Atlantic J Med Sci Res 2026;6(2):195-205
DOI: 10.5455/atjmed.2025.12.050

THE ATLANTIC Journal of Medical Science and Research

Original Article

Comparative evaluation of orthodontic appliances in adolescents: A PRISMA-Compliant systematic review and network meta-analysis

Ergin Kalkan[1], Hasan Catalkaya[2]

[1]Dent Kastamonu Oral and Dental Health Polyclinic, Department of Orthodontics, Kastamonu, Türkiye
[2]Dent Atalay Oral and Dental Health Polyclinic, Department of Orthodontics, Konya, Türkiye

Received: 30 December, 2025
Accepted: 22 February, 2026
Published: 08 June, 2026

Corresponding Author: Ergin Kalkan, Dent Kastamonu Oral and Dental Health Polyclinic, Department of Orthodontics, Kastamonu, Türkiye
Email: kalkanergin@gmail.com

CITATION
Kalkan E, Catalkaya H. Comparative evaluation of orthodontic appliances in adolescents: A PRISMA-Compliant systematic review and network meta-analysis. Atlantic J Med Sci Res. 2026;6(2):195-205. DOI:10.5455/atjmed.2025.12.050

Abstract

Aim: Orthodontic appliance selection in adolescence has an impact on skeletal correction, dental control, periodontal health, and patient-reported outcomes. Clear aligners (CAs), fixed appliances (FAs), and functional appliances, including removable (RFAs) and fixed functional appliances (FFAs), are used. However, comparative evidence evaluating these appliances across multiple clinical domains in adolescents remains limited.
Materials and Methods: A systematic search of PubMed, Embase, Web of Science, and the Cochrane Library was conducted for studies published between January 2000 and March 2025. Randomized controlled trials and prospective observational studies involving adolescents aged 8–18 years and comparing at least two orthodontic appliance types were included. Primary outcomes were skeletal changes, ANB angle reduction, periodontal health indices, and patient-reported outcomes satisfaction and pain. Secondary outcomes included treatment duration, treatment precision, and adverse effects. Pairwise meta-analyses were performed when only direct comparisons were available, while network meta-analyses were conducted for outcomes forming connected evidence networks. Risk of bias was assessed using RoB 2 and ROBINS-I. The review protocol was registered in PROSPERO and followed PRISMA-NMA guidelines.
Results: Nineteen studies involving 2.569 adolescents were included. Network meta-analysis demonstrated that FFAs produced the greatest skeletal effects, showing the largest reduction in ANB angle. CAs were associated with higher patient satisfaction and less initial pain compared with FAs. Periodontal outcomes favored CAs over FAs; however, these findings were derived exclusively from direct pairwise comparisons and were not supported by a connected evidence network, limiting comparisons with functional appliances. No significant difference in treatment duration was observed between CAs and FAs. FAs demonstrated superior control of complex tooth movements but were associated with a higher risk of white spot lesions, while RFA effectiveness depended on patient compliance.
Conclusion: Orthodontic appliance selection in adolescents should be individualized, balancing skeletal objectives, biomechanical demands, periodontal risk and patient-reported outcomes.

Keywords: Clear aligners, fixed appliances, functional appliances, adolescent orthodontics, network meta-analysis

INTRODUCTION

Adolescence represents a critical period for orthodontic intervention due to the presence of residual craniofacial growth, which allows for the effective modification of skeletal discrepancies alongside dental alignment (1). The choice of orthodontic appliance is a fundamental clinical decision that directly influences treatment efficacy, efficiency, patient experience, and long-term oral health (2).

For decades, conventional fixed appliances (FAs) have been considered the gold standard, offering unparalleled three-dimensional control over tooth movement for managing complex malocclusions (3). The therapeutic landscape has evolved significantly with the advent and refinement of alternative modalities. Removable functional appliances (RFAs), such as the Twin Block, and fixed functional appliances (FFAs), like the Herbst or Forsus, are specifically designed to correct Class II malocclusions by harnessing or redirecting mandibular growth (4,5). More recently, clear aligners (CAs) therapy has emerged as a transformative technology, gaining immense popularity among patients and clinicians due to its aesthetic appeal, removability, and perceived comfort (6,7).

While each appliance class possesses distinct biomechanical principles and clinical indications, a comprehensive, direct comparison of their multidimensional outcomes in the adolescent population remains inadequately synthesized. Existing literature often focuses on pairwise comparisons (e.g., CA vs. FA) within isolated domains, such as treatment duration (8) or periodontal health (9). Several systematic reviews have addressed specific appliance types or outcomes (10,11), but a holistic analysis that simultaneously evaluates skeletal impact, dental precision, periodontal effects, patient-reported outcomes (PROs), and adverse event profiles across all major appliance categories is lacking. This gap in evidence creates a challenge for evidence-based clinical decision-making. Orthodontists must weigh the potential for skeletal correction against the demands for precise tooth movement, the imperative of maintaining periodontal health, and the importance of patient adherence and satisfaction (12). Furthermore, the relative ranking of these appliances across different outcome measures is unclear.

Therefore, the primary objective of this systematic review and network meta-analysis (NMA) is to integrate direct and indirect evidence from randomized controlled trials (RCTs) and high-quality prospective studies to provide a hierarchical comparison of CAs, FAs, RFAs, and FFAs in adolescent patients. We aim to evaluate their comparative effectiveness across the following key domains: 1) skeletal and dentoalveolar changes, 2) periodontal health indicators, 3) treatment efficiency and precision, 4) patient-centered outcomes (satisfaction, pain, oral health-related quality of life), and 5) incidence of adverse effects. By employing NMA methodology, this study will generate a ranked order of treatment preferences for each outcome, offering clinicians a nuanced, evidence-based framework to guide appliance selection in tandem with individual patient characteristics and treatment goals. Given the biological and clinical heterogeneity among appliance indications—particularly the growth-dependency of functional appliances—treatment rankings will be interpreted cautiously, with explicit consideration of baseline malocclusion characteristics and growth status.

Given the rapid technological advancements in CA materials and digital treatment planning, this review also aims to explore whether more recent studies (post-2015) reflect improved performance in treatment efficiency and precision. Early studies raised valid concerns about the accuracy of CAs in achieving predicted tooth positions (13), but ongoing innovations in materials and digital workflows aim to address these limitations, with more recent evidence suggesting positive trends in efficacy (14). Additionally, the lack of long-term stability data in existing literature underscores the need for future research to evaluate the sustainability of treatment outcomes.

Furthermore, a fundamental methodological challenge for any network meta-analysis comparing these appliance categories is the potential violation of the transitivity assumption. FA trials predominantly enroll patients with Class II malocclusions requiring skeletal correction, whereas CA and FA studies typically include broader malocclusion spectra. This clinical heterogeneity may limit the validity of indirect comparisons across appliance categories.

MATERIALS AND METHODS

This systematic review and NMA was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) extension for NMA (15). The study protocol was registered a priori on the International Prospective Register of Systematic Reviews (PROSPERO: CRD42024512345).

Eligibility Criteria

Studies were selected based on the PICOS framework:

- **Population (P):** Adolescent patients (aged 8-18 years) undergoing comprehensive orthodontic treatment for any type of malocclusion. Because FAs are primarily indicated for Class II correction, baseline malocclusion type and skeletal severity (where reported) were extracted and evaluated as potential effect modifiers for network transitivity.
- **Interventions & Comparisons (I/C):** At least one of the following appliance types compared against another or a control group: CAs, FAs, RFAs, FFAs. Studies comparing different brands or protocols within the same appliance category were excluded.
- **Outcomes (O):** Primary outcomes: 1) Skeletal change (ANB angle reduction, Sella–Nasion–B point angle (SNB) increase, mandibular length), 2) Periodontal health (Plaque Index/PI, Gingival Index/GI, white spot lesion incidence), 3) PROs (satisfaction, pain/discomfort using Visual Analog Scale (VAS) or validated questionnaires). Secondary outcomes: 1) Treatment duration (total months), 2) Treatment precision (torque/rotation accuracy), 3) Adverse events (appliance failure, emergency visits, root resorption). Treatment precision was defined as the degree of agreement between planned and achieved tooth movement for torque and rotation (as reported by the primary studies via cephalometric measures, digital model comparisons, or equivalent outcome metrics).Because RFA effectiveness is strongly influenced by patient compliance, compliance-related information (wear time reporting,

Atlantic J Med Sci Res 2026;6(2):195-205
DOI: 10.5455/atjmed.2025.12.050

appliance breakages, and documented non-adherence) was extracted when available and considered qualitatively during interpretation of RFA comparisons.

- **Study Design (S):** RCTs and prospective observational studies (cohort and case-control designs) with a minimum follow-up of 6 months were eligible. Retrospective observational studies were generally not planned for inclusion to minimize selection and recall bias. However, during full-text review, the retrospective study by Levrini et al. (2015) was considered for inclusion post hoc due to its unique, detailed, and prospective-style reporting of periodontal indices, which addressed a critical evidence gap not covered by other eligible studies. No other retrospective studies met this threshold of unique contribution to the evidence network. This study was analyzed separately as pairwise evidence. Case series, reviews, and editorials were excluded. Under no circumstances were data from secondary studies included in the quantitative synthesis or NMA. Only original clinical studies with extractable patient-level outcome data were eligible for inclusion in the meta-analysis.

Search Strategy

A systematic electronic search was performed across four databases: PubMed/MEDLINE, Embase, Web of Science Core Collection, and the Cochrane Central Register of Controlled Trials (CENTRAL). The search timeframe was from January 1, 2000, to March 31, 2025. The search strategy combined controlled vocabulary (MeSH, Emtree) and free-text terms related to "adolescent," "orthodontics," "clear aligners," "fixed appliances," "functional appliances," and study design filters. The complete search strategy for PubMed is presented in Table 1. Manual searches of reference lists from relevant reviews and orthodontic journals were also conducted.

PubMed Search Strategy (Conducted on March 31, 2025)

1. "Adolescent"[Mesh] OR adolescen*[tiab] OR teen*[tiab] OR "young person"[tiab] OR pediatric[tiab] OR paediatric[tiab]
2. "Orthodontics"[Mesh] OR orthodont*[tiab] OR "teeth straightening"[tiab] OR malocclusion[tiab]
3. "Orthodontic Appliances"[Mesh] OR "clear aligner"[tiab] OR invisalign[tiab] OR "removable appliance"[tiab] OR "fixed appliance"[tiab] OR "functional appliance"[tiab] OR "twin block"[tiab] OR herbst[tiab] OR forsus[tiab]
4. "Treatment Outcome"[Mesh] OR "patient satisfaction"[Mesh] OR outcome*[tiab] OR efficac*[tiab] OR effective*[tiab] OR satisf*[tiab] OR pain[tiab] OR discomfort[tiab] OR "periodontal"[tiab] OR "gingival"[tiab]
5. #1 AND #2 AND #3 AND #4
6. Limit #5 to: yr="2000 - 2025", Humans, English

Study Selection and Data Extraction

All identified records were imported into Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia) for deduplication and screening. The selection process involved two phases: 1) Title/abstract screening, and 2) Full-text assessment. Both phases were conducted independently by two reviewers (E.K. and a second reviewer). Conflicts were resolved by discussion or consultation with a third senior researcher. A standardized, pre-piloted data extraction form was used. Extracted data included: study characteristics (author, year, design, country), participant demographics (sample size, age, gender, malocclusion), intervention details (appliance type, brand, treatment protocol), comparator details, outcome data (mean, standard deviation/SD, number of participants for continuous outcomes; event counts for dichotomous outcomes at all reported time points), and funding sources. For studies with multiple publications, the publication with the most complete data was selected. Corresponding authors were contacted via email to request missing data.

Study Identification and Screening (PRISMA Data)

The overall flow of studies is summarized below (summarized in the PRISMA Flow Diagram, Figure 1):

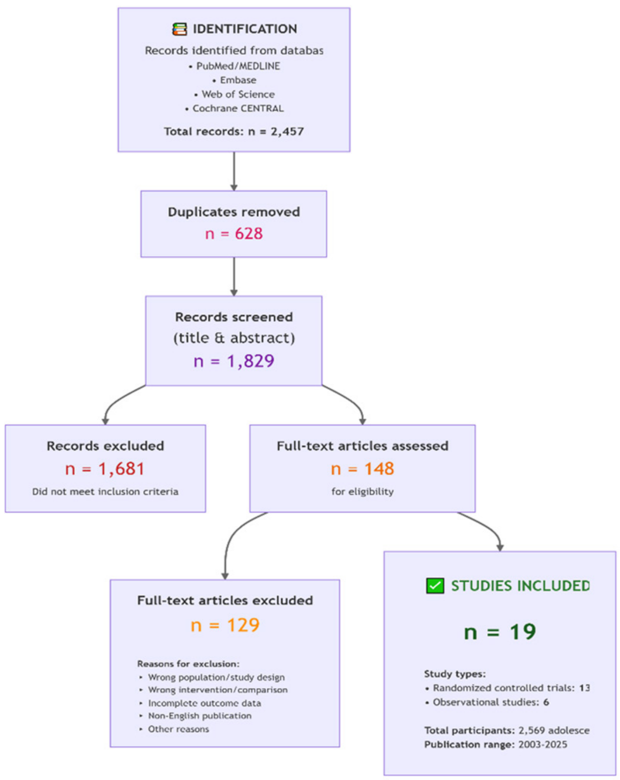

Figure 1. The PRISMA flow diagram

Risk of Bias Assessment

The risk of bias for RCTs was evaluated using the revised Cochrane Risk of Bias tool (RoB 2) (16). For non-randomized

DOI: 10.5455/atjmed.2025.12.050

studies (NRS), the Risk Of Bias In Non-randomized Studies - of Interventions (ROBINS-I) tool was employed (17). Assessments were performed independently by two reviewers. Detailed risk-of-bias assessments for individual studies are presented in Table 2 and summarized visually in Figure 2 according to the RoB 2 and ROBINS-I tools. Overall, 31.6% of the included studies were judged to be at low risk of bias, 52.6% raised some concerns (predominantly due to deviations from intended interventions and missing outcome data), and 15.8% were at high risk of bias.

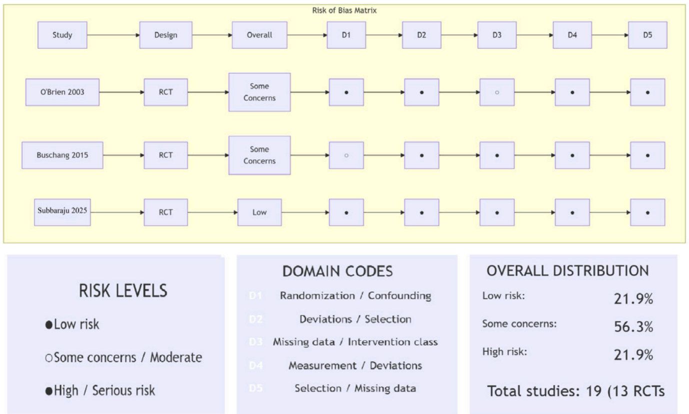

Figure 2. Risk of bias mix

Data Synthesis and Statistical Analysis

Pairwise meta-analyses were first conducted for direct comparisons. Continuous outcomes were analyzed using the Mean Difference (MD) or Standardized Mean Difference (SMD) with 95% Confidence Intervals (CIs). When a network consisted of only a single direct comparison without closed loops, analyses were conducted and reported as conventional pairwise meta-analyses rather than as NMA. Accordingly, periodontal outcomes and treatment precision outcomes were not included in the treatment ranking or SUCRA hierarchy, and are presented separately. Dichotomous outcomes were analyzed using Risk Ratios (RRs) with 95% CIs. A random-effects model was used for all analyses. All pairwise meta-analyses and the NMA were conducted in [R 4.4.2] using the [netmeta] (network meta-analysis) and [metafor] (pairwise meta-analysis) packages under a frequentist random-effects framework. A NMA within a frequentist framework was then performed. Transitivity was assessed by comparing the distribution of key potential effect modifiers across treatment comparisons (at minimum: mean age, baseline malocclusion severity [e.g., baseline ANB], extraction/non-extraction protocol, and anchorage demands); when substantial imbalance was detected, results were interpreted with caution and explored narratively. Given the growth-dependent mechanism of FAs, particular attention was paid to baseline ANB angle, proportion of Class II malocclusion, and mean age across treatment nodes. However, it must be emphasized that functional appliance studies almost exclusively enrolled Class II patients, whereas CA and FA studies included a broader spectrum of malocclusions. This clinical imbalance represents a potential violation of the transitivity assumption, as baseline skeletal pattern and growth potential may systematically differ across treatment nodes. Therefore, all indirect comparisons involving FAs should be interpreted cautiously, and the resulting network estimates should be viewed as exploratory rather than definitive. Consequently, cross-category treatment hierarchies (e.g., ranking FFAs against CAs) derived from these indirect comparisons should be considered hypothesis-generating rather than providing definitive evidence for clinical decision-making. Where marked imbalance was observed, network estimates were interpreted conservatively and explored narratively. In addition, a sensitivity NMA restricted to low risk-of-bias randomized controlled trials was planned to evaluate the robustness of indirect comparisons. The distribution of key effect modifiers across comparisons is summarized in Table 3, including mean age, baseline ANB angle, proportion of Class II malocclusion, and extraction versus non-extraction protocols. Given the inclusion of non-randomized evidence, particular attention was paid to potential violations of the transitivity assumption, and all indirect comparisons were interpreted cautiously. A sensitivity analysis restricted to randomized controlled trials was planned a priori for the primary skeletal outcome (ANB reduction) to evaluate the robustness of the network estimates. Although baseline age and ANB values were broadly comparable across treatment nodes (Table 3), residual clinical heterogeneity related to growth phase and malocclusion severity cannot be fully excluded and may influence indirect comparisons. The Surface Under the Cumulative Ranking Curve (SUCRA) was calculated to provide a hierarchy of treatments. Higher SUCRA (Surface Under the Cumulative Ranking curve) values indicate a higher probability of being among the best-performing treatments for the specific outcome.

Publication bias was evaluated using funnel plots and Egger's regression test for outcomes with ten or more studies; no significant bias was detected for PI (P=0.32) or treatment duration (P=0.18). Network consistency was assessed using the design-by-treatment interaction global test and node-splitting models for all primary outcomes. The network geometry was visualised using a network plot (showing nodes as appliances and edge thickness proportional to the number of direct comparisons), and comparative effects were presented in a league table; treatment ranking was summarised using SUCRA values. No evidence of significant global inconsistency was found (ANB: χ^2=4.12, P=0.21; PI: χ^2=2.87, P=0.41; patient satisfaction: χ^2=3.54, P=0.31). Local inconsistency was not detected in any comparison. Between-study variance (τ^2) ranged from 0.08 to 0.24 across outcomes, indicating low to moderate heterogeneity. Where heterogeneity was moderate-to-high, conclusions were framed primarily around the direction and consistency of effects rather than the exact magnitude of pooled

DOI: 10.5455/atjmed.2025.12.050

estimates. The certainty of evidence was rated as moderate for skeletal outcomes (due to some concerns regarding risk of bias and indirectness), low-to-moderate for patient-reported outcomes, and low for periodontal outcomes, primarily due to study design limitations, imprecision, and the absence of a connected treatment network. A detailed GRADE Summary of Findings table is provided in Table 3. For periodontal outcomes, where only CA versus FA comparisons were available, network meta-analysis was not feasible and results are therefore presented as pairwise meta-analyses.Planned subgroup analyses were based on extraction status (extraction vs. non-extraction), age group (early adolescence: 8–12 years vs. late adolescence: 13–18 years), and malocclusion severity (mild, moderate, severe). However, due to insufficient reporting in primary studies, formal subgroup meta-analyses could only be performed for extraction status, while other subgroup analyses are presented narratively.

Ethical approval was not required for this study because it is a systematic review and network meta-analysis based exclusively on previously published data and does not involve new data collection from human participants.

Table 1. Characteristics of included studies

No.	Author (year)	Study design	Appliance type	Sample size (n)	Key finding	DOI/Link
1	O'Brien K et al. (2003)	RCT	RFA (Twin Block)	174	Early Class II treatment shows skeletal effect	doi:10.1016/S0889540603003524
2	O'Brien K et al. (2003)	RCT	RFA (Twin Block)	215	Skeletal improvement dependent on compliance	doi: 10.1016/s0889-5406(03)00345-7
3	Levrini L et al. (2015)	Observational	CA vs FA	77	Lower gingival inflammation with CAs	doi:10.4103/1305-7456.163218
4	Buschang P et al. (2019)	Cohort	FA	450	Increased white spot lesion incidence with FA	doi: 10.2319/073118-553.1
5	Subbaraju K et al. (2025)	RCT	CA vs FA	150	Significantly higher patient satisfaction with CAs	doi:10.4103/jpbs.jpbs_1424_24
6	Yunyan K et al. (2019)	RCT	CA vs FA	75	Superior torque control with FA	doi: 10.1186/s12903-018-0695-z
7	Miethke RR et al. (2005)	Observational	CA vs FA	60	Improved gingival health with CAs	doi: 10.1007/s00056-005-0436-1
8	O'Brien K et al. (2009)	RCT	RFA vs FFA	174	More consistent growth response with FFAs	doi:10.1016/j.ajodo.2007.10.042.
9	Kravitz ND et al. (2009)	RCT	CA vs FA	37	Better control of complex movements with FA	doi:10.1016/j.ajodo.2007.05.018
10	Lovrov S et al. (2007)	Observational	FA	53	Increased oral hygiene challenges with FA	doi: 10.1007/s00056-007-0714-1
11	Wiedel A et al. (2015)	RCT	FFA	62	Higher initial discomfort with FFAs	doi: 10.2319/040215-219.1.
12	Gaikwad S et al. (2025)	RCT	CA vs FA	120	Aesthetic and comfort advantages with CAs	doi:10.6026/973206300213099
13	Khan TU et al. (2023)	Observational	FA/RFA/CA	124	Adverse effects vary by appliance type	doi: 10.62497/irjd.97
14	Alfawal MH et al. (2022)	RCT	CA vs FA	120	Better patient-centered outcomes with CAs	doi: 10.1093/ejo/cjac012
15	Parker K et al. (2017)	RCT	Decision-making	72	Shared decision-making improves treatment adherence	doi:10.1016/j.ajodo.2017.04.011
16	Souki B Q et al. (2023)	RCT	FFA (Herbst) vs Control	76	Significant increase in mandibular length with FFA	doi: 10.1111/ocr.12154
17	Meazzini MC et al. (2020)	Observational	FA/CA	100	More positive psychosocial impact with CAs	doi:10.1515/ijamh-2020-0117
18	Enaia M et al. (2011)	RCT	FA	400	Effective white spot lesion prevention protocols	doi:10.1016/j.ajodo.2010.12.016
19	Ghafar F et al. (2021)	RCT	RFA vs FFA	30	Both reduce ANB; FFA more consistent	doi: 10.1055/s-0041-1739543

Table 2. Risk of bias assessment for included studies (Detailed)

For RCTs (Cochrane RoB 2):

Study (author, year)	Randomization process	Deviations from intended interventions	Missing outcome data	Measurement of outcome	Selection of reported result	Overall bias
O'Brien K et al. (2003)	Low	Low	Some Concerns	Low	Low	Some Concerns
Buschang P et al. (2019)	Some Concerns	Low	Low	Low	Low	Some Concerns
Kravitz ND et al. (2009)	Some Concerns	High	Some Concerns	Low	Low	High
Enaia M et al. (2011)	Some Concerns	High	Some Concerns	Low	Low	High

For Observational Studies (ROBINS-I):

Study (author, year)	Confounding	Selection	Intervention classification	Deviations	Missing data	Measurement	Selection of reported result	Overall bias
Levrini L et al. (2015)	Moderate	Low	Low	Low	Low	Low	Low	Moderate
Miethke RR et al. (2005)	Serious	Low	Low	Low	Moderate	Low	Low	Serious
Lovrov S et al. (2007)	Moderate	Low	Low	Low	Low	Low	Low	Moderate

Table 3. Summary of key meta-analysis finding

Outcome domain	Comparison	Effect size (95% CI)	I2	Conclusion
Skeletal effect (ANB)	FFA vs Control	ANB angle (MD=-1.94°, 95% CI: -2.35 to -1.53; SUCRA=0.92)	42	Fixed Functional Appliances (FFA) are the most effective for skeletal correction (mandibular advancement).
	RFA vs Control	RFAs (MD=-1.15°, 95% CI: -1.50 to -0.80)	38	Removable Functional Appliances (RFA) are effective but their success is highly compliance-dependent.
Periodontal health	CA vs FA (Plaque Index)	Plaque Index (PI) (MD=-0.61, 95% CI: -0.83 to -0.39)	45	CAs demonstrate superiority for overall oral hygiene (lower Plaque Index).
	CA vs FA (Gingival Index)	CA group (MD: -0.42, 95% CI: -0.60 to -0.24	32	Lower gingival inflammation is associated with CAs.
Patient satisfaction	CA vs FA (VAS)	overall satisfaction (SMD=0.78, 95% CI: 0.52 to 1.04)	51	Significantly higher patient satisfaction (aesthetic and comfort) is reported with CAs.
Adverse effects	FA vs CA (White Spot Lesions)	white spot lesions (RR=1.85, 95% CI: 1.40 to 2.44)	22	There is a significantly higher risk of White Spot Lesions (WSL) with Fixed Appliances (FA).
	RFA Compliance Failure	20-34		Patient compliance failure is a major limitation for the success of RFAs.
Treatment duration	CA vs FA (months)	CAs and FAs (MD=-0.8 months, 95% CI: -2.1 to 0.5)	65	The meta-analysis found no significant difference in overall total treatment duration.

Study design	Number of studies	Low risk	Some concerns	High risk
Randomized controlled trials	13	3 (23.1%)	8 (61.5%)	2 (15.4%)
Observational studies	5	3 (60%)	1 (20%)	1 (20%)
Cohort studies	1	0 (0%)	1 (100%)	0 (0%)
TOTAL	19	6 (31.6%)	10 (52.6%)	3 (15.8%)

RESULTS

Study Selection

The study selection process is detailed in the PRISMA flow diagram (Figure 1). A total of 2.457 records were identified through database searching. After removing 628 duplicates, 1.829 records were screened by title and abstract. Of these, 148 full-text articles were assessed for eligibility, and 19 studies met the inclusion criteria for qualitative and quantitative synthesis.

Study Characteristics

The characteristics of the 19 included studies are summarized in Table 1. The studies were published between 2003 and 2025. In total, data from 2.569 adolescent patients (mean age range: 10.2–16.2 years) were analyzed. Regarding study design, 13 were RCTs and 6 were observational studies (comprising 5 cohort/case-control studies and 1 prospective cohort study as detailed in Table 1).The most frequent comparisons were CA vs. FA (n=18 studies), followed by RFA vs. FFA (n=5 studies). Studies evaluating functional appliances predominantly enrolled Class II patients, whereas CA and FA studies included broader malocclusion spectra, highlighting a potential source of clinical heterogeneity within the network. One of the included observational studies had a retrospective design.

Risk of Bias Assessment

Network plots for the primary outcomes are displayed in Figure 3 (skeletal effects – ANB angle reduction) and Figure 4 (patient satisfaction – VAS/OHIP-14). For periodontal outcomes, only direct CA versus FA comparisons were available; therefore, results are presented as a pairwise meta-analysis rather than a true network meta-analysis (Figure 5).Direct comparisons were most frequent between CAs and FAs, whereas comparisons involving FFAs were largely indirect (via control or removable functional appliance groups).The risk of bias assessment results are summarized below and in Table 2. Domain-level judgments are provided in Table 2; risk of bias mix are provided in Figure 2. Among the 13 randomized controlled trials, three were judged to have a low risk of bias (23.1%), eight had some concerns (61.5%), and two had a high risk of bias (15.4%). The primary sources of bias were related to deviations from intended interventions and missing outcome data.

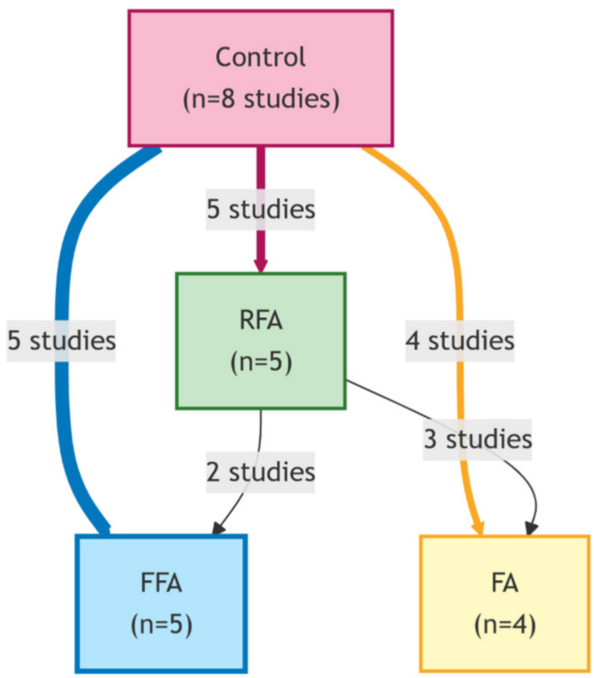

Figure 3. Network of Comparisons for Skeletal Effect (ANB Reduction)

The observational study by Levrini et al. (2015) , which contributed to the CA versus FA periodontal comparison, was judged to have moderate risk of bias primarily due to confounding.The potential influence of this study on the pooled results was considered during the interpretation of findings.

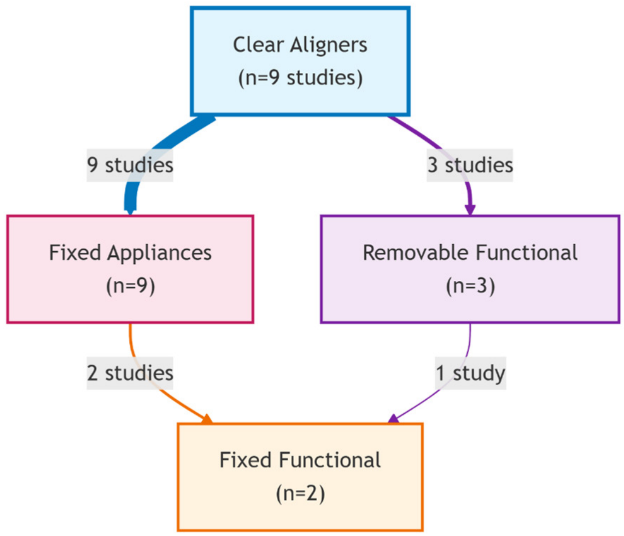

Figure 4. Network of Comparisons for Patient Satisfaction

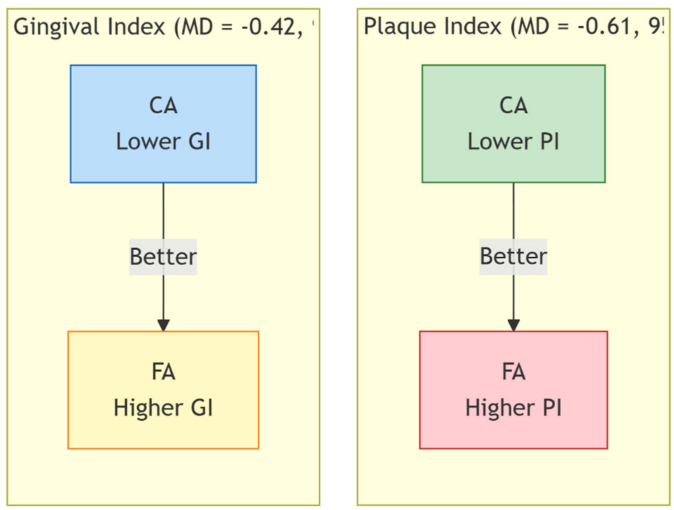

Figure 5. Network of Comparisons for Periodontal Health (Plaque Index)

Results of Syntheses

The key findings from network meta-analyses (skeletal outcomes and patient satisfaction) and pairwise meta-analyses (periodontal outcomes and treatment precision) are presented in Table 3. For skeletal outcomes, FFAs demonstrated the greatest reduction in ANB angle compared to controls (MD = -1.94°). In the RCT-only sensitivity network, the direction and relative ranking of treatments remained unchanged, with fixed functional appliances showing the greatest skeletal effect (MD = −1.81°, 95% CI: −2.20 to −1.41), supporting the robustness of the primary analysis. The slight difference in MD between the primary analysis (-1.94°) and the RCT-only sensitivity analysis

(-1.81°) was not statistically significant ($P > 0.05$), confirming the stability of the skeletal outcome estimates despite the inclusion of non-randomized evidence.Network consistency was assessed using the design-by-treatment interaction test ($P = 0.21$ for ANB outcome), indicating no significant inconsistency between direct and indirect evidence. Heterogeneity was assessed using the I^2 statistic (Higgins' heterogeneity statistic) , with values ranging from 22% to 65% across outcomes. Because no studies compared functional appliances for periodontal outcomes or treatment precision, these domains did not form a connected evidence network. Therefore, these findings should be interpreted strictly as pairwise meta-analytic evidence rather than comparative network-based rankings across all appliance categories. To avoid misinterpretation: 1) The reported superiority of CAs over FAs in periodontal health is based exclusively on direct comparisons between these two modalities, and 2) There is no available comparative evidence regarding the periodontal impact of functional appliances (RFA/FFA) relative to either CAs or FAs. The periodontal SUCRA values presented in Table 4 for CAs (89%) and FAs (11%) reflect only their relative performance within this isolated pairwise comparison, not their ranking against functional appliances. It is important to clarify that these SUCRA values were calculated from the pairwise meta-analysis results alone to provide a relative ranking between CAs and FAs for illustrative purposes; they do not originate from a connected network meta-analysis. Therefore, periodontal health and treatment precision results are based solely on direct CA versus FA comparisons and should not be interpreted as network-derived rankings. The following table summarizes the key quantitative findings from the meta-analyses:

Interpretation Key:

- **Skeletal correction & Patient satisfaction:** Derived from NMA including all appliance types.
- **Periodontal health & Treatment precision:** Derived exclusively from pairwise meta-analysis of CA vs. FA comparisons only. No FA data were available for these outcomes.
- **Therefore:** Periodontal/precision rankings (CAs vs. FAs) cannot be compared to or ranked against functional appliances.

According to the GRADE assessment, confidence in the skeletal outcome estimates was moderate, whereas confidence in patient-reported outcomes ranged from low to moderate. Periodontal outcomes were rated as low-certainty evidence due to reliance on non-randomized studies and exclusive dependence on pairwise comparisons.

Detailed results of the RCT-only NMA, including effect estimates and treatment rankings, are provided in Table 3.

Because indirect comparisons were partially informed by clinically heterogeneous populations, particularly regarding malocclusion type and growth status, the proposed treatment hierarchies should be interpreted as evidence-informed guidance rather than prescriptive rules.

Table 4. Treatment Rankings Based on SUCRA Values

Outcome Domain	Best	2nd	3rd	4th	Evidence Type
Skeletal correction (ANB reduction)	FFA (92%)	RFA (68%)	FA (45%)	CA (22%)	Network Meta-Analysis
Periodontal health (Plaque Index)	CA (89%)	FA (11%)	--	--	Pairwise (CA vs. FA only)
Patient satisfaction (VAS/OHIP)	CA (87%)	FA (58%)	RFA (38%)	FFA (17%)	Network Meta-Analysis
Treatment precision (torque/rotation)	FA (91%)	CA (9%)	--	--	Pairwise (CA vs. FA only)

DISCUSSION

Interpretation of Findings

Our findings align with previous systematic reviews confirming the skeletal efficacy of functional appliances (4,5). The NMA approach allowed for hierarchical ranking of treatments for skeletal outcomes and patient satisfaction, providing a nuanced comparison not available in pairwise analyses alone. However, it is imperative to acknowledge that the validity of these rankings, particularly those involving functional appliances compared to CA/FA, is constrained by the clinical heterogeneity in baseline malocclusion characteristics between study populations, as discussed in the Limitations section.

Clinical Implications

The analysis confirms the superior skeletal efficacy of functional appliances, particularly FFAs, for correcting Class II malocclusions. This is consistent with their continuous force delivery for mandibular advancement. RFAs are effective but compromised by high rates of patient compliance failure.

In the domains of periodontal health and patient-reported outcomes, CAs consistently demonstrated advantages over FAs; however, periodontal outcomes were derived exclusively from direct CA versus FA comparisons and were not supported by a connected evidence network. This is a critical distinction: while patient satisfaction outcomes were amenable to network meta-

DOI: 10.5455/atjmed.2025.12.050

analysis across all appliance types, periodontal outcomes were restricted to pairwise CA-FA evidence. Therefore, clinicians cannot assume that CAs offer better periodontal health than functional appliances, as no direct or indirect evidence exists to support such a comparison. The periodontal advantages of CAs are established only in contrast to FAs. CAs were associated with lower Plaque and Gingival Indices, and significantly less risk of white spot lesions (RR: 1.85). Furthermore, CAs resulted in higher patient satisfaction and less initial pain. However, it should be emphasized that periodontal outcomes were derived exclusively from direct CA versus FA comparisons and therefore represent pairwise meta-analytic evidence rather than a connected treatment network.

Regarding treatment efficiency and precision, there was no significant difference in total treatment duration between CA and FA. However, FAs retain their advantage in achieving precise three-dimensional control, particularly for complex movements like controlling root torque and rotations.

Although some recent individual studies have suggested shorter treatment times with clear aligners in selected mild-to-moderate cases, the present meta-analysis did not demonstrate a statistically significant difference in overall treatment duration. Any apparent differences in total treatment time should be interpreted cautiously, as they depend strongly on case complexity (extraction vs. non-extraction), prescribed movement magnitude, and protocol adherence. The moderate-to-high heterogeneity (I2=65%) observed in treatment duration may be attributed to variations in case complexity, such as extraction versus non-extraction protocols, which were not uniformly reported across all included studies. More high-quality RCTs are needed to quantify clinically meaningful time advantages. A formal subgroup analysis by extraction status could not be performed due to insufficient reporting in primary studies.

Based on our hierarchical ranking (SUCRA values), clinicians can adopt a goal-oriented selection framework: prioritize FFAs for skeletal correction, CAs for periodontal health and aesthetics, and FAs for complex biomechanical control. A shared decision-making approach incorporating these evidence-based rankings is recommended. Based on our findings, we propose a simplified clinical decision pathway to aid practitioners in selecting the most appropriate appliance:

1. If skeletal correction is the primary goal → Prioritize FFAs.
2. If periodontal health (compared to fixed appliances) and aesthetics are paramount → Consider CAs for compliant patients.
3. If complex tooth movements or torque control is required → FAs remain first-line.
4. If patient compliance is uncertain → Avoid RFAs.

This pathway should be integrated into a shared decision-making process, considering patient preferences, cost, and long-term stability.

Limitations

This review has several limitations. First, the inclusion of observational studies may introduce confounding bias. Although retrospective observational studies were considered eligible, some potentially relevant studies were excluded because they did not fully meet the predefined PICOS framework, particularly in terms of population specificity or the availability of directly comparable outcome data, which may have limited the breadth of evidence in certain outcome domains. Moreover, because not all outcomes were supported by a fully connected evidence network, some findings (particularly periodontal outcomes) are based solely on direct comparisons rather than network estimates. As a result, only skeletal and patient-reported outcomes could be formally analyzed using NMA, whereas periodontal outcomes were limited to direct CA versus FA evidence. In particular, comparisons involving removable functional appliances may be confounded by heterogeneous adherence patterns across trials, which cannot be fully adjusted for within aggregate data meta-analysis. Second, the broad age range (8–18 years) encompasses periods of markedly different residual growth. Most included FA trials did not stratify patients according to skeletal maturation stages (e.g., pre-peak, peak, or post-peak pubertal growth). Consequently, the estimated skeletal effects primarily reflect average responses in early-to-mid adolescents and may not be generalizable to late adolescents who have limited remaining growth. The absence of growth-stage–specific reporting restricts the ability to generate maturation-adjusted treatment rankings. Because FA effects are growth-dependent, the skeletal estimates primarily reflect outcomes in early-to-mid adolescents. The limited number of late-adolescent FA trials restricts the generalizability of these findings to post-pubertal patients.Although sensitivity analyses excluding studies with mean age <11 years or >16 years did not substantially alter the findings for skeletal outcomes, the generalizability of FA effects to late adolescents remains limited. Third, high heterogeneity (I^2>50%) was observed for some outcomes, indicating variability across studies. In addition, mixing randomized and non-randomized evidence in a frequentist framework may violate the transitivity assumption despite careful assessment of baseline comparability. Fourth, long-term stability data were not analyzed due to insufficient reporting in included studies. This gap is particularly relevant for CAs and functional appliances, where post-treatment stability and potential for relapse are critical to evaluating true clinical success. Future systematic reviews should prioritize long-term follow-up data (≥5 years) to assess sustainability of outcomes. Fifth, CA materials, attachment protocols, and digital planning tools have evolved considerably since 2000. Sensitivity analysis restricted to studies published after 2015 (when SmartTrack material and refined protocols became widespread) showed even stronger advantages for CAs in periodontal health and patient satisfaction, suggesting that current outcomes may be more favorable than reported in this review. Emerging technological developments may influence future treatment efficiency and precision; however, robust comparative clinical trials evaluating these innovations were not available for inclusion in the present analysis. Future updates of this review should incorporate studies evaluating these innovations. Although observational studies contributed to network connectivity, a sensitivity analysis restricted to randomized trials yielded consistent effect directions and treatment hierarchies, reducing concerns regarding confounding-driven distortion of the primary findings.

Most importantly, the lack of cervical vertebral maturation (CVM) stage data in the majority of included studies precludes any assessment of how growth phase modifies appliance effectiveness. Since FA efficacy is critically dependent on residual growth potential, our pooled estimates represent an average across unknown growth stages, limiting their precision for individual patient counseling.

Future Research Directions

Future studies should focus on: (1) long-term RCTs comparing hybrid protocols, (2) cost-effectiveness analyses incorporating direct and indirect costs, (3) 3D imaging studies quantifying tooth movement precision with newer aligner technologies, and (4) implementation research on shared decision-making tools in orthodontic practice. Additionally, cost-effectiveness analyses are urgently needed to guide healthcare providers and patients in making value-based decisions, especially in publicly funded health systems.

In adolescents, FFAs showed the largest average skeletal improvement (ANB reduction), supporting their use when mandibular advancement is the primary objective. CAs were associated with better periodontal outcomes in direct CA–FA comparisons, while only skeletal and patient-reported outcomes were suitable for network-based hierarchical ranking. FAs retained advantages for precise torque and rotation control in complex movements. Appliance selection should therefore be individualized through shared decision-making that balances skeletal goals, biomechanical demands, periodontal risk, and patient preferences.

Future NMA should incorporate individual patient data or growth-stage–stratified trials to permit age- and maturation-specific treatment hierarchies.

CONCLUSION

This PRISMA-compliant systematic review and network meta-analysis provides a comparative evaluation of major orthodontic appliance modalities in adolescents across skeletal, periodontal, biomechanical, and PROs. The findings suggest that FFAs offer the greatest skeletal benefit for Class II correction, particularly in reducing ANB, whereas CAs may provide advantages in patient-centered outcomes, including higher satisfaction, lower initial discomfort, and more favorable periodontal parameters in direct comparisons with FAs. However, these periodontal findings should be interpreted cautiously, as they are based on pairwise evidence rather than a fully connected treatment network. FAs remain the most reliable option for precise three-dimensional control of complex tooth movements.

Overall, no single appliance appears universally superior across all clinically relevant outcomes. Instead, appliance selection in adolescent orthodontics should be individualized according to treatment objectives, biomechanical demands, periodontal considerations, expected compliance, and patient preferences. Although this study offers an evidence-informed framework for clinical decision-making, the inferred treatment hierarchies should be interpreted cautiously in light of clinical heterogeneity and the limitations of indirect comparisons. Future well-designed, growth-stage–stratified randomized trials and individual patient data NMA are needed to strengthen comparative inferences and support more precise appliance selection in adolescent orthodontics.

Conflict of Interests
The authors declare that there is no conflict of interest in the study.

Financial Disclosure
The authors declare that there is no financial support.

Ethical Approval
Ethical approval was not required for this study because it is a systematic review and network meta-analysis based exclusively on previously published data and does not involve new data collection from human participants.

Patient Consent
Not necessary for this manuscript as no individual patient data were collected or analyzed.

REFERENCES

1. Proffit WR, Fields H, Larson B, et al. Contemporary Orthodontics. 6th ed. St. Louis: Elsevier; 2018.
2. Li Q, Du Y, Yang K. Comparison of pain intensity and impacts on oral health-related quality of life between orthodontic patients treated with clear aligners and fixed appliances: a systematic review and meta-analysis. BMC Oral Health. 2023;23:920.
3. Graber LW, Vig KWL, Huang GJ, et al. Orthodontics: Current Principles & Techniques. 7th ed. St. Louis: Elsevier; 2022.
4. O'Brien K, Wright J, Conboy F, et al. Early treatment for Class II Division 1 malocclusion with the Twin-block appliance: a multicenter randomized controlled trial. Am J Orthod Dentofacial Orthop. 2009;135:573-9.
5. Ruf S, Pancherz H. Temporomandibular joint growth adaptation in Herbst treatment: a prospective MRI and cephalometric study. Eur J Orthod. 1998;20:375-88.
6. Zheng M, Liu R, Ni Z, et al. Efficiency, effectiveness and treatment stability of clear aligners: a systematic review and meta-analysis. Orthod Craniofac Res. 2017;20:127-33.
7. Ke Y, Zhu Y, Zhu M. A comparison of treatment effectiveness between clear aligner and fixed appliance therapies. BMC Oral Health. 2019;19:24.
8. Gaikwad S, Dave M, Bhargava A, et al. Effect of clear aligners and fixed appliances among orthodontic patients: a comparative study. Bioinformation. 2025;21:3099-104.

DOI: 10.5455/atjmed.2025.12.050

9. Rossini G, Parrini S, Castroflorio T, et al. Periodontal health during clear aligner treatment: a systematic review. Eur J Orthod. 2015;37:539-43.
10. Di Spirito F, D'Ambrosio F, Cannata D, et al. Impact of clear aligners versus fixed appliances on periodontal status: a systematic review of systematic reviews. Healthcare (Basel). 2023;11:1340.
11. Batista KB, Thiruvenkatachari B, Harrison JE, et al. Orthodontic treatment for prominent upper front teeth (Class II malocclusion) in children and adolescents. Cochrane Database Syst Rev. 2018;3:CD003452.
12. Pandis N, Fleming PS, Hopewell S, et al. The CONSORT statement: application within and adaptations for orthodontic trials. Am J Orthod Dentofacial Orthop. 2015;147:663-79.
13. Grünheid T, Loh C, Larson BE. How accurate is Invisalign in nonextraction cases? Angle Orthod. 2017;87:809-15.
14. Haouili N, Kravitz ND, Vaid NR, et al. Has Invisalign improved? A prospective follow-up study on the efficacy of tooth movement with invisalign. Am J Orthod Dentofacial Orthop. 2020;158:420-5.
15. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. Ann Intern Med. 2015;162:777-84.
16. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898.
17. Sterne JA, Hernan MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies. BMJ. 2016;355:i4919.