

Artificial Intelligence in Endodontic Retreatment: Detection of “Missed” Root Canal Anatomy

Salah A Yousief^{1,2*}, Khaled N Alharbi¹, ME Al-Zubaidi¹, Hussam Alharbi¹, Reem Faihan Alotaibi¹ and Abdulrhman Youssef Alghamdi¹

¹Department of Restorative and Prosthetic Dental Sciences, College of Dentistry, Dar Al Uloom University, Riyadh, Saudi Arabia

²General Dentist, Saudi Arabia

***Corresponding Author:** Salah A Yousief, Department of Restorative and Prosthetic Dental Sciences, College of Dentistry, Dar Al Uloom University, Riyadh and General Dentist, Saudi Arabia.

Received: October 03, 2026; **Published:** October 30, 2026

Abstract

Background: Untreated canals are among the most consistent findings in teeth that fail after root canal treatment, and the second mesiobuccal (MB2) canal of the maxillary first molar is the most frequently implicated of them. Whether a missed canal is found before retreatment depends almost entirely on how an image is read, and reading is where human performance is weakest: intra- and inter-observer agreement for canal anatomy is modest, and in previously treated teeth the beam-hardening and streak artefacts generated by gutta-percha, posts and coronal restorations degrade precisely the region that must be inspected. Artificial intelligence (AI), and convolutional neural networks in particular, has been proposed as a second reader for this task.

Objectives: To evaluate systematically the diagnostic accuracy of AI models for the detection, localisation and segmentation of missed or unobturated root canal anatomy on cone-beam computed tomography (CBCT) and intraoral radiographs; to appraise the methodological quality of the primary studies; and to determine how much of the available evidence actually addresses the retreatment scenario rather than the untreated tooth.

Methods: A structured search of PubMed/MEDLINE, Scopus, Embase, Web of Science and the Cochrane Library was performed for studies published between January 2015 and September 2026, following PRISMA 2020. Eligible studies applied AI to the identification, localisation or segmentation of root canal anatomy — with emphasis on MB2 and other missed canals — on CBCT, periapical or panoramic images, and reported at least one quantitative performance metric against a defined reference standard. Study characteristics, model architecture, dataset size, validation strategy, reference standard and diagnostic metrics were extracted. Methodological quality was appraised with QUADAS-2 and reporting completeness against the CLAIM checklist.

Results: Reported performance for MB2 detection is high but heterogeneous. In the only study designed explicitly around the retreatment scenario — unobturated MB2 canals in already-obturated maxillary molars — a U-Net achieved sensitivity 0.80, specificity 1.00, positive predictive value 1.00 and accuracy 0.90, yet segmentation of the same canals was poor. Pooled estimates from a recent meta-analysis of four CBCT studies gave sensitivity 0.82 and specificity 0.74 (0.96 after correction for publication bias), with deep learning outperforming classical machine learning on every metric and heterogeneity exceeding 90%. Larger and more recent frameworks report AUC around 0.85 and, importantly, measurable improvement in the performance of junior clinicians when used as decision support. On two-dimensional periapical images the best models matched but did not surpass expert clinicians. Across the corpus, datasets are small, external validation is rare, reference standards are inconsistently defined, and only a minority of studies include previously treated teeth at all.

Conclusion: AI performs well enough on this task to be useful as a second opinion, and there is now direct evidence that it raises the performance of less experienced operators. What it does not yet have is the evidence base required for autonomous use: the field is dominated by single-centre, internally validated, retrospective models tested on curated images of untreated teeth. The clinically relevant question — can an algorithm find an unobturated canal in a metal-laden CBCT of a failing root filling and does finding it change the outcome — remains largely unanswered. Priorities are multicenter datasets that deliberately include retreatment cases, mandatory external validation, standardized reference standards and reader studies with outcome endpoints.

Keywords: *Artificial Intelligence; Deep Learning; Endodontic Retreatment; Missed Canals; Second Mesio Buccal Canal; Cone-Beam Computed Tomography; Diagnostic Accuracy; Systematic Review*

Introduction

Endodontic failure has a short list of usual suspects, and untreated anatomy sits near the top of it [1]. The observation is old, but the evidence supporting it became far more precise once cone-beam computed tomography entered routine endodontic practice. In a retrospective cohort of 1397 CBCT volumes from a North American population, teeth with an untreated canal carried a substantially higher burden of apical periodontitis than adequately treated teeth [2]. A cross-sectional study of endodontically treated teeth reported that those with a missed canal were roughly 4.4 times more likely to show periapical pathology [3], and comparable associations have been reported repeatedly from different populations [4,5]. Reported prevalences vary — missed canals in something between 12% and 19% of treated posterior teeth in most series, rising to more than half of maxillary molars in others — but the direction of effect is remarkably stable: where a canal is left untreated, apical periodontitis is usually present.

The anatomy responsible is well characterised. Weine and colleagues described the canal configuration of the mesiobuccal root of the maxillary first molar in 1969 [6], Vertucci systematised root canal morphology and its procedural consequences [7], and a multicentre CBCT analysis has since documented how much MB2 prevalence varies between populations [8]. The problem has never been that the anatomy is unknown. The problem is that it is difficult to see.

Two-dimensional periapical radiography superimposes the mesiobuccal root over the palatal root and hides the very structure of interest [9]. CBCT resolves this in principle, and a meta-analysis of its diagnostic accuracy for detecting a second canal in permanent teeth reports figures in the range of 69.2% to 96.6% [10]. Yet even with CBCT the result depends on the reader. Parker and co-workers showed both that CBCT improves a clinician’s ability to locate MB2 canals [11] and, in a companion study, that observers disagree with one another to a clinically meaningful degree when interpreting the same volumes [12]. Interpretation, not acquisition, is the bottleneck.

That bottleneck is worse in the situation this review is concerned with. In a tooth being assessed for retreatment, the mesiobuccal root is flanked by radiopaque gutta-percha, sometimes a metal post, and usually a cast or ceramic restoration. Beam hardening and streak artefacts arise from exactly these materials and project across exactly the region where an unobturated MB2 would appear. The reader must therefore distinguish a narrow, low-contrast, obliquely angled canal from an artefact, in a volume acquired at the lowest dose consistent with the ALARA principle. It is a hard visual task performed under poor conditions, and it is the one at which artificial intelligence has been proposed to help.

AI in endodontics is no longer a novelty. Reviews have mapped its application to periapical lesion detection, working length determination, vertical root fracture identification, morphological classification and outcome prediction [13,14], and the wider dental literature has begun to address the governance questions that follow [15]. What has not been done, to our knowledge, is a synthesis focused specifically on the retreatment problem: not “can AI find canals” in general, but can it find the canals that were missed, in the teeth where they were missed, under the imaging conditions that actually obtain when a clinician is deciding whether to retreat.

This review therefore asks three questions. How accurate are AI models at detecting missed or unobturated root canal anatomy? How trustworthy are the studies reporting that accuracy? And how much of the evidence is actually about retreatment?

Materials and Methods

Protocol and reporting

The review was conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement [16]. Methodological quality of the included diagnostic accuracy studies was appraised using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [17], and completeness of reporting was assessed against the Checklist for Artificial Intelligence in Medical Imaging (CLAIM) [18], with reference to PROBAST-AI [19] and STARD 2015 [20] for domain-specific items concerning model development and index test description.

Review question and eligibility

The question was framed using the PICO structure shown in table 1.

Element	Specification
Population (P)	Human permanent teeth imaged by CBCT, periapical or panoramic radiography, including previously root-treated teeth assessed for retreatment
Index test (I)	Any AI model (deep learning, convolutional neural network, transformer, or classical machine learning) applied to detection, localisation, classification or segmentation of root canal anatomy, with emphasis on missed or unobturated canals
Comparator (C)	Expert human interpretation, micro-computed tomography, direct clinical verification, or another AI model
Outcome (O)	Sensitivity, specificity, accuracy, PPV, NPV, AUC, Dice/IoU, F1-score; secondary outcomes: interpretation time, effect on clinician performance
Study designs	Diagnostic accuracy studies, retrospective and prospective cohorts, <i>in vitro</i> studies on extracted teeth, reader studies, randomised trials of AI assistance
Exclusions	Editorials and opinion pieces; studies without a quantitative performance metric; models addressing tasks unrelated to canal anatomy or periapical status; conference abstracts without full text; patents
Language/period	English; January 2015 to September 2026

Table 1: Review question and eligibility criteria.

Information sources and search strategy

Five electronic databases were searched (PubMed/MEDLINE, Scopus, Embase, Web of Science, Cochrane Library), supplemented by Google Scholar and by hand-searching the reference lists of all retrieved reviews. The search strings are given in table 2.

The four blocks were combined with AND. Filters were applied for language and publication date only; no restriction was placed on study design at the search stage.

Block	Terms combined with OR
Technology	“artificial intelligence”; “deep learning”; “machine learning”; “convolutional neural network”; “CNN”; “U-Net”; “YOLO”; “transformer”
Anatomy/target	“root canal”; “second mesiobuccal”; “MB2”; “missed canal”; “untreated canal”; “unobtained canal”; “canal morphology”; “periapical lesion”
Setting	“endodontic”; “retreatment”; “root canal treatment”; “apical periodontitis”
Imaging	“cone-beam computed tomography”; “CBCT”; “periapical radiograph”; “panoramic radiograph”

Table 2: Search strategy.

Study selection and data extraction

Records were screened in two stages — title and abstract, then full text — by two independent reviewers, with disagreements resolved by discussion or by a third reviewer. The selection process is summarised in the PRISMA flow diagram (Figure 1).

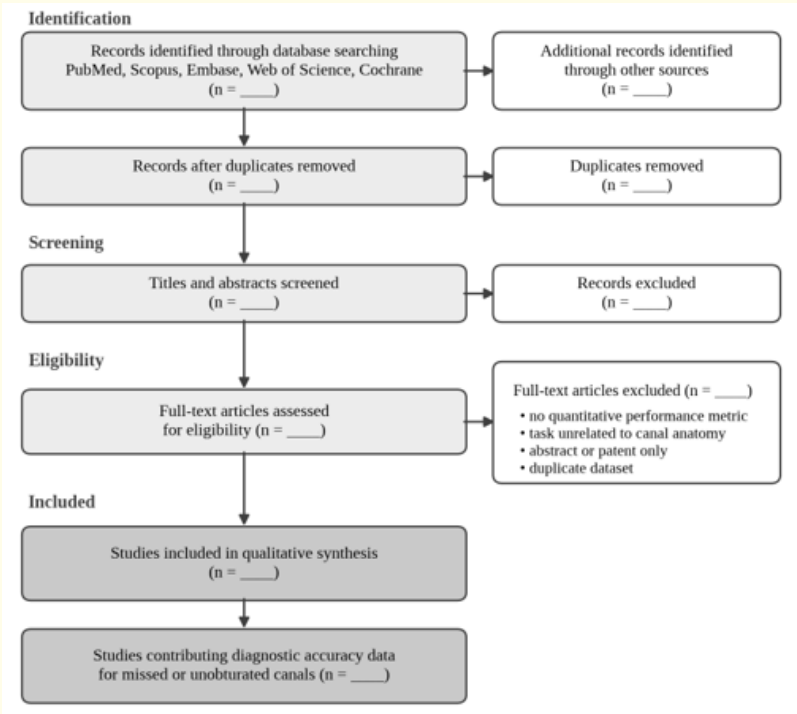


Figure 1: PRISMA 2020 flow diagram for the selection of studies. Counts are to be completed from the screening log of the review team; the structure, stage labels and exclusion reasons correspond to the eligibility criteria set out in table 1.

For each included study the following were extracted: authors, year, country, study design, imaging modality, clinical task, AI architecture, dataset size and partitioning, validation strategy (internal cross-validation versus external test set), reference standard, and all reported performance metrics. Where a study reported several models, the best-performing model was extracted, and the range across models was recorded.

Synthesis

A quantitative meta-analysis was not performed. Reference standards ranged from single-expert labelling to consensus reading to micro-CT; unit of analysis varied between the patient, the tooth, the root, the region of interest and the individual slice; and validation strategies were not comparable. Under these conditions a pooled estimate would be statistically obtainable and clinically meaningless. Findings are therefore synthesised narratively and tabulated with their reference standard and unit of analysis stated alongside each figure.

Results

The clinical problem in numbers

Before evaluating a diagnostic aid, it is worth quantifying the diagnostic gap it is meant to close (Table 3 and figure 2). Two features of the literature stand out. First, MB2 prevalence measured anatomically or on CBCT is high — commonly reported between roughly one third and more than nine tenths of maxillary first molars depending on population and method [8,21] — while the proportion actually located and treated in practice is considerably lower. Adjuncts such as the dental operating microscope and selective dentine removal narrow that gap but do not close it [22]. Second, when a canal is missed, apical periodontitis is the rule rather than the exception: reported prevalences of apical periodontitis among teeth with missed canals cluster between 85% and 96% across independent populations [2-5].

Finding	Reported value	Population / method	Source
MB2 prevalence, maxillary first molar	Approximately 33-96%	Multiple populations, CBCT and anatomical	[8,21]
CBCT accuracy for detecting a second canal	69.2-96.6%	Meta-analysis	[10]
Missed canals among treated posterior teeth	Approximately 12-19%	CBCT cross-sectional series	[4,5]
Missed canals, maxillary molars	Up to approximately 55%	CBCT cross-sectional	[5]
Apical periodontitis where a canal is missed	Approximately 85-96%	CBCT cross-sectional	[2,4,5]
Relative likelihood of periapical pathology with a missed canal	Approximately 4.4-fold	Cross-sectional	[3]
Inter-observer agreement, CBCT interpretation	Clinically meaningful disagreement reported	Reader study	[12]

Table 3: The diagnostic gap: missed canals and their consequences.

This is the case for a second reader. The task has a high prior probability of the target being present, a substantial rate of failure to detect it, a serious consequence when detection fails, and demonstrable observer variability.

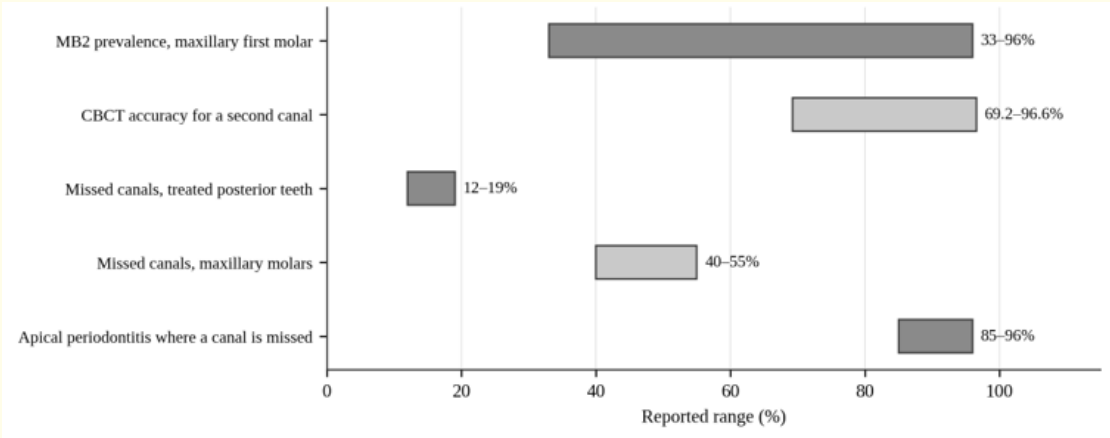


Figure 2: The diagnostic gap in numbers. Bars show the range of values reported across the studies cited in table 3. The mismatch between how often an MB2 canal exists, how often it is left untreated, and how consistently apical periodontitis follows, defines the problem that a diagnostic aid is being asked to solve.

Characteristics of the included evidence

The studies retrieved fall into four groups: models for MB2 detection or segmentation on CBCT; models for MB2 detection on two-dimensional periapical images; models addressing adjacent endodontic tasks that bear on retreatment planning (periapical radiolucency detection, root filling quality assessment, canal orifice identification); and evidence syntheses. Their characteristics and reported performance are summarised in table 4; the metrics used throughout are defined in figure 3.

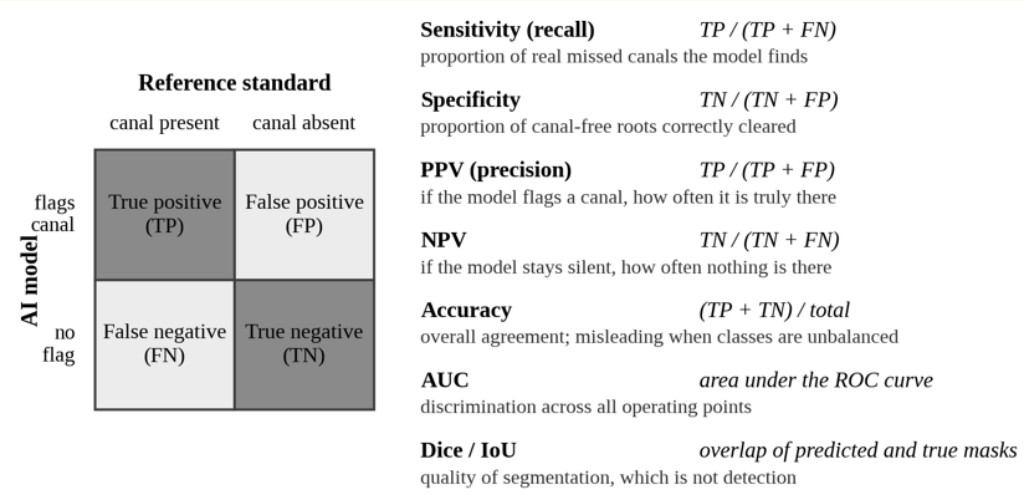


Figure 3: Diagnostic performance metrics and what each one means for this clinical task. Sensitivity and specificity depend on the operating point chosen, which most published studies do not report; Dice and IoU describe segmentation quality and should never be quoted as evidence of detection performance.

Study	Modality	Task	Architecture	Dataset	Reference standard	Key performance
Albitar, <i>et al.</i> 2022 [23]	CBCT	Detection and segmentation of unobturated MB2 in obturated molars	U-Net	57 CBCT studies; 102 roots	Expert consensus	Sens 0.80; Spec 1.00; PPV 1.00; NPV 0.83; Acc 0.90; segmentation metric 0.30
Duman, <i>et al.</i> 2024 [24]	CBCT	MB2 segmentation	YOLOv5	CBCT image set	Expert labelling	High detection performance for canal localisation
Shetty, <i>et al.</i> 2025 [25]	CBCT	MB2 orifice detection	Classical ML (SVM, neural network classifiers)	CBCT image set	Examiner labelling	High recall; performance below deep learning comparators
Ji, <i>et al.</i> 2024 [26]	CBCT (<i>in vitro</i>)	Resolution optimisation to improve canal visibility	AI super-resolution	Extracted teeth	Micro-CT	Improved visualisation of fine canal anatomy
Güllü, <i>et al.</i> 2026 [27]	CBCT	3D automatic segmentation of root canals with focus on MB2	nnU-Netv2	202 CBCT volumes	Expert annotation	Two-stage training with label dilation required for MB2
Two-stage DL framework 2026 [28]	CBCT	MB2 detection and clinician support	VGG16 localisation + ConvNeXt-Transformer classification	1022 patients; 1944 ROIs; independent test 100 ROIs	Expert consensus, volumetric	AUC 0.85; improved junior dentist performance
Two-step AI approach 2025 [29]	CBCT	MB2 detection and segmentation	Two-step deep learning	Institutional CBCT set	Experienced radiologists	Reported as accuracy, recall, precision, ROC
Ourang, <i>et al.</i> 2025 [30]	Periapical	Detection of clinically negotiable MB2	6 supervised CNNs + 3 self-supervised models	1504 pre-operative images	Clinical confirmation by specialists	Best models comparable to, but not better than, human experts
Karatas, <i>et al.</i> 2025 [31]	Microscope images	Canal orifice identification	CNN	Operating microscope images	Expert labelling	Accuracy approx. 91%
Jin, <i>et al.</i> 2025 [32]	Periapical	Root filling radiographic quality	Deep learning network	Radiographic dataset	Expert grading	Automated quality classification
Allihaibi, <i>et al.</i> 2025 [33]	CBCT	Periapical radiolucency detection in molars	Commercial AI platform	CBCT scans of molars	Expert reading	Diagnostic accuracy reported for platform

Sadr., <i>et al.</i> 2023 [34]	Mixed	Periapical radiolucent lesion detection	Systematic review and meta-analysis	—	—	Promising accuracy; bias and inconsistent reporting noted
Pul & Schwendicke 2024 [35]	Mixed	Periapical radiolucency detection	Systematic review (24 studies)	—	—	High overall sensitivity and specificity; marked heterogeneity
Pul., <i>et al.</i> 2025 [36]	Panoramic	Effect of AI assistance on dentists	Randomised controlled trial	Dentist readers	—	Effect of assistance on accuracy, confidence and treatment decisions
Dashti., <i>et al.</i> 2025 [37]	CBCT	AI for MB2 detection	Systematic review and meta-analysis (4 studies)	—	—	Pooled Sens 0.82; Spec 0.74 (0.96 corrected); Acc 0.77; I ² > 90%

Table 4: Characteristics and reported performance of included studies.

Diagnostic accuracy for missed and unobturated canals

The single most relevant study to the retreatment question remains that of Albitar and colleagues, because it is the only one that deliberately selected teeth with clinically confirmed unobturated MB2 canals in molars that had already been root filled [23]. A U-Net trained on 102 segmented roots achieved a sensitivity of 0.80, a specificity of 1.00, a positive predictive value of 1.00, a negative predictive value of 0.83 and an accuracy of 0.90 on the test set. Read clinically, that profile is favourable in one specific way: when the model says a canal is there, it is there. A positive flag is worth acting on. A negative flag is worth considerably less — one in five canals was missed by the model as well.

The same paper contains a finding that is easy to overlook and important to state. While detection performed well, segmentation of the unobturated canals was poor, with a mean custom segmentation metric of approximately 0.30 [23]. Detecting that a canal exists and delineating its course are different computational problems of very different difficulty, and the literature routinely conflates them. For retreatment planning, the second is what the clinician actually wants.

The meta-analysis by Dashti and colleagues pooled four CBCT studies and reported an overall sensitivity of 0.82 (95% CI 0.75-0.89) and specificity of 0.74 (95% CI 0.58-0.90), the latter rising to 0.96 after trim-and-fill correction for a significant publication bias [37]. Overall accuracy was 0.77 (95% CI 0.71-0.84). Deep learning outperformed classical machine learning on every metric — sensitivity 0.87 versus 0.80, specificity 0.90 versus 0.68, accuracy 0.84 versus 0.75, PPV 0.90 versus 0.77 and NPV 0.86 versus 0.73. Heterogeneity exceeded 90%, dataset sizes ranged from 10 scans to 277, and sensitivity analyses were not feasible with four studies. These pooled values are shown against the individual study results in figure 4.

More recent architectures report better and more balanced performance on larger datasets. A two-stage framework combining VGG16 region-of-interest localisation with ConvNeXt-Transformer classification, trained on scans from 1022 patients and evaluated on an independent 50/50 test set, achieved an AUC of 0.85 and, in the same study, improved the diagnostic performance of junior dentists when used as decision support [28]. Grad-CAM heatmaps localised model attention to the mesiobuccal root and the developmental groove between MB1 and the palatal canal — anatomically coherent behaviour, which is not a trivial observation given how often image models learn spurious features.

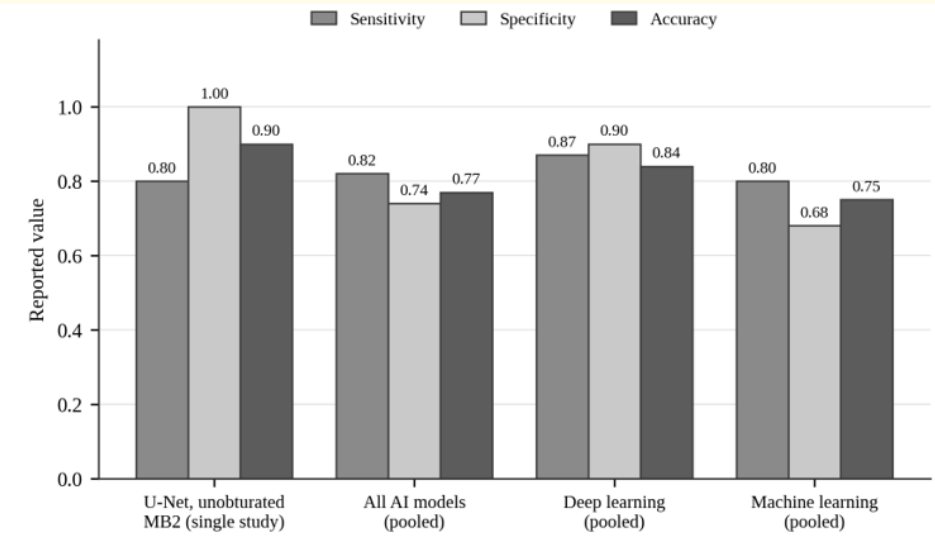


Figure 4: Reported diagnostic performance of AI models for MB2 canal detection on CBCT. The first group is the single study performed on unobturated canals in already-obturated molars; the remaining three are pooled estimates from a meta-analysis of four studies. Note the divergence in specificity, and that the pooled specificity rose from 0.74 to 0.96 after correction for publication bias. Data from references [23,37].

Segmentation has advanced in parallel. An nnU-Netv2 pipeline applied to 202 CBCT volumes achieved three-dimensional segmentation of the root canal system with concurrent MB2 delineation, but only after preprocessing the labels by binary dilation and training in two stages with different dilation levels [27]. The technical detail carries a clinical message: the MB2 canal is so narrow relative to CBCT voxel size that the ground truth itself has to be artificially thickened for the network to learn it. Any performance figure derived from dilated labels should be read with that in mind.

On two-dimensional imaging the ceiling is lower, as expected. Ourang and colleagues assembled 1504 pre-operative periapical radiographs of maxillary molars with clinical confirmation of whether the MB2 canal had proved negotiable, and fine-tuned nine architectures, six supervised and three self-supervised [30]. The best models matched expert clinicians on accuracy but did not surpass them. This is a more useful result than a higher number would have been, because the reference standard was clinical negotiability rather than radiographic appearance — that is, the outcome the clinician cares about.

Adjacent tasks that bear on retreatment decisions

Missed anatomy is rarely assessed in isolation. The retreatment decision also rests on whether a periapical lesion is present and on the technical quality of the existing root filling, and AI models exist for both. Systematic reviews of periapical radiolucency detection report high pooled sensitivity and specificity but also marked heterogeneity in modality, reference standard and metric definition, sufficient in one case to preclude valid meta-analysis [34,35]. Commercial platforms are now being evaluated prospectively against expert readers on CBCT scans of molars [33], and deep learning networks have been developed to grade root filling quality automatically on periapical images [32]. Intraoperatively, CNN analysis of dental operating microscope images has identified canal orifices with an accuracy of about 91% [31] — the only point in the workflow where an algorithm can assist after the access cavity has been cut.

The most clinically informative study in this group is not a model paper at all. Pul and colleagues conducted a randomised controlled trial of AI assistance for periapical radiolucency detection, measuring not only accuracy but also diagnostic confidence and the treatment decisions that followed [36]. This design — does the tool change what the clinician does, and for the better — is the one the MB2 literature almost entirely lacks.

Methodological quality and reporting

Applying QUADAS-2 across the retrieved studies, the recurrent problems are concentrated in two domains.

Patient selection: Several studies are *in vitro*, using extracted teeth [26]; these are informative about physics and about upper bounds on performance, but not about clinical applicability, and they were flagged as high risk of bias and high applicability concern in the published meta-analysis [37]. More importantly for this review, most *in vivo* studies drew their samples from untreated teeth. A model trained and tested on artefact-free volumes has not been evaluated for the task described in this paper’s title.

Reference standard: Ground truth ranges from a single examiner’s labelling to expert consensus to micro-CT to clinical negotiability. These are not interchangeable. A canal that is radiographically visible is not necessarily negotiable; a canal that is negotiable is not always radiographically obvious. None of the four studies in the published meta-analysis reported blinding of the reference standard assessors [37].

Reporting completeness against CLAIM is uneven. Items most often incomplete are: description of how the dataset was partitioned and whether images from the same patient could appear in both training and test sets; whether hyperparameter selection used a separate validation set; the operating point at which sensitivity and specificity were computed; and confidence intervals for the reported metrics. The consequences are not hypothetical — data leakage of this kind inflates apparent performance by allowing a model to memorise patient-specific features rather than learn generalisable ones, a failure mode recently and forcefully described in the endodontic literature [38].

Table 5 sets out the recurring weaknesses and the reporting items that would remedy them.

Weakness	Why it matters clinically	Remedy
Small, single-centre datasets	Model may encode scanner- and population-specific features	Multicentre datasets across CBCT units and voxel sizes
Absent external validation	Internal cross-validation overestimates real-world accuracy	Held-out external test set from a different institution
Patient-level data leakage	Inflated metrics; failure on new patients	Partition by patient, not by image or slice
Inconsistent reference standards	Metrics are not comparable between studies	Pre-specified standard; clinical negotiability where feasible
No blinding of reference readers	Bias in ground truth	Independent, blinded consensus labelling
Untreated teeth only	Does not address the retreatment scenario	Deliberate inclusion of obturated teeth with artefacts
Detection reported, segmentation neglected	Clinician needs the canal’s course, not just its presence	Report Dice/IoU alongside detection metrics
Operating point unreported	Sensitivity and specificity are meaningless without it	Report threshold, ROC curve and confidence intervals
Accuracy-only endpoints	Does not show clinical benefit	Reader studies and trials with decision or outcome endpoints

Table 5: Recurring methodological weaknesses and corresponding remedies.

Discussion

What the evidence supports

Three claims are supported by the current literature. AI models detect MB2 canals on CBCT with sensitivity in the region of 0.8 and, in the better-designed studies, high positive predictive value. Deep learning consistently outperforms classical machine learning for this task, which is unsurprising given that the target is a subtle spatial pattern rather than a set of hand-crafted features [37]. And AI assistance measurably improves the performance of less experienced operators [28] — arguably the most important finding in the field, because it identifies where the tool has value. An endodontist with a microscope, a CBCT and twenty years of experience has less to gain from a canal-detection algorithm than a general practitioner deciding whether a failing molar should be referred.

The clinical logic of the reported metric profile deserves attention. In the retreatment-specific study, specificity and PPV were maximal while sensitivity was 0.80 [23]. For this particular decision, that asymmetry is defensible. A false positive costs a few minutes of careful troughing under magnification; a false negative leaves an infected canal in place. If anything, a model tuned for this use should be pushed further toward sensitivity, accepting more false alarms — but the operating point is exactly what most papers fail to report.

What the evidence does not support

It does not support autonomous use, and it does not yet support strong claims about the retreatment scenario specifically. The great majority of MB2 models were developed on teeth that had never been treated. The retreatment case differs in ways that matter to a convolutional network: the dominant image gradients come from radiopaque filling material rather than from dentine boundaries; streak artefacts produce linear low-density features that resemble the target class; and the anatomically informative developmental groove may be obliterated by previous access preparation. A model that has never seen such images has not been tested on the problem.

Nor does the evidence support the accuracy figures at face value. Heterogeneity exceeding 90% across four studies [37], datasets of ten scans, publication bias significant enough that pooled specificity moved from 0.74 to 0.96 after correction, ground truth labels diluted to make a target learnable [27], and near-universal absence of external validation together mean that reported performance should be read as an upper bound obtained under favourable conditions. The recent call for greater rigour and relevance in endodontic AI research makes this argument in detail and should be required reading for anyone entering the field [38].

There is also a subtler issue about what is being predicted. Radiographic detection of an MB2 canal and clinical negotiation of that canal are not the same endpoint, and the study that used negotiability as its reference standard found that AI matched rather than exceeded experts [30]. This is the honest comparison, and it suggests that some of the more spectacular published accuracies reflect an easier surrogate task.

Where AI fits in the retreatment workflow

Read pragmatically, the evidence supports a defined and modest role (Figure 5).

At the assessment stage, an algorithm flagging a probable unobturated canal on a CBCT of a failing molar functions as a prompt to look harder — a triage signal for the clinician and, in a referral setting, an argument for specialist assessment. In parallel, models for periapical radiolucency detection and root filling quality grading contribute to the same decision from different directions [32-35]. Intraoperatively, image analysis of operating microscope views may assist orifice identification [31]. At no point in this chain does the algorithm make the decision, and at no point is a negative output a licence to stop searching.

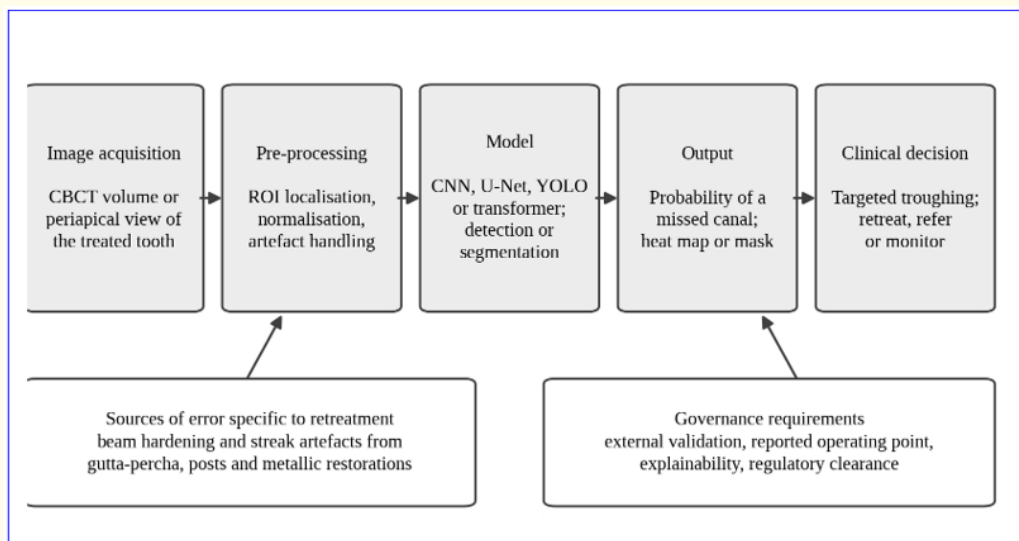


Figure 5: Where AI enters the retreatment workflow, and what constrains it at each step. The algorithm contributes a probability and a location; the clinical decision remains with the operator, and a negative output is not evidence that no canal is present.

The reservations that apply to any medical AI apply here as well: transparency of model development, guarding against algorithmic bias arising from geographically narrow training data — relevant given documented variation in MB2 prevalence between populations [8] — and regulatory clearance before clinical deployment [15,39].

A research agenda

Four priorities follow directly from the weaknesses identified in section methodological quality and reporting. First, a multicentre CBCT dataset that deliberately over-samples previously treated teeth with metallic restorations and posts, annotated by blinded consensus and, where possible, verified against clinical negotiation at retreatment. Second, mandatory external validation on data from institutions and scanners not represented in training. Third, standardised reporting following CLAIM and STARD-AI, with explicit operating points, patient-level partitioning and confidence intervals [18,20]. Fourth, and most important, a shift of endpoint: from “how accurate is the model” to “does using the model change what clinicians find and what happens to the tooth”. The randomised design already applied to periapical radiolucency detection [36] provides a workable template. The relationship between these stages and the density of published evidence at each is shown in figure 6.

Limitations of this Review

The synthesis is narrative and no meta-analysis was performed; while this decision was made on defensible grounds, it means no pooled estimate is offered here. The field is moving quickly, and studies published after the search date will not be represented. Restriction to English-language publications may have excluded relevant work, and the small number of studies addressing the specific retreatment scenario means that conclusions about that scenario rest on a narrow base — a limitation of the evidence rather than of the review, but a limitation nonetheless. Finally, publication bias in this domain is likely to be substantial: models that fail to perform are rarely written up.

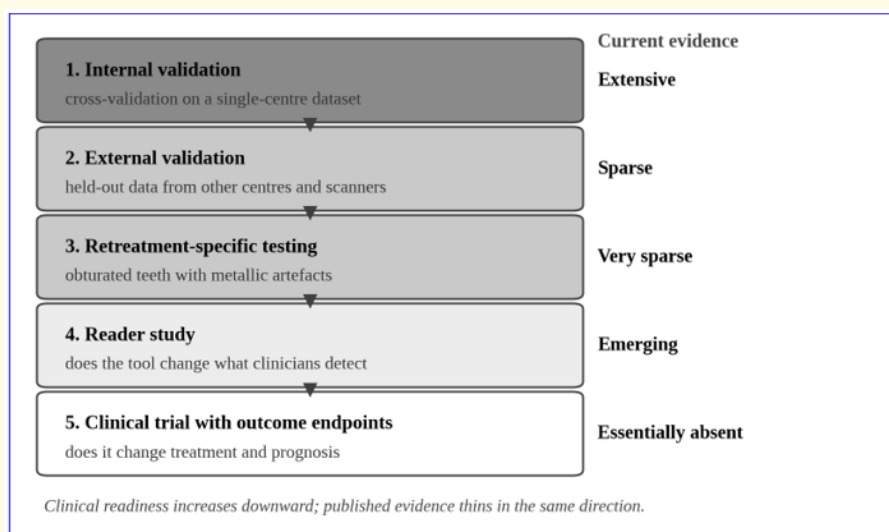


Figure 6: The translational ladder for diagnostic AI in endodontics. Published work is concentrated at the top of the ladder, where evidence is cheapest to generate, and thins sharply toward the bottom, where it would actually inform practice.

Conclusion

Missed canals cause endodontic failure, and they are missed because they are hard to see. Artificial intelligence addresses that specific weakness, and the current evidence shows it does so reasonably well: sensitivity around 0.8 for MB2 detection on CBCT, high positive predictive value in the best-designed study, consistent superiority of deep learning over classical machine learning, and a demonstrated improvement in the diagnostic performance of less experienced clinicians.

Those numbers should be read as what they are — upper bounds from small, mostly single-centre, internally validated, retrospective studies performed largely on teeth that had never been treated. The scenario that gives this review its title, an unobturated canal hiding beside a gutta-percha-filled root in an artefact-laden volume, has been directly studied only rarely.

The reasonable position for the clinician is therefore straightforward. Treat a positive AI flag as a reason to look harder with magnification and careful troughing; treat a negative flag as no evidence of absence; and do not allow either to displace clinical judgement. The reasonable position for the researcher is equally clear: the field does not need another single-centre model reporting 95% accuracy on curated images. It needs external validation, standardised reporting, retreatment-relevant data and trials that measure what happens to patients.

Bibliography

1. Tabassum S and Khan FR. “Failure of endodontic treatment: the usual suspects”. *European Journal of Dentistry* 10.1 (2016): 144-147.
2. Karabucak B, *et al.* “Prevalence of apical periodontitis in endodontically treated premolars and molars with untreated canal: a cone-beam computed tomography study”. *Journal of Endodontics* 42.4 (2016): 538-541.
3. Baruwa AO, *et al.* “The influence of missed canals on the prevalence of periapical lesions in endodontically treated teeth: a cross-sectional study”. *Journal of Endodontics* 46.1 (2020): 34-39.e1.

4. Costa FFNP, *et al.* “Association between missed canals and apical periodontitis”. *International Endodontic Journal* 52.4 (2019): 400-406.
5. Blanco Fuentes BY, *et al.* “Apical periodontitis in endodontically-treated teeth: association between missed canals and quality of endodontic treatment in a Colombian sub-population. A cross-sectional study”. *Acta Odontológica Latinoamericana* 37.1 (2024): 59-67.
6. Weine FS, *et al.* “Canal configuration in the mesiobuccal root of the maxillary first molar and its endodontic significance”. *Oral Surgery, Oral Medicine, Oral Pathology* 28.3 (1969): 419-425.
7. Vertucci FJ. “Root canal morphology and its relationship to endodontic procedures”. *Endodontic Topics* 10.1 (2005): 3-29.
8. Martins JNR, *et al.* “Worldwide analyses of maxillary first molar second mesiobuccal prevalence: a multicenter cone-beam computed tomographic study”. *Journal of Endodontics* 44.11 (2018): 1641-1649.
9. Setzer FC and Lee SM. “Radiology in endodontics”. *Dental Clinics of North America* 65.3 (2021): 475-486.
10. Aung NM and Myint KK. “Diagnostic accuracy of CBCT for detection of second canal of permanent teeth: a systematic review and meta-analysis”. *International Journal of Dentistry* 2021 (2021): 1107471.
11. Parker J, *et al.* “CBCT uses in clinical endodontics: the effect of CBCT on the ability to locate MB2 canals in maxillary molars”. *International Endodontic Journal* 50.12 (2017): 1109-1115.
12. Parker JM, *et al.* “Cone-beam computed tomography uses in clinical endodontics: observer variability in detecting periapical lesions”. *Journal of Endodontics* 43.2 (2017): 184-187.
13. Setzer FC, *et al.* “The use of artificial intelligence in endodontics”. *Journal of Dental Research* 103.9 (2024): 853-862.
14. Umer F and Habib S. “Critical analysis of artificial intelligence in endodontics: a scoping review”. *Journal of Endodontics* 48.2 (2022): 152-160.
15. Samaranayake L, *et al.* “The transformative role of artificial intelligence in dentistry: a comprehensive overview. Part 1: fundamentals of AI, and its contemporary applications in dentistry”. *International Dental Journal* 75.2 (2025): 383-396.
16. Page MJ, *et al.* “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews”. *BMJ* 372 (2021): n71.
17. Whiting PF, *et al.* “QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies”. *Annals of Internal Medicine* 155.8 (2011): 529-536.
18. Mongan J, *et al.* “Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers”. *Radiology: Artificial Intelligence* 2.2 (2020): e200029.
19. Collins GS, *et al.* “Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence”. *BMJ Open* 11.7 (2021): Article e048008.
20. Bossuyt PM, *et al.* “STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies”. *BMJ* 351 (2015): Article h5527.
21. Xiang Y, *et al.* “Morphology and classification of the second mesiobuccal canal in maxillary first molars: a cone-beam computed tomography analysis in a Chinese population”. *BMC Oral Health* 24.1 (2024): Article 568.
22. Manigandan K, *et al.* “Impact of dental operating microscope, selective dentin removal and cone beam computed tomography on detection of second mesiobuccal canal in maxillary molars: a clinical study”. *Indian Journal of Dental Research* 31.4 (2020): 526-530.

23. Albitar L., *et al.* “Artificial intelligence (AI) for detection and localization of unobturated second mesial buccal (MB2) canals in cone-beam computed tomography (CBCT)”. *Diagnostics* 12.12 (2022): Article 3214.
24. Duman SB., *et al.* “Second mesiobuccal canal segmentation with YOLOv5 architecture using cone beam computed tomography images”. *Odontology* 112.2 (2024): 552-561.
25. Shetty S., *et al.* “Machine learning models in the detection of MB2 canal orifice in CBCT images”. *International Dental Journal* 75.3 (2025): 1640-1648.
26. Ji Y., *et al.* “Construction and evaluation of an AI-based CBCT resolution optimization technique for extracted teeth”. *Journal of Endodontics* 50.9 (2024): 1298-1306.
27. Güllü DM, Orhan K and Kartal N. “Three-dimensional automatic segmentation of root canals with focus on the second mesiobuccal canal using nnU-Netv2 on CBCT images: deep learning approach”. *BMC Oral Health* 26.1 (2026): Article 973.
28. Lin J., *et al.* “Deep learning-based detection of the second mesiobuccal canal in maxillary first molars using cone-beam computed tomography”. *BMC Oral Health* 26.1 (2026): Article 974.
29. Mansour S., *et al.* “Two step approach for detecting and segmenting the second mesiobuccal canal of maxillary first molars on cone beam computed tomography (CBCT) images via artificial intelligence”. *BMC Oral Health* 25.1 (2025): Article 1404.
30. Ourang SAH., *et al.* “Artificial intelligence in the detection of clinically negotiable second mesio-buccal canals in periapical images of maxillary molars”. *International Endodontic Journal* (2025).
31. Karatas E., *et al.* “Artificial intelligence for root canal orifice identification using dental operating microscope images: a preliminary evaluation”. *Australian Endodontic Journal* (2025).
32. Jin L., *et al.* “Deep-learning network for automated evaluation of root-canal filling radiographic quality”. *European Journal of Medical Research* 30.1 (2025): Article 297.
33. Allihaibi M., *et al.* “Diagnostic accuracy of an artificial intelligence-based platform in detecting periapical radiolucencies on cone-beam computed tomography scans of molars”. *Journal of Dentistry* 160 (2025): Article 105854.
34. Sadr S., *et al.* “Deep learning for detection of periapical radiolucent lesions: a systematic review and meta-analysis of diagnostic test accuracy”. *Journal of Endodontics* 49.3 (2023): 248-261.
35. Pul U and Schwendicke F. “Artificial intelligence for detecting periapical radiolucencies: a systematic review and meta-analysis”. *Journal of Dentistry* 147 (2024): Article 105104.
36. Pul U., *et al.* “Impact of artificial intelligence assistance on diagnosing periapical radiolucencies: a randomized controlled trial”. *Journal of Dentistry* 160 (2025): Article 105868.
37. Dashti M., *et al.* “Use of artificial intelligence for detection of MB2 canals in maxillary first molars on CBCT: a systematic review and meta-analysis”. *BMC Oral Health* 25.1 (2025): 1860.
38. Mohammad-Rahimi H., *et al.* “Research that matters: a call for enhancing rigour and relevance in artificial intelligence research in endodontics”. *International Endodontic Journal* (2026).
39. Tuygunov N., *et al.* “The transformative role of artificial intelligence in dentistry: a comprehensive overview. Part 2: the promise and perils, and the International Dental Federation communiqué”. *International Dental Journal* 75.2 (2025): 397-404.

Volume 25 Issue 11 November 2026

©All rights reserved by Salah A Yousief., *et al.*