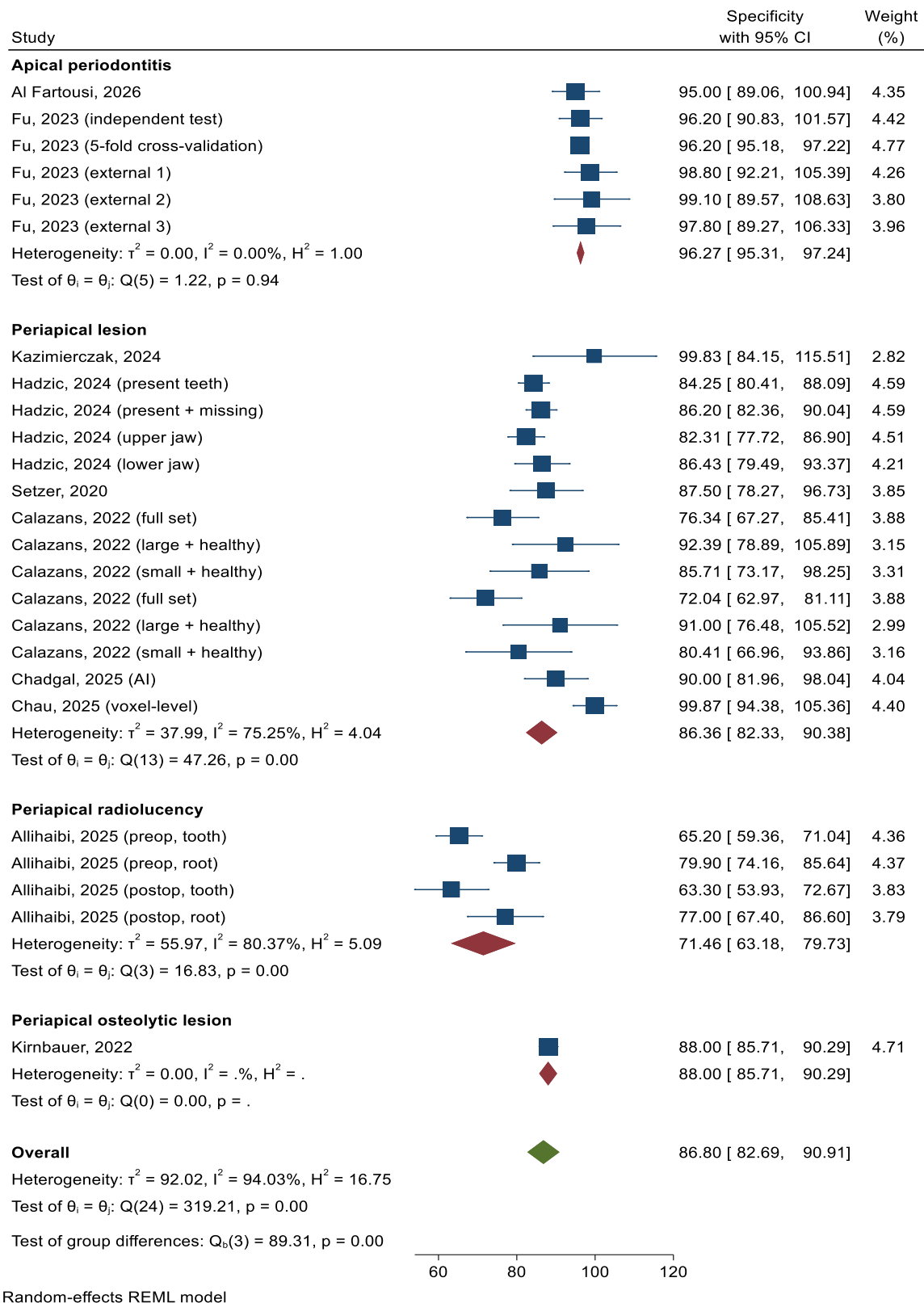


Figure 3. Forest plot of pooled specificity of artificial intelligence-assisted CBCT interpretation for the detection of endodontic pathologies and findings.

Diagnostic specificity

Subgroup analysis showed that the specificity of AI-assisted CBCT interpretation differed across the four endodontic finding categories. Across 25 diagnostic evaluations, the overall pooled specificity was 86.80% (95% CI, 82.69–90.91%). The highest pooled specificity was observed for apical periodontitis (96.27%; 95% CI, 95.31–97.24%), followed by periapical osteolytic lesions (88.00%; 95% CI, 85.71–90.29%) and periapical lesions (86.36%; 95% CI, 82.34–90.38%). The lowest pooled specificity was observed for periapical radiolucency, at 71.46% (95% CI, 63.18–79.73%) (Figure 3).

For apical periodontitis, specificity estimates ranged from 95.0% to 99.1%, with no evidence of statistical heterogeneity ($Q = 1.22$, $P = 0.943$; $I^2 = 0\%$). In contrast, substantial heterogeneity was observed for periapical lesions ($Q = 47.26$, $P < 0.001$; $I^2 = 75.25\%$; $\tau^2 = 37.988$) and periapical radiolucency ($Q = 16.83$, $P = 0.001$; $I^2 = 80.37\%$; $\tau^2 = 55.967$). The periapical osteolytic lesion subgroup included a single evaluation and therefore did not permit estimation of between-study heterogeneity. Overall heterogeneity was very high ($Q = 319.21$, $P < 0.001$; $I^2 = 94.03\%$; $\tau^2 = 92.020$) (Figure 3).

The test for subgroup differences was statistically significant ($Q_b = 89.31$, $\chi^2(3)$, $P < 0.001$), indicating that specificity varied substantially by endodontic finding. Within the periapical radiolucency subgroup, specificity ranged from 63.3% to 79.9%. Preoperative specificity was 65.2% at the tooth level and 79.9% at the root level, whereas postoperative specificity was 63.3% at the tooth level and 77.0% at the root level. These findings indicate that specificity varied more across the included evaluations for some diagnostic targets than for others, with the most consistent specificity observed for apical periodontitis (Figure 3).

Accuracy

Subgroup analysis showed that accuracy of AI-assisted CBCT interpretation varied by endodontic finding. Across 25 diagnostic evaluations, the overall pooled accuracy was 84.74% (95% CI, 80.53–88.96%). The highest pooled accuracy was observed for apical periodontitis, at 92.75% (95% CI, 83.94–101.57%), followed by periapical osteolytic lesions, at 88.90% (95% CI, 86.61–91.19%), and periapical lesions, at 82.67% (95% CI, 77.08–88.26%). The lowest pooled accuracy was observed for periapical radiolucency, at 78.89% (95% CI, 74.83–82.95%). The estimate for periapical osteolytic lesions was based on a single diagnostic evaluation.

Substantial heterogeneity was observed within the apical periodontitis subgroup, with accuracy estimates ranging from 70.7% to 98.3% ($Q = 71.32$, $P < 0.001$; $I^2 = 94.62\%$; $\tau^2 = 110.158$). The periapical lesion subgroup also showed considerable heterogeneity, with accuracy estimates ranging from 65.99% to 99.35% ($Q = 74.20$, $P < 0.001$; $I^2 = 87.90\%$; $\tau^2 = 90.754$). In contrast, heterogeneity within the periapical radiolucency subgroup was not statistically significant ($Q = 3.97$, $P = 0.265$; $I^2 = 21.67\%$; $\tau^2 = 3.783$). Overall heterogeneity across the 25 diagnostic evaluations was very high ($Q = 328.76$, $P < 0.001$; $I^2 = 94.35\%$; $\tau^2 = 97.517$) (Figure 4).

The test for subgroup differences was statistically significant ($Q_b = 21.42$, $\chi^2(3)$, $P < 0.001$), indicating that pooled accuracy differed across the four endodontic finding categories. Within the periapical radiolucency subgroup, accuracy was 79.1% for preoperative tooth-level detection, 82.6% for preoperative root-level detection, 71.6% for postoperative tooth-level detection, and 77.7% for postoperative root-level detection. Among individual evaluations in the apical periodontitis subgroup, accuracy ranged from 70.7% to 98.3%, whereas for periapical lesions it ranged from 65.99% to 99.35% (Figure 4).

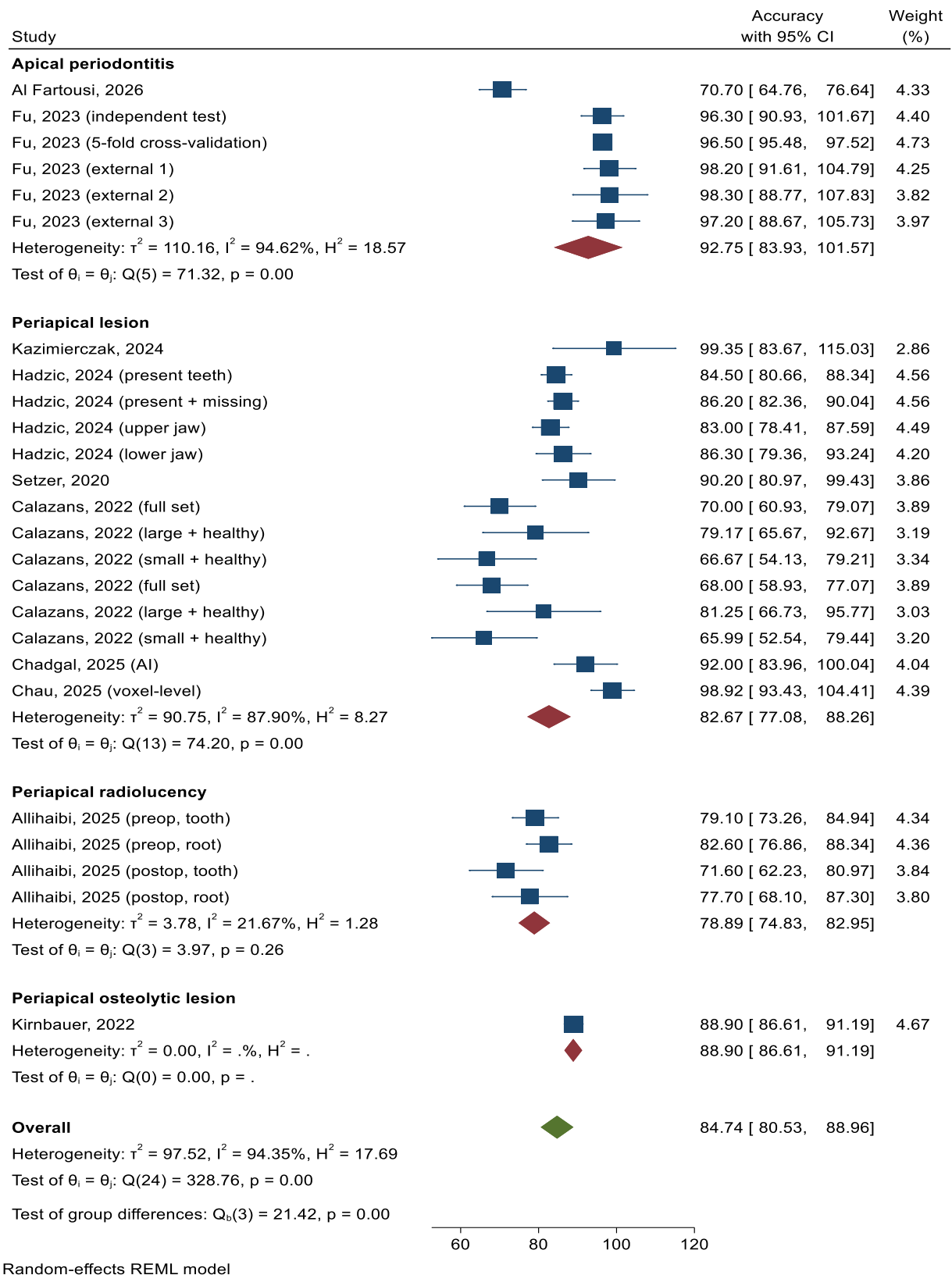


Figure 4. Forest plot of pooled accuracy of artificial intelligence-assisted CBCT interpretation for the detection of endodontic pathologies and findings.

Discussion

The present systematic review and meta-analysis evaluated the diagnostic performance of artificial intelligence-assisted interpretation of dental CBCT for endodontic pathologies and findings. Across 12 included studies, AI achieved a pooled sensitivity of 79.82%, specificity of 86.80%, and accuracy of 84.74%. Diagnostic performance varied significantly by target finding. Sensitivity was highest for periapical osteolytic lesions (97.10%) and apical periodontitis (88.27%), whereas periapical lesions showed lower sensitivity (74.38%). Specificity was highest for apical periodontitis (96.27%) and lowest for periapical radiolucency (71.46%). These differences suggest that AI performance depends on the diagnostic target rather than remaining consistent across endodontic CBCT applications. The pooled estimates were lower than those reported in the meta-analysis by Sadr et al., which found a sensitivity of 92.5% and specificity of 85.2% for deep learning-based detection of periapical radiolucent lesions (30). Pul and Schwendicke similarly reported pooled sensitivity and specificity of 94% and 96%, respectively, for AI-based detection of periapical radiolucencies (14). The lower sensitivity and specificity observed in the present analysis may partly reflect its broader scope. Unlike these reviews, which focused predominantly on periapical radiolucency, the present study included several endodontic targets and restricted the analysis to CBCT. Differences in diagnostic definitions, anatomical units, disease prevalence, and AI tasks may therefore have contributed to the lower pooled estimates.

The findings were more comparable with the CBCT-specific meta-analysis by Dashti et al. for detection of second mesiobuccal canals. That study reported pooled sensitivity of 82% and specificity of 74%, while deep learning models achieved sensitivity of 87%, specificity of 90%, and accuracy of 84% (31). These estimates are close to the overall performance observed in the present analysis, particularly for sensitivity and accuracy. However, Dashti et al. included only four studies and evaluated a specific anatomical finding, whereas the present review incorporated multiple endodontic pathologies and findings. This difference further supports the importance of considering the diagnostic target when comparing AI performance across studies. Subgroup analysis showed substantial differences

among endodontic findings. Apical periodontitis had relatively high sensitivity and specificity, whereas periapical radiolucency had lower specificity (71.46%). Periapical radiolucency sensitivity also varied by diagnostic unit. Tooth-level assessment showed higher sensitivity than root-level assessment in both preoperative and postoperative evaluations. This pattern suggests that the anatomical scale at which AI makes a prediction may affect diagnostic performance. Similarly, previous research on AI-based endodontic diagnosis has reported variation by imaging task, anatomical target, and evaluation strategy (14, 31). Heterogeneity was substantial in the overall analyses, with I^2 values of 99.15% for sensitivity, 94.03% for specificity, and 94.35% for accuracy. We also observed significant differences between diagnostic subgroups for sensitivity, specificity, and accuracy (all $P < 0.001$). Such heterogeneity is consistent with previous meta-analyses of AI in dental imaging. Dashti et al. reported considerable heterogeneity among CBCT-based studies, and broader dental AI meta-analyses have also shown substantial between-study variation (32). Differences in CBCT voxel size, acquisition protocols, sample composition, lesion prevalence, AI architecture, annotation procedures, and validation strategy are likely contributors to this variability. Methodological characteristics also require consideration. Six studies had unclear risk of bias in patient selection, and two had unclear risk for the index test, although all studies had low risk of bias for the reference standard and flow and timing. Previous reviews of AI-based endodontic diagnosis have reported similar methodological concerns (14, 30). In addition, most included studies were retrospective, and several relied on internal validation or cross-validation. Although these approaches are appropriate for model development, external validation using independent datasets from different institutions and CBCT systems is needed to establish whether reported performance is reproducible in routine clinical settings. Studies incorporating multicentre external validation, such as Fu et al., provide stronger evidence of generalizability than single-centre evaluations (25).

Overall, the findings indicate that AI-assisted CBCT interpretation can provide useful diagnostic performance for several endodontic findings. Still, the pooled estimates should not be interpreted as a uniform

performance measure for all clinical applications. The significant variation between diagnostic targets and the high heterogeneity across studies indicate that future research should use clearly defined diagnostic outcomes, standardized reference standards, and independent multicentre validation. At this stage of evidence, AI-based CBCT interpretation is better considered an adjunct to clinical and radiographic assessment rather than a replacement for expert diagnosis. This study has several limitations. Only 12 studies were included, and substantial heterogeneity was observed across the pooled analyses. Most studies were retrospective, with several relying on internal rather than external validation. Differences in AI models, diagnostic targets, CBCT parameters, reference standards, and diagnostic units may have contributed to the observed variability. In addition, unclear risk of bias in patient selection in several studies may limit the generalizability of the findings.

Conclusion

AI-assisted interpretation of dental CBCT demonstrated promising diagnostic performance for endodontic pathologies, with pooled sensitivity, specificity, and accuracy of 79.82%, 86.80%, and 84.74%, respectively. Performance varied across diagnostic targets, with higher estimates for apical periodontitis and periapical osteolytic lesions. However, substantial heterogeneity and methodological limitations across studies restrict the generalizability of the pooled estimates. Further multicentre studies using standardized diagnostic criteria and independent external validation are needed before AI-assisted CBCT interpretation can be routinely integrated into endodontic diagnosis.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

Not applicable.

Authors' Contributions

All authors equally contributed to this study.

Acknowledgments

We would like to express our appreciation and gratitude to all those who cooperated in carrying out this study.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

References

1. Sarsam W, Davies J, Al-Salehi SK. The role of imaging in endodontics. *British Dental Journal*. 2025;238(7):448-57. doi: 10.1038/s41415-025-8511-z.
2. Oliveira ML. Digital dental radiology and diagnostics: From 2D to 3D. *Australian Dental Journal*. 2025;70:S50-S66. doi: 10.1111/adj.70024.
3. Gliga A, Imre M, Grandini S, Marruganti C, Gaeta C, Bodnar D. The limitations of periapical X-ray assessment in endodontic diagnosis: a systematic review. *Journal of Clinical Medicine*. 2023;12(14):4647. doi: 10.3390/jcm12144647.
4. Pérez Alfayate R, Suárez García A, Algar Pinilla J. CBCT management of previously treated mandibular incisor with extensive internal root resorption: a case report. *International Journal of Scientific Research in Dental and Medical Sciences*. 2020;2(4):141-4.
5. Khanna AB. Applications of cone beam computed tomography in endodontics. *Evidence-Based Endodontics*. 2020;5(1):1. doi: 10.1186/s41121-020-00020-4.
6. Damor P, Wadhera RK, Parvez A, Gill S, Kawathe P. Three-dimensional imaging in vital pulp therapy: diagnostic yield, clinical decision-making, and risk of overdiagnosis. *Medical Research: Emerging Concepts and Applications Vol 22026*. p. 1-21.
7. Tsarenko M, Kalashnikova L. Optimization of scanning parameters for CT and CBCT: a systematic review. 2025.
8. Katsumata A. Deep learning and artificial intelligence in dental diagnostic imaging. *Japanese Dental Science Review*. 2023;59:329-33. doi: 10.1016/j.jdsr.2023.09.004.
9. Ding H, Wu J, Zhao W, Matinlinna JP, Burrow MF, Tsoi JK. Artificial intelligence in dentistry: a review. *Frontiers in Dental Medicine*. 2023;4:1085251. doi: 10.3389/fdmed.2023.1085251.
10. Patil S, Licari FW, Bennion C, Deshpande A, Nasseh AA, Bhandi S. Effectiveness of artificial intelligence models for

- detection of vertical root fractures: a systematic review. *Journal of Endodontics*. 2026. doi: 10.1016/j.joen.2026.02.012.
11. Setzer FC, Shi KJ, Zheng Z, Yan H, Yoon H, Mupparapu M. Artificial intelligence for the computer-aided detection of periapical lesions in cone-beam computed tomographic images. *Journal of Endodontics*. 2020;46(7):987-93. doi: 10.1016/j.joen.2020.03.025.
 12. Doumani M, Almaqboul F, Alduwaysan SSS, Alzahrani MA, Al Ghamdi SA, Alzahrani MN. Effectiveness of artificial intelligence in endodontic diagnosis and treatment evaluation: a systematic review. *Cureus*. 2025;17(11). doi: 10.7759/cureus.96091.
 13. Septina F, Kiswanjaya B. Evaluating deep learning AI for periapical lesion detection across panoramic, periapical, and CBCT radiographs: a systematic review. *The Open Dentistry Journal*. 2026;20(1). doi: 10.2174/0118742106391708260406104035.
 14. Pul U, Schwendicke F. Artificial intelligence for detecting periapical radiolucencies: a systematic review and meta-analysis. *Journal of Dentistry*. 2024;147:105104. doi: 10.1016/j.jdent.2024.105104.
 15. Goyal S. Effectiveness of artificial intelligence in detection of apical periodontitis: a systematic review. *International Journal of Special Education*. 2026;41(20s):20-36.
 16. Alfadley A, Shujaat S, Jamleh A, Riaz M, Aboalela AA, Ma H. Progress of artificial intelligence-driven solutions for automated segmentation of dental pulp space on cone-beam computed tomography images: a systematic review. *Journal of Endodontics*. 2024;50(9):1221-32. doi: 10.1016/j.joen.2024.05.012.
 17. Wang Y, Xia W, Yan Z, Zhao L, Bian X, Liu C. Root canal treatment planning by automatic tooth and root canal segmentation in dental CBCT with deep multi-task feature learning. *Medical Image Analysis*. 2023;85:102750. doi: 10.1016/j.media.2023.102750.
 18. Salameh JP, Moher D, McGrath TA, Frank RA, Sharifabadi AD, Islam N. Assessing adherence to the PRISMA-DTA guideline in diagnostic test accuracy systematic reviews: a five-year follow-up analysis. *The Journal of Applied Laboratory Medicine*. 2025;10(2):416-31. doi: 10.1093/jalm/jfae117.
 19. Al Fartousi M, Buscot A, Gernhardt CR. Diagnostic performance of AI-based cloud software regarding the detection of endodontic findings on CBCT: a single-centre cross-sectional validation study. *Journal of Clinical Medicine*. 2026;15(12):4839. doi: 10.3390/jcm15124839.
 20. Öksüzöğlu H, Topuz Ö, Bilgir E, Çelik Ö, Bayrakdar I. Fully automated 3D segmentation of periapical lesions on CBCT images using nnU-Net v2. *BMC Medical Imaging*. 2026. doi: 10.1186/s12880-026-02316-0.
 21. Chadgal S. Artificial intelligence-assisted interpretation of CBCT scans for early periapical pathology detection: a comparative endodontic study. *Multidisciplinary Innovations & Research Analysis*. 2025;6(4):111-6.
 22. Chau KK, Zhu M, AlHadidi A, Wang C, Hung K, Wohlgemuth P. A novel AI model for detecting periapical lesion on CBCT: CBCT-SAM. *Journal of Dentistry*. 2025;153:105526. doi: 10.1016/j.jdent.2024.105526.
 23. Hadzic A, Urschler M, Press JNA, Riedl R, Rugani P, Štern D. Evaluating a periapical lesion detection CNN on a clinically representative CBCT dataset: a validation study. *Journal of Clinical Medicine*. 2023;13(1):197. doi: 10.3390/jcm13010197.
 24. Kazimierczak W, Wajer R, Wajer A, Kiian V, Kloska A, Kazimierczak N. Periapical lesions in panoramic radiography and CBCT imaging: assessment of AI's diagnostic accuracy. *Journal of Clinical Medicine*. 2024;13(9):2709. doi: 10.3390/jcm13092709.
 25. Fu W, Zhu Q, Li N, Wang Y, Deng S, Chen H. Clinically oriented CBCT periapical lesion evaluation via 3D CNN algorithm. *Journal of Dental Research*. 2024;103(1):5-12. doi: 10.1177/00220345231201793.
 26. Calazans MAA, Ferreira FAB, Alcoforado M, Santos A, Pontual A, Madeiro F. Automatic classification system for periapical lesions in cone-beam computed tomography. *Sensors*. 2022;22(17):6481. doi: 10.3390/s22176481.
 27. Kimbauer 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. *Journal of Endodontics*. 2022;48(11):1434-40. doi: 10.1016/j.joen.2022.07.013.
 28. Orhan K, Bayrakdar I, Ezhov M, Kravtsov A, Özyürek T. Evaluation of artificial intelligence for detecting periapical pathosis on cone-beam computed tomography scans. *International Endodontic Journal*. 2020;53(5):680-9. doi: 10.1111/iej.13265.
 29. Allihabi M, Koller G, Mannocci F. Diagnostic accuracy of an artificial intelligence-based platform in detecting periapical radiolucencies on cone-beam computed tomography scans of molars. *Journal of Dentistry*. 2025;160:105854. doi: 10.1016/j.jdent.2025.105854.
 30. Sadr S, Mohammad-Rahimi H, Motamedian SR, Zahedrozegar S, Motie P, Vinayahalingam S. Deep learning for detection of periapical radiolucent lesions: a systematic review and meta-analysis of diagnostic test accuracy. *Journal of Endodontics*. 2023;49(3):248-61. doi: 10.1016/j.joen.2022.12.007.
 31. Alabdulkareem M, Atieh M, AbuMostafa A, Aldalaan K, Alturki N. Artificial intelligence in dental treatment planning and diagnostic decision-making: a systematic review and meta-analysis. *Clinical and Experimental Dental Research*. 2026;12(2):e70343. doi: 10.1002/cre2.70343.
 32. Dashti M, Khosraviani F, Ghadimi N, Baghaei K, Esmaeili S, Entezar-e-Ghaem M. Use of artificial intelligence for detection of MB2 canals in maxillary first molars on CBCT: a systematic review and meta-analysis. *BMC Oral Health*. 2025;25(1):1860. doi: 10.1186/s12903-025-07254-x.