International Platform of Registered Systematic Review and Meta-analysis Protocols

INPLASY

The accuracy of artificial intelligence in detecting success or failure of endodontic treatments. Systematic review

INPLASY2025110008
doi: 10.37766/inplasy2025.11.0008
Received: 4 November 2025
Published: 5 November 2025

Aqili, T.

Corresponding author:
Tariq Aqili

taqili@taibahu.edu.sa

Author Affiliation:
Taibah University.

ADMINISTRATIVE INFORMATION

Support - No external financial support was received for the preparation of this protocol.

Review Stage at time of this submission - Preliminary searches.

Conflicts of interest - None declared.

INPLASY registration number: INPLASY2025110008

Amendments - This protocol was registered with the International Platform of Registered Systematic Review and Meta-Analysis Protocols (INPLASY) on 5 November 2025 and was last updated on 5 November 2025.

INTRODUCTION

Review question / Objective The review was structured using the PICO framework:
• Population (P): Patient with surgical or non-surgical endodontics intervention.
• Intervention (I): Use of different types of AI in dental radiographs or clinical investigation for detection of success or failure.
• Comparator (C): Human or conventional methods.
• Outcome (O): Diagnostic performance metrics of AI systems, including accuracy, sensitivity, specificity, area under the curve (AUC), and predictive values.

The focused research question was:
“What is the accuracy of artificial intelligence in detecting success or failure of endodontic treatments?”

Condition being studied Usually, two-dimensional periapical radiographs or three-dimensional cone-beam computed tomography (CBCT) scans interpreted by clinicians are used to assess the success rate. However, human interpretation is subjective and subject to intr- and intera-observer variability. Additionally, even experienced dentists may miss superfine radiographic signs of periapical pathology or misdiagnose healing lesions.
In comparison, artificial intelligence (AI) systems provide several advantages over the traditional way. First of all, AI reduces the variability in inter- and intra agreement seen in human diagnosis by providing standardized, reproducible assessments. In addition, machine Learning Models can analyze numerous number of radiographs or CBCT slices in seconds, which is impossible for clinicians. Furthermore, AI systems can use clinical data with the the radiograohes (e.g., patient age, systemic health, treatment variables) to predict the success rate over the time. Also, AI is found to be acceptable tool as a second opinion. It can assist clinicians in treatment planing wether root canal retreatment, endodontic surgery, or extraction.

At the moment, there is no systematic review that has comprehensively synthesized the accuracy of AI in detecting success or failure of endodontic treatments. With rapid expansion of AI research, there is a crucial need to collect available evidence, evaluate the methodological quality published articles, and summerize clinical implications.

METHODS

Participant or population Patients who have undergone endodontic (root canal) treatments in clinical or retrospective study settings.

Intervention Use of artificial intelligence (AI)-based methods (e.g., machine learning classifiers, deep learning networks) for evaluation or predicting success or failure of root canal treatment.

Comparator Conventional diagnostic/predictive methods (such as clinician assessment, radiographic evaluation, traditional statistical models) or AI method.

Study designs to be included Retrospective/ prospective cohort studies, case-control studies, and machine learning validation studies evaluating AI methods regarding outcomes of endodontic treatments.

Eligibility criteria Inclusion criteria
1. Study type: Only original research that evaluated AI/ML models for detecting or predicting endodontic treatment outcomes were included (retrospective, prospective, cross-sectional, in vivo, ex vivo, or in silico).
2. Population: Human subjects (clinical or radiographic datasets) or human-derived imaging (including cadaveric samples).
3. Intervention: Any AI technique (CNNs, ANNs, hybrid ML models, commercial AI platforms).
4. Outcomes: Extractable diagnostic performance metrics (accuracy, sensitivity, specificity, AUC, F1-score, or predictive values).
5. Language: English only.

Information sources A comprehensive search of five electronic databases will be performed:
• PubMed/MEDLINE
• Scopus
• Web of Science
• Embase
• IEEE Xplore (to capture computer science/ engineering contributions)
The search will be supplemented by manual screening of reference lists of included articles and relevant reviews, as well as a grey literature search in , ProQuest Dissertations & Theses, and preprint servers (arXiv, medRxiv).
Search terms combine controlled vocabulary (MeSH, Emtree) and free-text keywords related to endodontics, AI, and outcomes.
The search will be conducted for articles published up to March 2025. Authors will be contacted for missing data.

Main outcome(s) Accuracy measures of AI models in detecting or predicting endodontic treatment outcomes, including sensitivity, specificity, area under the receiver-operating characteristic curve (AUC), positive predictive value (PPV), negative predictive value (NPV), and overall accuracy.

Quality assessment / Risk of bias analysis The QUADAS-2 tool (Quality Assessment of Diagnostic Accuracy Studies-2) will be used to evaluate the risk of bias in four domains:
1. Patient selection
2. Index test (AI model)
3. Reference standard (clinician or gold standard diagnosis)
4. Flow and timing
Each domain will be rated as “low risk,” “high risk,” or “unclear risk.”
For in vitro studies, risk of bias Wii be adapted from the modified QUADAS framework, with emphasis on dataset representativeness, blinding, and reproducibility.
Two reviewers independently performe risk of bias assessment. If tehre is discrepancies, they will be resolved by discussion.

Strategy of data synthesis Due to anticipated heterogeneity in AI models, imaging modalities, and outcome definitions, a narrative synthesis will be planned. Key performance indicators (accuracy, sensitivity, specificity, AUC) will be tabulated.
If sufficient homogeneous data are available, a meta-analysis using a random-effects model would be conducted to estimate pooled diagnostic accuracy metrics. Heterogeneity would be assessed using the I^2 statistic.

Subgroup analysis Subgroup analyses will be conducted by AI method type (machine learning vs. deep learning), study design (retrospective vs. prospective), and imaging modality (2D radiograph vs. 3D CBCT).

Sensitivity analysis Sensitivity analyses will exclude studies at high risk of bias, or studies using small sample sizes (<10 cases) to assess their impact on pooled estimates.

Country(ies) involved Saudi Arabia.

Keywords endodontics; root canal; artificial intelligence; machine learning; treatment outcome; diagnostic accuracy.

Contributions of each author

Author 1 - taqili@taibahu.edu.sa.
Email: taqili@taibahu.edu.sa