Iran J Orthod. 2026 December; 21(2):e1507. doi:10.32592/ijorth.21.2.1507.
Published online 2026 December 1 Research Article
Systematic Review **Open Access**

AI-Assisted Remote Monitoring Versus Conventional Orthodontic Follow-Up: A Systematic Review

Aneri Ashokkumar Patel [1*], Khyati Viral Patel [2], Alpesh Manilal Vaghela [3]

[1] Resident, Department of Orthodontics and Dentofacial Orthopedics, Narsinhbhai Patel Dental College and Hospital, Sankalchand Patel University, Visnagar, Gujarat, India.
[2] B.D.S., M.D.S., PhD, Professor and Head, Department of Orthodontics and Dentofacial Orthopedics, Narsinhbhai Patel Dental College and Hospital, Sankalchand Patel University, Visnagar, Gujarat, India.
[3] B.D.S., M.D.S., Reader, Department of Orthodontics and Dentofacial Orthopedics, Narsinhbhai Patel Dental College and Hospital, Sankalchand Patel University, Visnagar, Gujarat, India.

***Corresponding author:** Aneri Ashokkumar Patel

Address: Department of Orthodontics and Dentofacial Orthopedics, Narsinhbhai Patel Dental College and Hospital, Sankalchand Patel University, Visnagar, Gujarat, India.

Email: aneripatel244@gmail.com

Received: 2026 October 30; **Revised:** 2026 November 18; **Accepted:** 2026 November 25

Abstract

Aim: This study synthesizes evidence comparing AI-assisted remote monitoring (AIDRM) with conventional in-office orthodontic follow-up in terms of treatment efficiency, diagnostic accuracy, and patient-reported outcomes.
Data sources: PubMed/MEDLINE, Scopus, Web of Science, Cochrane CENTRAL, ScienceDirect, IEEE Xplore, Embase, and Google Scholar (publication-date window 1 January 2019–31 October 2025; one study published online in January 2026 was retained as a declared deviation), supplemented by hand-searching and reference list screening.
Study selection: Two independent reviewers screened records per PRISMA 2020 (κ=0.84); 30 records met inclusion criteria, of which 16 primary comparative/diagnostic accuracy studies in orthodontic patients served as evidentiary units for synthesis; the other 14 (reviews) provided context only.
Data extraction and synthesis: Two reviewers independently extracted study, outcome, and funding/conflict-of-interest data; risk of bias was assessed with QUADAS-2, RoB 2, and ROBINS-I, with GRADE certainty rating. Substantial heterogeneity precluded meta-analysis; synthesis was narrative.
Results: AI-assisted monitoring reduced in-office appointments by 23–33% without compromising occlusal outcomes, though treatment duration varied (RCT: +1.9 months; cohort: −1.7 months). Diagnostic accuracy for fixed-appliance emergencies was high (sensitivity: 91.1–98.4%, specificity: 88.3–99.6%) but rested substantially on two sponsor-linked studies (McCray et al., 2026 and 2025; refs 27 and 28), one funded by the manufacturer. Sensitivity for plaque, gingivitis, and recession was low (0.22–0.53). AI-monitored care improved oral hygiene and satisfaction.
Limitations: The evidence is dominated by one commercial platform; key diagnostic accuracy findings carry unresolved conflicts of interest; two studies lacked extractable sample sizes; heterogeneity precluded pooling.
Conclusions: AI-assisted remote monitoring is a promising but not yet independently validated adjunct to conventional follow-up. Independently funded, multi-center randomized trials with standardized outcomes are needed before routine substitution for in-person supervision.
Registration Information: Registered with PROSPERO (registration number CRD420261451606; registered 19 August 2026, after searching and screening had begun).

Keywords: Artificial intelligence, Corrective, Deep learning, Diagnostic accuracy, Orthodontics, Patient compliance, Telemedicine.

Copyright © 2023, Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/) which permits copy and redistribute the material just in noncommercial usages, provided the original work is properly cited.

Background

Orthodontic care built around fixed-appliance and clear aligner therapy has traditionally required visits every four to ten weeks for adjustments, monitoring tooth movement, reinforcing oral hygiene, and managing appliance-related incidents. The interval between these visits is poorly observed clinically: orthodontists depend on patient self-report for emergencies such as bracket debonding, lost aligner attachments, broken wires, or deteriorating oral hygiene, and the cumulative burden of visit frequency—in time, travel, and chairside resource allocation—has long been recognized as a structural limitation of orthodontic care delivery (1).

The COVID-19 pandemic exposed the fragility of this visit-dependent model. Routine orthodontic appointments were canceled or postponed worldwide while active tooth movement continued unsupervised, forcing the profession to adopt teledentistry—the use of information and communication technology for remote provision of dental services—almost overnight (2,3). Teleorthodontics, its orthodontic subset encompassing remote consultation, virtual treatment planning, and remote retention follow-up, was rapidly deployed to triage emergencies and maintain continuity of care (4,5).

This pragmatic acceleration coincided with the maturation of artificial intelligence (AI), and specifically of deep learning and convolutional neural networks (CNNs), for medical image analysis. CNN-based architectures have achieved diagnostic performance comparable to human experts in radiology, ophthalmology, and dermatology (6,7), and dentistry has followed a similar trajectory, with AI applied to caries detection, cephalometric landmark identification, periodontal bone-level assessment, and automated orthodontic model segmentation (8–10). For clarity, these terms are not used interchangeably in this review: artificial intelligence (AI) denotes the broad field; machine learning and, within it, deep learning and convolutional neural networks (CNNs) denote specific computational methods; teleorthodontics denotes a clinical care-delivery model (remote consultation, treatment planning, or follow-up, with or without AI); and AI-driven remote monitoring (AIDRM), the technology under review, denotes the specific convergence of the two—a teleorthodontics workflow in which patient-captured images are interpreted by a CNN-based algorithm rather than solely by a clinician. Each term is used in this precise sense throughout.

The convergence of teleorthodontics and AI has produced a distinct category of clinical technology—AI-driven remote monitoring (AIDRM). Patients capture intraoral images or video using a smartphone, with or without a dedicated scan box, most commonly through the DentalMonitoring® platform (Paris, France). Proprietary CNN-based algorithms analyze the uploaded images for bracket debonding, lost auxiliaries, aligner-tracking discrepancies, attachment loss, gingival inflammation, and plaque, and transmit alerts to the supervising orthodontist (11–13). Smartphone-only applications and AI-supported treatment-coaching tools such as StrojCHECK have followed in rapid succession (14). At the outset, the published evidence on AIDRM is disproportionately generated using a single commercial system, DentalMonitoring® (Dental Monitoring SAS, Paris, France); this concentration shapes both the interpretation of the findings below and is returned to explicitly as a limitation in the Discussion section.

Clinical adoption of these platforms has outpaced independent evaluation. Orthodontic practices across North America, Europe, the Middle East, and Asia increasingly integrate DentalMonitoring® or comparable systems into routine clear aligner therapy (15), with manufacturers claiming fewer in-person appointments, faster identification of tracking problems, improved oral hygiene, and higher patient satisfaction. The research community has, however, raised legitimate concerns regarding diagnostic accuracy validation, dependence on patient-generated image quality, algorithmic opacity, regulatory pathways, data privacy, and the medico-legal allocation of responsibility when an algorithm misses a clinical event (16–18).

For this review, AI-driven remote monitoring (AIDRM) is defined as any platform in which patient-captured intraoral images or video are analyzed by AI-based image analysis algorithms, with findings transmitted to the supervising clinician, in contrast to conventional care, defined as the prevailing standard of in-person orthodontic examination and adjustment without such algorithmic mediation. This distinction matters because several published studies compare AI-augmented teledentistry with non-AI telemonitoring rather than conventional in-person care, and these comparisons address different clinical questions.

The existing evidence base on AI-based remote dental and orthodontic monitoring is heterogeneous in design, outcome measurement, and risk of bias. Some studies report substantial reductions in face-to-face appointments without loss of clinical effectiveness; others describe longer overall treatment durations with remote monitoring; and still others identify concerning sensitivity deficits for plaque, gingivitis, and recession. Without a structured systematic synthesis, practicing orthodontists must rely on manufacturer literature, single-center studies, or anecdote when deciding whether to adopt AI-based remote monitoring platforms. A balanced systematic review grounded in PRISMA 2020 methodology, and framed explicitly around the orthodontic clinical question, is therefore both timely and necessary. This systematic review aimed to synthesize available

evidence comparing AI-assisted remote monitoring with conventional in-office orthodontic follow-up in terms of treatment efficiency, clinical and diagnostic outcomes, patient compliance, and visit frequency.

Methods

Study Design and Protocol Registration

This systematic review was conducted and reported in accordance with the PRISMA 2020 statement (19). The review protocol, including the research question, search strategy, eligibility criteria, and analysis plan, was developed and fixed before data extraction began. The protocol was registered with PROSPERO (registration number CRD420261451606; first submitted on 11 August 2026 and registered on 19 August 2026). The registered review start date was 15 January 2026, and the record shows that formal searching and title/abstract screening had begun by the time of registration; registration was therefore retrospective for those stages, although the record shows that data extraction had not yet started. Differences between the registered record and the review as conducted are listed under Protocol Deviations below.

Protocol Deviations

Compared with the registered record, the following differences apply. (i) The registered main sources (Embase, Google Scholar, LILACS, MEDLINE, PubMed and Scopus) were extended to include Web of Science, Cochrane CENTRAL, ScienceDirect and IEEE Xplore, and LILACS was not searched. (ii) The Newcastle–Ottawa Scale was listed among the registered risk-of-bias tools but was not used; all non-randomized comparative studies were appraised with ROBINS-I, and no Joanna Briggs Institute checklist was required because no cross-sectional study met the eligibility criteria. (iii) Adverse events were considered although they were not listed among the registered outcomes. (iv) The registered study design field stated that only randomized designs would be included, whereas the registered eligibility text listed a range of designs; the review followed the eligibility text. (v) Meta-analysis, planned where at least three studies of comparable design reported the same outcome using sufficiently homogeneous methods, was not performed because this condition was not met. (vi) One primary study (McCray et al.) (20) was published online on 29 January 2026, after the close of the publication-date eligibility window (31 October 2025); because the searches were run after the registered review start date (15 January 2026), this record was retrieved, and it was retained as a declared deviation from the publication-date window because it is the largest diagnostic accuracy evaluation of the platform under review and addresses the review question directly. Its inclusion is identified as such in the Search Strategy, the Results, Table 1, and the Limitations.

Focused Question (PICO/PECO)

Population: orthodontic patients of any age undergoing active fixed appliance or clear aligner therapy, or in the retention phase. Intervention/Exposure: AI-assisted remote monitoring systems, including DentalMonitoring® and comparable machine learning- or deep learning-based platforms. Comparator: conventional in-office orthodontic monitoring without AI-assisted remote surveillance; for diagnostic accuracy and agreement studies, the comparator was the conventional reference standard (clinical examination, orthodontist consensus, intraoral scan, or plaster models). Outcomes: treatment duration, number of in-office visits, patient compliance, oral hygiene indices, diagnostic accuracy for appliance- and hygiene-related findings, treatment accuracy (e.g., Peer Assessment Rating index change), patient satisfaction, and adverse events (not listed among the registered outcomes but considered given its clinical importance; as reported in the Results section, no included study provided extractable adverse-event data, and this is flagged as an evidence gap rather than silently omitted). A parallel PECO framework accommodated observational studies in which AI-based remote care was an exposure rather than a randomly allocated intervention.

Eligibility Criteria

Peer-reviewed studies published in English between 1 January 2019 and 31 October 2025 were eligible as primary (evidentiary) studies if they (i) enrolled orthodontic patients undergoing active fixed appliance or clear aligner therapy or in the retention phase; (ii) evaluated an AI, machine learning, deep learning, or CNN-based remote-monitoring system applied to patient-captured images or video; (iii) compared that system either with conventional in-office care (randomized controlled trials, prospective and retrospective cohorts, case–control and other controlled studies, cross-sectional surveys, and clinically relevant pilot studies) or, for diagnostic accuracy and agreement studies, with a conventional reference standard (clinical examination, orthodontist consensus, intraoral scan, or plaster models); and (iv) reported at least one measurable diagnostic, monitoring, compliance, clinical, or patient experience outcome. Systematic, scoping, and narrative reviews were eligible for contextual purposes only and were not treated as evidentiary units (see Data Synthesis). Studies unrelated to dentistry or to AI/teledentistry, studies in non-orthodontic populations, studies of in-clinic (non-remote) AI image analysis, studies without a conventional-care or reference-standard comparator, animal-only studies, purely technical AI reports without clinical relevance, non-peer-reviewed commentaries, and duplicate publications were excluded.

Information Sources and Search Strategy

Eight electronic sources were searched exhaustively against the full Boolean strings, with no results cap—PubMed/MEDLINE, Scopus, Web of Science, Cochrane CENTRAL, ScienceDirect, IEEE Xplore, Embase, and Google Scholar. Google Scholar alone was screened only to its first 200 hits sorted by relevance, a widely used and recommended threshold for this database given its lack of exportable, exhaustive search history functionality and its tendency toward diminishing relevance beyond this point; all other seven databases were searched in full. Electronic searching was supplemented by manual reference list screening of included studies and recent reviews (3–5), and by hand-searching of core orthodontic journals not fully indexed in the databases above—including the American Journal of Orthodontics and Dentofacial Orthopedics, Progress in Orthodontics, the Angle Orthodontist, and the Journal of the World Federation of Orthodontists—for the review period 2019–2025. The eligibility window of 1 January 2019 to 31 October 2025 refers to the publication date of eligible reports; the searches themselves were run after the registered review start date (15 January 2026), and therefore also retrieved one primary study (McCray et al., 2026) (20) that was published online on 29 January 2026, after the close of that window. This study is retained as a declared deviation from the publication-date window (Protocol Deviations, item vi) and is identified as such in the Results, Table 1 and the Limitations. Records identified from each source before de-duplication are itemized in the Supplementary Search Table. Three layered Boolean search strings combined terms for artificial intelligence/machine learning/deep learning, teledentistry/teleorthodontics/remote monitoring, and orthodontic-specific terms (clear aligners, fixed appliances, Invisalign);

Search Strategy (representative string)

("artificial intelligence" OR "AI" OR "machine learning" OR "deep learning" OR "convolutional neural network" OR "neural network") AND ("teledentistry" OR "Teleorthodontics" OR "remote dental care" OR "remote monitoring" OR "dental monitoring" OR "DentalMonitoring") AND ("diagnostic accuracy" OR "treatment planning" OR "clinical outcomes" OR "orthodontic monitoring" OR "patient compliance" OR "oral hygiene"). Database-specific syntactic adaptations were used where required (for example, MeSH terms in PubMed: "Artificial Intelligence"(MeSH) AND "Teledentistry"(MeSH); controlled-vocabulary mapping in Embase via Emtree). Two further layered strings targeted orthodontic-specific and diagnostic-accuracy/treatment-planning literature, respectively; full strings are available from the corresponding author on request.

Study Selection

Records were exported to Zotero and de-duplicated, then screened in Rayyan by two independent reviewers against the eligibility criteria, with inter-reviewer agreement quantified using Cohen's κ. Disagreements at the title/abstract stage were resolved by discussion or, where consensus could not be reached, by promotion to full-text review; full-text disagreements were resolved by a third senior reviewer (the dissertation guide).

Data Extraction

A piloted, standardized Excel extraction sheet captured study identifier, country, design, sample size, AI platform, comparator, follow-up duration, primary and secondary outcome measures, principal numerical findings, funding/conflicts of interest, methodological limitations, and risk-of-bias judgment, extracted independently by two reviewers with disagreements resolved by consensus.

Risk of Bias Assessment

Diagnostic accuracy and agreement studies were appraised with QUADAS-2 (21), the randomized trial with the Cochrane RoB 2 tool (22), and all non-randomized comparative studies (cohort, case–control and controlled designs) with ROBINS-I (23). No included primary study was a cross-sectional survey, so no Joanna Briggs Institute checklist was required, and the included reviews, which were used for context only and did not contribute outcome data, were not formally appraised. An overall judgment per study is reported using the categories of each tool (RoB 2: low, some concerns, high; ROBINS-I: low, moderate, serious, critical; QUADAS-2: low, high, unclear), with the domain driving the judgment stated in Table 4; domain-level judgments are available on request (see Data Availability). Certainty of evidence for prioritized outcomes was rated using the GRADE approach (24) and is summarized in Table 6. Two reviewers performed all risk-of-bias appraisals independently, with disagreements resolved by consensus or third reviewer adjudication.

Data Synthesis

Given anticipated clinical and methodological heterogeneity, narrative synthesis organized by clinical theme was the pre-specified primary analytic approach. Quantitative meta-analysis was reserved for outcomes reported by at least three comparable design studies with extractable summary statistics and acceptable homogeneity (I^2 <50%, where calculable); no outcome met these criteria, and the synthesis therefore remained qualitative throughout. Only the sixteen primary comparative and diagnostic accuracy studies (11-13,15,20,25-35)(Table 1) were used as evidentiary units in this synthesis; the two included systematic reviews and the remaining scoping/narrative reviews were consulted for

background and cross-checking purposes only and did not independently contribute outcome data, avoiding double-counting of evidence already synthesized elsewhere. Because the PECO outcomes span several conceptually distinct domains—care-delivery efficiency (visit frequency, treatment duration), diagnostic/monitoring accuracy (appliance-emergency and hygiene-finding detection), and patient-centered and biological outcomes (compliance, satisfaction, oral hygiene, rapid-palatal-expansion accuracy)—findings are grouped and reported by these three domains rather than pooled into a single composite conclusion; this reflects a deliberate scoping review-style breadth appropriate to a first synthesis of an emerging evidence base, with the trade-off of a broader research question discussed explicitly as a limitation.

Results

Study Selection

Database searching, journal hand-searching, and reference-list screening identified 313 records across all sources. After removing 68 duplicate records, 245 records were screened at the title-and-abstract stage, of which 133 were excluded (non-orthodontic AI application, n=20; no remote-monitoring content, n=8; reviews/editorials without primary data, n=13; animal or in vitro studies, n=17; other reasons, n=75). A total of 112 full-text articles were assessed for eligibility, of which 82 were excluded (no conventional-care or reference-standard comparator group, n=4; non-orthodontic population and in-clinic rather than remote AI image analysis, n=1; insufficient outcome data, n=4; conference abstract only, n=1; duplicate dataset, n=2; other reasons, n=70). Thirty records met the final eligibility criteria and were included (Figure 1); one of these, McCray et al. (2026) (20), was published online after the close of the publication-date window and is retained as a declared deviation (see Methods). Inter-reviewer agreement at the title-and-abstract stage was substantial (Cohen's κ=0.84).

Table 1. Characteristics of the 16 primary comparative and diagnostic-accuracy studies (the evidentiary units of this review)

Author (Year)	Country	Design	n	AI Platform / Comparator	Principal Outcome
Hansa et al. (11)	UAE/USA	Retrospective cohort	90	DentalMonitoring® (45) vs. conventional (45)	Appointments ↓ 33.1%; time to first refinement ↓ 1.7 mo
Lam et al. (12)	Australia	RCT	56	DentalMonitoring® vs. conventional (clear aligners)	1.5 fewer visits; 1.9 mo longer treatment
Snider et al. (13)	USA	Diagnostic accuracy	24 (232 timepoints)	DentalMonitoring® AI vs. clinical exam (reference)	Sn 0.53 plaque, 0.35 gingivitis, 0.22 recession; Sp 0.94–0.99
Hansa et al. (15)	UAE/USA	Retrospective cohort	155	DentalMonitoring® (88) vs. conventional (67)	Visits 7.56 vs 9.82 (↓ 23%); satisfaction 4.25/5
McCray et al. (20)†	USA (multicentric)	Diagnostic accuracy	1,014	DentalMonitoring® AI vs. orthodontist consensus	Sn 98.4% debonding, 91.1% open clip, 93.3% tie loss
Martínez Gil-Ortega et al. (25)	Spain	Case-control (pediatric)	39	DentalMonitoring® (20) vs. conventional (19), Invisalign First	Visits ↓ significantly (p<0.001); 85% easy/very easy scan
Marks et al. (26)	Australia	Retrospective cohort	51	DentalMonitoring® (25) vs. conventional (26); PAR index	No significant difference in weighted PAR change
McCray et al. (27)	USA (multicentric)	Diagnostic accuracy (retrospective)	719	DentalMonitoring® AI vs. orthodontist consensus	High Sn/Sp for aligner attachment and button loss
Homsi et al. (28)	USA	Prospective in vivo	24	DentalMonitoring® AI 3D model vs. intraoral scanner	No clinically significant difference over 13.4 mo
Ferlito et al. (29)	USA	Prospective study	30 + 24	DentalMonitoring® AI vs. intraoral scan gauge reliability	Gauge compatibility 44.7%; 0% tooth-level agreement
Moylan et al. (30)	USA	Pilot study	30 paired sets	DentalMonitoring® smartphone app vs. plaster models	Intercanine Δ 0.17 mm; intermolar Δ −0.02 mm; equivalent
Kuriakose et al. (31)	USA	Prospective study	—	DentalMonitoring® RPE monitoring vs. in-office measurement	Clinically acceptable remote RPE measurement
Sangalli et al. (32)	Italy	Prospective study	30	DentalMonitoring® (15) vs. conventional (15)	Plaque ↓ at 3 mo; caries 0/15 vs. 5/15 (P=0.049)
Sharma et al. (33)	India	Controlled study	40	DentalMonitoring® (20) vs. conventional (20)	OPI, MGI, WSL ↓ significantly in AI group
Snider et al. (34)	USA	Prospective study	24	DentalMonitoring® AI vs. clinical exam (paired)	Plaque and gingivitis reduced over time with AI monitoring
Manimegalan et al. (35)	India	Retrospective cohort	—	AI-assisted DentalMonitoring® vs. conventional (clear aligners)	Fewer visits; satisfaction and compliance ↑

Sample size could not be extracted from the published report for two studies (Manimegalan et al. (35); Kuriakose et al. (31)), shown as "—" above; findings from these studies are flagged and interpreted cautiously wherever cited in the narrative synthesis. †Published online on 29 January 2026, after the close of the publication-date eligibility window (31 October 2025); retained as a declared deviation (Methods, Protocol Deviations item vi).

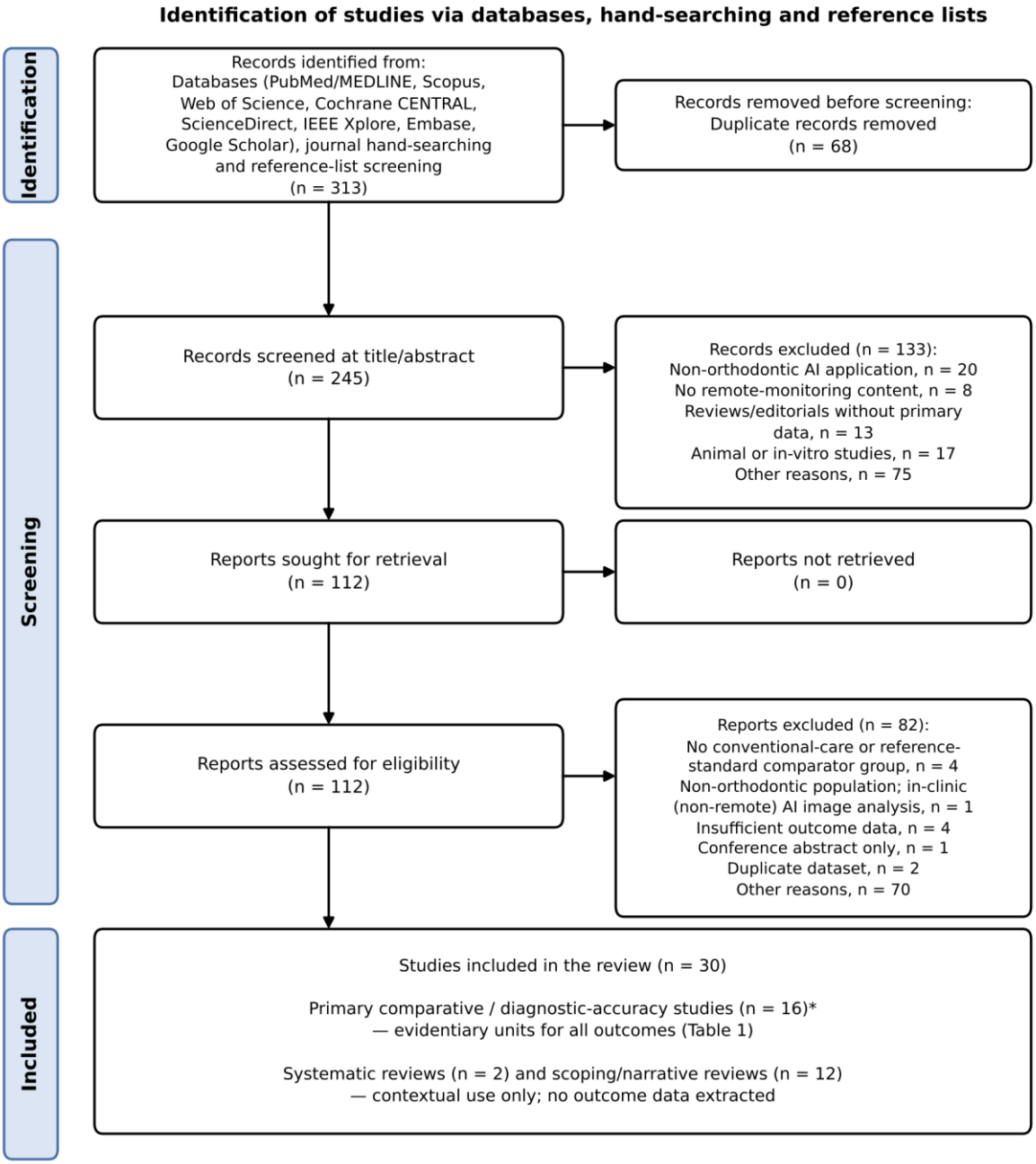


Figure 1. PRISMA 2020 flow diagram, with text summary.

The 30 included records comprised 16 primary comparative or diagnostic accuracy studies (Table 1: one randomized controlled trial, two case–control or controlled studies, four cohort studies, six prospective or pilot studies, and three diagnostic accuracy studies), two systematic reviews, and twelve scoping or narrative reviews that contributed contextual and discussion-level evidence only. Two reports that had been counted among the primary studies in an earlier version of this review—Thurzo et al. (2021) (14), in which an AI treatment-coaching application was evaluated without a concurrent conventional-care comparator, and Zhang et al. (2024) (36), an in-clinic evaluation of AI-assisted caries detection in a non-orthodontic population—did not meet eligibility criteria (i) and (iii) on re-examination and are excluded from the evidentiary set; Thurzo et al. is cited for background only.

Characteristics of Included Studies

The 16 primary studies were published between 2019 and 2026, with 11 (69%) published in 2023 or later, reflecting the recency of this evidence base. Eight originated from the United States, two from the

UAE/USA, two from Australia, two from India, and one each from Spain and Italy. Sample sizes in primary comparative studies ranged from 24 to 1,014 patients. DentalMonitoring® was the AI platform evaluated in the large majority of comparative and diagnostic accuracy studies, consistent with its position as the most widely deployed commercial system in clear aligner therapy, whose baseline efficacy for planned tooth movement was established in earlier prospective work (37). Table 1 summarizes the characteristics of the sixteen primary comparative and diagnostic accuracy studies that form the sole evidentiary unit for every outcome reported in Tables 2 to 6 and in the narrative synthesis below; the remaining fourteen included records (two systematic reviews and twelve scoping/narrative reviews) contributed background and discussion-level context only and were not treated as independent data points for any outcome, so no finding in this review double-counts evidence already synthesized in a prior review. Where a prior systematic or scoping review is cited alongside a primary study on the same topic, the citation illustrates convergent context, not an additional unit of quantitative evidence.

Appointment Reduction and Treatment Efficiency

Five studies reported the effect of AI-assisted remote monitoring on in-person visit frequency; all five reported reductions (with numerical estimates in three) in the AI-monitored group. Hansa et al. reported 7.56 versus 9.82 mean office visits (a 23% reduction, P<0.001) in a 155-patient retrospective cohort (15), with a companion 90-patient cohort reporting a 33.1% reduction and a 1.7-month reduction in time to first refinement (11). The single formally randomized trial (n=56) found 1.5 fewer visits with DentalMonitoring® (95% CI −3.3 to −0.7; P=0.02) but 1.9 months longer overall treatment duration (95% CI: 0.0–3.6; P=0.04), with no significant difference in refinements or total aligners (12). A pediatric case-control pilot (n=39) reported a significant reduction in visits (P<0.001), with 85% of children rating home scanning as “easy” or “very easy” (25). A retrospective cohort without an extractable sample size also reported fewer visits and is interpreted cautiously (39). End-of-treatment occlusal outcome, measured by the Peer Assessment Rating index in a 51-patient retrospective cohort, showed no statistically significant inter-group difference (P=0.181), supporting non-inferiority of AI-monitored treatment for occlusal change (26). Table 2 summarizes appointment- and treatment-outcome data across the six studies contributing these outcomes.

Table 2. Appointment frequency, treatment duration and occlusal outcomes in the primary studies

Study (year)	AI-group visits	Control visits	Reduction	Treatment duration	Occlusal outcome
Hansa et al. (11)	≈8.6	≈12.9	33%*	↓ 1.7 mo to 1st refinement	No difference in occlusal outcomes
Lam, Freer & Miles (12)	−1.5 vs control	Reference	Fewer visits*	+1.9 mo longer (P=0.04)	No difference in refinements/aligners
Hansa et al. (15)	7.56	9.82	23%*	No significant difference	No difference in refinements/aligners
Martínez Gil-Ortega et al. (25) (pediatric)	Significantly fewer*	Reference	Significant*	No significant difference	No difference in refinements/aligners
Marks et al. (26)	Not reported	Not reported	Not reported	Not primary outcome	PAR improved both groups; no inter-group diff. (P=0.181)
Manimegalan et al. (35)	Significantly fewer*	Reference	Not quantified	Not reported	High satisfaction, compliance

*Statistically significant reduction (P<0.05) reported by study authors. Manimegalan et al. (35): sample size not extractable; interpret with caution.

Table 3. Funding source and declared conflicts of interest of the 16 primary studies

Study	Funding source (as stated in the report)	Declared conflicts of interest/industry affiliation	Classification
Hansa et al. (11)	No funding statement is printed in the article.	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." Affiliations: European University College, Dubai; private practice, Sydney (treating orthodontist, V.K.). All patients were offered Dental Monitoring at the start of treatment (patient choice). No manufacturer support stated.	Not reported; no manufacturer involvement
Lam et al. (12)	"Funding: This research did not receive any grant from funding agencies." Trial registered ACTRN12620000475943; protocol published before commencement.	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." No manufacturer support stated. Affiliations: University of Queensland; private practice, Caloundra.	No specific funding; no manufacturer involvement
Snider et al. (13)	"Mohammed H. Elnagar was supported by funding from the American Association of Orthodontists Foundation, Robert M. Ricketts Sunflower Orthodontics Fellowship Award. The funder had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript."	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." Same eight authors as Snider et al. (2023), including M.H. Elnagar (see above).	Professional society (AAOF); senior author with a declared manufacturer relationship (elsewhere)
Hansa et al. (15)	"No funding was obtained." (Declarations)	"The authors declare that they have no competing interests." (Declarations). The practice offered dental monitoring to patients at no extra cost; the ScanBox is described as an optional auxiliary with an added cost to the orthodontist. No manufacturer support stated.	No funding; no manufacturer involvement
McCray et al. (20), Sci Rep, n=1,014	"This study was supported by DentalMonitoring." "The funder had no role in the design, data collection, analysis, interpretation of results, or writing of the manuscript." The dataset "was specifically compiled as part of an FDA regulatory submission;" expert readers were trained by DM's Clinical Affairs team.	"Dr. William Dabney reports receiving reimbursement for doctor time." "Dr. Rayan Skafi reports a DentalMonitoring share purchase." "Dr. Logan Smith reports receiving compensation for participation in data collection." "All other authors declare no competing interests." Skafi's affiliation: DentalMonitoring, Paris, France; "R.S. contributed to the writing of the manuscript."	Manufacturer-funded; manufacturer data and staff
Martínez Gil-Ortega et al. (25)	"This research was supported by the Institutional 16th call for research projects (granted number 1016016)." (Universidad Alfonso X El Sabio)	"The authors declare no competing interests." The article notes that Dental Monitoring and the ScanBox carry a monthly cost per patient; no manufacturer support stated.	Academic (university grant)
Marks et al. (26)	"This work was supported by the Australian Society of Orthodontists Foundation for Research and Education and the University of Queensland Research Support Fund."	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." Note: "The sample for this retrospective study was from a previous randomized controlled trial" (Lam et al. 2023; 51 of 56 randomized participants).	Professional society/university; no manufacturer involvement
McCray et al. (27), Semin Orthod, n = 719	Declaration of competing interest: "Julie Fahl McCray, William Dabney, Tanner Godfrey, Dylan Handlin, Logan Smith, Alexander Besmer, Ishita Johal, Rayan Skafi reports financial support was provided by Dental Monitoring." The study "was conducted in the context of a regulatory submission to the U.S. Food and Drug Administration (FDA);" cases "selected from DM's clinical database," with "system-generated clinical notifications" used for pre-screening; "DM's Clinical Affairs team provided standardized training to all participating sites and expert panel members." No separate funding section.	"Rayan Skafi reports a relationship with Dental Monitoring that includes employment. Mohammed H Elnagar reports a relationship with Dental Monitoring that includes speaking and lecture fees." Skafi's affiliation: "Clinical Education Manager, DentalMonitoring®, Paris, France."	Manufacturer-funded; manufacturer data and staff
Homsi et al. (28)	"This work was supported by the American Association of Orthodontists Foundation (Robert M. Ricketts Sunflower Orthodontics Fellowship Award; PI, Mohammed H. Elnagar). The funders had no role in the study design, data collection, analysis, the decision to publish, or the preparation of the manuscript." Methods: "The DM research department produced regenerated remote models from DM scans in STL format for this study."	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." Same University of Illinois Chicago team as Snider et al. (2023, 2024), including M.H. Elnagar.	Professional society (AAOF); manufacturer support in kind (regenerated models); senior author with a declared manufacturer relationship (elsewhere)
Ferlito et al. (29)	"This work was supported by the Northeastern Society of Orthodontists and the Moorrees and Lebret endowments. Furthermore, the Dental Monitoring units were donated to the orthodontic program by Dental Mind (Paris, France)."	"All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest, and none were reported." Affiliations: Harvard School of Dental Medicine, University of Minnesota, Howard University, private practice.	Professional society/endowment funding with manufacturer support in kind (donated units)
Moylan et al. (30)	"This study was supported in part by the VCU Alexander Fellowship, the Southern Association of Orthodontists, and the Virginia Orthodontic Education and Research Foundation." (Acknowledgments)	No conflict-of-interest statement is printed in the article; no relationship with Dental Monitoring is disclosed.	Academic/professional society
Kuriakose et al. (31)	"Funding: The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors." Acknowledgments: "The license for accessing the DM software and the tools (e.g., cheek retractors) needed for the study were provided by DM."	"Competing interests: Authors have completed and submitted the ICMJE Form for Disclosure of potential conflicts of interest. None declared." Acknowledgments: "None of the authors listed are affiliated with this company or have received any compensation (financial or otherwise) from DM. DM did not have any role in the conception or design of this study."	No grant; manufacturer support in kind (licenses and retractors)
Sangalli et al. (32)	"Dental Monitoring® partially sponsored the cost of the monthly monitoring." (Funding)	Competing interests section repeats: "Dental Monitoring® partially sponsored the cost of the monthly monitoring." No author affiliation with the manufacturer.	Manufacturer-supported (partial sponsorship of monitoring costs)
Sharma et al. (33)	Not stated: no financial support and sponsorship statement in the full text.	"There are no conflicts of interest." Ethics approval IEC No. DJD/IEC/2023/A-021. Platform: Dental Monitoring with ScanBox; no manufacturer support stated.	Not reported
Snider et al. (34)	Funding information: "American Association of Orthodontists Foundation." Acknowledgments: "MHE American Association of Orthodontists Foundation, Robert M. Ricketts Sunflower Orthodontics Fellowship Award. The funder had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript."	"None of the authors has a conflict of interest to disclose." Senior author M.H. Elnagar is the corresponding author of the three manufacturer-funded McCray et al. reports and, in the Semin Orthod report, "reports a relationship with Dental Monitoring that includes: speaking and lecture fees."	Professional society (AAOF); senior author with a declared manufacturer relationship (elsewhere)
Manimegalan et al. (35)	Not stated: the article has no financial support and sponsorship section.	"There are no conflicts of interest." The monitoring platform is not named ("AI-assisted apps").	Not reported

Table 4. Risk-of-bias judgments for all 16 primary comparative and diagnostic-accuracy studies

Study	Design/tool	Judgment	Principal concern
Hansa et al. (11)	Retrospective cohort/ROBINS-I	Moderate	Confounding by indication (patient choice of monitoring; non-randomized allocation)
Lam et al. (12)	RCT/RoB 2	Some concerns	No blinding of subjects or clinicians possible (deviations from intended interventions)
Snider et al. (13)	Diagnostic accuracy/QUADAS-2	Unclear	Clinical examination used as reference standard (reference-standard domain)
Hansa et al. (15)	Retrospective cohort/ROBINS-I	Moderate	Confounding by indication (non-randomized allocation)
Martínez Gil-Ortega et al. (25)	Case–control/ROBINS-I	Moderate	Small sample; non-blinded allocation
Marks et al. (26)	Retrospective cohort/ROBINS-I	Moderate	Retrospective analysis of 51 of 56 trial participants (selection of participants; missing data)
McCray et al. (20, 27)	Diagnostic accuracy/QUADAS-2	High	Patient-selection domain: evaluation datasets drawn from the manufacturer's own clinical data pool; manufacturer funding (2026 study) and manufacturer-affiliated co-author
Homsi et al. (28)	Prospective/QUADAS-2	Unclear	Single-center design (patient-selection domain); manufacturer regenerated the remote models in kind
Ferlito et al. (29)	Prospective/QUADAS-2	Unclear	Single-center sample (patient-selection domain); monitoring units donated by the manufacturer (in-kind support; no author conflicts declared)
Moylan et al. (30)	Pilot/QUADAS-2	Unclear	Small sample; single-center (patient-selection domain)
Kuriakose et al. (31)	Prospective agreement study/QUADAS-2	Unclear	Small single-center sample (patient-selection domain); software licence and tools provided by the manufacturer in kind
Sangalli et al. (32)	Prospective controlled study/ROBINS-I	Moderate	Non-randomized allocation; small sample; manufacturer partially sponsored monitoring costs
Sharma et al. (33)	Controlled study/ROBINS-I	Moderate	Small sample; allocation by clinician preference
Snider et al. (34)	Prospective agreement study/QUADAS-2	Unclear	Clinical examination used as reference standard; small single-center sample; no concurrent conventional-care group for the hygiene-change outcome
Manimegalan et al. (35)	Retrospective cohort/ROBINS-I	Serious	Confounding by indication; incomplete reporting of participant selection, sample size and outcome measurement

Overall judgment per study uses the categories of the tool applied (RoB 2: low, some concerns, high; ROBINS-I: low, moderate, serious, critical; QUADAS-2: low, high, unclear) and corresponds to the highest-risk domain; the domain driving the judgment is given in the principal-concern column. Domain-level judgments for every study are available from the corresponding author (Data Availability). Funding and conflict-of-interest information is reported in Table 3.

Diagnostic Accuracy for Appliance-Related Events

Two large diagnostic accuracy studies (n=1,014 and n=719) evaluated DentalMonitoring®'s AI against orthodontist-consensus reference standards for common fixed appliance and aligner emergencies, reporting sensitivities of 91.1–98.4% and specificities of 88.3–99.6% for bracket debonding, open self-ligating clips, tie loss, and aligner-attachment or button loss (20,27). Both evaluations drew images from the manufacturer's own clinical data pool and include a manufacturer-affiliated co-author (Table 3); the figures should therefore be read as sponsor-linked estimates awaiting independent replication. In contrast, a 24-patient (232 timepoint) diagnostic accuracy study of the same platform's oral hygiene module reported markedly lower sensitivity for plaque (0.53), gingivitis (0.35), and gingival recession (0.22) against clinical examination, despite high specificity (0.94–0.99) (13). An in vivo comparison of AI-derived and intraoral scanner three-dimensional models over a mean 13.4-month follow-up found no clinically significant difference between the two methods (28). A prospective evaluation of the AI's aligner GO/NO-GO instruction reported gauge compatibility of only 44.7% and 0% tooth-level agreement with clinician assessment, indicating limited repeatability of day-to-day tracking instructions (29). Remote smartphone-based monitoring of rapid palatal expansion produced intercanine and intermolar width measurements within 0.2 mm of direct plaster-model measurement (equivalence bounds: ±0.5 mm) (30,31). However, the sample size of one of these two studies could not be extracted from the published report (31), and this finding should accordingly be weighted cautiously pending independent confirmation.

Oral Hygiene, Compliance, and Patient Satisfaction

Despite modest diagnostic sensitivity for hygiene-related findings, several studies reported clinical improvement in plaque, gingival, or caries outcomes associated with AI-monitored care, including zero versus five carious lesions at 3–6 months in a 30-patient prospective study (P=0.049) (32) and significant reductions in plaque, gingival, and white-spot-lesion indices in a 40-patient Indian controlled study (33). Patient-reported satisfaction was generally favorable: mean satisfaction with DentalMonitoring® was 4.25/5, with 88% of patients preferring fewer office visits (15) and 85% of pediatric patients rating scanning as easy (26). Adverse events were considered as an outcome of interest in the PECO framework although they were not among the registered outcomes; no included study reported extractable adverse-event data (e.g., missed emergencies, delayed detection with a defined clinical consequence, or algorithm-attributable harm), and this absence of adverse-event reporting is treated in the Discussion and Future Research sections as a distinct evidence gap rather than as reassurance of safety.

Funding and Conflicts of Interest of Included Studies

Funding source and conflict-of-interest information were extracted for every included primary study, as specified in the data extraction sheet, and are tabulated in Table 3. The two largest diagnostic accuracy studies, which anchor the headline sensitivity and specificity figures for appliance emergency detection (n=1,014 and n=719) (20,27), are sponsor-linked. The n=1,014 study states that it was supported by DentalMonitoring, its dataset was extracted from the manufacturer's own database, and one co-author is affiliated with DentalMonitoring (Paris) and reports a DentalMonitoring share purchase (20). The n=719 study drew its sample from the manufacturer's U.S. clinical data pool and shared authorship with the first (27). Ferlito et al. reported funding from the North Eastern Society of Orthodontists and the Moorrees and Lebret endowments, stated that the DentalMonitoring units used in the study were donated to the orthodontic program by the manufacturer (Dental Mind, Paris, France), and declared no author conflicts of interest; the study is therefore classified as professional-society/endowment funded with manufacturer support in kind, and this in-kind support is reflected in its risk-of-bias judgment (Tables 3 and 4) (29). Martínez Gil-Ortega et al. (25) and Sharma et al.(33) declared no competing interests. Three further studies from one academic group (Snider et al. (13,34); Homsi et al. (28))were supported, according to the funder's published project reports, by an American Association of Orthodontists Foundation fellowship award; their senior author is also the senior author of the sponsor-funded n=1,014 study (20). Because the review's central diagnostic accuracy conclusion for fixed appliance and aligner-auxiliary emergencies rests substantially on the two sponsor-linked studies, this dependency is treated as a material risk-of-bias concern rather than a peripheral disclosure: it is reflected in the risk-of-bias judgment in Table 4, in the corresponding GRADE downgrade below, and in cautionary language in the Abstract, Discussion, and Conclusion sections.

Risk-of-bias judgments for all sixteen primary studies, each made with the tool appropriate to its design, are reported in Table 4. On QUADAS-2, the diagnostic accuracy and agreement studies were judged at unclear risk of bias, except the two sponsor-linked McCray studies, which were judged at high risk of bias in the patient-selection domain because their evaluation datasets were drawn from the manufacturer's own clinical data pool; the unclear judgments for the other studies related to the use of clinical examination as the reference standard for hygiene outcomes and to small, single-center samples, in three cases with equipment, software or models supplied by the manufacturer in kind. The single RCT was judged to have some concerns on RoB 2, principally because blinding of subjects and clinicians was not possible. On ROBINS-I, the retrospective cohorts and the non-randomized controlled studies were judged to be at moderate risk because of potential confounding by indication (allocation to AI monitoring by clinician or patient preference rather than randomization), except Manimegalan et al. (35), judged at serious risk because participant selection, sample size and outcome measurement were incompletely reported. GRADE certainty for the four prioritized outcomes is summarized in Table 6. Certainty was rated moderate for reduction in the number of in-office appointments with AI-assisted monitoring, a finding directionally consistent across all five studies that reported visit frequency, including the single RCT; certainty was downgraded one level from high because most contributing evidence came from retrospective cohorts at moderate risk of bias and blinding was not possible. By contrast, certainty for treatment duration was rated low (downgraded for serious inconsistency and risk of bias): the single RCT reported 1.9 months longer overall treatment with AI-assisted monitoring, whereas a retrospective cohort reported a 1.7-month reduction in time to first refinement. The designs and outcome definitions differ, but the effect directions conflict, so duration is rated as a separate, lower-certainty outcome rather than absorbed into the appointment-count judgment. Certainty for diagnostic accuracy of appliance emergency detection was rated low (downgraded two levels: once for indirectness relating to the scan-box-trained study populations, and a second time for serious risk of bias, since the two largest and most influential diagnostic accuracy studies underpinning this outcome were sponsor-linked (manufacturer funding and/or manufacturer-affiliated co-authorship)—see Funding and Conflicts of Interest above). Certainty was rated low for AI sensitivity in detecting plaque, gingivitis, and recession (downgraded for imprecision and indirectness). Table 5 summarizes major findings by clinical domain.

Table 5. Principal findings by clinical domain: AI-based remote monitoring versus conventional follow-up

Clinical domain	AI-based remote approach	Conventional approach
Bracket debonding detection	Sn 98.4%, Sp 99.6%	Patient self-report or detection at next visit
Open self-ligating clip/tie loss	Sn 91.1–93.3%, Sp 88.3–96.5%	Visual chairside inspection
Aligner attachment/button loss	High Sn and Sp	Visual chairside inspection
Aligner GO/NO-GO tracking decision	Gauge compatibility 44.7% (below industry standard)	Direct clinical evaluation by orthodontist
3D tooth movement tracking	No clinically significant difference from intraoral scanner	Intraoral scanner/plaster models
RPE intercanine/intermolar width	Mean Δ<0.2 mm vs plaster (equivalent, ±0.5 mm)	Direct plaster model measurement
Plaque detection	Sn 0.53, Sp 0.94, accuracy 0.60	Disclosing agent + Plaque Index scoring
Gingivitis detection	Sn 0.35, Sp 0.96, accuracy 0.49	Gingival Index / periodontal probing
Gingival recession detection	Sn 0.22, Sp 0.99, accuracy 0.72	Clinical examination with measurement
In-office appointment frequency (CAT)	Reduced 23–33% across cohorts and one RCT	Routine 6–10-week visits
Oral hygiene reinforcement	Improved PI, GI, WSL with AI notifications/app	Verbal reinforcement at scheduled visits
Patient satisfaction with monitoring	Mean 4.25/5; 88% prefer fewer visits	Variable; some preference for face-to-face

Table 6. GRADE summary of findings for the prioritized outcomes

Outcome	Contributing primary studies (design; participants)	Summary of effect	Downgrading	Certainty (GRADE)
Number of in-office appointments with AI-assisted monitoring vs. conventional follow-up	5 studies: 1 RCT (12), 3 retrospective cohorts (11,15,35) and 1 pediatric case–control pilot (25); n=39 to 155 where reported	Fewer visits in all five studies; 23–33% fewer visits in the cohorts; RCT: 1.5 fewer visits (95% CI −3.3 to −0.7)	Risk of bias (−1): most evidence from non-randomized studies at moderate risk; blinding not possible	Moderate
Treatment duration	2 studies: 1 RCT (12); 1 retrospective cohort (11)	RCT: 1.9 months longer overall treatment (95% CI 0.0–3.6); cohort: 1.7 months shorter time to first refinement	Risk of bias (−1); inconsistency (−1): conflicting effect directions and differing outcome definitions	Low
Diagnostic accuracy for fixed-appliance and aligner-auxiliary emergencies	2 diagnostic accuracy studies (20,27); n=1,014 and n=719	Sensitivity 91.1–98.4%; specificity 88.3–99.6% vs. orthodontist consensus	Indirectness (−1): scan-box-trained populations; risk of bias (−1): sponsor-linked studies at high risk in the patient-selection domain	Low
AI sensitivity for plaque, gingivitis and gingival recession	1 diagnostic accuracy study (13); n=24 (232 timepoints)	Sensitivity 0.53 (plaque), 0.35 (gingivitis), 0.22 (recession); specificity 0.94–0.99 vs. clinical examination	Imprecision (−1): single small study; indirectness (−1): single platform and setting	Low

GRADE, Grading of Recommendations Assessment, Development and Evaluation (24). Certainty was downgraded one level (−1) from a starting level of high for each serious concern; no outcome was upgraded. The same ratings and reasons are stated in the text (Funding and Conflicts of Interest of Included Studies). Study numbers refer to the reference list.

Discussion

The sixteen primary studies synthesized in this review (supplemented by 14 reviews used for context only) span highly heterogeneous designs—randomized trials, case–control and controlled studies, prospective and retrospective cohorts, and diagnostic-accuracy studies—yet point broadly in a consistent direction. AI-driven remote monitoring, dominated in the orthodontic literature by the DentalMonitoring® platform, reduces in-person appointments in clear aligner therapy by approximately one-quarter to one-third without compromising end-of-treatment occlusal outcomes, achieves high diagnostic accuracy for discrete mechanical incidents in fixed appliance and aligner therapy (on evidence that is substantially sponsor-linked); reproduces linear rapid-palatal-expansion measurements within clinically acceptable tolerances, and improves oral hygiene engagement through behavioral prompts even where its underlying diagnostic sensitivity for plaque, gingivitis, and recession is limited. These findings are broadly consistent with, and extend, the pooled directional conclusions of prior systematic and scoping reviews of teleorthodontics (4,5), while adding more granular diagnostic accuracy data than either prior synthesis.

Is AI-based remote monitoring more accurate than conventional examination? The answer is conditional, not categorical. For well-bounded, visually discrete diagnostic tasks—bracket debonding, open self-ligating clips, tie loss, attachment and button loss—AI achieved sensitivities of 91–98% and specificities of 88–99% against orthodontist-consensus reference standards (20,27), performance comparable to accepted screening benchmarks in radiology and dermatology. Because these events are otherwise identified only when the patient reports discomfort or attends the next scheduled visit, AI monitoring could represent a genuine absolute gain in early detection. However, these accuracy estimates rely heavily on two sponsor-linked studies that used the manufacturer's own data pool and await independent replication. For diffuse, multi-cue findings such as plaque, gingivitis, and gingival recession, however,

sensitivity fell to 0.22–0.53 (13), figures that would not justify replacing periodontal probing. AI in its current orthodontic application is therefore better characterized as accuracy-extending rather than accuracy-replacing: it widens the clinician's diagnostic reach between visits without matching the integrative judgment of direct examination.

Does AI-assisted monitoring translate operational efficiency into biological efficiency? The appointment-reduction signal is remarkably consistent—23–33% fewer visits across retrospective cohorts with differing sample sizes and geographies, and a statistically significant reduction in the sole RCT (12) and in a pediatric case-control pilot (25). Whether this operational gain shortens overall treatment is less certain: one cohort reported a 1.7-month reduction in time to first refinement (11), while the RCT reported 1.9 months longer overall treatment (12), and the PAR-index analysis found no significant inter-group difference in occlusal change (26). The most defensible synthesis is that AI-assisted remote monitoring reliably improves the efficiency of clinical contact without reliably shortening—and in some configurations, potentially extending—total treatment duration.

The oral hygiene findings illustrate an important dissociation between diagnostic accuracy and clinical effect. Despite the AI hygiene module's modest sensitivity (13), several controlled studies documented measurable improvements in plaque, gingival, and caries indices among AI-monitored patients (32,33), and a hygiene notification study reported behavioral improvement independent of the underlying algorithm's modest diagnostic sensitivity (34). The most plausible mechanism is behavioral rather than diagnostic: the requirement for regular image capture, paired with structured prompts, exerts an engagement effect analogous to the Hawthorne effect described in other areas of clinical research, rather than reflecting superior algorithmic detection of periodontal disease.

The reliability of AI-generated day-to-day aligner instructions is a further area requiring caution. While overall tooth movement accuracy in AI-monitored patients was comparable to conventionally monitored controls (11), a prospective evaluation reported gauge compatibility of only 44.7% and 0% tooth-level agreement for the platform's GO/NO-GO tracking instruction (29). Outcome-level equivalence and instruction-level repeatability are not contradictory, but together they indicate that AI output should function as an alert requiring clinician confirmation of the underlying image evidence, rather than as a stand-alone decision rule.

Several cross-cutting limitations temper enthusiasm for wholesale adoption. First, the leading AI platforms remain proprietary; algorithm architecture and training-data composition are not disclosed in the academic literature (5,38), which has implications for informed consent, clinical accountability, and the allocation of medico-legal responsibility when an algorithm misses a clinical event. Second, diagnostic accuracy figures were achieved on images from patients trained in scan box use; in unselected real-world populations with variable smartphone proficiency, lighting, and positioning, accuracy is likely to be modestly lower (29). Third, several diagnostic accuracy studies are linked, by authorship or by use of the platform's clinical data pool, to the technology under evaluation, limiting the independence of the evidence; this applies most notably to the two largest studies (n=1,014 and n=719), one of which was funded by the manufacturer (Table 3). Fourth, data privacy remains only partially resolved: patient images are typically uploaded to cloud servers, and regulatory frameworks for such health data are still maturing relative to established frameworks such as the EU General Data Protection Regulation (39).

For orthodontic practice in India and other geographically dispersed health systems, an access-related case for AI-assisted remote monitoring is plausible: orthodontists are concentrated in urban centers, and a 23–33% reduction in routine visits could, in principle, reduce the travel and opportunity cost borne by patients and families. This argument should be advanced cautiously rather than as an established finding; however, it rests on only two India-based studies in the whole synthesis, one with a modest sample (n=40) (33) and one for which the published report did not permit extraction of a sample size (35). Neither study was designed to measure access or travel-burden outcomes directly. Digital divide considerations remain real, and hybrid models pairing mid-level providers with AI triage under orthodontist supervision may be more contextually appropriate than direct patient-captured imaging for rural or resource-limited settings (40). The most acute periodontal presentations continue to require direct physical examination and palpation that cannot be delivered remotely, a boundary consistent with the wider AI remote-monitoring evidence base (41). Dedicated, adequately powered studies from India and other low-resource settings are needed before this access-related rationale can be regarded as evidence-based rather than plausible extrapolation.

Taken together, the evidence supports neither uncritical adoption nor reflexive rejection of AI-assisted remote monitoring. It supports structured, supervised, and transparent integration: AI as an adjunct that extends diagnostic reach and reduces routine visit burden, operating alongside—not in place of—periodic in-person clinical examination.

Clinical Implications

For practicing orthodontists, the evidence supports several practical positions. AI-assisted remote monitoring can reasonably be offered as an

adjunct during clear aligner therapy to reduce routine visit frequency, particularly for patients with demonstrated engagement and reliable smartphone access. AI-flagged mechanical incidents (bracket debonding, tie loss, attachment loss) can reasonably prompt clinician review, given the high reported accuracy. However, this estimate relies heavily on sponsor-linked studies, and its certainty is low pending independent validation. In contrast, AI-flagged hygiene findings should be treated as behavioral prompts rather than definitive periodontal diagnoses, given their modest sensitivity. A minimum interval of in-person review should be maintained irrespective of remote-monitoring status, since growth assessment, temporomandibular evaluation, and integrative clinical judgment are not remotely reproducible. Explicit, granular informed consent—covering platform identity, data-hosting jurisdiction, retention period, and secondary data use—should be obtained before enrolment, and the clinician's role in confirming AI-flagged events should be documented in the patient record.

Strengths

The principal strengths of this review are its structured PRISMA 2020 methodology, layered Boolean search strategy across multiple databases, independent dual-reviewer screening with quantified inter-rater agreement, and application of design-appropriate risk-of-bias tools (QUADAS-2, RoB 2, and ROBINS-I) with GRADE-based certainty rating. The orthodontic focus was maintained consistently throughout, and the synthesis integrated both comparative outcome and diagnostic accuracy evidence, providing a more granular clinical picture than prior narrative reviews of teleorthodontics. The distinction between primary evidentiary studies and contextual reviews was applied consistently, avoiding double-counting, and funding and conflict-of-interest data for each primary study are reported explicitly (Table 3).

Limitations

Quantitative pooling was not undertaken because of clinical and methodological heterogeneity across the included studies; pooled estimates, while potentially more compelling, would have required a degree of homogeneity the evidence base does not yet possess. The review question was deliberately broad, spanning care delivery efficiency, diagnostic accuracy, and patient-centered/biological outcomes across several distinct clinical tasks; this breadth suited a first, mapping-style synthesis of an emerging field but means the review does not offer the depth of a narrowly focused synthesis on any single outcome, and future reviews restricted to one outcome domain (for example, diagnostic accuracy alone) are recommended. Two of the sixteen primary studies did not report an extractable sample size, and their findings have been flagged accordingly throughout rather than weighted equally with better-reported studies. The protocol was registered with PROSPERO (registration number CRD420261451606) on 19 August 2026, after searching and screening had begun; registration was therefore retrospective for those stages, which limits the assurance that eligibility criteria, outcomes and methods were fixed before the evidence was seen, and differences from the registered record are reported in Section 2. The review's central diagnostic accuracy conclusion depends substantially on two large studies that are sponsor-linked: one was funded by the manufacturer of the evaluated platform (DentalMonitoring®), and both were co-authored by a manufacturer-affiliated investigator; this is treated as a material risk-of-bias concern, reflected in downgraded risk-of-bias and GRADE judgments, rather than as a minor caveat, and complete independence of these findings cannot be guaranteed despite formal appraisal. Finally, this single commercial platform still dominates the evidence base, limiting generalizability to other AI-based remote-monitoring systems. The publication-date eligibility window closed on 31 October 2025; because the searches were run in 2026, one study published online on 29 January 2026 (McCray et al.) (20) was retrieved and retained as a declared deviation from that window (Protocol Deviations, item vi), so the evidentiary set is not strictly bounded by the stated window, and other reports published after 31 October 2025 were not sought; Google Scholar was screened only to its first 200 relevance-ranked hits so that some grey literature may have been missed. No included study reported adverse events, so the safety of AI-assisted monitoring could not be assessed.

This heterogeneity also limits how far the findings can be generalized. The included studies differed in design (one randomized trial alongside case–control, cohort, pilot, and diagnostic accuracy studies), in comparator (conventional in-office monitoring, orthodontist consensus, plaster models or intraoral scans), and in outcome definitions and clinical tasks, with sample sizes ranging from 24 to 1,014 and most studies conducted in a small number of high-income practice settings, with limited data from low-resource and rural settings. The reported reductions in appointment frequency and the sensitivity and specificity estimates should therefore be read as context-specific findings for a particular platform, task and reference standard, not as effects that transfer automatically to other patient groups, appliance types, AI systems or health system settings. Because no common effect measure could be pooled, the extent of inter-study variability could not be quantified, and the direction and approximate magnitude of the findings should be interpreted cautiously until more uniformly designed, independent studies are available.

Future Research

Priorities for future research include independent, multi-center randomized controlled trials comparing AI-assisted remote monitoring with conventional care across diverse practice settings, using prospectively registered protocols and a standardized core outcome set; independent diagnostic accuracy studies using public or institutional datasets, reported according to QUADAS-2 and TRIPOD-AI standards; long-term outcome studies extending beyond two years to capture retention stability and relapse; cost-effectiveness analyses from clinician, patient, and health system perspectives; validation studies in under-represented populations, including Indian government and rural practice settings; and implementation research on hybrid teledentistry models combining mid-level providers with AI triage for vulnerable and underserved populations.

Conclusion

This systematic review synthesized 16 primary studies, contextualized by 14 reviews, examining AI-assisted remote monitoring in orthodontic care. The evidence indicates that AI-based remote monitoring is not categorically superior or inferior to conventional in-office follow-up, but performs well on specific, well-bounded tasks while still relying on it for others. AI-assisted platforms, principally DentalMonitoring®, reduce routine in-person appointments by approximately one-quarter to one-third without compromising end-of-treatment occlusal outcomes, detect discrete fixed appliance and aligner auxiliary incidents with high sensitivity and specificity (in sponsor-linked evaluations awaiting independent replication), reproduce rapid palatal expansion measurements accurately, and improve oral hygiene engagement through behavioral reinforcement. Conversely, diagnostic sensitivity for plaque, gingivitis, and recession remains low, day-to-day aligner tracking instructions show limited repeatability, and unresolved questions persist regarding algorithmic transparency, data privacy, and medico-legal responsibility. On current evidence, AI-assisted remote monitoring should be regarded as a supervised adjunct to, rather than a substitute for, periodic in-person clinical examination. Multi-center randomized trials with standardized outcome reporting, independent diagnostic validation, and profession-led guidance on consent and data governance are needed before AI-assisted remote monitoring can be recommended as a routine substitute for conventional orthodontic follow-up.

Authors' Contributions

Aneri Ashokkumar Patel (AAP): conceptualization, methodology, investigation, data curation, formal analysis, visualization, writing—original draft

Khyati Viral Patel (KVP): investigation, validation, writing—review and editing

Alpesh Manilal Vaghela (AMV): conceptualization, methodology, validation, supervision, writing—review and editing

All authors read and approved the final manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors received no funding for this work.

Conflicts of Interest

The authors declare no conflicts of interest with respect to the research, authorship, and/or publication of this article. This statement concerns the authors of this review; conflicts of interest identified for individual included studies are reported in Table 3.

Data Availability

The data extraction sheets, domain-level risk-of-bias assessments for every primary study, and full search strategies supporting this review's findings are available from the corresponding author upon reasonable request.

AI/Generative AI Disclosure Statement

The authors used an Anthropic Claude large language model (Claude Sonnet, Anthropic PBC) to organize literature-search results, draft and revise text in response to peer-review comments, and verify reference details against primary sources. All AI-assisted content was reviewed, fact-checked, and edited by the authors, who reviewed and verified it and take full responsibility for the accuracy, integrity, and scientific validity of the manuscript. The AI tool is not listed as an author.

References

1. Estai M, Kanagasingam Y, Tennant M, Bunt S. A systematic review of the research evidence for the benefits of teledentistry. J Telemed Telecare. 2018;24(3):147-156. doi: 10.1177/1357633X16689433.
2. Thalib B, Nurrahma R, Asmawati, Thalib AM, Taufik MNF. The use of teledentistry and Artificial Intelligence (AI) in dental and oral health services for the elderly in the era of the COVID-19 pandemic: a systematic review. J Dentomaxillofac Sci. 2022;7(3):135–40. doi:10.15562/jdmfs.v7i3.1406
3. Kaushik R, Rapaka R. AI-driven evolution in teledentistry: a comprehensive overview of

technology and clinical applications. Dent Rev. 2025,5(2):1-12.doi: 10.1016/j.dentre.2025.100154.
4. Torres DKB, Santos MCCD, Normando D. Is teledentistry effective to monitor the evolution of orthodontic treatment? A systematic review and meta-analysis. Dental Press J Orthod. 2023;28(4):e2322195. doi: 10.1590/2177-6709.28.4.e2322195.oar.
5. Polizzi A, Serra S, Leonardi R, Isola G. Clinical applications of artificial intelligence in Teleorthodontics: A scoping review. Medicina (Kaunas). 2025;61(7):1141. doi: 10.3390/medicina61071141.
6. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44. doi: 10.1038/nature14539.
7. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM. 2017;60(6):84–90. doi:10.1145/3065386
8. Schwendicke F, Samek W, Krois J. Artificial intelligence in dentistry: chances and challenges. J Dent Res. 2020;99(7):769-774. doi: 10.1177/0022034520915714.
9. Casalegno F, Newton T, Daher R, Abdelaziz M, Lodi-Rizzini A, Schürmann F, et al. Caries detection with near-infrared transillumination using deep learning. J Dent Res. 2019;98(11):1227-1233. doi: 10.1177/0022034519871884.
10. Bichu YM, Hansa I, Bichu AY, Premjani P, Flores-Mir C, Vaid NR. Applications of artificial intelligence and machine learning in orthodontics: a scoping review. Prog Orthod. 2021;22(1):18. doi: 10.1186/s40510-021-00361-9.
11. Hansa I, Katyal V, Ferguson DJ, Vaid N. Outcomes of clear aligner treatment with and without dental monitoring: A retrospective cohort study. Am J Orthod Dentofacial Orthop. 2021;159(4):453-459. doi: 10.1016/j.ajodo.2020.02.010.
12. Lam J, Freer E, Miles P. Comparative assessment of treatment efficiency and patient experience between dental monitoring and conventional monitoring of clear aligner therapy: A single-center randomized controlled trial. Am J Orthod Dentofacial Orthop. 2023;163(4):456-464. doi: 10.1016/j.ajodo.2022.12.004.
13. Snider V, Homsi K, Kusnoto B, Atsawasuwan P, Viana G, Allareddy V, et al. Clinical evaluation of Artificial Intelligence Driven Remote Monitoring technology for assessment of patient oral hygiene during orthodontic treatment. Am J Orthod Dentofacial Orthop. 2024;165(5):586-592. doi: 10.1016/j.ajodo.2023.12.008.
14. Thurzo A, Kurilová V, Varga I. Artificial intelligence in orthodontic smart application for treatment coaching and its impact on clinical performance of patients monitored with AI-TeleHealth system. Healthcare (Basel). 2021;9(12):1695. doi: 10.3390/healthcare9121695.
15. Hansa I, Semaan SJ, Vaid NR. Clinical outcomes and patient perspectives of Dental Monitoring® GoLive® with Invisalign®-a retrospective cohort study. Prog Orthod. 2020;21(1):16. doi: 10.1186/s40510-020-00316-6.
16. Schwendicke F, Krois J. Better reporting of studies on artificial intelligence: CONSORT-AI and beyond. J Dent Res. 2021;100(7):677-680. doi: 10.1177/0022034521998337.
17. Nordblom NF, Büttner M, Schwendicke F. Artificial intelligence in orthodontics: Critical review. J Dent Res. 2024;103(6):577-584. doi: 10.1177/00220345241235606.
18. Vashisht R, Sharma A, Kiran T, Jolly SS, Brar PK, Puri JV. Artificial intelligence in dentistry – a scoping review. J Oral Maxillofac Surg Med Pathol. 2024;36:579–592 doi:10.1016/j.ajoms.2024.04.009.
19. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71.
20. McCray JF, Smith L, Handlin D, Dabney W, Mietz C, Skafi R, et al. Accuracy of Dental Monitoring's artificial intelligence in detecting common orthodontic braces treatment related emergencies. Sci Rep. 2026;16(1):6627. doi: 10.1038/s41598-026-37329-w.
21. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36. doi: 10.7326/0003-4819-155-8-201110180-00009.
22. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. doi: 10.1136/bmj.l4898.
23. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi: 10.1136/bmj.i4919.
24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6. doi: 10.1136/bmj.39489.470347.AD.
25. Martínez Gil-Ortega A, Cintora-López P, Pérez Rodríguez LM, Viñas MJ, Aragoneses JM, Arrieta-Blanco P, et al. Pilot case control evaluation of artificial intelligence assisted orthodontic monitoring and pediatric patient perception. Sci Rep. 2025;15(1):40515. doi: 10.1038/s41598-025-24395-9.
26. Marks J, Freer E, Ong D, Lam J, Miles P. Evaluating Dental Monitoring effectiveness compared with conventional monitoring of clear aligner therapy using the Peer Assessment Rating index. Am J Orthod Dentofacial Orthop. 2024;166(4):350-355. doi: 10.1016/j.ajodo.2024.05.014.
27. McCray JF, Dabney W, Godfrey T, Handlin D, Smith L, Besmer A, et al. Accuracy of DentalMonitoring's artificial intelligence in detecting the loss of aligners' auxiliaries: a retrospective multi-centric study. Semin Orthod. 2025; 32(4): 1-7. doi:10.1053/j.sodo.2025.09.005.
28. Homsi K, Snider V, Kusnoto B, Atsawasuwan P, Viana G, Allareddy V, et al. In vivo evaluation of Artificial Intelligence Driven Remote Monitoring technology for tracking tooth movement and reconstruction of 3-dimensional digital models during orthodontic treatment. Am J Orthod Dentofacial Orthop. 2023;164(5):690-699. doi: 10.1016/j.ajodo.2023.04.019.

29. Ferlito T, Hsiou D, Hargett K, Herzog C, Bachour P, Katebi N, et al. Assessment of artificial intelligence-based remote monitoring of clear aligner therapy: A prospective study. Am J Orthod Dentofacial Orthop. 2023;164(2):194-200. doi: 10.1016/j.ajodo.2022.11.020.
30. Moylan HB, Carrico CK, Lindauer SJ, Tüfekçi E. Accuracy of a smartphone-based orthodontic treatment-monitoring application: A pilot study. Angle Orthod. 2019;89(5):727-733. doi: 10.2319/100218-710.1.
31. Kuriakose P, Greenlee GM, Heaton LJ, Khosravi R, Tressel W, Bollen AM. The assessment of rapid palatal expansion using a remote monitoring software. J World Fed Orthod. 2019;8(4):165–70. doi: 10.1016/j.ejwf.2019.11.001.
32. Sangalli L, Savoldi F, Dalessandri D, Bonetti S, Gu M, Signoroni A, et al. Effects of remote digital monitoring on oral hygiene of orthodontic patients: a prospective study. BMC Oral Health. 2021;21(1):435. doi: 10.1186/s12903-021-01793-9.
33. Sharma A, Sharma T, Tomer G, Pattanaik S, Khattar B, Shrinivaasan NR, et al. Assessment of AI-based monitoring method in assessing oral hygiene during orthodontic treatment. J Orthod Sci. 2025;14:17. doi: 10.4103/jos.jos_146_24.
34. Snider V, Homsi K, Kusnoto B, Atsawasuwan P, Viana G, Allareddy V, et al. Effectiveness of AI-driven remote monitoring technology in improving oral hygiene during orthodontic treatment. Orthod Craniofac Res. 2023 Dec;26 Suppl 1:102-110. doi: 10.1111/ocr.12666.
35. Manimegalan P, Pragnya B, Jain M, Tomy M, Chinnappa V, Ashwathi N. Digital orthodontic treatment monitoring and remote aligner therapy: a paradigm change in contemporary orthodontics. J Pharm Bioallied Sci. 2025;17(Suppl 3):S2623-S2625. doi: 10.4103/jpbs.jpbs_788_25.
36. Zhang JW, Fan J, Zhao FB, Ma B, Shen XQ, Geng YM. Diagnostic accuracy of artificial intelligence-assisted caries detection: a clinical evaluation. BMC Oral Health. 2024 Sep 16;24(1):1095. doi: 10.1186/s12903-024-04847-w.
37. Kravitz ND, Kusnoto B, BeGole E, Obrez A, Agran B. How well does Invisalign work? A prospective clinical study evaluating the efficacy of tooth movement with Invisalign. Am J Orthod Dentofacial Orthop. 2009;135(1):27-35. doi: 10.1016/j.ajodo.2007.05.018.
38. Mazin H. Artificial intelligence in dentistry: advancements in periodontology and other specialties, diagnosis, treatment planning, and ethical considerations. Dent Rev. 2025, 5(2):100157. doi: 10.1016/j.dentre.2025.100157.
39. Shaik T, Tao X, Higgins N, Li L, Gururajan R, Zhou X, et al. Remote patient monitoring using artificial intelligence: current state, applications, and challenges. WIREs Data Min Knowl Discov. 2023;13(2):e1485. doi: 10.1002/widm.1485.
40. Drafta S, Macris A, Petre AE. Innovations on the horizon: teledentistry, artificial intelligence, and hybrid models to improve oral health for vulnerable communities. Front Oral Health. 2025;6:1649715. doi: 10.3389/froh.2025.1649715.
41. Peacock ME, Ghaly M, Timothius CJC. Emergent periodontal conditions: an overview. Dent Rev. 2025;5(2):100155. doi:10.1016/j.dentre.2025.100155.

Appendix

Identification: records identified through database searching, journal hand-searching, and reference-list screening, n=313. Screening: 68 duplicate records removed; 245 records screened at title/abstract; 133 excluded (non-orthodontic AI application, n=20; no remote-monitoring content, n=8; reviews/editorials without primary data, n=13; animal/in-vitro studies, n=17; other reasons, n=75). Eligibility: 112 full-text articles assessed; 82 excluded (no conventional-care or reference-standard comparator group, n=4; non-orthodontic population and in-clinic rather than remote AI image analysis, n=1; insufficient outcome data, n=4; conference abstract only, n=1; duplicate dataset, n=2; other reasons, n=70). Included: 30 records (16 primary comparative/diagnostic-accuracy studies, one of which was published online after the close of the publication-date window and is retained as a declared deviation; 2 systematic reviews; 12 scoping/narrative reviews), of which only the 16 primary studies served as evidentiary units.