Systematic Review Article

Extraction vs nonextraction orthodontic treatment: a systematic review and meta-analysis

Karim Gamal Elias[a]; Gautham Sivamurthy[b]; David R. Bearn[c]

ABSTRACT

Objectives: To compare four first premolar extraction and nonextraction treatment effects on intra-arch width, profile, treatment duration, occlusal outcomes, smile aesthetics and stability.

Materials and Methods: An electronic search of the literature to June 2, 2023 was conducted using health science databases, with additional search of gray literature, unpublished material, and hand searching, for studies reporting nonsurgical patients with fixed appliances regarding sixteen sub-outcomes. Data extraction used customized forms, quality assessed with ROBINS-I (Risk Of Bias In Non-randomized Studies—of Interventions) and Cochrane RoB 2 (risk-of-bias) tool. GRADE (Grading of Recommendations Assessment, Development and Evaluation) assessed certainty of evidence.

Results: Thirty (29 retrospective studies, 1 randomized controlled trial) studies were included. Random-effect meta-analysis (95% CI) demonstrated maxillary (MD: −2.03 mm; [−2.97, −1.09]; $P < .0001$) and mandibular inter-first molar width decrease (MD: −2.00 mm; [−2.71, −1.30]; $P < .00001$) with four first premolar extraction; mandibular intercanine width increase (MD: 0.68 mm; [0.36, 0.99]; $P < .0001$) and shorter treatment duration (MD: 0.36 years; [0.10, 0.62]; $P = .007$) in the nonextraction group. Narrative synthesis included three and five studies for upper and lower lip-E plane, respectively. For American Board of Orthodontics Objective Grading System and maxillary/mandibular anterior alignment (Little's irregularity index), each included two studies with inconclusive evidence. There were no eligible studies for UK Peer Assessment Rating (PAR) score. Class I subgroup/sensitivity analyses favored the same results. Prediction interval indicated no significant difference for all outcomes.

Conclusions: Four first premolar extraction results in maxillary and mandibular inter-first molar width decrease and retraction of upper/lower lips. Nonextraction treatment results in mandibular intercanine width increase and shorter treatment duration. There was no significant difference between the two groups regarding maxillary intercanine width, US PAR score, and posttreatment smile esthetics. Further high-quality focused research is recommended. (*Angle Orthod*. 2024;94:83–106.)

KEY WORDS: Orthodontic extractions; Arch width; Profile; Treatment outcomes; Smile aesthetics; Stability

INTRODUCTION

The longest running debate in orthodontics, spanning more than a century, has been the effects of extraction and nonextraction treatment.[1] The main concern with extraction treatment has been the possible deleterious effect on facial profile and the main concern with nonextraction treatment being post-treatment stability.[2] Edward Angle's philosophy of preserving the full complement of teeth argued that extraction of teeth would cause an imbalance in facial harmony and abnormal function due to the change in arch width and form.[3] Unlike many of Angle's disciples, Calvin Case opposed this philosophy and defended

[a] Private Practice, Cairo, Egypt; and former postgraduate student, Orthodontics, School of Dentistry, University of Dundee, Dundee, United Kingdom.

[b] Clinical Lecturer and Honorary Consultant in Orthodontics, School of Dentistry, University of Dundee, Dundee, United Kingdom.

[c] Honorary Professor of Orthodontics, School of Dentistry, University of Dundee, Dundee; and Consultant in Orthodontics, NHS Fife, United Kingdom.

Corresponding author: Dr Karim Gamal Elias, School of Dentistry, University of Dundee, Dundee, United Kingdom (e-mail: kgnelias@outlook.com)

Accepted: September 2023. Submitted: February 2023.
Published Online: October 30, 2023

the extraction of teeth in treating malocclusion to avoid later relapse.[4] However, it was not until the 1940s, when more members of the orthodontic community (including Charles H. Tweed and Raymond Begg) also supported an extraction treatment approach, that it became a generally accepted option.[5,6]

Since then, the pendulum has swung between extraction and non-extraction treatment, reporting a peak extraction rate of 76% in 1968[7] declining to 17.6% in 2005[8] among University of North Carolina patients. At the University of São Paulo, nonextraction treatment continued with an upward trend from 14.29% (1973–1977) to 54.55% (2003–2007).[9]

The orthodontic literature has discussed this conundrum, with conflicting results. Bowman and Johnston[10] examined the effects on facial profile and concluded from a sample of 120 patients that extraction treatment had positive results for patients who had initial protrusion relative to the E plane, but it was detrimental for those who had retrusive lips before starting treatment. Boley et al.[11] studied profiles of 50 patients and concluded that no difference was found between the two groups as facial profile measurements (Holdaway H-line) were within normal limits. Konstantonis[12] attributed change in the soft tissue profile of extraction patients to greater incisor retraction, which could be controlled during treatment planning with less retraction mechanics and more mesial movement of posterior segments. These effects were more pronounced in patients with thin lips or high lip strain.

Little et al.[13] concluded that extraction did not guarantee long term stability and Rossouw et al.[14] reported no significant difference in stability between extraction and nonextraction groups, with similar amounts of relapse.

The literature has previously reported premolar extraction compared to nonextraction treatment focused on limited outcomes.[12,15–23] A recently published scoping review[24] outlined the weaknesses of published evidence across the breadth of the current literature but did not include any quantitative evaluation of the available data. This systematic review was, therefore, focused on four first premolar extraction, a broad range of outcomes and quantitative analysis, providing the orthodontist with the evidence required to inform clinical decisions.

The aim of this systematic review was to compare four first premolar extraction and nonextraction treatment effects on arch form, maxillary and mandibular intercanine width and first molar width, profile changes (upper and lower lip prominence to E plane), treatment duration, occlusal outcomes (end treatment UK and US weighted peer assessment rating [PAR] scores, American Board of Orthodontics Objective Grading System [ABO-OGS] score), posttreatment smile aesthetics (aesthetic score, maxillary intercanine width/smile width, visible dentition width/smile width, maxillary intercanine width/visible dentition width) and posttreatment changes of maxillary and mandibular anterior alignment (Little's irregularity index) to provide orthodontists with the best data available.

MATERIALS AND METHODS

This review was prepared in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 statement and the Cochrane Handbook for Systematic Reviews of Interventions (PROSPERO:CRD42021254523).

Eligibility criteria are in Table 1. Electronic databases were searched until June 2, 2023 without restrictions regarding publication year, study design, or language, with additional searching of gray literature, unpublished literature, and hand-searching of reference lists of included and excluded studies comparing premolar extraction to nonextraction treatment for outcomes of interest. Search strategies and publication date range of the search are in Table 2.

The articles resulting from the search were added to Zotero (version 6.0.26). Duplicates were identified and removed. Articles were manually checked during screening and further duplicates found and removed. Articles were first checked and excluded by title, with the resultant articles screened by their abstract and then full text articles checked for eligibility.

If there were any difficulties in getting the full text of an article, soft copies were obtained from the University of Dundee Library or The British Library. No contact was made with the authors.

Non-English studies without an English version were translated using Google Translate. Sample size and all reported data were checked and values were revised and recalculated whenever raw data were provided to ensure quality of included data.

Two reviewers undertook study selection and data extraction in duplicate using a customized data extraction form (Appendix 1 and 2) prepared by the third reviewer. When there was disagreement, discussion with the third reviewer reached the final decision.

Due to the ethical challenges in undertaking comparative prospective and randomized clinical trials in this subject area, retrospective studies were included. ROBINS-I tool was used to assess the quality of observational (prospective and retrospective) studies, with each outcome being individually judged. Cochrane RoB 2 tool was used for randomized trials.

Where appropriate, continuous data, with sample size, mean value, and standard deviation were available, RevMan (version 5.4.1) was used for quantitative

Table 1. Eligibility Criteria

	Arch Width	Profile	Treatment Duration	Occlusal Outcomes	Smile Aesthetics	Stability
Participants	Nonsurgical patients with fixed orthodontic appliances with no adjunctive procedures reported, such as expansion appliances or interproximal reduction.	Nonsurgical patients with fixed orthodontic appliances with no adjunctive procedures reported, such as expansion appliances or interproximal reduction and not reporting use of extraoral or functional orthopedic appliances.	Nonsurgical patients with fixed orthodontic appliances with no adjunctive procedures reported, such as expansion appliances or interproximal reduction.			
Intervention	Four first premolar extraction treatment					
Comparison	Nonextraction treatment					
Outcome measures/ Pooled data	Maxillary and mandibular intercanine and interfirst molar width. Pooled data: mean treatment changes. (Pretreatment/ end treatment).	Soft tissue cephalometric measurements: upper and lower lip prominence relative to E- plane Pooled data: mean treatment changes. (Pretreatment/ end treatment)	Pooled data: treatment duration in years.	UK and US weighted peer assessment rating [PAR] scores and American Board of Orthodontics Objective Grading System (ABO-OGS) Pooled data: end treatment for PAR scores and total score for ABO-OGS.	Aesthetic score (5-point scale), ratio of maxillary intercanine width/smile width, visible dentition width/smile width and maxillary intercanine width/ visible dentition width Pooled data: posttreatment measurements.	Maxillary and mandibular anterior alignment (Little's irregularity index from canine to canine). Posttreatment measurements taken 3 y or more after treatment. Pooled data: mean posttreatment changes. (End treatment/ posttreatment)
Study design	Prospective (randomized and nonrandomized) and retrospective studies.					
Language	Language restrictions were not applied.					
Exclusion criteria	Studies with more than one error in sample size/treatment data (inconsistent data) in any part of the manuscript were considered of low internal consistency, and excluded.					

synthesis and narrative synthesis reported when meta-analysis was not possible. Confidence interval (95%) with mean difference was used with significance level $P < .05$. I^2 statistic for random effects model meta-analysis was calculated using Comprehensive Meta-Analysis software (version 3.0). Prediction interval (95%) with mean difference, used to describe the distribution of true effect sizes, was calculated using an Excel spreadsheet (Microsoft Corp., Redmond, Washington) based on formulas by Borenstein et al.[25,26]

Outcome measures and time points of assessment are presented in Table 1. Studies reporting female and male subgroups were combined into a single group using RevMan. Random-effects model meta-analysis was used because of the amount of heterogeneity due to the difference in populations and study design. Heterogeneity was assessed by assessing overlap of the confidence intervals on Forest plots and I^2 statistic with threshold for interpretation as described in the Cochrane Handbook.

Subgroup and sensitivity analyses were carried out to deal with possible sources of heterogeneity of including different malocclusion classes together and differences in outcome measures to isolate their influence.

Publication bias was addressed by including unpublished literature. When more than 10 studies pooled together for an outcome in the meta-analysis, publication bias was identified through a funnel plot. GRADE was used to assess certainty of evidence for each outcome.

RESULTS

Study Selection and Characteristics

The search of databases (including gray literature and hand-searching) identified 2652 articles. Removal of duplicates, exclusion by title (Appendix 3), and screening by abstracts (Appendix 4), resulted in a total of 383 articles (Appendix 5).

Thirty (29 retrospective studies[27–55], 1 randomized controlled trial [RCT][56]) studies were included with some studies including multiple outcomes. Twenty-

Table 2. Search Strategy[a]

Database	Search Strategy/ Keywords
Cochrane Library (September 1993–June 2, 2023) DOSS (EBSCO) (April 1982–June 2, 2023) Medline Ultimate (EBSCO) (August 1975–June 2, 2023) PubMed (August 1975–June 2, 2023)	((orthodonti* OR "orthodontic treatment") AND (extract*) AND (nonextract* OR non-extract* OR "non extract*") AND ("arch width" OR width OR intraarch OR intra-arch OR intercanine OR intercanine OR intermolar OR inter-molar OR "occlusal outcome*" OR abo OR "objective grading system" OR "par index" OR "par score" OR "treatment duration" OR "treatment time" OR stability OR relapse OR "little's irregularity index" OR esthetic* OR aesthetic* OR smil* OR "hard tissue*" OR "soft tissue*" OR lip OR profile OR "facial profile"))
Scopus (May 1949–June 2, 2023)	TITLE-ABS-KEY (((orthodonti* OR "orthodontic treatment") AND (extract*) AND (nonextract* OR non-extract* OR "non extract*") AND ("arch width" OR width OR intraarch OR intra-arch OR intercanine OR inter-canine OR intermolar OR inter-molar OR "occlusal outcome*" OR abo OR "objective grading system" OR "par index" OR "par score" OR "treatment duration" OR "treatment time" OR stability OR relapse OR "little's irregularity index" OR esthetic* OR aesthetic* OR smil* OR "hard tissue*" OR "soft tissue*" OR lip OR profile OR "facial profile"))) AND (LIMIT-TO (SUBJAREA , "DENT"))
VHL Regional Portal (December 1974–June 2, 2023) Web of Science (January 1964–June 2, 2023)	((orthodonti* OR "orthodontic treatment") AND (extract*) AND (nonextract* OR non-extract* OR "non extract*") AND ("arch width" OR width OR intraarch OR intra-arch OR intercanine OR intercanine OR intermolar OR inter-molar OR "occlusal outcome*" OR abo OR "objective grading system" OR "par index" OR "par score" OR "treatment duration" OR "treatment time" OR stability OR relapse OR "little's irregularity index" OR esthetic* OR aesthetic* OR smil* OR "hard tissue*" OR "soft tissue*" OR lip OR profile OR "facial profile"))
ClinicalTrials.gov (until June 2, 2023)	Condition or disease: extraction in orthodontics Other terms: non extraction treatment
Google Scholar (until June 2, 2023) OpenGrey (DANS EASY) (until June 2, 2023)	((orthodonti* OR "orthodontic treatment") AND (extract*) AND (nonextract* OR non-extract* OR "non extract*") AND ("arch width" OR width OR intraarch OR intra-arch OR intercanine OR intercanine OR intermolar OR inter-molar OR "occlusal outcome*" OR abo OR "objective grading system" OR "par index" OR "par score" OR "treatment duration" OR "treatment time" OR stability OR relapse OR "little's irregularity index" OR esthetic* OR aesthetic* OR smil* OR "hard tissue*" OR "soft tissue*" OR lip OR profile OR "facial profile"))

[a] Note: For Cochrane Library, (title, abstract, keyword) selected in advanced search. No search limits were applied.

For Medline/DOSS (Dental and Oral Sciences Source), no limiters/expanders were applied in advanced search, with no field selected. Boolean/Phrase selected in search mode.

For PubMed, no filters were applied in basic search.

For Scopus, Field codes (TITLE-ABS-KEY) and (LIMIT-TO (SUBJAREA, "DENT") used in advanced search.

For VHL (Virtual Health Library) Regional Portal, (title, abstract, subject) selected in advanced search. No filters were applied.

For Web of Science, search documents within all fields. No filters were applied.

For ClinicalTrials.gov (classic website), Status (all studies) selected within basic search, (country) field left blank.

For Google Scholar, (include citations) and (include patents) selected. No filters were applied. Search resulted in 4440 hits; 990 hits were reviewed in 99 pages.

For OpenGrey, basic search applied within DANS EASY (Data Archiving and Networked Services Electronic Archiving SYstem) Archive.

four[27,29–33,35,36,38–44,46–51,54–56] were included in meta-analysis, of which one thesis[47] and two non-English articles in Chinese[54] and Korean[42] were included (Table 3). Narrative synthesis included three studies[31,45,53] for UL-E plane (upper lip), five studies[31,34,37,45,53] for LL-E plane (lower lip), two studies[28,52] for ABO-OGS, and two studies[32,33] for maxillary and mandibular anterior alignment (Little's irregularity index) (Table 4). Figure 1 shows the process of study identification and selection.

Risk of Bias Within Studies

The included RCT[56] was judged as being of high risk of bias (Figure 2). For retrospective studies, 23 were of serious risk[28–33,35–41,43,45–50,53–55] and six of moderate risk of bias[27,34,42,44,51,52] (Table 5) (Appendix 6).

Synthesis of Results

Results of meta-analyses, prediction interval, subgroup, and sensitivity analyses are presented in Table 6 with values rounded to two decimal places except for *P* value.

Arch Width

Intercanine Width (Figure 3). Nine retrospective studies[27,31,39,42,44,47,49,51,54] and one RCT[56] reported no statistically significant difference between four first premolar extraction and nonextraction treatment in maxillary intercanine width (MD: 0.02 mm; total 95% CI [−0.38, 0.43]; $I^2 = 0\%$; $P = .91$) with significant increase in mandibular intercanine width (MD: 0.68 mm; 95% CI [0.36, 0.99]; $I^2 = 0\%$; $P < .0001$) in the nonextraction group.

Table 3. Outcomes Included in Quantitative Synthesis (Meta-Analysis)[a]

Author/Y	Study Design	Participants				Intervention	Outcome
		Place/Country	Sex	Pretreatment Mean Age (Y)	Pretreatment Malocclusion		
Arch Width							
Aksu and Kocadereli,[27]	RS	University clinic	F 1st PE: 19 F, 11 M NE: 18 F, 12 M	F 1st PE: 14.3 ± 2.02 NE: 14.1 ± 2.9	Skeletal class I and Angle class I	F 1st PE: 30 NE: 30	Max. and Mand. ICW, IMW
Choi et al.[31]	RS	Seoul National University Bundang Hospital, Seongnam, Korea	All female	F 1st PE: 24.6 ± 5.8 NE: 28.6 ± 8.4	F 1st PE: 11 Class I, 4 Class II molar NE: 13 Class I, 4 Class II molar	F 1st PE: 15 NE: 17	Max. and Mand. ICW, IMW.
De Almeida et al.[56]	RCT	University of Lins, Dental School, SP, Brazil	F 1st PE: 9 M, 12 F NE: 10 M, 10 F	F 1st PE: 13.4 ± 1.0 NE: 13.1 ± 1.7	Angle Class I	F 1st PE: 21 NE: 20	Max. and Mand. ICW, IMW.
Dong et al.[54] (article in Chinese)	RS	Shandong University, China	Not mentioned	13–15 y	F 1st PE: 11 Class I, 10 Class II, 4 Class III NE: 15 Class I, 6 Class II, 4 Class III	F 1st PE: 25 NE: 25	Max. and Mand. ICW.
Işık et al.[39]	RS	University clinic	F 1st PE: 7 M, 20 F NE: 13 M, 29 F	F 1st PE: 13.57 ± 2.58 NE: 14.21 ± 2.79	Not mentioned	F 1st PE: 27 NE: 42	Max. and Mand. ICW, IMW.
Jeon et al.[42] (article in Korean)	RS	Kyung Hee University, South Korea	Not mentioned	F 1st PE: 14.3 NE: 14.1	F 1st PE: 20 Class I, 10 Class II NE: 20 Class I, 10 Class II	F 1st PE: 30 NE: 30	Max. and Mand. ICW, IMW.
Kim and Gianelly[44]	RS	University clinic	F 1st PE: 17 M, 13 F NE: 12 M, 18 F	F 1st PE: 14.1 NE: 14.2	F 1st PE: 18 Class I, 12 Class II division 1 NE: 18 Class I, 12 Class II division 1	F 1st PE: 30 NE: 30	Max. and Mand. ICW, IMW
MacKriel[47] (Thesis)	RS	Not mentioned	F 1st PE: 13 M, 13 F NE: 13 M, 13 F	F 1st PE: 13.93 ± 1.72 NE: 13.73±3.66	F 1st PE: 20 Class I, 3 Class II division 1, 3 Class II division 2 NE: 16 Class I, 5 Class II division 1, 5 Class II division 2	F 1st PE: 26 NE: 26	Max. and Mand. ICW, IMW.
Oz et al.[49]	RS	Ondokuz Mayıs University, Turkey	F 1st PE: 35 M, 45 F NE: 32 M, 48 F	F 1st PE: 14.3 ± 3.4 NE: 13.8 ± 2.1	Not mentioned	F 1st PE: 80 NE: 80	Max. and Mand. ICW, IMW.
Sumit and Ashima,[51]	RS	Manipal College of Dental Sciences, India	F 1st PE: 9 M, 16 F NE: 11 M, 14 F	F 1st PE: 18.2 ± 3.5 NE: 18.3± 3.8	Class I dental and skeletal.	F 1st PE: 25 NE: 25	Max. and Mand. ICW, IMW.
Treatment Duration							
Beit et al.[29]	RS	Private orthodontic offices and school of Dentistry of the National and Kapodistrian University of Athens, Greece	F 1st PE: 23 F, 18 M NE: 24 F, 18 M	F 1st PE: 13.71 ± 3.28 NE: 14.62 ± 3.84	Class I dental and skeletal	F 1st PE: 41 NE: 42	Treatment duration in y
Bishara et al.[30]	RS	University of Iowa, USA	F 1st PE: 21 M, 23 F NE: 20 M , 27 F	F 1st PE: M: 11.5 ± 1.6 y, F: 11.6 ± 1.6 NE: M: 12.1 ± 1.5, F: 10.9 ± 1.5	Class II division 1	F 1st PE: 44 NE: 47	Treatment duration in y
De Almeida et al.[56]	RCT	University of Lins, Dental School, SP, Brazil	F 1st PE: 9 M, 12 F NE: 10 M, 10 F	F 1st PE: 13.4 ± 1.0 NE: 13.1 ± 1.7	Angle Class I	F 1st PE: 21 NE: 20	Treatment duration in y
Francisconi et al.[32]	RS	Bauru Dental School, University of Sao Paulo, Brazil	F 1st PE: 15 M, 25 F NE: 17 M, 27 F	F 1st PE: 13.01 ± 0.99 NE: 12.96 ± 1.10	F 1st PE: 25 Class I, 15 Class II (6 1/2, 1 3/4, and 8 full unit). NE: 21 Class I, 23 Class II (4 1/2, 6 3/4 and 13 full unit).	F 1st PE: 40 NE: 44	Treatment duration in y
Freitas et al.[33]	RS	Bauru Dental School, University of Sao Paulo, Bauru, Brazil	F 1st PE: 44 M, 53 F NE: 24 M, 34 F	F 1st PE: 13.03 ± 1.09 NE: 12.83 ± 1.11	F 1st PE: 60 Class I, 37 Class II (7 1/4, 9 1/2, 5 3/4 and 16 full unit). NE: 29 Class I, 29 Class	F 1st PE: 97 NE: 58	Treatment duration in y

Table 3. Continued

Author/Y	Study Design	Participants				Intervention	Outcome
		Place/Country	Sex	Pretreatment Mean Age (Y)	Pretreatment Malocclusion		
					II (5 1/4, 7 1/2, 4 3/4 and 13 full unit).		
Gorucu-Coskuner et al.[36]	RS	Hacettepe University, Turkey	F 1st PE: 9 F, 6 M NE: 13 F, 3 M	F 1st PE: 13.89 ± 5.69 NE: 13.34 ± 1.82	Skeletal Class I	F 1st PE: 15 NE: 16	Treatment duration in y
Kouli et al.[46]	RS	Orthodontic offices and Department of Orthodontics National and Kapodistrian University of Athens, Greece	F 1st PE: 15 M, 19 F NE: 15 M, 19 F	F 1st PE: 13.94 ± 3.23 NE: 13.98 ± 3.37	Class I dental	F 1st PE: 34 NE: 34	Treatment duration in y
MacKriel[47] (thesis)	RS	Not mentioned	F 1st PE: 13 M, 13 F NE: 13 M, 13 F	F 1st PE: 13.93 ± 1.72 NE: 13.73 ± 3.66	F 1st PE: 20 Class I, 3 Class II division 1, 3 Class II division 2 NE: 16 Class I, 5 Class II division 1, 5 Class II division 2	F 1st PE: 26 NE: 26	Treatment duration in y
Mahmood[55]	RS	College of Dentistry, Mosul, University, Iraq	F 1st PE: 13 F, 7 M NE: 11 F, 9 M	F 1st PE: 13.18 ± 1.63 NE: 12.97 ± 1.76	Class I dental and skeletal	F 1st PE: 20 NE: 20	Treatment duration in y
				Occlusal Outcomes (US-Weighted PAR Score)			
Freitas et al.[33]	RS	Bauru Dental School, University of Sao Paulo, Bauru, Brazil	F 1st PE: 44 M, 53 F NE: 24 M, 34 F	F 1st PE: 13.03 ± 1.09 NE: 12.83 ± 1.11	F 1st PE: 60 Class I, 37 Class II (7 1/4, 9 1/2, 5 3/4, and 16 full unit). NE: 29 Class I, 29 Class II (5 1/4, 7 1/2, 4 ¾ & 13 full unit).	F 1st PE: 97 NE: 58	PAR score US weight
Holman et al.[38]	RS	Clinic of the co-author (MGH)	F 1st PE: 39 M, 61 F NE: 50 F, 50 M	F 1st PE: 13.5 ± 1.4 NE: 13.5 ± 1.2	F1st PE: 40 Class I, 35 Class II division 1, 4 Class II division 2, 14 Class II subdivision, 7 Class III NE: 54 Class I, 29 Class II division 1, 3 Class II division 2, 11 Class II subdivision, 3 Class III	F 1st PE: 100 NE: 100	PAR score US weight
Janson et al.[41]	RS	Bauru Dental School, University of Sao Paulo, Brazil	F 1st PE: 15 F, 15 M NE: 19 F, 11 M	F 1st PE: 13.10 ± 1.56 NE: 12.38 ± 1.22	Class II division 1	F 1st PE: 30 NE: 30	PAR score US weight
				Smile Aesthetics (mean age y)			
Ghaffar and Fida[35]	RS	Aga Khan University, Karachi, Pakistan.	F 1st PE: 10 M, 20 F NE: 11 M, 19 F	15–30 y	Not mentioned	F 1st PE: 30 NE: 30	Ratio: ICW/SW, VDW/SW, ICW/VDW
Işiksal et al.[40]	RS	Ege University, Izmir, Turkey	F 1st PE: 13 F, 12 M NE: 13 F, 12 M	F 1st PE: 19.08 ± 2.40 NE: 19.04 ± 1.97	Angle Class I	F 1st PE: 25 NE: 25	Aesthetic score Ratio: ICW/SW, VDW/SW, ICW/VDW
Johnson and Smith[43]	RS	Three private orthodontic practices	F 1st PE: 15 M, 15 F NE: 15 M, 15 F	F 1st PE: 16.4 ± 2.93 NE: 15.6 ± 1.45	Not mentioned	F 1st PE: 30 NE: 30	Aesthetic score Ratio: ICW/SW, VDW/SW, ICW/VDW
Naik et al.[48]	RS	College of Dental Sciences, Davangere, India	All female	F 1st PE: 21.07 ± 2.84 NE: 21.87 ± 1.68	Not mentioned	F 1st PE: 15 NE: 15	Aesthetic score
Prasad et al.[50]	RS	King George's University of Dental Sciences, Lucknow, India	F 1st PE: 20 M, 20 F NE: 20 M, 20 F	20.16 y	Not mentioned	F 1st PE: 40 NE: 40	Aesthetic score Ratio: ICW/SW, VDW/SW, ICW/VDW

[a] F 1st PE indicates four first premolar extraction; NE, nonextraction; Max., maxillary; Mand, mandibular; ICW, intercanine width; IMW, intermolar width; SW, smile width; VDW, visible dentition width; RCT, randomized controlled trial; RS, retrospective study, y: years.

Table 4. Outcomes Included In Qualitative Synthesis (Narrative Synthesis)[a]

Author/Y	Study Design	Participants				Intervention	Outcome
		Place/Country	Sex	Pretreatment Mean Age (Y)	Pretreatment Malocclusion		
				Profile			
Choi et al.[31]	RS	Seoul National University Bundang Hospital, Seongnam, Korea	All female	F 1st PE: 24.6 ± 5.8 NE: 28.6 ± 8.4	F 1st PE: 11 Class I, 4 Class II molar NE: 13 Class I, 4 Class II molar	F 1st PE: 15 NE: 17	UL- E plane, LL- E plane
Freitas et al.[34]	RS	Centro de Educação Continuada do Maranhão, São Luís/MA	F 1st PE: 6 F, 4 M NE: 5 F, 5 M	12.3 y	Angle Class I	F 1st PE: 10 NE: 10	LL- E plane
Hassan et al.[37]	RS	University Hospital in Karachi, Pakistan	All female patients	F 1st PE: 23.43 NE: 24.49	F 1st PE: 12 Class I, 18 Class II NE: 20 Class I, 10 Class II	F 1st PE: 30 NE: 30	LL- E plane
Konstantonis[45]	RS	Saint Louis University Graduate Orthodontic Clinic, USA	Not mentioned	Not mentioned	Class I dental and skeletal	F 1st PE: 30 NE: 32	UL- E plane, LL- E plane
Xu et al.[53]	RS	Orthodontic Department, Peking University School of Stomatology, China	F 1st PE: 4 M, 9 F NE: 6 M, 6 F	F 1st PE: 12.46 ± 1.71 NE: 12.08 ± 1.08	F 1st PE: 4 Class I, 8 Class II, 1 Class III NE: 7 Class I, 5 Class II	F 1st PE: 13 NE: 12	UL- E plane, LL- E plane
				Occlusal Outcomes (ABO-OGS)			
Anthopoulou et al.[28]	RS	University of Athens graduate clinic and private orthodontic practices, Greece	F 1st PE: 16 F, 9 M NE: 20 F, 10 M	F 1st PE: 16.3 ± 7.84 NE: 13.79 ± 3.99	Class I dental and skeletal	F 1st PE: 25 NE: 30	ABO-OGS
Vaidya et al.[52]	RS	JSS Dental College and Hospital, JSS University, Mysore, India	F 1st PE: 11 F, 9 M NE: 12 F, 8 M	F 1st PE: 15.2 ± 4.2 NE: 14.6 ± 2.7	Class I dental and skeletal	F 1st PE: 20 NE: 20	ABO-OGS
				Stability			
Francisconi et al.[32]	RS	Bauru Dental School, University of Sao Paulo, Brazil	F 1st PE: 15 M, 25 F NE: 17 M, 27 F	F 1st PE: 13.01 ± 0.99 NE: 12.96 ± 1.10	F 1st PE: 25 Class I, 15 Class II (6 1/2, 1 3/4, and 8 full unit). NE: 21 Class I, 23 Class II (4 1/2, 6 3/4, and 13 full unit).	F 1st PE: 40 NE: 44	Max. and Mand. LII.
Freitas et al.[33]	RS	Bauru Dental School, University of Sao Paulo, Bauru, Brazil	F 1st PE: 44 M, 53 F NE: 24 M, 34 F	F 1st PE: 13.03 ± 1.09 NE: 12.83 ± 1.11	F 1st PE: 60 Class I, 37 Class II (7 1/4, 9 1/2, 5 3/4, and 16 full unit). NE: 29 Class I, 29 Class II (5 1/4, 7 1/2, 4 3/4, and 13 full unit).	F 1st PE: 97 NE: 58	Max. and Mand. LII.

[a] F 1st PE, four first premolar extraction; NE, nonextraction; Max., maxillary; Mand, mandibular; UL, upper lip; LL, lower lip; LII, Little's irregularity index; RS, retrospective study, y: years.

Intermolar Width (Figure 4). Eight retrospective studies[27,31,39,42,44,47,49,51] and one RCT[56] reported significant decrease in maxillary (MD: −2.03 mm; total 95% CI [−2.97, −1.09]; $I^2 = 0\%$; $P < .0001$) and mandibular interfirst molar width (MD: −2.00 mm; total 95% CI [−2.71, −1.30]; $I^2 = 5.32\%$; $P < .00001$) with four first premolar extraction.

Profile

Three studies[31,45,53] were included for UL- E plane and five studies[31,34,37,45,53] for LL- E plane with vote counting indicating retraction of upper and lower lips with four first premolar extraction.

Choi et al.[31] compared 15 four first premolar extraction (UL- E plane: −1.61 ± 1.62, LL- E plane: −3.13 ± 1.97) with 17 nonextraction (UL- E plane: −0.07 ± 0.89, LL- E plane: −0.15 ± 0.70) Class I and II female patients and found significant retraction of upper and lower lips in the extraction group.

In an equally divided sample of 20 Class I patients, Freitas et al.[34] reported no significant difference between four first premolar extraction (LL- E plane: −0.2 ± 3.7) and nonextraction (LL- E plane: −0.05 ± 1.9) treatment.

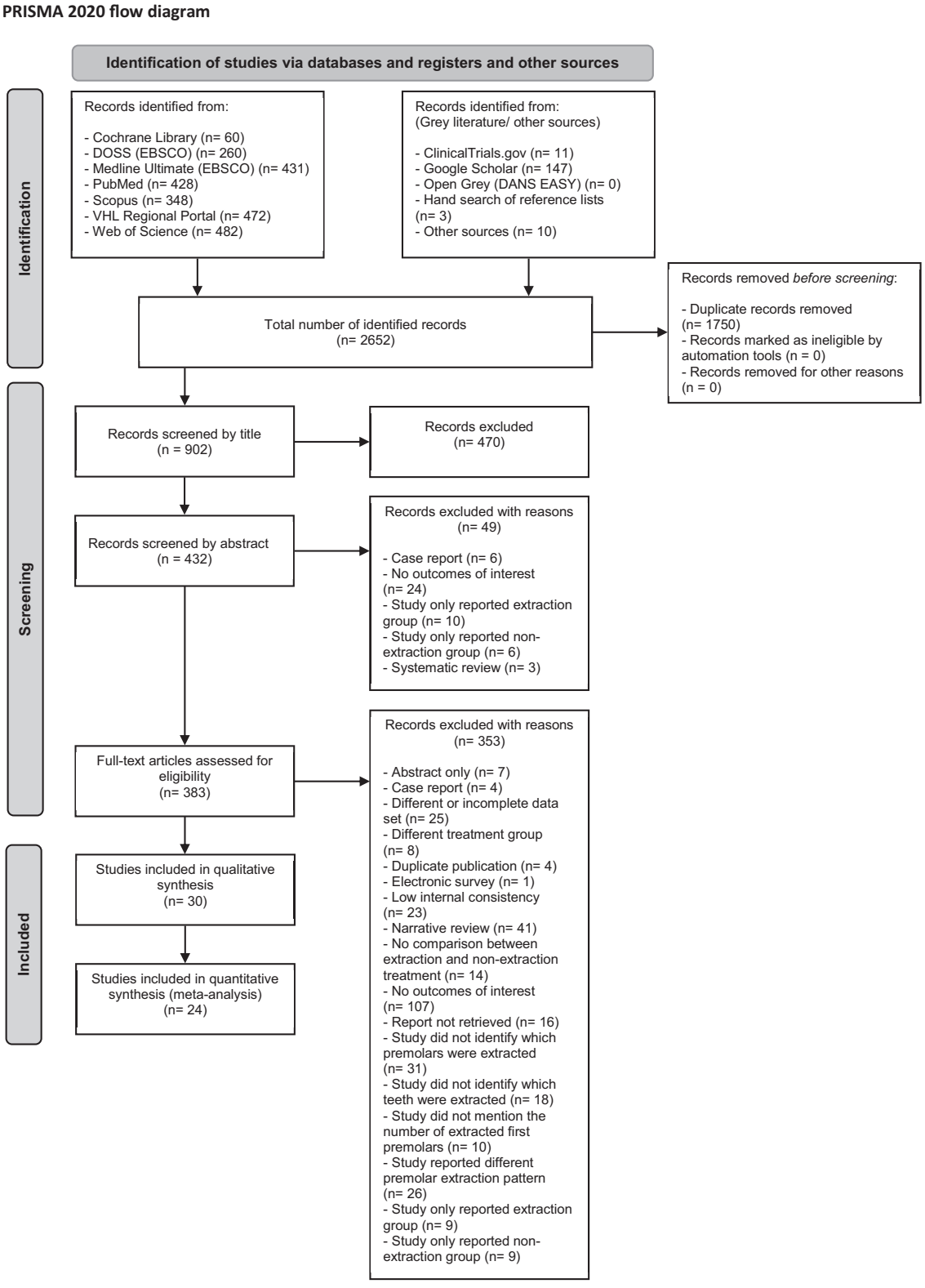


Figure 1. PRISMA Flow Diagram. PRISMA indicates Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

D1	D2	D3	D4	D5	Overall
-	!	+	!	+	-

D1 Randomization process was performed as follow: the first patient was placed into one of the two groups by the use of a coin toss and every following patient that was recruited was placed into every other group accordingly.
D2 Deviations from the intended interventions: it is likely that carers and people delivering the interventions were aware of participants' assigned intervention.
D3 Missing outcome data: Pre-tx, end-tx data and mean changes reported (mean and SD) for the two groups.
D4 Measurement of the outcome: Assessment of the outcomes could have been influenced by knowledge of the intervention received.
D5 Selection of the reported result: Maxillary and mandibular ICW and IMW width measurements were analyzed in accordance with a pre-specified plan and unlikely to have been selected from multiple outcomes/ analyses.

Figure 2. Risk of bias- De Almeida et al. (RCT).

Hassan et al.[37] reported no significant difference between four first premolar extraction (LL- E plane: −2.15 ± 3.38) and nonextraction (LL- E plane: −0.83 ± 2.75) treatment in a sample of 60 Class I and II Pakistani females.

Konstantonis[45] compared 30 four first premolar extraction (UL- E plane: −2.75 ± 1.5, LL- E plane: −3.34 ± 1.75) with 32 nonextraction (UL- E plane: −0.68 ± 1.89, LL- E plane: 0.67 ± 2.24) borderline Class I patients and found significant retraction of upper and lower lips in the extraction group.

Xu et al.[53] compared 13 four first premolar extraction (UL- E plane: –1.0 ± 1.9, LL- E plane: –2.6 ± 1.9) with 12 nonextraction (UL- E plane: –0.9 ± 2.4, LL- E plane: –0.4 ± 3.4) borderline Chinese patients of different malocclusion and found significant retraction of lower lip in the extraction group with no difference regarding upper lip retraction.

Treatment Duration (Figure 5)

Eight retrospective studies[29,30,32,33,36,46,47,55] and one RCT[56] reported shorter treatment duration in the nonextraction group (MD: 0.36 years; total 95% CI [0.10, 0.62]; $I^2 = 3.18\%$; $P = .007$) compared to the four first premolar extraction group.

Occlusal outcomes

PAR Score. No eligible studies were found for UK-weighted PAR score. Three retrospective studies[33,38,41] reported no statistically significant difference between four first premolar extraction and nonextraction treatment with US weighted PAR score. (MD: 0.33; total 95% CI [−0.21, 0.87]; $I^2 = 0\%$; $P = .23$) (Figure 6).

ABO-OGS. Two retrospective studies were included with inconclusive evidence. Anthopoulou et al.[28] compared 25 four first premolar extraction (total score: 27.04 ± 6.30) with 30 nonextraction (total score: 29.07 ± 7.11) Class I borderline patients and found no statistically significant difference between the two groups.

In 40 Class I borderline patients, Vaidya et al.[52] reported lower scores for four first premolar extraction group (total score: 22.0 ± 2.29), compared to nonextraction (total score: 26.8 ± 5.18).

Smile Aesthetics (Figure 7)

Four retrospective studies[40,43,48,50] for aesthetic score (MD: −0.09; total 95% CI [−0.24, 0.05]; $I^2 = 0\%$; $P = .21$) and four retrospective studies[35,40,43,50] for maxillary intercanine width/smile width (MD: 0.01; total 95% CI [−0.00, 0.02]; $I^2 = 0\%$; $P = .12$), visible dentition width/smile width (MD: −0.00; total 95% CI [−0.01, 0.01]; $I^2 = 0\%$; $P = .81$) and maxillary intercanine width/visible dentition width (MD: 0.00; total 95% CI [−0.02, 0.02]; $I^2 = 0\%$; $P = .94$) reported no statistically significant difference between four first premolar extraction and nonextraction treatment.

Stability

Two retrospective studies were included with inconclusive evidence. Francisconi et al.[32] compared 40 four first premolar extraction (maxillary Little index: 0.89 ± 1.48, mandibular Little index: 1.64 ± 1.75) with 44 nonextraction (maxillary Little index: 1.64 ± 1.37, mandibular Little index: 1.36 ± 1.33) patients of different malocclusions and found greater maxillary crowding relapse in the nonextraction group and no significant difference between the two treatment groups for mandibular crowding relapse.

Freitas et al.[33] compared 97 four first premolar extraction (maxillary Little index: 1.30 ± 1.75, mandibular Little index: 1.93 ± 2.06) with 58 nonextraction (maxillary Little index: 1.66 ± 1.42, mandibular Little index: 1.40 ± 1.18) patients with no significant difference between the two groups regarding maxillary crowding relapse, and more mandibular crowding relapse in the extraction group.

Table 5. ROBINS-I Tool

Study	Assessment by Outcome	Bias Due to Confounding	Bias in Selection of Participants into the Study	Bias in Classification of Interventions	Bias Due to Deviations From Intended Interventions	Bias Due to Missing Data	Bias in Measurement of Outcomes	Bias in Selection of the Reported Result	Overall Bias
Aksu and Kocadereli[27]	Max. ICW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. ICW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Max. IMW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. IMW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
Anthopoulou et al.[28]	ABO-OGS	Low	Serious	Low	Low	Low	Moderate	Low	Serious
Beit et al.[29]	Tx duration	NI	Serious	Low	Low	Low	Moderate	Serious	Serious
Bishara et al.[30]	Tx duration	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Choi et al.[31]	Max. ICW	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	Mand. ICW	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	Max. IMW	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	Mand. IMW	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	UL-E plane	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	LL-E plane	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
Dong et al.[54] [article in Chinese]	Max. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
Francisconi et al.[32]	Max. LII	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. LII	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Tx duration	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Freitas et al.[33]	Max. LII	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. LII	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	US weighted PAR score	Moderate	Serious	Low	Low	Low	Moderate	Serious	Serious
	Tx duration	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Freitas et al.[34]	LL-E plane	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
Ghaffar and Fida[35]	ICW/SW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	VDW/SW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	ICW/VDW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
Gorucu-Coskuner et al.[36]	Tx duration	NI	Serious	Low	Low	Low	Moderate	Serious	Serious
Hassan et al.[37]	LL-E plane	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Holman et al.[38]	US weighted PAR score	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
Işık et al.[39]	Max. ICW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. ICW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
	Max. IMW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. IMW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Işiksal et al.[40]	Esthetic score	Low	Low	Low	Low	Low	Serious	Low	Serious
	ICW/SW	Moderate	Low	Low	Low	Low	Moderate	Serious	Serious
	VDW/SW	Moderate	Low	Low	Low	Low	Moderate	Serious	Serious
	ICW/VDW	Moderate	Low	Low	Low	Low	Moderate	Serious	Serious
Janson et al.[41]	US weighted PAR score	Moderate	Low	Low	Low	Low	Moderate	Serious	Serious
Jeon et al.[42] [article in Korean]	Max. ICW	NI	Moderate	Low	Low	Low	Moderate	Low	Moderate
	Mand. ICW	NI	Moderate	Low	Low	Low	Moderate	Low	Moderate
	Max. IMW	NI	Moderate	Low	Low	Low	Moderate	Low	Moderate
	Mand. IMW	NI	Moderate	Low	Low	Low	Moderate	Low	Moderate
Johnson and Smith, 1995[43]	Esthetic score	Low	Serious	Low	Low	Low	Serious	Serious	Serious
	ICW/SW	NI	Serious	Low	Low	Low	Moderate	Serious	Serious
	VDW/SW	NI	Serious	Low	Low	Low	Moderate	Serious	Serious
	ICW/VDW	NI	Serious	Low	Low	Low	Moderate	Serious	Serious
Kim and Gianelly[44]	Max. ICW	NI	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. ICW	NI	Low	Low	Low	Low	Moderate	Low	Moderate
	Max. IMW	NI	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. IMW	NI	Low	Low	Low	Low	Moderate	Low	Moderate
Konstantonis[45]	UL-E plane	NI	Serious	Low	Low	Low	Moderate	Low	Serious

Table 5. Continued

Study	Assessment by Outcome	Bias Due to Confounding	Bias in Selection of Participants into the Study	Bias in Classification of Interventions	Bias Due to Deviations From Intended Interventions	Bias Due to Missing Data	Bias in Measurement of Outcomes	Bias in Selection of the Reported Result	Overall Bias
	LL-E plane	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Kouli et al.[46]	Tx duration	NI	Low	Low	Low	Low	Moderate	Serious	Serious
MacKriel[47]	Max. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
(thesis)	Mand. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Max. IMW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. IMW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Tx duration	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Mahmood[55]	Tx duration	NI	Low	Low	Low	Low	Moderate	Serious	Serious
Naik et al.[48]	Esthetic score	Low	Serious	Low	Low	Low	Serious	Low	Serious
Oz et al.[49]	Max. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. ICW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Max. IMW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
	Mand. IMW	Moderate	Serious	Low	Low	Low	Moderate	Low	Serious
Prasad et al.[50]	Esthetic score	Low	Serious	Low	Low	Low	Serious	Low	Serious
	ICW/SW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
	VDW/SW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
	ICW/VDW	NI	Serious	Low	Low	Low	Moderate	Low	Serious
Sumit and Ashima[51]	Max. ICW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. ICW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Max. IMW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
	Mand. IMW	Moderate	Low	Low	Low	Low	Moderate	Low	Moderate
Vaidya et al.[52]	ABO-OGS	NI	Low	Low	Low	Low	Moderate	Low	Moderate
Xu et al.[53]	UL-E plane	Low	Serious	Low	Low	Low	Moderate	Low	Serious
	LL-E plane	Low	Serious	Low	Low	Low	Moderate	Low	Serious

[a] Max., maxillary; Mand, mandibular; ICW, intercanine width; IMW, intermolar width; SW, smile width; VDW, visible dentition width; UL, upper lip; LL, lower lip; LII, Little's irregularity index; Tx: treatment, NI: no information.

Prediction Interval

Prediction interval, the range that in 95% of all populations the true effect size will fall within, was wider than 95% confidence intervals, but with no significant difference, for all outcomes.

Subgroup Analysis

Class I subgroup analyses of maxillary and mandibular intercanine width (Figure 3), maxillary and mandibular interfirst molar width (Figure 4) and treatment duration (Figure 5), favored the same results as the main analyses.

Sensitivity Analysis

For arch width, sensitivity analysis excluded three studies[27,47,56] in which measurement of intercanine (Figure 8) and intermolar (Figure 9) width was from the most labial aspect of buccal surfaces of teeth instead of canine tips and mesiobuccal cusp tips. Sensitivity analysis excluded two studies[30,32] for treatment duration reporting the use of extraoral appliances related to patient compliance (Figure 5), and two studies[35,48] for smile aesthetics (Figure 10), where it was not clear whether included data were end of treatment or posttreatment.

Risk of Bias Across Studies

No funnel plot generated, as no more than 10 studies were included in meta-analyses.

Quality of Evidence

GRADE evidence profile was completed for all outcomes (Table 7). No separate grading was undertaken for subgroup/sensitivity analyses.

DISCUSSION

This review only included studies with four first premolar extraction compared to nonextraction treatment-to-control heterogeneity, with the inclusion of gray literature, non-English studies and published theses to reduce reporting bias. However, eligible studies were of high level of bias except six studies of moderate risk of bias.

Eligibility criteria were set to reduce confounding variables related to effects due to different treatment approaches. For all outcomes, studies included were

Table 6. Synthesis of Results[a]

Time Points of Assessment	Outcome	Total 95% CI	Total 95% PI	Subgroup Analysis (Class I Subjects)	Sensitivity Analysis
Treatment changes (Pre/end Tx)	Max. ICW	(MD: 0.02 mm; total 95% CI [−0.38, 0.43]; $I^2 = 0\%$; P = 0.91)	(MD: 0.02 mm; total 95% PI [−1.00, 1.04])	(MD: 0.26 mm; subtotal 95% CI [−0.46, 0.98]; $I^2 = 8.65\%$; P = 0.47)	(MD: −0.06 mm; total 95% CI [−0.53, 0.41]; $I^2 = 0\%$; P = 0.80)
	Mand. ICW	(MD: 0.68 mm; total 95% CI [0.36, 0.99]; $I^2 = 0\%$; P <0.0001)	(MD: 0.68 mm; total 95% PI [−0.14, 1.50])	(MD: 0.77 mm; subtotal 95% CI [0.31, 1.24]; $I^2 = 0\%$; P = 0.001)	(MD: 0.61 mm; total 95% CI [0.17, 1.04]; $I^2 = 0\%$; P = 0.006)
	Max IMW	(MD: −2.03 mm; total 95% CI [−2.97, −1.09]; $I^2 = 0\%$; P < 0.0001)	(MD: −2.03 mm; total 95% PI [−5.26, 1.20])	(MD: −1.60 mm; subtotal 95% CI [−2.40, −0.80]; $I^2 = 0.32\%$; P < 0.0001)	(MD: −1.94 mm; total 95% CI [−3.16, −0.71]; $I^2 = 0\%$; P = 0.002)
	Mand. IMW	(MD: −2.00 mm; total 95% CI [−2.71, −1.30]; $I^2 = 5.32\%$; P < 0.00001)	(MD: -2.00 mm; total 95% PI [−4.44, 0.44])	(MD: −1.62 mm; subtotal 95% CI [−2.12, −1.12]; $I^2 = 0\%$; P < 0.00001)	(MD: −2.09 mm; total 95% CI [−3.17, −1.02]; $I^2 = 13.68\%$; P = 0.0001)
	UL- E plane	Narrative synthesis included three retrospective studies, with vote counting indicating retraction of upper lip with four first premolar extraction.			
	LL- E plane	Narrative synthesis included five retrospective studies, with vote counting indicating retraction of lower lip with four first premolar extraction.			
Treatment duration (y)	Treatment duration	(MD: 0.36 y; total 95% CI [0.10, 0.62]; $I^2 = 3.18\%$; P = 0.007)	(MD: 0.36 y; total 95% PI [−0.52, 1.24])	(MD: 0.57 y; subtotal 95% CI [0.22, 0.91]; $I^2 = 7.79\%$; P = 0.001	(MD: 0.38 y; total 95% CI [0.06, 0.69]; $I^2 = 0\%$; P = 0.02)
End treatment	UK weighted PAR score	No eligible studies were found.			
	US weighted PAR score	(MD: 0.33; total 95% CI [−0.21, 0.87]; $I^2 = 0\%$; P = 0.23)	(MD: 0.33; total 95% PI [−3.17, 3.83])	N/A	N/A
	ABO-OGS	Narrative synthesis included two retrospective studies with inconclusive evidence.			
Post treatment	Esthetic score	(MD: −0.09; total 95% CI [−0.24, 0.05]; $I^2 = 0\%$; P = 0.21)	(MD: −0.09; total 95% PI [−0.40, 0.22])	N/A	(MD: −0.11; total 95% CI [−0.28, 0.06]; $I^2 = 0\%$; P = 0.22)
	Max. ICW/SW	(MD: 0.01; total 95% CI [−0.00, 0.02]; $I^2 = 0\%$; P = 0.12)	(MD: 0.01; total 95% PI [−0.01, 0.03])	N/A	(MD: 0.01; total 95% [−0.00, 0.02]; $I^2 = 0\%$; P = 0.14)
	VDW/SW	(MD: −0.00; total 95% CI [−0.01, 0.01]; $I^2 = 0\%$; P = 0.81)	(MD: −0.00; total 95% PI [−0.02, 0.02])	N/A	(MD: −0.00; total 95% CI [−0.02, 0.02]; $I^2 = 9.31\%$; P = 0.91)
	Max. ICW/VDW	(MD 0.00; total 95% CI [−0.02, 0.02]; $I^2 = 0\%$; P = 0.94)	(MD 0.00; total 95% PI [−0.04, 0.04])	N/A	(MD: 0.00; total 95% CI [−0.02, 0.03]; $I^2 = 0\%$; P = 0.71)
Posttreatment changes (End/post Tx)	Max. LII Mand. LII	Narrative synthesis included two retrospective studies with inconclusive evidence.			

[a] Max., maxillary; Mand, mandibular; ICW, intercanine width; IMW, intermolar width; SW, smile width; VDW, visible dentition width; UL, upper lip; LL, lower lip; LII, Little's irregularity index.

limited to nonsurgical patients with fixed appliances with no adjunctive procedures. Studies reporting the use of functional appliances or extraoral appliances were excluded for profile.

Confounding variables were further controlled by exclusion of studies with incomplete data reporting to avoid imputations and studies with more than one error in sample size or treatment data, for greater consistency. However, this may have resulted in smaller sample sizes with a potential source of bias,[57] as not all studies provided raw data for recalculation. No study included in quantitative synthesis reported any error in outcomes of interest, and excluded studies are included in supplementary material with reason for their exclusions if needed for future analysis.

Ideally, age of subjects included would have been limited to 13 years of age or older for arch width[58] and 15 years of age or younger for profile changes[59] to exclude growth effects. Significant increase in maxillary and mandibular intercanine and intermolar width occurs due to growth between 3 and 13 years of age, and there is significant upper and lower lip retraction

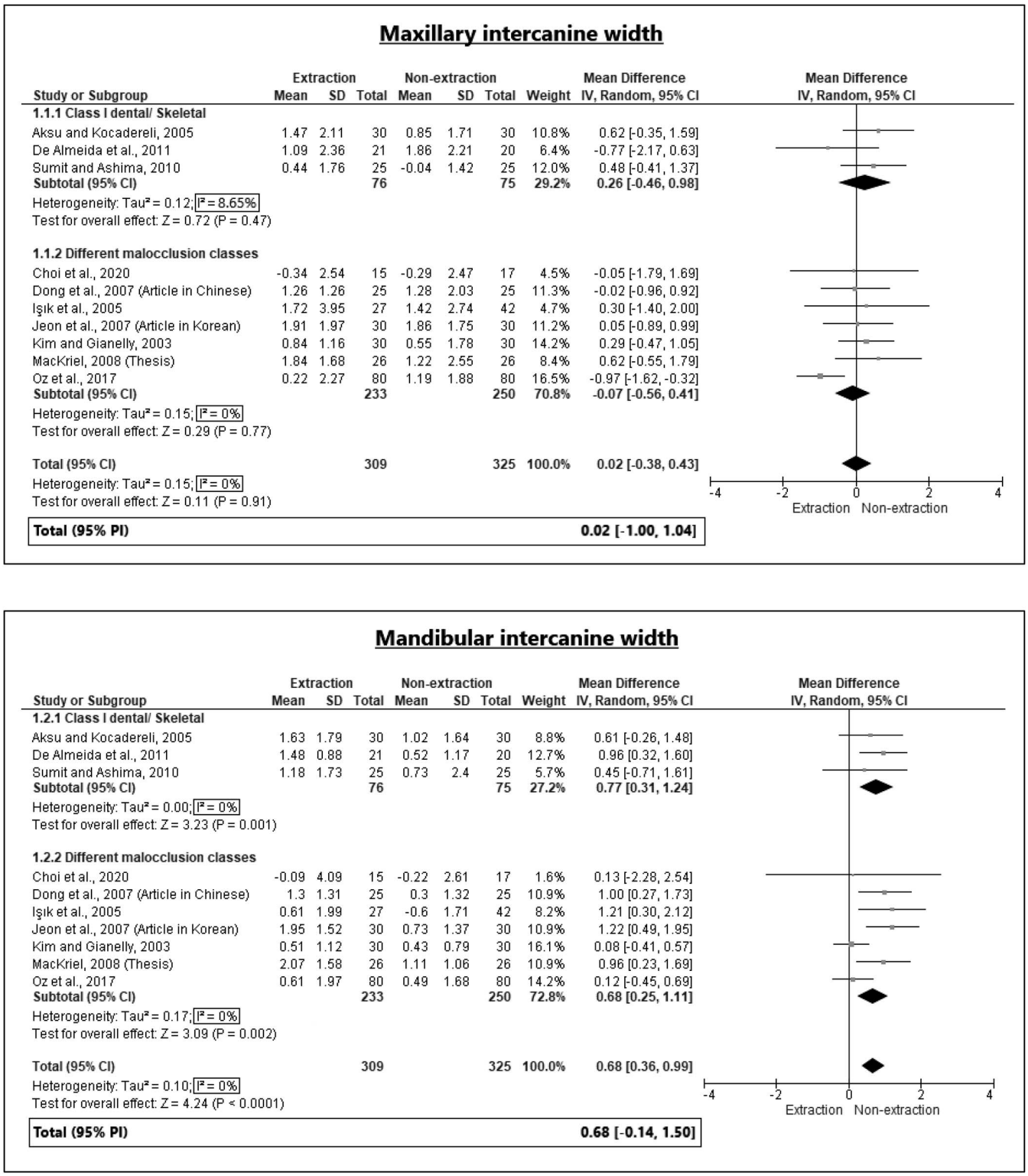
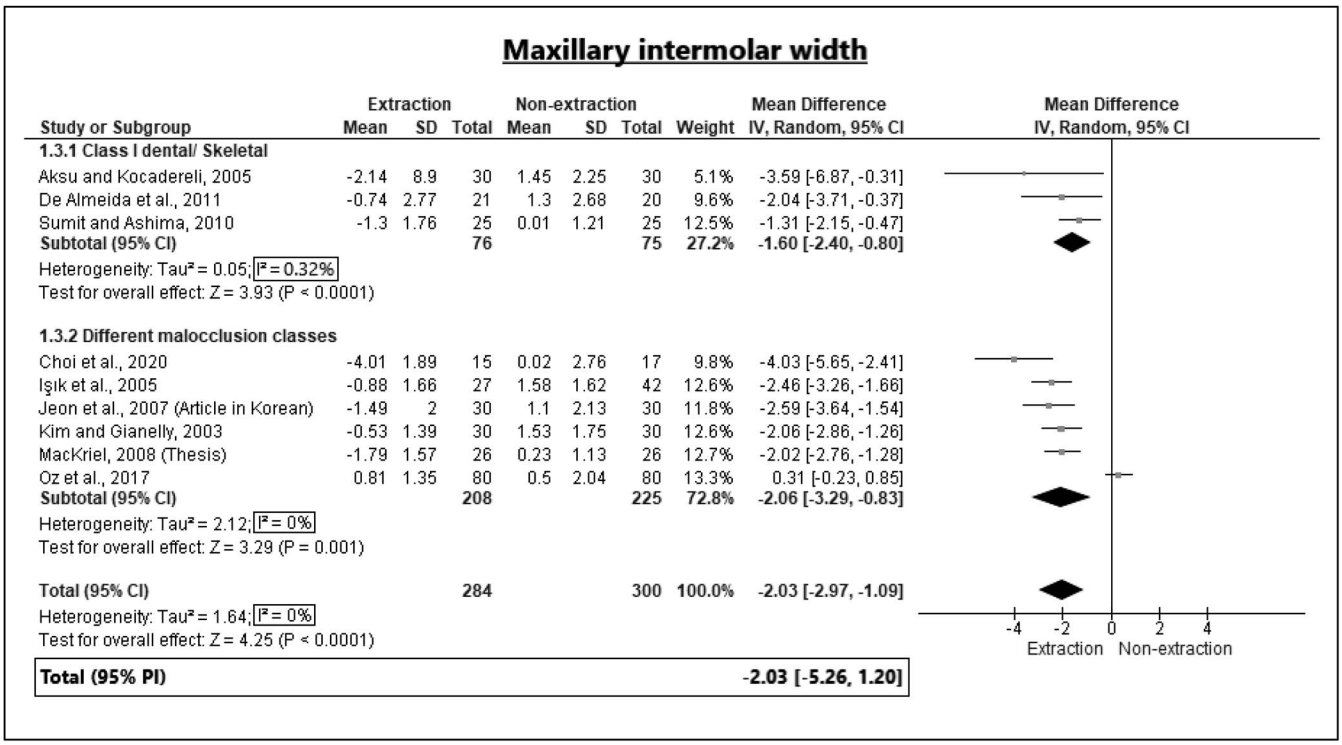


Figure 3. Forest plot, maxillary and mandibular intercanine width.

in relation to the E plane between 15 and 25 years of age. However, it was not possible to include an age threshold criterion based on these limits as raw age data were not provided and only mean age reported, a confounding variable of broad age range, increasing indirectness of the results.

There is conflicting evidence in the literature regarding arch width changes, with meta-analysis showing



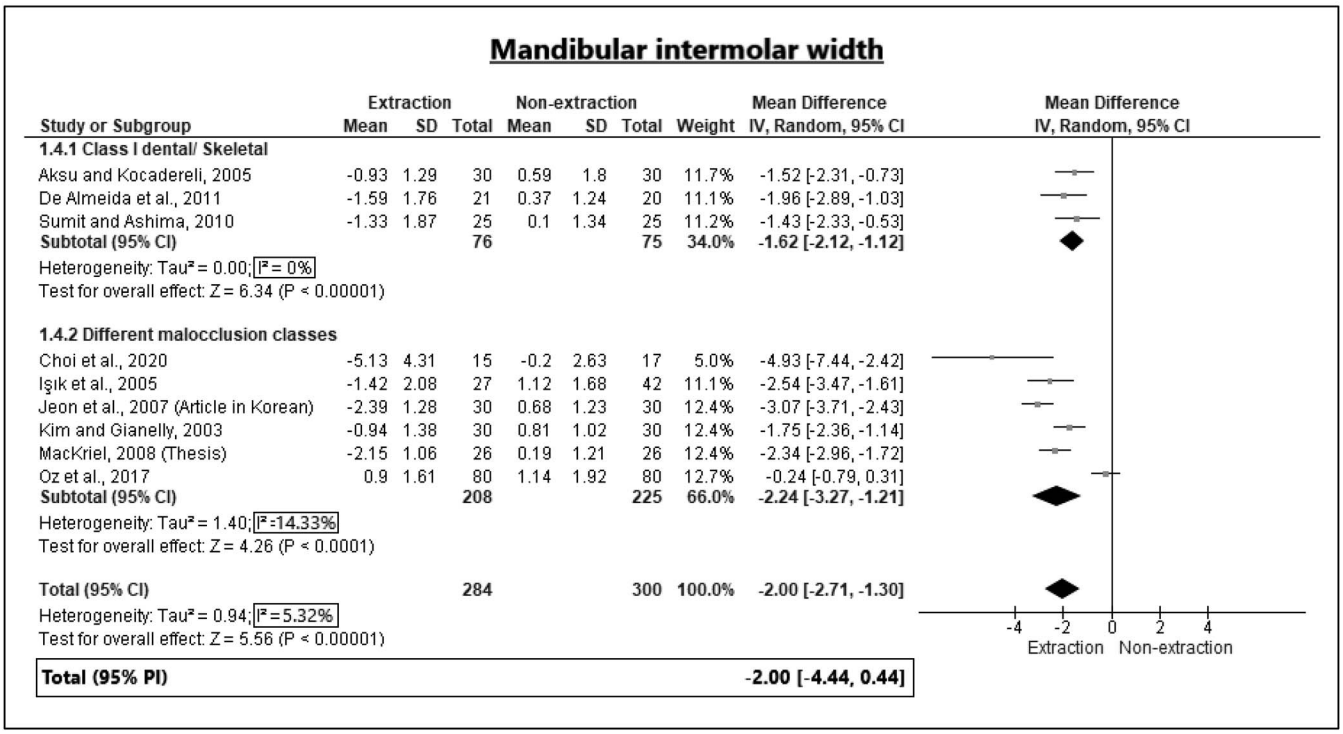
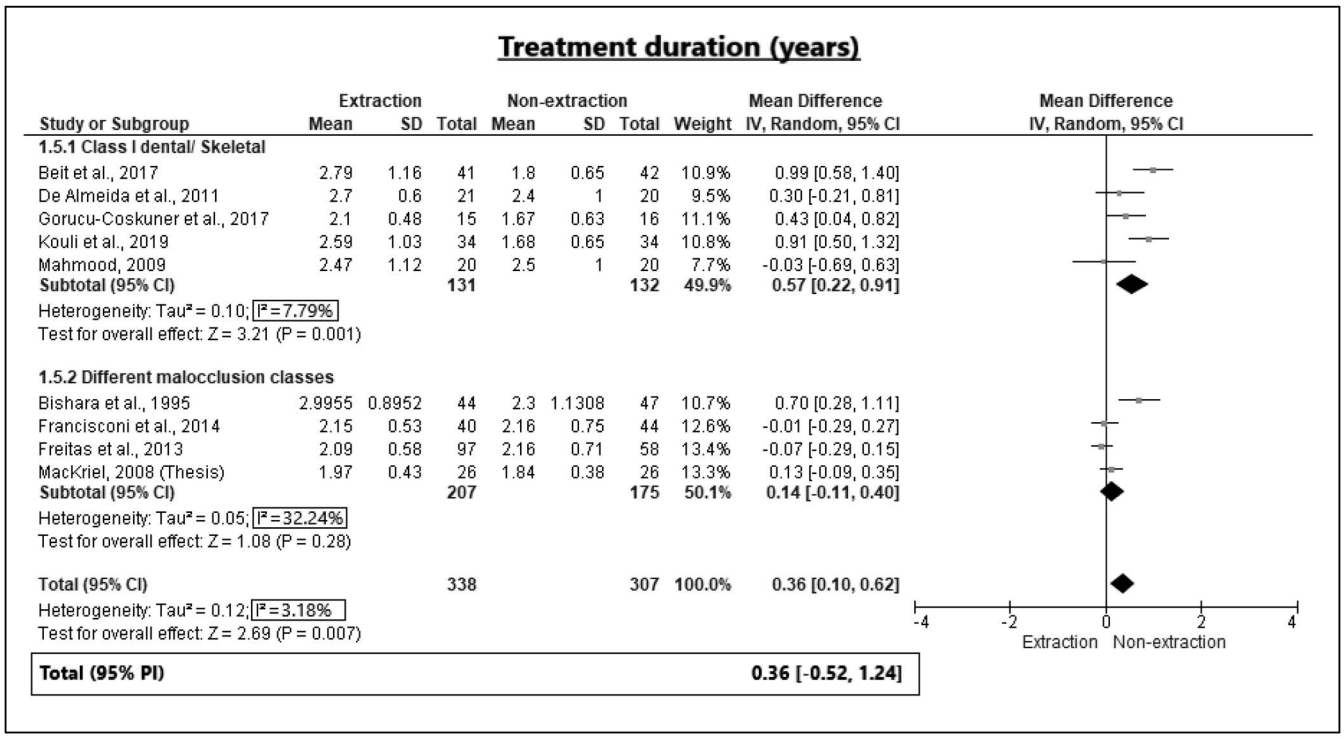


Figure 4. Forest plot, maxillary and mandibular intermolar width.

significant increase in mandibular intercanine width in the nonextraction group and no difference regarding maxillary intercanine width. A possible explanation might be related to greater variability in maxillary arch form, whereas mandibular arch form is more influenced by the soft tissue environment, meaning that the need to generate space for alignment in nonextraction treatment leads to mandibular arch width changes



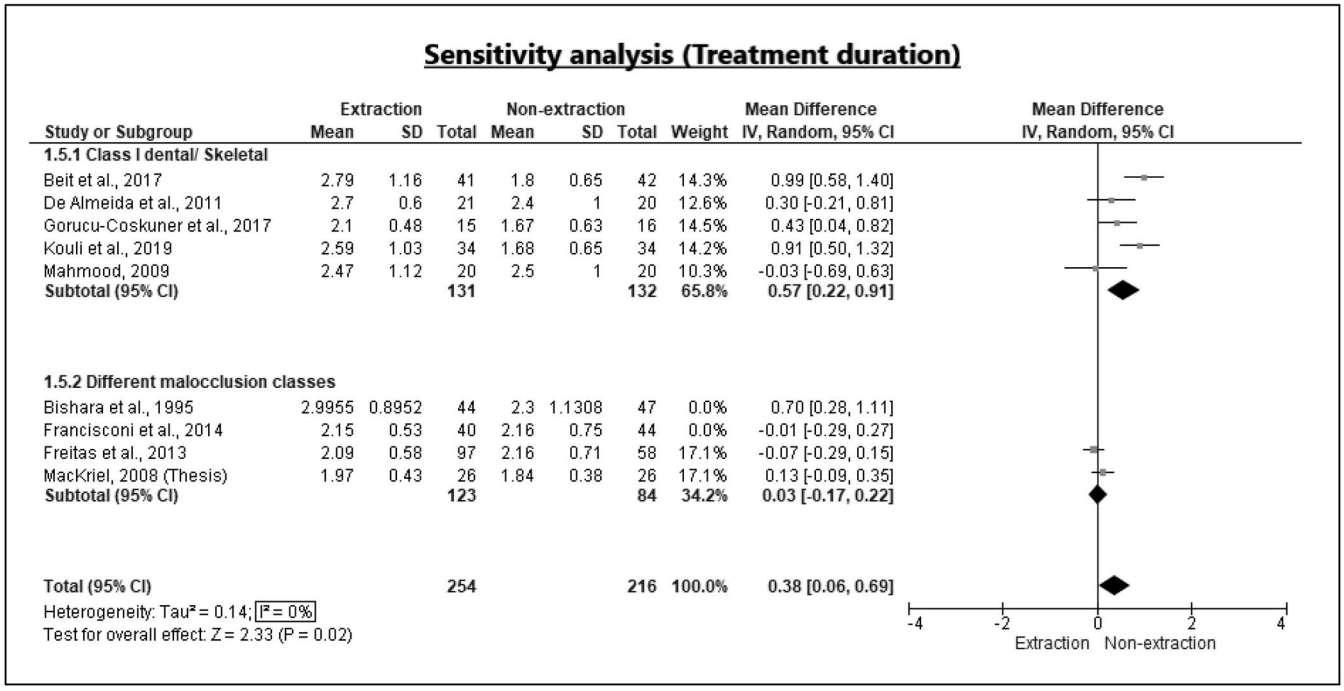


Figure 5. Forest plot and sensitivity analysis, treatment duration.

while the mandibular arch form is maintained in extraction treatment. The movement of posterior teeth mesially into narrower areas of the dental arch is the cause for maxillary and mandibular intermolar width decrease in the extraction group.[44]

The reported results of profile changes, ABO-OGS and stability should be carefully interpreted because of the imprecision of the results due to small sample sizes. Where two studies were included, no meta-analysis was undertaken, as with two studies and in the presence of

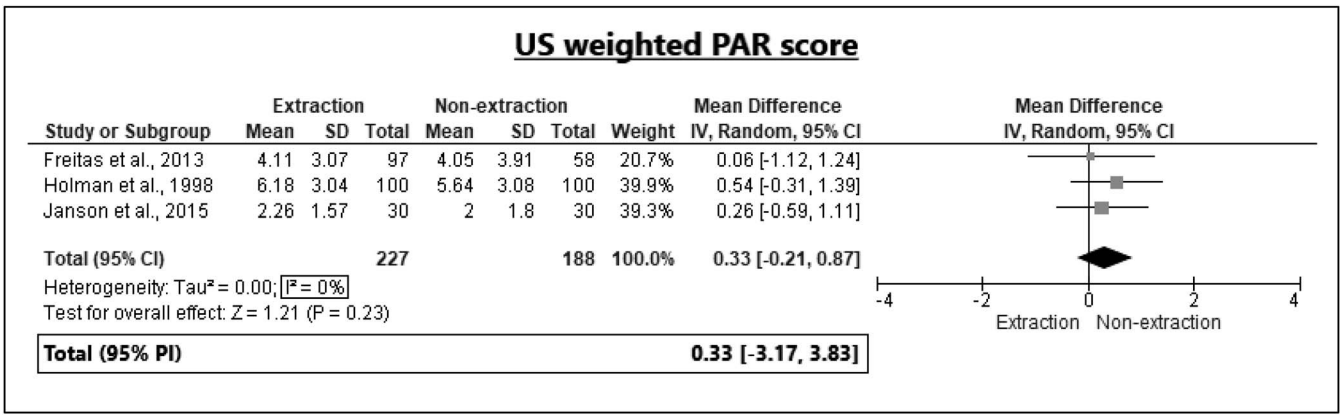


Figure 6. Forest plot, US-weighted PAR score. PAR indicates peer assessment rating.

heterogeneity, confidence intervals based on normal quantiles are not recommended.[60]

For profile changes, the current narrative synthesis indicated retraction of upper and lower lips relative to the E plane with four first premolar extraction, matching the findings of the meta-analysis by Konstantonis et al.[12] However, as ethnic differences were reported, indicating potential additional confounding factors for this outcome, no meta-analysis was undertaken.

Four first premolar extraction took an average of 0.36 years longer to complete in comparison to nonextraction treatment. This might be reasonable to assume as it allows for the time to complete space closure in extraction treatment, although it could also be related to more complex cases being treated with extractions. This difference is one of clinical significance for clinicians and patients. The evidence for this finding was graded as low certainty compared with most other outcomes, which were very low certainty.

Conflicting results on occlusal outcomes were reported in this review. No eligible studies were found for UK-weighted PAR score and no significant difference was reported with US-weighted PAR score. However, it should be noted that end treatment data were pooled rather than percentage improvement for ease of comparison with ABO-OGS.

The results of this review on smile aesthetics indicated no difference between the two treatment approaches in four different smile parameters. Furthermore, Işiksal et al.[40] stated that inadequate torque expression can affect smile aesthetics regardless of treatment modality.

There was no clear consensus whether four first premolar extraction or nonextraction treatment would provide greater posttreatment stability, as only two studies were included with conflicting results.

In summary, this review found low certainty evidence for a clinically significant difference in treatment duration. However, there are debatable clinical implications of differences found in arch width changes and no differences in occlusal outcomes and smile parameters. The decision whether to extract or not is very situational. Ruellas et al.[61] stated that clinicians should be aware of factors such as compliance, tooth-arch discrepancy, cephalometric discrepancy, facial profile, growth, anteroposterior relationships, dental asymmetry, and pathology in decision making.

In the context of existing data, more robust evidence for changes in outcome between extraction and nonextraction treatment approaches is still needed. However, this is an almost impossible aim, as one of the reasons for the lack of RCTs on this topic is ethical, with patient recruitment dilemmas of randomizing these treatments. Alternatively, high-quality observational studies may be most appropriate and a suggested protocol has been made available recently.[24]

Limitations

- Studies included in quantitative synthesis were retrospective in nature except one prospective randomized trial with high level of bias. The limitation due to the decision to include observational studies has been discussed in a Cochrane review[62] with little evidence for significant effect estimate differences between observational studies and RCTs. However, it is important to consider the level of heterogeneity in meta-analyses of RCTs or observational studies with control for confounding in observational studies.
- There is a possible source of bias due to exclusion of studies with incomplete dataset reporting, but exclusion provides greater confidence in results rather than imputation, with the same for studies with more than one error in sample size or treatment data.
- As malocclusion classes were not studied individually, this baseline characteristic caused an increase

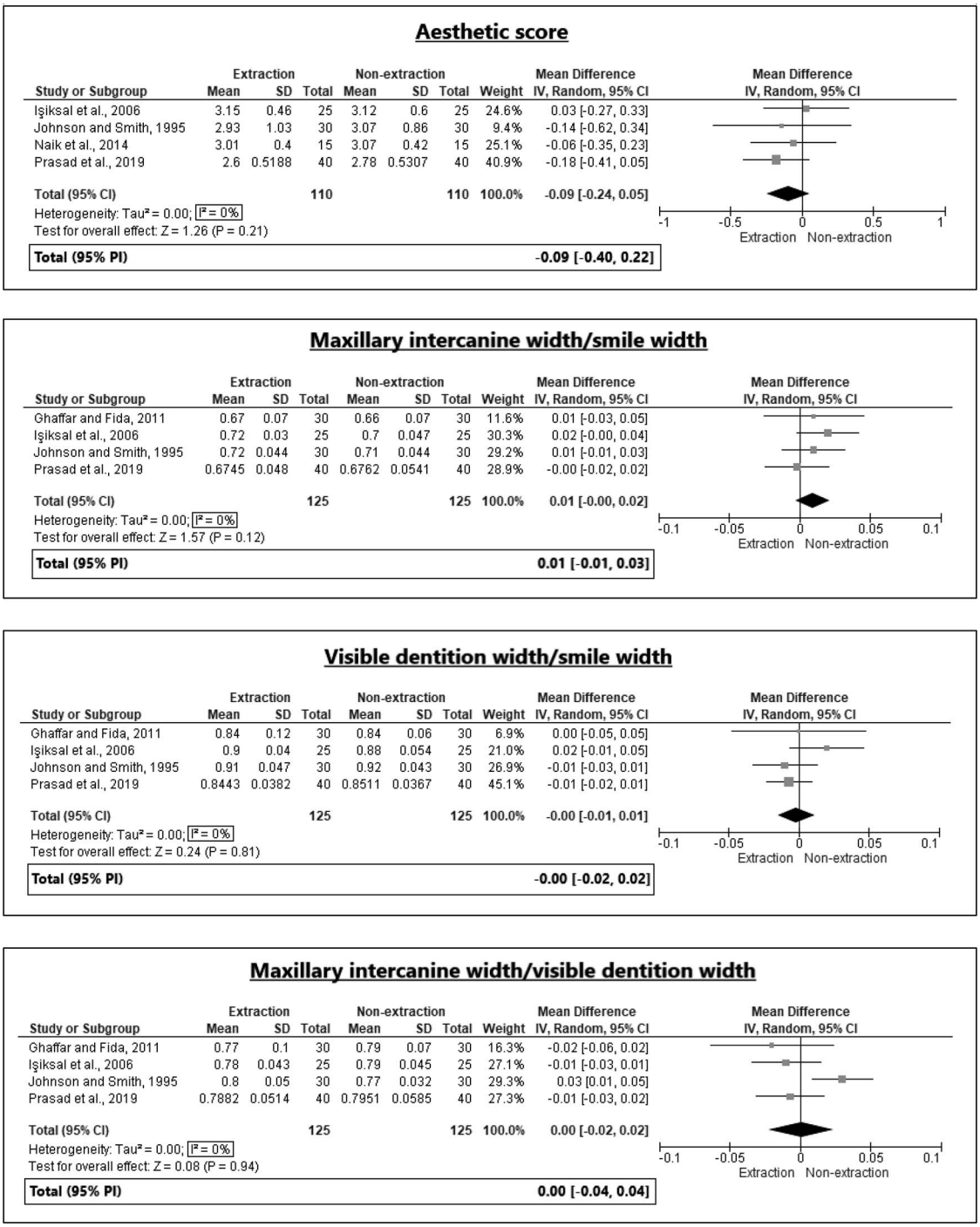
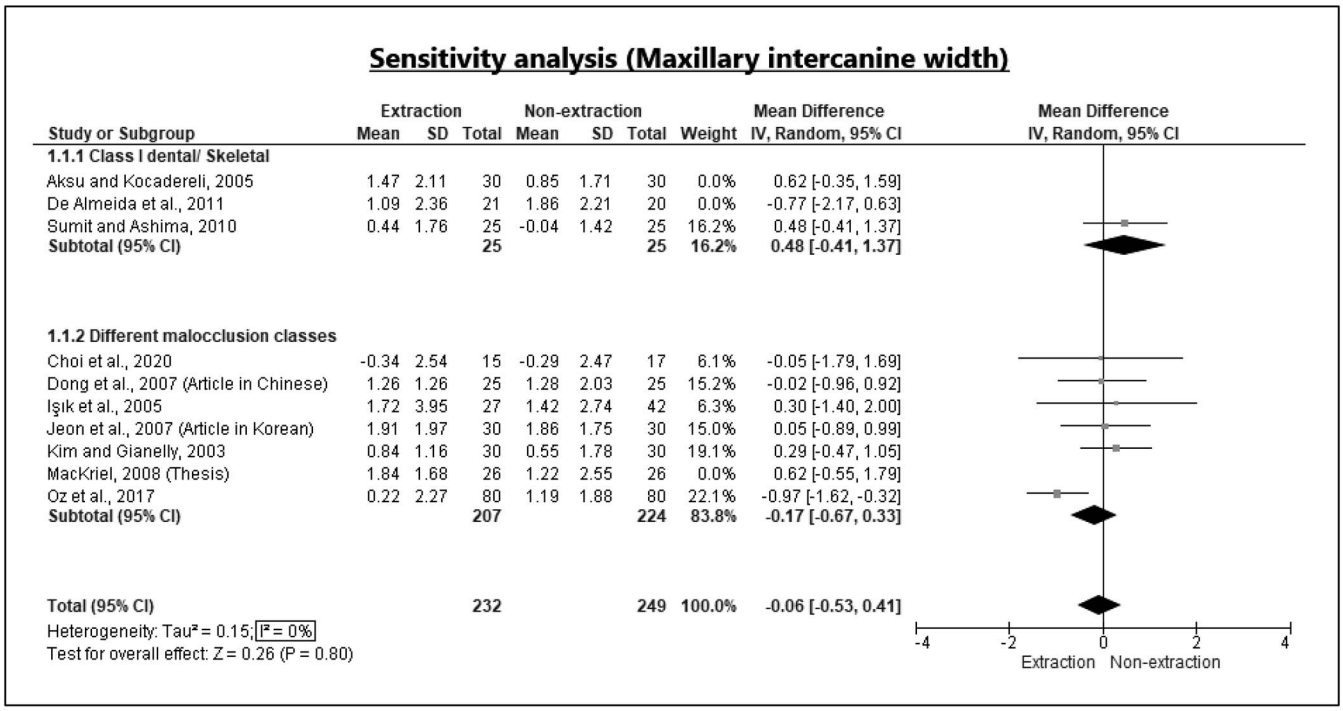


Figure 7. Forest plot, smile aesthetics.

in heterogeneity when studies were included for the meta-analysis. Clinical heterogeneity was controlled by limiting the intervention group to four first premolar extraction only and subgroup analyses were carried out where possible.

- As different ages were pooled on the same estimative, without a subgroup analysis, this baseline characteristic caused an increase in the degree of indirectness of the results.
- The subjective nature of the aesthetic score, especially since this was assessed by various raters in the different studies, the authors considered this as a possible source of heterogeneity due to the associated observational bias in the meta analysis.[17]
- Retention regimen might be a confounding factor related to stability while poor oral hygiene, poor patient compliance, and experience of the operator are factors related to treatment duration.[63]



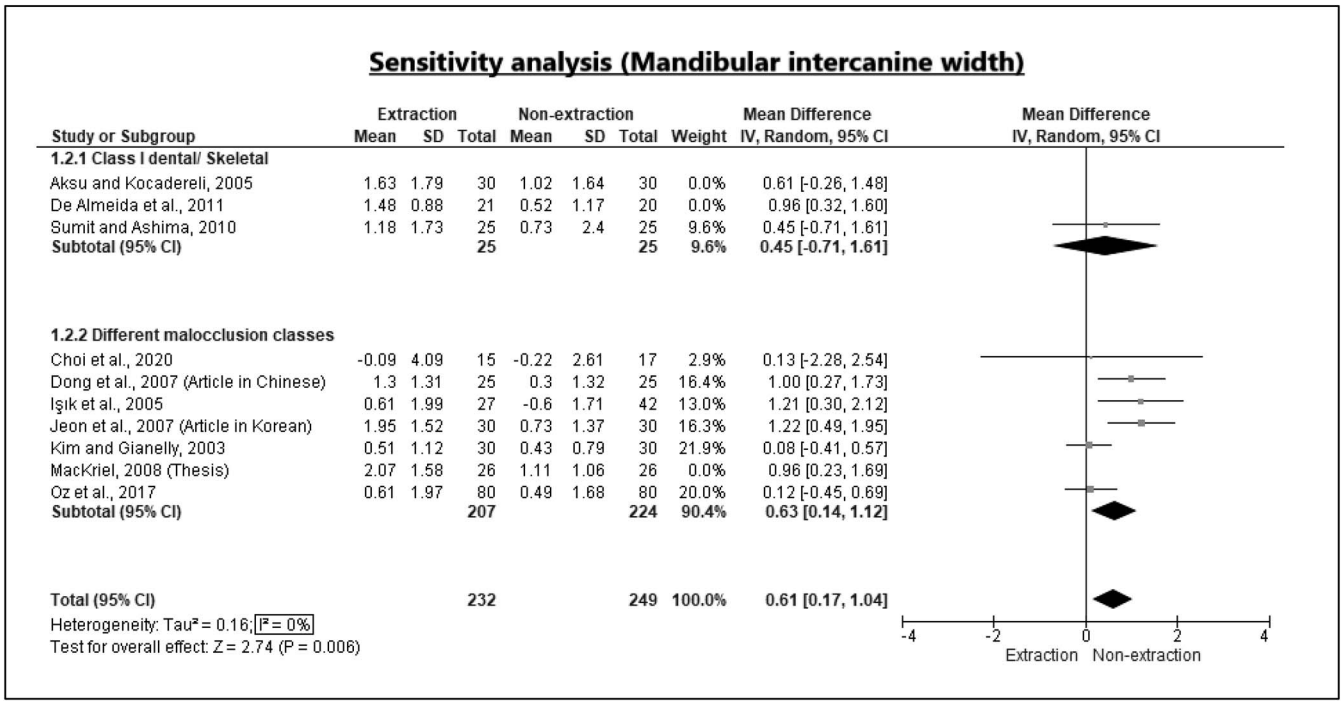


Figure 8. Sensitivity analysis, maxillary and mandibular intercanine width.

- There is an imprecision up to ±0.1 in recalculated data of included studies, which may affect the results.
- Non-English studies were translated using Google Translate, which might present a possible source of inaccuracy.
- GRADE does not allow inclusion of multiple study designs per outcome, a recognized limitation of GRADEpro. For outcomes reporting RCT and observational studies, changing study design in certainty assessment did not affect overall certainty.

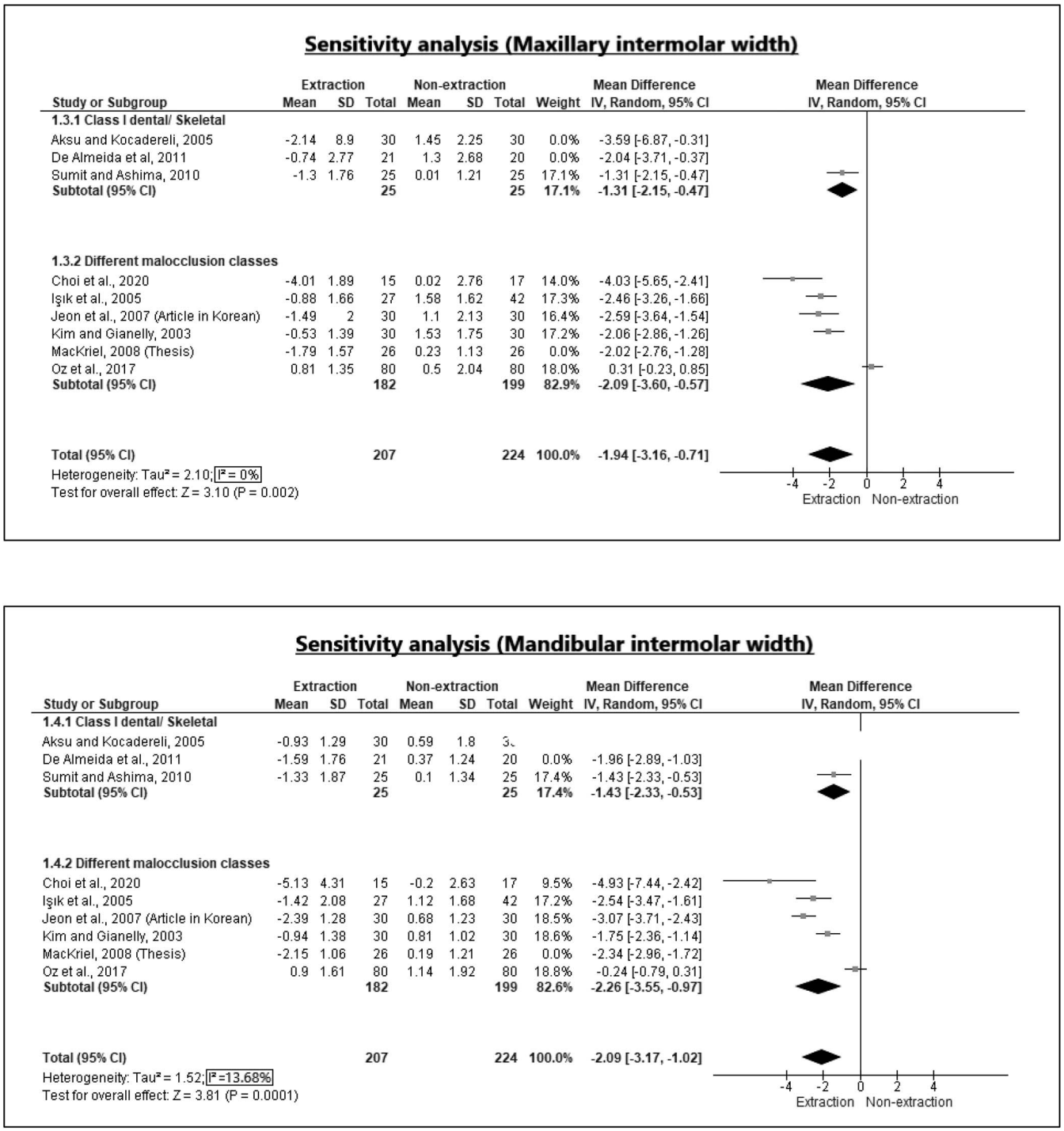


Figure 9. Sensitivity analysis, maxillary and mandibular intermolar width.

CONCLUSIONS

- Low level evidence indicates clinically significant, shorter treatment duration in the nonextraction group compared to four first premolar extraction.
- Low level evidence indicates mandibular intercanine width increase with nonextraction treatment and mandibular interfirst molar width decrease in the four first premolar extraction group.
- Very low evidence indicates no significant difference in maxillary intercanine width between the extraction and nonextraction groups and decrease of maxillary interfirst molar width with four first premolar extraction.

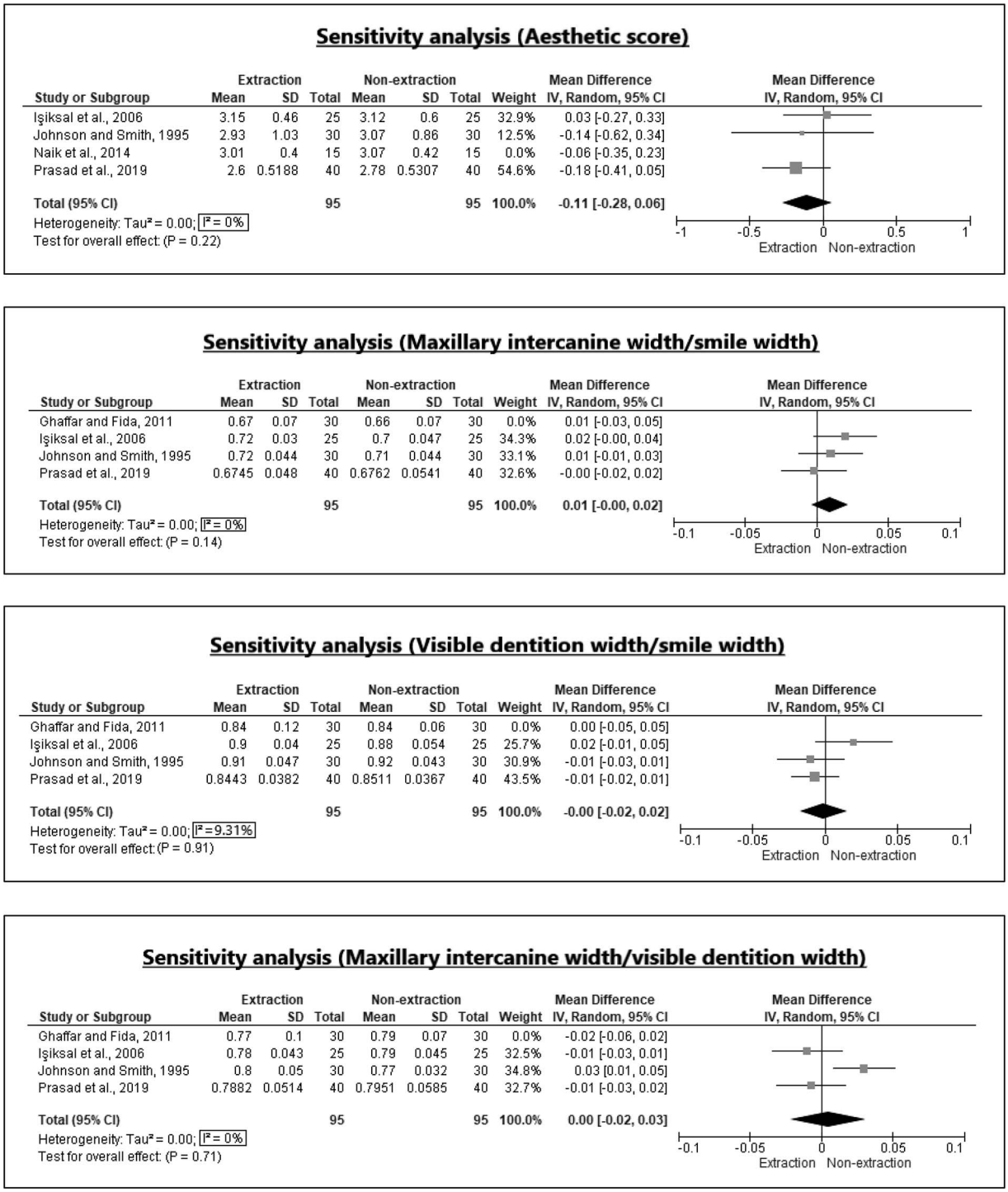


Figure 10. Sensitivity analysis, smile aesthetics.

- Very low level evidence indicates retraction of upper and lower lips-E plane in the four first premolar extraction group.
- Very low level evidence indicates no significant difference regarding US PAR score and posttreatment smile aesthetics.

ACKNOWLEDGMENTS

We would like to gratefully acknowledge Dr. Gavin Revie, Research Methodologist and Research Integrity Lead, School of Dentistry, University of Dundee, for his assistance with the meta-analysis.

Table 7. GRADE

	Certainty Assessment				
Outcome of Interest	Study Design	Risk of Bias	Inconsistency	Indirectness	Imprecision
Maxillary intercanine width	9 observational studies+ 1RCT	serious[a]	not serious[b]	serious[c]	serious[d]
Mandibular intercanine width	9 observational studies+ 1RCT	serious[a]	not serious[b]	serious[c]	not serious[f]
Maxillary intermolar width	8 observational studies+ 1RCT	serious[g]	serious[h]	serious[c]	not serious[i]
Mandibular intermolar width	8 observational studies+ 1RCT	serious[g]	not serious[b]	serious[c]	not serious[j]
(UL- E plane): Extraction: 58, nonextraction: 61	3 observational studies	serious[k]	not serious[l]	serious[c]	very serious[m]
(LL- E plane): Extraction: 98, nonextraction: 101	5 observational studies	serious[n]	not serious[l]	serious[c]	very serious[m]
Treatment duration	8 observational studies+ 1RCT	serious[o]	not serious[b]	serious[c]	not serious[p]
UK weighted PAR score US weighted PAR score	0 studies 3 observational studies	serious[q]	not serious[b]	serious[c]	serious[r]
(ABO-OGS): Extraction: 45, nonextraction: 50	2 observational studies	serious[s]	serious[t]	serious[c]	very serious[m]
Esthetic score	4 observational studies	serious[u]	not serious[b]	serious[c]	very serious[v]
Maxillary intercanine width/smile width	4 observational studies	serious[w]	not serious[b]	serious[c]	very serious[v]
Visible dentition width/smile width	4 observational studies	serious[w]	not serious[b]	serious[c]	very serious[v]
Maxillary intercanine width/visible dentition width	4 observational studies	serious[w]	not serious[b]	serious[c]	very serious[v]
(Maxillary anterior alignment): Extraction: 137, nonextraction: 102	2 observational studies	serious[x]	serious[t]	serious[c]	very serious[m]
(Mandibular anterior alignment): Extraction: 137, nonextraction: 102	2 observational studies	serious[x]	serious[t]	serious[c]	very serious[m]

[a] Data extracted from five studies with serious concerns regarding selection of participants, four studies with moderate concerns regarding measurement of outcomes and one randomized controlled trial (RCT) of high risk of bias with concerns regarding allocation concealment and no information regarding blinding.

[b] A random effects model was used; I-sq was low according to the rule of thumb. Confidence intervals overlap.

[c] Different age pooled on the same estimate without a subgroup analysis for adolescents or adults.

[d] Total number of patients is equal to 634 patients. Imprecise results due to wide confidence interval.

[e] No risk of publication bias as different sources were searched including key databases and grey literature.

[f] Confidence interval does not cross the null effect line, indicating significant increase in mandibular intercanine width with nonextraction group. A total of 634 participants included.

[g] Data extracted from four studies with serious concerns regarding selection of participants, four studies with moderate concerns regarding measurement of outcomes, and one RCT of high risk of bias with concerns regarding allocation concealment and no information regarding blinding.

[h] A random effects model was used; I-sq was low according to the rule of thumb. Oz et al.[49] not overlapping on confidence intervals.

[i] Confidence interval does not cross the null effect line, indicating significant decrease in maxillary intermolar width with four first premolar extraction group. A total of 584 participants included.

[j] Confidence interval does not cross the null effect line, indicating significant decrease in maxillary intermolar width with four first premolar extraction group. A total of 584 participants included.

[k] Data extracted from three studies with serious concerns regarding selection of participants.

[l] Results seem quite consistent.

[m] Narrative synthesis was conducted; estimates are not precise and small sample size.

[n] Data extracted from four studies with serious concerns regarding selection of participants and one study with moderate concerns regarding measurement of outcomes.

[o] Data extracted from six studies with serious concerns regarding selection of participants, two studies with serious concerns regarding selection of the reported results, and one RCT of high risk of bias with concerns regarding allocation concealment and no information regarding blinding.

[p] Confidence interval does not cross the null effect line, indicating significant shorter treatment duration in nonextraction group. A total of 645 participants included.

[q] Data extracted from three studies of serious risk of bias. Serious concerns regarding selection of participants in two studies and selection of reported results in one study.

[r] Total number of patients is equal to 415 patients, which is considered to be acceptable according to the rule of thumb. However, confidence interval is wide.

[s] Data extracted from one study with serious concerns regarding selection of participants and one study of moderate concerns regarding measurement of outcomes.

[t] Inconsistent results. Two studies included with conflicting evidence.

[u] Data extracted from four studies of serious risk of bias. Serious concerns regarding outcome measurement, being a subjective parameter.

[v] Small sample size.

[w] Data extracted from four studies of serious risk of bias.

[x] Data extracted from two studies of serious risk of bias with serious concerns regarding selection of participants.

Table 7. Extended

Certainty Assessment	No. of Patients		Effect		
Other Considerations	Extraction	Nonextraction	Relative (95% CI)	Absolute (95% CI)	Certainty
none[e]	309	325	–	MD 0.02 higher (0.38 lower to 0.43 higher)	⨁◯◯◯ Very low
none[e]	309	325	–	MD 0.68 higher (0.36 higher to 0.99 higher)	⨁⨁◯◯ Low
none[e]	284	300	–	MD 2.03 lower (2.97 lower to 1.09 lower)	⨁◯◯◯ Very low
none[e]	284	300	–	MD 2 lower (2.71 lower to 1.3 lower)	⨁⨁◯◯ Low
none[e]					⨁◯◯◯ Very low
none[e]					⨁◯◯◯ Very low
none[e]	338	307	–	MD 0.36 higher (0.1 higher to 0.62 higher)	⨁⨁◯◯ Low
none[e]	227	188	–	MD 0.33 higher (0.21 lower to 0.87 higher)	⨁◯◯◯ Very low
none[e]					⨁◯◯◯ Very low
none[e]	110	110	–	MD 0.09 lower (0.24 lower to 0.05 higher)	⨁◯◯◯ Very low
none[e]	125	125	–	MD 0.01 higher (0 to 0.02 higher)	⨁◯◯◯ Very low
none[e]	125	125	–	MD **0** (0.01 lower to 0.01 higher)	⨁◯◯◯ Very low
none[e]	125	125	–	MD **0** (0.02 lower to 0.02 higher)	⨁◯◯◯ Very low
none[e]					⨁◯◯◯ Very low
none[e]					⨁◯◯◯ Very low

This paper was based on a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science (Orthodontics) at University of Dundee.

No external funding was received to conduct this research. No conflicts of interest were declared.

SUPPLEMENTAL DATA

Appendices 1–6 are available online.

REFERENCES

1. Pollock HC. The extraction debate of 1911 by Case, Dewey, and Cryer. *Am J Orthod*. 1964;50(9):656–658.
2. The extraction debate of 1911 by Case, Dewey, and Cryer: discussion of Case: the question of extraction in orthodontia. *Am J Orthod*. 1964;50(11):843–851.
3. Angle EH. Treatment of Malocclusion of the Teeth. *Angle's System*. 7th ed., S.S. White dental manufacturing Co., Philadelphia, 1907.
4. Case CS. The question of extraction in orthodontia. *Am J Orthod*. 1964;50(9):660–691.
5. Begg PR. Stone age man's dentition: with reference to anatomically correct occlusion, the etiology of malocclusion, and a technique for its treatment. *Am J Orthod*. 1954;40(7):517–531.
6. Tweed CH. Indications for the extraction of teeth in orthodontic procedure. *Am J Orthod Oral Surg*. 1944;42:22–45.
7. Proffit WR. Forty-year review of extraction frequencies at a university orthodontic clinic. *Angle Orthod*. 1994;64(6):407–414.
8. Jackson TH, Guez C, Lin FC, Proffit WR, Ko CC. Extraction frequencies at a university orthodontic clinic in the 21st century: demographic and diagnostic factors affecting the likelihood of extraction. *Am J Orthod Dentofacial Orthop.* 2017;151(3):456–462.
9. Janson G, Maria FRT, Bombonatti R. Frequency evaluation of different extraction protocols in orthodontic treatment during 35 years. *Prog Orthod*. 2014;15(1):51.
10. Bowman SJ, Johnston LE. The esthetic impact of extraction and nonextraction treatments on Caucasian patients. *Angle Orthod*. 2000;70(1):3–10.
11. Boley JC, Pontier JP, Smith S, Fulbright M. Facial changes in extraction and nonextraction patients. *Angle Orthod*. 1998;68(6):539–546.
12. Konstantonis D, Vasileiou D, Papageorgiou SN, Eliades T. Soft tissue changes following extraction vs. nonextraction orthodontic fixed appliance treatment: a systematic review and meta-analysis. *Eur J Oral Sci*. 2018;126(3):167–179.

13. Little RM, Wallen TR, Riedel RA. Stability and relapse of mandibular anterior alignment-first premolar extraction cases treated by traditional edgewise orthodontics. *Am J Orthod*. 1981;80(4):349–365.
14. Rossouw PE, Preston CB, Lombard C. A longitudinal evaluation of extraction versus nonextraction treatment with special reference to the posttreatment irregularity of the lower incisors. *Semin Orthod*. 1999;5(3):160–170.
15. Almurtadha RH, Alhammadi MS, Fayed MMS, Abou-El-Ezz A, Halboub E. Changes in soft tissue profile after orthodontic treatment with and without extraction: a systematic review and meta-analysis. *J Evid-Based Dent Pract*. 2018;18(3): 193–202.
16. Burke SP, Silveira AM, Goldsmith LJ, Yancey JM, Van Stewart A, Scarfe WC. A meta-analysis of mandibular intercanine width in treatment and postretention. *Angle Orthod*. 1998;68(1):53–60.
17. Cheng HC, Wang YC, Tam KW, Yen MF. Effects of tooth extraction on smile esthetics and the buccal corridor: a meta-analysis. *J Dent Sci*. 2016;11(4):387–393.
18. Dai M Lu, Xiao M, Yu Z, Liu D Xu. [Effect of extraction and non-extraction treatment on frontal smiling esthetics: a meta-analysis]. *Shanghai Kou Qiang Yi Xue Shanghai J Stomatol*. 2015;24(4):499–504.
19. Mavreas D, Athanasiou AE. Factors affecting the duration of orthodontic treatment: a systematic review. *Eur J Orthod*. 2008;30(4):386–395.
20. Moon S, Mohamed AMA, He Y, Dong W, Yaosen C, Yang Y. Extraction vs. nonextraction on soft-tissue profile change in patients with malocclusion: a systematic review and meta-analysis. *Biomed Res Int*. 2021;2021:7751516.
21. Papageorgiou SN, Höchli D, Eliades T. Outcomes of comprehensive fixed appliance orthodontic treatment: a systematic review with meta-analysis and methodological overview. *Korean J Orthod*. 2017;47(6):401–413.
22. Swidi AJ, Griffin AE, Buschang PH. Mandibular alignment changes after full-fixed orthodontic treatment: a systematic review and meta-analysis. *Eur J Orthod*. 2019;41(6):609–621.
23. Ze-Ping L, Ting-Ting C, Gui-Feng L, Si-Cong J, Shu-Juan Z. Meta analysis of soft tissue profile repair in patients treated with and without premolar extraction. *Chin J Tissue Eng Res*. 2010;14(2):237.
24. Benson PE, Alshawy E, Fenton GD, et al. Extraction vs nonextraction of premolars for orthodontic treatment: a scoping review examining the extent, range, and characteristics of the literature. *Am J Orthod Dentofacial Orthop*. 2023;164:368–376.
25. Borenstein M, Higgins JPT, Hedges LV, Rothstein HR. Basics of meta-analysis: I2 is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017;8(1):5–18.
26. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. Chichester U.K.: John Wiley & Sons; 2009.
27. Aksu M, Kocadereli I. Arch width changes in extraction and nonextraction treatment in class I patients. *Angle Orthod*. 2005;75(6):948–952.
28. Anthopoulou C, Konstantonis D, Makou M. Treatment outcomes after extraction and nonextraction treatment evaluated with the American Board of Orthodontics objective grading system. *Am J Orthod Dentofacial Orthop*. 2014;146(6):717–723.
29. Beit P, Konstantonis D, Papagiannis A, Eliades T. Vertical skeletal changes after extraction and non-extraction treatment in matched class I patients identified by a discriminant analysis: cephalometric appraisal and Procrustes superimposition. *Prog Orthod*. 2017;18(1):44.
30. Bishara SE, Cummins DM, Jorgensen GJ, Jakobsen JR. A computer assisted photogrammetric analysis of soft tissue changes after orthodontic treatment. Part I: Methodology and reliability. *Am J Orthod Dentofacial Orthop*. 1995;107(6): 633–639.
31. Choi TH, Kim SH, Kim C, Kook YA, Larson BE, Lee NK. Changes in maximum lip-closing force after extraction and nonextraction orthodontic treatments. *Korean J Orthod*. 2020;50(2):120–128.
32. Francisconi MF, Janson G, Freitas KM, et al. Overjet, overbite, and anterior crowding relapses in extraction and nonextraction patients, and their correlations. *Am J Orthod Dentofacial Orthop*. 2014;146(1):67–72.
33. Freitas KM, Janson G, Tompson B, et al. Posttreatment and physiologic occlusal changes comparison. *Angle Orthod*. 2013;83(2):239–245.
34. Freitas BV, Rodrigues VP, Rodrigues MF, de Melo HVF, Dos Santos PCF. Soft tissue facial profile changes after orthodontic treatment with or without tooth extractions in Class I malocclusion patients: a comparative study. *J Oral Biol Craniofacial Res*. 2019;9(2):172–176.
35. Ghaffar F, Fida M. Effect of extraction of first four premolars on smile esthetics. *Eur J Orthod*. 2011;33(6):679–683.
36. Gorucu-Coskuner H, Atik, E, Kocadereli I. Effects of three different orthodontic treatment methods on the stability of mandibular incisor alignment. *J Clin Pediatr Dent*. 2017;41(6):486–493.
37. Hassan S, Shaikh A, Fida M. Esthetic impact of tooth extraction in Pakistani patients. *J Ayub Med Coll Abbottabad* 2014;26(3):263–268.
38. Holman JK, Hans MG, Nelson S, Powers MP. An assessment of extraction versus nonextraction orthodontic treatment using the peer assessment rating (PAR) index. *Angle Orthod*. 1998;68(6):527–534.
39. Işik F, Sayinsu K, Nalbantgil D, Arun T. A comparative study of dental arch widths: extraction and non-extraction treatment. *Eur J Orthod*. 2005;27(6):585–589.
40. Işiksal E, Hazar S, Akyalçin S. Smile esthetics: perception and comparison of treated and untreated smiles. *Am J Orthod Dentofacial Orthop*. 2006;129(1):8–16.
41. Janson G, Simao TM, Barros SE, Tompson BD, Freitas K. A 2-center comparison of maxillary incisor positioning with non-extraction, 2-maxillary premolar and 4-premolar extractions for Class II treatment. *Orthod WAVES*. 2015;74(4): 105–111.
42. Jeon JY, Kim SJ, Kang SG, Park YG. Change in arch width in extraction vs nonextraction treatment. *Korean J Orthod*. 2007;37(1):65–72.
43. Johnson DK, Smith RJ. Smile esthetics after orthodontic treatment with and without extraction of four first premolars. *Am J Orthod Dentofacial Orthop*. 1995;108(2):162–167.
44. Kim E, Gianelly AA. Extraction vs nonextraction: Arch widths and smile esthetics. *Angle Orthod*. 2003;73(4):354–358.
45. Konstantonis D. The impact of extraction vs nonextraction treatment on soft tissue changes in Class I borderline malocclusions. *Angle Orthod*. 2012;82(2):209–217.
46. Kouli A, Papagiannis A, Konstantoni N, Halazonetis DJ, Konstantonis D. A geometric morphometric evaluation of hard and soft tissue profile changes in borderline extraction versus non-extraction patients. *Eur J Orthod*. 2019;41(3): 264–272.

47. MacKriel EA. *Changes in Arch Dimensions After Extraction and Non-Extraction Orthodontic Treatment* [PhD thesis]. Bellville, South Africa: University of the Western Cape; 2008.
48. Naik P, Manohar MR, Shivaprakash G, Jabeen N. Smile esthetics: evaluation of differential perception among laypersons, dental professionals and orthodontics. *IOSR J Dent Med Sci*. 2014;13(1):35–43.
49. Oz AA, Oz AZ, Yazicioğlu S, Arici N, Ozer M, Arici S. Comparison of arch width changes following orthodontic treatment with and without extraction using three-dimensional models. *Niger J Clin Pract*. 2017;20(5):581–586.
50. Prasad V, Tandon P, Singh GK, Nagar A, Maurya RP. Comparison of smile esthetics after extraction and non-extraction orthodontic treatment. *IJODR*. 2018;4(4):182–189.
51. Sumit G, Ashima V. Arch width changes in canine and molar regions after extraction and non-extraction treatment. *J Pierre Fauchard Acad India Sect*. 2010;24(2):45–53.
52. Vaidya M, Jyothikiran H, Raghunath N, Pai P. Extraction versus non-extraction: a retrospective study. *Int J Orthod Rehabil*. 2018;9(1):23.
53. Xu TM, Liu Y, Yang MZ, Huang W. Comparison of extraction versus nonextraction orthodontic treatment outcomes for borderline Chinese patients. *Am J Orthod Dentofacial Orthop*. 2006;129(5):672–677.
54. Dong ZY, Liu DX, Wang TJ, Liu L. [Changes in arch width after extraction and nonextraction treatment]. *Shanghai Kou Qiang Yi Xue*. 2007;16(4):355–357 [Article in Chinese].
55. Mahmood DAD. Changes in third molar angulations after first premolar extractions during fixed orthodontic treatment. *Mustansiria Dent J*. 2009;6(4):356–360.
56. De Almeida MR, Almeida-Pedrin RRD. A prospective clinical trial of arch width changes evaluated through dental casts in consecutively treated non-extraction and extraction cases. *Dentistry*. 2011;01(02).
57. Kaji AH, Schriger D, Green S. Looking through the retrospectoscope: reducing bias in emergency medicine chart review studies. *Ann Emerg Med*. 2014;64(3):292–298.
58. Bishara SE, Jakobsen JR, Treder J, Nowak A. Arch width changes from 6 weeks to 45 years of age. *Am J Orthod Dentofacial Orthop*. 1997;111(4):401–409.
59. Bishara SE, Jakobsen JR, Hession TJ, Treder JE. Soft tissue profile changes from 5 to 45 years of age. *Am J Orthod Dentofacial Orthop*. 1998;114(6):698–706.
60. Friede T, Röver C, Wandel S, Neuenschwander B. Meta-analysis of two studies in the presence of heterogeneity with applications in rare diseases. *Biom J Biom Z*. 2017;59(4):658–671.
61. Ruellas AC de O, Ruellas RM de O, Romano FL, Pithon MM, Santos RL dos. Tooth extraction in orthodontics: an evaluation of diagnostic elements. *Dent Press J Orthod*. 2010;15:134–157.
62. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev*. 2014;2014(4):MR000034.
63. Skidmore KJ, Brook KJ, Thomson WM, Harding WJ. Factors influencing treatment time in orthodontic patients. *Am J Orthod Dentofacial Orthop*. 2006;129(2):230–238.