Changes in apical base sagittal relationship in Class II malocclusion treatment with and without premolar extractions: *A systematic review and meta-analysis*

Guilherme Janson[a]; Aron Aliaga-Del Castillo[b]; Ana Niederberger[b]

ABSTRACT
Objective: To evaluate the changes in apical base sagittal relationship in Class II treatment with and without premolar extractions.
Materials and Methods: Controlled studies evaluating ANB angle changes after Class II Division 1 malocclusion treatment with or without premolar extractions were considered. Electronic databases (PubMed, Embase, Web of Science, Scopus, The Cochrane Library, Lilacs, and Google Scholar) without limitations regarding publication year or language were searched. Risk of bias was assessed with Risk Of Bias in Non-randomized Studies—of Interventions tool of the Cochrane Collaboration. Mean difference (MD) and 95% confidence interval (CI) were calculated from the random-effects meta-analysis. Subgroup and sensitivity analyses were also performed.
Results: Twenty-five studies satisfied the inclusion criteria and were included in the qualitative synthesis. Eleven nonextraction and only one extraction Class II treatment studies presented untreated Class II control group. Therefore, meta-analysis was performed only for the nonextraction protocol. In treated Class II nonextraction patients, the average of the various effects was a reduction in the ANB angle of 1.56° (95% CI: 1.03, 2.09, $P < .001$) compared with untreated Class II subjects. Class II malocclusions treated with two maxillary-premolar extractions and four-premolar extractions produced estimated mean reductions in ANB of −1.88° and −2.55°, respectively. However, there is a lack of low-risk-of-bias studies.
Conclusions: According to the existing low quality evidence**,** the apical base sagittal relationship in nonextraction, two-maxillary and four-premolar extractions Class II treatments decreases −1.56°, 1.88° and 2.55°, respectively. Further studies are necessary to obtain more robust information. (*Angle Orthod.* 2017;87:338–355)

KEY WORDS: Orthodontics; Malocclusion, Angle Class II; Systematic review

INTRODUCTION

The ANB angle has frequently been used to evaluate the skeletal sagittal severity of Class II malocclusions.[1–6] However, it seems that there has been excessive importance given to the anteroposterior discrepancy depicted by a cephalometric variable to evaluate the actual treatment difficulty in correcting the Class II occlusal anteroposterior discrepancy.[7,8]

To correct a complete Class II malocclusion without premolar extractions (XP0) or with four-premolar extractions (XP4), Class I molar and canine relationships must be obtained.[2,5] If the correction is performed with 2 maxillary premolar extractions (XP2), the molars will finish in a complete Class II and the canines in Class I relationships.[9,10] A Class I canine relationship will allow correction of the severe pretreatment overjet to a normal overjet.[9] Therefore, the occlusal Class II anteroposterior discrepancy is more informative regarding treatment difficulty than a cephalometric variable.[11,12]

Additionally, in cases wherein the cephalometric anteroposterior discrepancy is accentuated, as evaluated by the ANB, even if the severe occlusal Class II

[a] Professor and Head, Department of Orthodontics, Bauru Dental School, University of São Paulo, São Paulo, Brazil.
[b] Graduate Student, Department of Orthodontics, Bauru Dental School, University of São Paulo, São Paulo, Brazil.
Corresponding author: Dr Guilherme Janson, Department of Orthodontics, Bauru Dental School, University of São Paulo, Alameda Octávio Pinheiro Brisolla 9-75, Bauru, São Paulo 17012-901, Brazil
(e-mail: jansong@travelnet.com.br)

Accepted: July 2016. Submitted: March 2016.
Published Online: September 6, 2016

 DOI: 10.2319/030716-198.1

anteroposterior discrepancy has been completely corrected, the ANB may not be significantly reduced to its standard value.[13–19] That is, reduction of the apical base anteroposterior discrepancy only by orthodontic means is very limited, as has been demonstrated by some studies.[2,4,10,13–20] However, to provide stronger scientific evidence to this fact, a systematic review including controlled clinical trials assessing the change in ANB angle in Class II malocclusion patients treated with or without premolar extractions was conducted.

MATERIALS AND METHODS

This systematic review is reported according to the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement.[21] The systematic review protocol was registered at PROSPERO database (http://www.crd.york.ac.uk/PROSPERO, CRD42015026677).

Eligibility Criteria

1. Participants: Growing patients with Class II Division 1 malocclusion.
2. Intervention: Class II treatment with or without premolar extractions, all using multibracket appliance (MBA) treatment.
3. Comparison: Class II subjects with or without orthodontic treatment, all after MBA treatment.
4. Outcome: ANB angle treatment changes.
5. Study design: Controlled clinical trials (randomized, prospective, or retrospective).

Exclusion Criteria

Studies of patients having craniofacial anomalies, adults, surgical-orthodontic treatment, or absence of a control group; and systematic reviews and meta-analyses.

Electronic databases (Pubmed, Embase, Web of Science, Scopus, The Cochrane Library, Lilacs) and a partial grey literature (academic literature that is not formally published) through Google Scholar without limitations regarding publication year or language were searched until June 20, 2016 (Appendix 1). In addition, the evaluators went through the reference lists of the selected articles to ensure that no potential articles were missed.

Two evaluators independently screened the titles and abstracts identified from the electronic database results after elimination of duplicates. Next, full articles were retrieved to confirm their eligibility. The same evaluators selected the articles for inclusion in the qualitative synthesis independently. Disagreements were resolved by verbal discussion between them and by consultation with another evaluator when necessary.

The following data were extracted independently by the two reviewers: study design, participants, interventions, initial and final ANB angle or ANB angle mean change, treatment duration, Class II diagnosis, and treatment timing. Factors for subgroup analyses were selected a priori to evaluate any influence of them on the ANB change after treatment. These factors included (1) patient's sex, (2) skeletal growth stage based on the cervical vertebral maturation method or on hand-wrist radiographs, (3) patient's growth pattern, (4) type of appliance used (functional appliances [FA] + MBA or headgear [HG] + MBA, (5) treatment time (up to or greater than 24 months).

The risk of bias (RoB) in individual studies was assessed using Cochrane Collaboration's ROBINS-I tool (Risk Of Bias in Non-randomized Studies—of Interventions).[22] For all included studies, the RoB for each domain and the overall RoB for each study were judged as Low, Moderate, Serious, Critical, or No information (Appendix 2).[22] When more than 10 studies were identified, standard funnel plots and Egger's test were planned to identify publication bias.[23]

The quality of evidence for the main outcome was rated by using the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) approach.[24] Evaluation of RoB in individual studies and across studies were independently performed by two evaluators. Any disagreement was resolved through verbal discussion between the evaluators and with another third evaluator.

Data were summarized in three groups according to the protocol used: XP0, XP4, and XP2, to independently evaluate the ANB change in each group. Performance of a meta-analysis was possible only in the XP0 comparison. When data were summarized from the premolar extraction studies, only one[5] had an untreated Class II control group (UCIICG). Nonetheless, the evaluators considered it important to show the data (without performing a meta-analysis) from the extraction protocols, as well.

For the XP0 group, mean difference and its 95% confidence interval (CI) were obtained. The random-effects model was chosen, supported by both clinical and statistical reasoning.[25,26] The between-study heterogeneity/inconsistency was assessed by inspecting the forest plot and calculating Tau^2, Chi^2, and I^2 statistics, respectively (the interpretation of I^2 was made together with the *P* value for the Chi^2 and the 95% CI for I^2).[23] The 95% prediction interval was also calculated.

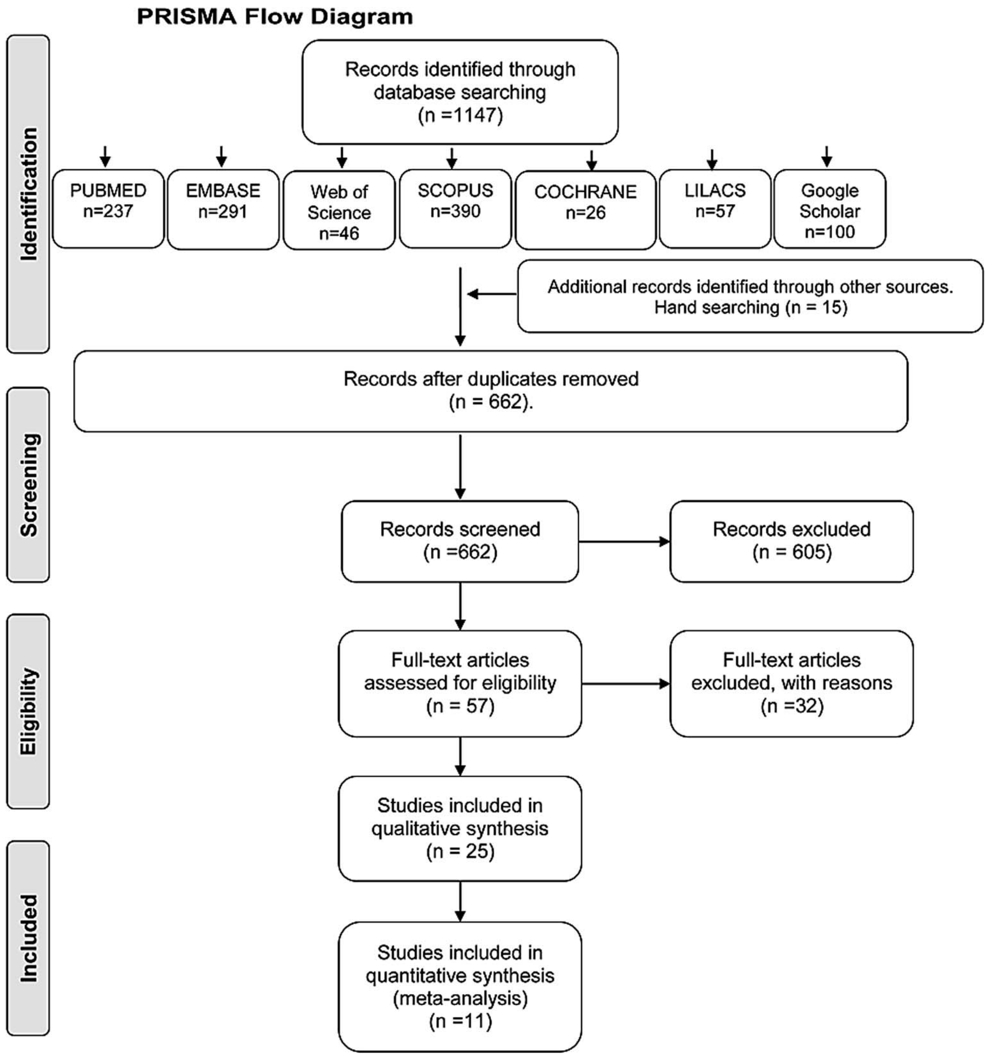


Figure 1. PRISMA flow diagram.

For the XP4 and XP2 groups, an estimated ANB angle mean change and standard deviation were obtained by combining studies data of each group.[23] Sources of heterogeneity in meta-analysis were evaluated with subgroup analyses. Differences between subgroups was assessed by considering the *P* value of the standard test for heterogeneity across subgroup results.[23] Sensitivity analyses based on study design and precision were performed to check the robustness of the results.[23]

The statistical analyses were performed using RevMan statistical software (version 5.3 for Windows; Nordic Cochrane Centre, Rigshospitalet, Denmark). Comprehensive meta-analysis statistical software (version 3.0; Biostat Inc, Englewood, NJ) was used only to identify publication bias (funnel plots and Egger's regression test).[23] All *P* values were two sided with $\alpha = 5\%$, except for the tests of between-studies or between-subgroups heterogeneity ($\alpha = 10\%$).[23]

RESULTS

Initially, 1147 records were identified, and 15 hand-searched articles were added. After exclusion of duplicates, 662 studies remained. The full texts of 57 articles were obtained and assessed for eligibility, and 32 articles were excluded with reasons, leaving 25 articles for qualitative analysis (Figure 1, Table 1). Of the 25 articles, only 1 was a prospective, controlled clinical trial[18] and 24 were retrospective, controlled clinical trials. No randomized, controlled clinical trials satisfying the inclusion criteria were found.

XP0

Twenty-three articles[2–6,10,14–19,27–37] were characterized by reporting samples with nonextraction treatment. Thirteen articles[3,5,14,17–19,27–29,31,33,35,36] reported a UCIICG. Three[3,17,19] did not report the ANB change standard deviation. Standard deviation could be estimated for only one study.[17] Therefore, 11 articles were included in the meta-analysis (nonextraction Class II vs UCIICG). All 11 studies reported one-phase treatment either with an FA or HG followed by MBA. The FA reported in seven studies were: Forsus fatigue resistant device,[27,28] Twin-block,[28] MARA,[17,29] AdvanSync,[29] Sydney Magnoglide,[18] Jasper Jumper,[33] and Herbst.[5] HG use was reported in five studies,[5,14,31,33,36] and one study reported the use of HG-activator combination.[35] Treatment timing was reported in five studies.[14,17,18,28,29] Two studies[18,28] reported combined data of various pubertal stages. Two other studies[17,29] reported data of pubertal subjects only and another[14] presented separate data of subjects in different pubertal stages. Only 2 studies[18,33] of the 11 had a treatment duration of 24 months or less. The mean treatment duration of the XP0 protocol was 30.39 months.

XP4

Nine[2,4,5,13,16,20,30,32,37] of the 25 articles included samples with XP4. Three articles[2,13,20] did not present standard deviations of the ANB change so they could not be estimated. The remaining six articles were included for analysis. Treatment timing was reported in three articles, two articles[30,32] presented data of pubertal patients, and one article[16] showed combined data of prepubertal and pubertal patients. They had a mean treatment duration of 34.9 months. Only one article[5] had an UCIICG; the other five did not. Although a meta-analysis should not be performed in this situation, the evaluators calculated the mean change of the six articles with the intention of evaluating the ANB angle behavior with this treatment protocol.

XP2

Three articles[10,15,20] of the 25 included samples with XP2. One article[20] was excluded because it did not present the standard deviation of ANB mean change and it could not be estimated. They did not report treatment timing information and they had a mean treatment duration of 27.96 months. They did not have an UCIICG, preventing a meta-analysis. Nevertheless, the mean change of the two articles was calculated.

After the assessment of RoB, an overall RoB for each study was assigned (Table 2). No study showed an overall Low RoB, so none could not be comparable to a well-performed randomized trial. Eighteen studies* showed overall Moderate RoB. Although these 18 studies presented Moderate RoB, a meta-analysis was performed with only 11 studies that reported data from UCIICG. Therefore, only a comparative analysis of nonextraction vs. UCIICG was performed. Seven studies[2,3,6,13,19,20,34] presented Serious RoB.

ANB angle mean changes and standard deviations of the studies included for quantitative analyses are summarized in Table 3. A meta-analysis was performed only regarding the XP0 protocol because of the absence of UCIICG in the studies involving premolar extractions.

XP0

In nonextraction Class II patients, the average of the various effects was a reduction in the ANB angle of 1.56° compared with UCIICG (Figure 2, Table 4). Multi-arm studies[5,28,29,33] were pooled prior to meta-analysis, except for one study,[14] which reported three different groups with specific control groups for each one, so 13 groups of nonextraction treatment present in the 11 studies were used.

Regarding the risk of publication bias, the funnel plots as Egger's regression test ($P = .16$) did not show asymmetry (Figure 3). The evidence rated by the GRADE approach was considered as Low quality (Table 5).

Subgroup analyses including patients' sex and growth pattern were not feasible due to lack of reporting data in the studies. The ANB change varied according to the skeletal growth spurt and type of appliance used. Treatment in prepeak and peak patients and the HG-activator combination + MBA induced greater decreases on ANB. (Table 4, Appendixes 3 and 4) Treatment duration did not produce significant differences between subgroups (Table 4, Appendix 5). Sensitivity analysis based on the study design did not find significance difference between the prospective and retrospective studies. Sensitivity analysis by precision, selecting studies[14,17,18,28,36] with narrower confidence intervals, showed a smaller decrease in ANB angle compared with the original one; however, it was statistically significant when compared with the UCIICG (Table 4).

XP4 and XP2

The estimated ANB reductions obtained without using data from UCIICG were 2.55° and 1.88° for the XP4 and for the XP2 treatment, respectively.

DISCUSSION

This systematic review is one of the few reviews[38,39] that include studies with completed Class II malocclusion treatment with MBA in growing patients using protocols with or without premolar extractions. The purpose was to obtain an estimate of anteroposterior

* References 4, 5, 10, 14–18, 27–33, 35–37.

Table 1. Characteristics of the 25 Articles Included in the Quality Assessment

		Participants			Outcome	
Study	D[a]	Age	N (M/F)	Intervention	ANB Change	SD
Dada et al., 2015	R	12	19 (11/8)	NE. Forsus (6 mo) + MBA. 0.018-inch slot; preadjusted brackets with −6° of torque on mandibular incisors. 0.016 × 0.022-inch SS AW and less than 6 months of Class II (Cl II) elastics use during MBA treatment.	−1.88	1.07
		Matched by age and sex	19 (–)	Untreated Cl II control group.	–0.33	0.98
Giuntini et al., 2015	R	12.4	28 (9/19)	NE. Twin-block (13.2 mo) + MBA. 0.022-inch slot; preadjusted brackets.	−2.6	1.3
		12.3	36 (20/16)	NE. Forsus (6 mo) + MBA. 0.022-inch slot; preadjusted brackets with −6° of torque on mandibular incisors. 0.019 × 0.025-inch SS AW.	−1.8	1.3
		12.2	27 (13/14)	Untreated Cl II control group.	−0.4	1
Zheng et al., 2015*	R	11.9	30 (13/17)	NE. HG (3.5 mo; more than 14 h/d, force of 300–350 g, traction direction depended on vertical facial types) + Cl II elastic traction + MBA.	−1.1	–
		11.9	30 (15/15)	NE. Cl II traction (elastics) + MBA.	−1.52	–
		11.9	30 (14/16)	Untreated Cl II control group.	0.16	–
Al-Jewair et al., 2012	R	11.6	40 (22/18)	NE. MARA (12 mo) + MBA. 0.022-inch slot; preadjusted brackets with a built-in labial root torque on mandibular incisors.	−2.4	1.6
		12.3	30 (13/17)	NE. AndvanSync (12 mo) + MBA. 0.022-inch slot; preadjusted brackets with a built-in labial root torque on mandibular incisors.	−2.6	1.9
		11.9	24 (13/11)	Untreated Cl II control group.	−0.4	1.5
Pangrazio et al., 2012**	R	11.35	30 (12/18)	NE. MARA (15.6 mo) + MBA. Edgewise brackets.	−0.63	0.81
		Matched by skeletal age	21(–)	Untreated Cl II control group.	−0.4	0.81
					$P = .35$	
Phelan et al., 2012	P	13.5	31 (19/12)	NE. Sydney Magnoglide (12 mo) + MBA. Straight-wire brackets. Cl II elastics (1/4-inch, 3.5 oz) were worn for 4 mo during MBA treatment.	−1	1
		13	30 (15/15)	Untreated Cl II control group.	0.3	1.1
Ye et al., 2012	R	13.4	70 (33/37)	NE. Twin-block (12 mo) + MBA.	−2.77	1.32
		13.2	76 (33/43)	4 first-premolar extractions + MBA. Straightwire edgewise brackets. Cl II elastics and high-pull HG connected to J-hook on maxillary AW distal to central incisors with a resultant force to control occlusal pane.	−3.04	2.13
Gkantidis et al., 2011	R	11.8	29 (13/16)	4 first-premolar extractions + MBA. Preadjusted edgewise brackets. Intrusive mechanics: Mesial molar movement, Nance and Goshgarian palatal arches without Cl II elastics, low-pull HG, or anterior biteplates.	−1.13	1.56
		11	28 (14/14)	NE + MBA. Preadjusted edgewise brackets. Extrusive mechanics: low-pull HG, anterior biteplates, and Cl II elastics or posterior crossbite elastics when necessary.	−1.96	1.6
Mann et al., 2011	R	10.57	10 (–)	NE. HG + MBA (protrusive maxilla group).	−2.1	–
		10.62	10 (–)	NE. HG + MBA (normal maxilla group).	−1.2	–
		10.87	10 (–)	NE. HG + MBA (retrusive maxilla group).	−1.18	–
				High-pull HGs, cervical-pull HGs, and combination high-pull plus cervical-pull HGs were used in 54%, 19%, and 27% of the patients, respectively, depending on their vertical type.		
de Almeida-Pedrin et al., 2009***	R	13.8	22 (7/15)	NE. Pendulum + MBA. Anchorage reinforcement was provided by HG at night and Cl II elastics during anterior retraction using heavier rectangular SS AW.	0	1.6
		13.3	30 (15/15)	NE. Cervical HG (outer bows of facebow tilted 15° to 20° upward from occlusal plane, 450 g of force on each side, 16 h/d) + MBA. Cl II elastics were used for canine and incisor retraction.	−1.8	1.2

Table 1. Extended

Total Treatment Duration or Follow Up (m)	Class II Diagnosis	Timing of Treatment
26	Cl II molar relationship	–
–	Skeletal Cl II malocclusion	–
27.6	Cl II dentoskeletal relationship (a full cusp or end-to-end Cl II molar relationship) in the 3 groups	At T1. Patients in the circumpubertal phase of skeletal development (18% prepubertal, 64% pubertal, and18% postpubertal for the patients treated with TB; 15% prepubertal, 70% pubertal, and 15% postpubertal for those treated with FDR; and 18% prepubertal, 64% pubertal, and 18% postpubertal for the control group). At T2. Patients in the postpubertal stage of skeletal development. CVM method.
27.6		
27.6		
19.2	Cl II Division 1 malocclusion, in the 3 groups.	–
19.2		–
17.8		–
39.6	Cl II molar relationship (in at least an end-to-end), in the 2 treated groups.	At T1. Peak growth spurt, CVM method.
27.6		
15.6	Skeletal Cl II	
42	Cl II molar relationship	At T1. CVM stage 2.7; T2. Immediately after MARA removal and prior to placement of full fixed edgewise appliances (CVM stage 4.2); and T3, at least 2 y after MARA removal and after completion of edgewise treatment (CVM stage 5.4).
–	Skeletal Cl II	
24	Cl II Division 1 malocclusion (of a half or full cusp Cl II molar relationship)	At T1. 77.4% of the treated subjects were at CVM stage 3 or 4, 12.9% were at CVM stage 2, and 9.7% were at CVM stage. 5. 90% of the control subjects were at CVM stage 3 or 4, 6.7% were at CVM stage 2, and 3.3% were at CVM stage 5. At T2 and T3, all subjects were at CVM stages 4–6
28.8	Dentoskeletal Cl II	
27.6	Skeletal Cl II Division 1 malocclusion (Cl II molar relationship)	At T1. Peak velocity in craniofacial growth (CVM stage 3 or CVM stage 4)
25.2		
38.4	Cl II Division 1 malocclusion (more than a half-cusp Cl II molar relationship. Unequal distribution of dental Cl II severity between groups). All with hyperdivergent facial type.	At T1. Skeletal maturation stage CVM 1 to CVM4
28.8		
36.6	Cl II molar relationships (end-on to full step), in the 3 groups.	–
32.76		–
39.84		–
45.6	Cl II molar relationship (at least a half cusp) in the 3 groups	–
38.4		–

Table 1. Continued

Study	D[a]	Participants: Age	Participants: N (M/F)	Intervention	Outcome: ANB Change	Outcome: SD
		13.6	30 (15/15)	2 maxillary-first-premolar extractions + MBA. Cl II elastics and cervical HG at night used during anterior retraction.	−1.8	2.3
Baccetti et al., 2009	R	11.3	23 (14/9)	NE. HG + MBA + Cl II elastics (prepeak group).	−1.2	1.4
		11.8	24 (11/13)	NE. HG + MBA + Cl II elastics (peak group).	−0.8	2
		13.7	13 (6/7)	NE. HG + MBA + Cl II elastics (postpeak group).	0	1.4
				All treated patients underwent the same protocol: 0.018-inch-slot brackets. Cervical-pull HG, 14 h/d for 12 mo. In vertical patterns, HG pull was more vertical. HG use was followed by Cl II elastics (5/16-inch, 4–6 oz), full-time wear for 6 to 10 mo.		
		10.2	17 (11/6)	Untreated Cl II control group (prepeak).	−0.2	0.8
		12.1	17 (11/6)	Untreated Cl II control group (peak).	−0.2	0.7
		14.3	13 (7/6)	Untreated Cl II control group (postpeak).	−0.3	0.6
Freitas et al., 2008	R	10.4	25 (5/20)	NE. Cervical HG (expanded inner bow and long outer bow bent upward 15° from horizontal in relation to inner bow, 12 h/d, 450 g of force per side, used for 18 mo) + MBA. 0.022-inch slot, straight-wire brackets. Biteplate used in deep bite patients for 3–5 mo. MBA placed in mandibular arch concomitant with HG; then maxillary teeth were included.	−2.36	1.75
		9.9	16 (4/12)	Untreated Cl II control group.	0.75	2.01
Marques et al., 2008	R	11.6	30 (–)	4 first-premolar extractions + MBA.	−2.97	1.63
		11.6	40 (–)	NE + MBA.	−2.52	1.74
				Patients in both groups were treated with simultaneous use of edgewise brackets and cervical HG (to promote molar retention, 14 h/d). Cl II elastics used in finishing procedures.		
de Oliveira Jr et al., 2007	R	11.86	25 (13/12)	NE. Jasper Jumper (6 mo) + MBA. 0.022-inch slot, standard edgewise brackets with 0.021 × 0.025-inch SS AW during Jasper Jumper therapy. Cl II elastics (5/16-inch) used for retention (4 mo)	−2.05	1.44
		12.29	25 (13/12)	NE. Cervical HG (to correct molar relationship, with outer bows tilted 15° to 20° upward from occlusal plane, for 8–12 mo, with 150–300 g of force per side, used 14–16 h/d) + MBA. HG and MBA used simultaneously. Sequential retraction of premolars and anterior teeth was performed with elastics, HG worn only at night, and Cl II elastics (5/16-inch).	−2.04	1.8
		11.82	25 (13/12)	Untreated Cl II control group.	−0.27	1.85
Janson et al., 2007	R	12.5	22 (10/12)	NE. HG + MBA. HG was used to distalize maxillary posterior teeth and to reinforce anchorage of posterior segment during anterior retraction with 0.018 × 0.025-inch SS AW.	−1.88	1.37
		12.86	22 (10/12)	2 maxillary-first-premolar extractions + MBA. Anterior-teeth retraction performed with 0.018 × 0.025-inch SS AW concurrently with a transpalatal bar and HG to reinforce anchorage of posterior segment.	−1.99	1.75
Janson et al., 2006*	R	14.04	19 (9/10)	2 maxillary-first-premolar extractions + MBA.	−2.06	–
		13.03	47 (20/27)	4-first-premolar extractions + MBA.	−2.53	–
				All patients treated with standard edgewise brackets. Mandibular anterior interproximal stripping was performed when necessary.		
LaHaye et al., 2006***	R	12.1	25 (12/13)	4 first-premolar extractions + HG + MBA. Typical Tweed edgewise mechanics with extensive use of tipback bends, anchorage preparation, and Cl II elastics. Various types of HG (high-pull, J-hook, combined-pull, high-pull bow, Hickham) were used.	−3.9	1.71
		12.7	23 (11/12)	NE. HG + MBA. Treated with Alexander straightwire appliance with cervical-pull HG use a minimum of 14 h/d.	−2.5	1.26
		12.7	19 (9/10)	NE. Herbst (12.7 mo) + MBA. Edgewise appliance.	−1.8	1.3
		12.4	29 (14/15)	Untreated Cl II control group.	0	2.42

Table 1. Continued, extended

Total Treatment Duration or Follow Up (m)	Class II Diagnosis	Timing of Treatment
26.4		–
28.8	Cl II Division 1 malocclusion (bilateral full cusp Cl II molar relationship) in the 3 treated groups	At T1. Before the pubertal growth spurt. CVM 1
30		At T1. During the pubertal growth spurt.CVM 3
30		At T1. postpubertal stage of development. CVM 5
24	Cl II Division 1 malocclusion (full-cusp or half-cusp Cl II molar relationship) in the control groups	At T1. Before the pubertal growth spurt. CVM 1
26.4		At T1. During the pubertal growth spurt. CVM 3
31.2		At T1. Postpubertal stage of development. CVM 5
30	Cl II Division 1 malocclusion (at least half-cusp Cl II molar relationship) in both groups	–
36.4		–
39.6	Cl II Division 1 malocclusion	At T1. Pubertal growth spurt confirmed with radiographs of the hand and wrist
39.6		
23.52	Cl II Division 1 malocclusion (full cusp and at least a half-cusp Cl II molar relationship) in the 3 groups.	–
22.56		–
23.4		–
31.44	Cl II malocclusion (complete bilateral Cl II molar relationship) in both groups	–
29.52		–
28.92	Cl II Division 1 malocclusion (at least half-cusp Cl II molar relationship) in both groups	–
28.92		–
34.2	Skeletal Cl II Division 1 malocclusion (at least half-step Cl II molar and canine relationship) in the treated groups	–
25.2		–
31.3	Cl II Division 1 malocclusion	–
26.4		–

Table 1. Continued

		Participants			Outcome	
Study	D[a]	Age	N (M/F)	Intervention	ANB Change	SD
Haralabakis and Sifakakis, 2004	R	10.41	31 (14/17)	NE. Cervical HG + MBA. (High-angle group).	−2.4	–
		11.73	29 (16/13)	NE. Cervical HG + MBA (Low-angle group).	−1.7	–
				All patients were treated with preadjusted brackets, cervical HG, Cl II elastics. Anterior biteplanes and posterior crossbite elastics were used when necessary.		
Janson et al., 2004	R	11.2	23 (9/14)	NE. High-pull HG-activator combination (10 mo) + MBA. HG-activator combination; HG associated with Cl II elastics and only HG were used for active retention (14.5 mo).	−2.57	1.21
		10.9	15 (8/7)	Untreated Cl II control group.	0	1.03
Ong and Woods, 2001*	R	13.6	15 (–)	4 first-premolar extractions + MBA.	−0.6	–
		13.8	30 (–)	2 maxillary-first-premolar and mandibular-second-premolar extractions + MBA.	−1	–
				All patients were treated with 0.018-inch slot preadjusted brackets. Interarch elastics were used when necessary. Adjunctive appliances (HGs, functional appliances, transpalatal arches) were not used.		
Schiavon Gandini et al., 2001***	R	11	45 (19/26)	NE. Cervical HG (external bow with 20° upward angulation, 400 g of force per side, 14–18 h/d until Class I relationship achieved, and to apply the same force 8–10 h/day thereafter) + MBA. Edgewise brackets.	−2.9	1.85
		10.2	30 (12/18)	Untreated Cl II control group.	−0.09	0.78
Zierhut et al., 2000	R	12.6	23 (11/12)	4 first-premolar extractions + cervical HG + MBA.	−2.12	1.6
		11.3	4 (19/21)	NE. Cervical HG + MBA.	−1.94	1.44
				All subjects treated with edgewise mechanotherapy. Cl II correction was achieved primarily using extraoral force (cervical HG) to redirect or inhibit maxillary anterior development with concurrent mandibular growth and maxillary incisor retraction.		
Bishara, 1998****	R	11.5	44 (21/23)	4 first-premolar extractions + MBA.	−1.83	1.55
		11.5	47 (20/27)	NE. HG + MBA.	−1.34	1.60
				All patients treated with fixed edgewise appliances, extraoral force, and Cl II elastics.		
Foley et al., 1997	R	12	36 (19/17)	NE. HG (cervical or combination type, 400–600 g of force per side, used 12–14 h/d) + MBA. Cl II elastics were used by almost every patient.	−2.2	–
		12	15 (11/4)	Untreated Cl II control group.	−0.4	–
Paquette et al., 1992	R	12.5	33 (13/20)	4 first-premolar extractions + MBA.	−1.8	–
		12.6	30 (19/11)	NE + MBA.	−1.3	–
				0.022-inch slot edgewise brackets. Edgewise mechanics. Backward-pulling HG and Cl II elastics used in both groups.		

[a] D indicates design; NE, nonextraction treatment; R, retrospective; P, prospective; SS, stainless steel; AW, archwire; mo, months; CVM, cervical vertebral maturation.

* ANB angle mean changes were calculated using pretreatment and posttreatment data. Imputing standard deviations for changes from baseline was not considered.

** When ANB angle mean changes were reported without their standard deviations, they were estimated using the *P* values for differences in means (when reported) as mentioned in the Cochrane handbook (Chapter 7, Section 7.7.3.3).[23]

*** ANB angle and SD mean changes were calculated using the complete treatment time mean of each group.

**** ANB angle and SD mean changes were calculated combining groups, using male and female data (reported separately in the article) as recommended in the Cochrane handbook (Chapter 7, Section 7.7.3.8).[23]

Table 1. Continued, extended

Total Treatment Duration or Follow Up (m)	Class II Diagnosis	Timing of Treatment	
33	Cl II Division 1 malocclusion (at least a half-step bilateral Cl II molar relationship)	–	
33.48		–	
29.88	Cl II Division 1 malocclusion	–	
32.16		–	
27.3	Cl II molar relationship	–	
26.3		–	
43.2	Cl II Division 1 malocclusion (Cl II molar and canine relationship) in both groups	–	
15.6		–	
34.8	Cl II Division 1 malocclusion (of at least end-on Cl II molar relationship)	–	
30			–
37.2	Cl II Division 1 malocclusion (Cl II molar relationship)	–	
27.6		–	
48	Cl II Division 1 malocclusion (Cl II molar relationship)	–	
–		–	
22.08	Cl II Division 1 malocclusion	–	
19.2		–	

apical base relationship changes during Class II malocclusion treatment with the different protocols.

The systematic review showed a lack of Low RoB studies. However, a meta-analysis was performed only in the nonextraction group including nonrandomized studies with Moderate RoB with matched UCIICGs (Figure 2, Table 4). This usually implies the use of historical Class II controls due to ethical issues. Therefore, we have included retrospective controlled trials, as performed in other systematic reviews.[38] Regarding the extraction protocols, the systematic review showed a lack of studies with UCIICGs data reporting and Low RoB. Only one extraction study[5] reported data of UCIICG (Tables 1 and 2). Nevertheless, in the absence of stronger evidence, they can provide information to orient clinicians (Table 3).

The meta-analysis results showed statistically significant improvement of the apical base sagittal relationship (ANB angle) in the XP0 protocol (Figure 2, Table 4). Treatment performed before or at the peak of growth showed greater ANB changes when compared with postpeak patients, as reported previously.[26,38] Differences in ANB changes between the use of FA + MBA or HG + MBA were minimal (Table 4), as expected.[39] However, a greater ANB change with the HG-activator combination, used in only one study, was reported.[35] Nevertheless, it is necessary to have more studies to support these findings.

Table 2. Risk of Bias (RoB) of Studies Included in the Qualitative Synthesis Based on ROBINS-I Tool

	Domains							
	Preintervention		At Intervention	Postintervention				
Author	Bias due to Confounding	Bias in Selecting Participants for the Study	Bias in Classifying Interventions	Bias due to Deviations From Intended Intervention	Bias due to Missing Data	Bias in Measuring Outcomes	Bias in Selecting Reported Result	Overall RoB Judgment
Dada et al., 2015	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Giuntini et al., 2015	Moderate	Moderate	Low	Low	Low	Low	Moderate	Moderate
Zheng et al., 2015	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
Al-Jewair et al., 2012	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Pangrazio et al., 2012	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Phelan et al., 2012	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Ye et al., 2012	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Gkantidis et al., 2011	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Mann et al., 2011	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
de Almeida-Pedrin et al., 2009	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Baccetti et al., 2009	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Freitas et al., 2008	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Marques et al., 2008	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
de Oliveira Jr et al., 2007	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Janson et al., 2007	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Janson et al., 2006	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
LaHaye et al., 2006	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Haralabakis and Sifakakis, 2004	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
Janson et al., 2004	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Ong and Woods, 2001	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
Schiavon Gandini et al., 2001	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Zierhut et al., 2000	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Bishara, 1998	Moderate	Moderate	Low	Low	Low	Moderate	Moderate	Moderate
Foley et al., 1997	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious
Paquette et al., 1992	Moderate	Moderate	Low	Low	Low	Moderate	Serious	Serious

Some reviews used annualized data to account for the different follow-up periods of the studies,[40] or because most of their samples reported annualized data.[38,39] In this systematic review, data were not annualized because the objective was to evaluate ANB changes in the complete period of treatment.

ANB reduction was greater in the XP4 protocol (2.55°) compared with the XP2 (1.88°) and with XP0 protocol (1.56°). Evidently, the results of the XP0 protocols have a greater reliability because they consisted of results of a meta-analysis, but they represent Low quality of evidence (GRADE), based on the nature of the included studies and must be regarded with caution. Although the results of the XP4 and XP2 protocols may provide some guidance to the clinician, they lack some consistency because they were derived from estimates.

Clinical Implications

The current results confirm previous speculations that Class II malocclusion anteroposterior apical base skeletal changes are small, especially regarding ANB.[2,4,10,15,16,20] Therefore, Class II malocclusion severity should be expressed primarily as the occlusal anteroposterior discrepancy and not as the skeletal discrepancy.[8,11,12] This would provide treatment strategies with more predictable results. A complete Class II anteroposterior discrepancy is likely to be corrected to a normal occlusion in most cases.[5,10,14] However, a severe cephalometric skeletal discrepancy, such as an ANB of 10° will on average be reduced only by about 2°,,resulting in an ANB of 8°, which is far from the ideal. This does not mean that the skeletal apical base discrepancy is not important in orthodontic diagnosis. It is, but its purpose is to provide additional information for diagnosis and not to define treatment planning.

Limitations

The great limitation found was the lack of Low RoB studies. No randomized, controlled clinical trial that satisfied our inclusion criteria was found. Meta-analysis was performed only with Moderate RoB studies for the

Table 3. Studies Included in the Quantitative Analyses

	Treatment	Mean	SD	n
Nonextraction				
Al-Jewair et al., 2012	NE. MARA + MBA	−2.4	1.6	40
	NE. AndvanSync + MBA	−2.6	1.9	30
Baccetti et al., 2009	NE. HG + MBA + Class II (Cl II) elastics (prepeak group)	−1.2	1.4	23
	NE. HG + MBA + Cl II elastics (peak group)	−0.8	2	24
	NE. HG + MBA + Cl II elastics (postpeak group)	0	1.4	13
Dada et al., 2015	NE. Forsus + MBA + Cl II Elastics	−1.88	1.07	19
de Oliveira et al., 2007	NE. Jasper Jumper + MBA + Cl II Elastics	−2.05	1.44	25
	NE.HG + MBA + Cl II elastics	−2.04	1.8	25
Freitas et al., 2008	NE. HG + MBA	−2.36	1.75	25
Giuntini et al., 2015	NE. Twin-block + MBA	−2.6	1.3	28
	NE. Forsus + MBA	−1.8	1.3	36
Janson et al., 2004	NE. HG-activator combination + MBA + Cl II elastics	−2.57	1.21	23
La Haye et al., 2006	NE. HG + MBA	−2.5	1.26	23
	NE. Herbst + MBA	−1.8	1.3	19
Pangrazio et al., 2012	NE. MARA + MBA	−0.63	0.81	30
Phelan et al., 2012	NE. Sydney Magnoglide + MBA + Cl II elastics	−1	1	31
Schiavon Gandini et al. 2001	NE. HG + MBA	−2.9	1.85	45
4 First-premolar Extractions				
Bishara et al., 1998	4 first-premolar extractions + HG + MBA + Cl II elastics	−1.83	1.55	44
Gkantidis et al., 2011	4 first-premolar extractions + MBA	−1.13	1.56	29
La Haye et al., 2006	4 first-premolar extractions + HG + MBA + Cl II elastics	−3.9	1.71	25
Marques et al., 2008	4 first-premolar extractions + HG + MBA + Cl II elastics	−2.97	1.63	30
Ye et al., 2012	4 first-premolar extractions + HG + MBA + Cl II elastics	−3.04	2.13	76
Zierhut et al., 2000	4 first-premolar extractions + HG + MBA	−2.12	1.6	23
	Mean Change of ANB Angle	−2.55	1.96	227
2 Maxillary-First-Premolar Extractions				
de Almeida-Pedrin et al., 2009	2 maxillary-first-premolar extractions + HG + MBA + Cl II elastics	−1.8	2.3	30
Janson et al., 2007	2 maxillary-first-premolar extractions + HG + MBA	−1.99	1.75	22
	Mean Change of ANB Angle	−1.88	2.06	52

XP0 protocol because of the lack of UCIICG data reporting in the extraction studies.

The use of nonannualized data for ANB changes was important to describe the ANB changes in the total treatment period. It may have provided greater ANB values in treatments with longer duration. However, the subgroup analysis showed no significant difference in ANB change considering this factor.

CONCLUSIONS

Overall, based on the low-quality evidence found:

• Class II malocclusion XP0 treatment produces an

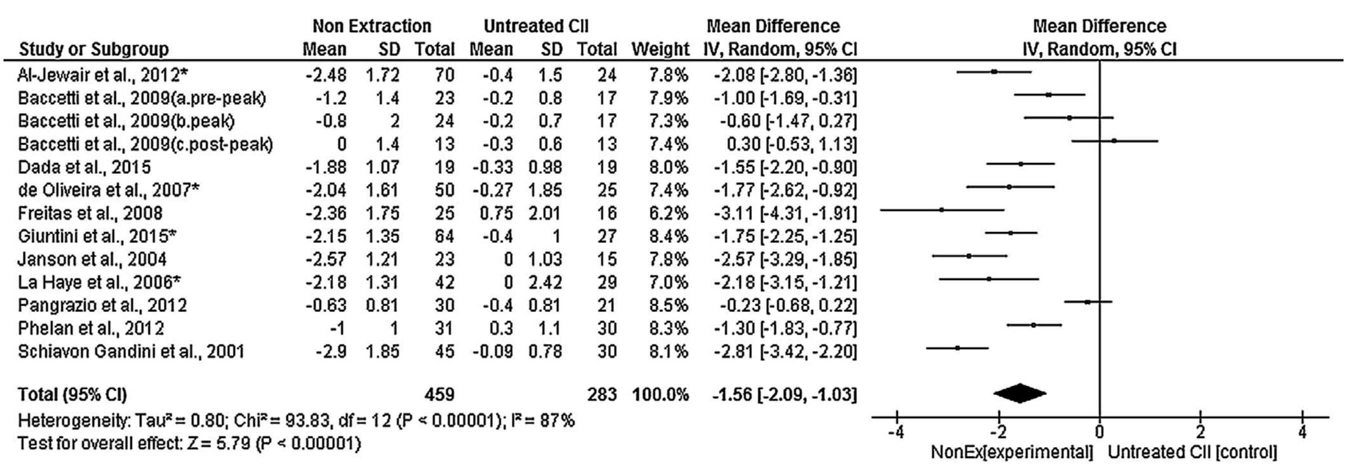


Figure 2. Forest plot (mean difference [MD] and 95% confidence interval [CI]) for the ANB angle mean changes between Class II nonextraction treatment and UCIICGs. *Multi-arm studies pooled before meta-analysis.

Table 4. Meta-analysis, Subgroup Analyses, and Sensitivity Analyses: Nonextraction vs Untreated Cl II

Meta-analysis

Variable	Studies (n)	Effect Size: MD[a] (95% CI)	Effect Size: *P* Value**
ANB angle	11 (13*)	−1.56 (−2.09, −1.03)	<.001

	Heterogeneity: Tau²	Chi²	*P* value***	I² (95% CI)	95% PI
	0.80	93.83	<.001	87% (80.92)	(−3.62, 0.50)

Subgroup Analyses

Factors	Studies (n)	MD (95% CI)	P_{SG}***
Skeletal growth stage			.05
Prepeak	1	−1.00 (−1.69, −0.31)	
Peak	3	−0.96 (−2.12, 0.21)	
Postpeak	1	0.30 (−0.53, 1.13)	
Type of appliance used			.01
Functional appliance + MBA	7	−1.03 (−1.79, −0.27)	
HG + MBA	5 (7*)	−1.62 (−2.57, −0.67)	
HG-activator combination + MBA	1	−2.57 (−3.29, −1.85)	
Treatment time			.72
Up to 24 mo	2	−1.43 (−1.88, −0,98)	
Greater than 24 mo	9 (11*)	−1.57 (−2.21, −0.94)	

Sensitivity Analyses

	Studies (n)	MD (95% CI)	*P* value**	P_{SG}***
By study design				.47
Prospective	1	−1.30 (−1.83, −0.77)	<.001	
Retrospective	10 (12*)	−1.59 (−2.18, −1.00)	<.001	
By precision				.36
All studies	11 (13*)	−1.56 (−2.09, −1.03)	<.001	
Studies with narrower 95% CI	5 (6*)	−1.08 (−1.95, −0.22)	<.001	

[a] MD indicates mean difference; CI, confidence interval; PI, prediction interval; MBA, multibracket appliances; $P_{SG,}$ *P* value for differences between subgroups.

* In this study,[14] groups were pooled as individual studies (they presented a different control group for each treated group).

** Statistical significance at $P < .05$.

*** Statistical significance at $P < .10$.

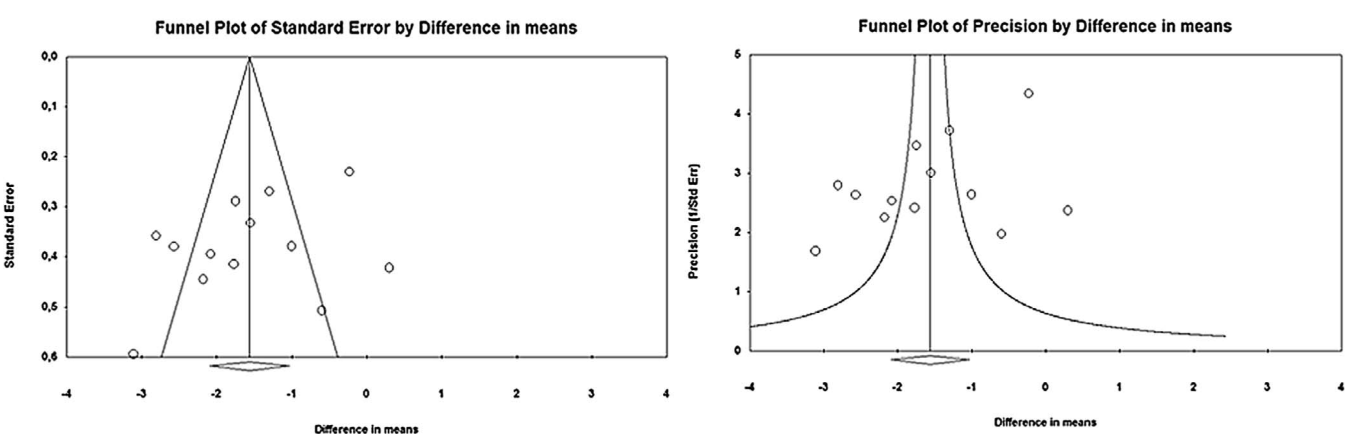


Figure 3. Funnel plots of ANB changes for the studies (11 studies, 13 groups) included in the meta-analysis.

Table 5. GRADE Summary of Findings Table for the ANB Mean Change Outcome of the Systematic Review and Meta-analysis Directly After Treatment Completion Without Premolar Extractions[a]

	Illustrative Comparative Risks (95% CI)				
	Assumed Risk	Corresponding Risk			
Outcomes	Untreated Cl II (Controls)	Nonextraction	No. of Participants (Studies)	Quality of Evidence (GRADE)	Comments
ANB mean changes (from beginning until treatment completion. Follow-up, 23.5–43.2 mo)*	ANB angle change ranged across control groups from $-0.4°$ to $+0.75°$	Mean ANB angle decreased in the nonextraction groups by 1.56° (95% CI: 1.03°–2.09° decrease) compared with control groups	742 (11)	⊕⊕⊖⊖ (Low, due to inclusion of nonrandomized studies)	Greater ANB angle mean changes (represented by smaller angle) indicates improvement in sagittal relationship of apical bases.

[a] Patient or population: growing patients with Cl II malocclusion; Settings: University clinics and private practice; Intervention: Cl II treatment without premolar extractions after multibracket-appliance treatment (MBA); Comparison: Untreated Class II subjects from follow-up or historical controls.

* From cephalometric analyses.

average reduction of 1.56° in ANB compared with untreated Class II subjects.
- Class II malocclusion treated with XP2 produces an estimated mean reduction of 1.88° in ANB.
- Class II malocclusion treated with XP4 produces an estimated mean reduction of 2.55° in ANB.
- However, further research is necessary to obtain the most robust results.

REFERENCES

1. Steiner CC. Cephalometrics for you and me. *Am J Orthod*. 1953;39:729–755.
2. Paquette DE, Beattie JR, Johnston LE Jr. A long-term comparison of nonextraction and premolar extraction edgewise therapy in "borderline" Class II patients. *Am J Orthod Dentofacial Orthop*. 1992;102:1–14.
3. Foley TF, Stirling DL, Hall-Scott J. The reliability of three sagittal reference planes in the assessment of Class II treatment. *Am J Orthod Dentofacial Orthop*. 1997;112:320–326; discussion 327–329.
4. Bishara SE. Mandibular changes in persons with untreated and treated Class II Division 1 malocclusion. *Am J Orthod Dentofacial Orthop*. 1998;113:661–673.
5. LaHaye MB, Buschang PH, Alexander RG, Boley JC. Orthodontic treatment changes of chin position in Class II Division 1 patients. *Am J Orthod Dentofacial Orthop*. 2006;130:732–741.
6. Mann KR, Marshall SD, Qian F, Southard KA, Southard TE. Effect of maxillary anteroposterior position on profile esthetics in headgear-treated patients. *Am J Orthod Dentofacial Orthop*. 2011;139:228–234.
7. Abdullah RT, Kuijpers MA, Berge SJ, Katsaros C. Steiner cephalometric analysis: predicted and actual treatment outcome compared. *Orthod Craniofac Res*. 2006;9:77–83.
8. Janson G, Sathler R, Fernandes TM, Zanda M, Pinzan A. Class II malocclusion occlusal severity description. *J Appl Oral Sci*. 2010;18:397–402.
9. Nangia A, Darendeliler MA. Finishing occlusion in Class II or Class III molar relation: therapeutic Class II and III. *Aust Orthod J*. 2001;17:89–94.
10. Janson G, Fuziy A, de Freitas MR, Castanha Henriques JF, de Almeida RR. Soft-tissue treatment changes in Class II Division 1 malocclusion with and without extraction of maxillary premolars. *Am J Orthod Dentofacial Orthop*. 2007;132:729.e1–729.e8.
11. Janson G, Barros SE, de Freitas MR, Henriques JF, Pinzan A. Class II treatment efficiency in maxillary premolar extraction and nonextraction protocols. *Am J Orthod Dentofacial Orthop*. 2007;132:490–498.
12. Durao AR, Alqerban A, Ferreira AP, Jacobs R. Influence of lateral cephalometric radiography in orthodontic diagnosis and treatment planning. *Angle Orthod*. 2015;85:206–210.
13. Ong HB, Woods MG. An occlusal and cephalometric analysis of maxillary first and second premolar extraction effects. *Angle Orthod*. 2001;71:90–102.
14. Baccetti T, Franchi L, Kim LH. Effect of timing on the outcomes of 1-phase nonextraction therapy of Class II malocclusion. *Am J Orthod Dentofacial Orthop*. 2009;136:501–509.
15. de Almeida-Pedrin RR, Henriques JFC, de Almeida RR, de Almeida MR, McNamara JA Jr. Effects of the pendulum appliance, cervical headgear, and 2 premolar extractions followed by fixed appliances in patients with Class II malocclusion. *Am J Orthod Dentofacial Orthop*. 2009;136:833–842.
16. Gkantidis N, Halazonetis DJ, Alexandropoulos E, Haralabakis NB. Treatment strategies for patients with hyperdivergent Class II Division 1 malocclusion: is vertical dimension affected? *Am J Orthod Dentofacial Orthop*. 2011;140:346–355.
17. Pangrazio MNK, Pangrazio-Kulbersh V, Berger JL, Bayirli B, Movahhedian A. Treatment effects of the mandibular anterior repositioning appliance in patients with Class II skeletal malocclusions. *Angle Orthod*. 2012;82:971–977.
18. Phelan A, Tarraf NE, Taylor P, Hönscheid R, Drescher D, Baccetti T, et al. Skeletal and dental outcomes of a new magnetic functional appliance, the Sydney Magnoglide, in Class II correction. *Am J Orthod Dentofacial Orthop*. 2012;141:759–772.
19. Zheng B, Jiang Z, Liu F, An N, Zheng Y, Zhang Y, et al. Effect of headgear and Class II traction on upper airway dimensions and hyoid bone position in non-extraction patients with Class II Division 1 malocclusion. *J Hard Tissue Biology*. 2015;24:1–6.
20. Janson G, Busato MC, Henriques JF, de Freitas MR, de Freitas LM. Alignment stability in Class II malocclusion treated with 2- and 4-premolar extraction protocols. *Am J Orthod Dentofacial Orthop*. 2006;130:189–195.

21. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol*. 2009;62:e1–e34.
22. Sterne JAC, Higgins JPT, Reeves BC. On behalf of the development group for ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions, Version 7. March 2016. Available from http://www.riskofbias.info. Accessed June 30, 2016.
23. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.1.0 [updated March 2011]: The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org. Accessed June 30, 2016.
24. Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the *Journal of Clinical Epidemiology*. *J Clin Epidemiol*. 2011;64:380–382.
25. Papageorgiou SN. Meta-analysis for orthodontists: part I—how to choose effect measure and statistical model. *J Orthod*. 2014;41:317–326.
26. Zymperdikas VF, Koretsi V, Papageorgiou SN, Papadopoulos MA. Treatment effects of fixed functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis. *Eur J Orthod*. 2016;38:113–126.
27. Dada DM, Galang-Boquiren MT, Viana G, Obrez A, Kusnoto B. Treatment effects of Forsus fatigue resistant device on Class II malocclusion cases: A cephalometric evaluation. *J World Fed Orthod*. 2015;4:14–17.
28. Giuntini V, Vangelisti A, Masucci C, Defraia E, McNamara JA Jr, Franchi L. Treatment effects produced by the Twin-block appliance vs the Forsus Fatigue Resistant Device in growing Class II patients. *Angle Orthod*. 2015;85:784–789.
29. Al-Jewair TS, Preston CB, Moll EM, Dischinger T. A comparison of the MARA and the AdvanSync functional appliances in the treatment of Class II malocclusion. *Angle Orthod*. 2012;82:907–914.
30. Ye C, Du X, Zhao Q, Tan Z. Dentoskeletal effects of 2-phase treatment versus 1-phase treatment with extraction of 4 first premolars in nonhypodivergent class II patients. *J Craniofac Surg*. 2012;23:1544–1547.
31. Freitas MR, Lima DV, Freitas KM, Janson G, Henriques JF. Cephalometric evaluation of Class II malocclusion treatment with cervical headgear and mandibular fixed appliances. *Eur J Orthod*. 2008;30:477–482.
32. Marques LS, Ramos-Jorge ML, Araujo MT, Bolognese AM. Class II Division 1 malocclusion with severe overbite: cephalometric evaluation of the effects of orthodontic treatment. *World J Orthod*. 2008;9:319–328.
33. de Oliveira Jr JN, Rodrigues de Almeida R, Rodrigues de Almeida M, de Oliveira JN. Dentoskeletal changes induced by the Jasper Jumper and cervical headgear appliances followed by fixed orthodontic treatment. *Am J Orthod Dentofacial Orthop*. 2007;132:54–62.
34. Haralabakis NB, Sifakakis IB. The effect of cervical headgear on patients with high or low mandibular plane angles and the “myth” of posterior mandibular rotation. *Am J Orthod Dentofacial Orthop*. 2004;126:310–317.
35. Janson G, Caffer Dde C, Henriques JF, de Freitas MR, Neves LS. Stability of Class II, Division 1 treatment with the headgear-activator combination followed by the edgewise appliance. *Angle Orthod*. 2004;74:594–604.
36. Schiavon Gandini MREA, Gandini LG Jr, Da Rosa Martins JC, Del Santo M Jr. Effects of cervical headgear and edgewise appliances on growing patients. *Am J Orthod Dentofacial Orthop*. 2001;119:531–538.
37. Zierhut EC, Joondeph DR, Artun J, Little RM. Long-term profile changes associated with successfully treated extraction and nonextraction Class II Division 1 malocclusions. *Angle Orthod*. 2000;70:208–219.
38. Perinetti G, Primozic J, Furlani G, Franchi L, Contardo L. Treatment effects of fixed functional appliances alone or in combination with multibracket appliances: a systematic review and meta-analysis. *Angle Orthod*. 2015;85:480–492.
39. Thiruvenkatachari B, Harrison J, Worthington H, O’Brien K. Early orthodontic treatment for Class II malocclusion reduces the chance of incisal trauma: results of a Cochrane systematic review. *Am J Orthod Dentofacial Orthop*. 2015;148:47–59.
40. Koretsi V, Zymperdikas VF, Papageorgiou SN, Papadopoulos MA. Treatment effects of removable functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis. *Eur J Orthod*. 2015;37:418–434.

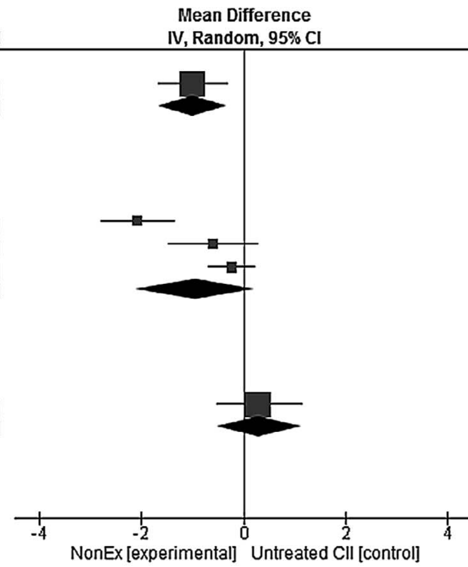
APPENDIX 1 Database and Search Strategy

Database	Key Words
Pubmed	(malocclusion, angle Class II OR (Class AND II AND malocclusion)) AND (orthodontic treatment OR orthopedic treatment OR orthopaedic treatment) AND (ANB OR a-n-b OR (SNA AND SNB)) AND (cephalo*)
Embase	#1 AND #2 AND #3 AND #4 AND [humans]/lim.#1: malocclusion AND angle AND Class AND II OR (Class AND II AND malocclusion). #2: orthodontic AND treatment OR (orthopedic AND treatment) OR (orthopaedic AND treatment). #3: anb OR (a AND n AND b) OR (sna AND snb). #4: cephalo*
Web of Science	TS = (malocclusion, angle Class II OR (Class AND II AND malocclusion)) AND TS = (orthodontic treatment OR orthopedic treatment OR orthopaedic treatment) AND TS = (ANB OR a-n-b OR (SNA AND SNB)) AND TS = (cephalo*)
Scopus	(malocclusion, angle Class II OR (Class AND II AND malocclusion)) AND (orthodontic treatment OR orthopedic treatment OR orthopaedic treatment) AND (ANB OR a-n-b OR (SNA AND SNB)) AND (cephalo*)
The Cochrane Library	(malocclusion, angle Class II OR (Class AND II AND malocclusion)) AND (orthodontic treatment OR orthopedic treatment OR orthopaedic treatment) AND (ANB OR a-n-b OR (SNA AND SNB)) AND (cephalo*)
Lilacs	(Má oclusão de angle Classe II) OR (Classe AND II AND má oclusão) [Words] and (ANB) OR (a-n-b) OR (SNA AND SNB) [Words]
Grey Literature (Google Scholar)	Any idiom; Without patents and citations; Classified by relevance; Search; malocclusion angle Class II treatment ANB SNA SNB anywhere in the article; At least one of following words: ANB SNA SNB; With exact phrase: class II treatment; 100 most relevant articles

APPENDIX 2 Criteria for Assessing Risk of Bias (Rob) With the Risk of Bias in Nonrandomized Studies of Interventions (ROBINS-I) Tool

Domains of Bias	Description
Preintervention	
Bias due to confounding	*Baseline confounding*. When one or more preintervention prognostic factors predict the intervention received at baseline (start of follow-up). *Time-varying confounding*. When the intervention received can change over time and when postintervention prognostic factors affect the intervention received after baseline.
Bias in selecting participants for study	When selection of participants is related to both intervention and outcome. *Lead time bias*. When some follow-up time is excluded from the analysis. *Immortal time bias*. When the interventions are defined in such a way that there is a period of follow-up during which the outcome cannot occur.
At Intervention	
Bias in classifying interventions	When intervention status is misclassified. *Nondifferential misclassification*. Is unrelated to the outcome. *Differential misclassification*. Is related to the outcome or to the risk of the outcome.
Postintervention	
Bias due to deviating from intended intervention	When there are systematic differences between intervention and comparator groups in the care provided.
Bias due to missing data	When attrition (loss to follow-up), missed appointments, incomplete data collection, and exclusion of participants from analysis by primary investigators occur.
Bias in measuring outcomes	When outcomes are misclassified or measured with error. *Nondifferential measurement error*. Is unrelated to the intervention received; it can be systematic or random. *Differential measurement error*. Is related to intervention status.
Bias in selecting reported result	Selective reporting of results, that should be sufficiently reported to allow the estimate to be included in a meta-analysis (or other synthesis) is considered. When selective reporting is based on the direction, magnitude, or statistical significance of intervention effect estimates. *Selective outcome reporting*. When the effect estimate for an outcome measurement was selected from among analyses of multiple outcome measurements for the outcome domain. *Selective analysis reporting*. When results are selected from intervention effects estimated in multiple ways.
Judgment for each domain	
Low RoB	Study is comparable to a well-performed, randomized trial with regard to this domain.
Moderate RoB	Study is sound for a nonrandomized study with regard to this domain but cannot be considered comparable to a well-performed, randomized trial.
Serious RoB	Study has some important problems in this domain.
Critical RoB	Study is too problematic in this domain to provide any useful evidence on the effects of intervention.
No information	No information on which to base a judgment about risk of bias for this domain.
Overall judgment for each study	
Low RoB	Study is judged to be at low risk of bias for all domains.
Moderate RoB	Study is judged to be at low or moderate risk of bias for all domains.
Serious RoB	Study is judged to be at serious risk of bias in at least one domain, but not at critical risk of bias in any domain.
Critical RoB	Study is judged to be at critical risk of bias in at least one domain.
No information	No clear indication that the study is at serious or critical risk of bias, and there is a lack of information in one or more key domains of bias (a judgment is required for this).

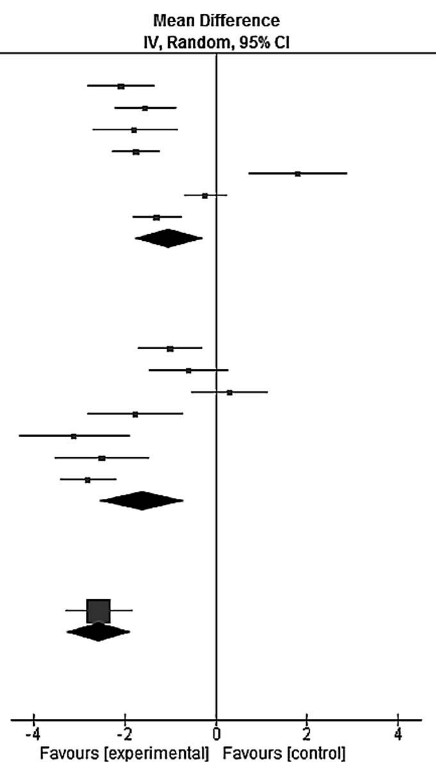
Study or Subgroup	Non Extraction Mean	SD	Total	Untreated CII Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
1.2.1 Pre-peak								
Baccetti et al., 2009(a.pre-peak)	-1.2	1.4	23	-0.2	0.8	17	100.0%	-1.00 [-1.69, -0.31]
Subtotal (95% CI)			**23**			**17**	**100.0%**	**-1.00 [-1.69, -0.31]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.85 (P = 0.004)								
1.2.2 Peak								
Al-Jewair et al., 2012*	-2.48	1.72	70	-0.4	1.5	24	33.0%	-2.08 [-2.80, -1.36]
Baccetti et al., 2009(b.peak)	-0.8	2	24	-0.2	0.7	17	31.2%	-0.60 [-1.47, 0.27]
Pangrazio et al., 2012	-0.63	0.81	30	-0.4	0.81	21	35.8%	-0.23 [-0.68, 0.22]
Subtotal (95% CI)			**124**			**62**	**100.0%**	**-0.96 [-2.12, 0.21]**
Heterogeneity: Tau² = 0.93; Chi² = 18.19, df = 2 (P = 0.0001); I² = 89%								
Test for overall effect: Z = 1.61 (P = 0.11)								
1.2.3 Post-peak								
Baccetti et al., 2009(c.post-peak)	0	1.4	13	-0.3	0.6	13	100.0%	0.30 [-0.53, 1.13]
Subtotal (95% CI)			**13**			**13**	**100.0%**	**0.30 [-0.53, 1.13]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.71 (P = 0.48)								



Test for subgroup differences: Chi² = 6.16, df = 2 (P = 0.05), I² = 67.6%

APPENDIX 3.

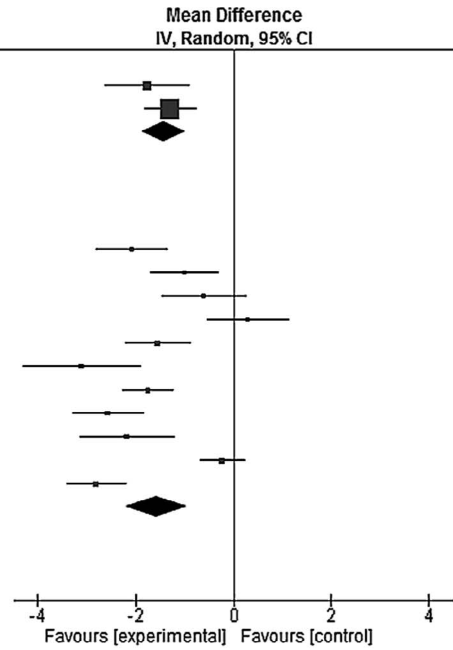
Study or Subgroup	Non Extraction Mean	SD	Total	Untreated CII Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
1.3.1 Functional appliances + MBA								
Al-Jewair et al., 2012*	-2.48	1.72	70	-0.4	1.5	24	14.2%	-2.08 [-2.80, -1.36]
Dada et al., 2015	-1.88	1.07	19	-0.33	0.98	19	14.5%	-1.55 [-2.20, -0.90]
de Oliveira et al., 2007(a.FA+MBA)	-2.05	1.44	25	-0.27	1.85	25	13.2%	-1.78 [-2.70, -0.86]
Giuntini et al., 2015*	-2.15	1.35	64	-0.4	1	27	15.2%	-1.75 [-2.25, -1.25]
La Haye et al., 2006(a.FA+MBA)	1.8	1.3	19	0	2.42	29	12.4%	1.80 [0.74, 2.86]
Pangrazio et al., 2012	-0.63	0.81	30	-0.4	0.81	21	15.4%	-0.23 [-0.68, 0.22]
Phelan et al., 2012	-1	1	31	0.3	1.1	30	15.1%	-1.30 [-1.83, -0.77]
Subtotal (95% CI)			**258**			**175**	**100.0%**	**-1.03 [-1.79, -0.27]**
Heterogeneity: Tau² = 0.93; Chi² = 61.15, df = 6 (P < 0.00001); I² = 90%								
Test for overall effect: Z = 2.64 (P = 0.008)								
1.3.2 Headgear + MBA								
Baccetti et al., 2009(a.pre-peak)	-1.2	1.4	23	-0.2	0.8	17	15.1%	-1.00 [-1.69, -0.31]
Baccetti et al., 2009(b.peak)	-0.8	2	24	-0.2	0.7	17	14.4%	-0.60 [-1.47, 0.27]
Baccetti et al., 2009(c.post-peak)	0	1.4	13	-0.3	0.6	13	14.6%	0.30 [-0.53, 1.13]
de Oliveira et al., 2007(b.HG+MBA)	-2.04	1.88	25	-0.27	1.85	25	13.7%	-1.77 [-2.80, -0.74]
Freitas et al., 2008	-2.36	1.75	25	0.75	2.01	16	13.0%	-3.11 [-4.31, -1.91]
La Haye et al., 2006(b.HG+MBA)	-2.5	1.26	23	0	2.42	29	13.8%	-2.50 [-3.52, -1.48]
Schiavon Gandini et al., 2001	-2.9	1.85	45	-0.09	0.78	30	15.4%	-2.81 [-3.42, -2.20]
Subtotal (95% CI)			**178**			**147**	**100.0%**	**-1.62 [-2.57, -0.67]**
Heterogeneity: Tau² = 1.44; Chi² = 52.64, df = 6 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 3.33 (P = 0.0009)								
1.3.3 Headgear-Activator combination + MBA								
Janson et al., 2004	-2.57	1.21	23	0	1.03	15	100.0%	-2.57 [-3.29, -1.85]
Subtotal (95% CI)			**23**			**15**	**100.0%**	**-2.57 [-3.29, -1.85]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 7.01 (P < 0.00001)								



Test for subgroup differences: Chi² = 8.46, df = 2 (P = 0.01), I² = 76.4%

APPENDIX 4.

	Non Extraction			Untreated CII				Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
1.4.1 Treatment time up to 24 months								
de Oliveira et al., 2007*	-2.04	1.61	50	-0.27	1.85	25	27.8%	-1.77 [-2.62, -0.92]
Phelan et al., 2012	-1	1	31	0.3	1.1	30	72.2%	-1.30 [-1.83, -0.77]
Subtotal (95% CI)			**81**			**55**	**100.0%**	**-1.43 [-1.88, -0.98]**
Heterogeneity: Tau² = 0.00; Chi² = 0.85, df = 1 (P = 0.36); I² = 0%								
Test for overall effect: Z = 6.25 (P < 0.00001)								
1.4.2 Treatment time greater than 24 months								
Al-Jewair et al., 2012*	-2.48	1.72	70	-0.4	1.5	24	9.2%	-2.08 [-2.80, -1.36]
Baccetti et al., 2009(a.pre-peak)	-1.2	1.4	23	-0.2	0.8	17	9.3%	-1.00 [-1.69, -0.31]
Baccetti et al., 2009(b.peak)	-0.8	2	24	-0.2	0.7	17	8.7%	-0.60 [-1.47, 0.27]
Baccetti et al., 2009(c.post-peak)	0	1.4	13	-0.3	0.6	13	8.9%	0.30 [-0.53, 1.13]
Dada et al., 2015	-1.88	1.07	19	-0.33	0.98	19	9.4%	-1.55 [-2.20, -0.90]
Freitas et al., 2008	-2.36	1.75	25	0.75	2.01	16	7.6%	-3.11 [-4.31, -1.91]
Giuntini et al., 2015*	-2.15	1.35	64	-0.4	1	27	9.8%	-1.75 [-2.25, -1.25]
Janson et al., 2004	-2.57	1.21	23	0	1.03	15	9.2%	-2.57 [-3.29, -1.85]
La Haye et al., 2006*	-2.18	1.31	42	0	2.42	29	8.4%	-2.18 [-3.15, -1.21]
Pangrazio et al., 2012	-0.63	0.81	30	-0.4	0.81	21	9.9%	-0.23 [-0.68, 0.22]
Schiavon Gandini et al., 2001	-2.9	1.85	45	-0.09	0.78	30	9.5%	-2.81 [-3.42, -2.20]
Subtotal (95% CI)			**378**			**228**	**100.0%**	**-1.57 [-2.21, -0.94]**
Heterogeneity: Tau² = 1.00; Chi² = 92.98, df = 10 (P < 0.00001); I² = 89%								
Test for overall effect: Z = 4.87 (P < 0.00001)								
Test for subgroup differences: Chi² = 0.13, df = 1 (P = 0.72), I² = 0%								



APPENDIX 5.