Treatment and stability of Class II division 2 malocclusion in children and adolescents: a systematic review

Declan T Millett[1] BDSc, DDS, FDS, DOrth, MOrth
Susan J Cunningham[2] BChD, MSc, PhD, FDS (Orth), MOrth
Kevin D O'Brien[3] BDS, MSc, PhD, FDS, DOrth
Philip E Benson[4] BDS, MSc, PhD, FDS (Orth), MOrth
Cesar M de Oliveira[5] BDS, MSc, PhD

[1]Professor of Orthodontics / Consultant Orthodontist, Oral Health and Development, Dental School, University College Cork, Ireland

[2]Professor / Honorary Consultant in Orthodontics, UCL Eastman Dental Institute, UK

[3]Professor of Orthodontics, Faculty of Medical and Human Sciences, University of Manchester, UK

[4]Reader / Honorary Consultant in Orthodontics, Academic Unit of Oral Health and Development, School of Clinical Dentistry, University of Sheffield, UK

[5]Research Fellow, Department of Epidemiology and Public Health, University College London, UK

Corresponding author:
Prof. D. Millett
Oral Health and Development
Dental School
University College Cork
Ireland
Tel: +353-21- 4901138
Fax: +353-21- 4901192
Email: d.millett@ucc.ie

Acknowledgements

We thank Sylvia Bickley and Anne Littlewood of the Cochrane Collaboration (Oral Health Group, School of Dentistry, University of Manchester, UK) for their invaluable assistance with developing and running the search strategies for the original Cochrane review; Una Ni Chonghaile (University College Cork, Ireland) is also thanked for her assistance with the search strategies.

Abstract

Objectives: To evaluate the evidence with regard to the effectiveness and stability of orthodontic treatment interventions for Class II division 2 malocclusion (II/2M) in children and adolescents.

Design: Systematic review conducted according to the PRISMA statement.

Methods: The Cochrane Oral Health Trials Register, the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE were searched to November 2011. Relevant conference abstracts were also screened. No language restrictions were applied. Inclusion criteria were: clinical studies with $\geq$ 20 II/2M subjects where comparisons were made with an untreated II/2 M group or another treated II/2 M group or neither. For included studies ranked best on the hierarchy of evidence, assessments of methodological quality and risk of bias were undertaken. Abstracts and, when appropriate, full papers were examined independently by two investigators. Disagreements were resolved through discussion.

Outcomes: Treatment changes and/or stability with or without retainers using the following measures: skeletal, soft tissue, dental, occlusal changes, gingival health, TMJ status and/or related muscular activity or quality of life.

Results: Of the 322 studies identified in the search, 20 met the final inclusion criteria. All were at high risk of bias.

Conclusions: Highly biased evidence exists with regard to management and stability of II/2 M. Guidelines are proposed based on current evidence.

Introduction

Class II division 2 malocclusion(II/2M), characterized by retroclination of the upper incisors and a deep overbite,[1] has a reported prevalence in UK children of 10%.[2] A prevalence of 5-12% has been reported in other European populations[3-6] and 3-4% in US populations[7] with the severe manifestation of "cover-bite" estimated at almost 2%.[8] Although controversy surrounds the accompanying dentofacial characteristics,[9,10] vertical skeletal factors make a greater contribution in more severe forms.[11,12] The high lower lip line with associated resting pressure (approximately 2.5 times greater than the upper lip resting pressure) has been shown to be linked with retroclination of the upper incisors.[13,14] A strong genetic input exists with regard to the underlying skeletal pattern and dental anomalies, especially the increased prevalence of impacted maxillary canines.[15]

Orthodontic treatment of II/2M is recognized as difficult and prone to relapse.[13,14] The randomized clinical trial provides the highest quality evidence with regard to effectiveness of treatment interventions and data from several trials have enabled meta-analysis to be undertaken on the effectiveness of growth modification for Class II division 1 malocclusion.[16] Retrospective controlled studies, have some benefit until results from prospective studies become available[17] and information from these studies have been included in orthodontic systematic reviews on Class III treatment and lingual arch space maintenance.[18,19] Although RCTs and controlled clinical trials (CCTs) have been considered in a previous review,[20] until now it would appear that no review has addressed all of the prospective and retrospective evidence regarding effectiveness of orthodontic treatment and its stability for II/2M.

The aim of this review was to evaluate the evidence with regard to the effectiveness of orthodontic treatment and stability for II/2M in children and adolescents and if possible to identify the most effective treatment strategies through the use of meta-analysis. The null hypothesis tested was that there is no difference in the effectiveness of any of the treatment interventions or in their ability to promote stability for II/2M. Reporting of this review is according to the PRISMA statement.[21]

Materials and methods

Criteria for selecting studies for the review were as follows:

Study design

Randomized and controlled clinical trials (RCTs and CCTs respectively) were included as these are likely to contain acceptable quality evidence. Prospective and retrospective studies with $\geq$ 20 subjects per intervention group were assessed also. A minimum sample of 20 was chosen based on the data from Stellzig et al.,[22] who found that patients with II/2 malocclusion treated with headgear and upper second molar extractions had a reduction in the interincisal angle of 12.6^{o} (SD 10.2^{o}) compared with an historical, untreated control group. Using these data we determined that a sample of 30 (i.e. 15 subjects in each group) would be sufficient to detect a significant difference between treated and untreated groups with a power 90% and $P<0.05$; however to account for the relatively wide SD in incisor inclination[23], which is especially relevant to II/2M outcome assessment, as well as to allow for drop-outs and withdrawals, a minimum sample size of 20 per group was chosen. Smaller samples are acknowledged to be of very limited use, particularly if cephalometric data are being evaluated.[24] Studies where comparisons were made with either an untreated II/2M group or to another treated II/2M or neither were included. Without a control group, very limited conclusions regarding outcomes of treatment can be made due to the increased susceptibility to bias.[25] Individual case reports were not considered for analysis due to the poor quality evidence provided.

Participants

Children and adolescents who had treatment for II/2M were included. Adults (where mean age pre-treatment was $\geq$ 18 years) were excluded due to lack of growth affecting treatment outcome. For studies with mixed child/adolescent and adult samples, only data for the former were considered.

Intervention

Cases treated with one arch or two arch full fixed appliances (with or without extractions) were accepted including those where Class II elastics were utilized without adjunctive appliances. In addition, removable, functional or headgear appliances, in isolation or in combination with fixed appliances, were included. Cases treated by a combined orthodontic-orthognathic approach were excluded as the focus was on orthodontic treatment only and the resultant stability. The type of appliance investigated was recorded to put studies into homogeneous groups, where meta-analysis was feasible.

Outcome measures

Studies were included if they reported data on treatment and/or stability of treatment with regard to one or more of the following measures: skeletal, soft tissue, dental, occlusal changes (preferably assessed with an occlusal index), gingival health, temporomandibular joint status or related muscular activity or quality of life. If stability was assessed, patients were followed up for a minimum of 12 months post-treatment, with or without retainer(s).

Search methods for study identification

Several sources were utilized as a search confined to Medline only is generally deemed to be inadequate. The Cochrane Oral Health Trials Register, the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE were searched to November 2011. Journals which are hand-searched for the Trials Register are given on the Cochrane Oral Health Group website (http://ohg.cochrane.org). To identify records, usually three basic sets of terms are used: those that identify records related to the health condition of interest (II/2M), those used to identify records related to the intervention being evaluated and finally those which identify the type of study design to be included. As a pilot run of the search strategy incorporating type of study design yielded no papers by any database, the search was confined to only two basic sets of search terms.

Details of the search strategies developed for all databases are given in Table I.

Other sources

Conference proceedings and abstracts from the British Orthodontic Conference, European Orthodontic Conference and the International Association for Dental Research Conference were searched up to November 2011.

The references quoted in the studies identified were screened for any further trials and international researchers potentially involved in II/2M clinical trials were contacted in an attempt to identify unpublished or ongoing RCTs and CCTs. No language restrictions were applied.

Study selection

The selection of papers, decision about eligibility, study classification and data extraction were undertaken independently and in duplicate by two assessors without blinding to the authors, appliance type or results obtained. All disagreements were resolved by discussion.

Data extraction and management

The following information was recorded for each eligible study on a customized data collection form: initials of reviewer, authors, year of publication, setting of the study, age and gender of the subjects, study design, defining criteria for the malocclusion, sample size calculation, treatment type and duration, drop outs and type of retention, outcome measures, method(s) of assessment, error study and study results.

The primary outcome measures in the identified studies were skeletal, soft tissue, dental, occlusal or gingival changes with treatment or during an observation period. Secondary outcome measures were temporomandibular joint status and quality of life.

Quality assessment / risk of bias assessment of eligible studies

For the eligible studies ranked highest on the hierarchy of evidence,[26,27] quality was assessed according to the following criteria:

- sample size reported
- sample based on power calculation
- eligibility criteria described
- random allocation
- allocation concealment
- baseline equivalence of groups
- blinding of participants/caregivers (where possible)
- blinding of outcome assessors
- point estimates and variability reported for primary outcome measures
- appropriateness of statistical analysis
- extent of drop outs / exclusions (trials using an intention-to-treat analysis were noted) and
- selective reporting.

For those eligible studies, a description of the quality items was tabulated, together with a judgment of low, high or uncertain risk of bias. Criteria for risk of bias judgments for allocation concealment were according to the Cochrane Handbook for Systematic Reviews of Interventions 5.0.2.[28]

Unit of analyses issues

All eligible studies were assessed for the appropriateness of their analyses.

Assessment of reporting biases

It was planned that if sufficient trials were identified, a funnel plot would be drawn and a formal investigation of the degree of funnel plot asymmetry undertaken using the method proposed by Egger.[29] Asymmetry may represent a true trial and effect size relationship, but may also indicate publication bias and other biases related to sample size.

Analyses

The characteristics of the eligible studies were used to evaluate their clinical heterogeneity. Following data extraction, it was intended to undertake Cochran's test for heterogeneity before any meta-analysis, to produce forest plots demonstrating the overall effects of the treatment interventions.

Results

As described in the PRISMA statement,[21] the review details are given in Figure I.

Studies identified

Of the 322 records resulting from the search strategies, only 23 full text articles (and one abstract) were retrieved for more detailed evaluation. Of these 3 (plus the abstract) were subsequently excluded (Table II[30-33]). Twelve studies (4 prospective; 8 retrospective) dealt with treatment (Table III[34-38], Table IV[39-47]) and 8 studies (all retrospective) dealt with stability (Table V[48-55]). The study types, with numbers per group, were as follows:

- prospective case-control of treatment (1)
- prospective cohort of treatment (1)
- prospective case series of treatment (2)
- retrospective cohort of treatment (4)
- retrospective case series of treatment (4)
- retrospective case-control of stability (1)
- retrospective cohort of stability (1)
- retrospective case series of stability (6).

Key methodological data are summarized in the Tables indicated. For those study designs ranked best on the hierarchy of evidence, a risk of bias assessment was undertaken. Both assessors, however, deemed those studies to be at high risk of bias (Table VI). All other designs were deemed to have inherent high risk of bias.

Prospective case-control study of treatment for II/2M (Table III)[38]

This was the only study where comparisons were made with a contemporaneous untreated II/2M control group. The comparisons were based on age and one of three mandibular growth directions; however, it was not clear, whether the "matching" of the treatment and control groups was done prospectively or retrospectively. Gender was closely, but not exactly, matched between treated and control groups. Non-extraction treatment, started in the late mixed dentition stage for overbite reduction in II/2M subjects with mesofacial or brachyfacial growth patterns (normal or reduced lower facial height respectively), led to a mean forward movement of B point of 4-5 mm compared with the controls during the treatment period (mean duration 2.4 yrs and 2.2 yrs respectively).

Prospective cohort of treatment for II/2M (Table III)[34]

One study followed II/2M cases treated using functional appliance therapy, preceded in some patients by a removable appliance to procline the upper incisors. A distal path of closure was found in 50 per cent of the II/2M sample pre-treatment and electromyographic assessment showed masseter and temporalis activity became more "normal" during functional appliance treatment. Additionally, a low gonial angle was associated with increased masseteric activity.

Prospective case series of treatment for II/2M (Table III)[35,37]

Two studies on the same cohort reported the results of upper removable appliance therapy to procline upper incisors and reduce the overbite, followed by functional appliance therapy. When the antero-posterior distance for the retruded to the intercuspal mandibular position was compared from start of treatment to post-incisor proclination, no statistically significant difference was found. The muscle activity was also unchanged on completion of treatment which contrasts with the study by Moss.[34] The findings led the authors to substantiate the view that the mandible does not position anteriorly during treatment of II/2M.

Retrospective cohort studies of treatment for II/2M (Table IV)[22,40,46,47]

Selection criteria for II/2M varied between studies so inter-study comparison is not appropriate. Extraction[22,40,47] and non-extraction[40,46] treatments were assessed. The only study to use an untreated II/2M

control group, compared extraction of four first premolars *versus* upper second permanent molars.[22] The control group was derived from the Belfast Growth Study, but it was not clear if there was gender matching with the treated groups. Furthermore, the treated and control groups were not equivalent at baseline although almost approximated on completion of treatment. The amount of crowding in either group was not specified. In addition, there was variation with regard to intrusion mechanics for overbite reduction, as well as the use of headgear as an adjunct to fixed appliances in many cases in both extraction groups. Four premolar extractions, rather than upper second permanent molar extractions, produced more retraction of the upper lip and less reduction in the interincisal angle; premolar extraction spaces also re-opened at the end of treatment in more than 40 percent of cases.

With upper premolar extractions only, wide variation existed in naso-labial angle changes; although there was a mean increase of ~2.5^0 for II/2M, this was not significantly different to the II/1M group.[47] Overbite was successfully reduced by several treatment approaches; the mean decrease varied from 1.9mm[46] to almost 5mm[40] and the mean decrease in interincisal angle varied from 6^0 to almost 22^0 respectively.

Retrospective case series of treatment for II/2M (Table IV) [39,41,44,45]

Again, selection criteria for II/2M varied between studies or were not specified.[39] Non-extraction treatment,[39,44] predominantly by functional appliances, appeared to be a common treatment approach but two studies did not specify whether extractions were undertaken or not.[41,45] Maxillary apical base size was the strongest predictor of occlusal change.[45]

Retrospective case-control study of stability for II/2M (Table V)[51]

At a mean time of 15.2 years out of retention, the mean relapse in overbite and interincisal angle correction was 40% and 59% respectively. The overbite relapse mirrored that of the reduction in lower anterior facial height (almost 40%). As the incisor segments uprighted, incisor crowding increased especially in the lower arch but this varied between individuals. Lower arch extractions did not appear to increase post-treatment overbite if appropriate treatment mechanics were used; rather the initial overbite was the best predictor of post-treatment overbite but predictability was not high ($R^2 = 0.42$). The chance of maintaining an overbite <4mm in the long term was deemed to be less than 50%. Post-treatment vertical facial growth contributed to maintenance of overbite correction. Molar relationship correction was very stable.

Retrospective cohort study of stability for II/2M (Table V)[55]

Following two-phase (Herbst and fixed appliances) non-extraction treatment and an average of 27 months retention, overbite correction was more stable in late (~86%) than in early adolescents (70%). Molar relationship relapsed minimally (5-7%).

Retrospective case series evaluating stability of II/2M treatment (Table V)[48-50,52-54]

The retention type, duration of retention and treatment approach varied, although non-extraction was again favoured across all studies. The overbite increased post-treatment[49,50] with an associated increase in interincisal angle[49] and relapse in maxillary incisor inclination correction;[48,52,54] the latter was found to be independent of retainer type and there was large inter-individual variation.[52,54] Incisor crowding increased simultaneously with overbite relapse[48,50] and was more marked in the lower arch,[50] supporting the findings of Kim and Little.[51] Lower incisor proclination and expansion of the intercanine width relapsed.[50] The former was regarded to be more stable than upper incisor proclination,[48] although both incisor segments uprighted.[48,49,54] The greater the treatment change in upper incisor inclination, the greater the relapse.[52,54] In a sample with a mix of removable and fixed appliance treatments (some combined), a mean value of 25% relapse in overbite was found at 2 years post-treatment.[52] In other post-retention studies, the mean overbite relapse varied from ~20% to ~30% (0.8-1.2mm) when assessed at 2 and 5 years respectively[49] to ~26% at a mean period of 7 years (0.96mm).[50] Time post-retention was correlated with the extent of overbite relapse

and lower incisor irregularity.[50] No variables were found to determine the prognosis for overbite stability,[50] with an anterior growth rotation of the mandible evident post-treatment, especially in male subjects.[49] Overcorrection of overbite did not show net improvement at a mean time of 7 years out of retention.[50]

The level of the lower lip post-treatment had a significant influence on the relapse tendency of the corrected incisors relationships.[52,53] Although recommended to reduce lower lip coverage to a maximum of 3mm,[52] a mean decrease of 0.6mm while statistically significant was not judged to be clinically significant.[54] At a minimum of 3 years post-retention, 10% of upper arches and 30% of lower arches had unacceptable irregularity.[50] Molar relationship was very stable after correction[50] which confirmed the findings of others.[51]

Due to the heterogeneity of all the included studies, it was not possible to undertake a meta-analysis to determine the most effective means of treatment or stability for this malocclusion. There is insufficient high quality evidence to reject the null hypothesis tested in this review.

Discussion

This systematic review found no RCTs investigating the effectiveness and stability of orthodontic treatment in children and adolescents with a II/2M, even though evidence of that quality is available for other malocclusions.[16,57] Unfortunately lack of RCT evidence is a consistent finding in both dental[58] and orthodontic systematic reviews.[59,60]

What are the implications of this review for clinical practice?

For ethical, administrative and financial reasons, RCTs are difficult to conduct for II/2M[61] as well as for other orthodontic questions.[17] Furthermore, they appear not to have provided extra knowledge to that already available from retrospective studies for treatment of Class II malocclusions,[62] but are likely to overestimate the treatment effect by approximately 30 per cent.[63] All the evidence found in this review was deemed to be at high risk of bias. So one is faced with a dilemma: on what does one base clinical practice for II/2M? Successful correction of II/2M has been achieved over many years (Tables III-V) and in the absence of well conducted clinical trials, it seems reasonable to re-visit this evidence in an attempt to develop broad guidelines for clinical practice. This approach adds some additional perspective on II/2M and appears to be all that is possible in the short term. The comments of Litt and Nielsen,[64] re-iterated by Stellzig et al,[22] seem to be echoed even in present times: "The clinical management of II/2M remains a "mystery" entailing problems of diagnosis, therapy and retention; when therapeutic aspects are condensed to a central topic, these publications are mainly restricted to general guidelines."

Guidelines for clinical practice and relevance

Based on the limited and highly biased evidence, it is possible to make recommendations only in the broadest sense.

Treatment timing and prediction of outcome

Prospective evidence indicates that to maximize favorable dentoalveolar and soft tissue changes, the facial growth pattern should be identified early and the deep overbite managed in a timely manner.[38] The relevance of this, as indicated by the author, is that the magnitude of mandibular growth will impact on the necessary amount of maxillary arch retraction during Class II correction.[38] Limited prospective evidence exists that growth modification by functional appliance therapy may also "normalize" muscle function,[34] which could assist with stability. As post-treatment vertical facial growth assists with maintenance of stability,[51] treatment should be timed to allow for this, where possible.

The size of the maxillary apical base has been suggested as a possible factor to influence the success of occlusal correction as it is "conceivable that the larger the transverse and sagittal extent of the maxillary apical base, the more unfavourable is the prospect of mechanically correcting II/2M."[45] The authors,

however, acknowledged that other factors, such as growth, patient compliance and operator proficiency are all relevant, but were not assessed. Further work on predictive models for treatment success of II/2M is necessary.

Extraction versus non-extraction

It is suggested that caution be exercised with regard to four first premolar extractions in view of the potential detrimental effects on facial profile, overbite and re-opening of the extraction spaces post-treatment.[22] Comparison, however, was made to an upper second permanent molar extraction group, so it is doubtful if the groups were comparable pre-treatment. Nonetheless, the propensity for maxillary extraction spaces (mainly first premolar) to re-open post-treatment was also found in another study and led to a greater relapse tendency of the corrected upper incisor inclination, particularly the more anterior the extraction;[52] this would appear to favour extraction further posteriorly in the arch or a non-extraction approach.

Only one study evaluated the soft tissue effects of Class II (II/2M versus II/1M) fixed appliance treatment with upper premolar extractions alone (camouflage treatment).[47] The authors concluded that such a treatment approach is likely to result in a wide variation in lip and upper incisor behavior. Furthermore, they suggested that negative effects on upper lip curve and nasolabial angle are more likely where the upper lip is thin pre-treatment and the nasolabial angle is increased. Upper premolar extractions would appear to be best avoided in such cases.

Extractions did not seem to influence the magnitude of overbite correction if adequate appliance therapy was used but the authors cautioned about making any inferences from this as the sample size did not allow breakdown into further subgroups according to treatment (extraction vs non-extraction), sex or age.[51] Non-extraction treatment,[34,44,50-52] however, seems favoured with functional appliances or a removable appliance used by several in advance of fixed appliances; this approach takes advantage of vertical facial growth to assist overbite correction.

Overbite reduction and correction of interincisal angle

Overbite may be reduced successfully by several means,[40,49,51] but success depends on alteration of interincisal angle,[48] moving the upper incisors from lower lip control possibly by intrusion [47,53,54] and vertical facial growth.[48] Correction of upper incisor inclination to as near normal as possible is advised; overcorrection is more prone to relapse.[53] Based on data from one study, overcorrection of the overbite would also appear not to be advisable as changes were not upheld long-term.[50] As with all malocclusion types, lower arch expansion, either antero-posteriorly or laterally is not recommended,[50] unless permanent retention is considered.

Retention planning and follow-up

To maintain overbite and interincisal correction, as well as incisor alignment, long-term retention is necessary.[54] In growing patients a bite-raising appliance, which also maintains the upper incisor inclination, is recommended to combat overbite relapse due to the tendency for anterior mandibular growth rotation.[49,54] Follow-up for 5 years minimum is advised as many skeletal, soft tissue and dental variables showed significant change from 2 to 5 years post-retention.[49] In addition, at 9 years post-treatment the lower lip level contributed to more of the relapse variability than at 2 years post-treatment.[54] A tighter control schedule or permanent palatal bonded retention is suggested,[52] but the type of upper retainer (Hawley or palatal bonded retainer) did not appear to influence upper incisor stability at a mean of 3.5 years post-treatment.[54]

Stability and prediction of relapse

Overbite correction appears reasonably stable, with on average ~20% relapse at 2 years post-retention;[49] but this doubled at 15 years follow-up.[51] Upper incisor inclination correction was also reasonably stable.[54] There is great inter-individual variation[53] and incisor crowding returned, especially in the lower arch.[52] There was a greater tendency to relapse of the corrected upper incisor inclination where the lower lip line was high post-treatment.[52,53] For maximum treatment stability, removal of excessive overlap of the upper

incisors by the lower lip has been highlighted as one of the most important treatment objectives.[53] Molar correction appeared to be particularly stable.[51] As it is not possible to predict overbite stability,[50] long-term retention is recommended. Clinicians should realise that those patients who present with the most upright upper and lower incisors tended to have a deeper initial overbite pre-treatment and a tendency to return to their original relationship post-retention;[51] particular vigilance should be paid to prescribing retention and monitoring occlusal change during and out of retention in these patients.

Implications for future research

Clear and reproducible defining criteria used in sample recruitment/selection should be reported in future studies to eliminate the variability of examiner classification of II/2M.[65] It is recommended that in order to be classified as II/2M, the upper incisor inclination to the maxillary plane or SN line should be greater than 1 SD beyond the mean for the ethnic group from which the sample is drawn.[22] Baseline matching of the study samples with regard to age, gender, skeletal pattern, amount of crowding and incisor inclinations is important in order to control for growth, in particular, as a confounding factor on treatment outcomes. Sample size calculations based on identified outcome measures should also be included in all future studies. More complete reporting of statistical analyses is also required with point and related variability data to be included for a limited number of relevant and clinically meaningful cephalometric landmarks. Patient reported outcome measures should also be included.[66]

A contemporaneous control group of II/2M subjects should be used in future comparative studies of treatment and stability and not subjects with a Class I occlusion/malocclusion; the former would eliminate bias due to secular changes affecting facial growth, which have been observed over recent decades.[67] The eight-year recruitment period in the study by Woods[38] highlights the difficulty of prospectively recruiting a treatment and control group for II/2M, despite offering a fee waiver. Similar difficulty has been encountered in a recent RCT.[62] In the light of this and the reported prevalence of II/2M, longitudinal *multicenter* international trials are required in order to achieve the sample size required for appropriate statistical analyses of treatment and stability outcomes (case–control and preferably randomized clinical trials). Recommended for orofacial clefting to recruit sufficient data,[68] such multicenter collaborative trials require control for racial growth variables for outcome analysis.

Conclusions

Highly biased prospective evidence exists with regard to the effect of late mixed dentition non-extraction treatment on facial growth in II/2M. Prospective and retrospective highly biased evidence appears to favour non-extraction treatment and indicates overbite correction to be reasonably stable in the short-term. International multicenter collaborative studies are required to gather appropriate epidemiological evidence regarding this condition. Prospective international studies are required (either case-control or randomized trials) to provide stronger evidence on treatment and stability for II/2M in children and adolescents.

References

1. van der Linden FPGM. Development of the dentition. Chicago: Quintessence Publishing Co. Inc.; 1983.
2. Foster TD, Day AJ. A survey of malocclusion and the need for orthodontic treatment in a Shropshire school population. Br J Orthod. 1974; 1:73-8.
3. Markovic M. A genetic study of Class II Division 2 malocclusions. Eur J Orthod. 2000; 22:453-4.
4. Mierut LL. Class II division 2 malocclusions: Frequency and upper incisor aspects. Eur J Orthod. 2000; 22:602.
5. Myllarniemi S. Malocclusion in Finnish rural children An epidemiological study of different stages of dental development. Suom Hammaslaak Toim. 1970; 66:219-64.
6. DeBruyne I, Willems G, Carels C, Fieuws S. Prevalence and characteristics of malocclusion in a Belgian orthodontic population. [abstract] Eur J Orthod. 2000; 22:575.

7. Massler M, Frankel JM. Prevalence of malocclusion in children aged 14 to 18 years. Am J Orthod. 1951; 37:751-68.
8. Peck S, Peck L, Kataja M. Class II Division 2 malocclusion: a heritable pattern of small teeth in well-developed jaws. Angle Orthod. 1998; 68:9-20.
9. Karlsen AT. Craniofacial characteristics in children with Angle Class II div. 2 malocclusion combined with extreme deep bite. Angle Orthod. 1994; 64:123-30.
10. Pancherz H, Zieber K. Dentoskeletal morphology in children with Deckbiss. J Orofac Orthop. 1998; 59:274-85.
11. Brezniak N, Arad A, Heller M, Dinbar A, Dinte A, Wasserstein A. Pathognomonic cephalometric characteristics of Angle Class II Division 2 malocclusion. Angle Orthod. 2002; 72:251-7.
12. Siriwat PP, Jarabak JR. Malocclusion and facial morphology is there a relationship? An epidemiologic study. Angle Orthod. 1985; 55:127-38.
13. Lapatki BG, Klatt A, Schulte-Monting J, Jonas IE. Dentofacial parameters explaining variability in retroclination of the maxillary central incisors. J Orofac Orthop. 2007; 68:109-23.
14. Lapatki BG, Mager AS, Schulte-Moenting J, Jonas IE. The importance of the level of the lip line and resting lip pressure in Class II, Division 2 malocclusion. J Dent Res. 2002; 81:323-8.
15. Basdra EK, Kiokpasoglou M, Stellzig A. The Class II Division 2 craniofacial type is associated with numerous congenital tooth anomalies. Eur J Orthod. 2000; 22:529-35.
16. Harrison JE, O'Brien KD, Worthington HV. Orthodontic treatment for prominent upper front teeth in children. Cochrane Database Syst Rev 2007, Issue 3. Art. No.: CD003452. DOI: 10.1002/14651858.CD003452.pub2
17. Johnston LE, Jr. Moving forward by looking back: 'retrospective' clinical studies. J Orthod. 2002; 29:221-6.
18. Toffol LD, Pavoni C, Baccetti T, Franchi L, Cozza P. Orthopedic treatment outcomes in Class III malocclusion. A systematic review. Angle Orthod. 2008; 78:561-73.20.
19. Viglianisi A. Effects of lingual arch used as space maintainer on mandibular arch dimension: a systematic review. Am J Orthod Dentofacial Orthop. 2010; 138:382 e1-4.
20. Millett DT, Cunningham S, O'Brien KD, Benson PE, Williams A, de Oliveira CM. Orthodontic treatment for deep bite and retroclined upper front teeth in children. Cochrane Database Syst Rev 2011, Issue 4. Art. No.: CD005972. DOI: 10.1002/14651858.CD005972.pub2.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 2009; 6:e1000097. doi:10.1371/journal.pmed.
22. Stellzig A, Basdra EK, Kube C, Komposch G. Extraction therapy in patients with Class II/2 malocclusion. J Orofac Orthop. 1999; 60:39-52.
23. Stabrun AE, Danielsen K. Precision in cephalometric landmark identification. Eur J Orthod 1982; 4: 185-96.
24. Proffit WR, Fields HW, Sarver DM. Contemporary Orthodontics 4th Edition ed. St Louis: Mosby/Elsevier; 2007.
25. Glenny A-M, Harrison JE. How to…interpret the orthodontic literature. J Orthod 2003; 30:159-64.
26. NHS Centre for Reviews and Dissemination. Undertaking Systematic Reviews of Research on Effectiveness: CRD's guidance for carrying out of or commissioning reviews, 2nd edn. York: NHS Centre for Reviews and Dissemination, University of York, 2001.
27. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RV. Evidence-based medicine: how to practice and teach EBM. Churchill Livingstone 2nd Ed; 2000.
28. Higgins JPT, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions 5.0.2 (updated September 2009). The Cochrane Collaboration. Available from **www.cochrane-handbook.org.2009**.
29. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997; 315:629-34.
30. Cleall JF, BeGole EA. Diagnosis and treatment of Class II division 2 malocclusion. Angle Orthod. 1982; 52:38-60.

31. Erickson LP, Hunter WS. Class II, division 2 treatment and mandibular growth. Angle Orthod. 1985; 55:215-24.
32. Pancherz H, von Bremen J. Outcome of Class II division 2 and Class II division 1 therapy using different treatment approaches. [abstract] Eur J Orthod. 2000; 22:455.
33. Kinzel J, Aberschek P, Mischak I, Droschl H. Study of the extent of torque, protrusion and intrusion of the incisors in the contect of Class II, Division 2 treatment in adults. J Orofac Orthop. 2002; 63: 283-99.
34. Moss JP. An investigation of the muscle activity of patients with Class II Division 2 malocclusion and the changes during treatment. Trans Eur Orthod Soc. 1975; 87-101.
35. Demisch A, Ingervall B, Thuer U. Mandibular displacement in Angle Class II, division 2 malocclusion. Am J Orthod Dentofacial Orthop. 1992; 102:509-18.
36. Ingervall B. Recording of retruded positions of mandible in children. *Odontol Rev*. 1968;19:65–82
37. Thuer U, Ingervall B, Burgin W, Demisch A. No posterior mandibular displacement in Angle Class II, division 2 malocclusion as revealed with electromyography and sirognathography. Eur J Orthod. 1992; 14:162-71.
38. Woods MG. Sagittal mandibular changes with overbite correction in subjects with different mandibular growth directions: late mixed-dentition treatment effects. Am J Orthod Dentofacial Orthop. 2008; 133:388-94.
39. Stefani E. Veranderungen der Weichgewebe nach funktionskiefer-orthopadischer Behandlung im Fernrontgen-Seitenbild - eine statistische Untersuchung bei der Anomalie Angle-Klasse II, 2. Fortschr Kieferorthop. 1984; 45:49-54.
40. Parker CD, Nanda RS, Currier GF. Skeletal and dental changes associated with the treatment of deep bite malocclusion. Am J Orthod Dentofacial Orthop. 1995; 107:382-93.
41. Eberhard H, Hirschfelder U. Treatment of Class II, Division 2 in the late growth period. 0 J Orofac Orthop. 1998; 59:352-61.
42. Björk A, Helm S. Prediction of the age of maximum puberal growth in body height. Angle Orthod. 1967; 37:134-42.
43. Droschl D. Die fernröntgen werte unbehandelter kinder zwischen dem 6. und 15. Berlin-Chicago-London-Rio de Janeiro-Tokio: Quintessenz; 1984.
44. Kalavritinos MK. Dentoskeletal and esthetic changes of facial profile following activators treatment of Class II, Division 2 malocclusions. Hellenic Orthod Rev. 2001; 4:21-36.
45. Zentner A, Peylo S, Brothag D. Predictive value of morphologic parameters for successful correction of Class II Division 2 malocclusion. Am J Orthod Dentofacial Orthop. 2003; 123:279-85.
46. Honn M, Schneider C, Dietz K, Godt A, Goz G. Treating Class II patients with removable plates and functional orthopedic appliances-the importance of anterior tooth inclination and direction of growth on treatment outcome. J Orofac Orthop. 2006; 67:272-88.
47. Tadic N, Woods MG. Incisal and soft tissue effects of maxillary premolar extraction in class II treatment. Angle Orthod. 2007; 77:808-16.
48. Mills JR. The problem of overbite in Class II, division 2 malocclusion. Br J Orthod. 1973; 1:34-48.
49. Binda SK, Kuijpers-Jagtman AM, Maertens JK, van 't Hof MA. A long-term cephalometric evaluation of treated Class II division 2 malocclusions. Eur J Orthod. 1994; 16:301-8.
50. Canut JA, Arias S. A long-term evaluation of treated Class II division 2 malocclusions: a retrospective study model analysis. Eur J Orthod. 1999; 21:377-86.
51. Kim TW, Little RM. Postretention assessment of deep overbite correction in Class II Division 2 malocclusion. Angle Orthod. 1999; 69:175-86.
52. Lapatki BG, Klatt A, Schulte-Monting J, Stein S, Jonas IE. A retrospective cephalometric study for the quantitative assessment of relapse factors in cover-bite treatment. J Orofac Orthop. 2004; 65:475-88.
53. Lapatki BG, Baustert D, Schulte-Monting J, Frucht S, Jonas IE. Lip-to-incisor relationship and postorthodontic long-term stability of cover-bite treatment. Angle Orthod. 2006; 76: 942-49.
54. Devreese H, De Pauw G, Van Maele G, Kuijpers-Jagtman AM, Dermaut L. Stability of upper incisor inclination changes in Class II division 2 patients. Eur J Orthod. 2007; 29:314-20.
55. Bock N, Ruf S. Post-treatment occlusal changes in Class II division 2 subjects treated with the Herbst appliance. Eur J Orthod. 2008; 30:606-13.

56. Hagg U, Taranger J. Skeletal stages of the hand and wrist as indicators of the pubertal growth spurt. Acta Odontol Scand. 1980;38:187-200.
57. Mandall N, DiBiase A, Littlewood S, Nute S, Stivaros N, McDowall R, et al. Is early Class III protraction facemask treatment effective? A multicentre, randomized, controlled trial: 15-month follow-up. J Orthod. 2010; 37:149-61.
58. Pjetursson BE, Tan WC, Tan K, Bragger U, Zwahlen M, Lang NP. A systematic review of the survival and complication rates of resin-bonded bridges after an observation period of at least 5 years. Clin Oral Implants Res. 2008; 19:131-41.
59. Fleming PS, DiBiase AT. Systematic reviews in orthodontics: what have we learned? Int Dent J. 2008; 58:10-4.
60. Bollen AM, Cunha-Cruz J, Bakko DW, Huang GJ, Hujoel PP. The effects of orthodontic therapy on periodontal health: a systematic review of controlled evidence. J Am Dent Assoc. 2008; 139:413-22.
61. Cunningham S, Bearn D, Benson P, Johal A, Millett D, O'Brien K, Luther F. In search of the sample: recent experiences of a trial team in orthodontics. Contemp Clin Trials 2011; 32: 530-4.
62. Meikle MC. What do prospective randomized clnical trials tell us about the treatment of Class II malocclusions? A personal viewpoint. Eur J Orthod 2005; 27:105-14.
63. O'Brien KD, Personal Communication. 2011.
64. Litt RA, Nielsen IL. Class II, division 2 malocclusion. To extract--or not extract? Angle Orthod. 1984; 54:123-38.
65. Williams AC, Stephens CD. A modification to the incisor classification of malocclusion. Br J Orthod. 1992; 19:127-30.
66. Vig KW, Weyant R, O'Brien K, Bennett E. Developing outcome measures in orthodontics that reflect patient and provider values. Semin Orthod 1999; 5:85-95.
67. Warren JJ, Bishara SE. Comparison of dental arch measurements in the primary dentition between contemporary and historic samples. Am J Orthod Dentofacial Orthop. 2001; 119:211-5.
68. Mossey P. Epidemiology underpinning research in the aetiology of orofacial clefts. Orthod Craniofac Res. 2007; 10:114-20.

Legends for Tables and Figures

Figure I

Search flow (as described in the PRISMA statement) [21]

Tables

Table I: Search strategies adopted for review of II/2M studies

Table II: Excluded studies of II/2M with reason(s)

Table III: Prospective studies of treatment for II/2M

Table IV: Retrospective studies of treatment for II/2M

Table V: Retrospective studies of stability for II/2M

Table VI: Quality assessment and indication of risk of bias of studies of treatment (prospective case-control, prospective cohort, prospective case series, retrospective cohort) and stability (retrospective case-control, retrospective cohort) for II/2M.

PRISMA 2009 Flow Diagram

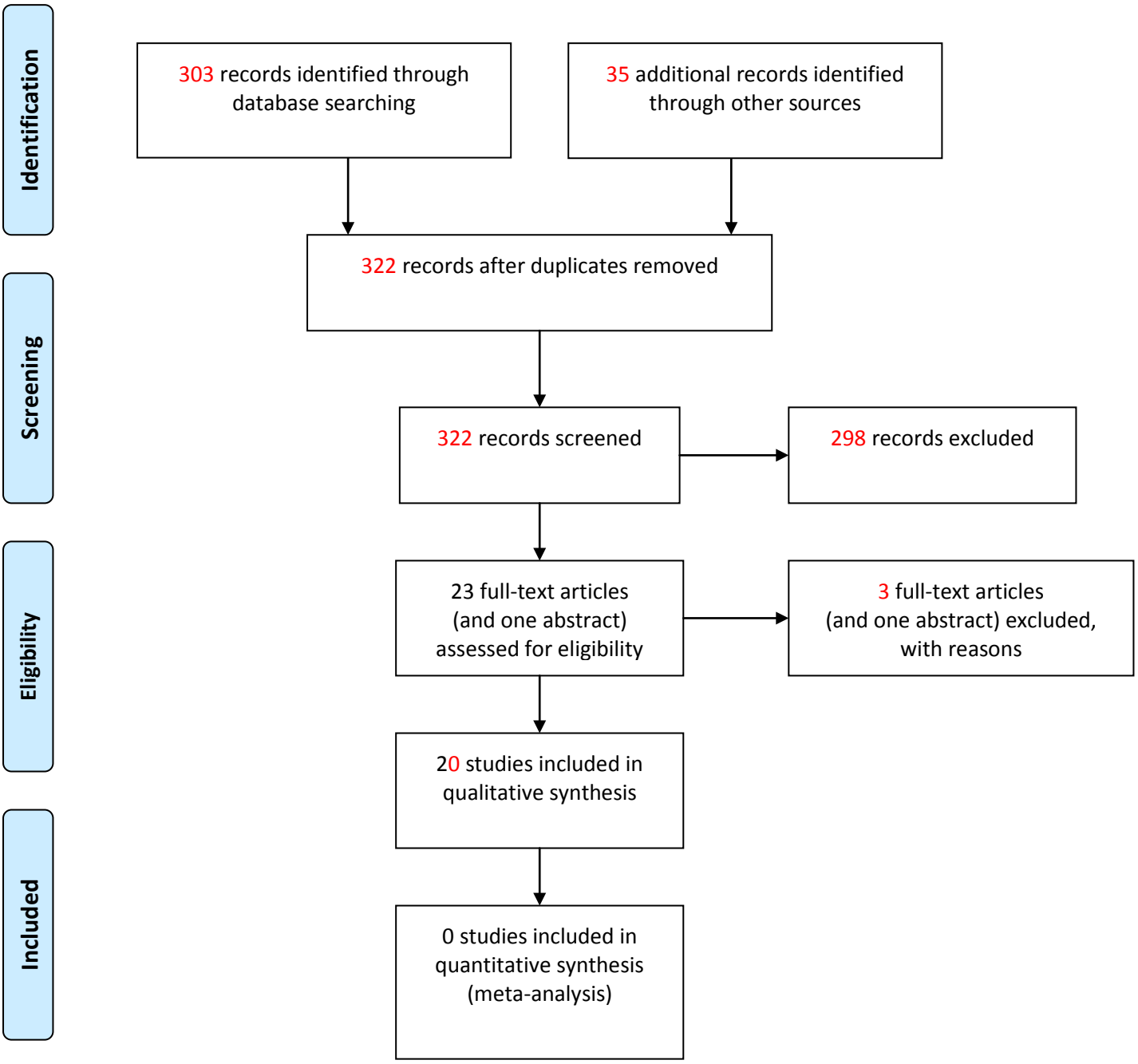

Table I Search strategies adopted for review of II/2M studies

Database		*Search history*	*Results*
OHG reg*	1	((Malocclusion* OR bite* OR Angle* OR class) AND ("division 2" OR "div* 2" OR "div* II" OR "div* II"))	2
CENTRAL**	1	MALOCCLUSION, ANGLE CLASS II (Single term)	1
	2	(“class II” AND (angle* OR malocclusion* OR bite*))	
	3	1 AND 2	
	4	“div* 2” OR “div* II”	
	5	3 AND 4	
MEDLINE	1	Malocclusion, Angle Class II/	264
	2	("Class II" and (angle$ or malocclusion$ or bite$)).mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
	3	or/1-2	
	4	("div$ 2" or "div$ II").mp. [mp=title, original title, abstract, name of substance word, subject heading word]	
	5	3 AND 4	
EMBASE	1	exp Malocclusion/	30
	2	malocclusion* OR bite*	
	3	1 AND 2	
	4	((angle* OR ('class 2' OR 'class ii')) AND ('division 2' OR 'division ii'))	
	5	3 AND 4	

***** OHG TRIALS REGISTER
****** Cochrane central register of controlled trials

Table II Excluded studies of II/2M with reason(s)

Author/year	*Reason for exclusion*
[RCS] Cleall and BeGole 1982[30]	Mean age not specified so unclear if subjects ≤18.
[RCoS] Erickson and Hunter 1985[31]	Only 15 II/2M control subjects *vs* 34 (groups of 6, 14, 14) II/2M treated subjects
[RCS] Pancherz and von Bremen 2000[32] (Abstract)	One of the two treated II/2M groups only 14 subjects *vs* 23 in other treated group; PAR assessment
[RCSTS] Kinzel et al 2002[33]	Adult treatment; 25 cases but only 11 at follow-up

[RCoS], Retrospective cohort study matched to some "controls"; [RCS], Retrospective case series; [RCSTS], Retrospective case series of treatment and stability

Table III Prospective studies of treatment for II/2M

Authors / year *Study design*	*No .of cases/ sex of subjects* *Mean age (range)*	*Sample size calculation* *Setting/Operator* *Drop-out*	*Definition of* ***malocclusion***	*Extraction/ non-extraction* *Treatment /duration* *Retention type/duration* ***Mean time out of retention***	***Error study***	***Outcome measures (OM) / Results (R)***
Moss 1975 [34] Cohort	23 II/2M; 11.9 SD 2.3 yrs 22 control: normally developing occlusion	No Not specified Ceph,3 drop out. EMG:8 drop out at end of tx and 4 more at least one year out of retention.	Not stated	Non-extraction Activator (some cases preceded by URA to procline UIs) 15 recorded at end of tx; 11 recorded at least 1yr out of retention Not stated	Yes	OM: ceph and EMG changes R:Pre-tx: anterior masseter activity > anterior temporalis; posterior temporalis slightly > posterior masseter Lower gonial angle associated with increased masseteric activity Distal displacement of jaw on closing in ~50% of cases Post-tx muscle activity more normal One case that relapsed: no change in muscle activity during tx or during retention
Demisch et al 1992[35] Case series	22 (11M, 11F) Median age pre-treatment 10yr 2mo	No University / 1 operator 1 failed to complete phase 2 and 3; tx discontinued due to lack of co-operation	Bilateral distal occlusion with retroclined UI`(central); Large OB; No symptoms or signs of mandibular dysfunction	Non-extraction 2-phase tx (Phase1) Proclination of UI and bite raising with URA; (Phase 2) Herren activator for night-only wear Median time (1) : 207 days Median time (2) : 270 days Phase 3: night only wear of modified retainer activator; median time 466 days	Yes for RCP-ICP positions	OM: ceph and recording of RCP/ICP in a modified gnatho-thesiometer. as described by Ingervall (1968)[36] R:median a-p difference between RCP and ICP: before tx 1.23 mm; after UI proclination 1mm; after activator 0.55mm; after retention 0.95mm NS difference between before and after UI proclination: so mandible not posteriorly displaced. OB: median decrease of 3mm with tx IIA: median decrease 15.2^0 with tx; median increase of 4.6^0 after retention UI: median increase 8.2^o; after tx and further 0.5^0 after retention. LI: median increase 4.9^0 with tx and median decrease 4.4^0 after retention

Thuer et al 1992[37] Case series	As for Demisch et al 1992	No As for Demisch et al 1992 As for Demisch et al 1992	As for Demisch et al 1992	Non-extraction As for Demisch et al 1992 Last recordings at 12 months with activator As for Demisch et al 1992	Yes	OM: EMG readings: 27 and 22 days before start of tx; end of phase 1 median time 207 days after start of treatment); 3 and 12 months after start of activator tx Sirognathograph recordings: position of mandible at rest, at intercuspation and during tooth contact during chewing and maximal mandibular movements R: Muscle activity unchanged during period of observation; activity decreased during maximal\biting, chewing and swallowing Positions of mandible at rest, at intercuspation, during chewing were stable during tx No signs recorded of anterior mandibular positioning during tx
Woods 2008[38] **Case-control**	**Results based on:** **C: 93; 49F, 44M** **T: 92; 51F, 41M** **Dolicofacial* (facial axis <870)** **C= 19; mean age 11.5 SD 1 yr)** **T= 15; mean age 11.6 SD 1.1 yr** **Mesofacial (facial-axis 87-930)** **C= 38; mean age 12 SD 1.2 yr** **T=39; mean age 12 SD 1.2 yr** **Brachyfacial (facial axis > 930)** **C=36; mean age 11.5 SD 1yr** **T= 38; mean age 11.2 SD 1.1yr**	**No** **Not stated** **All treated by an experienced orthodontist using consistent treatment regimen** **Drop outs for various reasons but not specified**	**UI to NA <180** **Overbite > 5mm;** **> 5mm Class II molars;** **<5mm mandibular arch crowding**	**Non-extraction** **U/L FA with HG (Ricketts mechanics)** **C= min 2.5yr** **Dolicofacial*** **T=2.5 yr** **Mesofacial** **T=2.4 yr** **Brachyfacial** **T=2.2 yr** **Retention not specified**	**Yes**	**OM: Ceph** **assessment of sagittal mandibular changes with overbite correction with different mandibular growth directions** **Mean changes during active treatment or a minimum 30-month control period** **No evidence of sexual dimorphism for point B or Pog change (control or treatment), results were pooled** **R: Mean point B change significant for Mesofacial (C=2.6 SD 1.6mm; T=6.8 SD 2.1mm; p<0.01)** **Brachyfacial (C=2.6 SD 2.3mm, T=8.2 SD 1.9mm; p<0.01)** **Mean Pog change significant for Brachyfacial 8.3 SD 2.4mm; p<0.05**

II/2M, Class II division 2 malocclusion; SD, standard deviation; yr, year; Ceph, lateral cephalometric radiograph; EMG, electro-myographic; URA, upper removable appliance; UI,upper incisor; tx, treatment; M, male; F, female; mo, month; OB, overbite; RCP-ICP, retruded contact position to intercuspal position; a-p, antero-posterior; NS, non-significant; IIA, inter-incisal angle; LI, lower incisor; C, Control; T, Treatment; *not considered for evaluation as $\leq$ 20 subjects in both C and T groups; UI to NA, upper incisor angulation to Nasion-A line; U/L FA, Upper and lower fixed appliances; HG, headgear; Pog, Pogonion.

Table IV Retrospective studies of treatment for II/2M

Authors / year *Study design*	*No. of cases* ***Mean age (range)/ sex of subjects***	*Sample size calculation* *Setting* ***Drop-out***	*Definition of* ***malocclusion***	*Extraction/ non-extraction* *Treatment / duration* *Retention type/duration*	***Error study***	***Outcome measures (OM) / Results (R)***
Stefani 1984[39] Case series	34 Mean age/gender of subjects n/g	No University clinic No	Not given	Non-extraction Removable functional appliances Not stated	No	OM: ceph changes. R: Soft tissue profile changes
Parker et al 1995[40] Cohort	44 II/2M 27 II/1M 61 Class 1 Age range entire sample: 11 yr to 15yr 9mo	No Private Practice No drop-outs recorded	Angle classification; At least 70% anterior OB	II/2M: 33 non-extraction; 11 extraction II/1M:17 non-extraction; 10 extraction Class I: 35 non-extraction; 26 extraction Six treatment modalities to reduce OB Average tx time entire sample: 31 mo (range 14 to 48) Not stated	Yes	OM: skeletal and dental changes with OB tx assessed on cephs and dental casts R: mean tx changes for II/2 sample OB: mean decrease ~5mm LI to NB: mean decrease ~7.6^0 UI to SN: mean increase ~14.7^0 IIA: mean decrease by ~22^0 LAFH increased significantly (mean 3.2mm) with all tx mechanics For extraction vs no-extraction II/2M : maxillary 6 to SN line in mm and maxillary 6 to perpendicular to the SN line at S were significant. LI intrusion only found in 38% of II/2M

Eberhard and Hirschfelder 1998[41] Case series	22 (12M; 10F) Mean age start: M 14 SD 0.9 yr; F 12.3 SD 0.4 yr)	No Not specified No drop-outs recorded	II/2M Ob>4mm Ml-NL <18 Ar-Go-Me < 118 Growth stage:DP3U (Bjork).[42]	Not specified. Herbst and UFA 6.4 SD 0.2 mo; then after Herbst removal, LFA fitted (before Herbst 6M, 4F unsuccessfully treated with removable appliances) -records taken at start of tx, at about 6mo later on removal of Herbst and again 6mo later Retention: Class II elastics worn for 24 hrs per day, later reduced if permissible.	Not stated	OM: changes with tx assessed with cephs/study casts/dpts R: IIA: mean decrease by 19^0 (after 6 mo removal of Herbst, mean change of 1.9^n) Significant improvement in vertical jaw base relationship stable after 12 months Intrusion of upper 6's and lower 4's observed; minor protrusion of lower incisors
Stellzig et al 1999[22] **Cohort**	**Gp1: 20 (8M; 12F)** **mean age start 11.1 yr** **mean age end 15.3 yr** **Gp2: 20 (8M; 12F)** **Mean age start 13.1yr** **Mean age end 15.3 yr** **Control: 20 untreated II/2M from Belfast Growth Study (ages 9-15 yrs)**	**No** **Not stated** **No drop-outs recorded**	**UI-SN of 1SD below mean of Class I occlusion patients of same age from study of Droschl 1984.[43]** **For control:** **II/2M** **ANB>4°** **UI-SN <96°**	**Gp1: extraction of 4x4** **19 treated with U/L FA** **(15 also wore H/G)** **1 treated with functional;** **Class II elastics also used with FA;** **intrusion mechanics used in 17 subjects for overbite reduction** **Gp2: extraction of 7's** **19 treated with U/L FA** **1 treated with functional;** **all wore HG** **intrusion mechanics used in 4 subjects for overbite reduction** **Mean tx duration** **Gp1: 4.2 yrs** **Gp2: 2.2 yrs** **Retention not stated**	**Yes**	**OM: Ceph/ study model/ dpt changes** **R: Treated vs untreated:** **Gp1: significant marked recession of upper lip** **IIA: only slight decrease (3.9°)** **OB: mean decrease 2.3mm** **Gp 2: Only slightly increased flattening of upper lip** **IIA: mean decrease (12.6 °) to value approaching controls (131.5°)** **OB: mean decrease 2.9mm** ***After premolar extraction, renewed spacing present in 41% of cases**

Kalavritinos 2001[44] Case series	39 (13M; 17F) Mean age start: M 10yr: F 8 yr 10mo	No University No drop-outs recorded.	UI palatal inclination; decreased values for facial angle, SNB, SNPg, reduced lower facial height	Non-extraction Activator only (Andresen-Haupl) with UI 0.7mm protrusion spring Cases treated 1962-1982 Mean duration: M, 3.7yrs; F, 3.5 yrs Not stated	Yes	OM: dentoskeletal and soft tissue changes with activator treatment assessed on cephs R: UI to maxillary plane: mean increase during tx ~4^0; LI to anterior margins of maxilla and mandible: mean increase during tx ~2^0 Mean increase 1.4^0 with tx in SNPg
Zentner et al 2003[45] Case series	96 Average age start: 12.16 SD 4.03 yr	No Not specified No drop-outs recorded.	Distal molar relationship, increased ob, retroclined UI	Not stated Various removable and fixed appliances and combinations of these. Tx and retention: 49.83 SD 18.66 mo Not stated	Yes	OM: Angular and linear changes on ceph / % PAR score change on study casts R: Maxillary apical base size strongest predictor of occlusal correction (%PAR reduction)
Honn et al 2006[46] Cohort	Gp1: 50 II/2M (Growth pattern: 14 horizontal; 11 neutral; 25 vertical) Mean age at start: 9yr 9mo (25M; 25F) 31 treated with U/L FA Gp2: 50 II/1 (growth pattern: 17 horizontal; 15 neutral;18 vertical) Mean age start: 9yr 11mo (25M,25F) 32 treated with U/L FA Classified by SN-MeGo angle: <31: horizontal >33: vertical In between: neutral	No University (30 patients) and two Private Practices (70 patients) No drop-outs recorded	Angle distocclusion least one premolar width in first molar region; UI-SN $\leq$ 95 degrees	Not stated. URA for expansion 7 mo maximum; then activator median tx time: Gp1:with URA/ activator 3yrs Gp2: 3yrs3mo Not reported	Yes	OM: anterior tooth inclination and growth pattern effects assessed on cephs R: OB: mean decrease with tx 1.9mm; 2.9mm; 1.6mm (for horizontal, neutral, vertical growth patterns); significant in neutral and vertical subgroups. IIA: mean decrease with tx 5.8^0, 9.3^0, 5.2^0 (horizontal, neutral, vertical growth patterns)

Tadic and Woods 2007[47] Cohort	22 II/2M (12F: 10M) Mean age start: 14.5 SD 2.4yr (F:14.5 SD 2.5yr M:15.5 SD 2.8yr) 39 II/1M (18F, 21M) Mean age start: 13.3 SD 1.9yr (F: 13.0 SD 2.1yr M: 14.0 SD 1.5yr)	No Not specified; one experienced operator using consistent biomechanics No drop-outs recorded	Minimal crowding upper and lower arches Class II Mesiobuccal cusps of both upper 6's occluded at least 5mm anterior to the midbuccal grooves of the lower 6's II/2M: UI $\leq 18^0$ to N-A line II/1M: UI $\geq 18^0$ to N-A line	Extraction of upper premolars U/L FA Duration not stated Not stated	Yes	OM:UI and soft tissue lip changes on cephs R: UI to N-A line: S diff in mean changes for II/2M (~+8.00^0) and II/1M (~-3.00^0) and for male and female groups per malocclusion. UI distance to N-A Line: S diff for II/2M (+1.3mm) and II/1M (-1.88mm) groups. NLA change: wide variation among subjects in all groups; mean increase for II/2M:2.53^0 and II/1M 4.28^0; NS difference between groups. Mean decrease in upper and lower lip curve (II/2M: 0.19mm; II/1M:- 0.21mm) respectively; NS difference between groups.	

n/g, not given; Ceph, lateral cephalometric; II/2M, Class II division 2 malocclusion; yr, year; mo, month; OB, overbite; II/1M, Class II division 1 malocclusion; Tx, treatment; LI, lower incisor; NB, Nasion-B point line; UI, upper incisor; SN, Sella-Nasion line; IIA, interincisal angle; LAFH, lower anterior facial height; M, male; F, female; SD, standard deviation; Ml-NL, maxillary-mandibular plane angle; Ar-Go-Me, Articulare-Gonion-Menton; DP3U, epiphyseal union of distal phalanx of the third finger; U/L FA, upper and lower fixed appliances; dpt, dental panoramic tomogram; ANB, measure of antero-posterior skeletal pattern; HG, headgear; SNB, Sella-Nasion-B Point; SNPg, Sella-Nasion-Pogonion; PAR, Peer Assessment Rating index; SN-MeGo, sella-nasion / menton to gonion angle; URA, upper removable appliance; N-A, Nasion to A point line; NLA, naso-labial angle

Table V Retrospective studies of stability for II/2M

Authors / year *Study/design*	*No .of cases/ sex of subjects* ***Mean age (range)***	*Sample size calculation* *Setting/ Operator* ***Drop-out***	*Definition of* ***malocclusion***	*Extraction/ non-extraction* *Treatment /duration* *Retention type/duration* ***Mean time out of retention***	***Error study***	***Outcome measures (OM) / Results (R)***
Mills 1973[48] Case series	60 (21M; 39F) Age 11.7 SD 2.4 yrs Age 17.9 SD 2.5 yrs Compared* with 9 year (13M; 10F) and 14 year (9M; 15F) control of London schoolchildren	No University N/R	UI ≤100 to maxillary plane OB ≥3mm UI in contact on their lingual side with lowers i.e. no overjet increase	Not stated Not stated Time out of retention: 3.13 SD 1.9 yrs	Not stated	OM: factors influencing successful treatment outcome R: Based on cases with ≥7^0 mean proclination (28UI and 30 LI cases, many cases of both) Mean UI proclination during tx (~13^0) of which mean 8° maintained after retention, relapse of ~37%; Mean LI proclination during tx of ~ 12^0 of which 10^0 maintained after retention, relapse of ~17% Large SD for both angulation changes; relapse usually associated with a return of incisor crowding Successful OB reduction correlated with reduction in IIA (~0.7 CC), relative lowering of lower lip line (~0.5 CC) and growth of lower face (~0.5 CC). Proclination LI generally more successful than UI.
Binda et al 1994[49] Case series	81 (42M; 39F) Mean age pre-tx 13.6M; 13.2F Post-retention (as above) 2yrs post-retention (33M; 32F) Mean age 19.4M; 19.5F 5 yrs post-retention (24M; 20F) Mean age 23.8M; 21.5F)	No University N/R	Disto-occlusion and retroclination of two or more UI	Not stated EOT or Functional and U/L FA (Edgewise) or sectional edgewise therapy Duration: not stated Type and duration: not stated Records at 2 and 5 years post-retention	Yes	OM: Post-retention changes in skeletal, dental, soft tissue variables R: OB: mean decrease 3.9mm (pre-tx to end of retention), mean increase 0.8mm (2 yrs post-retention) and 1.2mm (5 yrs post retention) IIA: mean decrease ~16^0 (pre-tx to end of retention), mean relapse 3.6^0 (2 yrs post-retention) and mean relapse ~ 6^0 (5 yrs post-retention) ALFH: mean increase 4.7mm (pre-tx to end of retention), mean increase 0.6mm (2yrs post-retention) and1mm (5 yrs post-retention). UI to SN: mean increase 10.6^0 (pre-tx to end of retention); mean decrease 0.9^0 (2 yrs post-retention) and 1.6^0 (5 yrs post-retention) LI to Mand plane: mean increase 5.8^0(pre-tx to end of retention); mean decrease 2.1^0 (2 yrs post-retention) and 2.6^0 (5 yrs post-retention) Skeletal variables changes: M> F; younger> older; anterior growth rotation indicated post-retention Lower lip cover of UI: mean decrease 1.2mm from start of tx to post-retention Post-retention: OB and IIA increased significantly Horizontal and vertical growth of soft tissues during tx but un-remarkable after tx except nose and chin became more prominent.

Canut and Arias 1999[50] Case series	30 (20F; 10M) Mean age pre-treatment 12 SD 2.8 yrs Mean age post-retention 22.2 SD3.6 yrs	No Not given N/R	Disto-occlusion and retroclination of ≥2 upper incisors	Non-extraction Routine edgewise 3 SD 1.2 years Type not given Mean duration 3 SD 1.2 yrs Minimum 3 years out of retention Mean period out of retention 7 SD 2.8 yrs	Yes	OM: long-term changes in occlusion, alignment and arch dimensions measured on study casts R: molar relationship stable: mean relapse 0.6 post tx OB: mean decrease 3.5mm during tx; mean increase 0.96mm post-tx; (~26% relapse) ; over-correction of OB relapsed Mean maxillary 3-3 expansion with tx: ~2.4mm; no decrease post-retention Mean maxillary intermolar increase with tx: 2.3mm; mean decrease 0.8mm post retention Anterior maxillary crowding: mean decrease with tx ~4mm; mean increase 0.8mm post retention 10% of maxillary arch unacceptable (Irregularity Index >4.5mm) anterior irregularities post-retention Mean mandibular 3-3 expansion with tx: 1.3mm; mean 1.1 decrease post retention Mean mandibular intermolar increase with tx: 1.2mm; mean decrease 0.6mm post retention. Increase in lower 3-3 width > 1mm and arch length always relapsed with associated crowding (30% unacceptable post-retention; Irregularity Index >2.5mm) OB overcorrection group, no net improvement long-term. Lower arch pre-tx crowding related to post-tx crowding. Number of years post-retention, OB relapse and lower post-retention crowding linked
Kim and Little 1999[51] Case-control	62 (31M; 31F) Mean age at start of tx 12.7yr (SD 2.6); at the end of tx 15.7yr (SD 2.4); long-term post-retention 30.9yr (SD 5) GP1: 33 OB≥ 4mm post-retention (mean 5.17, SD 0.87mm) GP2: 29 OB <4mm post-retention (mean 2.95, SD 0.87mm) cases selected at end of tx with initial deep OB and successful orthodontic tx judged clinically at end tx	No University N/R	II/ 2M, OB ≥4mm	29 non-extraction 23 extraction first premolars 10 extraction other than first premolars U/L FA Edgewise technique Duration: not stated Type: not given Duration: not given Mean time out of retention 15.2 years SD 4.5 yrs	Yes	OM: long-term stability of deep OB correction assessed on cephs and dental casts; predictors of post-retention OB R: mean OB decrease during tx 3.5mm ; mean relapse 1.4mm (40%) LAFH: mean increase during tx 4.1mm; mean relapse 1.4mm (~40%) IIA: mean decrease with tx 12.3^0; mean relapse 7.2^0 (~59%); Molar relationship: mean change with tx 3.1mm; mean relapse 0.1mm LII: mean decrease with tx 4mm; mean relapse 2.6mm (65%) UII: mean decrease with tx 7.7mm; mean relapse 1.2mm (~16%) Long term OB changes very variable; chance of maintaining OB < 4mm was 50% NS interaction between sex of patient and relapse. Tooth extraction does not seem to cause increase in post-tx OB. Initial OB best predictor of post-tx OB (R^2 = 0.42) but other factors could be involved.

Lapatki et al 2004[52] Case series	40 (8M, 22F) Mean age pre-tx 10.0 yrs (6.8 to 47.9 yrs); tx started in mixed dentition in 29 of 40 patients	No University clinic N/R	UI <98 degrees to anterior cranial base; OB ≤4mm	Non-extraction (26) U4's (11); U5 (1); U6 (1) U2 (1) Active tx started in mixed dentition (29) 6 URA solely; 28 removable and fixed; 6 fixed appliance only Duration: not stated Type not given Duration:\not given Median post-tx period: 2 yrs (1.1 to 5 years)	Yes	OM: relapse factors especially significance of a high lower lip line in comparison with other potential relapse factors assessed on cephs R: Only given for the 29 patients who started treatment in the mixed dentition.[1] median UI to SN change with treatment: 5.8^0; median change post-treatment $-.8^0$; relapse ~14%. OB: median change with treatment -2mm; median change post-treatment 0.5mm; relapse 25% Substantial inter-individual variation in UI and OB changes. Increased relapse tendency associated with: maxillary extractions, pronounced treatment induced UI inclination change, high post-tx lower lip line, poor compliance in retention phase
Lapatki et al 2006[53] Case series	31 (14M; 17F) Mean age pre-treatment 10.6 (7.0-33.9 yrs) (27 ≤ 18 yrs)[1]	No University clinic From 113 former patients 43 were contacted and 31 agreed to participate	Pre-tx retroclined maxillary incisors (<98 degrees to anterior cranial base), OB ≥ 3mm;	Non extraction (25) Extraction (6) Not stated Removable plates (13) ; Removable plates followed by vacuum formed stents (5); positioners or activators (5); bonded wire retainers (1) 7 pts discontinued retention shortly after tx was finished. Average retention period median) 1.3 yrs (0-15.2 yrs) Post-tx follow up median 9 yrs (3.4 to 15.2 yrs)	Yes	OM: impact of persisting high lip line and other potential relapse – inducing factors on long-term stability of orthodontic correction of retroclined UI assessed on dental casts and cephs. R: Only given for the 27 patients who started treatment in the mixed dentition.[1] median decrease in OB with treatment 1mm and median relapse 0mm median change of 7^0 in UIA to occlusal plane during tx; mean relapse 0.5^0 (~7%) Large inter-individual variability in post-tx UIA change (-5.25° to +6.75°). Increased tendency to relapse (a) where high post-tx lip line level is combined with UI and lower lip contact only in incisal crown area (b) marked tx induced UI inclination changes

Devresse et al 2007[54] Case series	61 (31M; 30F) Mean age: 13.4 SD 4.4 yrs	No University N/R	Class II molar relationship and obvious retroclination of both UI (central); competent lips on ceph	Not stated U/L FA Many tx 2-phase (removable/FA) Tx plan varied by malocclusion Mean treatment time 3.5 SD 1 yr Hawley retainer (18M; 13F) Upper lingual retainer (13M; 17F) Duration: not given Mean 3.5 years post-tx	Yes	OM: change in incisor inclination following tx; long-term stability after retention assessed on dental casts and cephs R: Mean change of 15.2^0 in UI inclination during tx; mean relapse of 2.2 ° (~14%) independent of type of retention appliance More UI proclined and torqued during tx, more relapse afterwards. Mean decrease in lip line height of 0.6mm at end of tx (statistically significant but deemed not clinically significant).
Bock and Ruf 2008[55] Cohort	37 (18M; 19F) 3 growth periods: Early adolescent: 7F 12.1-14.4yr; 3M 11.3-13.2yr Late adolescent: 6F 12.2-15 yr; 8M 14.1-16.4 yr Adult: 6F 16.8-36.5; 7M 16.3-25.6	No University N/R	Class II molar (>0.5mm) Bilaterally or >1 cusp width unilaterally; ob>3mm; retroclined UI (centrals)	Non-extraction 2 phase tx: Herbst/ Tip-Edge/multibracket Average tx duration: each phase 7.5 mo; 11 mo Activator(10), Upper Hawley(17) and fixed lower 3-3 lower fixed(27) Activator (5), U and L Hawley (1), positioner (3) Upper 3-3 / lower 3-3 fixed (1) Average retention time 27 SD 13.3mo	2 authors performed evaluations and mutually agreed	OM: post-tx occlusal changes on dental casts in early adolescent, late adolescent and adult subjects treated with Herbst (assessment of tx growth period using handwrist radiographs and method of Hagg and Taranger)[56] R: After 27 mo retention (average) and only those given for early and late adolescents. OB: early adolescent: mean decrease with tx ~3.6mm and increase post-tx ~0.8mm.(~22% relapse) Late adolescent: mean decrease with tx 4.7mm and increase post-tx 0.9mm (~19 % relapse) early and late adolescent combined: mean decrease with tx ~4.2mm and increase post-tx ~0.9mm.(~21% relapse)[2] Molar relationship: early adolescent 5%; late adolescent 7% combined. ~12% relapse.[2]

M, male; F, female; SD, standard deviation; *comparison only for aetiology but not for treatment; N/R, not relevant; UI, upper incisor; OB, overbite; yr, year; LI, lower incisors; CC, correlation coefficient; tx, treatment; EOT, extra-oral traction (headgear); U/L FA, upper and lower fixed appliances; IIA: inter-incisal angle; Ceph, lateral cephalometric radiograph; LAFH, lower anterior facial height; SN, Sella-Nasion,; Mand plane, mandibular plane; LII, lower irregularity index; UII, upper irregularity index; NS, non-significant; R^2, regression co-efficient; URA, upper removable appliance; [1], Results kindly supplied by Prof. Lapatki; UIA, upper incisor angulation; mo, months; [2], Results kindly supplied by Dr. Bock and Prof. Ruf.

Table VI (a) Quality assessment and (b) indication of risk of bias of studies of treatment (prospective case-control, prospective cohort, prospective case series, retrospective cohort) and stability (retrospective case-control, retrospective cohort) for II/2M

(a)

Quality assessment factors	*Moss 1975*[34]	*Demisch et al 1992*[35]	*Thuer et al 1992*[37]	*Parker 1995*[40]	*Kim and Little 1999*[51]	*Stellzig et al 1999*[22]	*Honn 2006*[46]	*Tadic and Woods 2007*[47]	*Woods 2008*[38]	*Bock and Ruf 2008*[55]
Sample size reported	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Sample size based on power calculation	No	No	No	No	No	No	No	No	No	No
Eligibility criteria described	No	Yes	Yes	Partial	Yes	Yes	Yes	Yes	No	Yes
Random allocation to groups	No	No	No	No	No	No	No	No	Unclear	No
Treatment allocation concealed	No	No	No	No	No	No	No	No	Unclear	No
Baseline equipoise between groups	Unclear	N/R	N/R	No	Age not gender	No	No	No	Yes	N/A
Blinding of treating clinician to treatment allocation	No	No	No	no	No	No	No	No	No	No
Blinding of patients to treatment allocation	No	No	No	No	No	No	No	No	No	No
Outcome assessors blinded to treatment allocation	No	No	No	No	No	Unclear	No	No	Unclear	No
Point estimates and measure of variability presented for primary outcome measures	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Appropriate statistical methods used to compare groups	Unclear	N/A	N/A	Unclear	Yes	Unclear	Unclear	Yes	Unclear	Unclear
Intention to treat analysis used	No	No	No	No	No	N/A	No	No	Unclear	No
Selective reporting	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Unclear	Yes

Key: N/R= not relevant; N/A= not applicable

(b)

Trial	*Adequate sequence generation*	*Allocation concealment*	*Blinding of participants*	*Incomplete outcome data addressed*	*Blinding of outcome assessors*	*Free of selective reporting*	*Free of other bias*	*Risk of bias*
Moss **1975** [34]	**No**	**No**	**No**	**No**	**No**	**No**	**No**	**High**
Demisch et al 1992[35]	No	No	No	No	No	No	No	High
Thuer et al 1992[37]	No	No	No	No	No	No	No	High
Parker 1995[40]	No	No	No	No	No	No	No	High
Kim and Little 1999[51]	No	No	No	No	No	No	No	High
Stellzig et al 1999[22]	No	No	No	No	Unclear	No	No	High
Honn 2006[46]	No	No	No	No	No	No	No	High
Tadic and Woods 2007[47]	No	No	No	No	Unclear	No	No	High
Woods 2008[38]	Unclear	Unclear	No	No	Unclear	Unclear	No	High
Bock and Ruf 2008[55]	No	No	No	No	For skeletal maturity	No	No	High