Orthodontic treatment for crowded teeth in children

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctorate of Dental Science by

Fyeza Janjua Sharif

January 2017

Contents

Orthodontic treatment for crowded teeth in children 1
1.ABSTRACT 5
1.INTRODUCTION 6
2. LITERATURE REVIEW 7
2.1 The Need for Systematic Reviews 7
2.2 Cochrane Reviews 9
2.3 Crowding 11
2.3.1 Definition of crowding 11
2.3.2 Incidence and prevalence of crowding 12
2.3.3 Factors which influence crowding 14
2.4 The impact of crowding on patients 17
2.5 Options for addressing crowding 17
2.5.1 Prevention 17
2.5.2 Treatment 20
3. AIMS, OBJECTIVES AND NULL HYPOTHESIS 23
3.1: Aims 23
3.2: Objectives 23
3.3: Null hypothesis 23
4. METHODS 24
4.1 Criteria for considering studies for this review 24
4.1.1 Types of studies 24
4.1.2 Types of participants 24
4.1.3 Types of interventions 24
4.1.4 Types of outcome measures 25
4.2 Search methods for identification of studies 26
4.2.1 Electronic searching 26
4.2.2 Databases searched 26
4.2.3 Searching other resources 26
4.2.4 Language 27
4.2.5 Unpublished studies 28
4.3 Data collection and analysis 28
4.3.1 Management of records produced by the searches 28
4.3.2 Selection of studies 28

4.3.3. Data extraction and management 29
4.3.4 Assessment of risk of bias in included studies 29
4.3.5 Measure of treatment effect 34
4.3.6 Unit of analysis issues 35
4.3.7 Dealing with missing data 35
4.3.8 Assessment of heterogeneity 35
4.3.9 Assessment of reporting bias 35
4.3.10 Data synthesis 36
4.3.11 Subgroup analysis and investigation of heterogeneity 36
4.3.12 Sensitivity analysis 36
4.3.13 Cross-over trials 36
5: RESULTS 37
5.1: Description of studies 37
5.1.1: Results of the search 37
5.1.2: Included studies 39
5.1.3: Characteristics of the trial designs and settings 39
5.1.4: Characteristics of the participants 40
5.1.6: Characteristics of the interventions and comparisons 41
5.1.7: Characteristics of the outcomes 42
5.1.8: Excluded Studies 45
5.2: Risk of bias in included studies 45
5.2.1: Allocation 45
5.2.2: Blinding 46
5.2.3: Incomplete outcome data 47
5.2.4: Selective reporting 47
5.2.5: Other potential sources of bias 47
5.2.6: Overall risk of bias 48
5.3: Effects of interventions 51
5.3.1: Comparison 1: Cervical pull headgear versus control 51
5.3.2: Comparison 2: Extraction of lower deciduous canines versus control 64
5.3.3: Comparison 3: Schwarz versus control 71
5.3.4: Comparison 4: Eruption Guidance appliance versus control 81
5.3.5: Comparison: Lower Lingual Arch versus control 89
5.3.6: Comparison: Lower Lip Bumper versus control 95

5.3.7: Comparison: Self-ligating brackets versus conventional brackets 102
5.3.8: Comparison: Active versus passive self-ligating brackets 112
5.3.9: Comparison: Copper nickel-titanium versus nickel-titanium archwires 116
5.3.10: Comparison: Coaxial nickel-titanium versus nickel-titanium 120
5.3.11: Comparison: Nitinol versus Titinol 124
5.3.12: Comparison: Nickel-titanium versus stainless steel 128
5.3.13: Comparison: Nickel-titanium versus multi-stranded stainless steel 132
5.3.14: Comparison: Multi-stranded stainless steel versus stainless steel 136
5.3.15: Comparison: Vibrational appliances versus control 140
6. DISCUSSION 144
6.1: Summary of main results 144
6.2: Potential biases and limitations of the review 151
6.3: Agreements and disagreements with other studies or reviews 151
7: AUTHORS CONCLUSIONS 153
7.1: Implications for practice 153
7.2: Implications for research 153
8: ACKNOWLEDGEMENTS 155
REFERENCES 156
APPENDIX 1: Characteristics of included studies 169
APPENDIX 2: Characteristics of excluded studies 229
APPENDIX 3: Search strategies 232
APPENDIX 4: Title and abstract screening form 238
APPENDIX 5: Eligibility form 239
APPENDIX 6: Data extraction form 241

1.ABSTRACT

Orthodontic treatment for crowded teeth in children

Fyeza N Janjua Sharif

Background: Crowded teeth develop when there is not enough space in the jaws into which the teeth can erupt. This gives the patient crooked or wonky teeth and cause them to present for treatment. Crowding can affect the baby teeth (primary dentition) and/or the adult teeth (permanent dentition). Crowding tends to increase with age, especially in the lower jaw, so that only a third of adults have well aligned lower front teeth (incisors)[1]. Crowding of the adult teeth can also occur when space is lost following the early loss of baby teeth either as a result of tooth decay or trauma. The Cochrane Oral Health Group undertook an extensive prioritisation exercise in 2014 to identify a core portfolio of titles that were the most clinically important ones to maintain on the Cochrane Library and this review was identified as a priority title by the orthodontic expert panel.
Objectives: To test the null hypotheses that there are no differences in outcomes between 1) the age at which orthodontic treatment for crowded teeth is carried out; 2) different orthodontic interventions for correcting/preventing crowded teeth against the alternative hypothesis that there are.
Search methods: We searched the following databases were searched up to July 2016: Cochrane Oral Health Group's Trials Register, Cochrane Central Register of Controlled Trials (CENTRAL) (*The Cochrane Library*), MEDLINE, EMBASE. No restrictions were placed on language or date of publication when searching databases.
Selection criteria: We included randomized controlled trials (RCTs) on any active interventions such as orthodontic braces (removable, fixed, functional), head-braces or extractions, against controls of no treatment, delayed treatment or another active intervention. The studies included had at least 80% of participants aged 16 years old and under.
Data collection and analysis: Two reviewers independently extracted information regarding methods, participants, interventions, outcomes, harms and results, independently and in duplicate. The Cochrane risk of bias tool was used to assess the methodological quality of the studies.
Main results: 19 RCTs were identified which included 1,101 participants. A meta-analysis was carried out on four papers, two that compared copper NiTi to NiTi and two that compared vibrational appliances to controls. No difference was found between either type of intervention. Additionally, subgroup analyses were carried out on thirteen other comparisons that revealed: there is low level evidence that Lower Lingual Arches and lip bumpers maintain space and prevent crowding, that the Schwarz appliance reduces crowding in the lower arch, that coaxial NiTi is better at treating crowding in the lower arch than single stranded NiTi and that self-ligating brackets are over a minute and a half quicker to untie and ligate than conventional brackets. No other statistically significant outcomes were found that were clinically significant, in any of the other comparisons. There were an insufficient number of studies to allow analysis between different age groups.
Authors' conclusions: 1) There is currently insufficient evidence to allow analysis and comparisons between different age groups. 2) There are three interventions that are effective at preventing crowding in the early dentition that are the Lower Lingual Arch, lip bumper and Schwarz appliance. There is some evidence to suggest that Coaxial NiTi is better at aligning teeth and reducing crowding than single-stranded NiTi. An additional outcome of clinical interest was that self-ligating brackets are quicker to untie and ligate than conventional brackets, saving clinical time. Further high-quality evidence is needed.

1.INTRODUCTION

Orthodontics is the branch of dentistry concerned with the growth of the jaws and face, the development of the teeth and the way the teeth and jaws bite together. It also involves treatment of the teeth and jaws when they are irregular and/or bite in an abnormal way. There are many reasons why the teeth may not bite together correctly. These include the position of the teeth, jaws, lips, tongue, and/or cheeks or may be due to a habit or the way people breathe.

Crowded teeth can develop in the primary dentition or the adult dentition and can affect how attractive an individual is perceived to be.[2] In fact, dental crowding has been found to significantly affect oral health-related quality of life[3] and self-perception[4] and so it is reasoned that by treating patients for crowding, we may create psychologically healthy adults with improved body images[5].

There are many different orthodontic interventions that can be utilised for the prevention and treatment of crowding and the question arises as to which intervention or interventions are the most effective. The purpose of this review was to address this question and assess the current body of available evidence.

2. LITERATURE REVIEW

2.1 The Need for Systematic Reviews

Orthodontics is a component of clinical dentistry which is centuries old with an interest in the anomalies of the face and teeth having been recorded by both physicians and philosophers, such as Hippocrates and Aristotle [6] as early as around 300 to 500 B.C. By 1728, Pierre Fauchard had written 'The Surgeon Dentist' containing notes on basic anatomy and more detailed notes on misaligned teeth.[6] He also described the first orthodontic appliance, known as a bandolet [6] which was used to expand the arches in order to allow tooth alignment. Modern-day orthodontic treatment typically involves the use of intra-oral or extra-oral appliances with extractions or enamel reduction as an adjunct to treatment, in order to achieve the desired aims set by both the patient and clinician. The General Dental Council has defined Orthodontics [7] as:

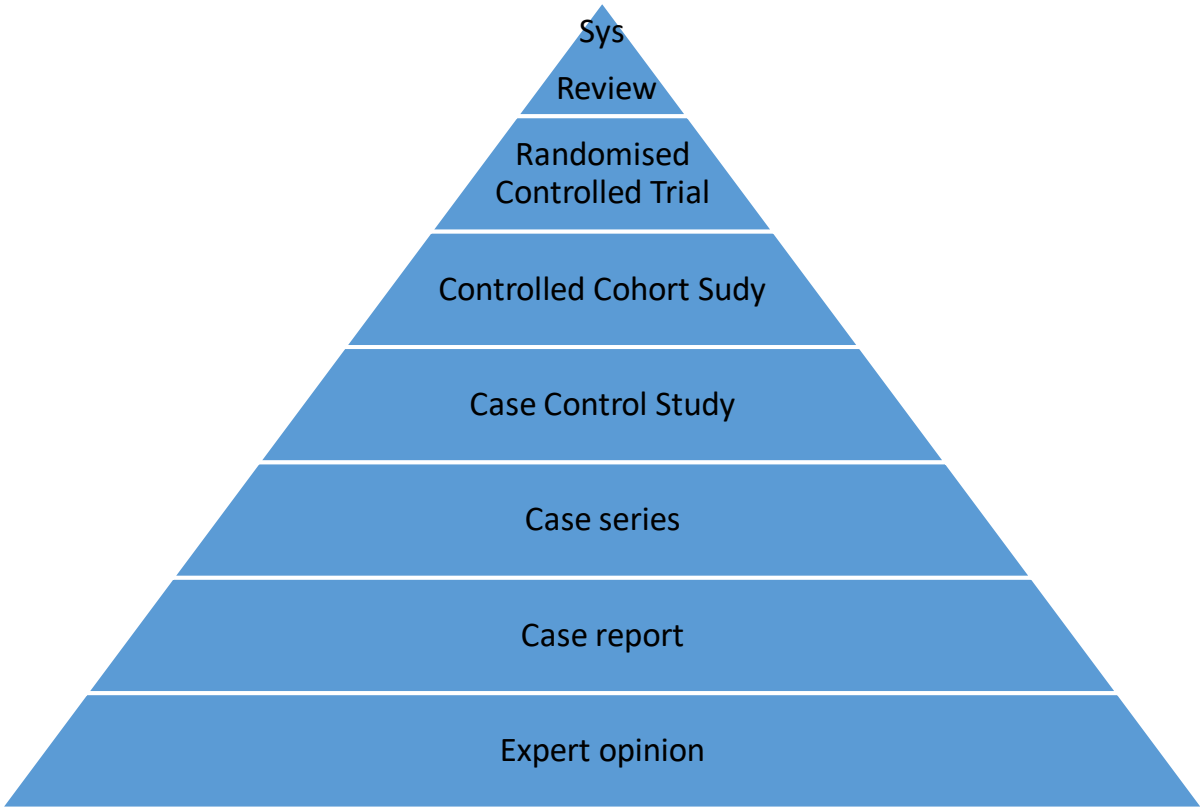
"The development, prevention, and correction of irregularities of the teeth, bite and jaw."

As with any branch of dentistry, the treatment provided should be evidence-based in order for the patient to receive the most appropriate care. Evidence based medicine has been described as:

"…the process of systematically reviewing, appraising and using clinical research findings to aid the delivery of optimum clinical care to patients." [8]

Whilst there may be many studies available on any one topic, being able to determine the quality of a study and of the evidence contained within it, is a skill that must be learnt. The hierarchy of evidence distinguishes between the many different study designs that are used to investigate interventions and is illustrated below[9,10]:

Figure 1: Hierarchy of evidence



Randomised controlled trials, if carried out to a high standard, are accepted as the gold standard for investigating interventions as they theoretically balance for known and unknown confounders through the process of randomisation. Systematic reviews of randomised controlled trials are at the peak of the hierarchy when investigating interventions (other methods are more suitable for other questions), as secondary research consisting of diligently, critically appraising all available data provides an overarching summary view of a topic.[11] The results can be combined statistically in a meta-analysis if the data are quantitative and in similar units. This is particularly useful in healthcare where the Department of Health and commissioners are constantly looking for treatments with a high level of efficacy and cost effectiveness. The Department of Health's executive agency, Public Health England, state that one of their responsibilities is "researching, collecting and analysing data to

improve our understanding of health and come up with answers to public health problems"[12].

In order to find the best available evidence on any intervention, a search strategy is constructed and run in appropriate databases. The results yielded may not be exhaustive, as there tends to be a bias towards publishing studies with significant results and to studies published in English[13], therefore supplemental searches that are not limited to English may also need to be performed in other databases. One such database for clinical trials is the Cochrane Controlled Trials Register, which is considered to be the single best source of published trials[13]. The number of studies and research on any one topic can be staggering; there are currently 231,756 clinical trials registered on the PubMed website [14] alone to date. For this reason, a method of condensing the information [15] from these search results into a manageable and coherent form is required.

Systematic reviews, if carried out to a high standard, provide answers to a specific research question by combining data from many studies. [16] Each well-conducted study builds upon previous work on a topic and expands upon it,[15] aiding the direction of future research and providing evidence-based treatments for patients[17].

2.2 Cochrane Reviews

The Cochrane Collaboration was founded in 1993 and consists of 53 groups that are responsible for carrying out and maintaining reviews in different areas of health care[18]. Their main aim is "to help people make well-informed decisions about health care by preparing, maintaining and promoting the accessibility of systematic reviews of the evidence that underpins them"[18]. Whilst there are many systematic reviews available, they may vary greatly depending upon the types of studies included and upon the methods used to assess the studies for quality and risk of bias. Cochrane reviews differ in this respect in that the methods used to assess studies is vigorous and standardised. [19] This allows for a consistent review on any topic, in the same format for ease of reading [15] and the reassurance that the review will have been checked and stages of it carried out in duplicate.

Initially, a protocol is submitted in a standardised format which includes a short background to the problem, the objectives and then the methods which include the types of studies and participants included, as well as the search strategy, then the results and analysis. [20]

Following on from this, the steps involved in carrying out a Cochrane review are[18]:

1. Defining the review question and developing criteria for including studies
2. Searching for studies
3. Selecting studies and collecting data
4. Assessing risk of bias in included studies
5. Analysing data and undertaking meta-analyses
6. Addressing reporting biases
7. Presenting results and 'Summary of findings' tables
8. Interpreting results and drawing conclusions

The review question is ordinarily set in the PICO format[21], which is to consider the Population, Interventions, Comparisons and Outcomes of interest. This format helps to focus the review and set the eligibility criteria for the studies to be included. The outcomes are not usually included as part of the eligibility criteria, as all studies on a particular topic would be included and all outcomes of the intervention and comparison would be evaluated. A search strategy is then developed with the Trial Co-Ordinator who searches databases such as EMBASE, MEDLINE and The Cochrane Central Register of Controlled Trials (CENTRAL), as well as journals and unpublished studies. The results of the searches yields a list of studies, which are screened using their titles and abstracts with specific criteria to see if they are relevant to the topic and research question. They are then assessed for eligibility using especially designed forms to ensure that they fulfil the inclusion criteria. Data are then extracted from these studies and they are assessed for bias using the Cochrane Collaboration's Risk of Bias tool[18]. These assess the methodological quality of the included studies and classifies them as 'low', 'high' or 'unclear' risk of bias both in a tabular and graphical format, in order to allow comparison between the included studies at a glance. The data are then analysed and if appropriate, a meta-

analysis is carried out combining outcomes to increase their power and precision. This can then be used to determine if there are any benefits or harms to a particular intervention, the strength, direction and magnitude of their effect.

2.3 Crowding

2.3.1 Definition of crowding

The British Standards Institute have described dento-alveolar disproportion and crowding as [22]:

"A disproportion between the size of the teeth and the space available in the arch for them."

In untreated malocclusions, such discrepancies can be a contributory factor in crowding. There are numerous methods that have been described in order to predict whether or not there is enough space for the permanent dentition, such as the Tanaka and Johnson Analysis [23] or the Hixon and Oldfather Analysis.[24] These analyses measure the amount of space available, from study models, and then use either radiographs of the unerupted teeth or regression formulae in order to estimate the size of the unerupted permanent canines and premolars. In this way, crowding can be predicted.

Additionally, crowding can also be defined in terms of severity as mild, moderate or severe with regards to the discrepancy between the size of the teeth and the space available in the arch. It is measured in millimetres, with mild being 0.1 to 4mm, moderate being 5mm to 9mm and severe being 10mm or greater [25]. There are various methods for measuring crowding mentioned in the literature, one of which is in the Royal London Space Analysis [26,27] which suggests that crowding is relative to the archform that is reflected by the positioning of the majority of the teeth[26]. The method advocated for measurement is the placement of a clear ruler over the occlusal or labial surface of study models in order to measure the mesiodistal width of misaligned teeth, not measuring more than 2 teeth at a time to allow measurements of chords in the arc, and the space available in the selected archform[26]. This analysis also considers vertical crowding and the space required to level the curve of Spee[27]. Other methods such as the use of callipers to measure

individual tooth widths and the brass wire technique are alternatives for crowding assessment and space analysis, but have been found to be less reliable[28].

Crowding can be assessed for the entire arch, or just for the anterior teeth, depending upon the requirements of each case and what the aims of treatment are. In 1975, Little [29] described one such index, which assessed the amount of contact point displacement of the mandibular anterior teeth, defining them into the following categories[29]:

- 0 - Perfect alignment
- 1-3 - Minimal irregularity
- 4-6 - Moderate irregularity
- 7-9 - Severe irregularity
- 10 - Very severe irregularity

However, a disadvantage of this method is that it can grossly overestimate the amount of crowding in simple cases where one tooth may be rotated, as this would register as a large contact point displacement.

2.3.2 Incidence and prevalence of crowding

In 2013, the Child Dental Health (CDH) Survey[30] found that 36% of 12 year olds and 20% of 15 year olds in England had unmet orthodontic needs having scored an Index of Orthodontic Treatment Need (IOTN) [31] Dental Health Component of 4 or above. Whilst this is an assessment of formative need, the level of expressed need and demand [32] was greater: 45% of 12 year olds wished for their teeth to be straightened, with this percentage dropping in 15 year olds as many may already have had orthodontic treatment at this point.

Values for incidences and prevalence of crowding are not readily available for national rates, perhaps due to the different methods and indices of measurement that are available. Whilst most crowding cases that qualify for National Health Service treatment would do so by scoring a 3d or 4d on the Index of Treatment Need dental health component[31], this is not strictly a reflection of crowding. These scores relate to greater than 2mm of contact point displacement, but less than or equal to

4mm for a 3d and greater than 4mms of contact point displacement for a 4d [31]. There are, however, some localised values and rates available, one of which is from the USA National Health and Nutrition Examination from 1988 to 1991[1]. Here, 7000 people were examined and it was found that the most common problem was incisor irregularity; 66% had mandibular incisor irregularity and 55% had maxillary incisor irregularity. In Dresden, Germany, between 1996 and 1997 a survey of 8768 school children from which a sample of 1975 children aged between 6 and 8 years old was used to estimate the prevalence of malocclusions using the IOTN[33]. It was found that 12% of teeth in the maxillary arch and 14.3% of teeth in the mandibular arch had a tooth width to arch length discrepancy[33]. In the UK, examinations of 924 and 996 schoolchildren in Manchester and Sheffield respectively in 1994 found that 26-28% had crowding as their worst occlusal trait. Only impacted teeth surpassed this feature of the malocclusion. [34]

Table 1: Prevalence of crowding

Study	Location	Date	N	Age	Outcome
Proffit et al1998	USA	1988-1991	7000	8-50	66% LI crowding 55% UI crowding
Tausche et al 2004	Germany (Dresden)	1996-1997	1975	6-8 years	12% upper arch crowding 14.3% lower arch crowding
Burden & Holmes 1994	UK (Manchester & Sheffield)	1994	1920	11-12	26-28% crowding

2.3.3 Factors which influence crowding

The aetiology of any malocclusion is multifactorial and consists of both genetic and environmental factors. The causative factors of dental crowding are discussed below.

2.3.3.1 Skeletal

Skeletal patterns and articulation of the jaws, as a result of mandibular length and face heights, have a substantial genetic element that has been demonstrated in monozygotic twins[35,36] and the overall contribution of genetics to malocclusion has been found to be about 40%. [37,38,39]

Skeletal patterns are considered in three planes of space: sagittal, vertical and transverse. Deviations from accepted norms can result in dento-alveolar compensation, a system that attempts to maintain normal inter-arch relationships under varying jaw relationships. [40] When considering sagittal discrepancies for example, in skeletal class III patients the upper incisors tend to be proclined and the lower incisors retroclined in order to try and achieve a positive or as close as possible to a positive overjet as possible, which can lead to crowding of the lower labial segment[41,42]. Sagittal jaw discrepancies can also directly affect the space available for teeth [25] and can therefore lead to crowding of teeth in order to use the available space.

Deviations in vertical skeletal growth can also have effects on the dentition and result in crowding. For example, one of the features of a class II division 2 malocclusion may be a forward mandibular growth rotation and reduction in vertical height, causing an increased curve of Spee[43]. This then results in vertical crowding, which requires space for alleviation during orthodontic treatment [26] or crowding of the lower incisors due to the increased over bite and retroclination of the lower incisors. [44] Conversely, in patients with increased vertical proportions, dental crowding may be present due to narrower arches and proportionally larger teeth. [44]

Transverse dimensional skeletal discrepancies also lead to crowding and impactions of teeth. In patients with a class 3 skeletal pattern, the aetiology can be due to maxillary hypoplasia, mandibular hyperplasia or a combination of both. In cases with

maxillary hypoplasia, crowding of the upper incisors and impaction or ectopic eruption of the upper canines is often seen due to tooth tissue to jaw size discrepancies, whereas in mandibular hyperplasia cases there may be retroclination and crowding of the lower incisors due to dentoalveolar compensation and the force of the lower lip musculature against the teeth.[44] In cases with a class 2 skeletal pattern, the aetiology can be due to maxillary hyperplasia, mandibular hypoplasia or a combination of both. Generally, crowding is seen in cases with mandibular hypoplasia cases, exhibiting as mandibular crowding or ectopic eruption due to a tooth size to jaw size discrepancy. [44]

2.3.3.2 Dental

There are many dental components that can contribute to the aetiology of crowding. They can be divided into malformed or supernumerary teeth, retention or early loss of deciduous teeth and ectopic eruption. Discrepancies in tooth size to jaw size have already been discussed, but each of the other factors will now be considered.

Malformed and supernumerary teeth

Disturbances during the development of the dentition can lead to malformations in the morphology, namely fusion or germination. Fusion of two teeth at dentine level, but with separate pulp chambers, results in a tooth with an increased mesio-distal width of the clinical crown and a reduced number of teeth in the associated arch. Gemination on the other hand occurs when a tooth bud has a common pulp chamber but has attempted to separate into 2 separate teeth, so that the total number of teeth within the arch remains the same. Reported prevalence of both these conditions varies in the literature, perhaps due to differences in detection as deciduous teeth are exfoliated or extracted early[45].

Supernumerary teeth form due to disturbances in initiation and proliferation of the tooth bud and have an incidence of 1 to 2% [46]. They can be conical or tuberculate in shape, or form complex or compound odontomes if they have abnormal morphology. If they resemble the crown morphology of adjacent teeth they can be considered as supplemental teeth. This type of supplemental tooth can lead to local crowding of the dentition if there is insufficient space for it to be accommodated.

Ectopic eruption

If a primary tooth is retained beyond its normal expectancy, eruption of the permanent successor can be delayed or can lead to ectopic eruption. In the U.S.A. 5-10% of children have at least one deciduous molar that suffers delayed exfoliation[25]. Conversely, premature loss of a deciduous tooth can lead to drift of the adjacent permanent teeth into the space combined with a more mesial path of eruption of the permanent tooth, causing either impaction or ectopic eruption which can lead to crowding. [25]

Ectopic eruption and resultant crowding can also be a consequence of a mal-positioned tooth bud. This most frequently occurs with the maxillary first permanent molar developing in a more mesial location and therefore taking the space of the second permanent premolar. [25]

2.3.3.3 Soft tissues

If 40% of the aetiology of malocclusion is inherited[37,38], the remaining 60% is due to environmental influences. The dentition is thought to lie in a zone of equilibrium between the periodontal tissues, soft tissues of the lips, tongue and cheeks[25]. The Equilibrium theory [25] suggests that although these soft tissues apply light forces to the dentition, they are in contact for a duration of over 6 hours which is enough to produce tooth movement. Masticatory forces, although much greater than those produced by the other mentioned soft tissues, are for much shorter duration and so do not seem to have the same effect on tooth position[25].

2.3.3.4 Habits

Digit-sucking or pacifier use is a normal stage of physiological development in infants [25] and will usually cease spontaneously, but if the habit persists at age 8 there is a statistically significant 11% increase in skeletal II base patterns [47] compared to non-digit-suckers. One of the other known effects of digit-sucking is retroclination of the lower incisors [48] which would result in contact point displacement and crowding as there is less space available in the arch.

2.4 The impact of crowding on patients

The socialisation theory puts forward the idea that the perceived attractiveness of an individual, by others, can cause differential treatment of them, as well as influence their development and interaction with others[49]. The appearance of teeth is a feature of facial attractiveness that is ranked as a priority by both males and females [50] and a malocclusion can affect an individual's self-esteem and their concept of body image into adulthood.[4,5] The prevalence of bullying in children, aged 10 to 14 years old, with an untreated malocclusion has been found to be 12.8%[51]; Guidelines produced by the British Psychological Society [52] estimate that 10% of children with malocclusions would have significant anxiety, emotional or behavioural problems. One of the characteristics identified as increasing the risk of bullying and affecting psychological development, by causing harm and distress to patients [53] is maxillary crowding[54]. Dental crowding has been found to affect oral health-related quality of life [3] and self-perception [2] significantly and so it is reasoned that by treating patients for crowding we may create psychologically healthy adults with improved body images[5]. In this vein, the American Academy of Pediatric Dentistry has produced 'Guidelines on the Management of the Developing Dentition and Occlusion in Pediatric Dentistry' [55] which advocate treatment for crowding as well as a list of objectives that includes interventions to prevent crowded incisors and decrease the ectopic eruption of canines.

2.5 Options for addressing crowding

2.5.1 Prevention

The prevention of crowding in the early or mixed dentition has been much debated in the literature with several methods having been described. Each of these will be discussed in turn.

2.5.1.1 Headgear

A number of studies have looked at the effect of cervical pull headgear on the early mixed dentition in order to relieve crowding, before considering a comprehensive course of treatment[56,57,58]. Whilst arch lengths and widths increased in the headgear groups and crowding reduced post treatment, long term the stability of early

treatment was not statistically different to those cases treated with extractions. The factors that were considered most important for long-term aesthetics were initial occlusion and upper anterior segment alignment, more so than lower incisor alignment.[57]

2.5.1.2 Extraction

The British Orthodontic Society states that "lack of spacing in the deciduous dentition may lead to crowding in the permanent dentition and so extraction of deciduous teeth that displace their permanent successors is advocated". [59] The term "driftodontics" is used to describe the extraction of teeth in order to allow some spontaneous improvement in the alignment of the permanent teeth. [59] They advise the extraction of all four deciduous canines in the following circumstances. [59]

1. To provide space so that a crowded but unerupted maxillary lateral incisor may erupt without being deflected into lingual occlusion. Once a positive overbite has been obtained, such teeth will not correct spontaneously even when space is made available. Early intervention is therefore crucial.

2. To provide space for crowded maxillary incisors, which are already in lingual occlusion, to be corrected in the early mixed dentition.

3. To provide space for severely crowded lower incisors to align spontaneously - if the crowding is more of a displacement than a rotation, and the lateral incisors are less than half erupted.

4. To ensure that incisors delayed by the presence of a supernumerary tooth have sufficient space to allow their full eruption.

5. To encourage a palatally ectopic maxillary canine to erupt.

Extraction of deciduous canines in order to alleviate lower incisor crowding has also been investigated [60,61] and found to be successful. However, long term there was a decrease in arch perimeter resulting in a reduced amount of space for the permanent canines[60], so space maintenance may be a consideration. The American Academy Of Pediatric Dentistry's Clinical Affairs Committee [55] also advise the consideration of

primary canine extractions in order to gain space for the permanent incisors to erupt through, as well as space maintenance.

2.5.1.3 Lingual arch

The use of a lingual arch in order to preserve arch perimeter has been examined in a systematic review [62] which included both prospective and retrospective studies. It identified 2 studies, one by Rebellato et al [63] and one by Villalobos et al.[64] Both studies found the use of lingual arches to be successful in preserving arch length as well as the vertical development of the mandibular molars, but at the expense of slight mandibular incisor proclination. The space that is maintained can be used to alleviate crowding or preserve it until the permanent teeth erupt, thus preventing or reducing crowding in the permanent dentition.

2.5.1.4 Lip bumper

The American Association of Pediatric Dentistry Clinical Affairs Committee advocates the use of lip bumpers to regain space in the mandibular arch.[55]; Lip bumpers are thought to work by alleviating the pressures on anterior teeth from the lower lip and cheeks. In a systematic review that included both prospective and retrospective studies[65], it was found that patients who had lip bumpers fitted had an increase in arch dimensions. This was accredited to incisor proclination as well as distal movement and tipping of the molars. Alongside these changes, there was an increase in arch width and intercanine widths. The lip bumper is therefore a treatment option that can both prevent and treat lower arch crowding.

2.5.1.5 Removable appliances

Several different designs of removable appliance have been used to alter arch dimensions in the early mixed dentition with the aim of preventing crowding in the permanent dentition. Appliances such as the Schwarz [66] and the Eruption Guidance Appliance [67] have shown statistically significant reductions in anterior crowding, in both the maxilla and mandible, compared to controls. Further studies to assess the

long-term stability of these outcomes are required as the follow up period was only one year and the patients were still in mixed dentition at this stage.

Disadvantages of removable appliances are that they require patient compliance and cannot produce complex tooth movements, as it is difficult to get two points of contact on the teeth. [25]

2.5.2 Treatment

Treatment of crowding normally occurs in the permanent dentition once prevention in the deciduous dentition has not been successful or not attempted. The methods that are ordinarily used are discussed below.

2.5.2.1 Distal movement and expansion

In cases where crowding is due to a tooth size to jaw size discrepancy, distal movement of the buccal segments or expansion of the arch/s can be considered as treatment options. Distal movement of the buccal dentition can be achieved with the use of headgear, [68] Ten Hoeve appliances, [69] repelling magnets [70], pendulum appliance, Jone's jig, [71] temporary anchorage devices (TADs) [72] or other skeletal anchorage devices. One recent review with meta-analysis found that the most effective method of distalising molars without anchorage loss was with the use of TADs, whereas other methods suffered up to 66% anchorage loss (measured as mesial movement of the premolars).[73] However, another (Cochrane) review found that headgear was the only method that was capable of producing distal movement without anchorage loss when compared to intra-oral methods, but it also produced the least distal movement. [72]

Arch expansion is limited by the amount of alveolar bone available into which the teeth can expand. As a result, the options depend on whether a patient is growing and how much expansion is required. For mild expansion of less than 5mm, expansion can be achieved by expanding the working archwire, with the use of a removable appliance with an expansion screw, expansion arches or a quad helix [74] whereas for more severe cases, rapid maxillary expansion (in growing patients) [75] and surgically assisted rapid maxillary expansion (in non-growing patients) [76] can be used. These methods are usually used to correct crossbites, but as a side product, could potentially be used to correct crowding.

2.5.2.2 Wires

Orthodontic wires can be made from different materials e.g. stainless steel, nickel titanium, titanium molybdenum alloy, to different dimensions e.g. 0.014" round, 0.016" x 0.016", 0.019" x 0.025" and can be used in several different arch wire sequence combinations during treatment which are mainly down to operator preference. Research suggests that there is no statistical difference between several of these sequences with regards to patient discomfort, root resorption, or time taken to reach the working arch wire. [77,78] Generally, the accepted convention is to use a small, round, nickel titanium (NiTi) arch wire in order to level and align and then build up to a stiffer stainless steel arch wire with rectangular dimensions for bodily tooth movements.

2.5.2.3 Brackets

Orthodontic brackets can be made from different materials e.g. stainless steel, ceramic and to different designs e.g. Siamese, mono-block, self-ligating. The combination of different materials and designs mean that there are many different types of brackets available and numerous claims made surrounding each type. Clinical trials suggest that there is no difference in the amount of time taken to alignment when conventional or self-ligating brackets are used, nor any difference in incisor inclination or transverse dimension changes. [79,80] In addition, no advantage to either with regards to pain reduction or treatment efficiency, whether metal or porcelain brackets are used, has been shown.[81,82] When comparing active and passive self-ligating brackets, there was once again no difference found between the two with reference to correction of anterior maxillary crowding. [83]

2.5.2.4 Aligners

Aligners are clear vacuum formed appliances that are used alone or in combination with clear attachments on the teeth, in a series of incrementally different shaped appliances, to align the teeth. A series of aligners are used to gradually move teeth in conjunction with bonded (normally composite) attachments for movement control.[25,84] The British Orthodontic Society states that aligners may be used in cases with mild to moderate dental irregularity where crowding can be corrected by

expansion or interproximal reduction, or in cases with mild spacing. [85]There is low-level evidence to suggest that aligners are effective at intruding teeth, controlling molar inclination and bodily movements up to 1.5mm, but are less effective at rotational movements. [86] It may be inferred that if they are able to produce small bodily movements then they may produce expansion that can then be used to alleviate mild to moderate crowding.

3. AIMS, OBJECTIVES AND NULL HYPOTHESIS

3.1: Aims

The aims of this review were to determine whether there is a difference in the outcomes:

1. between different orthodontic interventions for correcting/preventing crowding.
2. depending on the age at which orthodontic treatment for crowded teeth is carried out.

3.2: Objectives

The objectives of this review were to search and analyse the literature systematically surrounding the effectiveness of orthodontic interventions for crowding in children and to test the null hypothesis.

3.3: Null hypothesis

To test the null hypotheses that there are no differences in outcomes between:

1. Different orthodontic interventions for correcting/preventing crowded teeth;
2. The age at which orthodontic treatment for crowded teeth is carried out against the alternative hypothesis that there are.

4. METHODS

4.1 Criteria for considering studies for this review

4.1.1 Types of studies

All randomised controlled clinical trials (RCTs) of orthodontic treatments to correct or prevent crowding where one intervention was compared concurrently to a placebo or no intervention or another method of crowding correction/prevention.

4.1.2 Types of participants

Inclusion Criteria

Trials were eligible for inclusion in the review if they recruited at least 80% children and/or adolescents (age 16 years or less) who received orthodontic treatment to correct or prevent crowding.

Exclusion Criteria

Trials were excluded from the review if they recruited patients with a cleft lip and/or palate or other cranio-facial deformity/syndrome, as well as those trials that recruited less than 80% children or adolescents.

4.1.3 Types of interventions

Active interventions

The following active interventions to alleviate or prevent crowding were assessed:

- Orthodontic braces (removable, fixed, functional)
- Head braces
- Extractions
- Any adjunctive treatments

Any intervention or combination of treatments, at any time during treatment was evaluated.

Controls

Controls may consist of no treatment, delayed treatment, placebo or another active intervention.

4.1.4 Types of outcome measures

Primary Outcomes

The amount of crowding (measured in mm or by any index of malocclusion).

Secondary Outcomes

1. Size of the upper jaw (arch length),
2. Size of the lower jaw (arch length),
3. The relationship of the
    a. lower back teeth (molars) to the lower jaw (mandible);
    b. lower front teeth (incisors) to the lower jaw (mandible);
4. Self-esteem;
5. Patient satisfaction;
6. Jaw joint problems.
7. Harms were recorded and reported in descriptive terms and would have included:
    a. Health of the gums
    b. Damage to the teeth e.g. tooth decay.

    Outcomes were recorded at all ages reported. The results were reported according to the most common endpoints.

Comparisons

The following comparisons were included:

Table 2: Plan of comparisons of included studies

	Intervention Type	Intervention
Prevention	Head Gear vs. Interceptive treatment	Cervical head gear; several interceptive treatments e.g. XXXXXX
	Lip bumper	
	Lingual arch	
Treatment	Extraction of deciduous canines	
	Removable appliances	Swartz
		Guidance development
	Brackets	Convent / SLG
		Active / Passive SLG
	Arch wires	Stainless Steel / Multistranded Stainless Steel / Nickel Titanium (NiTi)
		Titanol / Nitanol

		NiTi/ Heat activated NiTi/Copper NiTi
		NiTi/Copper NiTi
		Coaxial NiTi / NiTi
	Vibrational appliance	

4.2 Search methods for identification of studies

4.2.1 Electronic searching

For the identification of studies included or considered for this review, a detailed search strategy was developed for each database searched by Cochrane's Oral Health Information Specialist. These were based on the search strategy developed for MEDLINE but revised appropriately for each database. The subject search strategy used a combination of controlled vocabulary and free text terms based on the search strategy for MEDLINE, in conjunction with phases of the Cochrane sensitive search strategy for RCTs: For the MEDLINE search, the subject search was run with the Cochrane Highly Sensitive Search Strategy (CHSSS) for identifying randomised trials in MEDLINE: sensitivity-maximizing version (2008 revision) as referenced in Chapter 6.4.11.1 and detailed in box 6.4.a. and b of the Cochrane Handbook for Systematic Reviews of Interventions version 5.1.0 (updated March 2011).

4.2.2 Databases searched

The following databases were searched:

- Cochrane Oral Health Group's Trials Register (searched 7 July 2016)
- Cochrane Central Register of Controlled Trials (CENTRAL; 2016, Issue 6) (*The Cochrane Library*, searched 7 July 2016)
- MEDLINE (from 1946 to 7 July 2016)
- EMBASE (from 1980 to 7 July 2016)

4.2.3 Searching other resources

A check was made to identify journals which had already been hand-searched as part of the Cochrane Journal Hand-searching Programme. These included:

- US National Institutes of Health Ongoing Trials Register ClinicalTrials.gov (clinicaltrials.gov; searched 7 July 2016) (Appendix 5);
- World Health Organization International Clinical Trials Registry Platform (apps.who.int/trialsearch; searched 7 July 2016) (Appendix 6).

The hand-searching of the following journals would have then been updated to the most current issue if appropriate:

- American Journal of Orthodontics and Dentofacial Orthopedics;
- The Angle Orthodontist;
- European Journal of Orthodontics;
- Journal of Orthodontics;
- Australian Orthodontic Journal;
- Seminars in Orthodontics;
- Orthodontics and Craniofacial Research;
- Clinical Orthodontics and Research;
- Journal of Orofacial Orthopaedics.

The bibliographies of the clinical trials identified were checked for references to trials published outside the hand-searched journals.

Personal references were checked.

Additionally, other resources such as The British Library EThOS service (http://ethos.bl.uk) were searched for relevant theses and ClinicalTrials.gov will be searched for otherwise unpublished and ongoing studies.

Conference proceeding of the European Orthodontic Congress, International Association of Dental Research, British Orthodontic Conference and American Association of Orthodontists were also searched to identify presented trials.

4.2.4 Language

Databases were searched to include papers and abstracts published in all languages and every effort was made to translate non-English papers.

4.2.5 Unpublished studies

The first named authors of all trial reports were contacted in an attempt to identify unpublished studies and to obtain any further information about the trials.

Trials databases were also searched to identify registered, ongoing trials.

4.3 Data collection and analysis

4.3.1 Management of records produced by the searches

All references were downloaded into Microsoft Excel to produce a single database to facilitate retrieval of relevant articles. Non-electronic references, that could not be downloaded, were entered into the database manually after which duplicates were removed.

4.3.2 Selection of studies

Two review authors (J Harrison and F Janjua Sharif (JH and FJS)) independently assessed the titles and abstracts (when available) of all reports that were identified as potentially relevant by the search both independently and in duplicate.

For studies with insufficient information in the title and abstract to make a clear decision to exclude, or studies where there was disagreement between the review authors about eligibility, a full report was obtained. These full reports were then assessed independently and in duplicate by two review authors (FJS and JH) to establish whether or not the studies met the inclusion criteria.

Disagreements were resolved by discussion between FJS and JH. A record of all decisions made about the potentially eligible studies was kept. Full reports were also obtained for those studies that were ultimately included in this review.

The review authors were not blinded to trial author(s), institution or site of publication.

4.3.3. Data extraction and management

Data extraction was carried out independently and in duplicate by two review authors (FJS and JH), recording year of publication, interventions assessed, outcomes, sample size and age of subjects, using a pre-designed and piloted data collection form and saved electronically. We contacted study authors for clarification on missing data where necessary and feasible and resolved any disagreements through discussion. We recorded the following key data for each included study in the Characteristics of Included Studies table:

- Trial design, setting, number of centres, source of participants, method of recruitment, recruitment period and study duration.
- Inclusion/exclusion criteria, age and gender and ethnicity of participants, and number selected and excluded.
- Description of the invention and comparison.
- Details of the outcomes reported, including method of assessment and time point(s) assessed.
- Details of sample size calculations, adverse effects, funding sources, declarations/conflicts of interest.

The primary outcome was the amount of crowding and the secondary outcomes were also recorded e.g. the relationship of the lower back teeth (molars) to the lower jaw (mandible); relationship of the lower front teeth (incisors) to the lower jaw (mandible); pain. Harmful outcomes studied were recorded for descriptive purposes e.g. health of the gums, damage to the teeth.

All outcome data were extracted. They were then grouped into the time-points which we felt were the most clinically relevant. If outcome data were reported at other time points, then consideration was given to examining those as well.

4.3.4 Assessment of risk of bias in included studies

The Risk of Bias was assessed independently and in duplicate by the two review authors (FJS and JH), using The Cochrane Risk of Bias tool, as described in Chapter 8 of the Cochrane Handbook for Systematic Reviews of Interventions. This

was utilised in order to assess the methodological quality of the included studies and looked at seven specific domains: random sequence generation; allocation concealment; blinding of participants and personnel, blinding of outcome assessment; incomplete outcome data; selective outcome reporting and 'other sources of bias.'

Each domain was assigned a judgement of high, low or unclear as an indication of its risk of bias according to the following criteria:

- Low risk of bias if plausible bias unlikely to seriously alter the results;
- High risk of bias if plausible bias that seriously weakens confidence in the results; or
- Unclear risk of bias if plausible bias that raises some doubt about the results.

Sequence generation, allocation concealment and selective outcome reporting were assessed for the study as a whole. Blinding and incomplete outcome data were assessed on the level of the study and for each outcome as appropriate.

Method of sequence generation

Adequate sequence generation

Using methods such as repeated coin-tossing, dice throwing, dealing previously shuffled cards/envelopes or more contemporary methods such as referring to a random numbers table or random assignments generated by a computer. Restricted or stratified randomisation can also be included.

Inadequate sequence generation

Using quasi-random methods such as alternation, or where assignment has been determined by date of birth, hospital or case number, date of presentation or clinician choice.

Unclear sequence generation

Where the allocation method is described as randomised but no further details are given as to the method used. In these cases, the authors were contacted for clarification.

Method of allocation concealment

Adequate concealment schemes

Central allocation (including telephone, web-based and pharmacy-controlled randomization), sequentially numbered drug containers or numbered opaque, sealed envelopes.

Inadequate concealment schemes

Using an open random allocation schedule, or enveloped that are unsealed or translucent, so that allocation of the next patient can be deciphered.

Unclear concealment schemes

Where the method of allocation concealment has not been described or is not described in sufficient detail in order to permit a fully informed judgement.

Blinding of participants, personnel and outcome assessors

Adequate blinding

Where blinding of participants and clinicians is ensured throughout the study, or where non-blinding is unlikely to have introduced any bias e.g. where strict treatment protocols are introduced to reduce any difference in management of the groups involved.

Inadequate blinding

Where there has not been any blinding or there has been incomplete blinding of the participants or personnel and so the outcomes are likely to be affected. This could also be in cases where blinding was broken therefore influencing the results.

Unclear

Cases where the method of blinding has not been described or there is not enough information to make an informed decision as to the adequacy of the blinding.

Incomplete outcome data

Adequately addressed

Where attrition or exclusions have been described and there are no dropouts, or no greater than a 20% attrition rate for any group. Also in cases where participants are randomised but then deemed ineligible for inclusion as long as the reason for ineligibility could not have been affected by the randomised intervention.

For dichotomous data, the proportion of missing outcomes should not have a clinically relevant effect on the effect estimate; for continuous data the plausible effect size of missing outcomes should not be large enough to have a clinically relevant impact on the observed size effect.

If outcome data is missing from both groups but the reasons are reported and balanced across both groups then this is still acceptable, providing they do not have different implications in the different groups. In studies examining time to event data, there is no bias if censoring is unrelated to prognosis.

For studies with missing data, it can sometimes be acceptable to input outcomes, such as “last observation carried forward”, but this would depend upon the type of study and needs to be assessed on a study to study basis, with the aid of a statistician.

Inadequately addressed

In studies where the missing outcomes are due to a true treatment effect, leading to an imbalance in numbers in the groups or differing reasons between groups for incomplete outcome data.

If a ‘per-treatment’ analysis is carried out when there has been a substantial dropout rate, this can over-estimate the treatment effect and lose the benefits of randomisation and therefore introduce a high level of bias.

For studies with missing data, if data is imputed without the advice of a statistician, this again can lead to over-estimation of treatment effects and severe bias.

Unclear

Where attrition/exclusions are not sufficiently reported in order for a judgement to be made, for example where dropouts are not reported on or reasons for attrition are not addressed.

Selective outcome reporting

Adequate outcome reporting

The study has addressed its aims and all its primary and secondary outcomes have been reported upon, as per the protocol. In cases where the protocol is unavailable, efforts have been made to include all expected outcomes such as those outlined in the methods section.

Inadequate outcome reporting

Where not all of a study's pre-specified outcomes are reported upon or they are reported using different measurements, methods of analysis or different subsets of data. Also cases where outcomes are reported upon when they were not initially described as part of the methodology or protocol are reported upon, unless they are as a direct result of one of the outcomes, such as an adverse event.

If outcomes are only partially reported upon, or are missing when you would ordinarily expect them to be reported, this increases the risk of bias and is considered inadequate.

Unclear

Where there is inadequate information in order for a judgement to be made.

Other potential threats to validity

Adequate

If no other obvious sources of bias can be detected.

Inadequate

The study has sources of bias relating to:

- Study design
- Baseline imbalances and population selection
- Blocked randomisation in unblended trials
- Differential diagnostic activity
- Departure from the study protocol that is unreflective of clinical practice
- Changes in protocol due to the effect of interim results
- Inaccurate measurement due to insensitive tools for measurement
- Fraudulent results

Unclear

Where there is inadequate information to allow a decision to be made as to whether an important risk of bias exists.

4.3.5 Measure of treatment effect

For continuous outcomes (e.g. crowding in millimetres) where studies used the same scale, we used the mean values and standard deviations (SDs) reported in the studies in order to express the estimate of effect as mean difference (MD) with 95% confidence interval (CI). Where different scales were used we would have considered expressing the treatment effect as standardised mean difference (SMD) with 95% CI.

4.3.6 Unit of analysis issues

The units of analysis were millimetres where crowding indices were used, degrees where the angulation of the incisors to the maxilla or mandible were used and time to alignment in days.

4.3.7 Dealing with missing data

Attempts were made to contact the author(s) in cases of missing data for all included studies, where it was feasible in order to gather details of outcomes that were measured but not reported upon, or for clarification and details. We were unable to use the methods described in Section 7.7.3 of the Cochrane Handbook for Systematic Reviews of Interventions to estimate missing SDs due to unclear or unavailable data. We did not use any other statistical methods or perform any further imputation to account for missing data.

4.3.8 Assessment of heterogeneity

Clinical heterogeneity was assessed by identifying the participants, interventions and outcomes and considering whether a meaningful summary to be produced by combining the results. We planned to assess statistical heterogeneity if there was a sufficient number of studies to allow meta-analysis to be carried out. This would be calculated using a chi-square (chi^2) test, where a P value < 0.1 indicated statistically significant heterogeneity, as advised in section 9.5.2 of the Cochrane Handbook for Systematic Reviews of Interventions. Statistical heterogeneity is quantified using the I^2 statistic, which is interpreted as follows:

- 0% to 40%: might not be important;
- 30% to 60%: may represent moderate heterogeneity;
- 50% to 90%: may represent substantial heterogeneity;
- 75% to 100%: considerable heterogeneity.

4.3.9 Assessment of reporting bias

Reporting bias is assessed via funnel plot asymmetry, as described in section 10.4 of the Cochrane Handbook for Systematic Reviews of Interventions. Whilst we had a

sufficient number of studies included in this review for the primary outcome, the results were split into many subheadings and so therefore we were unable to assess for publication bias and its possible causes.

4.3.10 Data synthesis

The Cochrane Collaboration statistical guidelines were followed and the data analysed using RevMan and reported upon according to Cochrane Collaboration criteria. Heterogeneity was assessed using Cochran's test and, if significant heterogeneity was detected, the significance of treatment effects was assessed using the random-effects model. Risk ratio, the number needed to treat and corresponding 95% confidence intervals, were calculated for dichotomous data. The weighted mean difference and 95% confidence intervals was calculated for continuous data. A subgroup analysis was be carried out on the age (stage of dental development) that treatment was carried out.

4.3.11 Subgroup analysis and investigation of heterogeneity

We carried out subgroup analyses according to type of crowding intervention and whether the intervention was for prevention or treatment.

4.3.12 Sensitivity analysis

If a sufficient number of studies had been found, then a sensitivity analysis would have been carried out.

4.3.13 Cross-over trials

There were no cross over trials identified for any given intervention.

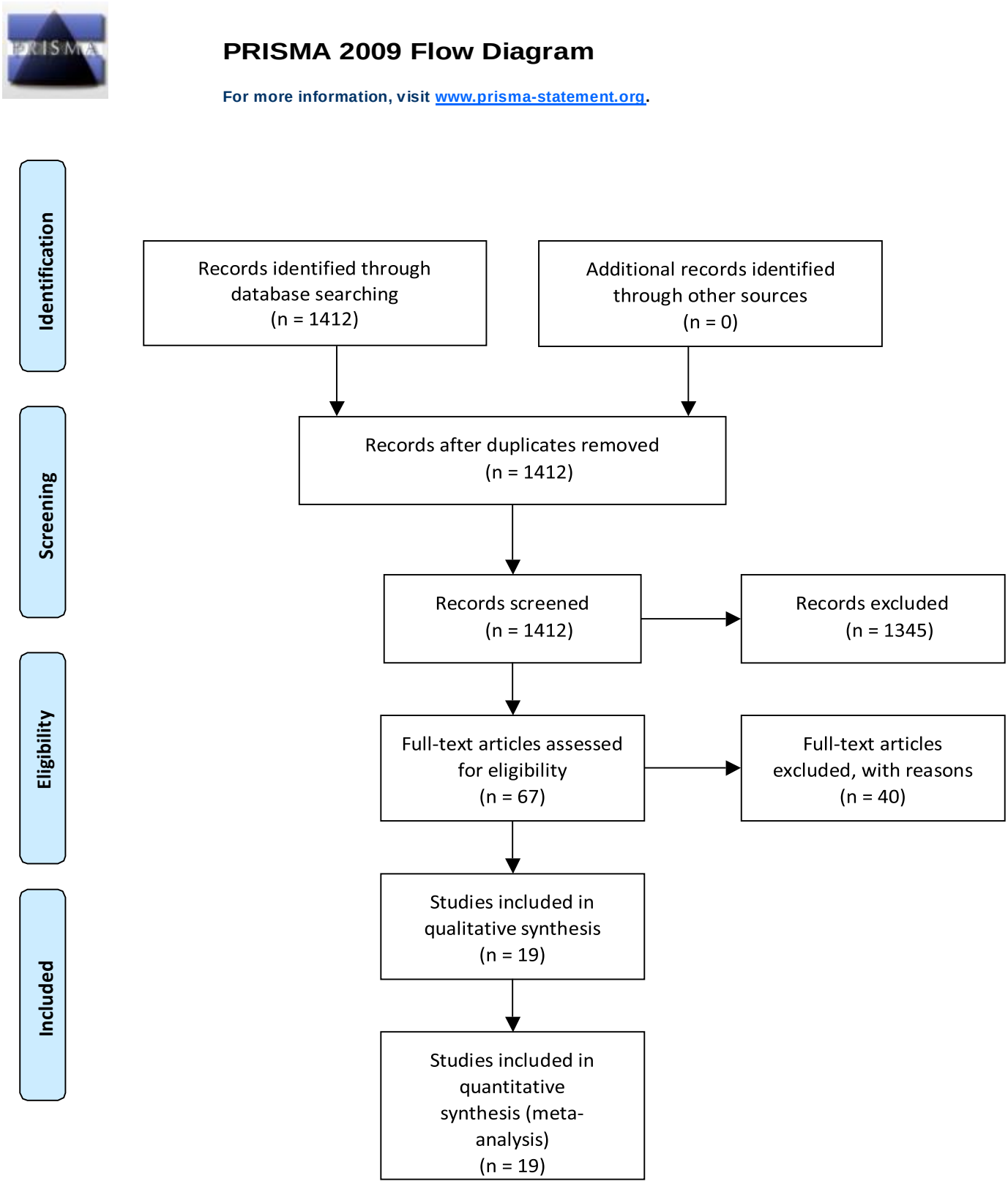
5: RESULTS

5.1: Description of studies

5.1.1: Results of the search

The electronic database search identified 1412 references to studies after the removal of duplicates. No additional articles were identified from additional sources. All titles and abstracts, where available, were screened and 1345 were discarded. Of the remaining 67 articles, we obtained full-text articles where possible, and excluded 40 of the studies at this stage. The remaining 27 studies appeared to meet our inclusion criteria and we were able to include 19 of the studies. Of the 8 studies that were excluded at this stage: 3 did not examine crowding as a primary outcome; 2 involved participants who fell outside the eligibility criteria; upon closer inspection, 1 was not an RCT; 1 presented insufficient data for inclusion in the analyses and 1 paper was not located despite having made attempts to contact the authors. This process is presented as a flow chart in Figure 2.

Figure 2. Study flow diagram.



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta- *A*nalyses The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

5.1.2: Included studies

Nineteen studies, involving a total of 1,101 participants (n =1,050) were included in this review.

5.1.3: Characteristics of the trial designs and settings

Design

Seventeen studies were of 2 arm parallel design (Atik 2014[80], Davidovitch 1997[87], Gravina 2013[88], Kau 2004[60], Krusinskiene 2008[57], Mantysaari 2004[89], Miles 2010[82], Miles 2012[90], Myrlund 2015[67], O'Brien 1990 [91], Pandis 2009[92], Pandis 2010[83], Pandis [79], Pirttiniemi 2005[56], Rebellato 1997[63], Sebastian 2012[93], Tai 2010[66]); the remaining two studies were of 3 arm parallel design (Ong 2011[94] and Woodhouse 2015[95])

Setting

One study was conducted in Turkey (Atik 2014[80]), one in the USA (Davidovitch 1997[87]), one in Brazil (Gravina 2013[88]), one was split over Italy, Germany and Wales (Kau 2004[60]), two in Finland (Krusinskiene 2008[57], Mantysaari 2004[89]), three in Australia (Miles 2010[82], Miles 2012[90], Ong 2011[94]), one in Norway (Myrlund 2015[67]), two in England (O'Brien 1990[91], Woodhouse 2011[95]), three in Greece (Pandis 2009[92], Pandis 2010[83], Pandis 2011[79]), one in India (Sebastian 2012[93], one in Japan (Tai 2010[66]) and for two the setting was not specified (Pirttiniemi 2005[56], Rebellato 1997[63]) although of these, one was linked to two other studies and so was most likely set in Finland (Pirttiniemi 2005[56]).

There were fifteen single centre studies (Atik 2014[80] Davidovitch 1997[87], Gravina 2013[88], Miles 2010[82], Miles 2012[90], Ong 2011[94], Myrlund 2015[67], O'Brien 1990[91], Pandis 2009[92], Pandis 2010[83], Pandis 2011[79], Rebellato 1997[63], Sebastian 2012[93] Tai 2010[66], Woodhouse 2011[95]) and one with three centres (Kau 2004[60]). Krusinskiene 2008[57], Mantysaari 2004 [89] and Pirttiniemi 2005[56]) report different outcomes and time points (1, 2, 4, 8 and 13 years) for one study and are collectively referred to as Finland 2004.

Funding

Four studies reported their funding source (Kau 2004[60], Krusinskiene 2008[57], Ong 2011[94], Rebellato 1997[63]), all of which were in the form of independent funding from

government, universities or dental/orthodontic charities and societies. The remaining fifteen studies did not report any funding source.

Conflict of interest

Three studies declared that there were no commercial or financial conflicts of interest (Gravina 2013[88] and Tai 2010[66], Woodhouse 2011[95]). However, one of these studies (Tai 2010[66]) later acknowledged an engineering company for developing the software which the authors used to disprove their null hypothesis and another study (Woodhouse 2011[95]) declared that their appliances were supplied by the development company of the appliances. Two other studies (Myrlund 2015[67], O'Brien 1990[91]) declared that various companies supplied them with materials or appliances. The remaining fourteen studies did not report on any conflicts of interest.

5.1.4: Characteristics of the participants

There were 1,101 participants randomized to interventions (including only the intervention groups relevant to this review), of which 1,050 were included in the studies' analyses. Ages ranged from 7.6 to 17 years. In general, there were comparable numbers of males and females in the studies however, two studies recruited only female participants, (Atik 2014[80], Sebastian 2012[93]) four studies did not report on the gender distributions of the respective groups (Davidovitch 1997[87], Kau 2004[60], O'Brien 1990[91], Rebellato 1997[63]) and one had large variations in the number of males to females in the different groups after statistical analysis (Miles 2012[90]). Another four studies also had variations in baseline gender distributions but appropriate statistical analysis was carried out to account for this in the results (Krusinskiene 2008[57], Mantysaari 2004[58], Pirttiniemi 2005[56], Ong 2011[94]).

5.1.5: Orthodontic interventions

The studies identified offered interventions either to prevent the perpetuation of crowding from the mixed dentition into the permanent dentition, or to treat crowding in the permanent dentition. In one study (Finland 2004, [56,57,88]), participants received cervical pull headgear. In one study participants had lip bumpers fitted (Davidovitch 1997[87]), in another study they had lingual arches placed (Rebellato 1997[63]) and another study's participants received extraction of the lower deciduous canines (Kau 2004[60]). In two studies, participants received removable appliances, namely the Schwarz and Guidance Development appliances (Tai 2010[66], Myrlund 2015[67]). In

four studies, participants received self-ligating brackets (Pandis 2010[83], Pandis 2011[79], Miles 2010[82], Atik 2014[80]) and in five studies they received nickel titanium archwires (Gravina 2013[88], O'Brien 1990[91], Ong 2011[94], Pandis 2009[92], Sebastian 2012[93]). In the final two studies, participants received vibrational appliances (Miles 2012[90] and Woodhouse 2015[95]).

5.1.6: Characteristics of the interventions and comparisons

Prevention

Cervical pull headgear

One study (Finland 2004) compared cervical pull headgear to a control group, which received interceptive procedures during the study duration to improve the alignment of the anterior teeth if deemed necessary (Mantysaari 2004[89], Pirttiniemi 2005[56], Krusinskiene 2008[57]). The interceptive procedures consisted of extraction of the upper deciduous canines, extraction of the lower deciduous canines or interdental stripping. These three studies were all carried out on the same group of participants in Finland, but each study reported on different outcomes at different time points.

Lip bumper

One study compared lip bumper therapy to a control group, who did not receive any active treatment to assess arch perimeter changes (Davidovitch 1997[87]).

Lingual arch

One study compared the lower lingual arch appliance against a control group who did not receive any active treatment during the study period. The aim was assess arch length and incisor inclinational changes (Rebellato 1997[63]).

Extraction of lower deciduous canines

One study compared lower deciduous incisor extractions against a control group who received no treatment during the study period (Kau 2004[60]).

Removable appliances

Two studies compared removable appliances against control groups in which participants received no treatment. The first study (Tai 2010[66]) expanded the upper and lower arches with the Schwarz appliance and the participants in the second study (Myrlund 2015[67]) received an Eruption Guidance Appliance for both the upper and lower arches.

Treatment

Brackets

Four studies compared different types of self –ligating and conventional brackets. Two studies compared conventional, metal Roth system brackets with passive, self-ligating metal brackets (Pandis 2011[79], Atik 2014[80]). In the Atik 2014 [81] study, the conventional group underwent treatment with a quadhelix before fixed appliance therapy. One study compared porcelain self-ligating brackets with conventional porcelain brackets (Miles 2010[82]) and one study compared active and passive self-ligating brackets (Pandis 2010[83]).

Archwires

A total of five studies compared different archwire types against one another. Two studies had three parallel arms; the first compared stainless steel, multi-stranded steel and super-elastic nickel-titanium archwires (Gravina 2013[88]). The second study compared nickel-titanium, heat activated nickel-titanium and copper nickel-titanium archwires (Ong 2011[94]). The remaining three studies had two parallel arms. One study compared stabilised nickel-titanium, Nitinol, against super–elastic nickel-titanium, Titinol (O'Brien 1990[91]); one compared nickel-titanium against copper nickel titanium (Pandis 2009[92]) and the last study compared coaxial nickel-titanium against single stranded nickel-titanium (Sebastian 2012[93]).

Vibrational appliances

Two studies investigated the effects of vibrational appliances on crowding. The first had two parallel arms and compared the vibrational appliance (Tooth Masseuse) and fixed appliances with fixed appliances alone (Miles 2012[90]). Another study had three parallel arms consisting of participants who underwent mandibular first premolar extractions and received the vibrational appliance (AcceleDent) and fixed appliances with a sham AcceleDent device and fixed appliances against fixed appliances only (Woodhouse 2015[95]).

5.1.7: Characteristics of the outcomes

5.1.7.1: Primary outcome

For the primary outcome of crowding, we were interested in the amount of crowding, measured in millimetres, measured by any index of malocclusion. Nine of the studies used Little's Irregularity Index in the mandible (Krusinskiene 2008[57], Kau 2004[60], Tai

2010[66], Myrlund 2015[67], Gravina 2013[88], Ong 2011[94], Pandis 2009[92], Sebastian 2012[93], Miles 2012[90], Woodhouse 2015[95]). Four studies reported on maxillary crowding (Myrlund 2015[67], Miles 2010[82], Pandis 2010[83], O'Brien 1990[91]) and two studies did not specify the region where crowding was measured (Davidovitch 1997[87], Tai 2010[66]). All fifteen of these studies used millimetres and recorded crowding in the anterior region of the maxilla or mandible, except for one study that recorded it from molar to molar in the mandible (Gravina 2013[88]).

A total of 19 different time points were recorded across these fifteen studies which ranged from pre-treatment records to a one thirteen-year follow-up. These time points varied greatly, with some readings in days, some in weeks and months and a few in years. For most interventions, there was only one study available, but for the comparison of vibrational appliances against a control we combined data from two studies (Miles 2012 [90] and Woodhouse 2015[95]) by converting the time points into weeks. This was considered to be the most clinically relevant time descriptor for the reduction of irregularity and also a reliable unit, as the days in a month can vary.

5.1.7.2: Secondary outcomes

Lower molars to mandible

This outcome was reported on in two studies investigating two different comparisons: lip bumper versus control in Davidovitch 1997 [87] and lingual arch versus control in Rebellato 1997[63]. It was measured in two different formats, molar inclinational change in degrees and molar anterior-posterior movement change in millimetres. In Rebellato 1997 [63] angular change was measured relative to the functional occlusal plane which was described as a line drawn through maximum inter-cuspation of the occlusion, whereas in Davidovitch 1997 [87] it was measured relative to the mandibular plane, described as the line drawn between gnathion and pogonion.

Lower incisors to mandible

Seven studies reported on this outcome (Mantysaari 2004[89], Davidovitch 1997[87], Rebellato 1997[63], Kau 2004[60], Tai 2010[66], Myrlund 2015[67], Gravina 2013[88]), but each study was investigating a different comparison and so it was not appropriate to combine the results.

Self esteem

No study reported on this outcome.

Patient satisfaction

No study reported on this outcome.

Jaw joint problems

No study reported on this outcome.

Harms

Four studies reported on the discomfort experienced for differing comparisons. Two of these used the 7-point Likert Scale (Miles 2010[82], Ong 2011[94]) and two used a 100 millimetre Visual Analogue Scale (VAS) (Atik 2014[80], Miles 2012[90]). One study also reported on Plaque Index, Gingival Index and probing depths (Atik 2014[80]).

Additional secondary outcomes

Upon data extraction, we decided that certain outcomes were of interest and relevant clinically. The data for these were extracted as an amendment to the original protocol and are mentioned below.

Upper incisors to maxilla

Three studies reported on this outcome for differing comparisons, so the results were not combinable. Two studies measured in degrees the angle formed between the upper incisor and the line between sella and nasion (Tai 2010[66], Myrlund 2015[67]) for the Schwarz appliance and the Eruption Guidance appliance. One study (Mantysaari 2004[89]) reported on the angle of the upper incisor to the maxillary plane (line between anterior and posterior nasal spines).

Arch length

This was reported by five studies as the change in arch length (Pirttiniemi 2005[56], Davidovitch 1997[87], Rebellato 1997[63], Kau 2004[60], Tai 2010[66]) for differing comparisons, so it was not appropriate to combine the results. Changes occuring from the start to the end of the individual studies were measured.

Time to alignment

Four studies reported on this outcome, two of which were both for copper nickel-titanium versus nickel-titanium archwire (Ong 2011[94], Pandis 2009[92]), so the results were combined. The other two studies (Pandis 2011[79], Pandis 2010[83]) looked at differing comparison so it was not appropriate to combine the results.

Ligation time

One study reported on this outcome (Miles 2010[82]) for the two different types of brackets investigated.

5.1.8: Excluded Studies

We excluded forty six studies (See Appendix 2: Characteristics of excluded studies) from this review for the following reasons:

- Not a randomized controlled trial (RCT):
  Abu Alhaija 2011[96], Barlin 2011[97], Baumrind 1996[98], Dai 2009[99], Freitas 2013[100], Fan 2009[101], Heiser 2004[102], Keski-Nisula 2008[103] Miyake 2008[104], Nagalakshmi 2014[105], Ong 2001[106], Owais 2011[107], Pandis 2007[108] Pandis 2010a[109], Ruf 1999[110] , Sucuru 1992[111], Vajaria 2011[112], Yu 2008[113].
- The patient's ages fell outside of the eligibility criteria:
  Almeida 2005[114], Clements 2003[115], Cobb 1998[116], DiBiase 2011[117], Fleming 2009[118], Harradine 1998[119], Kau 2013[120], Linqvist 1982[121], Scott 2008[122] Serafim 2015[123], Soldanova 2012[124], Talapaneni 2012[125], Taner 2000[126], Wahab 2012[127] , Wang 2010 [128].
- Crowding was not an outcome of the study:
  Altug 2005[129], Thickett 2009 [130], Bondemark 2005[131], Markovic 2015[132], Okay 2006 [133] , Rowland 2007[134], Sandhu 2013[135], Shawesh 2010[136], Silva 2012[137], Strahm 2009 [138] , Wortham 2009 [139]
- Insufficient information to allow inclusion of data:
  West 1995[140].

5.2: Risk of bias in included studies

5.2.1: Allocation

5.2.1.1: Random sequence generation

Ten studies described an adequate method of random sequence generation (Kau 2004[60], Krusinskiene 2008[57], Myrlund 2015[67], Ong 2011[94], Pandis 2009[92], Pandis 2010[109], Pandis 2011[79], Pirttiniemi 2005[56], Sebastian 2012[93] Woodhouse 2015[95]),

when assessed using information published in the papers and further information received via correspondence with the authors when required. These ten papers were therefore assessed as being at low risk of bias for this domain. The remaining twelve studies simply stated that participants were randomized, however either did not describe their methods, or the method remained unclear, so they were assessed as being at unclear risk of bias for this domain.

5.2.1.2: Allocation concealment

Eight studies described an adequate method of allocation concealment (Kau 2004[60], Krusinskiene 2008[57], Myrlund 2015[67], Ong 2011[94], Pandis 2009[92], Pandis 2010[109], Pandis 2011[79], Sebastian 2015[93]), when assessed using information published in the papers and further information received via correspondence with the authors when required. These eight papers were therefore assessed as being at low risk of bias for this domain. The eleven remaining studies did not mention any methods used to conceal the random sequence, and we assessed them as being at unclear risk of bias.

5.2.2: Blinding

5.2.2.1: Blinding of participants and personnel (performance bias)

Two studies described adequate methods of blinding of participants and personnel and were therefore assessed as being a low risk of bias for this domain (Ong 2011[94], Pandis 2009[92]).

In one study (Miles 2012[90]) the clinicians were blinded but it was not possible to blind participants to the type of intervention to which they were allocated, so this was assessed as being at unclear risk. In one study (Miles 2010[82]) the participants were blinded but it was not possible to blind the clinicians to the type of allocation to which the participants were allocated, so this was assessed as being at unclear risk.

Six studies did not mention blinding of participants and personnel (Atik 2014[80], Gravina 2013[88], Myrlund 2015[67], O'Brien 1990 [91], Pandis 2010 [109], Woodhouse 2015[95]) and so were assessed as being high risk.

The remaining nine studies were unable to blind both the participants and clinicians for the types of interventions used, so they were assessed as being high risk

(Davidovitch 1997[87], Kau 2004[60], Krusinskiene 2008[57], Mantysaari 2004[89], Pandis 2011[79], Pirttiniemi 2005[56], Rebellato 1997[63], Sebastian 2012[93], Tai 2010[66]).

5.2.2.2: Blinding of outcome assessment (detection bias)

Nine studies described an adequate method of blinding of outcome assessment (Kau 2004[60], Miles 2010[82], Miles 2012[90], Myrlund 2015[67], Ong 2011[94], Pandis 2011[109], Sebastian 2012[93], Tai 2010[66], Woodhouse 2015[95]). These nine studies were therefore assessed as being at low risk of bias for this domain.

Ten studies did not describe their methods and so were assessed as being at unclear risk of bias for this domain (Atik 2014[80], Davidovitch 1997[87], Gravina 2013[88], Krusinskiene 2008[57], Mantysaari 2004[89], O'Brien 1990[91], Pandis 2009[92], Pandis 2010 [109], Pirttiniemi 2005[56], Rebellato 1997[63]).

5.2.3: Incomplete outcome data

Eleven studies were assessed as being at low risk of attrition bias (Atik 2014[80], Kau 2004[60], Mantysaari 2004[58], Miles 2012[90], Myrlund 2015[67], O'Brien 1990[91], Pandis 2009[92], Pandis 2010[83], Pandis 2011[79], Sebastian 2012[93], Woodhouse 2015[95]).

Two studies were considered to be at high risk due to high numbers of attrition across the studies (Krusinskiene 2008[57], Pirttiniemi 2005[56]) which reported the 13-year and 8-year follow-up in Finland 2004.

The remaining six studies did not report on dropouts and so were assessed as being at unclear risk (Davidovitch 1997[87], Gravina 2013[88], Miles 2010[82], Ong 2011[94], Rebellato 1997[63], Tai 2010[66]).

5.2.4: Selective reporting

Two studies were at unclear risk of selective reporting bias as they did not state any specific outcomes in the methods section, but reported on appropriate outcomes in the results, so they were considered to be at unclear risk for bias (Tai 2010[66], Woodhouse 2015[95]).

5.2.5: Other potential sources of bias

Three studies were assessed as high risk of other sources of bias, two of which were as a result of gender bias in sampling, having recruited only female participants (Atik 2014[80], Sebastian 2012[93]) and one was as a result of having statistically significant

differences in baseline crowding between the two groups of participants (Kau 2004[60]).

Six studies were considered to have unclear risk of bias for this domain. One of these had statistically significant differences for the same outcome measured by two different methods in the same study (Davidovitch 1997[87]). Two studies carried out interceptive procedures on the control group, which were active treatments, whilst comparing against the main intervention and the effect of these on the final outcome is unknown for Mantysaari 2004 [58] but was statistically significant for Pirttiniemi 2005[56]; one study (Miles 2010[82]) removed the results for two participants in order to balance the two groups for numbers.

The remaining nine studies were not considered to have any other potential sources of bias and were therefore assessed as being at low risk of bias for this domain.

5.2.6: Overall risk of bias

None of the studies identified were completely free of bias. However, two studies were assessed as being at low risk of bias (Ong 2011[94], Pandis 2009[92]). Another two studies were assessed as being at unclear risk of bias (Miles 2010[82], Miles 2012[90]). The remaining fifteen studies were assessed as being at high overall risk of bias (Figure 3 and Figure 4).

Figure 3: Risk of bias summary: Review authors' judgements about each risk of bias item for each included study.

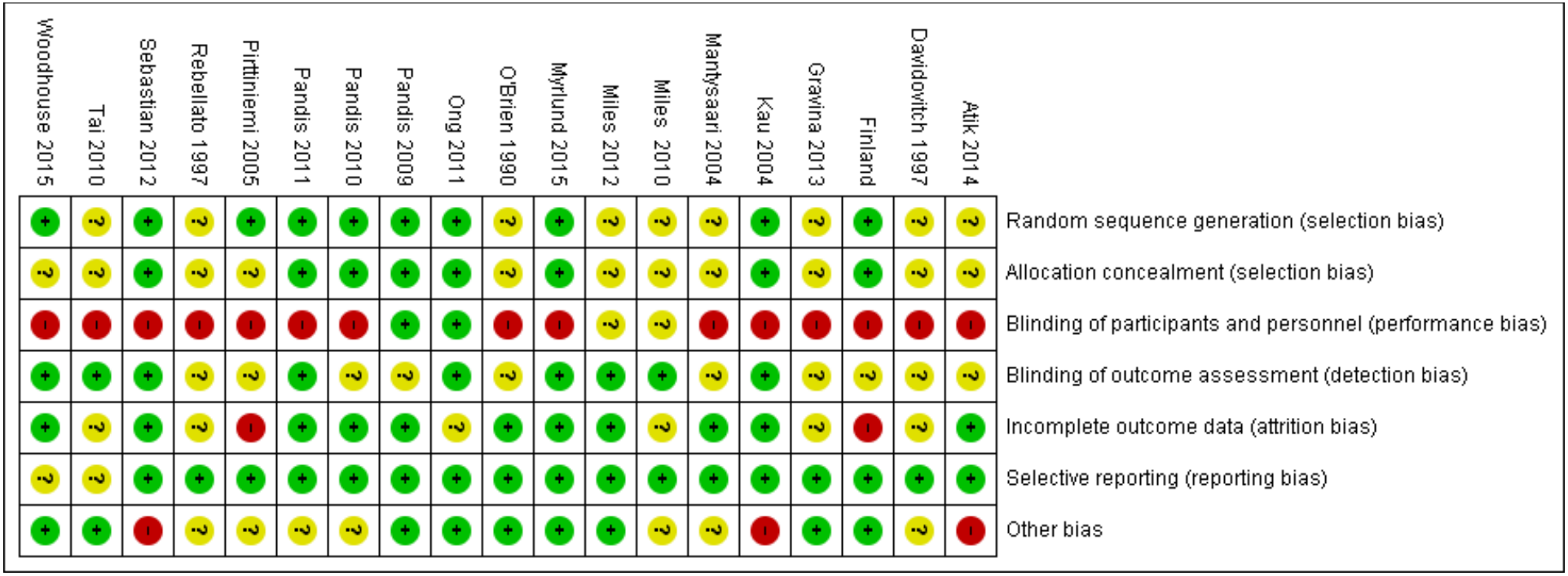
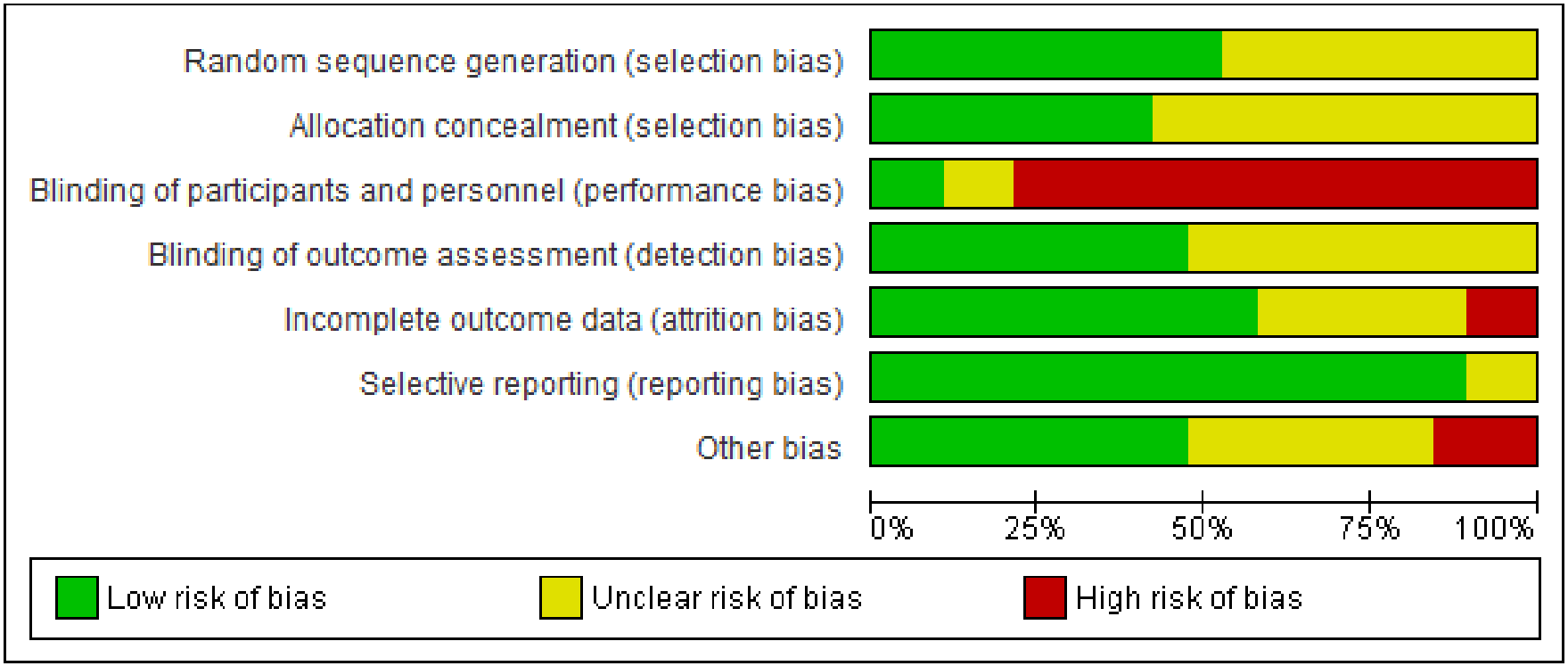


Figure 4: Risk of bias graph: Review authors' judgements about each risk of bias item presented as percentages across all included studies.



5.3: Effects of interventions

5.3.1: Comparison 1: Cervical pull headgear versus control

Crowding

One report of Finland 2004, investigated lower incisor crowding, in millimetres, over a thirteen-year follow-up period (Krusinskiene 2008[57]). The time points at which crowding was measured were: baseline, two, four, eight and thirteen years. This report was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible and at the thirteen-year recall, there was only a fifty-three per cent response rate meaning that there was high attrition bias. In total, sixty-four participants began the study and provided information on crowding for up to four years, fifty-four participants provided data for the eight-year recall but only thirty-four returned for the final recall at thirteen years.

We were unable to perform a meta-analysis, as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 5). This revealed:

- Baseline equivalence in crowding between the headgear and control groups with no statistical difference between them (P= 0.14)
- No statistically significant difference in the amount of lower incisor crowding between the two groups at any time point:
    - P= 0.48 at two years
    - P= 0.94 at four years
    - P= 0.24 at eight years
    - P= 0.75 at thirteen years
- No statistically significant difference in lower incisor crowding change between any of the time points (P= 0.42).

Time to alignment

This outcome was not relevant for this comparison.

Ligation time

This outcome was not relevant for this comparison.

Arch length

One report of this study (Pirttiniemei 2005[56]) reported on maxillary and mandibular arch length change, in millimetres, over an eight-year period. The time points at which the arch length was measured were at baseline, two years and eight years and the changes between these time points were reported. This report of the Finland 2004 study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In addition, although there was a 16% attrition rate overall, the attrition rate in the headgear group was 26.5%, so this study was assessed as being at high risk for this domain. In total, sixty-four participants began the study and provided information on crowding for up to two years and fifty-four (84.4%) participants provided data for the eight-year recall.

We were unable to perform a meta-analysis, as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figures 6 and 7). This revealed:

- For mandibular arch length:
    - Between baseline and the two-year recall, arch length increased statistically significantly more in the headgear group (1.3mm) compared to the control group (95% CI 1.17-1.43mm, P= < 0.00001).
    - Between the two and eight-year recall, arch length again increased statistically significantly more in the headgear group (1.52mm) compared to the control group (95% CI 1.3-1.74, P= < 0.00001).
    - There was no statistically significant difference in arch length change between two and eight-years for either group (P= 0.09).
- For maxillary arch length:
    - Between baseline and the two-year recall, arch length increased statistically significantly more in the headgear group (1.98mm) compared to the control group (95% CI 1.8-2.16, P= < 0.00001).
    - Between the two and eight-year recall, arch length again increased statistically significantly more in the headgear group (2.28mm) compared to the control group (95% CI 2.05-2.51, P= < 0.00001).
    - The change in arch length between two and eight-years was statistically significant (P= 0.04).

Lower incisors to mandible

One report of the Finland 2004 study reported on lower incisor inclination, in degrees, over a two-year period (Mantysaari 2004[89]). The time points at which lower incisor change was measured were at baseline, baseline to one-year and baseline to two years. This study was assessed as being at high risk of bias overall as blinding of personnel and participants was not possible. In total, sixty-four participants were included in the analysis.

We were unable to perform a meta-analysis, as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figures 8). This revealed:

- Baseline equivalence between the headgear and control groups with no statistical difference between the proclination of the lower incisors (P= 0.47)
- Statistically significant more proclination of the lower incisors (2.3 degrees) in the headgear group compared to the control group, between baseline and one-year (95% CI 0.67-3.93, P= 0.006).
- No-statistically significant difference in the proclination of the lower incisors (1.4 degrees) between baseline and two-years, (95% CI -0.42 to 3.22, P= 0.13).
- Overall, there is no statistical difference for in the change of the proclination of the lower incisors between either subgroup (P= 0.47).

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

One report of the Finland 2004 study reported on upper incisor inclination, in degrees, over a two-year period (Mantysaari 2004[89]). The time points at which upper incisor change was measured were baseline, baseline to one-year and baseline to two-years. This report was assessed as being at high risk of bias overall as blinding of personnel and participants was not possible. In total, sixty-four participants were included in the analysis.

We were unable to perform a meta-analysis, as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figures 9). This revealed:

- Baseline equivalence between the headgear and control groups with no statistical difference between their upper incisor inclination (P= 0.10).
- Statistically significant more proclination of the upper incisors in the headgear group, compared to the control group, of 4 degrees between baseline and one-year (95% CI 1.97-6.03, P= 0.0001).
- Statistically significant more proclination of the upper incisors between baseline and two years, of 4.5 degrees (95% CI 1.36-7.64, P= 0.005) at two years.
- Overall, there is no statistical difference for upper incisor change between either subgroup (P= 0.79).

Harms

No study in this subgroup presented data in a way that facilitated assessment of this outcome.

Self-esteem

No study in this subgroup reported on this outcome.

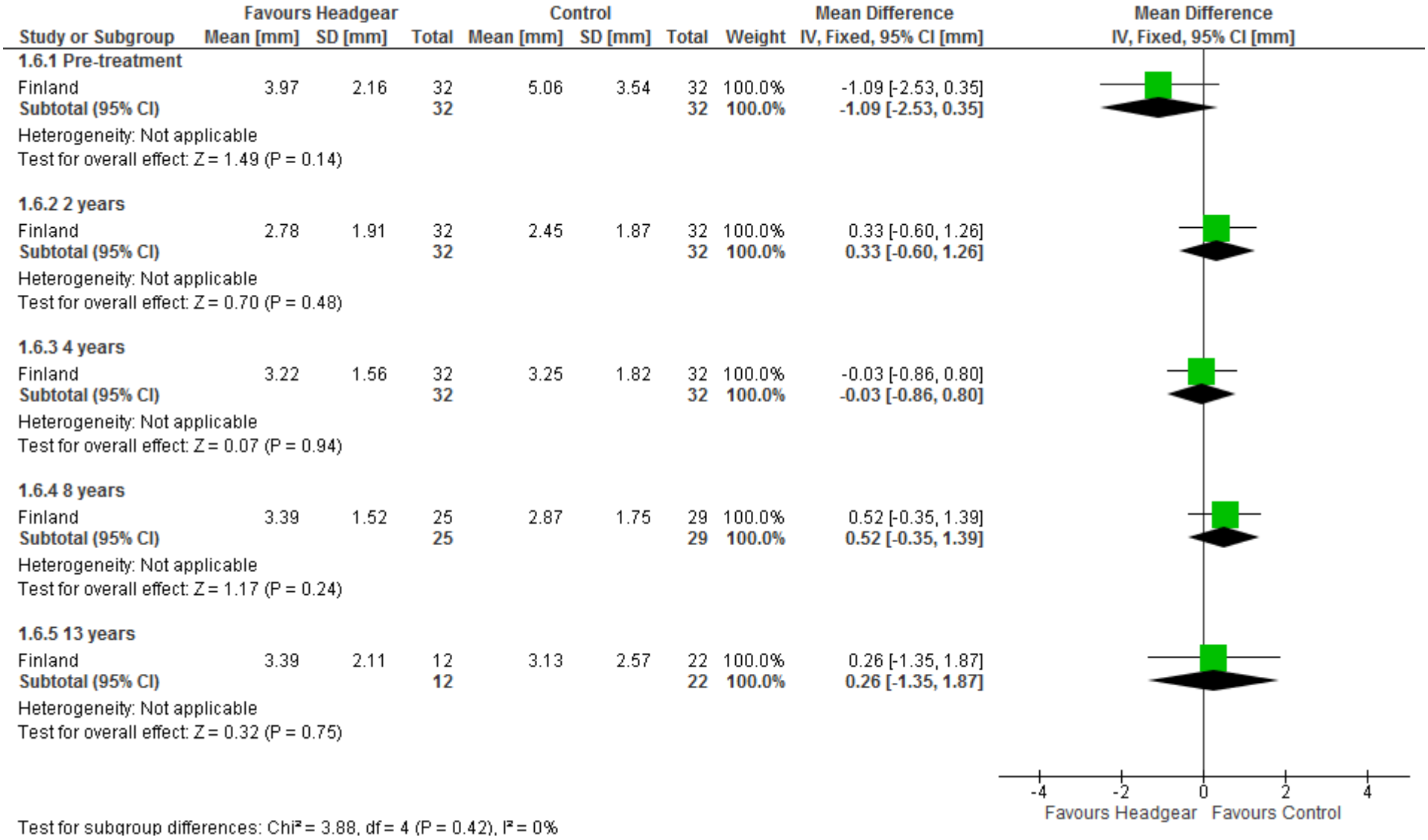
Patient satisfaction

No study in this subgroup reported on this outcome.

Jaw joint problems

No study in this subgroup reported on this outcome.

Figure 5: Forest plot of comparison 1: Cervical pull headgear versus control, outcome: Lower incisor crowding



Test for subgroup differences: Chi² = 3.88, df = 4 (P = 0.42), I² = 0%

Figure 6: Forest plot of comparison 1: Cervical pull headgear versus control, outcome: Mandibular arch length change

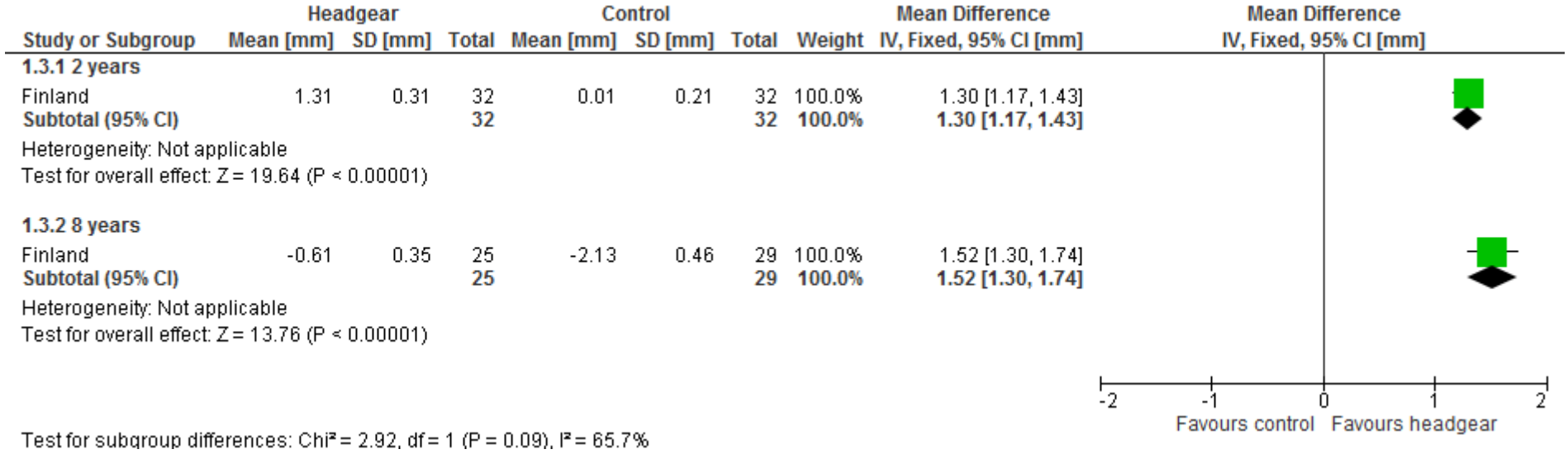


Figure 7: Forest plot of comparison 1: Cervical pull headgear versus control, outcome: Maxillary arch length change

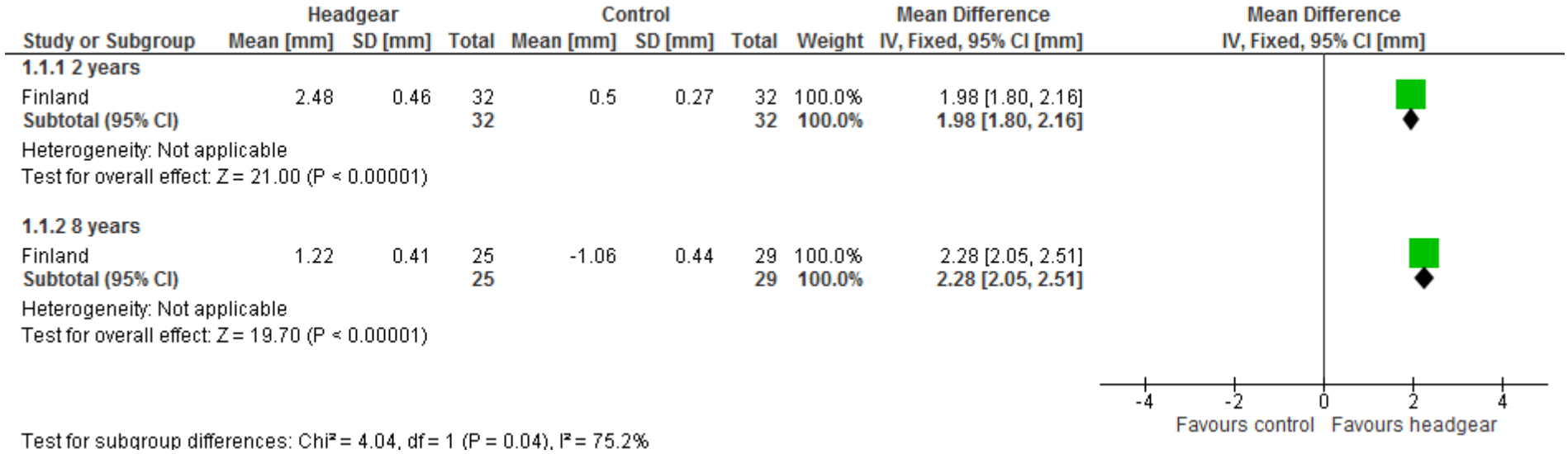


Figure 8: Forest plot of comparison 1: Cervical pull headgear versus control, outcome: Lower incisor inclinational change

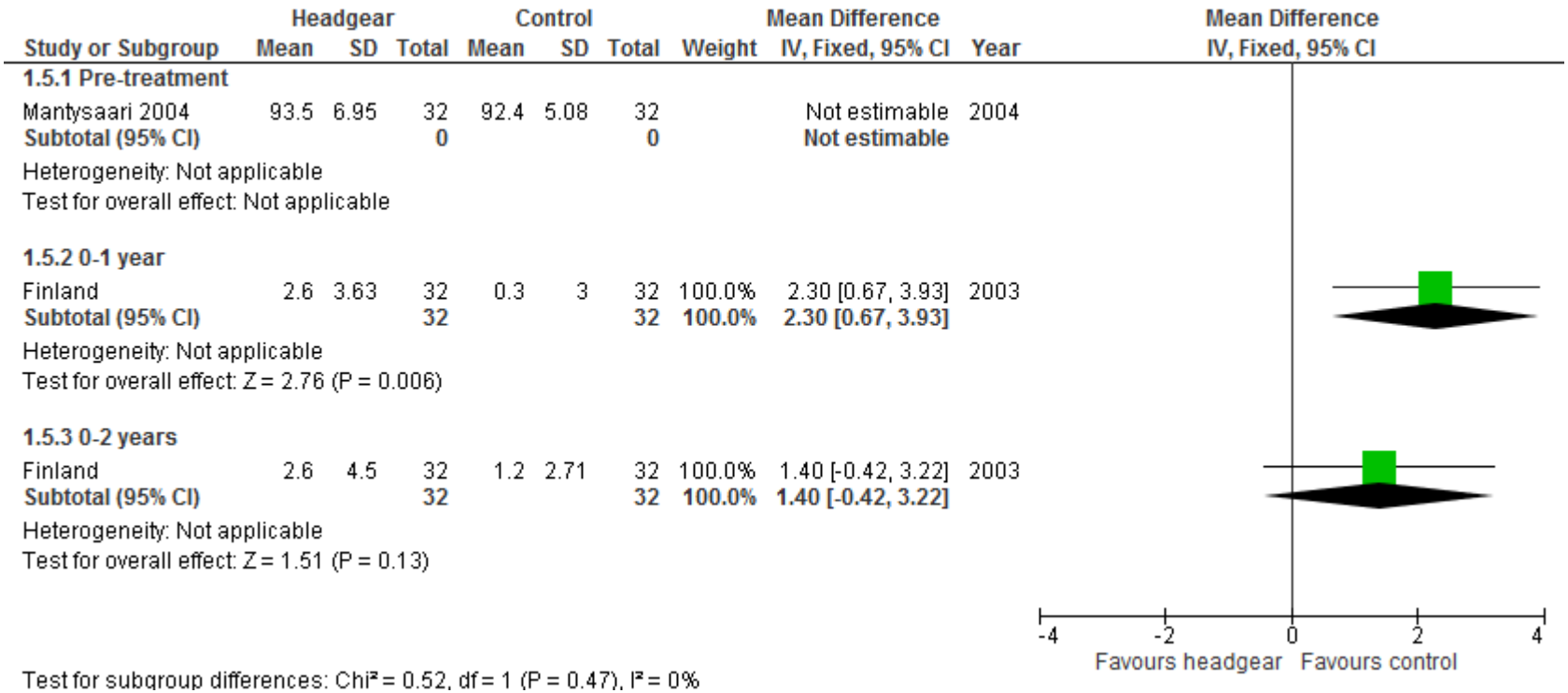

Headgear
Control
Mean Difference
Mean Difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Fixed, 95% CI Year
IV, Fixed, 95% CI
1.5.1 Pre-treatment
Mantysaari 2004 93.5 6.95 32 92.4 5.08 32 Not estimable 2004
Subtotal (95% CI) 0 0 Not estimable
Heterogeneity: Not applicable
Test for overall effect: Not applicable
1.5.2 0-1 year
Finland 2.6 3.63 32 0.3 3 32 100.0% 2.30 [0.67, 3.93] 2003
Subtotal (95% CI) 32 32 100.0% 2.30 [0.67, 3.93]
Heterogeneity: Not applicable
Test for overall effect: Z = 2.76 (P = 0.006)
1.5.3 0-2 years
Finland 2.6 4.5 32 1.2 2.71 32 100.0% 1.40 [-0.42, 3.22] 2003
Subtotal (95% CI) 32 32 100.0% 1.40 [-0.42, 3.22]
Heterogeneity: Not applicable
Test for overall effect: Z = 1.51 (P = 0.13)
-4
-2
0
2
4
Favours headgear
Favours control
Test for subgroup differences: Chi² = 0.52, df = 1 (P = 0.47), I² = 0%

Figure 9: Forest plot of comparison 1: Cervical pull headgear versus control, outcome: Upper incisor inclinational change

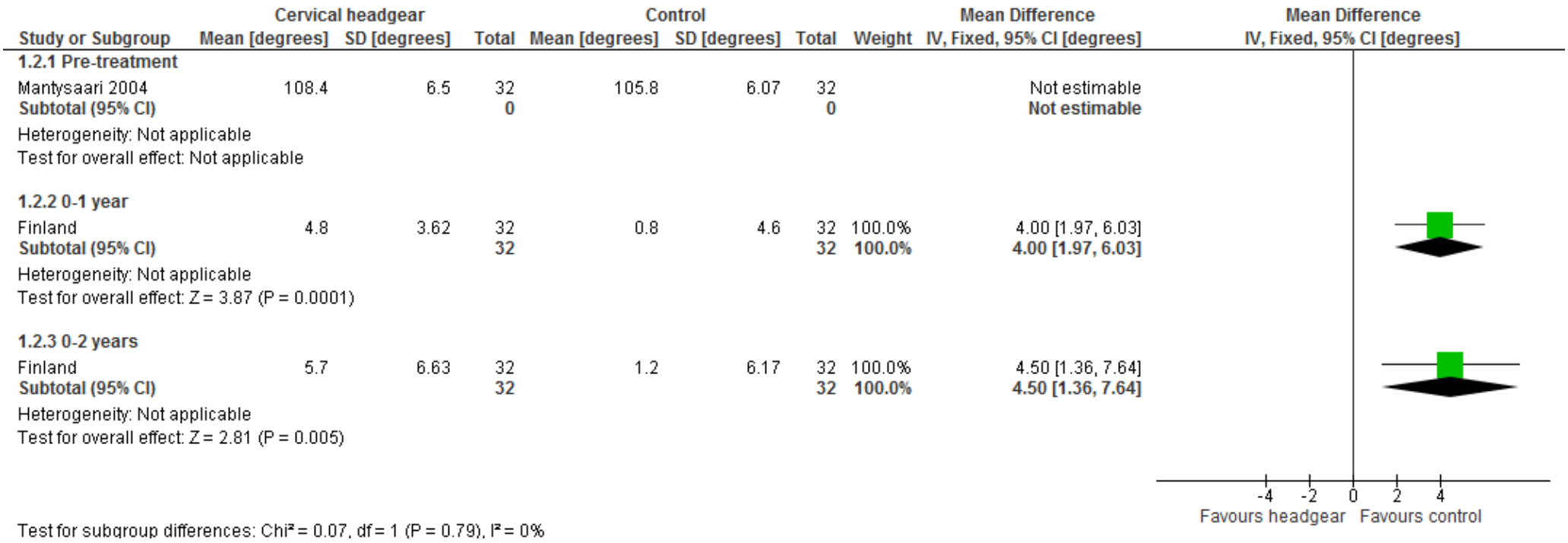


Table 3: Cervical pull headgear versus control, outcome: Lower incisor crowding

Time point	Mean HG	SD	N	Mean Control	SD	N	Weight	Mean diff (95% CI)
Pre-treatment								
Finland	3.97	2.16	32	5.06	3.54	32	100.0%	-1.09 [-2.53, 0.35]
Subtotal (95% CI)			32			32	100.0%	-1.09 [-2.53, 0.35]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.49 (P = 0.14)								
2 years								
Finland	2.78	1.91	32	2.45	1.87	32	100.0%	0.33 [-0.60, 1.26]
Subtotal (95% CI)			32			32	100.0%	0.33 [-0.60, 1.26]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.70 (P = 0.48)								
4 years								
Finland	3.22	1.56	32	3.25	1.82	32	100.0%	-0.03 [-0.86, 0.80]
Subtotal (95% CI)			32			32	100.0%	-0.03 [-0.86, 0.80]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.07 (P = 0.94)								
8 years								
Finland	3.39	1.52	25	2.87	1.75	29	100.0%	0.52 [-0.35, 1.39]
Subtotal (95% CI)			25			29	100.0%	0.52 [-0.35, 1.39]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.17 (P = 0.24)								
13 years								
Finland	3.39	2.11	12	3.13	2.57	22	100.0%	0.26 [-1.35, 1.87]
Subtotal (95% CI)			12			22	100.0%	0.26 [-1.35, 1.87]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.32 (P = 0.75)								
Test for subgroup differences: Chi^2 = 3.88, df = 4 (P = 0.42), I^2 = 0%								

Table 4: Cervical pull headgear versus control, outcome: Mandibular arch length change

Time point	Mean HG	SD	N	Mean Control	SD	N	Weight	Mean diff (95% CI)
2 years								
Finland	1.31	0.31	32	0.01	0.21	32	100.0%	1.30 [1.17, 1.43]
Subtotal (95% CI)			32			32	100.0%	1.30 [1.17, 1.43]
Heterogeneity: Not applicable								
Test for overall effect: Z = 19.64 (P < 0.00001)								
8 years								
Finland	-0.61	0.35	25	-2.13	0.46	29	100.0%	1.52 [1.30, 1.74]
Subtotal (95% CI)			25			29	100.0%	1.52 [1.30, 1.74]
Heterogeneity: Not applicable								
Test for overall effect: Z = 13.76 (P < 0.00001)								
Test for subgroup differences: Chi² = 2.92, df = 1 (P = 0.09), I² = 65.7%								

Table 5: Cervical pull headgear versus control, outcome: Maxillary arch length change

Time point	Mean HG	SD	N	Mean Control	SD	N	Weight	Mean diff (95% CI)
2 years								
Finland	2.48	0.46	32	0.5	0.27	32	100.0%	1.98 [1.80, 2.16]
Subtotal (95% CI)			32			32	100.0%	1.98 [1.80, 2.16]
Heterogeneity: Not applicable								
Test for overall effect: Z = 21.00 (P < 0.00001)								
8 years								
Finland	1.22	0.41	25	-1.06	0.44	29	100.0%	2.28 [2.05, 2.51]
Subtotal (95% CI)			25			29	100.0%	2.28 [2.05, 2.51]
Heterogeneity: Not applicable								
Test for overall effect: Z = 19.70 (P < 0.00001)								
Test for subgroup differences: Chi² = 4.04, df = 1 (P = 0.04), I² = 75.2%								

Table 6: Cervical pull headgear versus control, outcome: Lower incisor inclinational change

Time point	Mean HG	SD	N	Mean Control	SD	N	Weight	Mean diff (95% CI)
Pre treatment								
Mantysaari 2004	93.5	6.95	32	92.4	5.08	32	0.0%	1.10 [-1.88, 4.08]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
0-1 year								
Finland	2.6	3.63	32	0.3	3.0	32	100.0%	2.30 [0.67, 3.93]
Subtotal (95% CI)			32			32	100.0%	2.30 [0.67, 3.93]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.76 (P = 0.006)								
0-2 years								
Finland	2.6	4.5	32	1.2	2.71	32	100.0%	1.40 [-0.42, 3.22]
Subtotal (95% CI)			32			32	100.0%	1.40 [-0.42, 3.22]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.51 (P = 0.13)								
Test for subgroup differences: Chi² = 0.52, df = 1 (P = 0.47), I² = 0%								

Table 7: Cervical pull headgear versus control, outcome: Upper incisor inclinational change

Time point	Mean HG	SD	N	Mean Control	SD	N	Weight	Mean diff (95% CI)
Pre-treatment								
Mantysaari 2004	108.4	6.5	32	105.8	6.07	32	0.0%	2.60 [-0.48, 5.68]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
0-1 year								
Finland	4.8	3.62	32	0.8	4.6	32	100.0%	4.00 [1.97, 6.03]
Subtotal (95% CI)			32			32	100.0%	4.00 [1.97, 6.03]

Heterogeneity: Not applicable								
Test for overall effect: Z = 3.87 (P = 0.0001)								
0-2 years								
Finland	5.7	6.63	32	1.2	6.17	32	100.0%	4.50 [1.36, 7.64]
Subtotal (95% CI)			32			32	100.0%	4.50 [1.36, 7.64]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.81 (P = 0.005)								
Test for subgroup differences: Chi² = 0.07, df = 1 (P = 0.79), I² = 0%								

5.3.2: Comparison 2: Extraction of lower deciduous canines versus control

Crowding

One study investigated lower incisor crowding in the lower labial segment, in millimetres, at one-year post-treatment period (Kau 2004[60]). The time points at which crowding was measured were: baseline and one-year post-treatment. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, ninety-seven participants began the study and provided information on crowding at baseline, whilst eighty-three (85.6%) participants attended the one-year recall.

We were unable to perform a meta-analysis, as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 10). This revealed:

- Statistically significantly less lower incisor crowding (1.8mm) in the non-extraction group at baseline (95% CI 0.39- 3.21, P= 0.01)
- Statistically significantly less lower incisor crowding (2.96mm) in the extraction group at 1 year (95% CI -4.82, -1.1, P= 0.002).
- Statistically significantly greater reduction (4.76mm) in the lower incisor crowding in the extraction group between baseline and 1-year (95% CI -6.24, -3.28, P= 0.00001).
- Statistically significant difference between the subgroups at different time points, favouring extraction (P< 0.0001).

Time to alignment

This outcome was not relevant for this comparison.

Ligation time

This outcome was not relevant for this comparison.

Arch length

One study reported on this outcome (Kau 2004[60]). A total group analysis was carried out (Figure 11) and it revealed:

- The non-extraction group had 2.73mm longer arch length compared to the extraction group (95% CI 1.77-3.69, P< 0.00001), which was statistically significantly greater.

Lower incisors to mandible

One study reported on lower incisor inclination, in degrees, for up to one-year post-treatment (Kau 2004[60]). The time points at which lower incisor change was measured were at baseline and one-year post-treatment. This study was assessed as being at high risk of bias overall as blinding of personnel and participants was not possible. In total, ninety-seven participants began the study and provided information on crowding at baseline, eighty-three participants presented at the one-year recall.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 12). This revealed:

- No statistically significant difference in the change in lower incisor inclinational between baseline and one-year post-treatment, for any of the lower incisors (Table 9)
- Overall, there was no statistical difference in the change in lower incisor inclinational between either subgroup (P= 0.72)

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 10: Forest plot of comparison 2: Extraction of lower deciduous canines versus control, outcome: Lower incisor crowding

Figure 11: Forest plot of comparison 2: Extraction of lower deciduous canines versus control, outcome: Mandibular arch length change

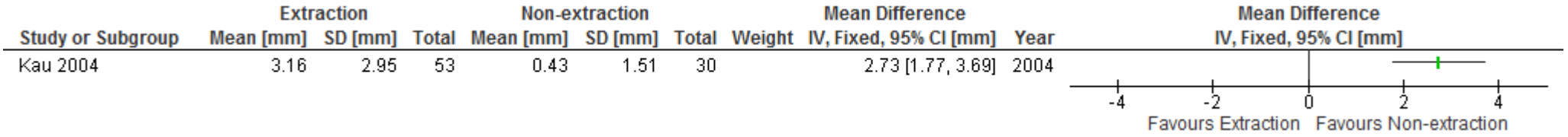


Figure 12: Forest plot of comparison 2: Extraction of lower deciduous canines versus control, outcome: Lower incisor inclination

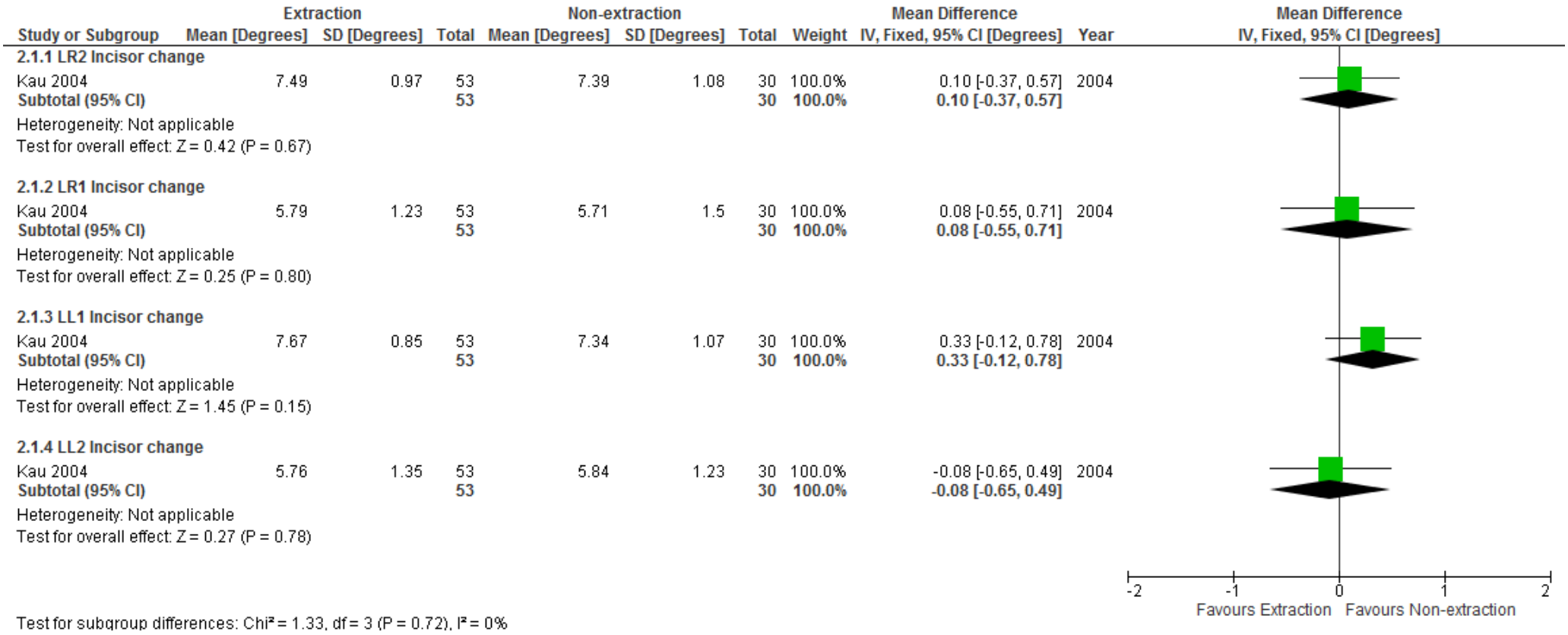


Table 8: Extraction of lower deciduous canines versus control, outcome: Lower incisor crowding

Time point	Mean Extraction	SD	N	Mean non-extraction	SD	N	Weight	Mean diff (95% CI)
Pre-treatment								
Kau 2004	11.82	4.12	55	10.02	2.95	42	100.0%	1.80 [0.39, 3.21]
Subtotal (95% CI)			55			42	100.0%	1.80 [0.39, 3.21]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.51 (P = 0.01)								
1 year								
Kau 2004	5.79	4.74	53	8.75	3.78	30	100.0%	-2.96 [-4.82, -1.10]
Subtotal (95% CI)			53			30	100.0%	-2.96 [-4.82, -1.10]
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.12 (P = 0.002)								
Change between baseline and 1 year								
Kau 2004	-6.03	4.44	53	-1.27	2.44	30	0.0%	-4.76 [-6.24, -3.28]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Test for subgroup differences: Chi² = 16.00, df = 1 (P < 0.0001), I² = 93.8%								

Table 9: Extraction of lower deciduous canines versus control, outcome: Mandibular arch length change

Time point	Mean Extraction	SD	N	Mean non-extraction	SD	N	Weight	Mean diff (95% CI)
Baseline to 1 year	3.16	2.95	53	0.43	1.51	30	100.0%	2.73 [1.77, 3.69]
Total (95% CI)			53			30	100.0%	2.73 [1.77, 3.69]
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.57 (P < 0.00001)								

Table 10: Extraction of lower deciduous canines versus control, outcome: Lower incisor inclination

Incisor	Mean Extraction	SD	N	Mean non-extraction	SD	N	Weight	Mean diff (95% CI)
2.1.1 LR2 Incisor change								
Kau 2004	7.49	0.97	53	7.39	1.08	30	100.0%	0.10 [-0.37, 0.57]
Subtotal (95% CI)			53			30	100.0%	0.10 [-0.37, 0.57]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.42 (P = 0.67)								
2.1.2 LR1 Incisor change								
Kau 2004	5.79	1.23	53	5.71	1.5	30	100.0%	0.08 [-0.55, 0.71]
Subtotal (95% CI)			53			30	100.0%	0.08 [-0.55, 0.71]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.25 (P = 0.80)								
2.1.3 LL1 Incisor change								
Kau 2004	7.67	0.85	53	7.34	1.07	30	100.0%	0.33 [-0.12, 0.78]
Subtotal (95% CI)			53			30	100.0%	0.33 [-0.12, 0.78]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.45 (P = 0.15)								
2.1.4 LL2 Incisor change								
Kau 2004	5.76	1.35	53	5.84	1.23	30	100.0%	-0.08 [-0.65, 0.49]
Subtotal (95% CI)			53			30	100.0%	-0.08 [-0.65, 0.49]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.27 (P = 0.78)								
Test for subgroup differences: Chi^2 = 1.33, df = 3 (P = 0.72), I^2 = 0%								

5.3.3: Comparison 3: Schwarz versus control

Crowding

One study investigated lower arch crowding, in millimetres, over a nine-month post-treatment follow-up period (Tai 2010[66]). The time points at which crowding was measured were baseline and nine months after expansion of the arches with a Schwarz appliance for six months. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, twenty-eight participants were included in this study.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 13). This revealed:

- Baseline equivalence in crowding with no statistically significant difference between the expansion (Schwarz) and control groups (P= 0.48).
- Statistically significantly less lower arch crowding (2.39mm) at the nine-month recall (95% CI -3.15, -1.63, P< 0.00001) in the expansion group.
- Statistically significantly more reduction in lower arch crowding (2.14mm) in the expansion group (95% CI -2.79, -1.48, P<0.00001).
- Statistically significant difference in crowding between the subgroups (P= 0.0001).

Time to alignment

This outcome was not relevant for this comparison.

Ligation time

This outcome was not relevant for this comparison.

Arch length

One study (Tai 2010[66]) reported mandibular arch length, in millimetres, as mentioned above. Twenty-eight participants were included in this analysis.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 14). This revealed:

- Statistically significant difference in baseline arch length (1.86mm) with the expansion group having a longer pre-treatment arch length (95% CI 0.23, 3.49, P= 0.03).
- Post-treatment, the arch length of the expansion group remained statistically significantly longer (1.97mm) than the control group (95% CI 0.28, 3.66 P= 0.02).
- No statistically significant difference in the change in arch length (0.11mm) between baseline and follow-up (95% CI -0.46, 0.68, P= 0.71).
- Overall, there was no statistically significant difference between the subgroups (P= 0.93).

Lower incisors to mandible

One study reported on lower incisor position, in millimetres (Tai 2010 (66)). The time points at which lower incisor position was measured are as above. This study was assessed as being at high risk of bias overall as blinding of personnel and participants was not possible. In total, twenty-eight participants were included in the analysis.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 15). This revealed:

- Baseline equivalence in the lower incisor position of the groups (P= 0.89).
- No statistically significant difference in lower incisor position post-treatment (P= 0.57).
- Statistically significantly more advancement of the lower incisors in the expanded group (P= 0.006).
- Overall, there was no statistically significant difference in the change in lower incisor positional between either subgroup (P= 0.63).

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

One study reported on upper incisor position, in degrees (Tai 2010[66]). The time points and risk of bias have been discussed earlier. In total, sixty-four participants were included in the analysis.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 16). This revealed:

- Baseline equivalence between the groups with no statistical difference in the incisor inclination (P= 0.94).
- No statistically significant difference between the expansion and non-expansion groups for upper incisor inclination at follow-up (P= 0.85).
- Overall, there was no statistically significant difference in the change in upper incisor inclination in each group (P= 0.8).
- No statistically significant difference between the subgroups (P= 0.95).

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 13: Forest plot of comparison 3: Schwarz appliance versus control, outcome: Lower incisor crowding

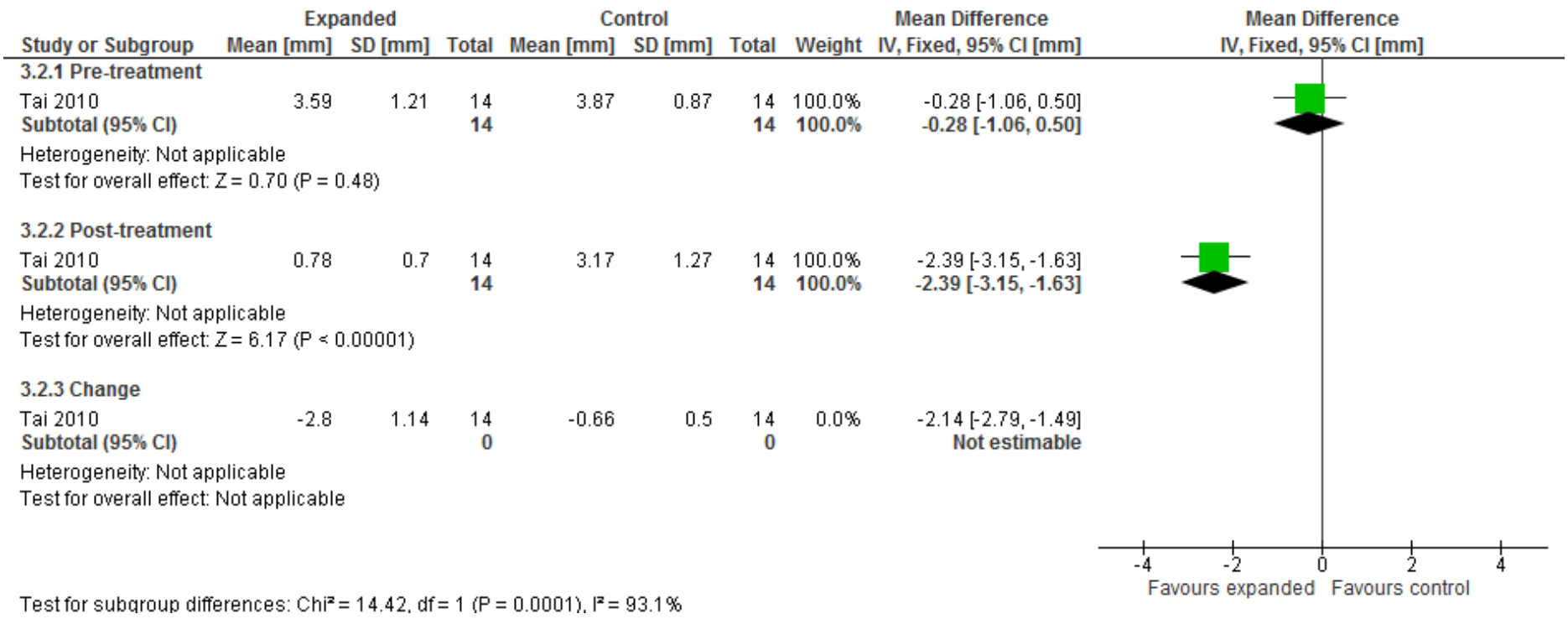


Figure 14: Forest plot of comparison 3: Schwarz versus control, outcome: Arch length

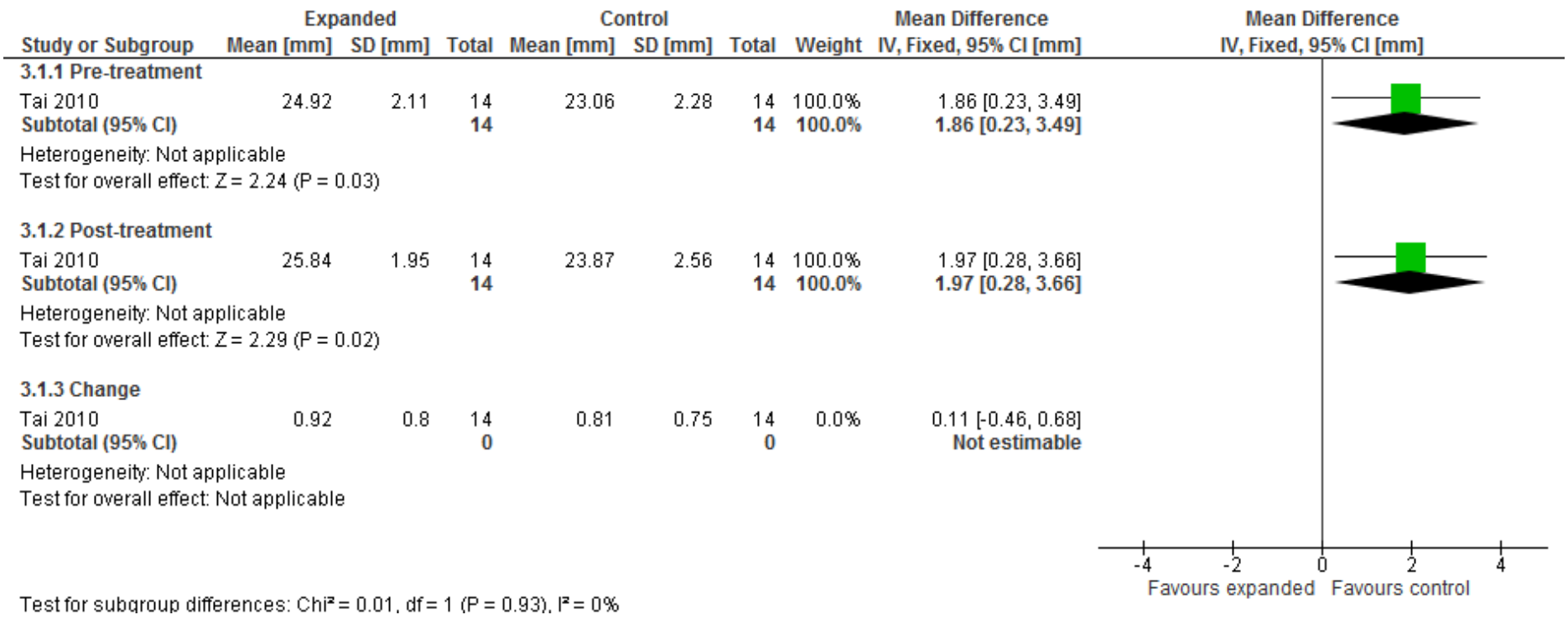


Figure 15: Forest plot of comparison 3: Schwarz versus control, outcome: Lower incisor to APog

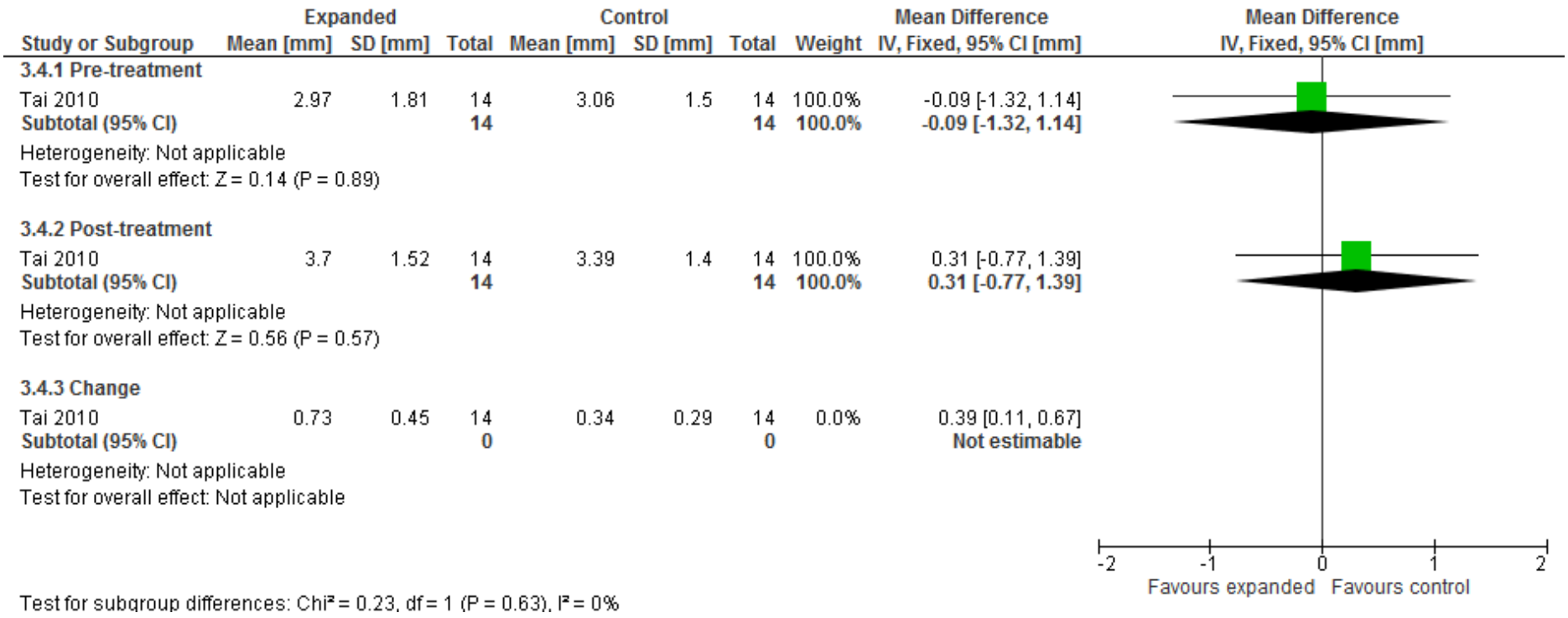


Figure 16: Forest plot of comparison 3: Schwarz versus control, outcome: Upper incisor to SN

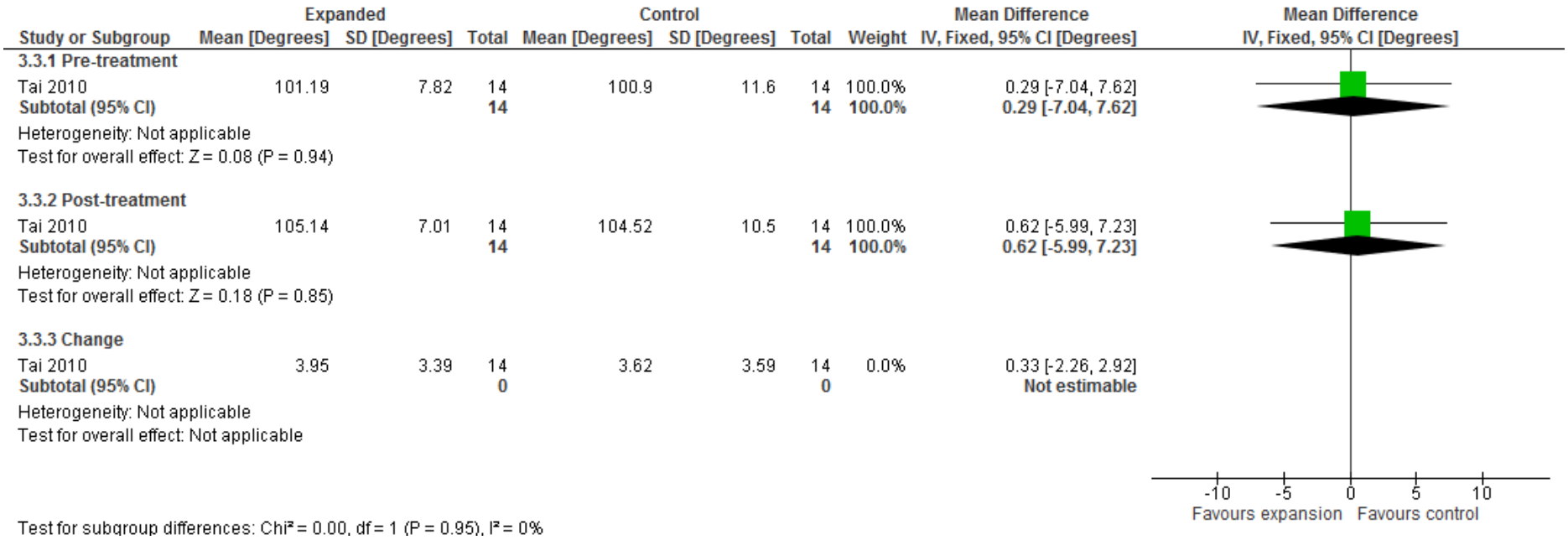


Table 11: Schwarz appliance versus control, outcome: Lower incisor crowding

Time point	Mean expansion	SD	N	Mean non-expansion	SD	N	Weight	Mean diff (95% CI)
Pre-treatment								
Tai 2010	3.59	1.21	14	3.87	0.87	14	100.0%	-0.28 [-1.06, 0.50]
Subtotal (95% CI)			14			14	100.0%	-0.28 [-1.06, 0.50]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.70 (P = 0.48)								
Post-treatment								
Tai 2010	0.78	0.7	14	3.17	1.27	14	100.0%	-2.39 [-3.15, -1.63]
Subtotal (95% CI)			14			14	100.0%	-2.39 [-3.15, -1.63]
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.17 (P < 0.00001)								
Change								
Tai 2010	-2.8	1.14	14	-0.66	0.5	14	0.0%	-2.14 [-2.79, -1.49]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Test for subgroup differences: Chi² = 14.42, df = 1 (P = 0.0001), I² = 93.1%								

Table12: Schwarz versus control, outcome: Arch length

Time point	Mean expansion	SD	N	Mean non-expansion	SD	N	Weight	Mean diff (95% CI)
3.1.1 Pre-treatment								
Tai 2010	24.92	2.11	14	23.06	2.28	14	100.0%	1.86 [0.23, 3.49]
Subtotal (95% CI)			14			14	100.0%	1.86 [0.23, 3.49]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.24 (P = 0.03)								
3.1.2 Post-treatment								
Tai 2010	25.84	1.95	14	23.87	2.56	14	100.0%	1.97 [0.28, 3.66]
Subtotal (95% CI)			14			14	100.0%	1.97 [0.28, 3.66]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.29 (P = 0.02)								
3.1.3 Change								

Tai 2010	0.92	0.8	14	0.81	0.75	14	0.0%	0.11 [-0.46, 0.68]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.93), I² = 0%								

Table 13: Schwarz versus control, outcome: Lower incisor to APog

Time point	Mean expansion	SD	N	Mean non-expansion	SD	N	Weight	Mean diff (95% CI)
3.4.1 Pre-treatment								
Tai 2010	2.97	1.81	14	3.06	1.5	14	43.6%	-0.09 [-1.32, 1.14]
Subtotal (95% CI)			14			14	43.6%	-0.09 [-1.32, 1.14]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.14 (P = 0.89)								
3.4.2 Post-treatment								
Tai 2010	3.7	1.52	14	3.39	1.4	14	56.4%	0.31 [-0.77, 1.39]
Subtotal (95% CI)			14			14	56.4%	0.31 [-0.77, 1.39]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.56 (P = 0.57)								
3.4.3 Change								
Tai 2010	0.73	0.45	14	0.34	0.29	14	0.0%	0.39 [0.11, 0.67]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			28			28	100.0%	0.14 [-0.68, 0.95]
Heterogeneity: Chi² = 0.23, df = 1 (P = 0.63); I² = 0%								
Test for overall effect: Z = 0.33 (P = 0.74)								
Test for subgroup differences: Chi² = 0.23, df = 1 (P = 0.63), I² = 0%								

Table 14: Schwarz versus control, outcome: Upper incisor to SN

Time point	Mean expansion	SD	N	Mean non-expansion	SD	N	Weight	Mean diff (95% CI)
3.3.1 Pre-treatment								
Tai 2010	101.19	7.82	14	100.9	11.6	14	44.9%	0.29 [-7.04, 7.62]
Subtotal (95% CI)			14			14	44.9%	0.29 [-7.04, 7.62]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.08 (P = 0.94)								
3.3.2 Post-treatment								
Tai 2010	105.14	7.01	14	104.52	10.5	14	55.1%	0.62 [-5.99, 7.23]
Subtotal (95% CI)			14			14	55.1%	0.62 [-5.99, 7.23]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.18 (P = 0.85)								
3.3.3 Change								
Tai 2010	3.95	3.39	14	3.62	3.59	14	0.0%	0.33 [-2.26, 2.92]
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			28			28	100.0%	0.47 [-4.44, 5.38]
Heterogeneity: $Chi^2 = 0.00$, df = 1 (P = 0.95); $I^2 = 0\%$								
Test for overall effect: Z = 0.19 (P = 0.85)								
Test for subgroup differences: $Chi^2 = 0.00$, df = 1 (P = 0.95), $I^2 = 0\%$								

5.3.4: Comparison 4: Eruption Guidance appliance versus control

Crowding

One study investigated incisor crowding in the maxilla and the mandible, in millimetres, for a one-year follow-up period (Myrlund 2015[67]). The time points at which crowding was measured were: baseline and one-year post-treatment. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, forty-six participants began the study and provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figures 17 and 18). This revealed:

- Baseline equivalence in crowding between the Eruption Guidance appliance (EGA) and control groups for upper and lower anterior crowding, with no statistically significant difference between them (P= 0.15 in maxilla, P= 0.26 in mandible).
- In the maxilla:
    - No statistically significant difference in crowding between the EGA and control groups at one-year post-treatment (P= 0.74).
- No statistically significant difference in crowding between either subgroup (P= 0.21).
- In the mandible:
    - Statistically significantly fewer participants with crowding in the EGA group at one-year post-treatment, (odds ratio 0.19, 95% CI 0.05-0.68, P= 0.01).
    - Statistically significant difference in crowding between subgroups (P= 0.008).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant for this comparison.

Arch length

This outcome was not reported

Lower incisors to mandible

One study reported on lower incisor inclination to the mandibular plane, in degrees, for up to one-year post-treatment (Myrlund 2015[67]). The time points at which lower incisor inclination was measured were at baseline and one-year post-treatment. This study was assessed as being at high risk of bias overall as blinding of personnel and participants was not possible. In total, forty-six participants were included in the analysis. We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figures 19). This revealed:

- Statistically significant proclination of the lower incisors in the EGA group between baseline and one-year post-treatment, of 4.1 degrees (95% CI -6.99, -1.21, P= 0.005).

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

One study reported on upper incisor inclination, in degrees, over a two-year period (Myrlund 2015[67]). We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 20). This revealed:

- No statistically significant difference in upper incisor between pre and post-treatment (P= 0.80).

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 17: Forest plot of comparison 4: Eruption Guidance appliance versus control, outcome: Maxillary crowding

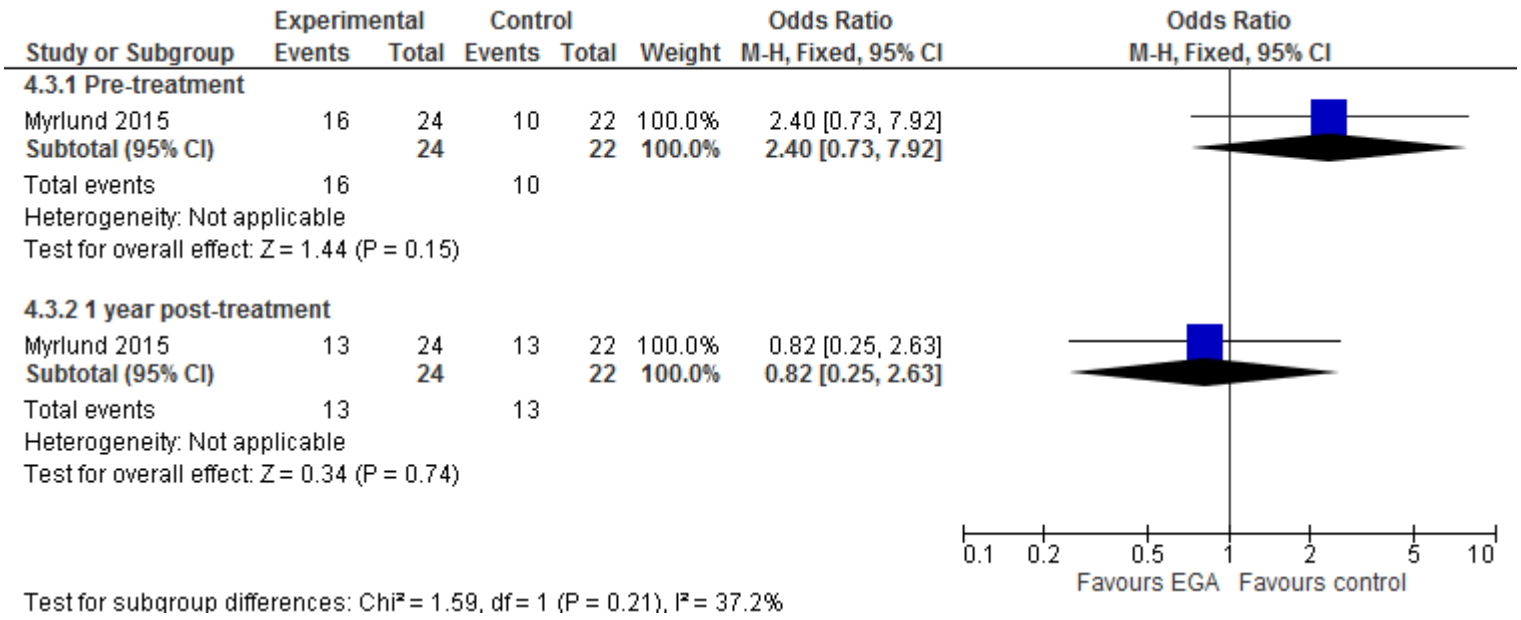


Figure 18: Forest plot of comparison 4: Eruption Guidance appliance versus control, outcome: Mandibular crowding

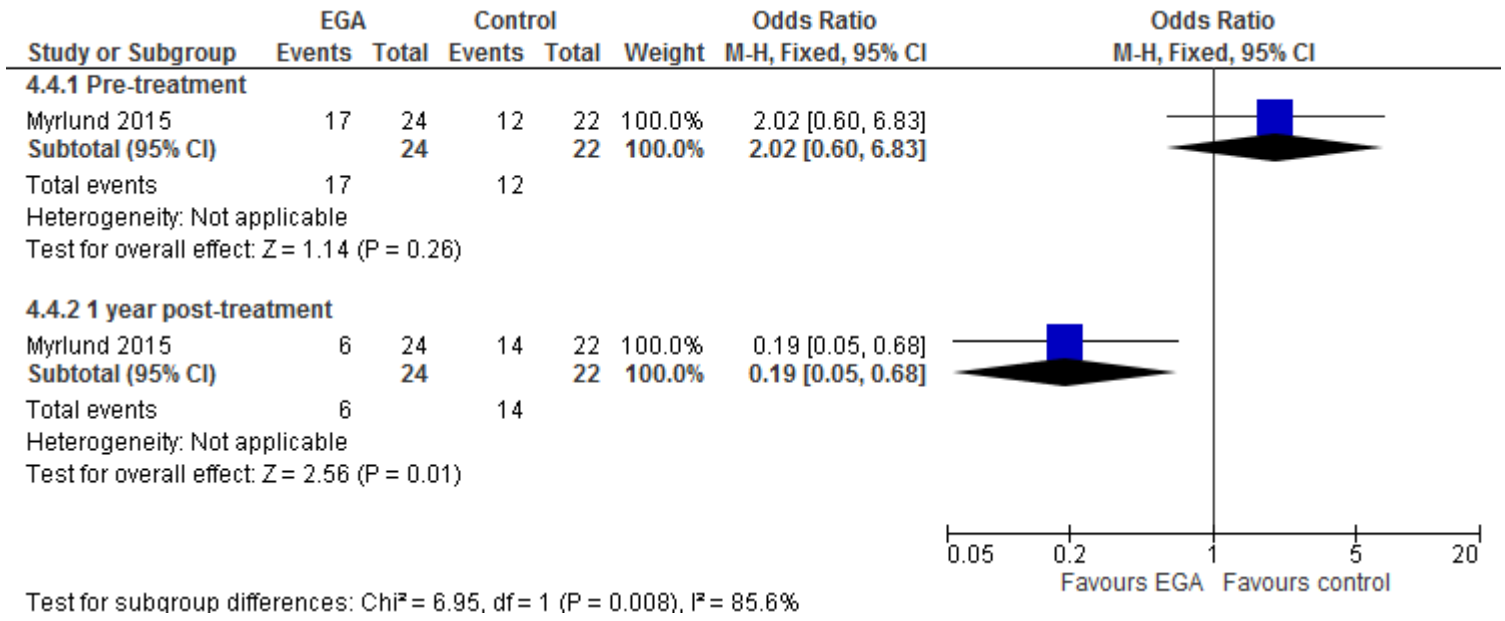


Figure 19: Forest plot of comparison 4: Eruption Guidance appliance, pre-treatment versus 1 year post-treatment, outcome: Lower incisor to mandibular plane

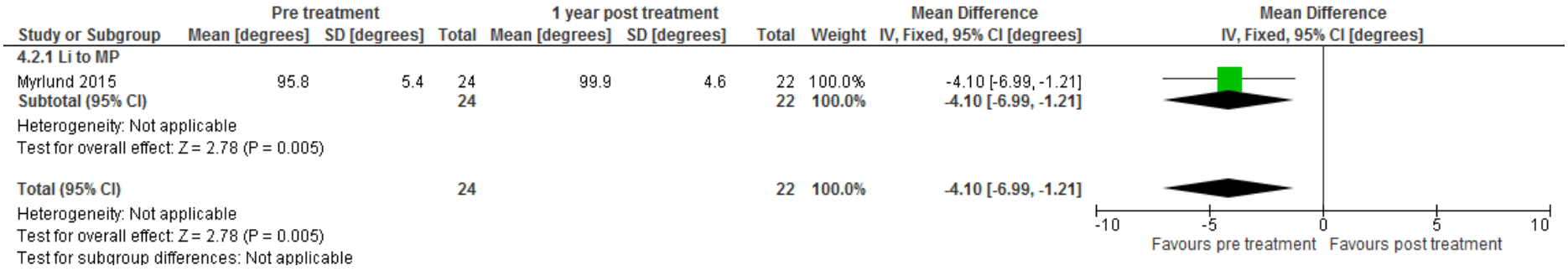


Figure 20: Forest plot of comparison 4: Eruption Guidance appliance, pre-treatment versus 1 year post-treatment, outcome: Upper incisor to SN

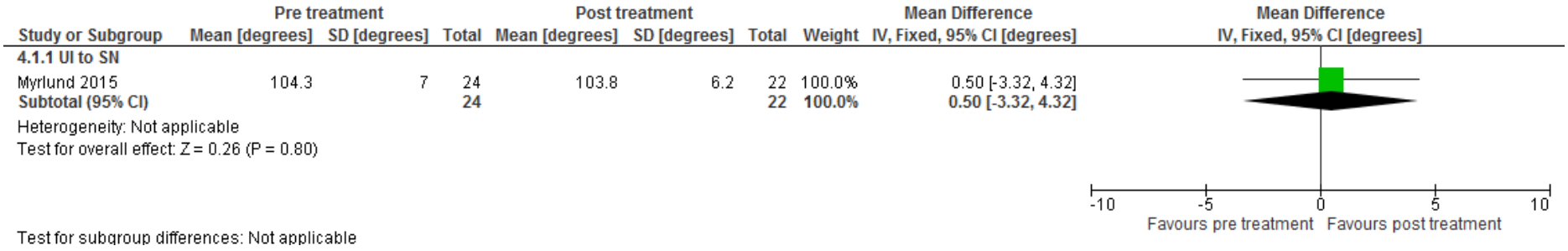


Table 15: Eruption Guidance appliance versus control, outcome: Maxillary crowding

Time point	EGA		Control		Weight	Odds ratio (95% CI)
	Events	Total	Events	Total		
Pre-treatment						
Myrlund 2015	16	24	10	22	100.0%	2.40 [0.73, 7.92]
Subtotal (95% CI)		24		22	100.0%	2.40 [0.73, 7.92]
Total events	16		10			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.44 (P = 0.15)						
1-year post-treatment						
Myrlund 2015	13	24	13	22	100.0%	0.82 [0.25, 2.63]
Subtotal (95% CI)		24		22	100.0%	0.82 [0.25, 2.63]
Total events	13		13			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.34 (P = 0.74)						
Test for subgroup differences: Chi² = 1.59, df = 1 (P = 0.21), I² = 37.2%						

Table 16: Eruption Guidance appliance versus control, outcome: Mandibular crowding

Time point	EGA		Control		Weight	Odds ratio (95% CI)
	Events	Total	Events	Control		
Pre-treatment						
Myrlund 2015	17	24	12	22	100.0%	2.02 [0.60, 6.83]
Subtotal (95% CI)		24		22	100.0%	2.02 [0.60, 6.83]
Total events	17		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.14 (P = 0.26)						
1-year post-treatment						
Myrlund 2015	6	24	14	22	100.0%	0.19 [0.05, 0.68]
Subtotal (95% CI)		24		22	100.0%	0.19 [0.05, 0.68]
Total events	6		14			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.56 (P = 0.01)						
Test for subgroup differences: Chi² = 6.95, df = 1 (P = 0.008), I² = 85.6%						

Table 17: Eruption Guidance appliance versus control, outcome: Lower incisor to mandibular plane

Li to MP	Mean	SD	N	Mean	SD	N	Weight	Mean difference (95% CI)
Myrlund 2015	95.8	5.4	24	99.9	4.6	22	100.0%	-4.10 [-6.99, -1.21]
Subtotal (95% CI)			24			22	100.0%	-4.10 [-6.99, -1.21]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.78 (P = 0.005)								
Total (95% CI)			24			22	100.0%	-4.10 [-6.99, -1.21]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.78 (P = 0.005)								

Table 18: Eruption Guidance appliance versus control, outcome: Upper incisor to SN

UI to SN	Mean	SD	N	Mean	SD	N	Weight	Mean difference (95% CI)
Myrlund 2015	104.3	7.0	24	103.8	6.2	22	100.0%	0.50 [-3.32, 4.32]
Subtotal (95% CI)			24			22	100.0%	0.50 [-3.32, 4.32]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.26 (P = 0.80)								

5.3.5: Comparison: Lower Lingual Arch versus control

Crowding

This outcome was not reported

Time to alignment

This outcome was not relevant for this comparison.

Ligation time

This outcome was not relevant for this comparison.

Arch length

One study investigated arch length change in the mandible, in millimetres, up to a one-year follow-up period (Rebellato 1997[63]). The time points at which crowding was measured were: baseline and at ten to twelve months post-treatment. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, thirty participants provided information on arch length.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figure 21). This revealed:

- Statistically significant increase in arch length of 2.61mm (95% CI 1.83, 3.39 P < 0.00001) in the Lower Lingual Arch appliance (LLA) group.

Lower incisors to mandible

One study reported on the lower incisors to the mandibular plane, in degrees and in millimetres, for up to a one-year follow-up period (Rebellato 1997[63]). Details of the trial have been discussed above. We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figures 22 and 23). This revealed:

- The lower incisors moved mesially by 0.32mm in the LLA, compared to a 0.34mm distal movement in the control group and that this difference of 0.66mm was statistically significant (95% CI 0.46, 0.86, P < 0.00001).

- The lower incisors proclined by 0.73 degrees in the LLA, compared to 2.28 degrees of retroclination in the control group and that this difference of 3.01 degrees was statistically significant (95% CI 1.71, 4.31, P < 0.00001).

Lower molars to mandible

One study reported on the lower molar to the mandibular plane, in degrees and in millimetres, for up to a one-year follow-up period (Rebellato 1997[63]). Details of the trial have been discussed above. We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figures 24 and 25). This revealed:

- The lower molar moved mesially by 0.33mm in the LLA, compared to a 1.44mm mesial movement in the control group and that this difference of 1.11mm was statistically significant (95% CI -1.51, -0.71, P < 0.00001).
- The lower molars tipped distally by 0.54 degrees in the LLA, compared to 2.19 degrees of mesial tipping in the control group and that this difference of 2.73 degrees was statistically significant (95% CI -4.29, -1.17, P= 0.0006).

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 21: Forest plot of comparison 5: Lower Lingual appliance versus control, outcome: Arch length

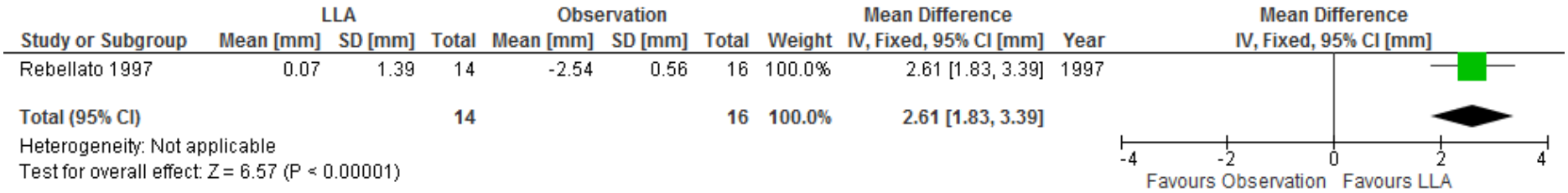


Figure 22: Forest plot of comparison 5: Lower Lingual appliance versus control, outcome: Lower incisor A-P positional change

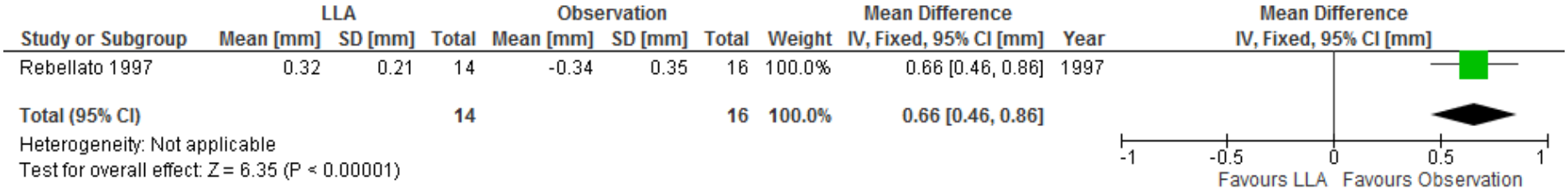


Figure 23: Forest plot of comparison 5: Lower Lingual appliance versus control, outcome: Lower incisor inclinational change

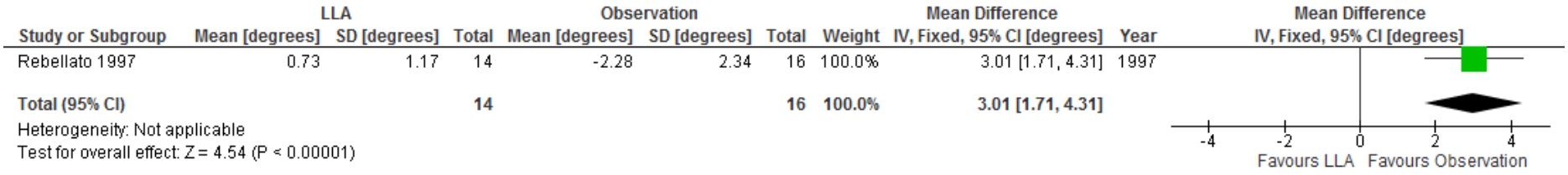


Figure 24: Forest plot of comparison 5: Lower Lingual appliance versus control, outcome: Lower molar A-P positional change

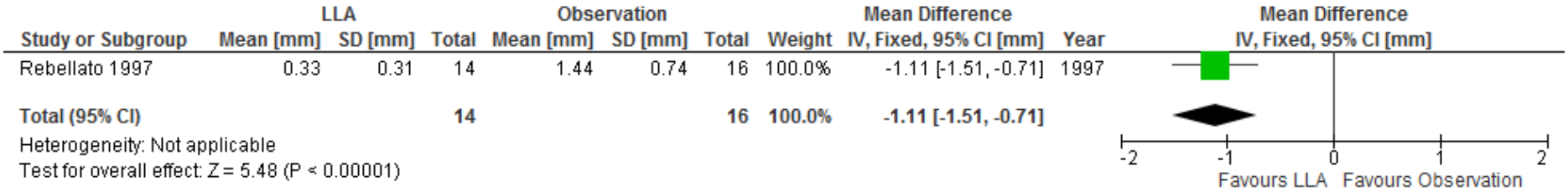


Figure 25: Forest plot of comparison 5: Lower Lingual appliance versus control, outcome: Lower molar inclinational change

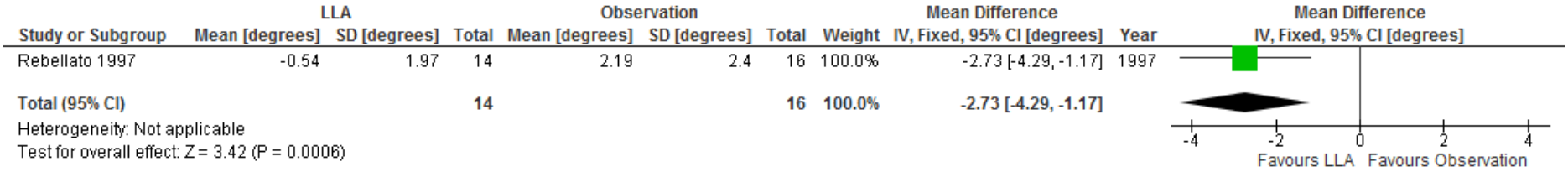


Table 19: Lower Lingual appliance versus control, outcome: Arch length

Arch length change (baseline to follow-up)	LLA			Control				
	Mean length	SD	N	Mean length	SD	Total	Weight	Mean difference 95% CI
Rebellato 1997	0.07	1.39	14	-2.54	0.56	16	100.0%	2.61 [1.83, 3.39]
Total (95% CI)			14			16	100.0%	2.61 [1.83, 3.39]
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.57 (P < 0.00001)								

Table 20: Lower Lingual appliance versus control, outcome: Lower incisor A-P positional change

Lower incisor A-P position change (baseline to follow-up)	LLA			Control				
	Mean LI positional change	SD	N	Mean LI positional change	SD	Total	Weight	Mean difference 95% CI
Rebellato 1997	0.32	0.21	14	-0.34	0.35	16	100.0%	0.66 [0.46, 0.86]
Total (95% CI)			14			16	100.0%	0.66 [0.46, 0.86]
Heterogeneity: Not applicable								
Test for overall effect: Z = 6.35 (P < 0.00001)								

Table 21: Lower Lingual appliance versus control, outcome: Lower incisor inclinational change

Lower incisor inclinational change (baseline to follow-up)	LLA			Control				
	Mean LI angular change	SD	N	Mean LI angular change	SD	Total	Weight	Mean difference 95% CI
Rebellato 1997	0.73	1.17	14	-2.28	2.34	16	100.0%	3.01 [1.71, 4.31]
Total (95% CI)			14			16	100.0%	3.01 [1.71, 4.31]
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.54 (P < 0.00001)								

Table 22: Lower Lingual appliance versus control, outcome: Lower molar A-P positional change

Lower molar A-P position change (baseline to follow-up)	LLA			Control				
	Mean Lower Molar positional change	SD	N	Mean Lower Molar positional change	SD	Total	Weight	Mean difference 95% CI
Rebellato 1997	0.33	0.31	14	1.44	0.74	16	100.0%	-1.11 [-1.51, -0.71]
Total (95% CI)			14			16	100.0%	-1.11 [-1.51, -0.71]
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.48 (P < 0.00001)								

Table 23: Lower Lingual appliance versus control, outcome: Lower molar inclinational change

Lower molar inclinational change (baseline to follow-up)	LLA			Control				
	Mean Lower Molar angular change	SD	N	Mean Lower Molar angular change	SD	Total	Weight	Mean difference 95% CI
Rebellato 1997	-0.54	1.97	14	2.19	2.4	16	100.0%	-2.73 [-4.29, -1.17]
Total (95% CI)			14			16	100.0%	-2.73 [-4.29, -1.17]
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.42 (P = 0.0006)								

5.3.6: Comparison: Lower Lip Bumper versus control

Crowding

One study investigated change in mandibular incisor crowding, in millimetres, for a six-month follow-up period (Davidovitch 1997[87]). The time points at which crowding was measured were: baseline and six months into treatment. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, thirty-four participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the different time points was carried out (Figure 26). This revealed:

- Statistically significantly greater reduction in lower incisor crowding (4.39mm) in the lip bumper group, when compared to the control group, at six months, (95% CI -5.07, -3.71, P< 0.00001).

Time to alignment

This outcome was not relevant for this comparison.

Ligation time

This outcome was not relevant for this comparison.

Arch length

One study investigated arch length change in the mandible, in millimetres, up to a six-month follow-up period (Davidovitch 1997[87]). The time points at which crowding was measured were: baseline and six months into treatment. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, thirty-four participants provided information on arch length.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figure 27). This revealed:

- Statistically significantly greater increase in arch length (3.34mm) in the lip bumper group (95% CI 2.71, 3.97, P < 0.00001) when compared with the control group, at six months.

Lower incisors to mandible

One study reported on the relationship on the lower incisors to the mandible, in degrees to the mandibular plane and in millimetres to A-Pogonion, for up to a six month follow-up period (Davidovitch 1997[87]). Details of the trial have been discussed above. We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figures 28 and 29). This revealed:

- Statistically significantly less labial movement of the lower incisors (0.49mm) between baseline and six months, in the control group (95% CI 0.09, 0.89).
- The lower incisors proclined 3.14 degrees more in the lip bumper group than the control, which was statistically significant (95% CI 1.73, 4.55, P < 0.0001) and favoured the control group.

Lower molars to mandible

One study reported on the lower molar to the mandibular plane, in degrees and in millimetres, for up to a six-month follow-up period (Davidovitch 1997[87]). Details of the trial have been discussed above. We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figures 30 and 31). This revealed:

- The lower molar moved distally by 0.61mm in the lip bumper group, compared to a 0.3mm mesial movement in the control group. This difference of 0.91mm was statistically significantly different (95% CI -1.58, -0.24, P= 0.008) and favoured the lip bumper group.
- The lower molars tipped distally by 3.38 degrees in the lip bumper group, compared to 0.75 degrees of mesial tipping in the control group. This difference of 4.14 degrees was statistically significantly different (95% CI -6.09, -2.17, P< 0.0001) and favoured the lip bumper group.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 26: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower incisor crowding change

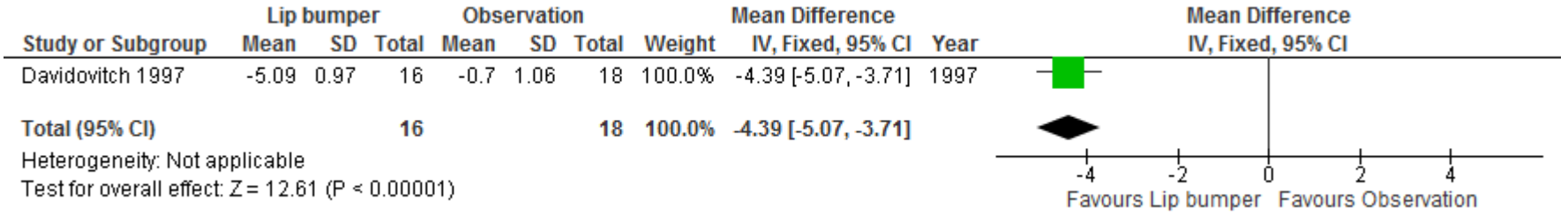


Figure 27: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower arch length change

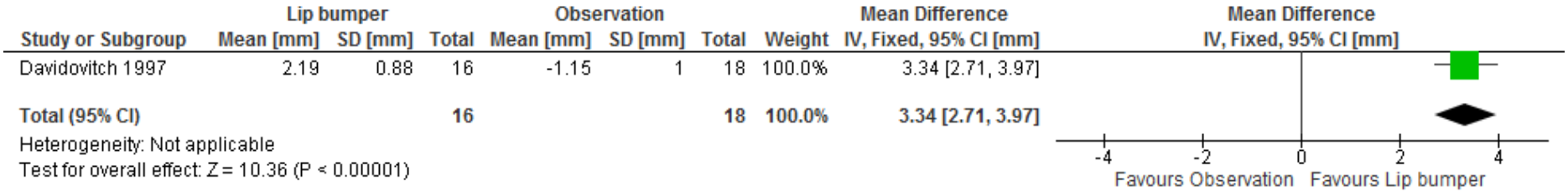


Figure 28: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower incisor A-P change

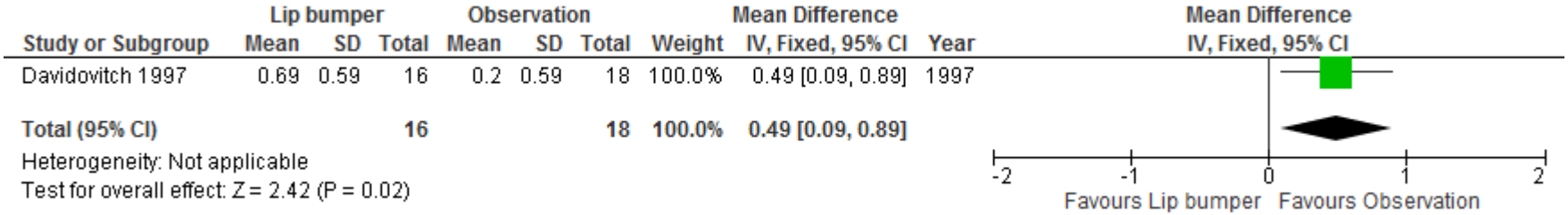


Figure 29: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower incisor inclinational change

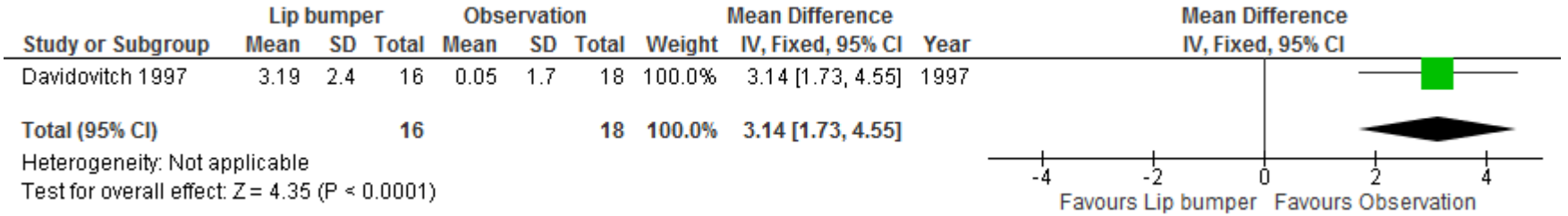


Figure 30: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower molar A-P change

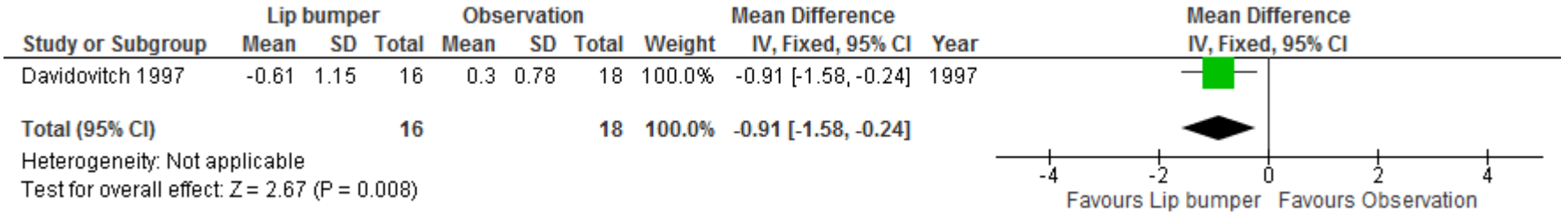


Figure 31: Forest plot of comparison 6: Lower Lip Bumper versus control, outcome: Lower molar inclinational change

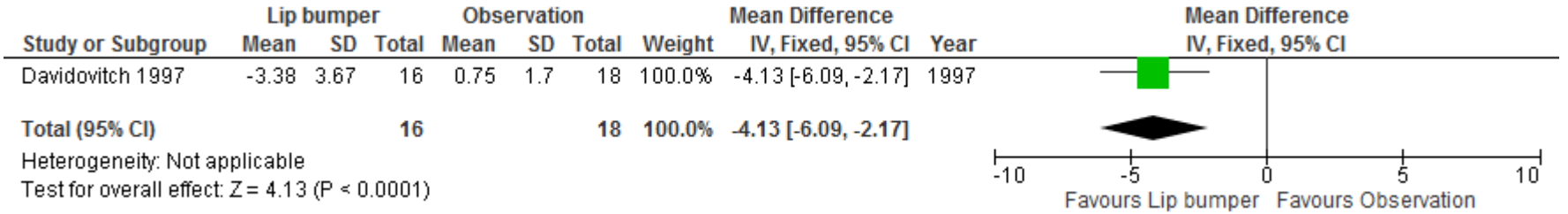


Table 24: Lower Lip Bumper versus control, outcome: Lower incisor crowding change

Lower incisor crowding change	Lip Bumper			Control				
	Mean LI crowding change	SD	N	Mean LI crowding change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	-5.09	0.97	16	-0.7	1.06	18	100.0%	-4.39 [-5.07, -3.71]
Total (95% CI)			16			18	100.0%	-4.39 [-5.07, -3.71]
Heterogeneity: Not applicable								
Test for overall effect: Z = 12.61 (P < 0.00001)								

Table 25: Lower Lip Bumper versus control, outcome: Lower arch length change

Lower arch length	Lip Bumper			Control				
	Mean lower arch length change	SD	N	Mean lower arch length change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	2.19	0.88	16	-1.15	1.0	18	100.0%	3.34 [2.71, 3.97]
Total (95% CI)			16			18	100.0%	3.34 [2.71, 3.97]
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.36 (P < 0.00001)								

Table 26: Lower Lip Bumper versus control, outcome: Lower incisor A-P positional change

Lower incisor A-P change	Lip Bumper			Control				
	Mean lower incisor A-P change	SD	N	Mean lower incisor A-P change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	0.69	0.59	16	0.2	0.59	18	100.0%	0.49 [0.09, 0.89]
Total (95% CI)			16			18	100.0%	0.49 [0.09, 0.89]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.42 (P = 0.02)								

Table 27: Lower Lip Bumper versus control, outcome: Lower incisor inclinational change

Lower incisor inclinational change	Lip Bumper			Control				
	Mean LI inclinational change	SD	N	Mean LI inclinational change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	3.19	2.4	16	0.05	1.7	18	100.0%	3.14 [1.73, 4.55]
Total (95% CI)			16			18	100.0%	3.14 [1.73, 4.55]
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.35 (P < 0.0001)								

Table 28: Lower Lip Bumper versus control, outcome: Lower molar A-P positional change

Lower molar A-P change	Lip Bumper			Control				
	Mean lower molar A-P change	SD	N	Mean lower molar A-P change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	-0.61	1.15	16	0.3	0.78	18	100.0%	-0.91 [-1.58, -0.24]
Total (95% CI)			16			18	100.0%	-0.91 [-1.58, -0.24]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.67 (P = 0.008)								

Table 29: Lower Lip Bumper versus control, outcome: Lower molar inclinational change

Lower molar inclinational change	Lip Bumper			Control				
	Mean LM inclination change	SD	N	Mean LM inclination change	SD	Total	Weight	Mean difference 95% CI
Davidovitch 1997	-3.38	3.67	16	0.75	1.7	18	100.0%	-4.13 [-6.09, -2.17]
Total (95% CI)			16			18	100.0%	-4.13 [-6.09, -2.17]
Heterogeneity: Not applicable								
Test for overall effect: Z = 4.13 (P < 0.0001)								

5.3.7: Comparison: Self-ligating brackets versus conventional brackets

Crowding

One study investigated incisor crowding in the anterior maxilla, in millimetres, for a ten week period (Miles 2010[82]). The time points at which crowding was measured were: baseline and ten weeks into treatment. This study was assessed as being at overall unclear risk of bias as blinding of participants was carried out, but blinding of personnel was not possible. The study states that participants were randomly allocated but no further details on the methods used were given. Additionally, there were dropouts in the conventional ligation group, so analysis was not performed on two of the self-ligating group. In total, sixty-eight participants provided baseline information and sixty (88.2%) were analysed at follow-up.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the different time points was carried out (Figure 32). This revealed:

- Baseline equivalence between the groups with regards to crowding (P= 0.90).
- No statistically significant difference in lower incisor crowding between the self-ligating and conventional groups at ten weeks (P= 0.94).
- No statistically significant difference between the subgroups (P= 0.72)

Time to alignment

One paper (Pandis 2011[79]) reported on time to ligation, in days, reporting up to one hundred and eighty-eight day follow-up period. The time points reported are the mean number of days it took for alignment in each group. Alignment is described as the point at which a rectangular copper nickel-titanium archwire could be placed passively. This study was considered to be at a high overall level of bias as it was not possible to blind participants and personnel as to which bracket type each participant received. In total, fifty participants provided information on time to alignment.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figure 33). This revealed:

- No statistically significant difference in time to alignment between the groups (P= 0.30)

Ligation time

One study reported on the time to tie, ligate and untie six brackets, in seconds (Miles 2010[82]). This study was assessed as being at overall unclear risk of bias, as the methods of randomisation and allocation concealment were not described, the participants were blinded but the clinicians were not and not all the participants who completed the study were analysed in order to keep the groups equal in size at the follow-up. In total, sixty-eight participants provided information on the time taken to untie and sixty participants provided information on the time taken to ligate, for six brackets.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between the untying and ligating was carried out (Figure 34). This revealed:

- Statistically significantly quicker (22.3 seconds) untying, in the self-ligating group (95% CI -25.83, -18.77 P < 0.00001).
- Statistically significant quicker (78.8 seconds) ligation, in the self-ligating group (95% CI -85.86, -75.74 P < 0.00001).
- Statistically significantly differences between the subgroups with untying being significantly quicker than ligation (P 0.00001).

Arch length

This outcome was not reported.

Lower incisors to mandible

One study reported on the lower incisors to the mandibular plane, in degrees (Atik 2014[80]) until the stage of treatment when a 0.019”x0.025” stainless steel archwire was placed. This study was assessed as being at overall high risk of bias as neither the participants nor the clinicians were blinded and the entire sample consisted of female participants. A total of thirty-three participants provided information for lower incisor inclination. We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between the groups was carried out (Figure 35). This revealed:

- No statistically significant baseline equivalence in lower incisor inclination despite 3.38 degrees less proclination in the conventional bracket group (95% CI -0.04, 6.8 P= 0.05).
- No statistically significant differences in the post-treatment lower incisor inclination despite the conventional bracket group being 4.53 degrees less proclined than the self-ligating group, (95% CI -0.02, 9.08 P= 0.05)
- No statistically significant difference in the change (1.29 degrees) in inclination between the self-ligating and conventional bracket groups (95% CI -1.77, 4.35 P= 0.41).

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

Two studies (Atik 2014[80], Miles 2010[82]) reported on discomfort and one reported on Plaque Index, Gingival Index and Probing depth (Atik 2014[80]).

Discomfort

Atik 2014 [80] described discomfort using a 100mm Visual Analogue Scale (VAS) over the first month. The participants were asked to keep a diary and record discomfort in the maxilla and mandible at 4 hours, 24 hours, 3 days, 1 week, and 1 month using the terms “very comfortable” and “very uncomfortable” at the ends of the scale. No statistically significant difference was found in the study for discomfort scores between self-ligating and conventional brackets.

Miles 2010 [82] described discomfort using a 7 point Likert Scale for the first week. The participants were given a questionnaire and asked to record discomfort in the upper arch at 4 hours, 24 hours, 3 days and 1 week. Again, there was no statistically significant difference for discomfort scores between self-ligating and conventional brackets.

Periodontal Index, Gingival Index, Probing depth

Atik 2014 [80] described the periodontal and gingival health of all 24 maxillary and mandibular teeth and estimated the mean value per participant. They did not find any statistically significant differences in any of the scores between self-ligating and conventional brackets, from baseline measurement to the end of the study.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 32: Forest plot of comparison 7: Self-ligating brackets versus Conventional brackets, outcome: Upper incisor crowding

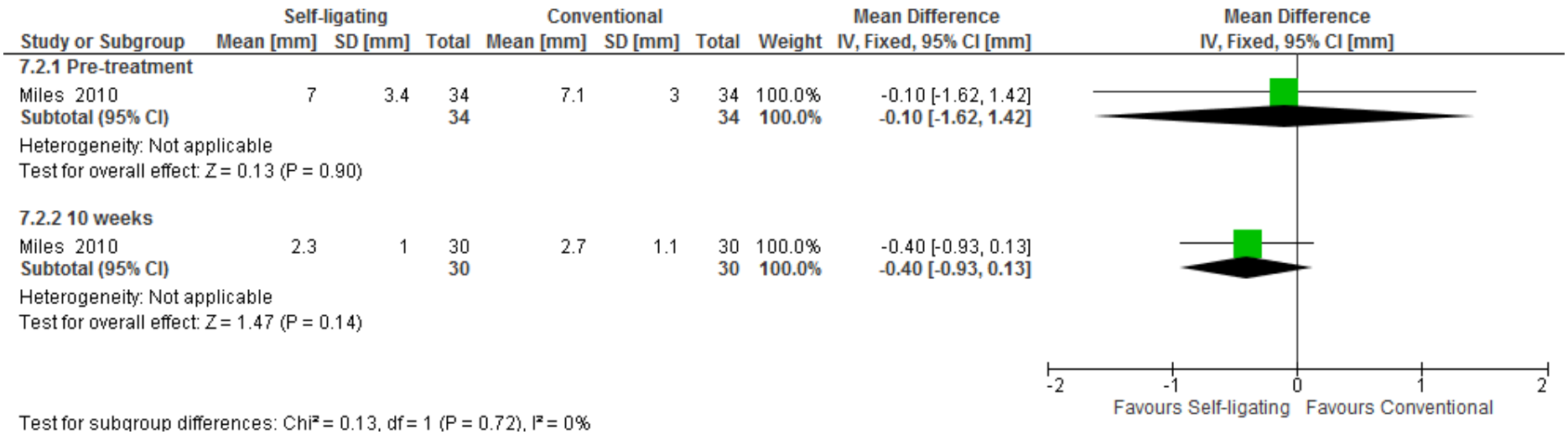


Figure 33: Forest plot of comparison 7: Self-ligating brackets versus Conventional brackets, outcome: Time to alignment

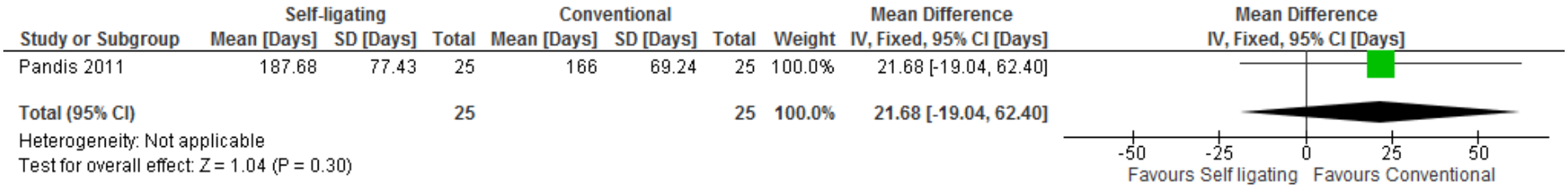


Figure 34: Forest plot of comparison 7: Self-ligating brackets versus Conventional brackets, outcome: Ligation time

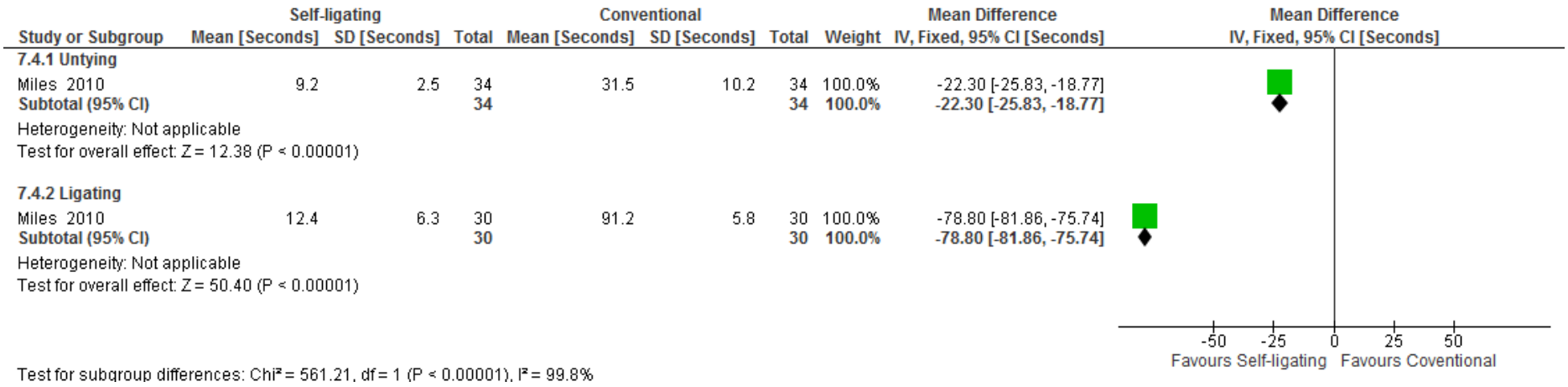


Figure 35: Forest plot of comparison 7: Self-ligating brackets versus Conventional brackets, outcome: Lower incisor inclination

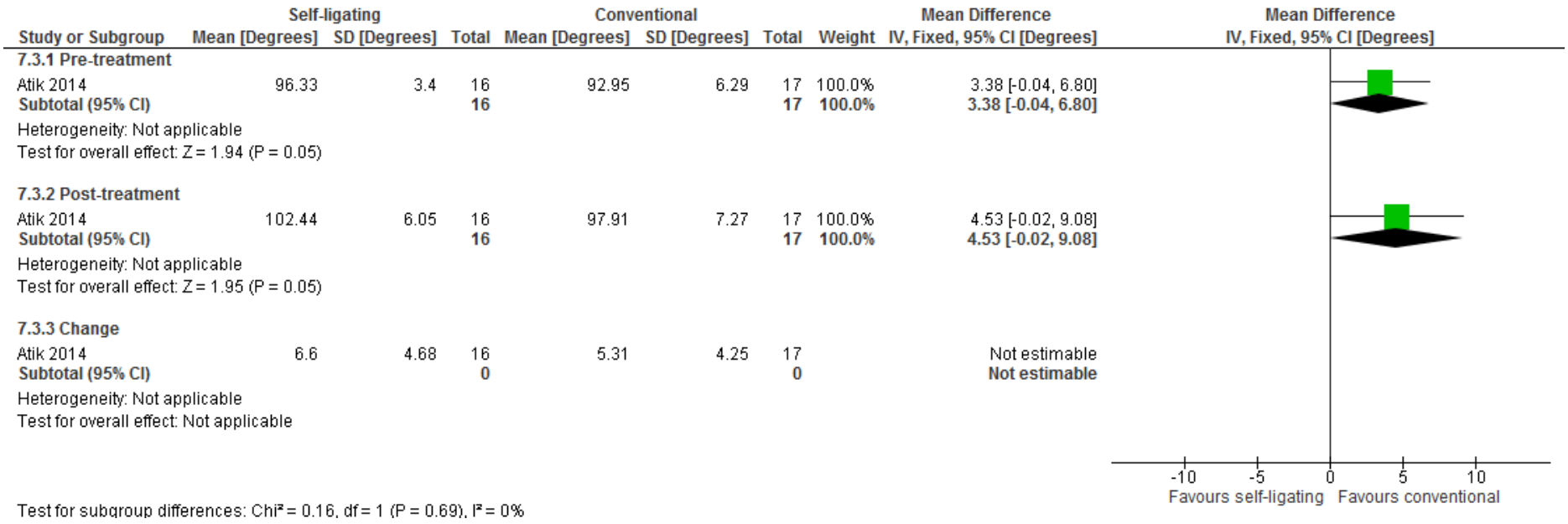


Table 30: Self-ligating brackets versus Conventional brackets, outcome: Upper incisor crowding

	Self Ligating			Conventional				
Time point	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (95% CI)
Pre-treatment								
Miles 2010	7.0	3.4	34	7.1	3.0	34	100.0%	-0.10 [-1.62, 1.42]
Subtotal (95% CI)			34			34	100.0%	-0.10 [-1.62, 1.42]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.13 (P = 0.90)								
10 weeks								
Miles 2010	2.3	1.0	30	2.7	1.1	30	100.0%	-0.40 [-0.93, 0.13]
Subtotal (95% CI)			30			30	100.0%	-0.40 [-0.93, 0.13]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.47 (P = 0.14)								
Test for subgroup differences: Chi² = 0.13, df = 1 (P = 0.72), I² = 0%								

Table 31: Self-ligating brackets versus Conventional brackets, outcome: Time to alignment

Days to alignment	**Self Ligating**			**Conventional**				
	Mean days	SD	N	Mean days	SD	N	Weight	Mean difference (95% CI)
Pandis 2011	187.68	77.43	25	166.0	69.24	25	100.0%	21.68 [-19.04, 62.40]
Total (95% CI)			25			25	100.0%	21.68 [-19.04, 62.40]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.04 (P = 0.30)								

Table 32: Self-ligating brackets versus Conventional brackets, outcome: Ligation time

Untying	**Self Ligating**			**Conventional**				
	Mean time (secs)	SD	N	Mean time (secs)	SD	N	Weight	Mean difference (95% CI)
Miles 2010	9.2	2.5	34	31.5	10.2	34	100.0%	-22.30 [-25.83, -18.77]
Subtotal (95% CI)			34			34	100.0%	-22.30 [-25.83, -18.77]
Heterogeneity: Not applicable								
Test for overall effect: Z = 12.38 (P < 0.00001)								
Ligating								
Miles 2010	12.4	6.3	30	91.2	5.8	30	100.0%	-78.80 [-81.86, -75.74]
Subtotal (95% CI)			30			30	100.0%	-78.80 [-81.86, -75.74]
Heterogeneity: Not applicable								
Test for overall effect: Z = 50.40 (P < 0.00001)								
Test for subgroup differences: Chi² = 561.21, df = 1 (P < 0.00001), I² = 99.8%								

Table 32: Self-ligating brackets versus Conventional brackets, outcome: Ligation time

Pre-treatment	Self Ligating			Conventional				
	Mean LI inclination	SD	N	Mean LI inclination	SD	N	Weight	Mean difference (95% CI)
Atik 2014	96.33	3.4	16	92.95	6.29	17	100.0%	3.38 [-0.04, 6.80]
Subtotal (95% CI)			16			17	100.0%	3.38 [-0.04, 6.80]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.94 (P = 0.05)								
Post-treatment								
Atik 2014	102.44	6.05	16	97.91	7.27	17	100.0%	4.53 [-0.02, 9.08]
Subtotal (95% CI)			16			17	100.0%	4.53 [-0.02, 9.08]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
Test for subgroup differences: Chi² = 0.16, df = 1 (P = 0.69), I² = 0%								

Table 33: Self-ligating brackets versus Conventional brackets, outcome: Lower incisor inclination

Pre-treatment	Self Ligating			Conventional				
	Mean LI inclination	SD	N	Mean LI inclination	SD	N	Weight	Mean difference (95% CI)
Atik 2014	96.33	3.4	16	92.95	6.29	17	100.0%	3.38 [-0.04, 6.80]
Subtotal (95% CI)			16			17	100.0%	3.38 [-0.04, 6.80]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.94 (P = 0.05)								
Post-treatment								
Atik 2014	102.44	6.05	16	97.91	7.27	17	100.0%	4.53 [-0.02, 9.08]
Subtotal (95% CI)			16			17	100.0%	4.53 [-0.02, 9.08]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.95 (P = 0.05)								
Test for subgroup differences: Chi² = 0.16, df = 1 (P = 0.69), I² = 0%								

5.3.8: Comparison: Active versus passive self-ligating brackets

Crowding

One study investigated baseline upper anterior crowding in millimetres (Pandis 2010[83]). This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not possible. In total, seventy participants provided information on baseline crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between baseline crowding in the two groups was carried out (Figure 36). This revealed:

- Statistically significant difference in crowding at baseline between the two bracket groups (P= 0.04) (Figure 36).

Time to alignment

One study (Pandis 2010[83]) reported on this outcome, in days, for up to one hundred and seven days. Alignment was considered complete when the maxillary incisors were visually regarded as aligned. In total, sixty-six participants completed the study and provided information on time to alignment.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis between baseline crowding in the two groups was carried out (Figure 37). This revealed:

- No statistically significant difference in time to alignment between the two bracket groups (P= 0.15) (Figure 36).

Ligation time

This outcome was not reported.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 36: Forest plot of comparison 8: Active versus passive self-ligating brackets, outcome: Upper anterior segment alignment

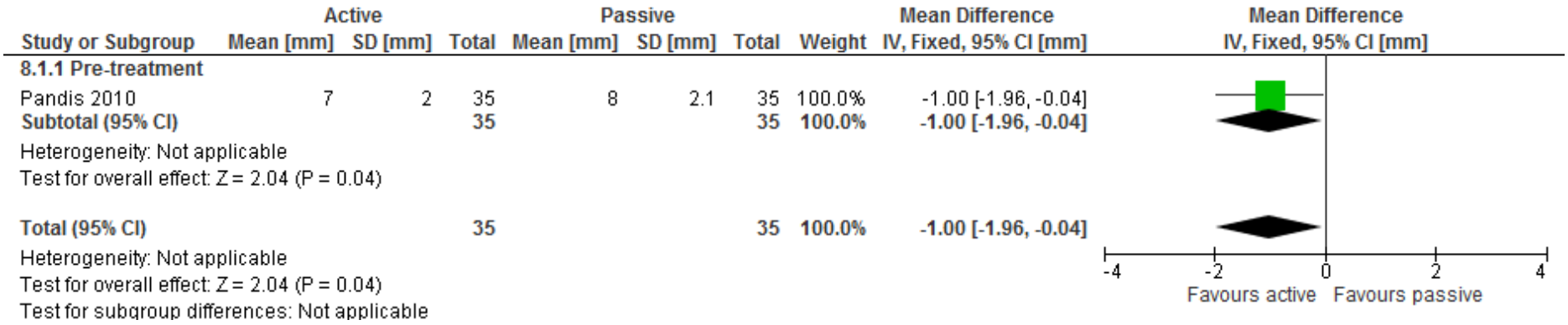


Figure 37: Forest plot of comparison 8: Active versus passive self-ligating brackets, outcome: Time to alignment

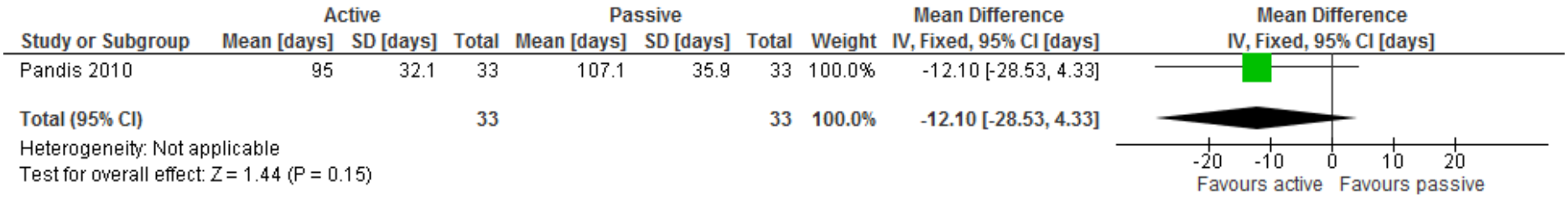


Table 34: Active versus passive self-ligating brackets, outcome: Upper anterior segment alignment

Pre-treatment	Active SL			Passive SL				
	Mean irregularity	SD	N	Mean irregularity	SD	N	Weight	Mean difference (95% CI)
Pandis 2010	7.0	2.0	35	8.0	2.1	35	100.0%	-1.00 [-1.96, -0.04]
Subtotal (95% CI)			35			35	100.0%	-1.00 [-1.96, -0.04]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.04 (P = 0.04)								
Total (95% CI)			35			35	100.0%	-1.00 [-1.96, -0.04]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.04 (P = 0.04)								
Test for subgroup differences: Not applicable								

Table 35: Active versus passive self-ligating brackets, outcome: Time to alignment

	Active SL			Passive SL				
Number of days till alignment	Mean number of days	SD	N	Mean number of days	SD	N	Weight	Mean difference (95% CI)
Pandis 2010	95.0	32.1	33	107.1	35.9	33	100.0%	-12.10 [-28.53, 4.33]
Total (95% CI)			33			33	100.0%	-12.10 [-28.53, 4.33]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.44 (P = 0.15)								

5.3.9: Comparison: Copper nickel-titanium versus nickel-titanium archwires

Crowding

Two studies investigated baseline mandibular anterior crowding in millimetres and they were combined in a meta-analysis (Ong 2011[94], Pandis 2009[92]). Both were considered to be at low risk of overall bias and in total, one hundred and forty-eight participants were analysed for baseline crowding.

The meta-analysis revealed:

- Overall, there was no statistically significant difference in baseline crowding in either study (MD -0.61mm, 95% CI -1.37, 0.16, P= 0.44).
- There was no heterogeneity (I^2= 0%).

Time to alignment

The same two studies were combined in a meta-analysis to assess time to alignment in days. This revealed:

- Overall, there was no statistically significant difference in time to alignment based on either archwire (MD -4.87mm, 95% CI -22.47, 12.72, P= 0.19).
- There was moderate heterogeneity (I^2= 41%); this may be because the studies showed opposite treatment effects although the differences were not statistically significant. Pandis 2009 [92] favoured NiTi whereas Ong 2011 [94] favoured CuNiTi.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

Discomfort

One study investigated the discomfort experienced over a period of on a 7-point Likert scale (Ong 2011[94]). The participants were given a questionnaire and asked to record discomfort in the upper arch at 4 hours, 24 hours, 3 days and 1 week after each archwire was changed. There was no significant difference in the overall discomfort levels between archwire sequences.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 38: Forest plot of comparison 9: CuNiTi versus NiTi archwires, outcome: Baseline crowding

Study or Subgroup	CuNiTi sequence Mean [mm]	SD [mm]	Total	NiTi sequence Mean [mm]	SD [mm]	Total	Weight	Mean Difference IV, Fixed, 95% CI [mm]
Pandis 2009	5.3	2.3	30	5.6	2	30	48.9%	-0.30 [-1.39, 0.79]
Ong 2011	2.8	2.4	44	3.7	2.7	44	51.1%	-0.90 [-1.97, 0.17]
Total (95% CI)			**74**			**74**	**100.0%**	**-0.61 [-1.37, 0.16]**

Heterogeneity: Chi² = 0.59, df = 1 (P = 0.44); I² = 0%
Test for overall effect: Z = 1.56 (P = 0.12)

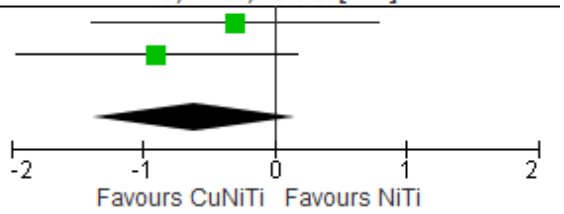


Figure 39: Forest plot of comparison 9: CuNiTi versus NiTi archwires, outcome: Time to alignment in days

Study or Subgroup	CuNiTi sequence Mean [Days]	SD [Days]	Total	NiTi sequence Mean [Days]	SD [Days]	Total	Weight	Mean Difference IV, Random, 95% CI [Days]
Pandis 2009	129.46	49.21	30	122.6	45.72	31	36.9%	6.86 [-17.00, 30.72]
Ong 2011	117.33	35.2	42	129.07	35.2	44	63.1%	-11.74 [-26.62, 3.14]
Total (95% CI)			**72**			**75**	**100.0%**	**-4.87 [-22.47, 12.72]**

Heterogeneity: Tau² = 70.07; Chi² = 1.68, df = 1 (P = 0.19); I² = 41%
Test for overall effect: Z = 0.54 (P = 0.59)

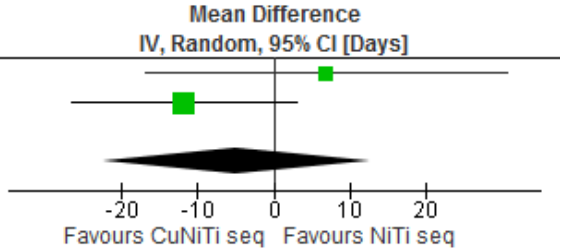


Table 36: CuNiTi versus NiTi archwires, outcome: Baseline crowding

Baseline crowding	CuNiTi			NiTi				
	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (95% CI)
Pandis 2009	5.3	2.3	30	5.6	2.0	30	48.9%	-0.30 [-1.39, 0.79]
Ong 2011	2.8	2.4	44	3.7	2.7	44	51.1%	-0.90 [-1.97, 0.17]
Total (95% CI)			74			74	100.0%	-0.61 [-1.37, 0.16]
Heterogeneity: Chi² = 0.59, df = 1 (P = 0.44); I² = 0%								
Test for overall effect: Z = 1.56 (P = 0.12)								

Table 37: CuNiTi versus NiTi archwires, outcome: Time to alignment in days

Days to alignment	CuNiTi			NiTi				
	Mean days to alignment	SD	N	Mean days to alignment	SD	N	Weight	Mean difference (95% CI)
Pandis 2009	129.46	49.2	30	122.6	45.7	31	36.9%	6.86 [-17.00, 30.72]
Ong 2011	117.33	35.2	42	129.07	35.2	44	63.1%	-11.74 [-26.62, 3.14]
Total (95% CI)			72			75	100.0%	-4.87 [-22.47, 12.72]
Heterogeneity: Tau² = 70.07; Chi² = 1.68, df = 1 (P = 0.19); I² = 41%								
Test for overall effect: Z = 0.54 (P = 0.59)								

5.3.10: Comparison: Coaxial nickel-titanium versus nickel-titanium

Crowding

One study investigated baseline lower anterior crowding in millimetres (Sebastian 2012[93]) for up to eight weeks. The time points at which reduction in crowding, or tooth movement, was reported were 4, 8 and 12 weeks. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not carried out and the sample consisted of only female participants. In total, twenty-four participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis on crowding in the two groups was carried out (Figures 40 and 41). This revealed:

- Baseline equivalence in crowding between coaxial NiTi and NiTi groups (MD 0.1mm, 95% CI -1.14, 1.34, P= 0.87).
- Statistically significant more tooth movement, or reduction in crowding, in the coaxial NiTi group compared to the NiTi group at 4, 8 and 12 weeks (P < 0.00001 for all three time points).
- Subgroup analysis also revealed a statistically significant difference favouring coaxial NiTi (P< 0.00001).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 40: Forest plot of comparison 10: Coaxial NiTi vs NiTi archwires, outcome: Baseline crowding

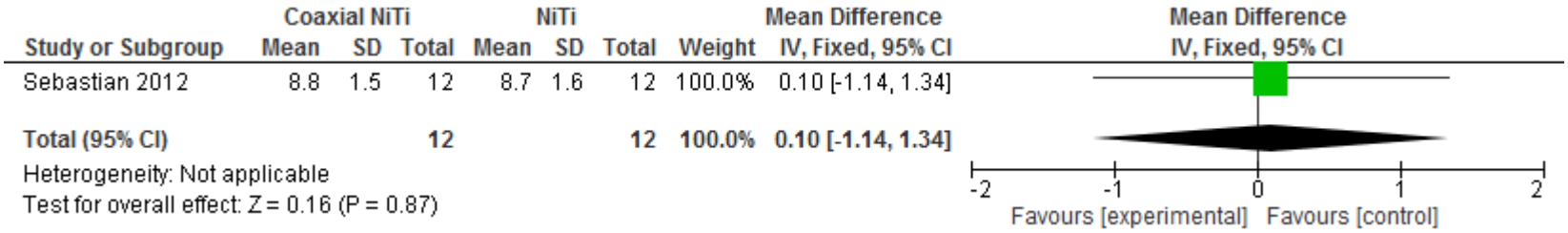


Figure 41: Forest plot of comparison 10: Coaxial NiTi vs NiTi archwires, outcome: tooth movement

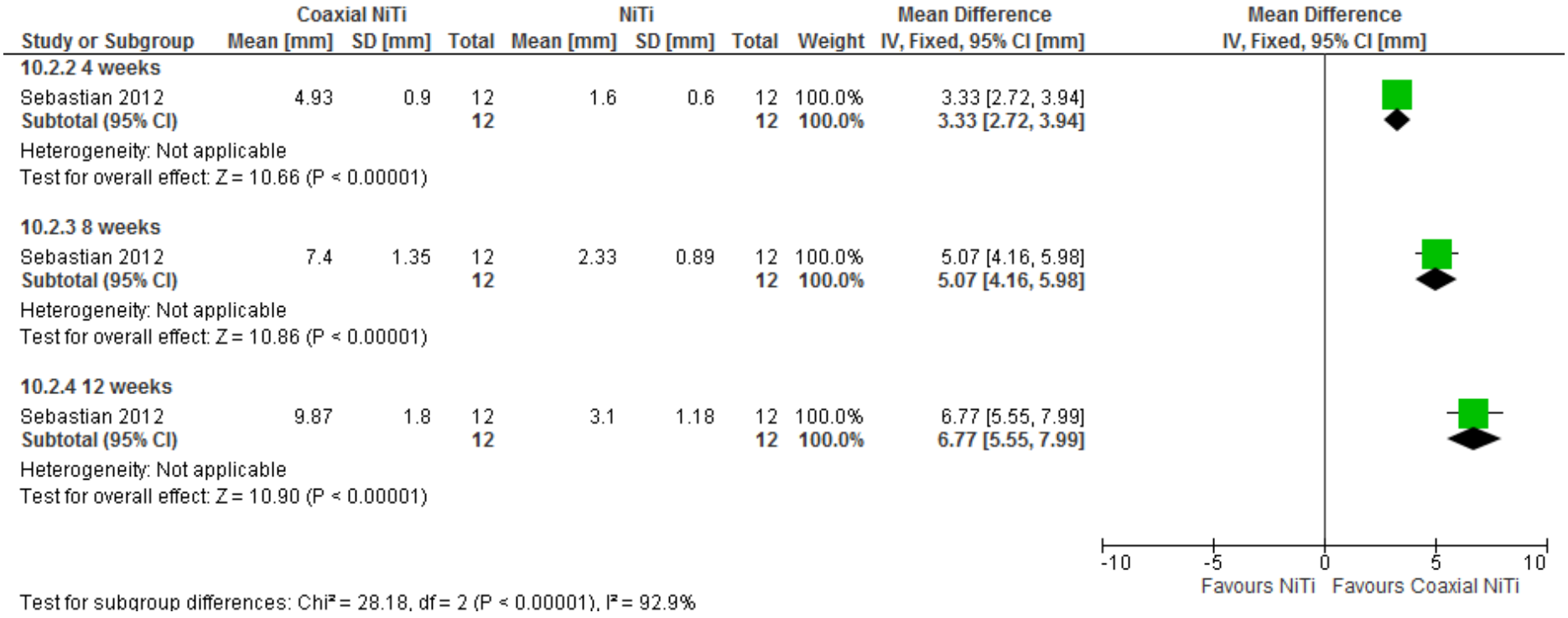


Table 38: Coaxial NiTi vs NiTi archwires, outcome: Baseline crowding

	Coaxial NiTi			NiTi				
	Mean crowding	SD	Mean crowding	Mean crowding	SD	N	Weight	Mean difference (95% CI)
Pre-treatment crowding	8.8	1.5	12	8.7	1.6	12	100.0%	0.10 [-1.14, 1.34]
Total (95% CI)			12			12	100.0%	0.10 [-1.14, 1.34]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.16 (P = 0.87)								

Table 39: Coaxial NiTi vs NiTi archwires, outcome: tooth movement

	Mean tooth movement	SD	N	Mean tooth movement	SD	N	Weight	Mean difference (95% CI)
4 weeks								
Sebastian 2012	4.93	0.9	12	1.6	0.6	12	100.0%	3.33 [2.72, 3.94]
Subtotal (95% CI)			12			12	100.0%	3.33 [2.72, 3.94]
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.66 (P < 0.00001)								
8 weeks								
Sebastian 2012	7.4	1.3	12	2.33	0.9	12	100.0%	5.07 [4.16, 5.98]
Subtotal (95% CI)			12			12	100.0%	5.07 [4.16, 5.98]
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.86 (P < 0.00001)								
12 weeks								
Sebastian 2012	9.87	1.8	12	3.1	1.2	12	100.0%	6.77 [5.55, 7.99]
Subtotal (95% CI)			12			12	100.0%	6.77 [5.55, 7.99]
Heterogeneity: Not applicable								
Test for overall effect: Z = 10.90 (P < 0.00001)								
Test for subgroup differences: Chi^2 = 28.18, df = 2 (P < 0.00001), I^2 = 92.9%								

5.3.11: Comparison: Nitinol versus Titinol

Crowding

One study investigated baseline upper anterior crowding in millimetres (O'Brien 1990[91]) for up to thirty-seven days. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not mentioned. In total, forty participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a total group analysis on crowding at baseline, and on the change in crowding from start to finish, was carried out (Tables 38 and 39). This revealed:

- Baseline equivalence between Nitinol and Titinol groups (MD 3.31mm, 95% CI -0.73, 7.35, P= 0.11).
- No statistically significant difference between groups in the change in crowding for either group (MD 0.28, 95 CI -0.89, 0.33, P= 0.37).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 42: Forest plot of comparison 11: Nitinol versus Titinol archwires, outcome: Crowding

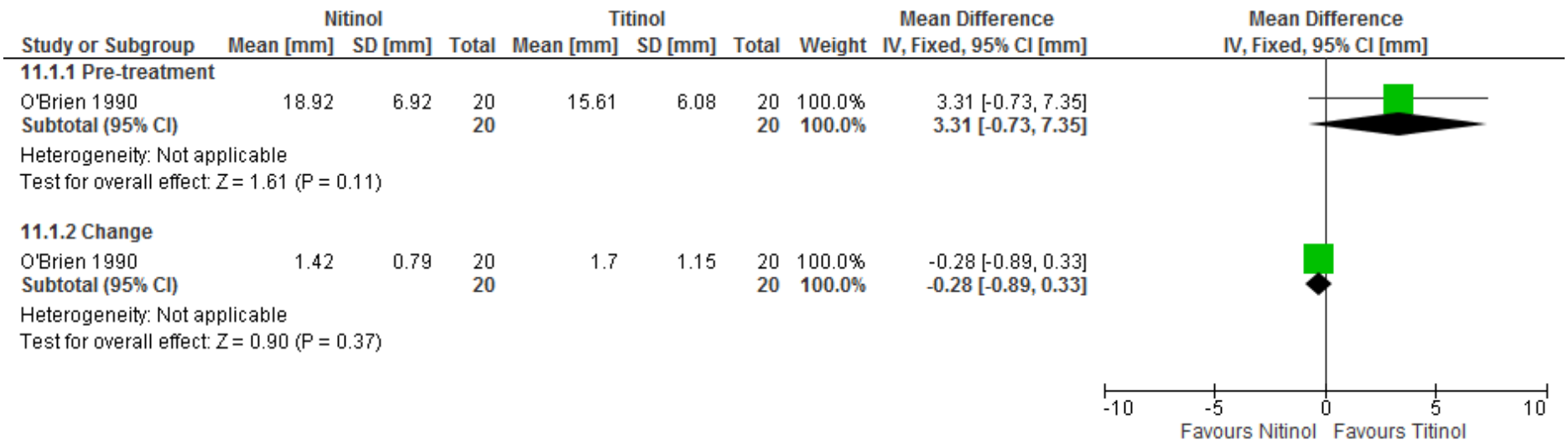


Table 40: Nitinol versus Titinol, outcome: Baseline crowding

	Nitinol			Titinol				
Pre-treatment	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mead difference (95% CI)
O'Brien 1990	18.92	6.92	20	15.61	6.08	20	100.0%	3.31 [-0.73, 7.35]
Subtotal (95% CI)			20			20	100.0%	3.31 [-0.73, 7.35]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.61 (P = 0.11)								

Table 41: Titinol versus Nitinol, outcome: Change in crowding

	Nitinol			Titinol				
Change	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mead difference (95% CI)
O'Brien 1990	1.42	0.79	20	1.7	1.15	20	100.0%	-0.28 [-0.89, 0.33]
Subtotal (95% CI)			20			20	100.0%	-0.28 [-0.89, 0.33]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.90 (P = 0.37)								

5.3.12: Comparison: Nickel-titanium versus stainless steel

Crowding

One study investigated baseline mandibular crowding in millimetres (Gravina 2013[88]) for up to eight weeks. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not mentioned. In total, twenty-four participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however, a subgroup analysis between baseline and eight weeks was carried out (Figure 43). This revealed:

- Baseline equivalence between NiTi and stainless steel groups (MD -26.0mm, 95% CI -0.60, 0.07, P= 0.13).
- Statistically significantly less crowding in the NiTi group at eight weeks, (MD -0.43mm, 95 CI -0.78, -0.08, P= 0.02).
- No statistically significant difference in the change in crowding between NiTi and stainless steel (MD -0.17mm, 95 CI -0.42, 0.09)
- Overall, there is no statistically significant difference between the subgroups (P= 0.5).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No study in this subgroup presented data in a way which facilitated assessment of this outcome.

Self-esteem

No study in this subgroup reported on this outcome.

Patient satisfaction

No study in this subgroup reported on this outcome.

Jaw joint problems

No study in this subgroup reported on this outcome.

Figure 43: Forest plot of comparison 12: Nickel-titanium versus stainless steel, outcome: Crowding

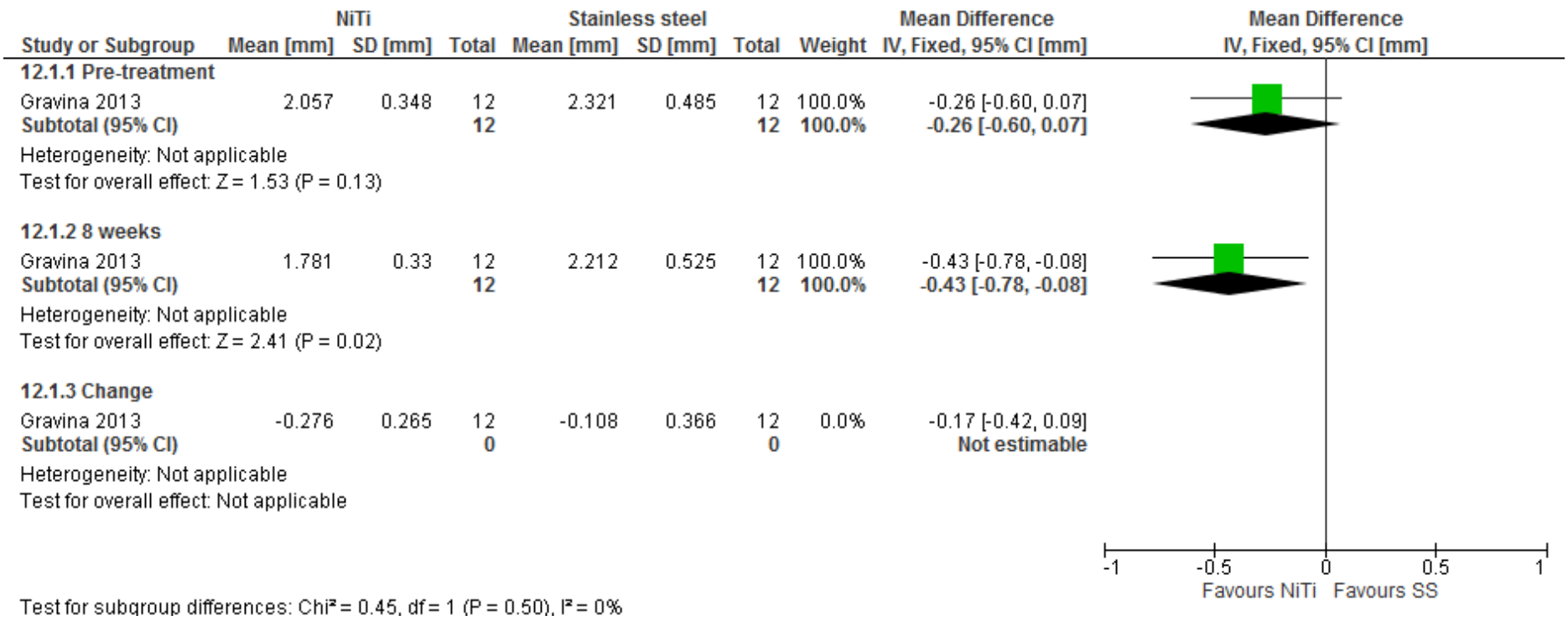


Table 42: Nickel-titanium versus stainless steel, outcome: Lower arch crowding

	NiTi			Stainless steel				
Pre-treatment	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (9% CI)
Gravina 2013	2.057	0.348	12	2.321	0.485	12	100.0%	-0.26 [-0.60, 0.07]
Subtotal (95% CI)			12			12	100.0%	-0.26 [-0.60, 0.07]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.53 (P = 0.13)								
8 weeks								
Gravina 2013	1.781	0.33	12	2.212	0.525	12	100.0%	-0.43 [-0.78, -0.08]
Subtotal (95% CI)			12			12	100.0%	-0.43 [-0.78, -0.08]
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.41 (P = 0.02)								
Test for subgroup differences: Chi^2 = 0.45, df = 1 (P = 0.50), I^2 = 0%								

5.3.13: Comparison: Nickel-titanium versus multi-stranded stainless steel

Crowding

One study investigated baseline mandibular crowding in millimetres (Gravina 2013[88]) for up to eight weeks. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not mentioned. In total, twenty four participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between baseline and eight weeks was carried out (Figure 44). This revealed:

- Baseline equivalence of crowding between NiTi and multi-stranded stainless steel groups (MD -0.04mm, 95% CI -0.33, 0.24, P= 0.76).
- No statistically significant difference in crowding between groups at eight weeks, (MD -0.03mm, 95 CI -0.39, 0.33, P= 0.88).
- Overall, there was no statistically significant difference between correction of crowding with NiTi and multi-stranded stainless (P= 0.94).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 44: Forest plot of comparison 13: Nickel-titanium versus multi-stranded stainless steel, outcome: Crowding

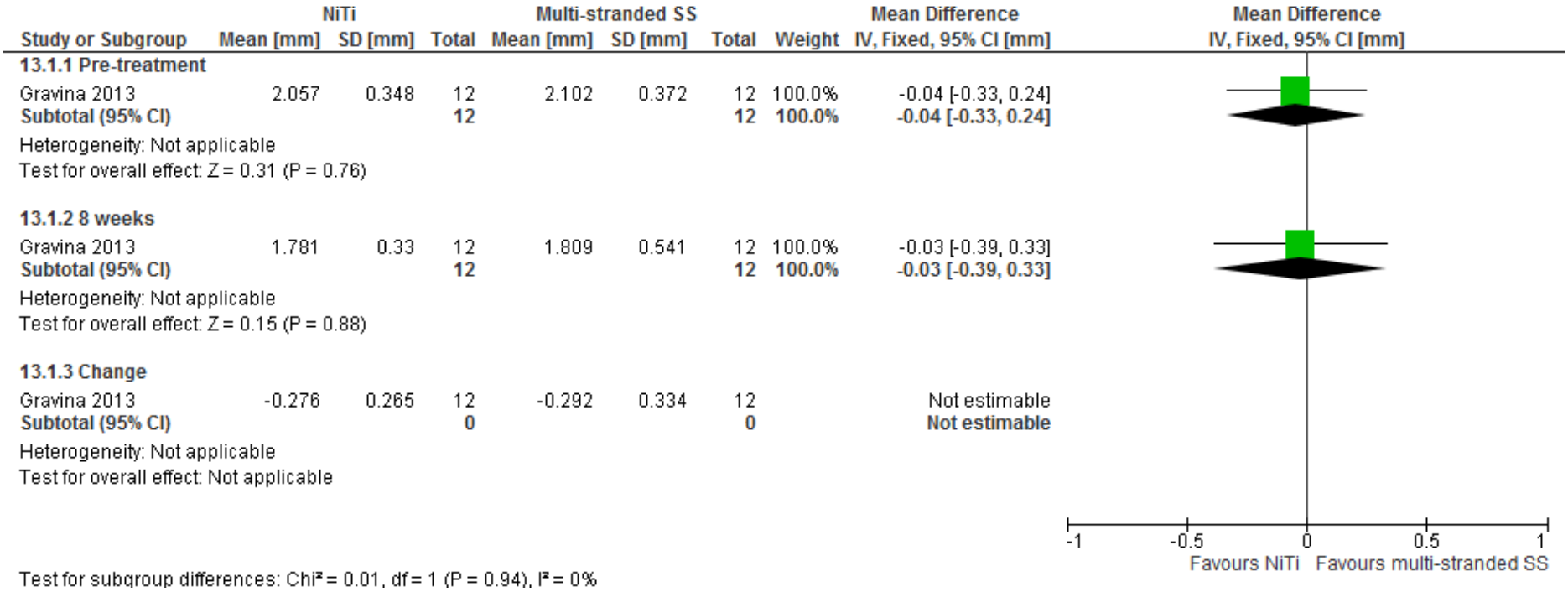


Table 43: Nickel-titanium versus multi-stranded stainless steel, outcome: Lower arch crowding

	NiTi			**Multi stranded stainless steel**				
Pre-treatment	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (9% CI)
Gravina 2013	2.057	0.348	12	2.102	0.372	12	100.0%	-0.04 [-0.33, 0.24]
Subtotal (95% CI)			12			12	100.0%	-0.04 [-0.33, 0.24]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.31 (P = 0.76)								
8 weeks								
Gravina 2013	1.781	0.33	12	1.809	0.541	12	100.0%	-0.03 [-0.39, 0.33]
Subtotal (95% CI)			12			12	100.0%	-0.03 [-0.39, 0.33]
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.15 (P = 0.88)								
Test for subgroup differences: Chi^2 = 0.01, df = 1 (P = 0.94), I^2 = 0%								

5.3.14: Comparison: Multi-stranded stainless steel versus stainless steel

Crowding

One study investigated baseline mandibular crowding in millimetres (Gravina 2013[88]) for up to eight weeks. This study was assessed as being at overall high risk of bias as blinding of personnel and participants was not mentioned. In total, twenty-four participants provided information on crowding.

We were unable to perform a meta-analysis as only one study assessed this comparison, however a subgroup analysis between baseline and eight weeks was carried out (Figure 44). This revealed:

- Baseline equivalence in crowding between multi-stranded stainless steel and stainless steel groups (MD -0.22mm, 95% CI -0.56, 0.13, P= 0.21).
- No statistically significant difference in crowding between groups at eight weeks, (MD -0.40mm, 95 CI -0.83, 0.02, P= 0.06).
- Overall, there is no statistically significant difference between the subgroups (P= 0.51).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

No study reported on this outcome.

Upper incisors to maxilla

This outcome was not reported.

Harms

No harms were reported.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 45: Forest plot of comparison 14: Multi-stranded stainless steel versus stainless steel, outcome: Crowding

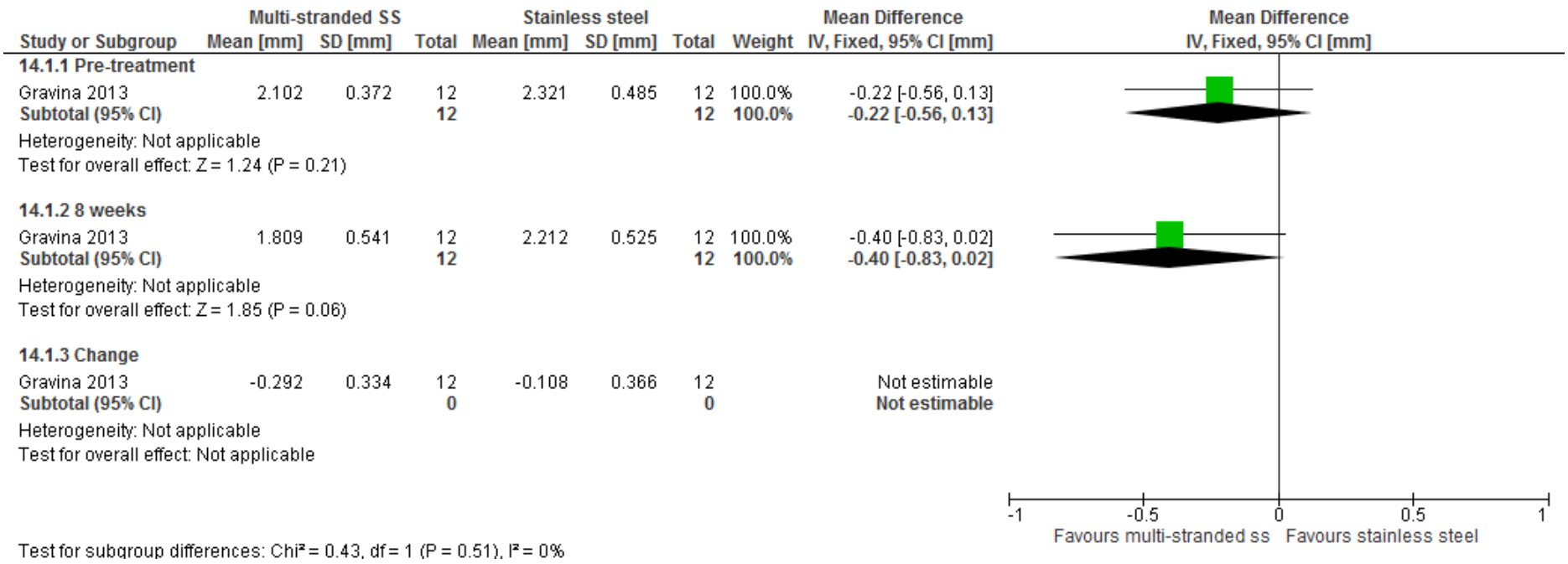


Table 44: Multi-stranded stainless steel versus stainless steel, outcome: Crowding

	Multi stranded stainless steel			**Stainless steel**				
Pre-treatment	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (9% CI)
Gravina 2013	2.102	0.372	12	2.321	0.485	12	100.0%	-0.22 [-0.56, 0.13]
Subtotal (95% CI)			12			12	100.0%	-0.22 [-0.56, 0.13]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.24 (P = 0.21)								
8 weeks								
Gravina 2013	1.809	0.541	12	2.212	0.525	12	100.0%	-0.40 [-0.83, 0.02]
Subtotal (95% CI)			12			12	100.0%	-0.40 [-0.83, 0.02]
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.85 (P = 0.06)								
Test for subgroup differences: Chi^2 = 0.43, df = 1 (P = 0.51), I^2 = 0%								

5.3.15: Comparison: Vibrational appliances versus control

Crowding

Two studies investigated baseline mandibular anterior crowding in millimetres and they were combined in a meta-analysis (Miles 2012[90], Woodhouse 2015[95]). Miles 2012 [90] was assessed as being at overall unclear risk of bias, whereas Woodhouse 2015 [95] was at a high level of bias, mainly as the clinicians and participants were not blinded to the intervention. In total, one hundred and twenty-two participants provided information on baseline crowding. One hundred and nineteen participants provided information for crowding at 8-10 weeks.

The meta-analysis revealed:

- Baseline equivalence in irregularity between the two groups in each study and between them (MD 0.69, 95% CI -0.56, 1.95, P= 0.28)
- No statistically significant difference in irregularity between vibrational appliances and controls at 8-10 weeks (MD 0.42, 95% CI -0.05, 0.9, P= 0.08)
- Overall, there is no statistically significant difference in crowding in either group (P= 0.69).
- There is no heterogeneity (I^2= 0%).

When reviewing the change in crowding, the meta-analysis revealed:

- No statistical significance to the change in crowding between groups (P= 0.46).

Time to alignment

This outcome was not reported.

Ligation time

This outcome was not relevant to this comparison.

Arch length

This outcome was not reported.

Lower incisors to mandible

This outcome was not reported.

Lower molars to mandible

This outcome was not reported.

Upper incisors to maxilla

This outcome was not reported.

Harms

Discomfort

One study reported on discomfort (Miles 2012[90]) using a 100mm Visual Analogue Scale (VAS) over the first week. The participants were asked to keep a diary and record discomfort at 4 time-points: bond-up, 6-8 hours after appliance placement, 1 day after, 3 days after and 7 days after at the appliances were placed. There was no statistical difference found in the study for discomfort scores between the vibrational appliance and control group.

Self-esteem

This outcome was not reported.

Patient satisfaction

This outcome was not reported.

Jaw joint problems

This outcome was not reported.

Figure 46: Forest plot of comparison 15: Vibrational appliances versus control, outcome: Crowding

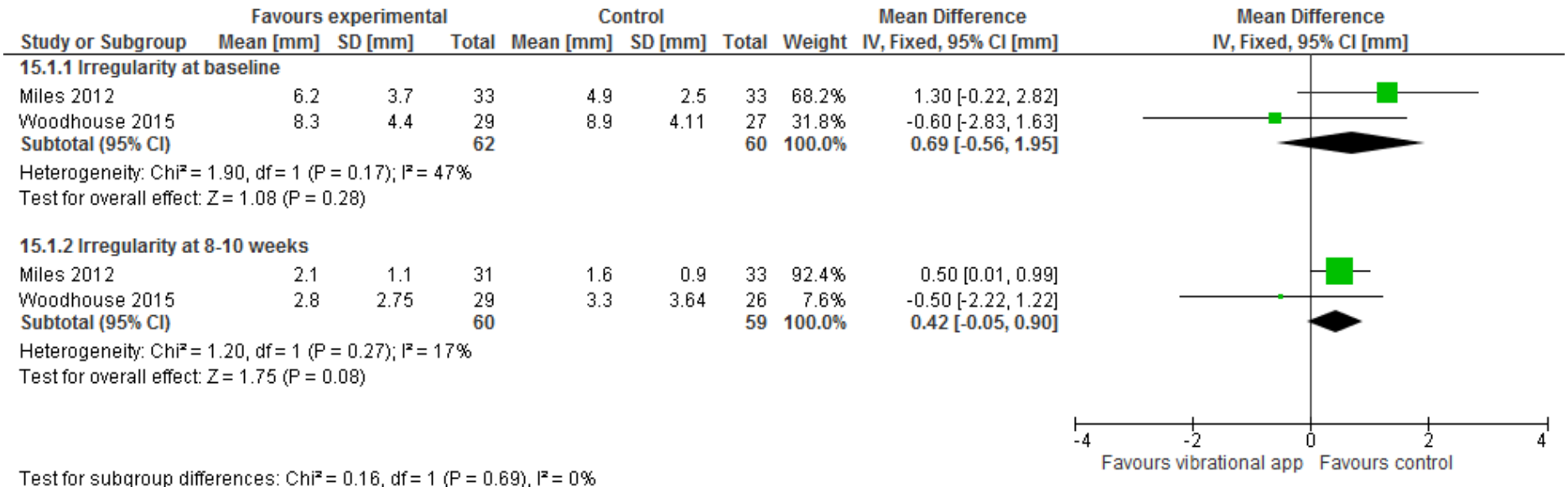


Figure 47: Forest plot of comparison 15: Vibrational appliances versus control, outcome: Change in crowding

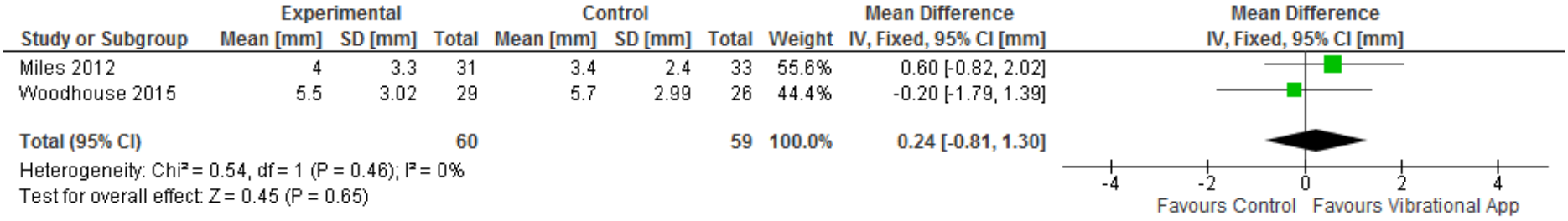


Table 45: Vibrational appliances versus control, outcome: Crowding

	Vibrational appliance			Control				
Irregularity at baseline	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (95% CI)
Miles 2012	6.2	3.7	33	4.9	2.5	33	68.2%	1.30 [-0.22, 2.82]
Woodhouse 2015	8.3	4.4	29	8.9	4.11	27	31.8%	-0.60 [-2.83, 1.63]
Subtotal (95% CI)			62			60	100.0%	0.69 [-0.56, 1.95]
Heterogeneity: Chi² = 1.90, df = 1 (P = 0.17); I² = 47%								
Test for overall effect: Z = 1.08 (P = 0.28)								
Irregularity at 8-10 weeks								
Miles 2012	2.1	1.1	31	1.6	0.9	33	92.4%	0.50 [0.01, 0.99]
Woodhouse 2015	2.8	2.75	29	3.3	3.64	26	7.6%	-0.50 [-2.22, 1.22]
Subtotal (95% CI)			60			59	100.0%	0.42 [-0.05, 0.90]
Heterogeneity: Chi² = 1.20, df = 1 (P = 0.27); I² = 17%								
Test for overall effect: Z = 1.75 (P = 0.08)								
Test for subgroup differences: Chi² = 0.16, df = 1 (P = 0.69), I² = 0%								

Table 46: Vibrational appliances versus control, outcome: Change in crowding

	Vibrational appliance			Control				
Change in irregularity over 8-10 weeks	Mean crowding	SD	N	Mean crowding	SD	N	Weight	Mean difference (95% CI)
Miles 2012	4.0	3.3	31	3.4	2.4	33	55.6%	0.60 [-0.82, 2.02]
Woodhouse 2015	5.5	3.02	29	5.7	2.99	26	44.4%	-0.20 [-1.79, 1.39]
Total (95% CI)			60			59	100.0%	0.24 [-0.81, 1.30]
Heterogeneity: Chi² = 0.54, df = 1 (P = 0.46); I² = 0%								
Test for overall effect: Z = 0.45 (P = 0.65)								

6. DISCUSSION

6.1: Summary of main results

Nineteen reports of seventeen randomised controlled trials (RCTs) met our eligibility criteria and were included in this review. In total, fifteen comparisons were identified and a meta-analysis was performed for two comparisons. Further analysis was not possible, as there was no standardised method for outcome measurement or an accepted, uniform outcome across the studies.

Prevention

Cervical headgear versus control

We found that headgear increased arch length in the upper arch by up to 2.28mm more than the control group, but this was at the expense of 2.5 degrees of upper incisor proclination, therefore resulting in anterior anchorage loss. Subgroup analysis revealed that there was no statistically significant change between 0-1 or 0-2 years, so the majority of the proclination occurred in the first year.

Lower arch length also increased by up to 1.52mm and the majority of the increase was in the first two years, as subgroup analysis did not reveal a statistical difference between 2 and 8 years. The lower incisors also initially proclined, but this effect was not maintained at 2 years, indicating that cervical pull headgear therapy does not result in long term lower incisor proclination. It was also not found to affect lower incisor crowding. Upper incisor crowding was not examined, but this would be an outcome worth investigating to determine whether cervical pull headgear does alleviate crowding and if so, whether this is by increasing the maxillary arch length by incisor proclination, by distal movement of the buccal segments or a combination of both.

Extraction of lower deciduous canines versus control

Extracting the lower deciduous canines statistically significantly reduced the arch length by 2.73mm and also reduced crowding by 4.76mm more when compared to the observation group. There was no statistically significant difference in the inclination of the lower incisors between the extraction and observation groups, so correction of crowding did not occur by proclination of the lower labial segment and was more likely to be as a direct result of alignment into the extraction space.

However, as arch length was also reduced in the extraction group, this would imply that space is then lost for either the permanent canines or the premolars, thus transferring the crowding into the adult dentition.

The study (Kau 2004[60]) discusses the impact of the crowding in more detail; crowding was considered to have improved if there was a 50% reduction, or if there was an irregularity score of less than 2.5mm at the end of the study. In the extraction group, only 28% of cases demonstrated an improvement against these criteria. In fact, when arch length loss was also considered alongside the crowding, only 6% of cases showed an improvement, meaning that overall there was only a 5% chance of improving crowding by extracting the lower deciduous canines.

These results should be interpreted with caution, as there was a statistically significant difference in baseline crowding between the groups, with the extraction group having statistically significantly more crowding pre-treatment. This suggests that the treatment effect may have been overestimated, as the more severely crowded teeth were able to align into the available space.

Schwarz versus control

Our analysis found that the improvement in lower arch crowding was statistically significant, with the Schwarz appliance reducing crowding by 2.14mm more in the than the control group. Although there were statistically significant differences in baseline arch lengths favouring the Schwarz group, the change in arch length overall was not statistically significant. There was also no difference in maxillary incisor advancement, but subgroup analysis revealed that the change in mandibular incisor advancement was double in the Schwarz appliance group compared to the control group. However, whilst this advancement was statistically significant (P= 0.006), at only 0.39mm, it was not thought to be clinically significant and could also be attributed to tracing error[141].

This suggests that the Schwarz appliance is an effective treatment option for the alleviation of crowding in the mandibular arch in the short term however, there was no follow-up beyond 9 months so it is not possible to say whether this improvement in crowding was maintained into the permanent dentition.

Eruption Guidance appliance versus control

The Eruption Guidance appliance improved crowding in the lower labial segment with an odds ratio of 0.19, so that the likelihood of remaining crowded in the control group were 5.3 times more than in the treatment group. This was also confirmed by subgroup analysis that demonstrated a statistically significant difference in crowding levels post treatment, despite baseline equivalence. However, the treated group also had 4.1 degrees of lower incisor proclination post-treatment, which was statistically significantly different from the pre-treatment inclination and would suggest that resolution of crowding was, in part, due to the space gained from proclination. As the control group's incisor inclination was not reported, we are unable to determine if this effect is due to the appliance.

In the maxilla, the appliance made no difference to crowding or to upper incisor proclination.

Further research with data for the control group is required for this appliance.

Lower Lingual Arch versus control

The LLA is traditionally used to maintain arch length and the maintain leeway space following the loss of deciduous molar(s). Our analysis revealed that the LLA held the position of the molar by 1.11mm and 2.73 degrees more than the control group, so that it keeps the molar upright and reduces its mesial movement, and that these findings were statistically significant. More detailed analysis revealed that the LLA only permitted 0.33mm of mesial movement of the molar and resulted in 0.54 degrees of distal tipping, which would be 0.66mm for the entire lower arch. In comparison, the control group had 1.44mm of mesial movement, giving a total of 2.88mm for the arch and 2.19 degrees of mesial inclination.

Anteriorly, in the LLA group, the lower incisors advanced by 0.66mm and 3.01 degrees more than the control group, which can be interpreted as anterior anchorage loss. However, upon closer inspection, the advancement in LLA was only 0.32mm and by 0.73 degrees. In contrast, the control group had retroclination and distal movement of the anterior teeth, so whilst the difference was statistically significantly different, it may not be clinically significant.

Overall, it seems that the LLA keeps the molars upright and preserves space, despite a small amount of anterior anchorage loss, compared to the control group

who suffered mesial movement of the lower molars and retroclination of the lower incisors. These findings were confirmed by the statistically significant difference in the arch lengths: the LLA increased mandibular arch length by 0.07mm, whereas the control group suffered a 2.54mm reduction.

These findings could be clinically significant and impact treatment plans with regards to being an extraction or non-extraction plan. The LLA could be considered a method of preserving the leeway space and preventing crowding in the mixed dentition from being perpetuated into the permanent dentition.

Lip bumper versus control

The lip bumper reduced crowding by 4.29mm more than the control group and increased the arch length by 2.19mm. In the control group, the arch length decreased by 1.15mm so that a difference of 3.34mm was seen between the lip bumper and control groups. All of these findings were statistically significant.

The increase in arch length for the lip bumper group appeared to be due to incisor advancement and distal movement of the molar, which were 0.69mm and 0.61mm respectively thus potentially gaining 1.3mm. Additionally, the lower incisors proclined by 3.14 degrees and the molars tipped distally by 4.13 degrees more than in the control group.

These findings suggest that the lip bumper keeps the molars upright and increases arch length at the expense of anterior anchorage loss, and whilst these findings were statistically significant, their clinical significance would have to be assessed on a case-by-case basis.

The findings were similar to those resulting from treatment with the LLA; both appliances increased the arch length by around 3mm, both suffered anterior anchorage loss of about 0.5mm and both appliances prevented mesial movement of the molars and caused them to tip distally compared to the control groups.

Treatment

Brackets

Self-ligating versus conventional brackets

Our analysis of two studies (Miles 2010[82], Pandis 2011[79]) revealed that there was no statistical difference in relief of crowding or time to alignment between self-ligating and conventional brackets. Additionally, the irregularity scores had an I^2 value of 0%, so they demonstrated homogeneity and a consistent treatment effect.

Whilst the lower incisors appeared to show more statistically significant proclination in the self-ligation group, this was attributed to significant baseline differences between the groups pre-treatment where the self-ligation group already had greater incisor proclination (Atik 2014[80]). Further confirmation was provided by the lack of statistically significant change in inclination of the lower incisors between the groups.

There was, however, a statistically significant and clinically significant difference in ligation and untying time between self-ligating and conventional brackets (Miles 2010[82]). Overall, self-ligating brackets saved 101.1 seconds per case (where we assume a patient has an archwire change and so requires untying and ligation). If this is extrapolated to 30 patients, which is an estimation of the number of patients seen in primary care practice a day, this is a saving of 50.55 minutes. This is clinically significant as this time could be utilised to see more patients, take a break or undertake other activities however, the increased cost of the self-ligating bracket systems also needs to be considered.

No statistically significant difference was found in discomfort between the two types of brackets from 4 hours to 1 month after bond-up (Atik 2014[80], Miles 2010[82]).

There was also no difference found between the Periodontal Index, Gingival Index or pocketing depths between both groups of participants between the start and end of the study.

Active versus passive self-ligating brackets

There was no statistical difference in the time to alignment between active and passive self-ligating brackets, despite the active brackets being 12 days quicker to

align. This lack of significance is attributed to the study (Pandis 2010[83]) being potentially underpowered with a relatively small sample size and resultant large standard deviations; had the sample sizes been increased to 60 per group, a statistical difference may have been found. Additionally, there was a statistically significant difference of 1mm less crowding in the active self-ligation group at baseline, so alignment in this group may have been quicker, so the difference in treatment effect may have been overestimated.

Archwires

Copper nickel-titanium versus nickel-titanium

Our search identified two papers (Pandis 2009[92], Ong 2011[94]) that were combined in a meta-analysis to assess differences in baseline crowding and time to alignment. This revealed that there was baseline equivalence with regards to pre-treatment crowding and there was no heterogeneity as I^2 was 0%.

For time to alignment, there was again no statistical difference found between the groups in the meta-analysis, although there was moderate heterogeneity with I^2 at 41%. There may have been other differences in baseline characteristics that were not reported.

When investigating discomfort, there was no difference found between the archwire groups for up to 1 week after the archwire was changed (Ong 2011[94]).

Coaxial nickel-titanium versus nickel-titanium

One study provided evidence that there is greater resolution of irregularity with coaxial NiTi than NiTi and that the treatment effect increases over time for up to 12 weeks, which was the end point of the study (Sebastian 2012[93]). However, the sample consisted entirely of females and so the study was at high risk of bias.

Nitinol versus titinol

One study provided evidence that showed no difference in the relief in crowding with either archwire, after exhibiting baseline equivalence between the groups (O'Brien 1990[91]). Again, this would suggest that there is no advantage to using one wire or the other for faster resolution of crowding.

Nickel-titanium versus stainless steel

We assessed one study for this comparison (Gravina 2013[88]) and found that whilst initially there appeared to be a benefit to using NiTi between baseline and 8 weeks, there was no statistical difference in the change in crowding over time. This study appeared to be under-powered and had the sample size been increased to 30, a statistical difference may have been discovered.

Nickel-titanium versus multi-stranded stainless steel

We assessed the same study (Gravina 2013[88]) and again found no difference in the change in crowding over time potentially a false negative due to the small sample size.

Multi-stranded stainless steel versus stainless steel

Again, there was no difference found in the change in crowding over 8 weeks (Gravina 2013[88]) and this was once again attributed to the study being under-powered.

Vibrational appliances versus control

Two studies were identified and assessed in a meta-analysis for this comparison (Pandis 2009[92], Ong 2011[94]). This revealed:

- Baseline equivalence in samples
- No difference in the change in irregularity between the start and at 8-10 weeks, or any subgroup differences between the vibrational appliance group and the control.
- The studies were potentially under-powered and had the sample sizes been increased to 50 per group, the outcome would have favoured the control group.

Additionally, there was no reduction in discomfort when using the vibrational appliances, so overall they provided no benefit with regards to crowding or pain reduction.

6.2: Potential biases and limitations of the review

Bias has been reduced in this systematic review by using a broad, sensitive search of multiple databases with no restrictions on language to identify reports of potentially eligible studies. We have also searched for unpublished studies and data, and have included studies reported in all languages. It was also reduced by following Cochrane guidelines and undertaking study selection and data extraction both independently and in duplicate. Despite this, potential biases have been detected whilst conducting this review, both within studies and between them. Another source of bias may have been due to the fact that the original protocol for the review was published before the year 2000 and since then treatment modalities have changed and additional outcomes have been considered relevant.

There were many different treatment options identified for both the prevention and treatment of crowding, but there was a lack of good quality RCTs available for each comparison. This meant that in many instances, there was only one study included per comparison. There were also a variety of outcome measures reported, making it difficult to draw parallels between the outcomes of different comparisons. As a result, only two meta-analyses were carried out with two studies per comparison.

Several studies had small sample sizes and/or had not carried out sample size calculations (Pandis 2009[92], Pandis 2010[83], Ong 2011[94], Gravina 2013[88]). Additionally, one study suffered a high level of dropout at the 13-year follow up, which led to attrition bias (Krusinskiene 2008[57]).

6.3: Agreements and disagreements with other studies or reviews

Four other reviews were found that reported on similar comparisons and outcomes to this review (Fleming 2016[142], El-Angbawi 2015[143], Yu et al 2013[144], Jian et al 2013[77]).

Fleming et al 2016 investigated 'Non-pharmacological interventions for alleviating pain during orthodontic treatment' and assessed Miles 2012 [90] as part of their review. They agreed with our findings that vibrational appliances do not reduce discomfort or pain at any of the time points investigated.

El-Angbawi et al 2015 [143] investigated 'Non-surgical adjunctive interventions for accelerating tooth movement in patients undergoing fixed orthodontic treatment' and

also assessed Miles 2012[90]. They also found that there was no statistical difference between the two groups.

The third review by Yu et al 2013[144] entitled “Interventions for managing relapse of the lower front teeth after orthodontic treatment” did not uncover any studies during their searches and so were unable to include any.

The final review by Jian et al 2013 [77] on ‘Initial archwires for alignment during orthodontic treatment with fixed appliances’ compared stabilsed NiTi against superelastic NiTi. As part of this comparison, they assessed O’Brien 1990 [91] and agreed with our findings: there was no statistically significant difference between Nitinol and Titinol in terms of tooth movement. Additionally, this review compared single stranded NiTi against other types of NiTi and concluded that there was very weak evidence from one study (Sebastian 2012[93]) that coaxial NiTi produces greater tooth movement than single stranded NiTi. Again, this is in agreement with our findings.

7: AUTHORS CONCLUSIONS

7.1: Implications for practice

Overall, the quality of the evidence for the prevention and treatment of crowding in children was low and suffered from methodological bias.

Three interventions were identified as being effective for the prevention of crowding with low quality evidence to support their use. These were the Schwarz appliance, the lower lingual arch and the lip bumper.

The Schwarz appliance was found to reduce incisor crowding in the mandible by expanding the arch and without causing clinically significant proclination of the lower incisors. The lower lingual arch and lip bumper were both found to be effective at maintaining space by increasing the arch length by around 3mm and preventing mesial movement of the molars, but at the expense of anterior anchorage loss of around 0.5mm.

For the treatment of crowding, there was one intervention, with low-quality evidence, that was identified as being more effective than its comparison. This was the use of coaxial NiTi when compared to single-stranded NiTi for initial alignment. It was also found that self-ligating brackets were quicker to untie and ligate than conventional brackets by over a minute and a half per case per treatment episode.

There was an insufficient number of studies identified per intervention for an analysis to be carried out for the treatment effect at varying ages.

7.2: Implications for research

As the overall quality of research was low to moderate, the results highlight the need for a uniform and systematic way of assessing, recording and measuring crowding in future research. This would allow comparison and combination of the results in a meta-analysis and would provide a stronger level of evidence.

The results of this review imply that there is a need for more long-term, well designed and reported randomised controlled clinical studies to assess the interventions and treatment options for crowding in children. This is particularly pertinent for interventions that are used in the mixed dentition with the aim of preventing and/or reducing crowding in the permanent dentition.

When designing future studies, the following need to be considered:

- Clear inclusion and exclusion criteria should be set.
- Use a standardised index for measuring crowding outcomes.
- An a priori sample size calculation should be carried out.
- Longer follow-up times especially when interventions are carried out in the mixed dentition.
- Reporting of outcomes in a format that is clinically useful. We suggest the following:
    - Changes in irregularity or angulation of teeth, as well as raw values alone of irregularity at different time points.
    - Time to alignment, and time for ligation and untying, particularly for interventions such as different archwires or brackets.
- Adverse effects or the absence of them should be reported in all studies.
- Reports of clinical trials would be improved by following the guidelines produced by the CONSORT group to ensure that all relevant information is provided.

8: ACKNOWLEDGEMENTS

We would like to thank Anne Littlewood (Trials Search Co-ordinator, Cochrane Oral Health Group) for developing the search strategy and Janet Lear (Cochrane Oral Health Group) for her support with locating papers and coordinating translations.

We are grateful to Professor Chengge Hua, Tian Ye and Chunjie Li for their translations of Dai 2009 [99] and Wang [145] and to Fang Hua for the translation of Fan [101] and Yu[113].

We would also express our sincere thanks to the following study authors who provided extra information on their studies:

- Victorija Krusinskiene: provided details of the number of subjects at each time point for Pirttiniemi 2005[56], Krusinskiene 2008[57].
- Nik Pandis: provided details of individual irregularity scores Pandis 2009 [92] and the standard deviations for time to alignment Pandis 2011 [79]
- Peter Miles: provided details of ages of subjects in the study sample for Ong 2011[94].
- Marcio Campos: provided corrected data for irregularity scores for Gravina 2013 [88]

REFERENCES

1. Proffit WR, Fields HW Jnr, Moray LJ. Prevalence of malocclusion and orthodontic treatment need in the United States: estimates from the NHANES III survey. *International Journal of Adult Orthodontics and Orthognathic Surgery*, 1998;13(2):97-106.

2. Helm S, Kreiborg S, Solow B. Psychosocial implications of malocclusion: A 15-year follow-up study in 30-year-old Danes. *American Journal of Orthodontics*, 1985 Feb;87(2):110-8.

3. Obilade OA, Sanu OO, da Costa OO. Impact of three malocclusion traits on the quality of life of orthodontic patients. *International Orthodontics*, 2016; 14(3): 366–385.

4. Helm S, Kreiborg S, Solow B. Psychosocial implications of malocclusion: A 15-year follow-up study in 30-year-old Danes. *American Journal of Orthodontics*, 1985 Feb;87(2):110-8.

5. DiBiase AT, Sandler PJ. Malocclusion, Orthodontics and Bullying. *Dental Update*, 2001 Nov;28(9):464-6.

6. British Orthodontic Society. Early Orthodontics. *British Orthodontic Society*. [Online] 2014. [Cited: November 7, 2016.] http://www.bos.org.uk/Museum-and-Archive/History-of-Orthodontics/Early-orthodontics.

7. General Dental Council. Look for a Specialist. *General Dental Council*. [Online] 2016. [Cited: November 7, 2016.] http://www.gdc-uk.org/membersofpublic/lookforaspecialist/Pages/default.aspx.

8. Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. *British Medical Journal*, 1995 Apr 29;310(6987):1122-6.

9. National Institute for Health and Clinical Excellence. Methods for development of NICE public. London : *NHS*, 2006. p. 36.

10. Hujoel, P. Grading the Evidence: The Core of EBD. J*ournal of Evidence Based Dental Practice*, September 2009, 9:122–124.

11. Greenhalgh, T. How to read a paper. London : *BMJ Books*, 2001.

12. gov.uk. About Us. *Public Health England*. [Online] [Cited: November 8, 2016.] https://www.gov.uk/government/organisations/public-health-england/about#who-we-are.

13. Egger M, Smith GD. Meta-analysis: Bias in location and selection of studies. Bristol : *British Medical Journal*, 1998 Jan 3;316(7124):61-6.

14. Pubmed. National Center for Biotechnology Information. [Online] U.S. *National Library of Medicine*, December 7, 2016. [Cited: December 7, 2016.] https://www.ncbi.nlm.nih.gov/pubmed.

15. CD, Mulrow. Systematic Reviews: Rationale for systematic reviews. *British Medical Journal*, 1994 Sep 3;309(6954):597-9.

16. Belsey, J. What is evidence-based medicine? http://www.bandolier.org.uk/. [Online] May 2009. [Cited: December 26, 2016.] http://www.bandolier.org.uk/painres/download/whatis/ebm.pdf.

17. Gelber RD, Goldhirsch A. Boston Meta-analysis: the fashion of summing-up. *Annals of Oncology*, 1991 2(7): 461-468.

18. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. *The Cochrane Collaboration*. [Online] March 2011. [Cited: November 9, 2016.] www.handbook.cochrane.org.

19. Cochrane. http://www.cochrane.org. [Online] The Cochrane Collaboration, 2016. [Cited: December 26, 2016.] http://www.cochrane.org/what-is-cochrane-evidence.

20. The Cochrane Collaboration. Chapter 4: Guide to the contents of a Cochrane protocol and review. [book auth.] *The Cochrane Collaboration*. The Cochrane Handbook, version 5.0.1. 5.0.1. s.l. : The Cochrane Collaboration, 2011.

21. Huang X, Lin J, Demner-Fushman D. Evaluation of PICO as a knowledge representation for clinical questions. *AMIA Annual Symposium Proceedings Archive*, 2006.

22. British standards, British Standards Institution. British standard glossary of dental terms. Michigan : *British Standards Institution*, 1983. 4492.

23. Tanaka MM, Johnston LE . The prediction of the size of unerupted canines and premolars in a contemporary orthodontic population. *Journal of Americal Dental Association*, 1974 Apr 1;88(4):798-801.

24. Oldfather, E. H. Hixon and R. E. Estimation Of The Sizes Of Unerupted Cuspid And Bicuspid Teeth. *The Angle Orthodontist*, **October 1958**, 28(4): 236-240.

25. Proffit WR, Fields HW, Sarver DM. Contemporary Orthodontics, 5th edition. *Elsevier Health Sciences*, 2014.

26. Kirschen RH, O'Higgins EA, Lee RT. The Royal London Space Planning: An integration of space analysis and treatment planning, Part I: Assessing the space required to meet treatment objectives. London : *American Journal of Orthodontics and Dentofacial Orthopedics*, 2000 Oct;118(4):448-55.

27. Kirschen RH, O'Higgins EA, Lee RT. The Royal London space planning: An integration of space analysis and treatment planning, Part II: The effect of other treatment procedures on space. London : *American Journal of Orthodontics and Dentofacial Orthopedics*, 2000 Oct;118(4):456-61.

28. Johal AS, Battagel JM. Dental crowding: a comparison of three methods of assessment. London : *European Journal of Orthodontics*, 1997 19: 543-551.

29. RM, Little. The Irregularity Index: A quantitative score of mandibular anterior alignment. Seattle : *American Journal of Orthodontics and Dentofacial Orthopaedics*, 1975 Nov;68(5):554-63.

30. Steele J, White D, Rolland S, Fuller E. Children's Dental Health Survey 2013 : Health and Social Care Information Centre, 2013.

31. Brook PH, Shaw WC. The development of an index of orthodontic treatment priority. *European Journal of Orthodontics*, 1989 Aug;11(3):309-20.

32. NHS England. Guide for comissioning dental specialties-Orthodontics. *NHS England*, 2015.

33. Tausche E, Luck O, Harzer W. Prevalence of malocclusions in the early mixed dentition and orthodontic treatment need. Dresden : *European Journal of Orthodontics*, 2004 26(3):237-44 ·

34. Burden DJ, Holmes A. The need for orthodontic treatment in the child population of the United Kingdom. *European Journal of Orthodontics*,1994,16 (5): 395-399.

35. Horowitz SL, Osborne RH, DeGeorge FV. A cephalometric study of craniofacial variation in adult twins. *Angle Orthodontics*, 1960;30:1–5.

36. Harris EF, Johnson MG. Heretability of craniometric and occlusal variables: A longitudinal sib analysis. *American Journal of Orthodontics and Dentofacial Orthopaedics*, 1991 Mar;99(3):258-68.

37. Lauweryns I, Carels C, Vlietinck R. The use of twins in dentofacial genetic research. *American Journal of Orthodontics and Dentofacial Orthopaedics*, 1993 Jan;103(1):33-8

38. Lundström, A. Nature vs nurture in dentofacial variation. *European Journal of Orthodontics*, 1984 6(1): 77-91.

39. The Heritability of Malocclusion: Part 2. The Influence of Genetics in Malocclusion. Mossey, PA. 3, 1999, *British Journal of Orthodontics*, 26:195-203.

40. Solow, B. The dentoalveolar compensatory mechanism: background and clinical implications. *British Journal of Orthodontics*, 1980 Jul;7(3):145-61.

41. Ahn HW, Baek SH. Skeletal anteroposterior discrepancy and vertical type effects on lower incisor preoperative decompensation and postoperative compensation in skeletal Class III patients. *Angle Orthodontics*, 2011 Jan;81(1):64-74.

42. Ishikawa H, Nakamura S, Iwasaki H, Kitazawa S, Tsukada H, Sato Y. Dentoalveolar compensation related to variations in sagittal jaw relationships. *Angle Orthodontics*, 1999 Dec;69(6):534-8.

43. Peck S, Peck L, Kataja M. Class II division 2 malocclusion: A heritable pattern of small teeth in well-developed jaws. *Angle Orthodontics*, 1998 Feb;68(1):9-20.

44. Sassouni, V. A classification of skeletal facial types. February 1969, *American Journal of Orthodontics*, 1969 Feb;55(2):109-23.

45. Venkataraghavan K, Anantharaj A, Nihal NK. Supplemental Primary Tooth: A Review & Report Of A Rare Occurrence. *International Journal of Clinical Dental Science*, 2011, Vol. 2.

46. Garvey MT, Barry HJ, Blake M. Supernumerary teeth--an overview of classification, diagnosis and management. *Journal of Canadian Dental Association*, 1999 Dec;65(11):612-6.

47. Bowden, BD. A longitudinal study of the effects of digit- and dummy-sucking. *American Journal of Orthodontics*, **December 1966** 52(12): 887–901

48. Mistry P, Moles DR, O'Neill J, Noar J. The occlusal effects of digit sucking habits amongst schoolchildren in Northamptonshire (UK). *Journal of Orthodontics*, 2010;37:87-92.

49. Langlois JH, Kalakanis L, Rubenstein AJ, Larson A, Hallam M, Smoot M. Maxims or Myths of Beauty? A Meta-Analytic and Theoretical Review . Texas : *Psychological Bulletin* , 2000 May;126(3):390-423.

50. Lerner RM, Karabenick SA, Stuart JL. Relations among physical attractiveness, body attitudes, and selfconcept in male and female college students. *Journal of Psychology*, 1973 Sep;85(1st Half):119-29.

51. Seehra J, Newton T, Dibiase A T. MSc thesis: Bullying in orthodontic patients and its relationship to malocclusion, self-esteem and oral health related quality of life. London : *King's College London*, 2009.

52. Lindsay, S. A guide to purchasers of clinical psychology services. Briefing paper no.11: *Clinical Psychology in Dentistry*. Leicester: Division of Clinical Psychology. The British Psychological Society, 1996.

53. Seehra J, Fleming PS, Newton T, DiBiase AT. Bullying in orthodontic patients and its relationship to malocclusion,self-esteem and oral health-related quality of life. *Journal of Orthodontics*, 2011 Dec;38(4):247-56

54. Seehra J, Newton JT, DiBiase AT. Bullying in schoolchildren – its relationship to dental appearance and psychosocial implications: an update for GDPs. *British Dental Journal*, 2011 May 14;210(9):411-5

55. Guideline on Management of the Developing Dentition and Occlusion in Pediatric Dentistry. Clinical Affairs Committee. *American Academy Of Pediatric Dentistry*, 2014, Vol. 37. 2005-2006;27(7 Suppl):143-55.

56. Pirttiniemi P, Kantomaa T, Mäntysaari R, Aila Pykäläinen A, Krusinskiene V, Laitala T, Karikko J. The effects of early headgear treatment on dental arches and craniofacial morphology: an 8 year report of a randomized study. *European Journal of Orthodontics,* 2005 27(5): 429-436.

57. Krušinskiene˙ V, Kiuttu P, Julku J , Silvola AS, Kantomaa T, Pirttiniemi P. A randomized controlled study of early headgear treatment on occlusal stability — a 13 year follow-up. *European Journal of Orthodontics*, 2008 30 (4): 418-424.

58. Mäntysaari R, Kantomaa T, Pirttiniemi P, Pykäläinen A. The effects of early headgear treatment on dental arches and craniofacial morphology: a report of a 2 year randomized study. *European Journal of Orthodontics*, 2005 27 (5): 429-436.

59. British Orthodontic Society. Managing the Developing Occlusion: A guide for dental practitioners. www.bos.org.uk. [Online] April 2010. [Cited: December 12, 2016.] http://www.bos.org.uk/Portals/0/Public/docs/Making%20a%20Referral/Managing-the-Developing-Occulusion-Updated-Apr10.pdf.

60. Kau CH, Durning P, Richmond S, Miotti FA, Harzer W. Extractions as a form of interception in the developing dentition: a randomized controlled trial. *Journal of Orthodontics*, 2004 Jun;31(2):107-14.

61. Sayın MO, Turkkahraman H. Effects of Lower Primary Canine Extraction on the Mandibular Dentition. *Angle Orthodontist*, 2006, January 2006, 76(1): 31-35.

62. Viglianisi, A. Effects of lingual arch used as space maintainer on mandibular arch dimension: a systematic review. *American Journal of Orthodontics and Dentofacial Orthopaedics*., 2010 Oct;138(4):382.e1-4

63. Rebellato J, Lindauer SJ, Rubenstein LK, Isaacson RJ, Davidovitch M, Vroom K. Lower arch perimeter preservation using the lingual arch. *American Journal of Orthodontics and Dentofacial Orthopedics* , 1997 Oct;112(4):449-56.

64. Villalobos FJ, Sinha PK, Nanda RS. Longitudinal assessment of vertical and sagittal control in the mandibular arch by the mandibular fixed lingual arch. *American Journal of Orthodontics and Dentofacial Orthopedics*, 2000 Oct;118(4):366-70.

65. Hashisha DI, Mostafa YA. Effect of lip bumpers on mandibular arch dimensions. Cairo : *American Journal of Orthodontics and Dentofacial Orthopedics*, 2009 Jan;135(1):106-9.

66. Tai K, Hotokezaka H, Park JH, Tai H, Miyajima K, Choi M, Kai LM, Mishima K. Preliminary cone-beam computed tomography study evaluating dental and skeletal changes after treatment with a mandibular Schwarz appliance : *American Journal of Orthodontics and Dentofacial Orthopedics*, 2010 Sep;138(3):262.e1-262.e11

67. Myrlund R, Dubland M, Keski-Nisula K, Kerosuo H. One year treatment effects of the eruption guidance appliance in 7- to 8-year-old children: a randomized clinical trial. *European Journal of Orthodontics* , 2014 37 (2): 128-134

68. Sandler J, Benson PE, Doyle P, Majumder A, O'Dwyer J, Speight P, Thiruvenkatachari B, Tinsley D. Palatal implants are a good alternative to headgear: a randomized trial. American Journal of Orthodontics and Dentofacial Orthopedics, January 2008, *American Journal of Orthodontics and Dentofacial Orthopedics*, 2008 Jan;133(1):51-7

69. Cetlin NM, Ten Hoeve A. Nonextraction treatment. *Journal of Clinical Orthodontics*, 1983 17(6): 396-413.

70. Bondemark, L. A comparative analysis of distal maxillary molar movement produced by a new lingual intra-arch Ni-Ti coil appliance and a magnetic appliance. *European Journal of Orthodontics*, 2000 22(6):683-95

71. Sfondrini MF, Cacciafesta V, Sfondrini G. Upper molar distalization: a critical analysis. *Orthodontics and Craniofacial Research*, 2002, 5(2): 114–126

72. Jambi S, Thiruvenkatachari B, O'Brien KD, Walsh T. Orthodontic treatment with appliances which move the upper molar teeth backwards. *Cochrane Library*. [Online] October 23, 2013. [Cited: December 26, 2016.] http://www.cochrane.org/CD008375/ORAL_orthodontic-treatment-with-appliances-which-move-the-upper-molar-teeth-backwards_.

73. Grec RH de C, Janson G, Branco NC, Moura-Grec PG, Patel MP, Henriques JFC. Intraoral distalizer effects with conventional and skeletal anchorage: A meta-analysis. *American Journal of Orthodontics and Dentofacial Orthopaedics*, 2013 May;143(5):602-15.

74. McNally MR, Spary DJ, Rock WP. A randomized controlled trial comparing the quadhelix and the expansion arch for the correction of crossbite. *Journal of Orthodontics*, 2005 Mar;32(1):29-35.

75. Lagaravere MO, Majorb PW, Flores-Mirc C. Long-term Skeletal Changes with Rapid Maxillary Expansion: A Systematic Review. *Angle Orthodontist*, 2005 75 (6): 1046-1052.

76. Magnusson A, Bjerklin K, Nilsson P, Marcusson A. Surgically assisted rapid maxillary expansion: long-term stability. *European Journal of Orthodontics*, 2008 31 (2): 142-149.

77. Jian F, Lai W, Furness S, McIntyre GT, Millett DT, Hickman J, Wang Y. Initial arch wires for tooth alignment during orthodontic treatment with fixed appliances. www.cochrane library.com. [Online] 2013. [Cited: December 26, 2016.]

78. Mandall NA, Lowe C, Worthington HV, Sandler J, Derwent S, Abdi-Oskouei M, Ward S. Which orthodontic archwire sequence? A randomized clinical trial. *The European Journal of Orthodontics*, 2006, 28(6): 561-566.

79. Pandis N, Polychronopoulou A, Katsaros C, Eliades T. Comparative assessment of conventional and self-ligating appliances on the effect of mandibular intermolar distance in adolescent nonextraction patients: A single-center randomized controlled trial. A*merican Journal of Orthodontics and Dentofacial Orthopedics*, September 2011,140(3):e99-e105

80. Atik E, Ciger S. An assessment of conventional and self-ligating brackets in Class I maxillary constriction patients. *Angle Orthodontist*, 2014, 84(4): 615-622.

81. Fleming PS, Johal A. Self-Ligating Brackets in Orthodontics: A systematic review. *Angle Orthodontist*, 2010 May;80(3):575-84.

82. Miles P, Weyant R. Porcelain brackets during initial alignment: are self-ligating cosmetic brackets more efficient? *Australian Orthodontic Journal*, May 2010 26(10): 21-26.

83. Pandis N, Polychronopoulou A, Eliades T. Active or passive self-ligating brackets? A randomized controlled trial of comparative efficiency in resolving maxillary anterior crowding in adolescents. *American Journal of Orthodontics and Dentofacial Orthopaedics*, January 2010, 137: 1-6.

84. Phan X, Ling PH. Clinical Limitations of Invisalign. *Journal of Canadian Dental Association*, April 2007 73(3): 263-266.

85. Society, British Orthodontic. British Orthodontic Society- Clear Aligners. *British Orthodontic Society*. [Online] 2014. [Cited: December 26, 2016.] http://www.bos.org.uk/Public-Patients/Orthodontics-for-Children-Teens/Treatment-brace-types/Removable-appliances/Clear-aligners.

86. Rossinia G, Parrini S, Castroflorio T, Deregibus A, Debernardi CL. Efficacy of clear aligners in controlling orthodontic tooth movement: A systematic review. *Angle Orthodontist,* 2015, 85: 881-889.

87. Davidovitch M, McInnis D, Lindauer SJ. The effects of lip bumper therapy in the mixed dentition. 1997, American Journal of Orthodontics and Dentofacial Orthopedics , 111: 52-58.

88. Gravina MA, Brunharo IHVP, Fraga MR, Artese F, da Silva Campos MJ, Vitral RWF, Quintão CCA. Clinical evaluation of dental alignment and leveling with three different types of orthodontic wires. *Journal of Orthodontics*, December 2013 18(3): 31-37.

89. Mantysaari R, Kantomaa T, Pirttiniemi P, Pykalainen A. The effects of early headgear treatment on dental arches and craniofacial morphology: a report of a 2 year randomized study. *European Journal of Orthodontics,* 2005 27(5): 429-436.

90. Miles P, Smith H, Weyant R, Rinchuse DJ. The effects of a vibrational appliance on tooth movement and patient discomfort: a prospective randomised clinical trial. Australian Orthodontic Journal, November 2012 Nov;28(2):213-8.

91. O'Brien K, Lewis D, Shaw W, Combe E. A clinical trial of aligning archwires. 1990, *European Journal of Orthodontics*, 12(4)*,* 380-384.

92. Pandis N, Polychronopoulou A, Eliades T. Alleviation of mandibular anterior crowding with copper-nickel-titanium vs nickel-titanium wires: a double blind randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopaedics*, 2009 136(2):152.e1-7

93. Sebastian, B. Alignment efficiency of superelastic coaxial nickel-titanium vs superelastic single-stranded nickel-titanium in relieving mandibular anterior crowding A randomized controlled prospective study. *Angle Orthodontist*, 2012 Jul;82(4):703-8.

94. Ong E, Ho C, Miles P. Alignment efficiency and discomfort of three orthodontic archwire sequences: a randomized clinical trial. *Journal of Orthodontics*, 2011 Mar;38(1):32-9

95. Woodhouse NR, DiBiase AT, Johnson N, Slipper C, Grant J, Alsaleh M, Donaldson ANA, Cobourne MT. Supplemental Vibrational Force During Orthodontic Alignment: A Randomized Trial. *Journal of Dental Research* , 2015, 94:682 –689

96. Abu Alhaija E, Al-Khateeb S. Skeletal, dental and soft tissue changes in Class III patients treated with fixed appliances and lower premolar extractions. *Australian Orthodontic Journal* , May 2011, 27: 40-45.

97. Barlin S, Smith R, Reed R and Sandy J, Ireland AJ. A retrospective randomized double-blind comparison study of the effectiveness of Hawley vs vacuum-formed retainers. *Angle Orthodontist*, 2011, 81: 404-409.

98. Baumrind S, Korn EL, Boyd RL, Maxwell R .The decision to extract: Part II. Analysis of clinicians'stated reasons for extraction. *American Journal of Orthodontics and Dentofacial Orthopedics*, April 1996, 109: 393-402.

99. Dai JY, Zhang MM, Sun M, Ni H. Treating high angle bimaxillary protrusion with three kinds of extraction method: a clinical study. *China Journal of Stomatology,* 2009 , China Journal of Stomatology, 27: 268-271.

100. Freitas KM, Janson G, Tompson B, de Freitas MR, Simao TM, Valarelli FP, Cancado RH. Post treatment and physiologic occlusal changes comparison. *Angle Orthodontist*, 2013, 83: 239-245.

101. Fan CX, Sun R, Wu LP. The effect of first premolar extraction on vertical dimension in skeletal Class I patients .*Shanghai journal of Stomatology,* 2009, 18: 592-595 .

102. Heiser W, Niederwanger A, Bancher B, Bittermann G, Neunteufel N, Kulmer S. Three-dimensional dental arch and palatal form changes after extraction and nonextraction treatment. Part 1. Arch length and area. *American Journal of Orthodontics and Dentofacial Orthopedics* , July 2004, 126: 71-81.

103. Keski-Nisula K, Hernesniemi R, Heiskanen M, Keski-Nisula L, Varrela, J. Orthodontic intervention in the early mixed dentition: a prospective, controlled study on the effects of the eruption guidance appliance. *American Journal of Orthodontics and Dentofacial Orthopedics* , February 2008, 133: 254-260.

104. Miyake H, Ryu T, Himuro, T. Effects on the dental arch form using a preadjusted appliance with premolar extraction in Class I crowding. Angle Orthodontist, 2008, 78:1043-1049.

105. Nagalakshmi S, Sriram G, Balachandar K, Dhayanithi D. A comparative evaluation of mandibular incisor decrowding with coaxial and optiflex arch wires and their load-deflection rates . *Journal of pharmacy & bioallied sciences,* 2014, 6:118-121.

106. Ong HB, Woods, MG. An occlusal and cephalometric analysis of maxillary first and second premolar extraction effects. *Angle Orthodontist,* 2001, 71: 90-102.

107. Owais AI, Rousan ME, Badran SA, Abu Alhaija ES. Effectiveness of a lower lingual arch as a space holding device. *European Journal of Orthodontics*, 2011, 33: 37–42 .

108. Pandis N, Polychronopoulou A, Eliades T. Self-ligating vs conventional brackets in the treatment of mandibular crowding: a prospective clinical trial of treatment duration and dental effects . *American Journal of Orthodontics and Dentofacial Orthopedics* , 2007, 132: 208-215.

109. Pandis N, Polychronopoulou A, Makou M, Eliades T. Mandibular dental arch changes associated with treatment of crowding using self-ligating and conventional brackets. *European Journal of Orthodontics*, 2010, 32: 248-253.

110. Ruf S, Pancherz H. Dentoskeletal effects and facial profile changes in young adults treated with the Herbst appliance. *Angle Orthodontist*, 1999, 69: 239-246.

111. Sucuru R, Atagun C, Suytarhan A. Poster abstract: The effects of intermaxillary lip bumper in cases with class II div 1 malocclusion. *68th Congress, European Orthodontic Society Venice*. 1992. Vol. 113.

112. Vajaria R, BeGole E, Kusnoto B, Galang MT, Obrez A. Evaluation of incisor position and dental transverse dimensional changes using the Damon system. *Angle Orthodontics,* July 2011 , 81: 647-52.

113. Yu YL, Tang GH, Gong FF, Chen LL, Qian YF. A comparison of rapid palatal expansion and Damon appliance on non-extraction correction of dental crowding. *Shanghai journal of stomatology*, 2008, 17: 237-242 .

114. Almeida NV, Silveira GS, Pereira DM, Mattos CT, Mucha JN Interproximal wear versus incisors extraction to solve anterior lower crowding: a systematic review. *Journal of orthodontics* , 2015, 20: 66-73.

115. Clements KM, Bollen AM, Huang G, King G, Hujoel P, Ma, T. Activation time and material stiffness of sequential removable orthodontic appliances. Part 2: dental improvements. *American Journal of Orthodontics and Dentofacial Orthopedics* , 2003, 124: 502-508.

116. Cobb NW, Kula KS, Phillips C, Proffit WR. Efficiency of multi-strand steel, superelastic Ni-Ti and ion-implanted Ni-Ti archwires for initial alignment. *Clinical Orthodontic Research*, 1998 1(1): 12-19.

117. DiBiase AT, Nasr IH, Scott P, Cobourne MT. Duration of treatment and occlusal outcome using Damon3 self-ligated and conventional orthodontic bracket systems in extraction patients: a prospective randomized clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics*, 2011, 139: 111-116.

118. Fleming PS, DiBiase AT, Sarri G, Lee RT. Efficiency of mandibular arch alignment with 2 preadjusted edgewise appliances. *American Journal of Orthodontics and Dentofacial Orthopedics*, May 2009, 135: 597-602.

119. Harradine N, Pearson MH, Toth B. The effect of extraction of third molars on late lower incisor crowding: a randomized controlled trial. *British Journal of Orthodontics,* 1998, 25: 117-122.

120. Kau C, Kantarci A, Shaughnessy T, Vachiramon A, Santiwong P, de la Fuente A, Skrenes D, Ma D, Brawn P. Photobiomodulation accelerates orthodontic alignment in the early phase of treatment. *Progress in Orthodontics,* 2013, 14:1-9.

121. Lindqvist B, Thilander B. Extraction of third molars in cases of anticipated crowding in the lower jaw. *American Journal of Orthodontics,* 1982, 81: 130-139.

122. Scott P, DiBiase AT, Sherriff M, Cobourne MT. Alignment efficiency of Damon 3 self-ligating and conventional orthodontic bracket systems: a randomized clinical trial. *American Journal of Orthodontics and Dentofacial Orthopedics*, 2008 134: 1-8.

123. Serafim CMC, Gurgel JA, Tiago CM, Tavarez RRJ,Filho EMM . Clinical efficiency of two sequences of orthodontic wires to correct crowding of the lower anterior teeth. T*he Scientific World Journal,* June 2015, pp1-5.

124. Soldanova M, Leseticky O, Komarkova L, Dostalova T, Smutny V, Spidlen M. Effectiveness of treatment of adult patients with the straightwire technique and the lingual two-dimensional appliance. *European Journal of Orthodontics,* 2012, 34: 674–680.

125. Talapaneni AK, Supraja G, Prasad M, Kommi PB. Comparison of sagittal and vertical dental changes during first phase of orthodontic treatment with MBT vs ROTH prescription. *Indian Journal of Dental Research*, 2012, 23:182-186.

126. Taner T, Haydar B, Kavuklu I, Korkmaz A. Short-term effects of fiberotomy on relapse of anterior crowding. *American Journal of Orthodontics and Dentofacial Orthopedics* , December 2000, 118: 617-623.

127. Wahab RM1, Idris H, Yacob H, Ariffin SH. Comparison of self- and conventional-ligating brackets in the alignment stage. *European Journal of Orthodontics,* April 2012, 34: 176-181.

128. Wang Y, Jian F, Lai W, Zhao Z, Yang Z, Liao Z, Shi Z, Wu T, Millett DT, McIntyre GT, Hickman J. Initial arch wires for alignment of crooked teeth with fixed orthodontic braces (Review). www.cochrane.co.uk. [Online] 2010. [Cited: December 28, 2016.]

129. Altug H, Bengi O, Akin E, Karacay S. Dentofacial Effects of Asymmetric Headgear and Cervical Headgear with Removable Plate on Unilateral Molar Distalization. *Angle Orthodontist,* 2005, Angle Orthodontist, 175: 585-592.

130. Thickett E, Power S. A randomized clinical trial of thermoplastic retainer wear. European Journal of Orthodontics , 2010, 32:1–5 .

131. Bondemark L, Karlsson I. Extraoral vs Intraoral Appliance for Distal Movement of Maxillary First Molars: A Randomized Controlled Trial. *Angle Orthodontist*, 2005 175: 699-706.

132. Markovic E, Fercec J, Scepan I, Glisic B, Nedelkkovic N, Juloski J, Rudolf R. The Correlation between Pain Perception among Patients with Six Different Orthodontic Archwires and the Degree of Dental Crowding. *Srp Arh Celok Lek,* 2015, 143:134-140.

133. Okay C, Gulsen A, Keykubat A, Ucem TT, Yuksel S. A comparison of the effects of 2 mandibular anchorage systems used with a 3 dimensional bimetric maxillary distalizing arch. *World Journal of Orthodontics*, 2006, 7:125-133.

134. Rowland H, Hichens L, Williams A, Hills D, Killingback N, Ewings P, Clark S, Ireland AJ, Sandy JR. The effectiveness of Hawley and vacuum-formed retainers: a single-center randomized controlled trial. *American Journal of Orthodontics and Dentofacial Orthopedics,* December 2007, 132: 730-737.

135. Sandhu SS, Sandhu J. A randomized clinical trial investigating pain associated with superelastic nickel-titanium and multistranded stainless steel archwires during the initial leveling and aligning phase of orthodontic treatment. *Journal of orthodontics,* 2013, 40: 276-285 .

136. Shawesh M, Bhatti B, Usmani T, Mandall N.Hawley retainers full- or part-time? A randomized clinical trial. *European Journal of Orthodontics* , 2010 , 32:165–170.

137. Silva RGC, Kaieda AK, Paranhos LR, Angelieri F, Torres FC, Scanavini MA. A Comparative Study Between Lip Bumper and Headgear as Maxillary Molar Retainers Following Distalization. *International Journal of Orthodontics,* 2012, 23: 29-34.

138. Strahm C, De Sousa AP, Grobéty D, Mavropoulos A, Kiliaridis S. Is bodily advancement of the lower incisors possible? *European Journal of Orthodontics*, August 2009 , 31: 425-431.

139. Wortham JR, Dolce C, McGorray SP, Le H, King GJ, Wheeler TT . Comparison of arch dimension changes in 1 phase vs 2 phase treatment of class II malocclusion. *American Journal of Orthodontics and Dentofacial Orthopedics* 2009 Jul;136(1):65-74.

140. West AE, Jones ML, Newcombe RG. Multiflex versus superelastic : a randomized clinical trial of the tooth alignment ability of arch wires . *American Journal of Orthodontics and Dentofacial Orthopedics* , 1995, 108(5): 464–471

141. Houston, WJB. The analysis of errors in orthodontic measurements. *American Journal of Orthodontics,* 1983, 83: 382-390.

142. Fleming PS, Strydom H, Katsaros C, MacDonald LCI, Curatolo M, Fudalej P, Pandis N. Non-pharmacological interventions for alleviating pain during orthodontic treatment. *The Cochrane Library*. [Online] December 23, 2016. [Cited: January 1, 2017.] http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD010263.pub2/full.

143. El-Angbawi A, McIntyre GT, Fleming PS, Bearn DR. Non-surgical adjunctive interventions for accelerating tooth movement in patients undergoing fixed orthodontic treatment. *The Cochrane Library*. [Online] November 18, 2015. [Cited: January 1, 2017.] http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD010887.pub2/full.

144. Yu Y, Sun J, Lai W, Wu T, Koshy S, Shi Z. Interventions for managing relapse of the lower front teeth after orthodontic treatment. *The Cochrane Library*. [Online] September 6, 2013. [Cited: JAnuary 1, 2017.] http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD008734.pub2/full.

145. Wang X, Wang R, Zhang DL. Evaluation on effectiveness of individual lingual orthodontics and labial straight wire orthodontics by X-ray cephalometrics. *Journal of Jilin University(Medicine Edition)* 2014, pp. 861-865.

146. Straus SE, Glasziou P, Richardson WS. Evidence Based Medicine, Fourth Edition. *Elsevier,* 2011.

147. Pencheon, D. Managing demand: Matching demand and supply fairly and efficiently. *British Medical Journal*, 1998 May 30; 316(7145): 1665–1667.

148. Brook P, Shaw WC. The development of an index of orthodontic treatment priority. Manchester : *European Journal of Orthodontics,* 1989 Aug;11(3):309-20.

149. Bolton, W. Disharmony In Tooth Size And Its Relation To The Analysis and Treatment of Malocclusion. Seattle *Angle Orthodontics*, July 1958, 28(3): 113-130.

150. Bolton, W. The Clinical Application of a Tooth Size Analysis. Seattle : *American Journal of Orthodontics*, July 1962 48(7): 504–529

APPENDIX 1: Characteristics of included studies

Atik 2014

METHODS	**Setting:** Department of Orthodontics, Hacettepe University, Ankara, Turkey. **Design:** Parallel (2 arms) **No. of centres:** 1 **Study duration:** Not mentioned
PARTICIPANTS	**Inclusion criteria**: “between 13 and 17 years of age at the start of the treatment, moderate maxillary and mandibular crowding, a Class I malocclusion, and a dentally constricted maxillary arch.” **Exclusion criteria:** None stated **Orthodontic intervention:** Fixed appliances **Patient sampling:** N=33 selected Gp 1= 17 females (mean age of 14.5 6 1.2 years) Gp 2= 16 females (mean age of 14.8 6 1.0 years) No dropouts
INTERVENTIONS	**Quad helix and conventional Roth fixed app vs self-ligating Damon fixed app**

OUTCOMES	Upper incisor to NA, Upper incisor to Frankfurt Plane, Lower incisor to NB all pre and post treatment Pain (VAS) at 4-24hrs, 24-3 days, 3-7 days, 7-30 days Gingival Index, Bleeding Index and Probing Depths all at 6 months after-prebond up, end of treatment-pre bond up, end of treatment-6 months after
NOTES	

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Quote: "The subjects were randomly allocated to either treatment system."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Not reported Quote: "The subjects were randomly allocated to either treatment system."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not reported Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Inadequate information regarding blinding Quote: " The pretreatment and

		posttreatment lateral cephalograms of each patient were traced by one examiner"
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: No apparent dropouts Quote: " "
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: Reported on all outcomes Quote:
Other bias	High/Low/ Unclear risk	Comment: All patients female so not a true reflection of orthodontic population

Davidovitch 1997

METHODS	**Setting:** Medical College of Virginia, postgraduate orthodontic clinic **Design:** Parallel (2 arms) **No. of centres:** 1 **Study duration:** 6 months
PARTICIPANTS	**Inclusion criteria**: (1) white ethnicity, (2) 3 to 8 mm mandibular arch length deficiency, (3) presence of the mandibular deciduous second molars, and (4) Class I, Division 2 malocclusion. **Exclusion criteria:** None stated **Orthodontic intervention:** Lip bumper **Patient sampling:**

	N=34 selected Gp 1= 16 Gp 2= 18 7.9 to 13.1 years (mean = 10.2) Dropouts not reported upon
INTERVENTIONS	**Lip bumper vs control**
OUTCOMES	Crowding (mm) Lower molar angulation Lower incisor angulation All at start to 6/12 tomographically and cephalometrically
NOTES	Funding source not cited

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Inadequate information on how randomisation was carried out therefore unable to make judgement on appropriateness Quote: "Subjects were randomly assigned to either the experimental (N = 16), or control (N = 18) group." “Assignment of each subject to either of the populations was random,”

Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Not reported Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not reported; unable to blind. Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not specifically reported whether lip bumper removed for impressions / radiographs Quote: “All data were independently measured by two observers.” “In addition, interobserver reliability was gauged."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Unclear; number reported not quoted Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: Expected outcomes reported Quote: "Comparisons were made for changes in arch length and perimeter, intercanine and deciduous molar distances, crowding, and linear and angular changes in molar and incisor positions."
Other bias	High/Low/ Unclear risk	Comment: Difference between tomographic and lateral ceph measurements – changes sig diff with one and not the other.

		Quote: Tomographic data (-6.31 ° _+ 1.28 °) showed approximately twice the angulation change as that measured from lateral cephalometric radiographs (-3.38 ° + 3.67°). The average change in molar angulation of experimental versus control subjects was found to be statistically significant when observed tomographically (p < 0.02).

Gravina 2013

METHODS	**Setting:** Brazil, Rio de Janeiro **Design:** Parallel (3 arms) **No. of centres:** 1 **Study duration:** 8 weeks

PARTICIPANTS	**Inclusion criteria**: presence of all erupted permanent teeth except for second and third molars; no previous orthodontic treatment; no indications for tooth extraction; overbite and overjet that allowed brackets to be placed on the lower teeth without occlusal interferences; level of crowding and teeth position that allowed a maximum deflection of 2 mm in the archwire when inserted in the bracket slots, and good conditions of oral hygiene and health. **Exclusion criteria:** None stated **Orthodontic intervention:** Different archwires **Patient sampling:** N=36 selected Group I (n = 11): stainless steel 0.014-inch lower; Group II (n = 12): multistranded stainless steel 0.015-inch lower Group III (n = 13): superelastic nickel-titanium 0.014-inch lower Ages: 14+/- 2 yrs 18 Males, 18 females Dropouts not reported upon
INTERVENTIONS	**SS vs multi stranded SS vs NiTi**

OUTCOMES	Irregularity Index (mm) Lower incisor to gonian-gnathion
NOTES	The authors report no commercial, proprietary or financial interest in the products or companies described in this article

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Inadequate information on how randomisation was carried out therefore unable to make judgement on appropriateness Quote: " The patients were randomly divided into 3 groups according to the type of precontoured archwire used:“
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Inadequate information regarding how allocation concealment was carried out therefore unable to make judgement on appropriateness Quote: " The patients were randomly divided into 3 groups according to the type of precontoured archwire used:“
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not reported. Not possible to blind clinicians Quote:

Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: : Inadequate information regarding blinding therefore unable to make judgement on appropriateness Quote: "On lateral cephalometric radiographs, the structures directly related to the position of the lower incisors were traced and linear and angular measurements were obtained for T1 and T2 " "The irregularity index (II) and the depth of the curve of Spee (CS) were measured in the lower dental casts for T1 and T2."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Not reported Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: Expected outcomes reported Quote: " Means and standard deviations are presented for all cephalometric variables, irregularity index and curve of Spee at T1 (Table 1) and T2 (Table 2)."
Other bias	High/Low/ Unclear risk	Comment: No obvious bias noted Quote:

Kau 2004

METHODS	**Setting:** Dental clinics in Italy, Germany, Wales **Design:** Parallel (2 arms)

	No. of centres: 3 **Study duration:** 1 year, follow up 2 years
PARTICIPANTS	**Inclusion criteria**: patients should be Caucasian aged between 8 and 9 years old; crowding of the lower incisors greater than or equal to 6 mm, according to the irregularity index of Little (1975); Class I type occlusion as indicated by the molar relationship; the lower molars should have a good long-term prognosis; overbite should be within normal limits. **Exclusion criteria:** None stated **Orthodontic intervention:** Extraction of lower Cs **Patient sampling:** N=97 selected Group I (n = 55): XLA lower Cs Group II (n = 42): No treatment Ages: 8-9 yrs Gender not specified Caucasian 14 Dropouts
INTERVENTIONS	**Extraction of lower Cs vs Control**

OUTCOMES	Irregularity Index (mm) Lower incisor angle Arch length change
NOTES	**Funding source:** “This study was supported by a general research grant from the Wales Office of Research and Development for Health and Social Care.”

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate method of randomisation used Quote: "Simple randomization was the method of allocation treatment. A restricted randomization of allocation was used in blocks of 50 to ensure that equal numbers of patients were allocated to each of the treatment groups."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate method of allocation concealment Quote: "The random allocation was then concealed in envelopes labeled with the study identification number and held in a central place."

Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Pts & clinician unable to be blinded; assessors was blind. Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Outcome assessors blinded Quote: " Observer bias was reduced by ensuring that the examiner was blind to whether the patient had received an extraction or non-extraction treatment "
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Adequate response rate Quote: "53/55 (96%) Xn group; 30/42 (71%) non-Xn group followed up; overall 83/97 86% "
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported. Quote: "*place corresponding quote here*"
Other bias	High/Low/ Unclear risk	Comment: Unsure as to how many patients came from each centre and characteristics of participants from each centre, statistically signifiant differences in baseline groups

Krusinskiene 2008

METHODS	**Setting:** Finland, Department of Oral Development and Orthodontics, Institute of Dentistry **Design:** 2 arms **No. of centres: 3** **Study duration:** 13 years
PARTICIPANTS	**Inclusion criteria**: A need for orthodontic treatment due to moderate crowding and a Class II tendency. The crowding was clinically diagnosed as moderate, based on the degree of space deficiency in the anterior regions of the dental arches **Exclusion criteria:** None stated **Orthodontic intervention:** Cervical pull HG **Patient sampling:** N= 64 children of both sexes (40 males and 28 females) Group I (n=34): Low pull HG Group II (n = 34): Minor interceptive tx Ages: aged 7.6 years [standard deviation (SD) 0.3 years] 54 analysed, 14 dropouts/refusals
INTERVENTIONS	**Cervical pull HG vs Control**

OUTCOMES	Irregularity Index (mm)
NOTES	**Funding source:** Research grant from the European Orthodontic Society.

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate method of randomisation used Quote: "This was undertaken by one author (TK) using random numbers."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate method of allocation concealment Quote: "To conceal the allocation, most of the practitioners who undertook the treatment were not given information concerning the aim or rationale of the study."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Pts & clinician unable to be blinded; assessors not mentioned Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not mentioned Quote: "All measurements were made by one author (VK) directly on dental casts using a digital calliper with an accuracy of 0.01 mm. Dental aesthetics was evaluated in the patients by two calibrated

		observers (PK and ASS) using the AC of IOTN scores (Evans and Shaw, 1987) at the last follow-up (T4)."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Low response rate Quote:" Fifty-three patients (83 per cent of the total study group) who continued to the second phase of treatment at T2 completed the follow-up at T3 and full records were available. Thirty-four subjects (53 per cent of the total study group) attended a recall appointment at T4 at the mean age of 20.6 years."
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported. Quote: "Little's Irregularity Index (LII; Little, 1975) was measured as the sum of the linear displacements of the anatomical contact points of each mandibular incisor from the adjacent tooth anatomical points for all time periods." "No significant differences were found between the HG and control groups in LII at any time period (Table 2, Figure 4a)"
Other bias	High/Low/ Unclear risk	Comment:

Mantysaari 2004

METHODS	**Setting:** Finland
	Design: 2 arms
	No. of centres: 3?
	Study duration: 16 months
PARTICIPANTS	**Inclusion criteria**: None specifically stated
	"Children in need of orthodontic treatment due to moderate crowding and a class II tendency were selected for comprehensive orthodontic examination."
	Exclusion criteria: None stated
	Orthodontic intervention: Cervical pull HG
	Patient sampling:
	N= 68 children of both sexes (40 males and 28 females)
	Group I (n=34): Low pull HG
	Group II (n = 34): Minor interceptive tx
	Ages: aged 7.6 years [standard deviation (SD) 0.3 years]
	No dropouts
INTERVENTIONS	**Cervical pull HG vs Control**
OUTCOMES	Upper incisor to max plane angle
	Lower incisor to mand plane angle

NOTES **Funding source:** not stated

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Method not stated Quote: "The children were randomly divided into two groups of equal size, matched according to gender."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment Not reported Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment : Pts & clinician unable to be blinded; assessors not mentioned Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not reported Quote:
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: 30/34 reported in both groups (88%) Quote: “Angular measurements made on the cephalograms of children in the headgear (n=30) and control (n=30) groups”

Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported Quote:
Other bias	High/Low/ Unclear risk	Comment: It is unclear what impact these 'interceptive' treatments will have had [In the second group, which served as the control, only interceptive procedures were performed during the follow-upperiod". "During the period T0–T2, treatment procedures in the control group included any necessary interceptive procedures. These included extraction of the upper primary canines in 38 per cent of the subjects and of the lower primary canines in 35 per cent, to ease the eruption of the lateral incisors. In addition, in 19 per cent of the patients in the control group, some interdental stripping was carried out.."]

Miles 2010

METHODS	**Setting:** Caloundra, Queensland, Australia **Design:** 2 arms **No. of centres: 1** **Study duration:** 10 weeks

PARTICIPANTS	**Inclusion criteria: none stated** **Exclusion criteria: none stated** **Orthodontic intervention:** SL Innovation brackets **Patient sampling:** N= 68 consecutive subjects Group I (n= 34) Group II (n = 34) Ages: Overall mean age at the conclusion of the trial was 13.5 + 1.5 years 19M and 11F per group
INTERVENTIONS	**SL (Innovation) vs Conventional (Clarity) brackets**
OUTCOMES	Irregularity (mm) Discomfort (Likert scale) Untying and ligating 6 brackets (secs)
NOTES	**Funding source:** not stated

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Method used not mentioned Quote: "The subjects were randomly allocated to one of two groups."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Not mentioned Quote: "The subjects were randomly allocated to one of two groups."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Participants – blind; Clinicians – not possible Quote: "The subjects were not informed which bracket was the newer design."
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: SMs assessed blind; Quote: "…….the operator blinded to the identity of each cast."
Incomplete outcome data (attrition bias)	High - discomfort Low - irreg Unclear risk	Comment: **Discomfort:** 42/60 (70%) returned discomfort questionnaires – 30% missing - HIGH; **Irregularity:** SL porcelain – 2/34 (5.9%) models missing; Convent porc – 4/34 (11.8%) missing - LOW Quote: "Of the 68 patients enrolled in the study, follow-up impressions were missed for two subjects in Group 1 and four subjects in Group 2"

Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment: Because of the 2 extra missed imps for in group 2: “Two subjects, matched for age, gender and incisor irregularity with two subjects in Group 2, were dropped from Group 1 to keep the same number of subjects in each group.

Miles 2012

METHODS	**Setting:** Caloundra and University of Queensland Department of Orthodontics, Australia **Design:** 2 arms **No. of centres: 1** **Study duration:** 10 weeks
PARTICIPANTS	**Inclusion criteria:** children aged between 11-15, a non-extraction treatment plan in the lower arch, no impactions/unerupted teeth, fixed appliances bonded from first molar to first molar in both arches, and living locally to allow for additional appointments for impressions. **Exclusion criteria:** none stated

	Orthodontic intervention: vibrational appliance (Tooth Masseuse) **Patient sampling:** N= 66 Group I (n= 33 for irregularity, 31 for discomfort): fixed app Group II (n = 33, 29 for discomfort): fixed app + vibrational app Ages: The age range for the study was 11.1 to 15.7 years with an average age of 13.1 years for the control group and 13.0 years for the experimental group Control: 14M 19F Intervention: 12M 21F Dropouts for irregularity: Group 1: 2 Group 2: 0 Dropouts for discomfort: Group 1: 1 Group 2: 1
INTERVENTIONS	**Fixed vs fixed with vibrational appliance**
OUTCOMES	Irregularity (mm) Discomfort (VAS scale 100mm)
NOTES	**Funding source:** not stated

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Method used not mentioned Quote: "Patients who met the inclusion criteria were randomly assigned in blocks of six to ensure even numbers in the control and experimental groups"
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Not mentioned Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Participants – not possible; Clinicians – blind Quote: "The clinician was blinded to the study participants at all appointments."
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: SMs assessed blind Quote: "Identification numbers were assigned to the models prior to measurement to ensure blinding. The irregularity index was measured by one of the authors (PM). After data collection was complete, the model numbers were matched back to the corresponding patients."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: **Discomfort:** low. 100% response rate **Irregularity:** low. 94% response rate

		Quote: "Sixty-six patients were enrolled in the study, of whom 64 patients reported for all 4 impression appointments. Pain scores were recorded by 60 patients, with 58 completing all 5 time points"
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote: "Table II shows the mean irregularity indices for both groups at the 4 time points." "Table IV shows the mean VAS score for both groups at the 5 time points following appliance placement."
Other bias	High/Low/ Unclear risk	Comment:

Myrlund 2015

METHODS	**Setting:** Tromsø, Norway; Public Dental Service Competence Centre of Northern Norway (TkNN) and the University student clinic (UTK). **Design:** 2 arms **No. of centres: 1** **Study duration:** 1 year
PARTICIPANTS	**Inclusion criteria:** early mixed dentition with upper central incisors and first molars fully erupted; Angle Class I or

	Class II occlusion with one or more of the following characteristics: deep bite (≥2/3 overlapping of the incisors), increased overjet ≥5 mm, moderate anterior crowding in combination with an overjet of ≥4 mm. **Exclusion criteria:** Children with Angle Class III malocclusion, crossbites, or retroclined upper incisors were not included. **Orthodontic intervention:** Eruption guidance appliance **Patient sampling:** N= 48 Group I (n= 25): EGA Group II (n = 23): control Ages: The mean age of the children in the treatment group was 7.7 years [standard deviation (SD) 0.6] and in the control group 7.7 years (SD 0.5) The treatment group consisted of 13 boys and 12 girls and the control group of 12 boys and 11 girls. Dropouts: 2, one from each group
INTERVENTIONS	**Eruption Guidance Appliance vs Control**
OUTCOMES	Anterior crowding maxilla (mm)

	Anterior crowding mandible (mm) Lower incisor to mandible
NOTES	**Funding source:** “LM-Instruments Oy, Finland, has supplied the study with free LM activators for the patients”

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate Quote: “For the randomization, each subject was given an identification number. The numbers were written on a closed raffle ticket and put in a hat from where 25 subjects were blindly drawn to the experimental group, the remaining 23 subjects comprising the control group. Drawing was performed by an independent person (HK)."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate Quote: “To avoid any allocation bias, all clinical characteristics and personal data of the patients were concealed at this point.”
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Participants – not possible; Clinicians – not mentioned Quote:

Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: SMs and cephs assessed blind Quote: "Before measuring, all study casts were pooled together and labelled by only numbers to hide any identification of group, patient name, or date of the model from the investigator. Similarly, all cephalograms were blinded before tracing by numbering the X-rays randomly."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: low risk Quote: "After 1 year, one boy from the treatment group (refused treatment after 6 months) and one girl from the control group (moved) had dropped out, resulting in 24 and 22 subjects in the treatment and control groups, respectively." Drop outs: 1/25 (4%) in treatment group; 1/23 (4.3%) in control group""
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment:

O'brien 1990

METHODS

Setting: Manchester

Design: 2 parallel arms

	No. of centres: 1 **Study duration:** Up to 37 days
PARTICIPANTS	**Inclusion criteria:** None stated **Exclusion criteria:** None stated **Orthodontic intervention:** Titinol or Nitinol 016 archwire. They were all fitted with identical edgewise brackets. **Patient sampling:** N= 40 Group I (n= 20): 016 Titinol Group II (n = 20): 016 Nitinol Ages and male:female ratio: 12.95 years (sd = 3.2), 11F 9M Titinol 13.4 years (sd = 3.12) Nitinol Dropouts: No dropouts
INTERVENTIONS	**0.016 Titinol archwire vs 0.016 Nitinol archwire**
OUTCOMES	Primary outcome: Little's Irregularity Index of upper canine to canine Secondary outcome: Bending moment/angular deflection characteristics .

NOTES	**Funding source**"Thomas Bolton & Johnson Limited, Stoke-on- Trent, England, for supplying the archwires."

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Method not described Quote: " Forty patients who were attending for routine Edgewise fixed appliance therapy were randomly allocated "
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment Not reported Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Not reported Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment Not reported Quote:
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Data presented for 20 groups of 20 pts; no loss to follow-uo Quote: "Means of twenty subjects with standard deviations in parentheses."

Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on
Other bias	High/Low/ Unclear risk	Comment: No obvious sources of bias. No Sample size calculation

Ong 2011

METHODS	**Setting:** Private orthodontic practice, Caloundra, QLD, Australia **Design:** 3 parallel arms **No. of centres:** 1 **Study duration:** 1 year
PARTICIPANTS	**Inclusion criteria:** All patients who required both upper and lower orthodontic appliances were included. There were no restrictions regarding age, previous orthodontic experience, or extraction/non-extraction treatment. No patients had craniofacial abnormalities. **Exclusion criteria:** Patients were excluded if they had asymmetrically missing or extracted premolars, missing or unerupted lower incisors or canines, or teeth blocked out that did not allow for placement of all brackets at the initial bonding appointment. **Orthodontic intervention:**

	1. 3M Unitek: Orthoform II archforms N 0.014 inch Nitinol N 0.01760.017 inch heat activated Ni–Ti 2. GAC International: Medium Ideal archform N 0.014 inch Sentalloy N 0.01660.022 inch Bioforce 3. Ormco: Damon archform N 0.014 inch Damon Copper Ni–Ti N 0.014x0.025 inch Damon Copper Ni–Ti. **Patient sampling:** N= 132 Group I (n= 44): Unitek Group II (n = 44): Ormco Group III (n=44): GAC Ages and male:female ratio: 14.4 (4.4) Unitek 14:30 15.5 (6.4) GAC 19:25 16.1 (8.3) Ormco 19:25 Dropouts: 1 from GAC group
INTERVENTIONS	**Unitek vs GAC vs Ormco archwire sequences**
OUTCOMES	Primary outcome: time required to reach the working archwire (months.) Secondary outcomes: discomfort experienced at each archwire change

	and total discomfort for each archwire sequence. In addition, Little's irregularity index score after removal of the first and second archwire was determined, as well as the mandibular intercanine width after removal of the first and second archwires.
NOTES	**Funding source:** "This study was supported by a grant from the Australian Society of Orthodontists' Foundation for Research and Education"

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "A restricted randomization process was employed, where patients were randomized in blocks of 12 to ensure equal allocation of patients to the treatment groups. Randomization was performed using computer generated numbers."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "A list of numbers was kept by the laboratory staff, who on the day of bracket placement would then assign them to the

		appropriate group, so the treating clinician was not involved."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Adequate for patients, not possible for clinicians Quote: "The patients were blinded to their group allocation throughout treatment; however, the treating clinician could not be masked."
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Adequate method Quote: "The clinician's staff assigned an identification number to each model prior to measurement in order to mask the principal researcher (EO) to the patient name, treatment allocation, time-point and archwire group during study model analysis. The models were rematched to the patient and archwire group after data collection was complete."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk Low – LTFU Low – irregularity High - pain	Comment: Quote: "**Patients** LTFU: 3M 0/44; Ormco 0/44; GAC 1/44 = 2.3%. **Missing models** 3M 7/44 = 16%; Ormco 2/44 = 4.5%; GAC 3/44 = 6.8%. **Irregularity** LTFU 3M 3/44 = 6.8%; Ormco 2/44 = 4.5%; GAC 1/44 = 2.3%. **Discomfort** "Some discomfort data were also lost because 63 questionnaires were not returned" Analysed: 3M 23/44 = 52.3%;

		Ormco 20/44 = 45.5%; GAC 20/44 = 45.5%
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported. Quote:
Other bias	High/Low/ Unclear risk	Comment:

Pandis 2009

METHODS	**Setting:** Corfu, Greece, private practice of author **Design:** 2 parallel arms **No. of centres:** 1 **Study duration:** 6 months
PARTICIPANTS	**Inclusion criteria:** nonextraction treatment on the mandible, eruption of all mandibular teeth, no spaces in the mandibular arch, no crowding in the posterior segments, mandibular irregularity index greater than 2, and no therapeutic intervention planned involving intermaxillary or other intraoral or extraoral appliances including intra-arch or interarch elastics, lip bumpers, maxillary expansion appliances, or headgears

	Exclusion criteria: none stated **Orthodontic intervention:** 0.016-in CuNiTi 35C **Patient sampling:** N= 60 Group I (n= 30): CuNiTi Group II (n = 30): NiTi Ages: CuNiTi 13.4 (1.8) NiTi 12.8 (1.7) Gender: CuNiTi F: M 70:30 NiTi F:M 83.4:16.6 Dropouts: 0
INTERVENTIONS	**0.016-in CuNiTi vs 0.016-in NiTi wire**
OUTCOMES	Primary outcome: time to alignment of the mandibular anterior dentition, determined as the time from first archwire placement to complete alignment, according to the operator. The observation period ended after 6 months of intervention for all patients; for patients not aligned after 6 months of active treatment, the remaining crowding was recorded. In this case, the irregularity index was measured intraorally, and the mean of the 2 measurements was recorded.

NOTES **Funding source:** none stated

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "Randomization was done using random permuted blocks of size 6"
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: " Opaque envelopes were used to allocate treatment"
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Adequate for patients and clinicians Quote: " Allocation of wires was concealed from the investigator and the participants during the observation period, and no other wire was used throughout the study"
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not mentioned Quote: " Measurements were made intraorally twice by the same clinician using a fine-tip digital caliber (Digimatic NTD12-6-in C, Mitutoyo, Kanagawa, Japan), and the means of the 2 measurements were entered into an Excel spreadsheet (Microsoft, Redmond, Wash)."

Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: no dropouts Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported. Quote:
Other bias	High/Low/ Unclear risk	Comment:

Pandis 2010

METHODS	**Setting:** Corfu, Greece, private practice of author **Design:** 2 parallel arms **No. of centres:** 1 **Study duration:** 175 days
PARTICIPANTS	**Inclusion criteria**: nonextraction treatment in both arches, eruption of all maxillary teeth, no spaces in the maxillary arch, no high canines, maxillary irregularity index greater than 4 mm, and no therapeutic intervention planned involving intermaxillary or other intraoral or extraoral appliances including elastics, maxillary expansion appliances, or headgear**.** **Exclusion criteria:** none stated

	Orthodontic intervention: Active self ligating brackets. Active self-ligating group was bonded with the Roth prescription In-Ovation R bracket (GAC, Central Islip, NY), with a 0.022-in slot. Passive self ligating brackets. Passive self-ligating group received the high-torque version of the Damon MX (Ormco,Glendora, Calif), with a 0.022-in slot. **Patient sampling:** N= 70 Group I (n= 35): Active SL Group II (n = 35): Passive SL **Ages**: 13.8 (1.8) overall **Gender**: Active 57% F Passive 60% F Dropouts: 4, 2 per group
INTERVENTIONS	**Active vs Passive SL brackets**
OUTCOMES	Primary outcome: time to alignment of the mandibular anterior dentition, determined as the time from first archwire placement to complete alignment, according to the operator. The observation period ended after 6 months of intervention for all patients; for patients not aligned after 6 months of active treatment, the remaining

	crowding was recorded. In this case, the irregularity index was measured intraorally, and the mean of the 2 measurements was recorded.
NOTES	**Funding source:** none stated

Risk of Bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Randomisation carried out Quote: "Randomization was accomplished by generating random permuted blocks of variable size; this ensured equal patient distribution between the 2 trial arms."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: Numbered, opaque, sealed envelopes were prepared before the trial containing the treatment allocation card. After patient selection, the secretary of the practice was responsible for opening the next envelope in sequence."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians, not mentioned for patients Quote:

Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not mentioned Quote: Measurements were made twice on the initial casts by the first author with a digital caliper (Digimatic NTD12-6''C, Mitutoyo, Tokyo, Japan)."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: 4 dropouts, 2 per group; 4/70= 6% Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment: single centre, private practice; per protocol analysis

Pandis 2011

METHODS	**Setting:** Corfu, Greece, private practice of author **Design:** 2 parallel arms **No. of centres:** 1 **Study duration:** 16 months
PARTICIPANTS	**Inclusion criteria**: nonextraction treatment on both arches, eruption of all

mandibular teeth, no spaces in the mandibular arch, mandibular irregularity index greater than 2 mm (canine to canine), and no therapeutic intervention planned involving intermaxillary or other intraoral or extraoral appliances including elastics,

maxillary expansion appliances, or headgears before the end of the observation period.

Exclusion criteria: none stated

Orthodontic intervention:
conventional brackets: Roth prescription microarch bracket

Self-ligating group received the DamonMX

both with a 0.022-in slot size

Patient sampling:

N= 50

Group I (n= 25): conventional

Group II (n = 25): self ligating

Ages: 13.4 (1.6) conventional

13.2 (1.6) self ligating

Gender: 64F 36M conventional

68F 32M self ligating

No dropouts

INTERVENTIONS	**Conventional vs self ligating brackets**
OUTCOMES	Main outcome was intermolar width at the time of the passive placement of the 0.01630.025-in stainless steel wire on the mandibular arch. Secondary outcomes included intercanine width and days to reach the time to place the final 0.016x0.025-in stainless steel wire. Did not look at crowding. No values for end results. Time to alignment?
NOTES	**Funding source:** none stated

Risk of Bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology used Quote: "Fifty patients were randomized to either a conventional or a self-ligating appliance. The statistical software package was used by the first author, and the user-written ralloc command was implemented to generate the random allocation sequence. Randomization was accomplished by generating random permuted blocks of variable size (2 and 4), which assured

		equal patient distribution between the 2 trial arms.”
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: “Sequentially numbered, opaque, sealed envelopes were prepared before the trial containing the treatment allocation cards. After patient selection and recording of baseline information, the secretary in the practice was responsible for opening the next envelope in sequence and allocation.”
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians or patients Quote: “Blinding of either the patient or the orthodontist during delivery of the interventions and assessment of whether the 0.01630.025-in wire could be placed was not possible”
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	“Impressions were taken, plaster models were prepared, and the brackets from second premolar to second premolar were scraped off with a curving knife by the laboratory technician to facilitate blind measurements. Each model had an identification number so that the measurements could be correctly entered into the spreadsheet”
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: no dropouts Quote:

Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment: single centre, private practice.

Pirttiniemi 2005

METHODS	**Setting:** Not specified **Design:** 2 parallel arms **No. of centres:** 3 **Study duration:** 16 months of CPHG then follow ups of 2 yrs and 8 yrs
PARTICIPANTS	**Inclusion criteria**: a need for orthodontic treatment due to moderate crowding and a Class II tendency. The crowding was clinically diagnosed as moderate, based on the degree of space deficiency in the anterior regions of the dental arches **Exclusion criteria:** None stated **Orthodontic intervention**: In the HG group, the maxillary first molars were banded and cervical HG was used, but no other appliances were used. The long outer bows of the HG were bent 10 degrees upwards in relation to the inner

bow. The inner bow of the HG was expanded and was constantly held 10 mm wider than the dental arch. The patients were instructed to wear the HG during sleep, for 8–10 hours. The interceptive procedures in the control group were extraction of the upper primary canines in 38 per cent and lower primary canines in 35 per cent, to ease the eruption of the lateral incisors. In addition, in 19 per cent of the subjects in the control group, interdental stripping was carried out.

Between T1 and T2, Orthodontic treatment, if needed, during this phase comprised fixed appliance treatment, including extractions of permanent premolars due to crowding.

Patient sampling:

N= 68 children of both sexes (40 males and 28 females) aged 7.6 years [standard deviation (SD) 0.3 years].

Group I (n= 32): CPHG

Group II (n = 32): control

Dropouts:

Intervention: 7

3 moved

3 refused

	1 missing permanent teeth Control: 3 1 moved 2 refused
INTERVENTIONS	**Low pull Head gear vs Control**
OUTCOMES	SM: Fifteen linear dimensions were measured on the dental casts Ceph: Five angular measurements, describing skeletal changes, were chosen to represent skeletal variables: SNA, SNB, ANB, NL/ML (the angle between the line intersecting anterior and posterior nasal spine and the line from the inferior surface of the symphysis to the antegonial notch), and SN/NL (the angle between the line from nasion to sella and the line intersecting anterior and posterior nasal spine).
NOTES	**Funding source:** none stated

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology used Quote: The children were randomly divided into two groups of equal size, matched

		according to gender. This was undertaken by one author (TK) using random numbers.
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Some attempt made Quote: To conceal the allocation, most of the practitioners who undertook the treatment were not given information concerning the aim or rationale of the study
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians or patients Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not reported
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Dropouts described. 26.5% in intervention group, 9% control group
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment:

Rebellato 1997

METHODS	**Setting:** Not specified **Design:** 2 parallel arms **No. of centres:** 1 **Study duration:** 10.5 months for lingual arch, 12.5 months for control
PARTICIPANTS	**Inclusion criteria**: (1) Both mandibular second deciduous molars were present with some clinical mobility, (2) mandibular crowding was 3 mm or more, (3) permanent molar relationships were end-on to Class I (end-on molars would have flush mesial planes and Class I mandibular molars were up to 4 mm mesial of flush mesial plane15), (4) overbite was 1 mm or greater, (5) mandibular plane inclination was average (MP-SN) of 32 ° + 6 °, and (6) the lower lip was less than 4 mm ahead of Rickett's E line. **Exclusion criteria:** Patients were excluded from the study if they had any congenitally or prematurely missing teeth. Only European American patients were selected, because ethnic differences in mean skeletal patterns 17,1s and mean differences in arch length and tooth sizes between European Americans and African Americans a9 have been reported

	Orthodontic intervention: The lingual arch appliance used in the treatment group was a passive 0.032-inch stainless steel wire, which contacted the cingulae of the lower incisors. **Patient sampling:** N= 30 Group I (n= 14): lingual arch Group II (n = 16): control **Ages:** 11.5 lingual arch (no mean) 11.3 control (no mean) Gender: unknown Dropouts unclear
INTERVENTIONS	**Lingual arch vs control**
OUTCOMES	Ceph: changes in centre of rotation, rotation of teeth SM: Changes in intermolar width, arch depth, and total arch length
NOTES	**Funding source: This research was supported by USPHS Research Grant 1 R03 DE10002-1 from the National Institute of Dental Research, National Institutes of Health, Bethesda, Md.**

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Methodology not described Quote: "Subjects were randomly assigned to two groups"
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Not reported Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians or patients Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Not reported
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: no dropouts reported Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment: no sample size calculation, no mention of source of patients or their genders, blinding of assessor, allocation concealment

Sebastian 2012

METHODS	**Setting:** Noorul Islam College of Dental Sciences, Trivandrum, Kerala, India **Design:** 2 parallel arms **No. of centres:** 1 **Study duration:** 12 weeks
PARTICIPANTS	**Inclusion criteria**: Female patients in postmenarche period between 13 and 15 years of age with crowding in the lower anterior segment and having a mandibular irregularity index greater than 6 Class I skeletal pattern Nonextraction treatment in mandibular arch Eruption of all mandibular teeth with no spacing between them No relevant medical history No recent history of intake of drugs such as nonsteroidal anti-inflammatory drugs (NSAIDs) Patients who may have experienced periodontal disease and hence loss of attachment was avoided

No previous active orthodontic treatment

Full arch mechanics, preadjusted edgewise appliance therapy

No therapeutic intervention planned involving intermaxillary or other intraoral or extraoral appliances during the study period

Exclusion criteria: Participants were told to avoid intake of drugs during the study period. If they had taken any drugs because of unavoidable circumstances, they were requested to report the matter. The intention was to exclude them from the study in such instances

Orthodontic intervention: 0.016-inch coaxial superelastic NiTi and 0.016-inch single-stranded superelastic NiTi

Patient sampling:

N= 24

Group I (n= 12): single stranded NiTi

Group II (n = 12): Coaxial NiTi

Ages: 13.8 (0.7) (single strand)

13.6 (0.6) (coaxial)

Gender: all female

No dropouts

INTERVENTIONS	**Coaxial NiTi vs single stranded NiTi**
OUTCOMES	SM: Data for intertooth distances (3-2, 2-1, 1-1, 1-2, 2-3) were collated, and changes at all intertooth distances were summated to represent overall tooth movement, thus deriving a figure for the average alignment of the lower anterior segment at each stage.
NOTES	**Funding source:** none stated

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology used Quote: "Randomization was done using computer software–generated numbers."
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "Opaque envelopes were used to allocate the archwires to two groups, each consisting of 12 participants. Allocation thus was concealed from the investigator and from participants during the study."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians or patients Quote:

Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Comment: "All readings were measured by an expert single operator, who was not aware of the archwire specimen used for the arches being measured."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: no dropouts Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: All expected outcomes reported on Quote:
Other bias	High/Low/ Unclear risk	Comment: All females

Tai 2010

METHODS	**Setting:** private orthodontic office Osaka Japan? **Design:** 2 parallel arms **No. of centres:** 1? **Study duration:** Up to 21 months
PARTICIPANTS	**Inclusion criteria**: Angle Class I

	malocclusions with crowding and normal vertical dimensions and no posterior crossbites. **Exclusion criteria:** No posterior crossbites **Orthodontic intervention**: Mandibular Swartz appliance **Patient sampling:** N= 28 Group I (n= 14): Swartz Group II (n = 14): Control **Ages:** 7 y 11m@ T0 9y 1m @ T1 Swartz 8y @ T0 9y 8m @ T1 Control Gender: (6 boys, 8 girls) Swartz (6 boys, 8 girls) Control No dropouts? Japanese?
INTERVENTIONS	**Swartz appliance vs control**
OUTCOMES	No obvious outcomes described. Ceph: Sixteen points were digitized on each cephalometric radiograph, and 12 cephalometric measurements were made.

	SM: arch crowding, arch perimeter, and arch length (digital calipers). CBCT: A slice plane perpendicular to the occlusal plane, passing through both sides of the mesiobuccal cusp tips of the mandibular first molars. Ten points of interest were measured including mandibular firstmolar crowns, cementoenamel junctions (CEJs), roots, buccal and lingual alveolar processes, inner and outer surfaces of the mandibular bodies, zygomatic bones, condylar heads, and antegonial notches
NOTES	**Funding source: The authors report no commercial, propriety,or financial interest in the products or companies described in this article.** However, “We thank Toyohisa Tanijiri (Medic Engineering, Kyoto, Japan) for developing the software"

Risk of Bias

Bias	Authors' judgement	Support for judgement

Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Not reported Quote: "After initial recording of the data, the patients were randomized to 2 groups"
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: not reported Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians or patients Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Comment: "to prevent bias in the measurement of the expanded and nonexpanded groups, the investigator was blinded."
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: not reported Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: No outcomes specified; some appropriate outcomes reported Quote:
Other bias	High/Low/ Unclear risk	Comment: sample size based on pilot study. Ethnicity?

Woodhouse 2011

METHODS	**Setting:** King's College London Dental Institute (Guy's Hospital); the Royal Alexander Children's Hospital, Brighton, Sussex; and William Harvey Hospital, Ashford, Kent **Design:** 3 parallel arms **No. of centres:** 1 **Study duration:** 209 ± 65 days
PARTICIPANTS	**Inclusion criteria** (1) <20 y old at start of treatment, (2) no medical contraindications, (3) in the permanent dentition, (4) mandibular arch incisor irregularity, and (5) extraction of mandibular first premolars included in the orthodontic treatment plan. **Exclusion criteria:** No posterior crossbites **Orthodontic intervention:** Accel appliance **Patient sampling:** N= 81 Group I (n= 29): Accel Group II (n = 25): Accel sham Group III (n=27): Fixed only Ages: 13.9 ± 1.6 y Accel

	14.1 ± 1.9 y Accel sham 14.4 ± 1.8 y Fixed only Gender: 40M 41F 4 dropouts: 1 Accel 2 Sham Accel 1 Fixed only
INTERVENTIONS	**Accel appliance vs Accel sham vs Fixed alone**
OUTCOMES	Primary outcome: initial rate of tooth alignment in the mandibular arch. Secondary outcome: time to achieve complete alignment
NOTES	**Funding source:** none stated

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "The randomization sequence was generated using GraphPad online software (http://www.graphpad.com/quickcalcs/index.cfm)"
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Quote: "participant allocation undertaken centrally at King's College London, independently from the clinical operators,

		following recruitment (allocation concealment; Schulz and Grimes 2002)."
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment: Not possible for clinicians, possible for participants but not directly expressed Quote: "The sham device was identical to the active in all respects, except that it did not vibrate."
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Adequate methodology Comment: "Dental casts were coded so that measurements were undertaken blind"
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Adequate reporting Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: No outcomes described, results reported No outcomes specified; some appropriate outcomes reported?
Other bias	High/Low/ Unclear risk	Comment: sample size based on pilot study. Ethnicity?

APPENDIX 2: Characteristics of excluded studies

STUDY	REASON FOR EXCLUSION
Abu Alhaija 2011	Not RCT
Almeida 2015	Patients ages fall outside of eligibility criteria
Altug 2005	Not on crowding
Barlin 2011	Not RCT
Thickett 2009	Not on crowding
Baumrind 1996	Not RCT
Bondemark 2005	Not on crowding
Clements 2003	Patients ages fall outside of eligibility criteria
Cobb 1998	Patients ages fall outside of eligibility criteria
Dai 2009	Not RCT ELIGIBILITY
DiBiase 2011	Patients ages fall outside of eligibility criteria

Fleming 2009	Patients ages fall outside of eligibility criteria
Freitas 2013	Not RCT
Fan 2009	Not RCT, patients ages fall outside of eligibility criteria ELIGIBILITY
Harradine 1998	Patients ages fall outside of eligibility criteria
Heiser 2004	Not RCT
Kau 2013	Patients ages fall outside of eligibility criteria
Keski-Nisula 2008	Not RCT
Lindqvist 1982	Patients ages fall outside of eligibility criteria
Markovic 2015	Not on crowding
Miyake 2008	Not RCT
Nagalakshmi 2014	Not RCT
Okay 2006	Not on crowding
Ong 2001	Not RCT
Owais 2011	Not RCT
Pandis 2007	Not RCT
Pandis 2010a	Not RCT
Rowland 2007	Not on crowding
Ruf 1999	Not RCT
Sandhu 2013	Not on crowding

Scott 2008	Patients ages fall outside of eligibility criteria
Serfim 2015	Patients ages fall outside of eligibility criteria
Shawesh 2010	Not on crowding
Silva 2012	Not on crowding
Soldanova 2012	Patients ages fall outside of eligibility criteria
Strahm 2009	Not on crowding
Sucuru 1992	Not RCT
Talapaneni 2012	Patients ages fall outside of eligibility criteria
Taner 2000	Patients ages fall outside of eligibility criteria
Vajaria 2011	Not RCT
Wahab 2012	Patients ages fall outside of eligibility criteria
Wang 2010	Patients ages fall outside of eligibility criteria
Wasserman 2010	Unable to locate paper
Wortham 2009	Not on crowding
West 1995	Insufficient information to allow inclusion of data
Yu 2008	Not RCT ELIGIBILITY

APPENDIX 3: Search strategies

1 .Cochrane Oral Health's Trials Register search strategy

#1 (("class i" and (angle* or malocclusion or bite))) AND (INREGISTER)

#2 (("class ii" and (angle* or malocclusion or bite))) AND (INREGISTER)

#3 (("class iii" and (angle* or malocclusion or bite))) AND (INREGISTER)

#4 ((crowd* AND teeth)) AND (INREGISTER)

#5 (#1 or #2 or #3 or #4) AND (INREGISTER)

#6 (orthodontic*) AND (INREGISTER)

#7 (appliance*) AND (INREGISTER)

#8 (("lip bumper*" OR lip-bumper*)) AND (INREGISTER)

#9 (("arch develop*" AND (jaw* OR mandib* OR maxill*))) AND (INREGISTER)

#10 (((expansion OR expand) AND (jaw* OR maxill*))) AND (INREGISTER)

#11 ("leeway space*") AND (INREGISTER)

#12 (("two-phase treatment*" OR "two-phase therap*" or "space maintain*" OR "space maintenance")) AND (INREGISTER)

#13 (#6 or #7 or #8 or #9 or #10 or #11 or #12) AND (INREGISTER)

#14 (#5 and #13) AND (INREGISTER)

2. Cochrane Central Register of Controlled Clinical Trials (CENTRAL) search strategy

#1 MALOCCLUSION ANGLE CLASS I Single term (MeSH)

#2 MALOCCLUSION ANGLE CLASS II Single term (MeSH)

#3 MALOCCLUSION ANGLE CLASS III Single term (MeSH)

#4 ((class next i) and ((angle or angle*) or malocclusion or bite))

#5 ((class next ii) and ((angle or angle*) or malocclusion or bite))

#6 ((class next iii) and ((angle or angle*) or malocclusion or bite))

#7 (crowd* near teeth)

#8 (#1 or #2 or #3 or #4 or #5 or #6 or #7)

#9 ORTHODONTIC APPLIANCES, FUNCTIONAL Explode all trees (MeSH)

#10 ORTHODONTIC APPLIANCES, REMOVABLE Explode all trees (MeSH)

#11 ORTHODONTICS PREVENTIVE, Explode all trees (MeSH)

#12 ORTHODONTICS INTERCEPTIVE, Explode all trees (MeSH)

#13 TOOTH EXTRACTION Explode all trees (MeSH)
#14 (leeway next space*)
#15 ((extraoral or (extra next oral) or extra-oral) and appliance*)
#16 (lip next bumper*)
#17 (((two next phase next treatment) or (two next phase next therapy)) and (orthodontic* or malocclusion))
#18 ((arch next development) and (jaw* or mandible or maxilla*))
#19 ((extraction* and (dental or teeth or tooth)) and orthodontic*)
#20 (expansion and (jaw or maxilla*))
#21 ((serial next extract*) and (teeth or orthodontic*))
#22 ((space next maintenance) and orthodontic*)
#23 (#9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22)
#24 (#8 and #23)

3. MEDLINE Ovid search strategy

1. exp MALOCCLUSION, ANGLE CLASS I/
2. exp MALOCCLUSION, ANGLE CLASS II/
3. exp MALOCCLUSION, ANGLE CLASS III/
4. (("Class 1" or "Class I") and (Angle or Angle's or Angles or malocclusion$ or bite$)).mp.
5. (("Class 2" or "Class II") and (Angle or Angle's or Angles or malocclusion$ or bite$)).mp.
6. (("Class 3" or "Class III") and (Angle or Angle's or Angles or malocclusion$ or bite$)).mp.
7. ((crowd$ or overcrowd$) adj6 (teeth or dentition)).mp.
8. or/1-7
9. exp Orthodontic Appliances, Functional/
10. exp Orthodontic Appliances, Removable/
11. exp Orthodontics, Preventive/
12. exp Orthodontics, Interceptive/
13. exp Tooth Extraction/
14. ((extraoral or "extra oral" or extra-oral) and appliance$).mp

15. ("Lip bumper$" or lip-bumper$).mp.
16. ("arch develop$" and (jaw or mandib$ or maxill$)).mp.
17. (extract$ and (dental or teeth or tooth) and orthodontic$).mp.
18. ((expansion or expand) and (jaw$ or maxill$)).mp.
19. ("serial extract$" and (teeth or orthodontic$)).mp.
20. "leeway space$".mp.
21. ("two-phase" adj3 treat$).mp.
22. ("two-phase" adj3 therap$).mp.
23. (21 or 22) and (orthodontic$ or malocclusion$).mp.
24. ((space adj maintenance) and orthodontic$).mp.
25. ((space adj3 maintain$) and orthodontic$).mp.
26. 24 or 25
27. (orthodontic$ and (functional or removable) and appliance$).mp.
28. ((interceptive or preventive) and orthodontic$).mp.
29. ((activator adj4 appliance$) and orthodontic$).mp.
30. 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 23 or 26 or 27 or 28 or 29
31. 8 and 30

This subject search was linked to the Cochrane Highly Sensitive Search Strategy (CHSSS) for identifying randomised trials in MEDLINE: sensitivity- maximising version (2008 revision) as referenced in Chapter 6.4.11.1 and detailed in box 6.4.c of *The Cochrane Handbook for Systematic Reviews of Interventions,* Version 5.1.0 [updated March 2011](Lefebvre 2011).

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. or/1-8

10. exp animals/ not humans.sh.
11. 9 not 10

4. Embase Ovid search strategy

1. MALOCCLUSION/
2. (("Class 1" or "class I") and (Angle or Angle's or Angles or malocclusion$ or bite$)).mp.
3. (("Class 2" or "Class II") and (Angle or Angles or Angle's or malocclusion$ or bite$)).mp.
4. (("Class 3" or "Class III") and (Angle or Angle's or Angles or Malocclusion$ or bite$)).mp.
5. ((crowd$ or overcrowd$) adj6 teeth).mp.
6. or/1-5
7. exp Orthodontic Device/
8. orthodontics.mp. or exp ORTHODONTICS/
9. 8 and ((removable or functional) and appliance$).mp.
10. 8 and (preventive or interceptive).mp.
11. Tooth Extraction/
12. "leeway space$".mp.
13. ((extraoral or "extra oral" or extra-oral) and appliance$).mp.
14. ("lip bumper$" or lip-bumper$).mp.
15. ("arch develop$" and (jaw$ or mandib$ or maxilla$)).mp.
16. (extract$ and (dental or teeth or tooth) and orthodontic$).mp.
17. ((expansion or expand$) and (jaw$ or maxilla$)).mp.
18. ("serial extraction$" and (teeth or orthodontic$)).mp.
19. ("two-phase" adj3 treat$).mp.
20. ("two-phase" adj3 therap$).mp.
21. (19 or 20) and (orthodontic$ or malocclusion$).mp.
22. ("space maintenance" and orthodontic$).mp.
23. ((space adj3 Maintain$) and orthodontic$).mp.
24. 7 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 21 or 22 or 23
25. 6 and 24

This subject search was linked to an adapted version of the Cochrane Embase Project filter for identifying RCTs in Embase Ovid (see http://www.cochranelibrary.com/help/central-creation-details.html for information:)

1. Randomized controlled trial/
2. Controlled clinical study/
3. Random$.ti,ab.
4. randomization/
5. intermethod comparison/
6. placebo.ti,ab.
7. (compare or compared or comparison).ti.
8. ((evaluated or evaluate or evaluating or assessed or assess) and (compare or compared or comparing or comparison)).ab.
9. (open adj label).ti,ab.
10. ((double or single or doubly or singly) adj (blind or blinded or blindly)).ti,ab.
11. double blind procedure/
12. parallel group$1.ti,ab.
13. (crossover or cross over).ti,ab.
14. ((assign$ or match or matched or allocation) adj5 (alternate or group$1 or intervention$1 or patient$1 or subject$1 or participant$1)).ti,ab.
15. (assigned or allocated).ti,ab.
16. (controlled adj7 (study or design or trial)).ti,ab.
17. (volunteer or volunteers).ti,ab.
18. trial.ti.
19. or/1-18
20. (exp animal/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans).ti.)
21. 19 not 20

5 US National Institutes of Health Ongoing Trials Register (ClinicalTrials.gov) search strategy

crowded and teeth

6 World Health Organization International Clinical Trials Registry Platform search strategy

crowded and teeth

APPENDIX 4: Title and abstract screening form

Author	Title	Year	NOT an RCT?			NOT a review with relevant references?			NOT primarily to do with crowding?			Do NOT have >80% participants <16?			Notes	Include?		
			Yes	No	?	Yes	no	?	Yes	No	?	Yes	No	?		Yes	No	?
1																		
2																		
3																		
4																		
5																		
6																		

APPENDIX 5: Eligibility form

Study name	(first author and date)			
		Yes	Unclear	No
Type of Study	Is the study a randomised controlled clinical trial?			
		Go to next question		Exclude
Participants in the study	Are 80% or more of the participants aged under 16 years old?			
		Go to next question		*Exclude*
	Were the participants receiving orthodontic treatment to correct or prevent crowding?			
		Go to next question		*Exclude*
	Were the participants without a cleft lip and/or palate or cranio-facial anomaly?			
		Go to next question		Exclude
Types of intervention	Was one or more groups treated with an orthodontic appliance (removable, fixed or functional), headgear or extractions?			
		Go to next question		Exclude
	Did the other group(s) receive the same care together with no treatment, delayed treatment or treatment with another active intervention?			
		Go to next question		*Exclude*

Types of outcomes	Did the study report the degree of crowding or incisor relationship to the mandible or maxilla at the end of treatment?			
		Include *(subject to clarification of 'unclear' points)*		*Exclude*
		Include	***Unclear***	***Exclude***
Reviewer	(initials)			

APPENDIX 6: Data extraction form

Time start extraction_____ Time finish extraction_____Time taken__

_____

Paper Number _____ Extractor Initials _____

*First Author _____ *Year of publication _____

**Combine these to give a unique name to the paper*

Number of trials included in this paper _____

If more than one, complete separate extraction forms for each and add letters A, B, C etc to the paper name

Paper number(s) that this report links with _____

If other papers report further results of this trial, incorporate them onto this form and note what has been done here e.g. time points, outcomes.

_____

_____

_____

Details of trial

Location of trial centre(s) _____

Source of participants _____

Funding source _____

Method of recruitment _____

Dates for recruitment _____

Study duration _____

Maximum duration of follow-up _____

Time points at which follow-up are reported _____

Notes

Conflict of interests/funding:

Adverse events/Harm:

Quality assessment

Controlled clinical trial

***NB**. Both criteria need to be fulfilled to be a CCT*

Yes / No / Unclear

A study where the individuals (or other units) followed in the trial were definitely or possibly assigned prospectively to one of two (or more) alternative forms of healthcare

Trials employing treatment allocation methods such as coin flips, odd/even numbers, patient social security numbers, days of the week, medical record numbers, or other such pseudo- or quasi-random processes are designated as controlled clinical trials.

Randomised controlled clinical trial

***NB**. All criteria need to be fulfilled to be an RCT*

Yes / No / Unclear

Human; prospective; 2 or more interventions; random allocation.

A study that uses the play of chance to assign participants to test or control treatments in a trial, e.g. by using a random numbers table or a computer-generated random sequence.

NOT randomly selected; allocation for clinical reasons; participants selected own intervention; intervention & control groups different e.g. sick vs. healthy, practice vs. hospital; matched unless matched **prior** to randomisation.

Method of allocation concealment

Adequate (A) / Unclear (B) / Inadequate (C) / Not used (D)

Adequate concealment schemes.

Centralised (e.g. allocation by a central office unaware of subject characteristics) or pharmacy-controlled randomisation

Pre-numbered or coded identical containers which are administered serially to participants

On-site computer system combined with allocations kept in a locked unreadable computer file that can be accessed only after the characteristics of an enrolled participant have been entered

Sequentially numbered, sealed, opaque envelopes

Inadequate concealment schemes:

Alternation;

The use of case record numbers, dates of birth or day of the week,

Any procedure that is entirely transparent before allocation, e.g. open list of random numbers.

Unclear concealment schemes

Studies that do not report any specific concealment approach, e.g. merely stating that a randomisation list or table was used, only specifying that sealed envelopes were used and reporting an apparently adequate concealment scheme in combination with other information that leads the reviewer to be suspicious.

Reporting of withdrawals	***Yes / No / Unclear / No dropouts***
Double-blind	***Yes / No / Unclear / Not applicable***
Outcomes assessed blind	***Yes / No / Unclear / Not applicable***
Intention to treat analysis.	***Yes / No / Unclear***

Risk of bias assessment

Please **CIRCLE / HIGHLIGHT** response as appropriate

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High/Low/ Unclear risk	Comment: Quote:
Allocation concealment (selection bias)	High/Low/ Unclear risk	Comment Quote:
Blinding of participants and personnel (performance bias)	High/Low/ Unclear risk	Comment Quote:
Blinding of outcome assessment (detection bias)	High/Low/ Unclear risk	Comment: Quote:
Incomplete outcome data (attrition bias)	High/Low/ Unclear risk	Comment: Quote:
Selective reporting (reporting bias)	High/Low/ Unclear risk	Comment: Quote:
Other bias	High/Low/ Unclear risk	Comment: Quote:

Comparisons evaluated

Removable appliance Y / N Type _____

Fixed appliance Y / N Type_____

Functional	Y / N	Type _____
Headgear	Y / N	Type _____
Extractions	Y / N	Type _____
No or delayed Rx	Y / N	Type _____

Details of the interventions

	Group 1	Group 2	Group 3	Overall
Group Name (for trial ID)				
Group randomised	**Yes / No**	**Yes / No**	**Yes / No**	**Yes / No**
Intervention (Removable / Fixed / Functional / Headgear / Extractions / No or delayed treatment)				
Schedules				
Number recruited				

Characteristics of participants

	Group 1	Group 2	Group 3	Group 4	Overall / Total
Age					
Sex					
Ethnicity					

Outcomes

Primary: Amount of crowding (measured in mm or by any index of malocclusion).

Secondary: Relationship of the lower back teeth (molars) to the lower jaw (mandible)

Relationship of the lower front teeth (incisors) to the lower jaw (mandible)

Relationship of the top back teeth (molars) to the upper jaw (maxilla)

Relationship of the top front teeth (incisors) to the upper jaw (maxilla)

Self-esteem

Patient satisfaction

Jaw joint problems.

Harms: Health of the gums; damage to the teeth e.g. tooth decay.

Results

Amount of crowding (measure in mm or by any index of malocclusion)

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Relationship of the lower back teeth (molars) to the lower jaw (mandible)

Outcome measure _____

	Results at each time point			
Time point				
Group 1				

Group 2				
Group 3				
Group 4				

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Relationship of the lower front teeth (incisors) to the lower jaw (mandible)

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Relationship of the top back teeth (molars) to the upper jaw (maxilla)

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Relationship of the top front teeth (incisors) to the upper jaw (maxilla)

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Self-esteem

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				

Group 3				
Group 4				

Patient satisfaction

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Jaw joint problems

Outcome measure _____

	Results at each time point			
Time point				
Group 1				
Group 2				
Group 3				
Group 4				

Harms (Health of the gums; damage to the teeth e.g. decalcification, caries, gingivitis etc.).

Outcome measure(s) _____

Group 1	
Group 2	
Group 3	
Group 4	
Group 4	

Comments

_____

_____