European Journal of Orthodontics, 2015, 418–434
doi:10.1093/ejo/cju071
Advance Access publication 14 November 2014

Systematic Review

Treatment effects of removable functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis

Vasiliki Koretsi*, Vasileios F. Zymperdikas, Spyridon N. Papageorgiou***,****,***** and Moschos A. Papadopoulos******

*Department of Orthodontics, School of Dentistry, University Medical Centre Regensburg, Germany, **Dental Unit, Medical Company, 71st Airmobile Brigade, Nea Santa, Greece, ***Department of Orthodontics, ****Department of Oral Technology, and *****Clinical Research Unit 208, University of Bonn, Germany, ******Department of Orthodontics, School of Health Sciences, Faculty of Dentistry, Aristotle University of Thessaloniki, Greece.

Correspondence to: Moschos A. Papadopoulos, Department of Orthodontics, School of Health Sciences, Faculty of Dentistry, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece. E-mail: mikepap@dent.auth.gr

Summary

Objective: To assess the treatment effects of removable functional appliances (RFAs) in treated versus untreated patients with Class II malocclusion by means of lateral cephalometric radiographs.
Search methods: Unrestricted electronic search of 18 databases and manual searches up to October 2013.
Selection criteria: Prospective randomized and non-randomized controlled trials reporting on cephalometric angular measurements of Class II patients treated with RFAs and their matched controls.
Data collection and analysis: Skeletal, dental, and soft tissue changes were annualized and stratified to short- and long-term effects. Methodological limitations were evaluated with the Cochrane Risk of Bias tool and the Downs and Black checklist. Mean differences (MDs) with their 95% confidence intervals (CIs) were calculated from random-effects meta-analyses. Patient- or appliance-related subgroup analyses and sensitivity analyses were performed with mixed-effects models.
Results: Seventeen studies were included (1031 patients; mean age: 10.6 years), with most of them originating from university clinics and reporting short-term effects (directly after the removal of RFAs). Treatment was associated with minimal reduction of SNA angle (11 studies, MD = –0.28 degree/year, 95% CI: –0.44 to –0.12 degree/year), minimal increase of SNB angle (11 studies, MD = 0.62 degree/year, 95% CI: 0.36–0.88 degree/year), and small decrease of ANB angle (10 studies, MD = –1.14 degree/year, 95% CI: –1.52 to –0.77 degree/year) compared to untreated Class II patients. RFAs caused significant dentoalveolar changes (predominantly retroclination of the upper incisors) and significant soft tissue changes. Skeletal changes were more pronounced with the Twin Block appliance. Various patient- or appliance-related factors influenced the results of the subgroup analyses, while the sensitivity analyses indicated robustness. Existing evidence was inadequate to assess the long-term effectiveness of RFAs.
Conclusions: The short-term evidence indicates that RFAs are effective in improving Class II malocclusion, although their effects are mainly dentoalveolar, rather than skeletal.

Introduction

Rationale

Class II malocclusion appears with high frequency and constitutes a significant proportion of patients seeking orthodontic treatment (1–3). Various factors can contribute to the development of Class II malocclusion and their differential diagnosis can help in the selection of the most appropriate treatment approach. Among these factors, mandibular retrognathism shows a prevailing frequency (4, 5). In these cases, the use of functional appliances for mandibular growth stimulation has been an appealing perspective in growing patients.

Although early animal studies demonstrated that a true stimulation of mandibular growth is feasible with the use of functional appliances (6–9), subsequent human research did not fully confirm this. While some researchers reported favourable treatment effects on mandibular growth, either as an increase in the length of the mandible (10–12) or as an effective growth of the condyle (13–16), others reported that the effect of functional treatment on the mandible was not significant (17–19). Similarly, for the maxilla, some researchers found a restriction effect on the maxilla caused by functional appliances (20–22), while others disputed this (23, 24). Moreover, dentoalveolar components of functional treatment might account equally or even more than the skeletal effects to the successful treatment outcome (25, 26).

Functional treatment can be carried out either with removable functional appliances (RFAs) or with fixed functional appliances (FFAs). An essential difference between them is the factor of patient compliance, which can strongly influence the effectiveness of functional treatment (27). Therefore, there is need to differentiate between RFAs and FFAs in order to evaluate their effectiveness, as well as to identify the factors that could influence the respective treatment outcomes. Previous reviews have assessed various types of functional appliances (19, 28–36), but most of them focused on a single appliance or were associated with serious methodological flaws.

Objectives

Aim of this study was to summarize current evidence exclusively from randomized controlled trials (RCTs) and prospective controlled clinical trials (pCCTs) evaluating the effectiveness of RFAs in the treatment of Class II malocclusion by means of lateral cephalometric radiographs in comparison with untreated individuals and to identify any parameters influencing the treatment outcome.

Materials and methods

Protocol and registration

The protocol for this systematic review was made *a priori* based on Cochrane Handbook for Systematic Reviews of Interventions 5.1.0 (37) and is available upon request (no registration was made). This systematic review is reported according to the PRISMA statement (38) and its extension for abstracts (39).

Information sources and search

Electronic databases were systematically searched up to November 2011 and updated in October 2013. In order to avoid omission of relevant studies, a broad search strategy was undertaken that covered the whole spectrum of functional appliances and was then carefully restricted on RFAs. MESH terms and relative keywords were used accordingly for each database (Supplementary Table 1). No limitations were applied regarding publication year, status or language. The reference lists of included articles and relevant reviews were checked. Grey literature was searched through appropriate databases and registers. Authors were contacted when necessary for additional data or clarifications. The search was conducted by two review authors independently (VK and VFZ).

Eligibility criteria and study selection

Inclusion and exclusion criteria were determined *a priori* (Table 1). A study was judged as eligible when at least one treatment arm with an RFA was included, while none of the exclusion and all of the inclusion criteria were fulfilled. After removal of duplicates, articles were screened on the basis of title, abstract, and full text. When the decision on the basis of title and abstract was inconclusive, the full-text article was acquired. Additional reports of the same trial were grouped together. When an identical study was published in more than one language, the English publication was considered.

Data collection process and data items

Data extraction was performed independently by the same two review authors in a predetermined and piloted collection form. Any disagreement was resolved by discussion with the last author (MAP). In order to evaluate effectiveness of Class II treatment with RFAs, angular measurements from lateral cephalometric analyses were considered as valid outcomes, as radiographic magnification does not affect their precision (40, 41). Linear measurements were excluded because of their susceptibility to magnification bias (40, 41). As in many instances, various terms of the same cephalometric variables were reported by authors, all equivalent names of the same variable were grouped together, and only one term was used throughout the article (Supplementary Table 2).

If two or more included studies reported the same cephalometric variable, data were extracted and classified into skeletal (sagittal and vertical), dental, and soft tissue variables. Reported outcomes were stratified according to the time of evaluation in effects: 1. after the removal of RFAs, 2. after the completion of the subsequent fixed-appliance treatment (when applicable), and 3. after the retention phase.

The factors for the subgroup analyses were selected *a priori* to examine their possible influence on the RFAs, if five or more eligible studies could contribute to the analysis of this factor. These factors were classified as patient-related (i.e. specific characteristics of the patients) and appliance-related (i.e. specific characteristics of the design of the appliance or the treatment plan). The patient-related factors included 1. patients' gender (males or females), 2. skeletal growth stage (pre-peak, peak, or post-peak), and 3. patients' growth pattern (horizontal, vertical, or average). The discrimination among various maturation stages was based on the cervical vertebral maturation index or on hand-wrist radiographs, which are considered reliable methods of assessing skeletal maturity (42, 43). Due to the fact that the chronological or the dental age are not valid methods for identifying skeletal age (44, 45), studies reporting only on them were not included in the assessment of skeletal maturity. If the growth pattern was used as a prerequisite in the inclusion criteria of the original study, it was classified accordingly. In cases where no statement with regard to growth pattern existed, this was classified by taking into account the mean values of either the FH-ML or SN-ML angles, as provided by the baseline characteristics in the original studies.

Appliance-related factors included 1. appliance used (i.e. exact type and design of RFA used), 2. mode of action (bite-jumping or periosteal pull), 3. number of appliance components (one- or two-piece), and 4. construction bite (single step or stepwise mandibular

Table 1. Eligibility criteria used for the study selection.

Category	Inclusion criteria	Exclusion criteria
Participant characteristics	Studies on human patients with Class II malocclusion of any age or gender	Patients with craniofacial syndromes and/or cleft lip palate Patients with temporomandibular joint disorders Animal studies
Intervention	Orthodontic treatment with removable functional appliances	Patients with Class II malocclusion treated with extractions, Class II elastics, orthognathic surgery, or fixed functional appliances
Comparison	Untreated patients with Class II malocclusion matched for age and gender	Studies without an untreated Class II control group
Outcome	Skeletal, dentoalveolar, and soft tissue variables on lateral cephalometric radiographs	Any other outcome reported
Principal outcome measures	Studies providing angular skeletal, dentoalveolar, and soft tissue cephalometric measurements	Studies providing only linear cephalometric measurements Electromyographic evaluation Evaluation employing 3D imaging techniques Cost–benefit analysis
Study design	Randomized controlled clinical trials Prospective controlled clinical trials	Unsupported opinion of expert Editor' s choices Replies to the author/editor Interviews Commentaries Books'/conferences' abstracts Summaries Cross-sectional surveys Case series without a control Case reports or reports of cases Case–control observational studies Cohort studies Retrospective clinical trials Narrative reviews* Systematic reviews* Meta-analyses*

*After checking the reference lists for relevant articles.

advancement). After the data extraction, no included study was found to report treatment effects separately for boys and girls, except for one (46). Therefore, contrary to the review protocol, the patient gender ratio (male patients/female patients) was used to give an insight into the influence of gender distribution on the outcome but is not analysed in detail.

Risk of bias in individual studies

The risk of bias of included RCTs was assessed with the Cochrane Collaboration's risk of bias tool (37). The following domains were considered: 1. random sequence generation, 2. allocation sequence concealment, 3. blinding of outcome assessment, 4. incomplete outcome data, 5. selective outcome reporting, and 6. other sources of bias. For all included trials, the risk of bias for each domain was judged as 'low risk', 'high risk', or 'unclear risk'. Each RCT was assigned an overall risk of bias in terms of 'low risk' (low for all key domains), 'high risk' (high for ≥1 key domain), and 'unclear risk' (unclear for ≥1 key domain).

The risk of bias of the included pCCTs was assessed with a modified Downs and Black (47) checklist. In general, the criteria were grouped in five main domains: 1. reporting, 2. external validity, 3. internal validity—bias, 4. internal validity—confounding, and 5. power. All items were given one point when the respective criterion was fulfilled, except for the 'power' domain, where up to five points could be given, summing up to a maximum of 30 points per article. Serious methodological limitations were judged to exist when a pCCT collected less than 17 points on the checklist (48). Study selection and risk of bias assessment were made without blinding of the assessors (49).

Risk of bias across studies

If a sufficient number of trials were identified ($n > 10$), analyses were planned to identify reporting biases ('small-study effects' and/or publication bias), through the inspection of a contour-enhanced funnel plot (50), Begg's rank correlation test (51), and Egger's weighted regression test (52). When the tests hinted towards the existence of publication bias, the Duval and Tweedie's trim and fill procedure (53) was performed.

The overall quality of evidence (confidence in effect estimates) for each of the main outcomes was rated by using the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach (54), based on the following interpretations—'high quality': further research is very unlikely to change our confidence in the estimate of effect, 'moderate quality': further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate, 'low quality': further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate, and 'very low quality': very uncertain about the estimate.

The minimal clinical important effect was conventionally defined (55) as half a standard deviation (SD) of the Caucasian cephalometric norm plus 1 degree to allow for the methodological error (as reported on average by the identified studies). Likewise, large effects were defined as 1 SD of the norm plus 1 degree and very large effects as 2 SDs of the norm plus 1 degree. The optimal information size (i.e. required meta-analysis sample size) was calculated for each outcome independently in order to be able to identify a minimal clinical important effect with an average variance (based on this review's study sample), α at 5 per cent and β at 80 per cent.

The risk of bias within studies was assessed independently by two review authors (VK and VFZ) and across studies by a third author (SNP). Any disagreement was resolved by discussion with the last author (MAP). Inter-reviewer agreement for all three stages (study selection, data extraction, and risk of bias assessment) was evaluated with the un-weighted Cohen's kappa (56).

Summary measures and synthesis of results

Data were summarized and considered suitable for pooling if similar control groups of untreated Class II patients were used and the same cephalometric angular outcomes were reported. In cases of inadequate reporting, the missing data were calculated or requested from the authors. In order to account for the different follow-up periods of the included studies, the treatment (or observation) changes were annualized.

Mean differences (MDs) and their corresponding 95% confidence intervals (CIs) were used. A random-effects model as proposed by DerSimonian and Laird (57) was chosen *a priori* as the primary method to estimate all pooled estimates, since the observed treatment effect was expected to differ across studies due to differences in the sample (i.e. patient's skeletal maturity, growth pattern, gender) and implementation (i.e. various operators and settings, appliances utilized). This model takes into account existing heterogeneity and can be considered more conservative than the fixed-effect model in the presence of heterogeneity. The extent and impact of between-study heterogeneity was assessed by inspecting the forest plots and by calculating the τ^2 and the I^2 statistic, respectively. The 95% CIs around I^2 were calculated according to the non-central χ^2 approximation of Q. In case of great-unexplained heterogeneity ($I^2 > 75$ per cent), individual trials were omitted to achieve homogeneity. If homogeneity could not be achieved by excluding one or two trials and heterogeneity remained great, data were not pooled. For meta-analyses with ≥3 trials, 95% prediction intervals (PIs) (58, 59) were calculated to predict treatment effects in a future setting.

Additional analyses

Possible sources of heterogeneity were sought through pre-specified mixed-effects subgroup analyses and random-effects meta-regression with the Knapp–Hartung adjustment (60). In order to minimize the risk of excessive significance testing, the aforementioned analyses were performed only for meta-analyses with ≥5 trials.

Robustness of the results was *a priori* to be checked with sensitivity analyses for meta-analyses with ≥10 trials according to 1. the method error of the cephalometric analysis (if reported), 2. design of each study, and 3. the improvement of the GRADE classification by omitting trials.

All analyses were performed in Stata version 10 (StataCorp LP, College Station, Texas, USA) with the macros metan, metabias, heterogi, and confunnel. All P values were two sided with a level of significance at $\alpha = 0.05$, except for the test of between-studies or between-subgroups heterogeneity ($\alpha = 0.10$) (61).

Results

Study selection

From the initially identified 8175 records, 5633 records remained after exclusion of duplicates, to which 6 additional manually identified articles were added (Figure 1). A total of 5403 records were excluded on the basis of screening. The full texts of 236 articles were obtained and assessed for eligibility and 201 articles were excluded for various reasons, leaving 35 articles for further evaluation, from which 8 studies were again excluded, as they did not report cephalometric angular measurements. Finally, from the remaining 27 articles, 8 articles evaluated the effectiveness of FFAs and were excluded (Table 2). Consequently, 19 studies (46, 62–79) remained for final evaluation. In two instances, two studies pertaining to the same trial were grouped together (68 and 69; 46 and 72). Two studies (67, 71) were excluded due to missing appropriate data, which could not be obtained after communication, thus leaving 15 data-sets for inclusion in the meta-analysis.

In total, 39 authors were contacted, some of them more than once, with various questions regarding 44 articles and 13 of them responded by sending the requested data. Four of the authors answered, but could not send the requested data, while 22 of them did not reply. Finally, two e-mails could not be delivered (details available upon request). The kappa scores before reconciliation for the selection, data extraction, and risk of bias assessment procedures were 0.836, 0.883, and 0.904, respectively (with asymptotic standard errors 0.051, 0.093, and 0.045), indicating almost perfect agreement.

Study characteristics and risk of bias within studies

The characteristics of the 17 included studies are summarized in Table 3. Fifteen of them took place at a university setting and two at a hospital setting, including a total of 1031 patients with a mean age of 10.6 years. Most patients were treated with the original design of the respective RFA, while in seven studies, the appliances were modified and/or incorporated screw/spring elements for maxillary expansion. Almost all studies provided skeletal and dentoalveolar cephalometric outcomes and five (29 per cent) provided additionally soft tissue cephalometric outcomes.

Out of the seven included RCTs (65, 66, 70, 73, 76, 78, 79), five (65, 70, 73, 78, 79) were judged as being in high risk of bias, while two were considered as being in unclear (66) and low (76) risk of bias (Supplementary Table 3a). Nevertheless, the 10 pCCTs (46, 62–64, 67–69, 71, 72, 74, 75, 77) scored an average of 21.8 points on the modified Downs and Black tool, with one of them (62) demonstrating serious methodological limitations (i.e. less than 17 points) (Supplementary Table 3b).

Results of individual studies, synthesis of results, and risk of bias across studies

Effectiveness of treatment after removal of RFAs

Meta-analyses were performed regarding the short-term effectiveness of RFAs compared to natural growth (from untreated patients) for 28 cephalometric variables, including 15 skeletal (6 sagittal and 9 vertical), 8 dental, and 5 soft tissue variables (Table 4). The treatment effects of RFAs (without using data from untreated patients) and the effects of untreated patients alone are given separately as overview in Table 4 but are not discussed in detail. In general, RFA treatment was shown to have a statistically significant effect on skeletal, dental, and soft tissue relationships.

As far as skeletal changes in the sagittal plane are concerned, Class II malocclusion of treated patients was moderately improved (MD = −1.14 degree/year, $P < 0.001$, and MD = −3.16 degree/year, $P < 0.001$, for ANB and NAPg angles, respectively) compared to the untreated group (Figure 2). The skeletal component of the mandible was only minimally affected, with the SNB angle being on average

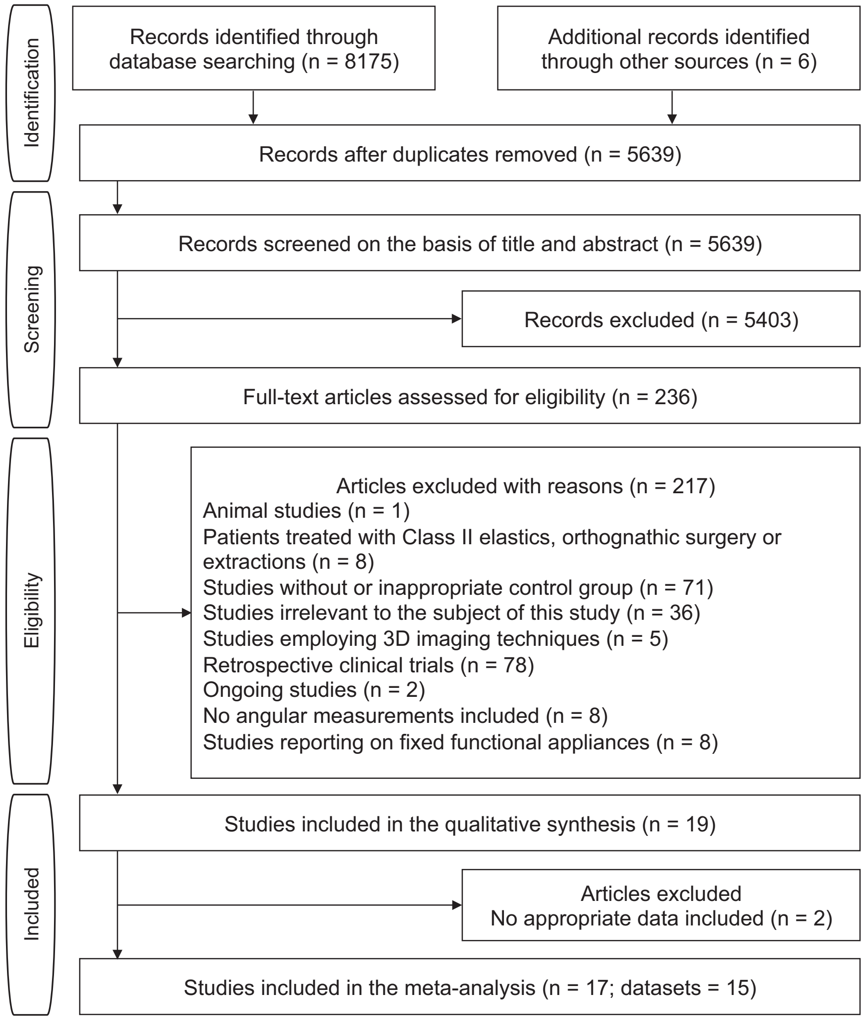

Figure 1. PRISMA flow diagram for the selection of studies.

Table 2. Number of excluded studies with reasons.

Reason for exclusion	Excluded articles on the basis of title and abstract	Excluded articles on the basis of full text
Animal studies	1	1
Patients treated with extractions, Class II elastics, or orthognathic surgery	73	8
Studies without or with inappropriate comparison group	129	71
Investigation not relevant to the subject of this study	4992	36
Evaluation employing 3D imaging techniques	5	5
Books'/conferences' abstracts	65	—
Cross-sectional surveys	2	—
Case reports or reports of cases	51	—
Retrospective clinical trials	59	78
Narrative reviews	7	—
Systematic reviews	17	—
Meta-analyses	2	—
Ongoing studies	—	2
No angular cephalometric measurements	—	8
Fixed functional appliances	—	8
No appropriate data provided	—	2
Sum	5403	219

Table 3. Characteristics of the 17 studies included in the current systematic review. CCT, prospective controlled trial; CCT*, prospective controlled trial with historical control; Ctr, control group; Gp1/2/3, treatment group; INCL, modified inclination; ME, maxillary expansion; NR, not reported; RCT, randomized clinical trial.

A/A	Study	Design	Setting	Characteristics of patients	Interventions	No. of patients (M/F)	Age in years (SD)	Skeletal growth stage	Treatment time* (m)	Outcomes	Risk of bias	Conflict of interest
1	Almeida-Pedrin *et al.* (62)	CCT*	University; Brazil	Class II/1 malocclusion, bilateral distal molar relationship > one-half cusp, ANB ≥ 4.5°	Bionator	Gp: 30 (15/15) Ctr: 30 (15/15)	Gp: 10.3 (NR) Ctr: 10.0 (NR)	Pre-peak	Gp: 18.2 Ctr: 17.9	Skeletal Dental	16/30	External; non-profit
2	Andreoli (63)	CCT	University; Brazil	Class II/1 malocclusion, retrognathic mandible, no previous orthodontic treatment, no posterior crossbite, mixed dentition	FR-II	Gp: 28 (15/13) Ctr: 18 (10/8)	Gp: 10.9 (NR) Ctr: 10.7 (NR)	NR	Gp: 18.0 Ctr: 18.0	Skeletal Dental Soft tissue	25/30	NR
3	Angelieri *et al.* (64)	CCT*	University; Brazil	Class II/1 malocclusion, mixed dentition with full-cusp or end-to-end molar relationship; late mixed dentition with clinically mandibular skeletal retrusion, distal step, overjet ≥5 mm, ANB > 2°	FR-II	Gp: 17 (10/7) Ctr: 17 (9/8)	Gp: 10.8 (0.6) Ctr: 11.3 (0.6)	Pre-peak and peak	Gp: 20.4 Ctr: 20.4	Skeletal Dental	19/30	External; non-profit
4	Baysal and Uysal (65)	RCT	University; Turkey	ANB > 4°, SNB < 78°, overjet ≥ 5 mm, SN-GoGn 32° ± 6°, crowding in dental arches ≤ 4 mm, bilateral Class II molar and canine relationship ≥ 3.5 mm	Twin Block	Gp: 20 (10/10) Ctr: 20 (11/9)	Gp: 13.0 (1.3) Ctr: 12.2 (1.5)	Pre-peak and peak	Gp: 16.2 Ctr: 15.6	Skeletal Dental Soft tissue	High risk	Internal
5	Brunharo *et al.* (66)	RCT	University; Brazil	ANB > 4°, molar Class II relationship, overjet > 6 mm, no previous orthodontic treatment, prepubertal growth spurt phase of skeletal maturity	Modified Twin Block (ME)	Gp: 19 (12/7) Ctr: 19 (12/7)	Gp: 9.5 (0.8) Ctr: 9.7 (1.0)	Pre-peak	Gp: 12.0 Ctr: 12.0	Skeletal Dental	Unclear	None
6	Cevidanes *et al.* (67)	CCT	University; Brazil	White Brazilians, 9–12 years, at the end of the mixed dentition and beginning of the pubertal growth spurt, no early loss of primary teeth, no missing permanent teeth, Class II/1 malocclusion with ≥ three-fourths cusp Class II molars	FR-II	Gp: 28 Ctr: 28 (matched)	Gp: 10.3 (0.9) Ctr: 10.9 (0.7)	Pre-peak	Gp: 18.0 Ctr: 18.0	Skeletal Dental	24/30	External; non-profit

(*Continued*)

Table 3. Continued

A/A	Study	Design	Setting	Characteristics of patients	Interventions	No. of patients (M/F)	Age in years (SD)	Skeletal growth stage	Treatment time* (m)	Outcomes	Risk of bias	Conflict of interest
7	Courtney *et al.* (68) and Nelson *et al.* (69)	CCT	University; New Zealand	From 10 to 13 years, Class II/1 malocclusion	Gp1: FR-II Gp2: Activator	Gp1: 13 (7/6) Gp2: 12 (7/5) Ctr: 17 (11/6)	Gp1: 11.7 (0.7) Gp2: 11.7 (0.9) Ctr: 11.5 (0.9)	NR	Gp1: 18.0 Gp2: 18.0 Ctr: 18.0	Skeletal Dental	21/30	External; non-profit
8	Dolce *et al.* (70)	RCT	University; USA	Class II malocclusion, fully erupted permanent first molars, positive overbite and overjet, absence of active dental or periodontal pathology, good general health, free from systemic illness/dysfunction	Bionator	Gp: 70 (46/24) Ctr: 74 (46/28)	Gp: 9.6 (1.1) Ctr: 9.5 (0.8)	NR	Gp: 26.4 Ctr: 26.4	Skeletal Dental	High risk	External; non-profit
9	Ehmer *et al.* (71)	CCT	University; Germany	Caucasian, overjet > 7 mm, first molar and incisors erupted, ≥ 1 year before pubertal growth maximum	Activator (ME)	Gp: 26 (13/13) Ctr: 61 (35/26)	Gp: 9.8 (0.7) Ctr: 9.5 (1.2)	Pre-peak	Gp: 15.0 Ctr: 15.0	Skeletal	21/30	NR
10	Illing *et al.* (46) and Morris *et al.*(72)	CCT	Hospital; UK	Caucasian patients, Class II/1 malocclusion with overjet > 7 mm, ANB > 6°, no previous orthodontic treatment, no adverse medical history	Gp1: Bass Gp2: Bionator Gp3: Twin Block (INCL)	Gp1: 13 (7/6) Gp2: 18 (9/9) Gp3: 16 (6/10) Ctr: 20 (13/7)	Gp1: 12.5 (1.8) Gp2: 11.8 (1.5) Gp3: 11.5 (1.5) Ctr: 11.2 (1.7)	NR	Gp1: 9.0 Gp2: 9.0 Gp3: 9.0 Ctr: 9.0	Skeletal Dental Soft tissue	22/30	NR
11	Jakobsson (73)	RCT	University; Sweden	Class II/1 malocclusion	Activator	Gp: 20 (NR) Ctr: 20 (NR)	Gp: 8.5 (NR) Ctr: 8.5 (NR)	NR	Gp: 18.0 Ctr: 18.0	Skeletal Dental	High risk	Internal
12	Janson *et al.* (74)	CCT*	University; Brazil	Class II/1 malocclusion with ≥ 1/2 cusp or ≥ end-to-end molar relationship	FR-Ib	Gp: 18 (10/8) Ctr: 23 (13/10)	Gp: 9.3 (NR) Ctr: 9.3 (NR)	NR	Gp: 28.0 Ctr: 28.0	Skeletal Dental	23/30	External; non-profit
13	Lund and Sandler (75)	CCT	Hospital; UK	10–14 years, white, ANB > 5°, Class II/1 incisor relationship, overjet > 6 mm	Twin Block (ME/INCL)	Gp: 36 (19/17) Ctr: 27 (13/14)	Gp: 12.4 (NR) Ctr: 12.1 (NR)	NR	Gp: 10.8 Ctr: 14.4	Skeletal Dental	25/30	NR
14	Martina *et al.* (76)	RCT	University; Italy	Full Class II molar relationship, overjet ≥ 6 mm, 10–13 years for boys and 9–12 for girls	Sander bite-jumping appliance (ME)	Gp: 23 (15/8) Ctr: 23 (11/12)	Gp: 10.9 (1.3) Ctr: 10.5 (1.2)	Pre-peak and peak	Gp: 14.5 Ctr: 12.0	Skeletal Dental	Low risk	External; non-profit
15	Quintão *et al.* (77)	CCT	University; Brazil	ANB > 4°, overjet ≥ 6 mm, Class II canine and molar relationship, no previous orthodontic treatment	Twin Block (ME)	Gp: 19 (12/7) Ctr: 19 (12/7)	Gp: 9.5 (0.8) Ctr: 9.9 (1.1)	Pre-peak	Gp: 12.0 Ctr:12.0	Skeletal Dental Soft tissue	22/30	NR

(*Continued*)

Table 3. Continued

A/A	Study	Design	Setting	Characteristics of patients	Interventions	No. of patients (M/F)	Age in years (SD)	Skeletal growth stage	Treatment time* (m)	Outcomes	Risk of bias	Conflict of interest
16	Tulloch *et al.* (78)	RCT	University; USA	Overjet ≥ 7 mm, all permanent incisors and first molars erupted, all permanent teeth (excluding third molars) developing, 1 year pre-peak-height velocity	Bionator (modified)	Gp: 53 (30/23) Ctr: 61 (35/26)	Gp: 9.4 (1.0) Ctr: 9.4 (1.2)	Pre-peak	Gp:15.0 Ctr:15.0	Skeletal Dental	High risk	External; non-profit
17	Varlik *et al.* (79)	RCT	University; Turkey	Class II molar relationship, overjet ≥ 5 mm, ANB > 4°, SNB < 78°, SN-MP 32° ± 2°, maximal pubertal growth, no previous orthodontic therapy	Gp1: Twin Block Gp2: Activator	Gp1: 25 (12/13) Gp2: 25 (13/12) Ctr: 25 (13/12)	Gp1: 11.9 (0.2) Gp2: 11.9 (0.2) Ctr: 10.1 (0.9)	Peak	Gp1: 8.0 Gp2: 9.0 Ctr: 8.0	Dental Soft tissue	High risk	NR

*Control group received no intervention, it refers to the observation time.

0.62 degree/year greater than the untreated group. RFA treatment also resulted in a similar minimal restriction effect of −0.28 degree/year on the maxillary growth. Finally, in the vertical plane, most of the measurements were statistically insignificant, except for SN-ML and ArGoMe angles (MD = 0.66 and 0.94 degree/year, respectively).

As far as dentoalveolar changes are concerned, treatment effects were more pronounced on the upper dentition, where the upper central incisors were significantly retroclined (1s-SN angle: MD = −3.29 degree/year; 1s-NL angle: MD = −6.33 degree/year; 1s-NA angle: MD = −5.21 degree/year). The 95% PI of −14.38 to 1.71 degree/year for 1s-NL indicated that the effect could be expected to be considerably large in almost every future clinical use of RFAs. Moreover, significant proclination of the lower incisors was observed via the 1i-ML angle (MD = 1.37 degree/year) and 1i-NB angle (MD = 1.81 degree/year).

Finally, the influence of RFA treatment on the soft tissues was evident, as indicated by the significant changes of all cephalometric variables and, especially, the mentolabial angle (MD = 22.60 degree/year).

The results of the included studies varied considerably and various patient- and appliance-related factors influenced the treatment results of RFAs (Table 5). Subgroup analyses with respect to the patients' growth pattern were not feasible because of the limited reporting in the original studies. Additionally, comparisons with regard to the patients' skeletal growth stage were feasible only between the pre-peak and peak stages. The skeletal sagittal effects of RFAs varied significantly according to both patient-related characteristics (patients' gender ratio) and appliance-related characteristics (appliance utilized, mode of action, number of RFA's pieces, and construction bite).

Among the RFAs studied, the Twin Block appliance induced significantly greater skeletal changes (through the SNA, SNB, and ANB angles) and greater dental effects (through the 1s-NL angle) than the other RFAs.

Effectiveness of treatment after removal of fixed appliances and after retention

Due to insufficient number of identified studies, no meta-analyses could be performed for the effects of RFAs after the period of fixed appliances (that followed functional treatment) or after the period of retention.

Risk of bias across studies and additional analyses

Regarding the risk of publication bias, evidence of funnel plot asymmetry was found through the Egger and the Begg test only for the ANB angle (Table 6 and Figure 3). However, no missing studies were found by the 'fill and trim' procedure and no conflict of interest could explain the asymmetry. Therefore, it was concluded that the funnel plot asymmetry indicated a small-study effect, i.e. that small studies tend to overestimate the decrease of ANB due to imprecision. A sensitivity analysis was conducted by including only the 5 most precise studies from the 10 included (Supplementary Table 4) and found that the decrease in ANB angle was minimized compared to the original one but was still statistically significant.

The sensitivity analysis on the basis of the design of the included studies is presented in Supplementary Table 4. No significant difference could be found between RCTs and pCCTs for any of the variables assessed.

The GRADE assessments for the main outcomes after removal of the RFAs ranged from 'very low' to 'high' (Table 7). The quality of clinical recommendations was upgraded due to the magnitude of treatment effects for the sagittal position of the maxilla (via the SNA angle) and the retroclination of the maxillary incisors (via the 1s-NL angle). Further sensitivity analyses on the basis of the GRADE

Table 4. Details of the performed meta-analyses with tests on heterogeneity. CI, confidence interval; Ctr, control group; MD, mean difference; NA, non-applicable; NT, not tested; PI, prediction interval; RFA, removable functional appliance.

					Effect size			Heterogeneity			
No.	Variable	Studies	Post–pre in RFA*	Post–pre in Ctr*	MD	95% CI	*P* value	*P* value	τ^2	I^2 (%)	95% PI
01	SNA	11**	↓(−0.10)	↑(0.19)	−0.28	−0.44, −0.12	**0.001**	0.334	0.009	12 (0,57)	−0.56, 0.00
02	SNB	11**	↑(0.96)	↑(0.29)	0.62	0.36, 0.88	**<0.001**	0.001	0.109	67 (26,81)	−0.19, 1.42
03	SNPg	1**	↑(0.77)	–(0.00)	0.73	−0.31, 1.77	0.168	NA	NA	NA	NA
04	ANB***	10**	↓(−1.30)	↓(−0.14)	−1.14	−1.52, −0.77	**<0.001**	<0.001	0.283	85 (72,90)	−2.45, 0.16
05	NAPg	2	↓(−2.40)	↓(−0.39)	−3.16	−4.32, −2.00	**<0.001**	0.211	0.274	36 (NA)	NA
06	NSBa	3**	↑(0.25)	↓(−0.11)	0.49	0.03, 0.94	**0.037**	0.037	0.000	0 (0,73)	−2.47, 3.45
07	SN-ML	8**	↑(0.20)	↓(−0.49)	0.66	0.19, 1.12	**0.006**	0.006	0.264	65 (0,82)	−0.73, 2.04
08	FH-ML****	4	↑(0.16)	↓(−0.46)	0.14	−0.25, 0.54	0.478	0.077	0.085	56 (0,83)	−1.39, 1.68
09	SN-SGn	2	↓(−0.17)	↓(−0.09)	0.20	−0.55, 0.94	0.608	0.203	0.136	38 (NA)	NA
10	NL-ML	3**	↓(−0.56)	↓(−0.42)	0.30	−0.94, 1.53	0.638	0.083	0.713	60 (0,87)	−13.10, 13.69
11	SN-NL	6**	↑(0.26)	↑(0.22)	−0.09	−0.43, 0.24	0.583	0.043	0.093	56 (0,80)	−1.06, 0.88
12	FH-NL	2	↓(−0.02)	↑(0.10)	−0.12	−0.31, 0.07	0.213	0.610	0.000	0 (NA)	NA
13	SN-OP	2	↓(−0.20)	↓(−0.84)	0.65	−0.09, 1.39	0.083	0.560	0.000	0 (NA)	NA
14	ML-OP	1**	↓(−0.23)	↓(−0.67)	0.58	−1.95, 3.11	0.654	NA	NA	NA	NA
15	ArGoMe	4**	↑(0.58)	↓(−0.46)	0.94	0.49, 1.40	**0.001**	0.309	0.048	15 (0,77)	−0.06, 1.95
16	1s-SN	4**	↓(−3.32)	↑(0.35)	−3.29	−4.69, −1.89	**<0.001**	0.018	1.304	70 (0,87)	−9.09, 2.50
17	1s-NL***	5**	↓(−6.51)	↑(0.35)	−6.33	−8.53, −4.14	**<0.001**	<0.001	5.125	87 (69,93)	−14.38, 1.71
18	1i-ML***	12**	↓(−0.10)	↑(0.19)	1.37	0.64, 2.10	**<0.001**	<0.001	1.159	81 (66,88)	−1.16, 3.91
19	1s-1i***	4**	↑(4.07)	↑(1.26)	2.55	−0.70, 5.80	0.125	<0.001	8.603	83 (40,92)	−11.95, 17.04
20	1s-NA***	5	↓(−4.63)	↑(0.65)	−5.21	−7.46, −2.96	**<0.001**	<0.001	5.420	85 (62,92)	−13.47, 3.05
21	1i-NB	5	↑(2.40)	↑(0.38)	1.81	0.82, 2.81	**<0.001**	0.084	0.640	51 (0,80)	−1.20, 4.83
22	6s-NL*****	3**	↑(0.14)	↓(−0.62)	0.44	−0.77, 1.65	0.477	0.016	0.812	76 (0,91)	−13.43, 14.31
23	6i-ML	2	↓(−0.40)	↓(−0.25)	−0.14	−0.69, 0.40	0.603	0.771	0.000	0 (NA)	NA
24	N′PnPg′	2**	↑(1.37)	↓(−1.58)	1.96	0.80, 3.13	**0.001**	0.389	0.000	0 (NA)	NA
25	Nasolabial angle	5**	↑(2.87)	↑(0.16)	2.78	1.02, 4.55	**0.002**	0.263	0.981	24 (0,72)	−1.48, 7.04
26	Mentolabial angle	3**	↑(18.12)	↑(0.19)	22.60	18.31, 26.90	**<0.001**	0.474	0.000	0 (0,73)	−5.23, 50.44
27	H angle	3**	↓(−2.70)	↑(0.17)	−2.76	−4.28, −1.23	**<0.001**	0.085	1.076	59 (0,87)	−19.24, 13.73
28	Z angle******	2**	↑(6.02)	↑(0.02)	5.24	0.70, 9.78	**0.024**	0.004	9.461	88 (NA)	NA

Values in bold are significant at the 5% level.

*Results from random-effects meta-analysis of the post–pre differences in each group to provide an overview of the effect's direction.

**Multiple functional appliance trial arms identified from one or more included studies. Trial arms were pooled prior to the meta-analysis to avoid double-counting patients.

***High inconsistency between studies, which was accounted for in subgroup analyses.

****Initial analysis included five studies (MD = 0.73; 95% CI = −0.15 to 1.61; P = 0.104; τ^2 = 0.870; I^2 = 93%), but the study Andreoli (63) (Fränkel) was omitted to achieve homogeneity.

*****Inconsistency just above the 75% margin, but our confidence in the I^2 is very low (judged from the 95% uncertainty interval) and therefore no study was omitted.

******High inconsistency, which however affects only the magnitude of the effect, but not the direction (both studies on the same side of the forest plot). Effect might not be accurately estimated.

quality of recommendations were not feasible, as the only reason for downgrading was reporting bias for the ANB angle.

Discussion

Summary of evidence

This review included data from 1031 patients and 17 RCTs and pCCTs, which assessed cephalometric changes induced by Class II treatment with RFAs. As far as the ANB angle is concerned, results of the random-effects meta-analyses indicated that RFAs were statistically significantly successful in improving Class II skeletal relationships (as indicated by the 1.14 degree/year decrease in the ANB angle of the treated patients in comparison to untreated individuals). The contribution of mandibular advancement was approximately twice the amount of maxillary growth restriction. However, this 'growth stimulation' or 'orthopaedic effect' on the mandible could be considered clinically negligible (0.62 degree/year). The skeletal vertical effects of RFAs were similarly minimal.

It seems that the effects of RFAs are more pronounced on the dentition, with a considerable amount of maxillary incisor retroclination and small mandibular incisor proclination. The inclination of the upper and lower molars on the other hand was not significantly affected by RFA treatment. Finally, the soft tissues were also favourably affected from the use of RFAs.

Various factors seem to influence the outcome of RFA treatment. The influence of patient's gender on the effects of RFA treatment could not be formally assessed in this study due to the incomplete reporting of separate data for boys and girls. Current results indicate that treatment effects of RFAs on the SNB, ANB, and 1s-NA angles might differ according to patient's gender, but additional studies are needed to confirm or refute this. The amount of mandibular growth has been reported to be related to patient's gender (80), while different treatment responses to RFAs by boys and girls have also been documented (81, 82).

Study	RFAs N	RFAs Mean	RFAs SD	Control N	Control Mean	Control SD	MD (95% CI)	Weight
SNA angle change								
Almeida-Pedrin 2007	30	-0.22	0.68	30	-0.33	0.89	0.11 (-0.29, 0.51)	3.52
Andreoli 2008	28	0.15	0.83	18	0.31	0.45	-0.16 (-0.53, 0.21)	3.56
Angelieri 2013	17	-0.24	0.65	17	0.18	1.00	-0.42 (-0.99, 0.15)	3.26
Baysal 2011	20	-0.56	0.76	20	0.15	0.53	-0.71 (-1.12, -0.30)	3.51
Brunharo 2011	19	0.71	1.54	19	1.16	3.05	-0.45 (-1.99, 1.09)	1.68
Illing 1998-pooled	34	-0.24	2.74	20	0.40	2.00	-0.64 (-1.91, 0.63)	2.04
Janson 2003	18	-0.06	0.67	23	-0.01	1.01	-0.05 (-0.57, 0.47)	3.34
Lund 1998	36	-0.11	1.78	27	0.25	0.67	-0.36 (-0.99, 0.27)	3.14
Nelson 1993-pooled	25	-0.11	0.51	17	0.23	0.49	-0.34 (-0.65, -0.03)	3.64
Quintao 2006	19	0.05	1.07	19	0.95	2.37	-0.90 (-2.07, 0.27)	2.20
Tulloch 1997	53	0.11	1.26	61	0.26	1.17	-0.15 (-0.60, 0.30)	3.45
Total SNA (I^2=12%; P=0.334)							-0.28 (-0.44, -0.12)	33.31
with estimated predictive interval							(-0.56, 0.00)	
SNB angle change								
Almeida-Pedrin 2007	30	0.81	0.85	30	-0.12	0.92	0.93 (0.48, 1.38)	3.45
Andreoli 2008	28	0.81	0.47	18	0.41	0.35	0.40 (0.16, 0.64)	3.71
Angelieri 2013	17	0.41	0.59	17	0.29	0.65	0.12 (-0.30, 0.54)	3.49
Baysal 2011	20	1.53	1.01	20	0.35	0.73	1.18 (0.63, 1.73)	3.29
Brunharo 2011	19	1.83	1.32	19	1.15	2.71	0.68 (-0.68, 2.04)	1.92
Illing 1998-pooled	34	0.93	2.33	20	-0.27	1.73	1.20 (0.11, 2.29)	2.33
Janson 2003	18	0.43	0.67	23	0.19	0.73	0.24 (-0.19, 0.67)	3.47
Lund 1998	36	2.11	2.22	27	0.33	0.83	1.78 (0.99, 2.57)	2.86
Nelson 1993-pooled	25	0.47	0.69	17	0.29	0.47	0.18 (-0.17, 0.53)	3.58
Quintao 2006	19	1.38	1.05	19	0.92	2.01	0.46 (-0.56, 1.48)	2.44
Tulloch 1997	53	1.07	0.91	61	0.43	0.90	0.64 (0.31, 0.97)	3.61
Total SNB (I^2=67%; P=0.001)							0.62 (0.36, 0.88)	34.15
with estimated predictive interval							(-0.19, 1.42)	
ANB angle change								
Almeida-Pedrin 2007	30	-1.02	0.85	30	-0.30	0.62	-0.72 (-1.10, -0.34)	3.55
Andreoli 2008	28	-0.85	0.57	18	-0.11	0.58	-0.74 (-1.08, -0.40)	3.60
Angelieri 2013	17	-0.71	0.41	17	-0.12	0.53	-0.59 (-0.91, -0.27)	3.62
Baysal 2011	20	-2.11	0.93	20	-0.17	0.64	-1.94 (-2.43, -1.45)	3.37
Brunharo 2011	19	-1.16	0.94	19	0.02	1.53	-1.18 (-1.99, -0.37)	2.82
Illing 1998-pooled	34	-2.36	2.50	20	0.53	1.60	-2.89 (-3.98, -1.80)	2.32
Janson 2003	18	-0.51	0.63	23	-0.21	0.67	-0.30 (-0.70, 0.10)	3.52
Lund 1998	36	-2.22	2.11	27	-0.08	0.67	-2.14 (-2.87, -1.41)	2.96
Quintao 2006	19	-1.33	0.68	19	0.03	1.20	-1.36 (-1.98, -0.74)	3.16
Tulloch 1997	53	-0.93	0.99	61	-0.17	0.73	-0.76 (-1.08, -0.44)	3.62
Total ANB (I^2=85%; P=0.000)							-1.14 (-1.52, -0.77)	32.54
with estimated predictive interval							(-2.45, 0.16)	

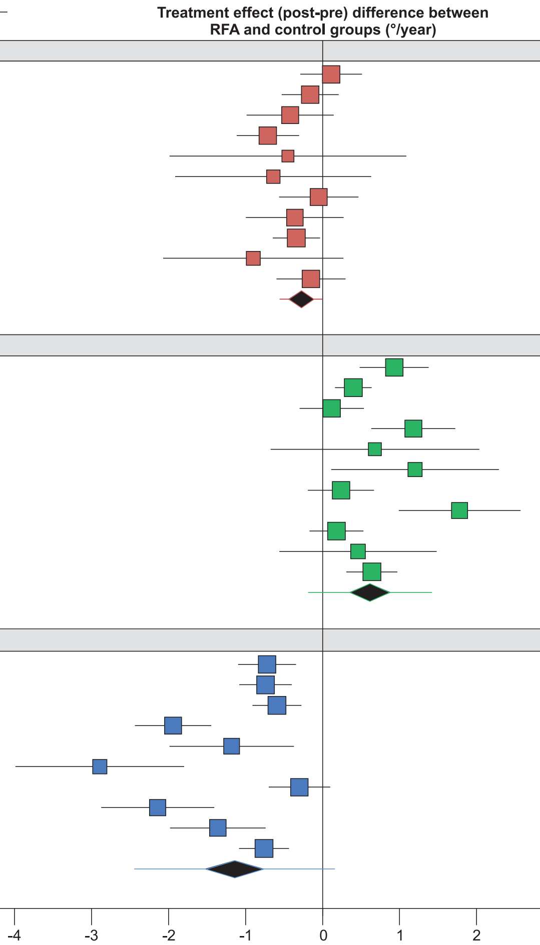

Figure 2. Forest plot of the mean difference of the SNA, SNB, and ANB angles (in °/year) between RFAs and control groups based on the random-effects model together with the 95% confidence interval and the 95% prediction interval. Studies on the right indicate that removable functional treatment results in decreased SNA angle/increased SNB angle/decreased ANB angle compared to the normal growth, while studies on the right indicate increased SNA angle/decreased SNB angle/increased ANB angle compared to the normal growth.

An analysis regarding the treatment effects of RFAs according to the patient's growth pattern (i.e. average, vertical, and horizontal) was originally planned but was not possible. However, the retrospective studies of Greco *et al.* (83) and Hönn *et al.* (84) report that RFA treatment produces more evident results on patients with a horizontal growth pattern compared to patients with average and vertical growth patterns.

The timing of functional treatment is a subject of controversy among investigators (85–88). The present results, including a dental variable (1i-ML) and a soft tissue variable (nasolabial angle), did not indicate that the patient's skeletal growth stage (pre-peak versus peak) at the start of RFA treatment had a significant effect on any of the variables included. Due to insufficient data, the present study could not provide results concerning the influence of the growth stage on the skeletal component. However, a recent report (89) indicated that patients at their pubertal peak treated with Activator and Bionator presented significantly greater increases in total mandibular length and ramus height associated with a significant advancement of the bony chin compared with treatment before their pubertal peak. Moreover, treatment with Twin Block during or slightly after the onset of the pubertal peak produced more favourable effects than before growth peak (86). It must be noted that the analysis for these two subgroups included only a limited number of studies.

A plethora of RFAs was used in the included studies including Activator, Bionator, the Fränkel appliance, the Sander appliance, and Twin Block. Twin Block as a bite-jumping appliance consisting of two bite blocks induced the greatest changes in skeletal sagittal relationships and maxillary incisor retroclination, followed by Bionator. Moreover, Toth and McNamara (11) reported that Twin Block, as compared to Fränkel appliance, produced greater changes in regard to SNB and ANB angles and posterior tipping of the upper incisors. Furthermore, in comparison to Bionator, Twin Block was more effective in the treatment of Class II malocclusion (90). The supremacy of Twin Block over other RFAs might derive from additionally utilizing the mastication forces to treat Class II malocclusion since Twin Blocks are more stably attached on the dentition and designed for full-time wear (91). If so, the factor of full-time wear, which is strongly related to patients' cooperation, is once more highlighted. However, despite its' importance, patient compliance to RFA is under-reported in clinical trials (92). The supremacy of Twin Block over the other RFAs is also depicted in Supplementary Figures 1–5, which represent meta-analyses of annual changes of three skeletal (Supplementary Figures 1–3) and two dentoalveolar (Supplementary Figures 4 and 5) caused by the various RFAs assessed in the present study.

Table 5. Subgroup analyses for meta-analyses containing five or more studies. CI, confidence interval; MD, mean difference; NA, non-applicable; P_{SG}, P value for difference between subgroups; SG, subgroup.

Factor	SG	*n*	MD (95% CI), SNA (*n* = 13)	P_{SG}	*n*	MD (95% CI), SNB (*n* = 13)	P_{SG}	*n*	MD (95% CI), ANB (*n* = 11)	P_{SG}	*n*	MD (95% CI), SN-ML (*n* = 9)	P_{SG}	*n*	MD (95% CI), FH-ML (*n* = 5)	P_{SG}
Patient related																
Gender	*	13		0.721	13		**0.043**			**0.090**			0.996	4		0.678
Appliance related																
Appliance	Activator	1	−0.39 (−0.77, −0.01)	**0.053**	1	0.15 (−0.38,0.68)	**0.006**	—	—	**0.032**	1	1.06 (0.34,1.78)	0.867	—	—	0.913
	Bionator	3	0.03 (−0.27,0.32)		3	0.75 (0.49,1.02)		3	−0.83 (−1.22, −0.44)		1	0.38 (−0.88,1.64)		1	0.16 (−0.05,0.37)	
	Fränkel	4	−0.22 (−0.44, −0.01)		4	0.29 (0.13,0.46)		3	−0.56 (−0.80, −0.33)		3	0.81 (−0.21,1.82)		3	0.90 (−0.78,2.58)	
	Sander	—	—	—	—	—	—	—	—		1	0.94 (−0.39,2.27)		1	0.98 (−0.44,2.40)	
	Twin Block	5	−0.79 (−1.28, −0.29)		5	1.19 (0.76,1.61)		5	−1.96 (−2.58, −1.33)		3	0.29 (−0.36,0.95)		—	—	
Mode of action	Bite-jumping	9	−0.40 (−0.74, −0.05)	0.577	9	0.86 (0.54,1.19)	**0.015**	8	−1.62 (−2.17, −1.06)	**0.067**	6	0.64 (0.22,1.07)	0.641	2	0.26 (−0.26,0.78)	0.782
	Periosteal pull	4	−0.22 (−0.44, −0.01)		4	0.29 (0.13,0.46)		3	−0.56 (−0.80, −0.33)		3	0.81 (−0.21,1.82)		3	0.90 (−0.78,2.58)	
Number of components	1 piece	8	−0.18 (−0.34, −0.02)	**0.018**	8	0.41 (0.21,0.60)	**0.008**	6	−0.66 (−0.87, −0.45)	**0.005**	5	0.80 (0.12,1.47)	0.461	4	0.69 (−0.29,1.67)	0.876
	2 pieces	5	−0.79 (−1.28, −0.29)		5	1.19 (0.76,1.61)		5	−1.96 (−2.59, −1.33)		4	0.42 (−0.17,1.01)		1	0.98 (−0.44,2.40)	
Construction bite	Single step	2	−0.13 (−1.57,1.32)	0.855	2	1.17 (0.65,1.68)	**0.095**	2	−1.97 (−2.44, −1.50)	**0.075**	1	0.29 (−0.71,1.29)	0.890	—	—	
	Stepwise	2	−0.42 (−0.96,0.11)		2	0.17 (−0.23,0.57)		2	−0.76 (−1.29, −0.24)		2	0.43 (−0.66,1.53)		—	—	

Factor	SG	*n*	MD (95% CI), SN-NL (*n* = 8)	P_{SG}	*n*	MD (95% CI), ArGoMe (*n* = 5)	P_{SG}	*n*	MD (95% CI), 1s-SN (*n* = 5)	P_{SG}	*n*	MD (95% CI), 1s-NL (*n* = 7)	P_{SG}	*n*	MD (95% CI), 1i-ML (*n* = 15)	P_{SG}
Patient related																
Gender	*	8		0.731	4		0.918	5		0.767	7		0.232	14		0.497
Skeletal growth stage	Pre-peak	—	—	—	—	—	—	—	—	—	—	—	—	2	1.42 (0.37,2.47) (0.63,2.55)	0.162
	Peak	—	—		—	—		—	—		—	—		2	3.10 (2.29,3.91)	
Appliance related																
Appliance	Activator	1	0.37 (−0.34,0.28)	0.467	2	1.01 (0.35,1.67)	0.540	1	−4.13 (−4.67, −3.59)	0.668	1	−5.16 (−8.10, −4.29)	**0.050**	3	2.38 (0.66,4.10)	0.465
	Bionator	2	−0.19 (−0.63,0.25)		—	—		—	—		1	−7.60 (−11.22, −3.98)		2	4.11 (−2.17,10.39)	
	Fränkel	3	−0.26 (−0.75,0.23)		2	0.65 (−0.08,1.38)		1	−1.94 (−3.41, −0.47)		3	−4.27 (−5.20, −3.34)		4	0.49 (−0.20,1.18)	
	Sander	—	—		—	—		1	−4.41 (−6.91, −1.91)		—	—		1	2.43 (0.07,4.79)	
	Twin Block	2	0.41 (−0.37,1.19)		1	1.81 (0.61,3.01)		2	−3.70 (−5.43, −1.97)		2	−11.39 (−13.83, −8.95)		5	3.46 (1.16,5.76)	
Mode of action	Bite-jumping	5	0.16 (−0.20,0.52)	0.207	3	1.20 (0.62,1.78)	0.277	4	−4.15 (−4.55, −3.76)	**0.065**	4	−8.53 (−11.95, −5.12)	**0.050**	11	2.85 (1.78,3.92)	**0.062**
	Periosteal pull	3	−0.26 (−0.75,0.23)		2	0.65 (−0.08,1.38)		1	−1.94 (−3.41, −0.47)		3	−4.27 (−5.20, −3.34)		4	0.49 (−0.20,1.18)	
Number of components	1 piece	6	−0.11 (−0.45,0.23)	0.283	4	0.81 (0.35,1.26)	0.222	2	−3.15 (−5.28, −1.01)	0.625	5	−4.55 (−5.51, −3.58)	**0.006**	9	1.58 (0.44,2.71)	0.245
	2 pieces	2	0.41 (−0.37,1.19)		1	1.81 (0.62,3.01)		3	−4.10 (−4.89, −3.32)		2	−11.39 (−13.83, −8.95)		6	3.24 (1.32,5.15)	
Construction bite	Single step	—	—	—	—	—	—	3	−4.14 (−4.57, −3.71)	0.171	—	—	—	5	2.47 (1.10,3.83)	0.540
	Stepwise	—	—		—	—		2	−2.96 (−5.34, −0.58)		—	—		3	1.41 (−0.25,3.07)	

(*Continued*)

Table 5. Continued

Factor	SG	n	MD (95% CI), 1s-1i (n = 5)	P_{SG}	n	MD (95% CI), 1s-NA (n = 5)	P_{SG}	n	MD (95% CI), 1i-NB (n = 5)	P_{SG}	n	MD (95% CI), nasolabial (n = 7)	P_{SG}	n	MD (95% CI), mentolabial (n = 5)	P_{SG}
Patient related																
Gender	*	5		0.761	5		**0.024**	5		0.679	7		0.762	5		0.692
Skeletal growth stage	Pre-peak	—	—	—	—						1	3.79 (−1.43,9.01)	0.989			
	Peak	—	—		—						2	3.95 (−6.21,14.11)				
Appliance related																
Appliance	Activator	—	—	**0.090**	—	—	0.713	—	—	0.213	1	−1.22 (−2.72,0.28)	0.769	1	21.74 (14.24,29.24)	0.542
	Bionator	2	−0.50 (−2.03,1.03)		1	−3.41 (−4.95, −1.87)		1	2.58 (1.24,3.92)		1	2.00 (−4.61,8.61)		1	12.93 (−1.94,27.80)	
	Fränkel	1	1.47 (−0.65,3.59)		1	−3.66 (−4.87, −2.45)		1	0.49 (−0.56,1.54)		1	2.16 (−0.56,4.88)		—	—	
	Sander	—	—		—	—		—	—		—	—		—	—	
	Twin Block	2	7.33 (4.28,10.38)		3	−6.62 (−11.12, −2.12)		3	2.15 (1.06,3.24)		4	4.24 (−1.41,9.89)		3	23.36 (18.12,28.60)	
Mode of action	Bite-jumping	4	2.94 (−1.73,7.60)	0.780	4	−5.74 (−8.97, −2.50)	0.619	4	2.32 (1.48,3.17)	**0.076**	6	2.81 (−2.42,8.05)	0.895	—	—	—
	Periosteal pull	1	1.47 (−0.65,3.59)		1	−3.66 (−4.87, −2.45)		1	0.49 (−0.56,1.54)		1	2.16 (−0.56,4.88)		—	—	
Number of components	1 piece	3	0.20 (−1.14,1.53)	**0.029**	2	−3.56 (−4.52, −2.61)	0.350	2	1.49 (−0.55,3.54)	0.522	3	0.45 (−2.26,3.17)	0.285	2	19.77 (12.57,26.97)	0.490
	2 pieces	2	7.33 (4.28,10.38)		3	−6.62 (−11.12, −2.12)		3	2.15 (1.06,3.24)		4	4.24 (−1.41,9.89)		3	23.36 (18.12,28.60)	
Construction bite	Single step	1	−0.67 (−5.22,3.88)	0.535	1	−2.38 (−4.65, −0.11)	NA	1	1.62 (0.01,3.23)	NA						
	Stepwise	2	4.45 (−1.70,10.60)		1	−9.71 (−12.46, −6.96)		1	2.82 (0.59,5.05)		—	—		—		

Values in bold are significant at the 5% level.
*Coefficient and its 95% CI from random-effects meta-regression are given.

Table 6. Assessment of reporting biases in the performed meta-analyses. NA, non-applicable; NT, not tested.

No.	Variable	Studies	Egger's P value	Begg's P value	Duval's imputed studies	Adjusted estimate
01	SNA	11	0.377	0.583	NT	NT
02	SNB	11	0.136	0.200	NT	NT
03	SNPg	1	NT	NT	NT	NT
04	ANB	10	**0.007**	**0.043**	0	NA
05	NAPg	2	NT	NT	NT	NT
06	NSBa	3	NT	NT	NT	NT
07	SN-ML	8	NT	NT	NT	NT
08	FH-ML	4	NT	NT	NT	NT
09	SN-SGn	2	NT	NT	NT	NT
10	NL-ML	3	NT	NT	NT	NT
11	SN-NL	6	NT	NT	NT	NT
12	FH-NL	2	NT	NT	NT	NT
13	SN-OP	2	NT	NT	NT	NT
14	ML-OP	1	NT	NT	NT	NT
15	ArGoMe	4	NT	NT	NT	NT
16	1s-SN	4	NT	NT	NT	NT
17	1s-NL	5	NT	NT	NT	NT
18	1i-ML	12	0.911	0.276	NT	NT
19	1s-1i	4	NT	NT	NT	NT
20	1s-NA	5	NT	NT	NT	NT
21	1i-NB	5	NT	NT	NT	NT
22	6s-NL	3	NT	NT	NT	NT
23	6i-ML	2	NT	NT	NT	NT
24	N′PnPg′	2	NT	NT	NT	NT
25	Nasolabial angle	5	NT	NT	NT	NT
26	Mentolabial angle	3	NT	NT	NT	NT
27	H angle	3	NT	NT	NT	NT
28	Z angle	2	NT	NT	NT	NT

Values in bold are significant at the 5% level.

The mode of action of the RFAs (bite-jumping or periosteal pull) seems to influence their treatment effects. Bite-jumping appliances were the most effective in advancing the mandible in a forward position and improving Class II skeletal relationship. However, they were at the same time the most inclined to dentoalveolar compensation since they presented the highest upper incisor retroclination and lower incisor proclination. A previous study comparing Twin Block and Fränkel appliances is in accordance with the present findings (11). However, this comparison might be confounded by the fact that the periosteal pull subgroup included only the Fränkel appliance.

Finally, the construction bite or the activation of the RFA (single step versus stepwise) played a role on RFA treatment only with respect to SNB and ANB angles, with mandibular advancement in a single step demonstrating more pronounced results than stepwise advancement. A study conducted with Fränkel appliances reported more favourable response with regard to sagittal jaw discrepancy, with incremental mandibular advancement than single step advancement (93). However, an RCT conducted with Twin Blocks did not detect differences between maximum and incremental mandibular advancement (94).

Strengths and limitations

The strengths of the present systematic review include the pre-defined protocol, the extensive literature search, and the strict methodology implemented during every stage of it, according to existing guidelines (37–39, 54). In addition, the 17 included studies provided data to perform adequately powered meta-analyses of many important treatment outcomes. Since a random-effects model was used for data synthesis, our results provide the average effects of RFAs across studies. Heterogeneity was explained in most cases by the pre-defined subgroup analyses, while publication bias diagnostics and sensitivity analyses indicated that the results were fairly robust. Data from a minimum of five studies was considered as adequate to perform subgroup analyses, in order to minimize as much as possible multiple testing. Although the majority of studies were conducted in university clinics, the results could be probably generalized to the average patient, due to the generic patient inclusion criteria and the various countries of origin of the original studies. Finally, many pCCTs were included, which represent more pragmatic settings (comparable to every day clinical practice) than RCTs, adding further to the results' applicability.

Nevertheless, although no serious methodological limitations were found in the included original studies, their quality was moderate. For example, the patients' compliance, which is a crucial parameter for successful treatment (95), was reported as adequate in only five of the pCCTs (Supplementary Table 3b). In addition, the number of operators and their experience, which are potential sources for introducing bias, were not reported in many studies. Moreover, a possible limitation of the present study could be pooling RCTs and pCCTs together (37) although, according to the sensitivity analyses (Supplementary Table 4), no statistically significant differences were found regarding the different study designs for any of the examined variables. Finally, only a limited number of studies was found assessing any outcomes in the long term (i.e. post-retention), which precluded extensive assessment about the stability of RFA treatment.

Conclusions

According to existing evidence on treatment effects of RFAs in the short term, i.e. directly after their removal, the following conclusions can be drawn:

Table 7. GRADE summary of findings table for the main outcomes of the systematic review directly after treatment with removable functional appliances. CI, confidence interval; Ctr, untreated control group; mo, month; RFA, removable functional appliance.

Patients: receiving orthodontic treatment to improve Class II malocclusion
Settings: university clinics (Brazil, Italy, New Zealand, Turkey, UK, USA) and hospitals (UK)
Intervention: RFAs (Activator, Bionator, Fränkel appliance, Sander appliance, Twin Block)
Comparison: untreated patients from follow-up or historical controls

Outcomes	Illustrative comparative risks (95% CI)		No. of patients (trials)	Quality of evidence (GRADE)	Comments
	Assumed risk Untreated (Ctr) patients	Corresponding risk RFA patients			
Annualized SNA change from baseline (follow-up: 9.0–28.0 months)*	The SNA increased on average by 0.19° per year in the Ctr groups (range −0.33° to 1.16°)	The mean SNA decreased in the RFA groups by 0.28° per year (95% CI: 0.12°–0.44° decrease) compared to the Ctr groups	570 (11)	⊕⊕⊕⊖ Moderate**	Effect magnitude affected by (i) appliance and (ii) number of appliance components (one or two)
Annualized SNB change from baseline (follow-up: 9.0–28.01 months)*	The SNB increased on average by 0.29° per year in the Ctr groups (range −0.27° to 1.15°)	The mean SNB additionally increased in the RFA groups by 0.62° per year (95% CI: 0.36°–0.88° increase) compared to the Ctr groups	570 (11)	⊕⊕⊖⊖ Low	Effect magnitude affected by (i) patient's gender, (ii) appliance, (iii) mode of action (bite-jumping or periosteal pull), (iv) number of appliance components (one or two), and (v) construction bite (single step or stepwise)
Annualized ANB change from baseline (follow-up: 9.0–28.0 months)*	The ANB decreased on average by 0.14° per year in the Ctr groups (range −0.30° to 0.53°)	The mean ANB additionally decreased in the RFA groups by 1.14° per year (95% CI: 0.77°–1.52° decrease) compared to the Ctr groups	528 (10)	⊕⊖⊖⊖ Very low***	Effect magnitude affected by (i) patient's gender, (ii) appliance, (iii) mode of action (bite-jumping or periosteal pull), (iv) number of appliance components (one or two), and (v) construction bite (single step or stepwise)
Annualized SN-ML change from baseline (follow-up: 10.8–28.0 months)*	The SN-ML decreased on average by 0.49° per year in the Ctr groups (range −0.79° to −0.24°)	The mean SN-ML increased in the RFA groups by 0.67° per year (95% CI: 0.17°–1.18° increase) compared to the Ctr groups	376 (8)	⊕⊕⊖⊖ Low	—
Annualized 1s-NL change from baseline (follow-up: 9.0–28.0 months)*	The 1s-NL increased on average by 0.35° per year in the Ctr groups (range −2.53° to 1.16°)	The mean 1s-NL decreased in the RFA groups by 6.33° per year (95% CI: 4.14°–8.53° decrease) compared to the Ctr groups	246 (5)	⊕⊕⊕⊕ High****	Effect magnitude affected by (i) appliance, (ii) mode of action (bite-jumping or periosteal pull), and (iii) number of appliance components (one or two)
Annualized 1i-ML change from baseline (follow-up: 8.0–28.0 months)*	The 1i-ML increased on average by 0.19° per year in the Ctr groups (range −2.27° to 0.67°)	The mean 1i-ML increased in the RFA groups by 1.37° per year (95% CI: 0.64°–2.10° increase) compared to the Ctr groups	539 (12)	⊕⊕⊖⊖ Low	Effect magnitude affected by (i) mode of action (bite-jumping or periosteal pull)
Annualized nasolabial angle change from baseline (follow-up: 8.0–18.0 months)*	The nasolabial angle increased on average by 0.16° per year in the Ctr groups (range: −2.09° to 1.81°)	The mean nasolabial angle additionally increased in the RFA groups by 2.78° per year (95% CI: 1.02°–4.55° increase) compared to the Ctr groups	253 (5)	⊕⊕⊖⊖ Low	—

All judgements start from 'low', due to the vast inclusion of non-randomized studies.

*From cephalometric analysis.

**Upgraded by one for (absence of) residual confounding; even though the studies were heterogeneous, our confidence in the effect direction is considerable: the 95% prediction interval (−0.56° to −0.001°) includes only favourable values.

***Downgraded by one for reporting bias; both the Egger and the Begg test confirm asymmetry of the funnel plot. Even though no sign of publication bias can be found, there is clear evidence of effect overestimation by small studies (small-study effects).

****Upgraded by two for effect magnitude; very large effect (cephalometric norm + 2 SDs + 1° for method error), which was included in the mean effect, the confidence interval and the prediction interval, while no serious limitations were found.

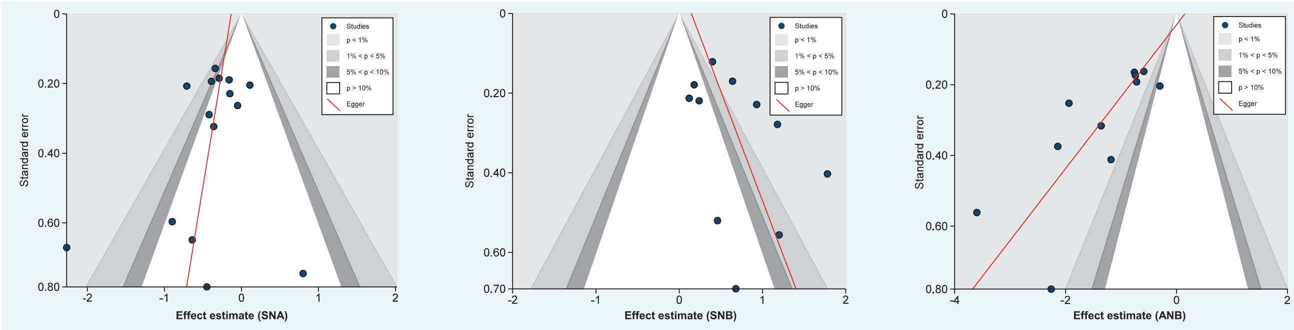

Figure 3. Contour-enhanced funnel plot of the SNA, SNB, and ANB angles for the studies included in the analysis for the comparison between RFAs and control groups.

- The skeletal effects of RFAs are minimal and, probably, of negligible clinical importance, when the natural growth is taken into account, i.e. when these were compared with untreated individuals.
- The treatment of Class II malocclusion with RFAs is associated with a minimal stimulation of mandibular growth, a minimal restriction of maxillary growth, and, to a much larger extent, with dentoalveolar and soft tissue changes.
- Patient- and appliance-related factors might influence the outcomes of RFA treatment.
- No assessment of the long-term effects of RFAs is possible at the present time.

As far as the clinical recommendations with the GRADE framework for the use of RFAs are concerned, only the changes in SNA and 1s-NL angles after the completion of RFA treatment can be adequately substantiated (high GRADE assessments):

- Clinicians might expect an average improvement of −0.28 degree per year in the SNA angle compared to untreated patients with the use of RFAs.
- Clinicians should expect a retroclination of the maxillary incisors of −6.33 degree per year through the 1s-NL angle with the use of RFAs compared to untreated patients.

Recommendations about the effects of RFAs on the skeletal growth of the mandible and the inclination of the mandibular incisors and the nasolabial angle are weaker and further research might influence them.

It seems that RFAs are not as effective as believed in correcting Class II malocclusion skeletally. More studies are needed to assess the skeletal, dental, and soft tissue outcomes in long-term, both after the period of fixed appliances that usually follows RFAs and the retention period. Ideally, these studies should report complete data of patient-related characteristics (gender, skeletal growth stage, and growth pattern of the patients), appliance-related characteristics (details of the appliance design, patient wear of the RFA), and details of the retention scheme. Finally, in order to minimize the magnification bias and be able to also use linear cephalometric measurements, authors should always report the magnification factor of the lateral cephalometric radiographs.

Supplementary material

Supplementary material is available at *European Journal of Orthodontics* online.

Acknowledgements

We thank the following authors for providing missing articles, clarifications, and/or additional data for the systematic review: C. Chabre (Université Paris Diderot, Paris, France), A. DiBiase (East Kent Hospitals NHS Trust, Kent, UK), L. Franchi (University of Florence, Florence, Italy), L. Frye (RWTH Aachen University, Aachen, Germany), A. Jamilian (Islamic Azad University, Tehran, Iran), G.R. Janson (University of São Paulo, Bauru, São Paulo, Brazil), Y. Luo (Zhejiang University School of Medicine, Zhejiang Province, China), J.R.C. Mew (Sussex, United Kingdom), P. Ngan (West Virginia University, Morgantown, USA), J. Philippe (Chartres, France), P.J. Sandler (Chesterfield and North Derbyshire Royal Hospital NHS Trust, Chesterfield, United Kingdom), C. Sharp (Auckland, New Zealand), P. Stüber (Kreiskrankenhaus Königs Wusterhausen, Königs Wusterhausen, Germany), and S. Tecco (University of Chieti, Italy).

References

1. Josefsson, E., Bjerklin, K. and Lindsten, R. (2007) Malocclusion frequency in Swedish and immigrant adolescents-influence of origin on orthodontic treatment need. *European Journal of Orthodontics*, 29, 79–87.
2. Ast, D.B., Carlos, J.P. and Cons, D.C. (1965) Prevalence and characteristics of malocclusion among senior high school students in upstate New York. *American Journal of Orthodontics*, 51, 437–445.
3. Goldstein, M.S. and Stanton, F.L. (1936) Various types of occlusion and amounts of overbite in normal and abnormal occlusion between two and twelve years. *International Journal of Orthodontia and Oral Surgery*, 22, 549–569.
4. McNamara, J.A., Jr (1981) Components of class II malocclusion in children 8–10 years of age. *The Angle Orthodontist*, 51, 177–202.
5. Pancherz, H., Zieber, K. and Hoyer, B. (1997) Cephalometric characteristics of Class II division 1 and Class II division 2 malocclusions: a comparative study in children. *The Angle Orthodontist*, 67, 111–120.
6. Stockly, P. and Willert, H. (1971) Tissue reactions in the temporomandibular joint resulting from anterior displacement of the mandible in the monkey. *American Journal of Orthodontics*, 60, 142–155.
7. McNamara, J.A., Jr and Bryan, F.A. (1987) Long term mandibular adaptations to protrusive function: an experimental study in *Macaca mulatta*. *American Journal of Orthodontics and Dentofacial Orthopedics*, 92, 98–108.
8. Charlier, J.P., Petrovic, A.G. and Stutzmann, J. (1969) Effects of mandibular hyperpropulsion on the prechondroblastic zone of rat condyle. *American Journal of Orthodontics and Dentofacial Orthopedics*, 55, 71–74.
9. Woodside, D.G., Metaxas, A. and Altuna, G. (1987) The influence of functional appliance therapy on glenoid fossa remodeling. *American Journal of Orthodontics and Dentofacial Orthopedics*, 92, 181–198.
10. de Almeida, M.R., Henriques, J.F., de Almeida, R.R. and Ursi, W. (2002) Treatment effects produced by Frankel appliance in patients with Class II, division 1 malocclusion. *The Angle Orthodontist*, 72, 418–425.

11. Toth, L.R. and McNamara, J.A., Jr (1999) Treatment effects produced by the twin-block appliance and the FR-2 appliance of Fränkel compared with an untreated Class II sample. *American Journal of Orthodontics and Dentofacial Orthopedics*, 116, 597–609.
12. Marsico, E., Gatto, E., Burrascano, M., Matarese, G. and Cordasco, G. (2011) Effectiveness of orthodontic treatment with functional appliances on mandibular growth in the short term. *American Journal of Orthodontics and Dentofacial Orthopedics*, 139, 24–36.
13. Pancherz, H., Ruf, S. and Kohlhas, P. (1998) Effective condylar growth and chin position changes in Herbst treatment: a cephalometric roentgenographic long-term study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 114, 437–446.
14. Ruf, S., Baltromejus, S. and Pancherz, H. (2001) Effective condylar growth and chin position in activator treatment: a cephalometric roentgenographic study. *The Angle Orthodontist*, 71, 4–11.
15. Serbesis-Tsarudis, C. and Pancherz, H. (2008) "Effective" TMJ and chin position changes in Class II treatment. *The Angle Orthodontist*, 78, 813–818.
16. Voudouris, J.C., Woodside, D.G., Altuna, G., Angelopoulos, G., Bourque, P.J., Lacouture, C.Y. and Kuftinec, M.M. (2003) Condyle-fossa modifications and muscle interactions during Herbst treatment, Part 2. Results and conclusions. *American Journal of Orthodontics and Dentofacial Orthopedics*, 124, 13–29.
17. Creekmore, T.D. and Radney, L.J. (1983) Fränkel appliance therapy: orthopedic or orthodontic? *American Journal of Orthodontics and Dentofacial Orthopedics*, 83, 89–108.
18. Giannelly, A., Bronson, P., Martignoni, M. and Bernstein, L. (1983) Mandibular growth, condylar position and Fränkel appliance therapy. *The Angle Orthodontist*, 53, 131–142.
19. Cozza, P., Baccetti, T., Franchi, L., De Toffol, L. and McNamara, J.A., Jr (2006) Mandibular changes produced by functional appliances in Class II malocclusion: a systematic review. *American Journal of Orthodontics and Dentofacial Orthopedics*, 129, 599.e1–599.e12.
20. Pancherz, H. (1984) A cephalometric analysis of skeletal and dental changes contributing to Class II correction in activator treatment. *American Journal of Orthodontics*, 85, 125–134.
21. Vargervik, K. and Harvold, E.P. (1985) Response to activator treatment in Class II malocclusions. *American Journal of Orthodontics and Dentofacial Orthopedics*, 88, 242–251.
22. Hashim, H.A. (1991) Analysis of activator treatment changes. *Australian Orthodontic Journal*, 12, 100–104.
23. Sidhu, M.S., Kharbanda, O.P. and Sidhu, S.S. (1995) Cephalometric analysis of changes produced by a modified Herbst appliance in the treatment of Class II division 1 malocclusion. *British Journal of Orthodontics*, 22, 1–12.
24. McNamara, J.A., Bookstein, F.L. and Shaughnessy, T.G. (1985) Skeletal and dental changes following functional regulator therapy on Class II patients. *American Journal of Orthodontics and Dentofacial Orthopedics*, 88, 91–111.
25. Wieslander, L. and Lagerström, L. (1979) The effect of activator treatment on Class II malocclusions. *American Journal of Orthodontics*, 75, 20–26.
26. Robertson, N.R.E. (1983) An examination of treatment changes in children treated with the function regulator of Fränkel. *American Journal of Orthodontics*, 83, 299–310.
27. Sahm, G., Bartsch, A. and Witt, E. (1990) Micro-electronic monitoring of functional appliance wear. *European Journal of Orthodontics*, 12, 297–301.
28. Bertl, M.H., Mandl, C. and Crismani, A.G. (2011) Do functional orthodontic appliances stimulate mandibular growth in class II division 1 patients? *International Journal of Stomatology & Occlusion Medicine*, 4, 45–53.
29. Chen, J.Y., Will, L.A. and Niederman, R. (2002) Analysis of efficacy of functional appliances on mandibular growth. *American Journal of Orthodontics and Dentofacial Orthopedics*, 122, 470–476.
30. Harrison, J.E., O'Brien, K.D. and Worthington, H.V. (2007) Orthodontic treatment for prominent upper front teeth in children. *Cochrane Database of Systematic Reviews*, 3, CD003452.
31. Niu, Y. and Zhou, H. (2011) Effect on functional appliances on mandibular growth on skeletal Class II malocclusion: a systematic review. *West China Journal of Stomatology*, 29, 384–388.
32. Perillo, L., Cannavale, R., Ferro, F., Franchi, L., Masucci, C., Chiodini, P. and Baccetti, T. (2011) Meta-analysis of skeletal mandibular changes during Frankel appliance treatment. *European Journal of Orthodontics*, 33, 84–92.
33. Barnett, G.A., Higgins, D.W., Major, P.W. and Flores-Mir, C. (2008) Immediate skeletal and dentoalveolar effects of the crown- or banded type Herbst appliance on Class II division 1 malocclusion. *The Angle Orthodontist*, 78, 361–369.
34. Flores-Mir, C. and Major, P.W. (2006) A systematic review of cephalometric facial soft tissue changes with the Activator and Bionator appliances in Class II division 1 subjects. *European Journal of Orthodontics*, 28, 586–593.
35. Flores-Mir, C. and Major, P.W. (2006) Cephalometric facial soft tissue changes with the twin block appliance in Class II division 1 malocclusion patients. A systematic review. *The Angle Orthodontist*, 76, 876–881.
36. Antonarakis, G.S. and Kiliaridis, S. (2007) Short-term anteroposterior treatment effects of functional appliances and extraoral traction on Class II malocclusion. *The Angle Orthodontist*, 77, 907–914.
37. Higgins, J.P.T. and Green, S. (eds) (2011) Cochrane handbook for systematic reviews of interventions. Version 5.1.0, updated March 2011. The Cochrane Collaboration. http://handbook.cochrane.org/ (18 January 2014, date last accessed).
38. Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D. (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Journal of Clinical Epidemiology*, 62, e1–e34.
39. Beller, E.M., Glasziou, P.P., Altman, D.G., Hopewell, S., Bastian, H., Chalmers, I., Gøtzsche, P.C., Lasserson, T. and Tovey, D.; PRISMA for Abstracts Group (2013) PRISMA for abstracts: reporting systematic reviews in journal and conference abstracts. *PLOS Medicine*, 10, e1001419.
40. Athanasiou, A.E. (1995) *Orthodontic Cephalometry*. Mosby-Wolfe, London, UK.
41. Adams, J.W. (1940) Correction of errors in cephalometric roentgenograms. *The Angle Orthodontist*, 10, 3–13.
42. Gandini, P., Mancini, M. and Andreani, F. (2006) A comparison of hand-wrist bone and cervical vertebral analyses in measuring skeletal maturation. *The Angle Orthodontist*, 76, 984–989.
43. Rasool, G., Bashir, U. and Kundi, I. (2010) Comparative evaluation between cervical vertebrae and hand-wrist maturation for assessment of skeletal maturity orthodontic patients. *Pakistan Oral & Dental Journal*, 30, 85–95.
44. Alkhal, H.A., Wong, R.W.K. and Rabie, B.M. (2008) Correlation between chronological age, cervical vertebral maturation and Fishman's skeletal maturity indicators in Southern Chinese. *The Angle Orthodontist*, 78, 591–596.
45. Bala, M., Pathak, A. and Jain, R.L. (2010) Assessment of skeletal age using MP 3 and hand-wrist radiographs and its correlation with dental and chronological ages in children. *Journal of the Indian Society of Pedodontics and Preventive Dentistry*, 28, 95–99.
46. Illing, H.M., Morris, D.O. and Lee, R.T. (1998) A prospective evaluation of Bass, Bionator and Twin Block appliances. Part I—The hard tissues. *European Journal of Orthodontics*, 20, 501–516.
47. Downs, S.H. and Black, N. (1998) The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health*, 52, 377–384.
48. Harmelink, K. (2010) Effects of neurolytic blocks (botulinum toxin type A and motor branch block) in patients with a Stiff Knee Gait: A systematic review. Master Thesis, Utrecht University.
49. Morissette, K., Tricco, A.C., Horsley, T., Chen, M.H. and Moher, D. (2011) Blinded versus unblended assessments of risk of bias in studies included in a systematic review. *The Cochrane Database of Systematic Reviews*, 9, MR000025.
50. Peters, J.L., Sutton, A.J., Jones, D.R., Abrams, K.R. and Rushton, L. (2008) Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology*, 61, 991–996.
51. Begg, C.B. and Mazumdar, M. (1994) Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50, 1088–1101.

52. Egger, M., Smith, G.D., Schneider, M. and Minder, C. (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315, 629–634.
53. Duval, S. and Tweedie, R. (2000) Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56, 455–463.
54. Guyatt, G.H., Oxman, A.D., Schünemann, H.J., Tugwell, P. and Knottnerus, A. (2011) GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. *Journal of Clinical Epidemiology*, 64, 380–382.
55. Sloan, J., Symonds, T., Vargas-Chanes, D. and Fridley, B. (2003) Practical guidelines for assessing the clinical significance of health-related quality of life changes within clinical trials. *Drug Information Journal*, 37, 23–31.
56. Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46.
57. DerSimonian, R. and Laird, N. (1986) Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7, 177–188.
58. Higgins, J.P., Thompson, S.G. and Spiegelhalter, D.J. (2009) A re-evaluation of random effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172, 137–159.
59. Graham, P.L. and Moran, J.L. (2012) Robust meta-analytic conclusions mandate the provision of prediction intervals in meta-analysis summaries. *Journal of Clinical Epidemiology*, 65, 503–510.
60. Ioannidis, J.P. (2008) Interpretation of tests of heterogeneity and bias in meta-analysis. *Journal of Evaluation in Clinical Practice*, 14, 951–957.
61. Borenstein, M., Hedges, L.V., Higgins, J.P.T. and Rothstein, H.R. (2009) *Introduction to Meta-Analysis*. John Wiley and Sons Ltd., Chichester, UK.
62. Almeida-Pedrin, R.R., de Almeida, M.R., de Almeida, R.R., Pinzan, A. and Carvalho Ferreira, F.P. (2007) Treatment effects of headgear biteplane and bionator appliances. *American Journal of Orthodontics and Dentofacial Orthopedics*, 132, 191–198.
63. Andreoli, F.A.M. (2008) Avaliação cefalométrica da estabilidade do tratamento com o regulador de função Fränkel-2. Master Thesis, Methodist University of São Paulo.
64. Angelieri, F., Franchi, L., Cevidanes, L.H., Scanavini, M.A. and McNamara, J.A. (2014) Long-term treatment effects of the FR-2 appliance: a prospective evalution 7 years post-treatment. *European Journal of Orthodontics*, 36, 192–199.
65. Baysal, A. and Uysal, T. (2011) Soft tissue effects of Twin Block and Herbst appliances in patients with Class II division 1 mandibular retrognathy. *European Journal of Orthodontics*, 35, 71–81.
66. Brunharo, I.H.V.P., Quintão, C.A., Almeida, M.A.D.O., Motta, A. and Barreto, S.Y.N. (2011) Dentoskeletal changes in Class II malocclusion patients after treatment with the Twin Block functional appliance. *Dental Press Journal of Orthodontics*, 16, e1–e8.
67. Cevidanes, L.H.S., Franco, A.A., Scanavini, M.A., Vigorito, J.W., Enlow, D.H. and Proffit, W.R. (2003) Clinical outcomes of Fränkel appliance therapy assessed with a counterpart analysis. *American Journal of Orthodontics and Dentofacial Orthopedics*, 123, 379–387.
68. Courtney, M., Harkness, M. and Herbison, P. (1996) Maxillary and cranial base changes during treatment with functional appliances. *American Journal of Orthodontics and Dentofacial Orthopedics*, 109, 616–624.
69. Nelson, C., Harkness, M. and Herbison, P. (1993) Mandibular changes during functional appliance treatment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 104, 153–161.
70. Dolce, C., Babh, L.K., McGorray, S.P., Taylor, M.G., King, G.J. and Wheeler, T.T. (2002) Vertical skeletal and dental changes in early treatment of class II malocclusion. *Seminars in Orthodontics*, 8, 141–148.
71. Ehmer, U., Tulloch, C.J., Proffit, W.R. and Phillips, C. (1999) An international comparison of early treatment of angle Class-II/1 cases. Skeletal effects of the first phase of a prospective clinical trial. *Journal of Orofacial Orthopedics*, 60, 392–408.
72. Morris, D.O., Illing, H.M. and Lee, R.T. (1998) A prospective evaluation of Bass, Bionator and Twin Block appliances. Part II—The soft tissues. *European Journal of Orthodontics*, 20, 663–684.
73. Jakobsson, O.S. (1967) Cephalometric evaluation of treatment effect on Class II, Division 1 malocclusions. *American Journal of Orthodontics*, 53, 446–457.
74. Janson, G.R.P., Toruño, J.L.A., Martins, D.R., Henriques, J.F.C and de Freitas, M.R. (2003) Class II treatment effects of the Fränkel appliance. *European Journal of Orthodontics*, 25, 301–309.
75. Lund, D.I. and Sandler, P.J. (1998) The effects of Twin Blocks: a prospective controlled study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 113, 104–110.
76. Martina, R., Cioffi, I., Galeotti, A., Tagliaferri, R., Cimino, R., Michelotti, A., Valletta, R., Farella, M. and Paduano S. (2013) Efficacy of the Sander bite-jumping appliance in growing patients with mandibular retrusion: a randomized controlled trial. *Orthodontics & Craniofacial Research*, 16, 116–126.
77. Quintão, C., Helena, I., Brunharo, V.P., Menezes, R.C. and Almeida, M.A. (2006) Soft tissue facial profile changes following functional appliance therapy. *European Journal of Orthodontics*, 28, 35–41.
78. Tulloch, J.F.C., Phillips, C., Koch, G. and Proffit, W.R. (1997) The effect of early intervention on skeletal pattern in Class H malocclusion: a randomized clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics*, 111, 391–400.
79. Varlik, S.K., Gültan, A. and Tümer, N. (2008) Comparison of the effects of Twin Block and activator treatment on the soft tissue profile. *European Journal of Orthodontics*, 30, 128–134.
80. Bishara, S.E., Jamison, J.E., Peterson, L.C. and DeKock, W.H. (1981) Longitudinal changes in standing height and mandibular parameters between the ages of 8 and 17 years. *American Journal of Orthodontics*, 80, 115–135.
81. Jakobsson, S.O. and Paulin, G. (1990) The influence of activator treatment on skeletal growth in Angle Class II: 1 cases. A roentgenocephalometric study. *European Journal of Orthodontics*, 12, 174–184.
82. Luder, H.U. (1982) Skeletal profile changes related to two patterns of activator effects. *American Journal of Orthodontics*, 81, 390–396.
83. Greco, M., Fichera, G., Caltabiano, M., Barbato, E. and Leonardi, R. (2010) Short-term effects of the activator in skeletal class II division 1 patients with different vertical skeletal pattern. A retrospective study. *Minerva Stomatologica*, 59, 61–74.
84. Hönn, M., Schneider, C., Dietz, K., Godt, A. and Göz, G. (2006) Treating Class II patients with removable plates and functional orthopedic appliances—the importance of anterior tooth inclination and direction of growth on treatment outcome. *Journal of Orofacial Orthopedics*, 67, 272–288.
85. Hansen, K., Pancherz, H. and Hägg, U. (1991) Long-term effects of the Herbst appliance in relation to the treatment growth period: a cephalometric study. *European Journal of Orthodontics*, 13, 471–481.
86. Baccetti, T., Franchi, L., Toth, L.R. and McNamara, J.A., Jr (2000) Treatment timing for twin block therapy. *American Journal of Orthodontics and Dentofacial Orthopedics*, 118, 159–170.
87. Faltin, K.J., Faltin, R.M., Baccetti, T., Franchi, L., Ghiozzi, B. and McNamara, J.A., Jr (2003) Long-term effectiveness and treatment timing for Bionator therapy. *The Angle Orthodontist*, 73, 221–230.
88. Dolce, C., McGorray, S.P., Brazeau, L., King, G.J. and Wheeler, T.T. (2007) Timing of Class II treatment: skeletal changes comparing 1-phase and 2-phase treatment. *American Journal of Orthodontics and Dentofacial Orthopedics*, 132, 481–489.
89. Franchi, L., Pavoni, C., Faltin, K., Jr, McNamara, J.A., Jr and Cozza, P. (2013) Long-term skeletal and dental effects and treatment timing for functional appliances in Class II malocclusion. *The Angle Orthodontist*, 83, 334–340.
90. Jena, A.K., Duggal, R. and Parkash, H. (2006) Skeletal and dentoalveolar effects of Twin-block and bionator appliances in the treatment of Class II malocclusion: a comparative study. *American Journal of Orthodontics and Dentofacial Orthopedics*, 130, 594–602.
91. Clark, W.J. (1988) The twin block technique. A functional orthopedic appliance system. *American Journal of Orthodontics and Dentofacial Orthopedics*, 93, 1–18.
92. Collett, A.R. (2000) Current concepts on functional appliances and mandibular growth stimulation. *Australian Dental Journal*, 45, 173–178.
93. Falck, F. and Fränkel, R. (1989) Clinical relevance of step-by-step mandibular advancement in the treatment of mandibular retrusion using the Fränkel appliance. *American Journal of Orthodontics and Dentofacial Orthopedics*, 96, 333–341.
94. Banks, P., Wright, J. and O'Brien, K. (2004) Incremental versus maximum bite advancement during twin-block therapy: a randomized controlled clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics*, 126, 583–588.
95. Casutt, C., Pancherz, H., Gawora, M. and Ruf, S. (2007) Success rate and efficiency of activator treatment. *European Journal of Orthodontics*, 29, 614–621.