

Systematic review and meta-analysis of artificial intelligence approaches for dental age assessment using orthopantomographic images

Cristiana Palmela Pereira, Alexandre Francisco, Cátia Vaz, Arlindo Oliveira, Henrique Oliveira, José Figueira, Francisco Salvado, Rui Santos, Andreas Schmeling, Hrvoje Brkić, Chandra Gracias, Rui Cardoso, Paulo Guerra, Ana Rodrigues, Diana Augusto, Valon Nushi, Renato Marcelo, António Figueiras, João Martins

Citation

Cristiana Palmela Pereira, Alexandre Francisco, Cátia Vaz, Arlindo Oliveira, Henrique Oliveira, José Figueira, Francisco Salvado, Rui Santos, Andreas Schmeling, Hrvoje Brkić, Chandra Gracias, Rui Cardoso, Paulo Guerra, Ana Rodrigues, Diana Augusto, Valon Nushi, Renato Marcelo, António Figueiras, João Martins. Systematic review and meta-analysis of artificial intelligence approaches for dental age assessment using orthopantomographic images. PROSPERO 2025 CRD420251117151. Available from <https://www.crd.york.ac.uk/PROSPERO/view/CRD420251117151>.

REVIEW TITLE AND BASIC DETAILS

Review title

Systematic review and meta-analysis of artificial intelligence approaches for dental age assessment using orthopantomographic images

Condition or domain being studied

Adolescence; Artificial intelligence; Machine learning; Convolutional neural network ; Orthopantomogram; Full Mouth X-ray Of Teeth; Assessment Using Assessment Scale; Regression analysis ; Sensitivity

This systematic review and meta-analysis focuses on the accuracy and performance of artificial intelligence (AI) approaches for dental age assessment (DAA), based on orthopantomographic images (OPGs). Dental age estimation plays a fundamental role in forensic contexts, particularly when assessing the chronological age of minors, undocumented individuals, asylum seekers, and juvenile offenders, where legal age thresholds (e.g., 14, 16, 18, and 21 years) are critical.

The review aims to synthesize and compare various AI architectures—such as convolutional neural networks (CNNs)—regarding their ability to estimate age through dental development stages. These approaches are evaluated against traditional manual methods performed by

forensic dental experts. Performance indicators include accuracy, mean absolute error (MAE), sensitivity, specificity, and F1-score, among others. This review addresses the need for standardized, reproducible, and accurate medical assessment methods to support equitable forensic decision-making and to reduce geographic and methodological bias in age estimation.

Rationale for the review

This systematic review and meta-analysis aims to evaluate whether current artificial intelligence (AI) models used for dental age assessment (DAA) based on orthopantomographic images (OPGs) are sufficiently accurate and reliable to be applied as decision-support tools in forensic and legal contexts.

While traditional manual assessments performed by forensic dental experts remain the reference standard, AI models may offer enhanced performance in terms of diagnostic accuracy, reproducibility, and speed, especially in scenarios involving high caseloads or limited expert availability. However, it remains unclear whether these AI approaches consistently outperform or match the diagnostic accuracy of manual methods.

This review will fill that gap by systematically comparing the performance of various AI architectures to manual assessment, assessing whether they meet the requirements to be safely integrated into forensic decision-making processes. The results may guide the development of standardized, evidence-based AI tools for dental age estimation, with implications for forensic odontology, legal medicine, and judicial systems.

Review objectives

The review aim is to assess the efficacy and performance of the artificial intelligence approaches, when assessing the dental age, through dental development stages techniques using orthopantomogram images.

First objective: to evaluate the performance and accuracy of artificial intelligence approaches in assessing the dental age, through dental development stages techniques using orthopantomogram images.

Second objective: to compare the artificial intelligence approaches in dental age assessment with performance of traditional manual methods by forensic experts.

How accurate is the performance of artificial intelligence approaches in assessing dental age and how is comparable to the traditional manual age assessment methods applied by forensic experts?

Keywords

Artificial intelligence; Dental Age Assessment; Orthopantomogram; Forensic Odontology; Diagnostic Accuracy

Country

Portugal

ELIGIBILITY CRITERIA

Population

Included

Population inclusion criteria:

Individuals undergoing the dental age assessment (DAA) process, regardless of age, sex, or geographic origin. There will be no restrictions on the age of the subjects. Studies must involve the use of orthopantomographic images (OPGs) for age estimation using artificial intelligence approaches.

Excluded

Studies that apply artificial intelligence models for medical purposes unrelated to dental age assessment will be excluded. Additionally, studies will be excluded if they lack full-text availability, or if they are review articles, systematic reviews, or case reports.

Intervention(s) or exposure(s)***Included***

Convolutional neural network ; Machine learning; Artificial intelligence; Orthopantomogram; Radiography Of Teeth

The index test includes various artificial intelligence (AI) approaches applied to orthopantomographic images (OPGs) for the purpose of dental age assessment (DAA). Eligible studies must apply AI techniques—such as convolutional neural networks (CNNs), deep learning, machine learning, or other neural network-based models—specifically for the estimation of chronological age based on dental development stages.

Studies must clearly describe the AI model architecture, training process, and performance metrics related to age estimation. Only studies using orthopantomograms (panoramic dental radiographs) as the input imaging modality will be included. The use of hyperparameter optimization, loss functions, model interpretability tools (e.g., heatmaps, Grad-CAM), and validation procedures (internal or external) will also be considered as relevant exposure elements.

Excluded

Studies applying artificial intelligence models for medical purposes other than dental age assessment will be excluded. This includes AI models used for diagnosis, treatment planning, or classification of other dental, craniofacial, or systemic conditions unrelated to chronological age estimation. Additionally, studies not based on orthopantomographic (panoramic dental) images as the primary input for the AI model will be excluded.

Comparator(s) or control(s)***Included***

PICO tags selected: Functional Assessment; Assessment Using Assessment Scale; Forensic Examination

Comparisons among various artificial intelligence (AI) approaches for dental age assessment, including different neural network architectures, and comparisons between AI-based methods and traditional manual dental age estimation performed by forensic dental experts.

Excluded

Studies will be excluded only if they do not include any form of comparative analysis, whether between artificial intelligence models or between AI and manual dental age assessment

methods. Studies that do not report any diagnostic performance metrics (e.g., accuracy, MAE, sensitivity) for the models or methods being compared will also be excluded.

Study design

Only nonrandomized study types will be included.

Included

This review will include diagnostic accuracy studies with a computational approach, in which dental age assessment is performed using artificial intelligence models applied to orthopantomographic images (OPGs). Eligible studies include retrospective observational studies, computational experimental studies, and model performance evaluations that compare either different AI architectures or AI against manual forensic methods. Only primary studies published between 2020 and 2025 will be included.

Excluded

Studies will be excluded if they are review articles, systematic reviews, case reports, or conference abstracts without full-text availability. Additionally, the exclusion criteria of this study will be use of artificial intelligence models for other medical purposes.

Context

This review focuses on studies that apply artificial intelligence (AI) models to orthopantomographic (panoramic dental) images for the purpose of dental age assessment in living individuals. The context includes forensic, legal, clinical, humanitarian, and administrative settings, where accurate estimation of chronological age is required — such as in cases of asylum seeking, undocumented individuals, juvenile justice, human trafficking, or legal age verification.

There are no restrictions on country, sex, age, or institutional setting. Studies from any geographic or economic region will be included, provided they meet the methodological criteria. Only studies using AI for age estimation will be included; those applying AI to other medical diagnostics will be excluded.

TIMELINE OF THE REVIEW

Date of first submission to PROSPERO

22 August 2025

Review timeline

Start date: 31 July 2025. End date: 31 October 2025.

Date of registration in PROSPERO

27 August 2025

AVAILABILITY OF FULL PROTOCOL

Availability of full protocol

A full protocol has been written and uploaded to PROSPERO. The protocol will be made available after the review is completed.

SEARCHING AND SCREENING

Search for unpublished studies

Only published studies will be sought.

Main bibliographic databases that will be searched

The main databases to be searched are *CLIB - The Cochrane Library*, *PubMed* and *Scopus*.

Search language restrictions

There are no language restrictions.

Search date restrictions

Databases will be searched for articles published from 1 January 2020 and before by 29 July 2025.

Other methods of identifying studies

No other methods will be used.

Link to search strategy

A full search strategy is available in the full protocol as described in the *Availability of full protocol* section

Selection process

Studies will be screened independently by at least two people (or person/machine combination) with a process to resolve differences.

Other relevant information about searching and screening

In this systematic review, the criteria of the "Centre for Evidence-Based Medicine (CEBM)" of Oxford will be used to analyze and establish the quality of the articles included in the review.

Two reviewers will assess the risk of bias independently by duplicate. In this review two independent investigators will review the titles and abstracts initially. Subsequently, after the initial selection, they will read the full text of the selected articles to select those studies that meet the inclusion and exclusion criteria. Thus, the Kappa coefficient will be calculated at each step to determine interobserver agreement. Disagreements between the reviewers will be resolved by discussion. If necessary, a third reviewer will be involved in the selection. If there are needed article's lack information, the author will be contacted.

DATA COLLECTION PROCESS

Data extraction from published articles and reports

Data will be extracted independently by at least two people (or person/machine combination) with a process to resolve differences.

Authors will be asked to provide any required data not available in published reports.

Study risk of bias or quality assessment

Risk of bias will be assessed using:

The risk of bias and applicability of the included studies will be assessed using the QUADAS-AI tool, which is specifically designed for systematic reviews of diagnostic accuracy studies

involving artificial intelligence.

Sunderajah, V., et al. (2021). A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nature Medicine*, 27(10),1663–1665.

Data will be assessed independently by at least two people (or person/machine combination) with a process to resolve differences.

Additional information will be sought from study investigators if required information is unclear or unavailable in the study publications/reports.

Reporting bias assessment

If more than ten studies are included in the meta-analysis, publication bias and risk of bias due to missing results will be assessed through funnel plots, as well as Egger's and Begg's tests, in accordance with established guidelines for diagnostic accuracy reviews.

Certainty assessment

The certainty of evidence for each key outcome will be assessed using the GRADE (Grading of Recommendations, Assessment, Development and Evaluations) framework. The following domains will be considered: risk of bias (using QUADAS-AI), inconsistency (heterogeneity across studies), indirectness (relevance of the population, intervention, comparator, and outcomes to the review question), imprecision (based on sample size and confidence intervals), and publication bias (assessed via funnel plots and Egger/Begg tests when applicable).

Each outcome will be rated as having high, moderate, low, or very low certainty, and summary tables will be developed to transparently report the level of confidence in the evidence base to support forensic and clinical decision-making.

OUTCOMES TO BE ANALYSED

Main outcomes

To evaluate the performance of dental age assessment using artificial intelligence approaches applied to orthopantomogram (OPG) images.

Additional outcomes

Secondary outcome(s)

1. Comparison between different artificial intelligence architectures in terms of diagnostic performance.
2. Comparison between artificial intelligence and manual forensic expert assessments (human-based estimations).
3. Analysis of model generalizability (e.g., external validation, cross-population datasets).
4. Subgroup and threshold analyses based on:
 - Age thresholds of 14, 16, 18, and 21 years (legal and forensic relevance);
 - Sex (when available);
 - Geographic origin of dataset.
5. Assessment of model explainability, including visual techniques such as heatmaps, Grad-CAM or saliency maps, to support forensic transparency.

6. Computational efficiency of artificial models (e.g., training/inference time, number of parameters).
7. Dataset quality indicators, such as image resolution, number of OPGs, and annotation reliability.
8. Bias or fairness analysis, e.g., performance discrepancies by sex, population or ethnic group (if reported).

PLANNED DATA SYNTHESIS

Strategy for data synthesis

Data will be combined through both traditional and network meta-analysis approaches.

For direct pairwise comparisons where at least two studies are available, a traditional meta-analysis will be performed using relative risk (RR) for categorical outcomes and mean difference for continuous outcomes (e.g., MAE, accuracy).

When multiple interventions (AI models) are compared across studies, a Frequentist network meta-analysis will be conducted using a common comparator. Heterogeneity will be assessed using the I^2 statistic and τ^2 , and consistency between direct and indirect estimates will be evaluated through inconsistency models. Ranking probabilities of model performance will be estimated using the Frequentist analogue of SUCRA (surface under the cumulative ranking curve).

All analyses will be conducted using the "meta" and "netmeta" packages in R (version 4.4.0) and IBM SPSS version 30.0.

CURRENT REVIEW STAGE

Stage of the review at this submission ^{1 change}

Review stage	Started	Completed
Pilot work	✓	✓
Formal searching/study identification	✓	✓
Screening search results against inclusion criteria	✓	✓
Data extraction or receipt of IPD	✓	✓
Risk of bias/quality assessment	✓	✓
Data synthesis	✓	✓

Review status

The review is completed.

Publication of review results

Results of the review will be published in English.

REVIEW AFFILIATION, FUNDING AND PEER REVIEW

Review team members

Professor Cristiana Palmela Pereira (review guarantor and contact) ORCID: 0000-0002-9164-7189. Faculdade de Medicina Dentária Universidade de Lisboa/ CEAUL. Portugal.

No conflict of interest declared.

Professor Alexandre Francisco. Instituto Superior Técnico, Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor Cátia Vaz. Lisbon School of Engineering (ISEL). Portugal.

No conflict of interest declared.

Professor Arlindo Oliveira. Instituto Superior Técnico, Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor Henrique Oliveira. Instituto Superior Técnico, Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor José Figueira. Instituto Superior Técnico, Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor Francisco Salvado. Faculdade de Medicina da Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor Rui Santos. School of Technology and Management of Leiria, Polytechnic Institute of Leiria. Portugal.

No conflict of interest declared.

Professor Andreas Schmeling. Institute of Legal Medicine University Hospital Munster, Germany. Study Group on Forensic Age Diagnostics (AGFAD). Germany.

No conflict of interest declared.

Professor Hrvoje Brkić. University of Zagreb School of Dental Medicine. Croatia.

No conflict of interest declared.

Dr Chandra Gracias. Family Court. Portugal.

No conflict of interest declared.

Dr Rui Cardoso. Procuradoria da República. Portugal.

No conflict of interest declared.

Dr Paulo Guerra. Tribunal da Relação. Portugal.

No conflict of interest declared.

Dr Ana Rodrigues. Faculdade de Medicina Dentária da Universidade de Lisboa. Portugal.

No conflict of interest declared.

Dr Diana Augusto. Faculdade de Medicina Dentária da Universidade de Lisboa. Portugal.

No conflict of interest declared.

Dr Valon Nushi. Faculdade de Medicina da Universidade de Lisboa. Portugal.

No conflict of interest declared.

Dr Renato Marcelo. Instituto Superior Técnico Universidade de Lisboa. Portugal.

No conflict of interest declared.

Dr António Figueiras. Instituto Superior Técnico Universidade de Lisboa. Portugal.

No conflict of interest declared.

Professor João Martins. Escola Superior de Saúde do Politécnico do Porto. Portugal.

No conflict of interest declared.

Named contact

Professor Cristiana Palmela Pereira (cpereira@campus.ul.pt). ORCID: 0000-0002-9164-7189. Faculdade de Medicina Dentária Universidade de Lisboa/ CEAUL. Portugal.

Review affiliation

Centro de Estatística e Aplicações da Universidade de Lisboa- CEAUL

Instituto de Engenharia de Sistemas e Computadores - INESC ID

Funding source

Grant number

2024.07444.IACDC

Additional non-commercial funding information

Artificial Intelligence and its Application by Forensic Science Service Providers: Migrant Unidentified Age Estimation - reference Grant 2024.07444.IACDC.

<https://doi.org/10.54499/2024.07444.IACDC>. FCT.

Peer review

Within the scope of project 2024.07444.IACDC, supported by the measure "RE-C05-i08.M04 – Support for the launch of a R&D project programme aimed at the development and implementation of advanced systems in cybersecurity, artificial intelligence and data science in public administration, as well as a scientific training programme", under the Recovery and Resilience Plan – PRR, and framed by the funding agreement signed between the Recovery Portugal Task Force (EMRP) and the Foundation for Science and Technology (FCT), as the intermediate beneficiary.

ADDITIONAL INFORMATION

Additional information 1 change

The present study is designed as a systematic review with meta-analysis, following a predefined protocol registered in PROSPERO (CRD420251117151). It incorporates a formal assessment of risk of bias using QUADAS-AI and evaluates the certainty of evidence using the GRADE framework, ensuring methodological rigor and reproducibility.

This review has not been commissioned by a specific guideline-developing body; however, it is conducted within the scope of the funded research project "Artificial Intelligence and its Application by Forensic Science Service Providers: Migrant Unidentified Age Estimation" (Grant 2024.07444.IACDC, FCT/PRR). As such, it is intended to inform the development of evidence-based frameworks and decision-support tools for forensic and judicial applications, particularly

in the context of age estimation of living individuals, where legal thresholds (e.g., 14, 16, 18, and 21 years) are critical for judicial decision-making .

The review distinguishes itself from existing literature by:

(i) restricting inclusion to primary diagnostic accuracy studies using orthopantomographic images and AI-based methods;

(ii) applying comparative analyses between AI models and traditional forensic expert assessments;

(iii) incorporating advanced synthesis methods, including network meta-analysis and ranking of model performance;

(iv) addressing key dimensions such as model generalizability, explainability (e.g., Grad-CAM), and fairness across populations.

In addition to the listed research team, the review benefits from interdisciplinary collaboration involving experts in forensic odontology, artificial intelligence, statistics, legal medicine, and judicial practice, including contributors affiliated with academic institutions, forensic institutes, and judicial bodies, as detailed in the PROSPERO registration.

Review conflict of interest

Declared individual interests are recorded under team member details.. No additional interests are recorded for this review.

Medical Subject Headings

Artificial Intelligence; Age Determination by Teeth; Deep Learning; Machine Learning; Neural Networks, Computer; Refugees

Revision note ¹ change

The review has progressed from the protocol stage to completion, and the full manuscript has now been submitted for publication (Research Square preprint: <https://www.researchsquare.com/article/rs-8535308/v1>). The update reflects the transition to a completed review, including final data synthesis and reporting of results.

SIMILAR REVIEWS

Check for similar records already in PROSPERO

PROSPERO identified a number of existing PROSPERO records that were similar to this one (last check made on 31 July 2025). These are shown below along with the reasons given by that the review team for the reviews being different and/or proceeding.

- Artificial Intelligence Methods in Forensic Dentistry Identification Dentomaxillofacial (DMF) Imaging approach. A systematic Review [published 5 July 2025] [CRD420250653013]. The review was judged **not to be similar**
- Assessing Age in Adults: A Systematic Review and Meta-Analysis of Dental Radiographic Techniques [published 5 December 2023] [CRD42023481814]. The review was judged **not to be similar**
- Comparative Analysis of Manual and AI Image Segmentation Accuracy in Panoramic Dental Imaging: A Systematic Review and Meta-Analysis [published 25 December 2023] [CRD42023493455]. The review was judged **not to be similar**

PROSPERO version history 1 change

- [Version 1.1, published 26 Mar 2026](#)
- [Version 1.0, published 27 Aug 2025](#)

Disclaimer

The content of this record displays the information provided by the review team. PROSPERO does not peer review registration records or endorse their content.

PROSPERO accepts and posts the information provided in good faith; responsibility for record content rests with the review team. The guarantor for this record has affirmed that the information provided is truthful and that they understand that deliberate provision of inaccurate information may be construed as scientific misconduct.

PROSPERO does not accept any liability for the content provided in this record or for its use. Readers use the information provided in this record at their own risk.

Any enquiries about the record should be referred to the named review contact