Journal of Oral Rehabilitation

Journal of Oral Rehabilitation 2017

Review
Skeletal and dental effects of Class III orthopaedic treatment: a systematic review and meta-analysis

R. RONGO*, V. D'ANTÒ*[†], R. BUCCI*, I. POLITO*, R. MARTINA* & A. MICHELOTTI* *Department of Neurosciences, Reproductive Sciences and Oral Sciences, Division of Orthodontics, University of Naples 'Federico II', Naples, Italy and [†]Division of Dentistry, Department of Pediatric Surgery, Bambino Gesù Children's Hospital, Rome, Italy

SUMMARY To summarise the skeletal, dental and soft tissue effects of orthopaedic treatment on growing skeletal class III patients compared with a concurrent untreated similar control group and to evaluate whether the design of the primary studies may affect the results. A literature search was performed up to the end of February 2016. No restrictions were applied concerning language and appliances. Once the quality score was assessed, a meta-analysis was performed for the appliances used in more than three studies. A moderator analysis for study design was performed. The level of evidence was evaluated by means of the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) tool. The search resulted in 21 papers. The quality of most of the studies was medium. Each study reported skeletal sagittal improvement and overjet correction. Fourteen studies reported a significant increase in lower facial height. Follow-up data showed slight relapses in about 15% of patients. Meta-analyses were performed for the facemask and chin cup. The two appliances were efficient for correcting the sagittal discrepancy, increasing the divergence. In the analysis for study design, the retrospective studies showed a more efficient appliance than RCTs for 6 of 13 variables. The level of evidence was between very low and moderate. There is very low to low evidence that orthopaedic treatment is effective in the correction of Class III skeletal discrepancies and moderate evidence for the correction of the overjet. A common side effect is mandibular clockwise rotation in older subjects.

KEYWORDS: malocclusion, angle class III, orthodontic appliances, meta-analysis, evidence-based dentistry, child, growth and development

Accepted for publication 15 February 2017

Background

Skeletal class III malocclusion is one of the most investigated topics in orthodontics (1–9). The skeletal and dental components of Class III malocclusions are usually established since early childhood (1, 2) and may worsen with growth (3–5). In the majority of cases, without any treatment during childhood, orthognathic surgery is the only option to establish a correct occlusion (6, 7). However, early interception of this malocclusion may represent an opportunity to avoid or reduce the risk of surgery (8, 9).

Several orthopaedic appliances are used in the early treatment of this malocclusion, such as the Frankel III (FR-III) (10), chin cup (CC) (11), mandibular headgear (MHG) (12), reverse headgear (RPHG) (13) or facemask (FM) (14), and bone-anchored maxillary protraction (BAMP) (15).

Actually, there is still a lack of evidence concerning the effectiveness of orthopaedic Class III treatments, particularly regarding the changes in sagittal and vertical dimension and long-term stability.

Six systematic reviews and meta-analysis have recently been published on skeletal Class III malocclusion treatment (16–22), only focusing on RCTs or a single appliance.

doi: 10.1111/joor.12495

The aim of the present systematic review was to summarise the skeletal, dento-alveolar and soft tissue effects of orthopaedic treatment on growing patients with skeletal Class III malocclusion by answering the following question according to the PICO schema (23): 'Do growing skeletal Class III patients (P) treated with an orthopaedic appliance (I) show improvement in skeletal, dental or soft tissues outcomes in the short and long term (O), as opposed to an untreated concurrent control group of growing skeletal Class III patients (C)?' The meta-analysis aimed to assess the efficacy of the treatment and verify if the design of the primary studies affects the reported results.

Methods

Search strategy and study collection

This systematic review is based on the PRISMA guidelines for systematic reviews and meta-analyses (24). A literature survey was performed up to the end of February 2016 using the following search engines: PubMed, Literature in the Health Sciences in Latin America and the Caribbean (LILACS), Scientific Electronic Library Online (SciELO), Cochrane Central Register of Controlled Trials, Scopus, Web of Knowledge, Cambridge Scientific Abstracts, UMI ProQuest metaRegister of Controlled Trials. No restrictions were applied concerning language and appliances. Article abstracts were reviewed to select papers in which an orthopaedic treatment device was used. To minimise the risk of omitting any relevant literature, two authors (IP, RR) independently performed the first step of the screening procedure. The reference lists of the selected articles were hand searched for possible missing articles. Furthermore, the same authors carried out the hand search of American Journal of Orthodontics and Dentofacial Orthopedics, The Angle Orthodontist, Orthodontic and Craniofacial Research and European Journal of Orthodontics.

The research focused on orthopaedic therapy for growing Class III patients and each database had a specific research strategy (Table S1).

Selection criteria

The study included randomised clinical trials (RCT), prospective non-randomised clinical trials (CCT) and retrospective non-randomised clinical trials (Ret), with or without follow-up. Other studies, such as systematic reviews, reviews, case reports, case series, opinion articles or letters to the editor, were excluded (Table S2).

Two reviewers (IP, RR) independently assessed the studies. Disagreements on the selection were resolved through discussion and if necessary consulting a third reviewer (VD).

Quality scores of the included studies

Two reviewers (AM, RM) independently evaluated the quality scores using a modified Downs and Black checklist (25). In this checklist, there are five main domains: reporting, external validity, internal validity-bias, internal validity-confounding and power. The maximum score is 27. A study was judged of low quality if the score was lower than 16, medium quality from 17 to 20, medium-high quality from 21 to 23, and of high quality from 24 to 27. The concordance level was assessed by means of Cohen's k. Moreover, the Cochrane risk of bias tool (26) was used to assess the RCTs. Disagreements were solved by discussion or after consulting the third reviewer (VD).

Data extraction

Two examiners (RB, VD) extracted the data independently using a customised form. The following data were extracted: author and year of publication, study design, ethnic group, sample size, treatment, full observational period, class III diagnosis, inclusion criteria, treatment time, success description, success rate, main treatment effects (skeletal, dental, soft tissue), side effects, follow-up, cephalometric values (SNA, SNB, ANB, maxillary length, mandibular length, mandibular divergence, upper incisors inclination, lower incisors inclination, overjet, overbite, upper lip position, lower lip position and profile variation).

Statistical analysis

Two different meta-analyses were performed: one including studies on FM/RPHG/MBPA (maxillary bow protraction appliance) and the other including studies on CC. All the data were annualised, and the random-effects model was chosen as the observed effect was expected to differ across studies due to sample differences. Statistical heterogeneity was explored using a test for heterogeneity (I^2), the level of significance was set at $P < 0{\cdot}10$ two-sided and a moderator analysis was performed. Subgroup analysis or meta-regression was performed, when possible, for study design, mean patient age at the beginning of the treatment, presence of an expansion phase, treatment duration and gender. For continuous data, the Cohen's d coefficient or standard mean difference (SMD) with a 95% confidence interval, as well as the standard error, was calculated, using different methods according to the data available from the primary studies or after a request of the authors. The statistical significance of the hypothesis test was set at $P < 0{\cdot}05$ (two-tailed Z-tests). Egger's test was chosen to detect publication bias if the number of included studies exceeded 10 ($P < 0{\cdot}10$).

Evaluation of the level of evidence

The level of evidence was calculated using the Grades of Recommendation, Assessment, Development and Evaluation Pro (GRADEpro GDT: GRADEpro Guideline Development Tool [Software], Available from gradepro.org.) software. This approach considers five aspects for overall risk of bias: directness of the evidence, consistency of the results, precision of the estimates, risk of publication bias and magnitude of the effect (27). The quality of the body of evidence was categorised as high, moderate, low or very low. Only RCTs were included in this analysis.

Results

Search results

The search results and the flow chart of the studies included for the analysis are shown in Table S1 and Table 1. The literature search resulted in 3745 published articles, of which 109 were considered potentially relevant. An independent review of the 109 full-text articles led to the exclusion of 88. The most common causes were absence of control [15], control group not of Class III [19], historical control group [23] and the absence of cephalometric analysis [17] (Table S3). The final sample consisted of 21 articles (12–14, 28–46) and comprised seven RCT, eight CCT and six Ret studies (Table S4).

Of 21 studies, 13 were analysed in the meta-analysis of FM with 15 independent treated groups examined (13, 14, 30, 31, 33–35, 38, 40, 42–45), and three studies (12, 28, 41) with four independent treated groups were analysed in the meta-analysis of CC.

Characteristics of the studies

The age range of the treated group was between 5·6 ± 1·0 years (34) to 12·5 ± 0·7 years (45) while the control group ranged from 4·8 ± 1·4 years (34) to 11·5 ± 1·1 years (30). The treatment time variation was from 5·2 months (34) to 60 months (32) although it was not reported in one paper (28). The full observational time was often similar to the treatment time because only two studies reported follow-up data (14, 30). Many appliances were used in these studies: the CC was used in three studies (12, 28, 41), and 16 studies used only FM (13, 30, 33–35, 40, 42, 43, 45), or FM associated with expansion (RME+FM) (14, 38, 40, 43, 44), with Bionator III (FM+BIO) (31), with miniplates (FM+MP) (38, 39) or with splints (FM+splint) (44) in at least one group. Only 13 of 21 studies based the skeletal Class III diagnosis on cephalometric criteria. Success description was reported in 16 studies, including mainly dental outcomes (Table S4).

Quality assessment

The quality assessment is shown in Table 2. In particular, only two studies (14, 37) had a blinded design and four studies (14, 37, 39, 44) examined the study power (*a priori* or *a posteriori*). Furthermore, three studies (14, 37, 44) analysed the methodological error and 11 studies had an adequate statistical analysis. Interestingly, only three studies (32, 39, 46) did not have a sufficient external validity due to the distributions of the confounding factors. Two RCTs of seven (14, 37) achieved the maximum score in both internal validity domains. The two examiners presented a high level of concordance ($k = 0{\cdot}88$). The Cochrane risk of bias tool results are shown in Table 3.

Skeletal effects

All the studies reported skeletal effects of the treatment. On the sagittal plane, most of the studies showed effects on both the maxilla and mandible, depending on the appliance used. The best improvement of the ANB, SNA and SNB angles was in the Kajiyama (34) study (+7·83°, +4·16°, −3·66°). The

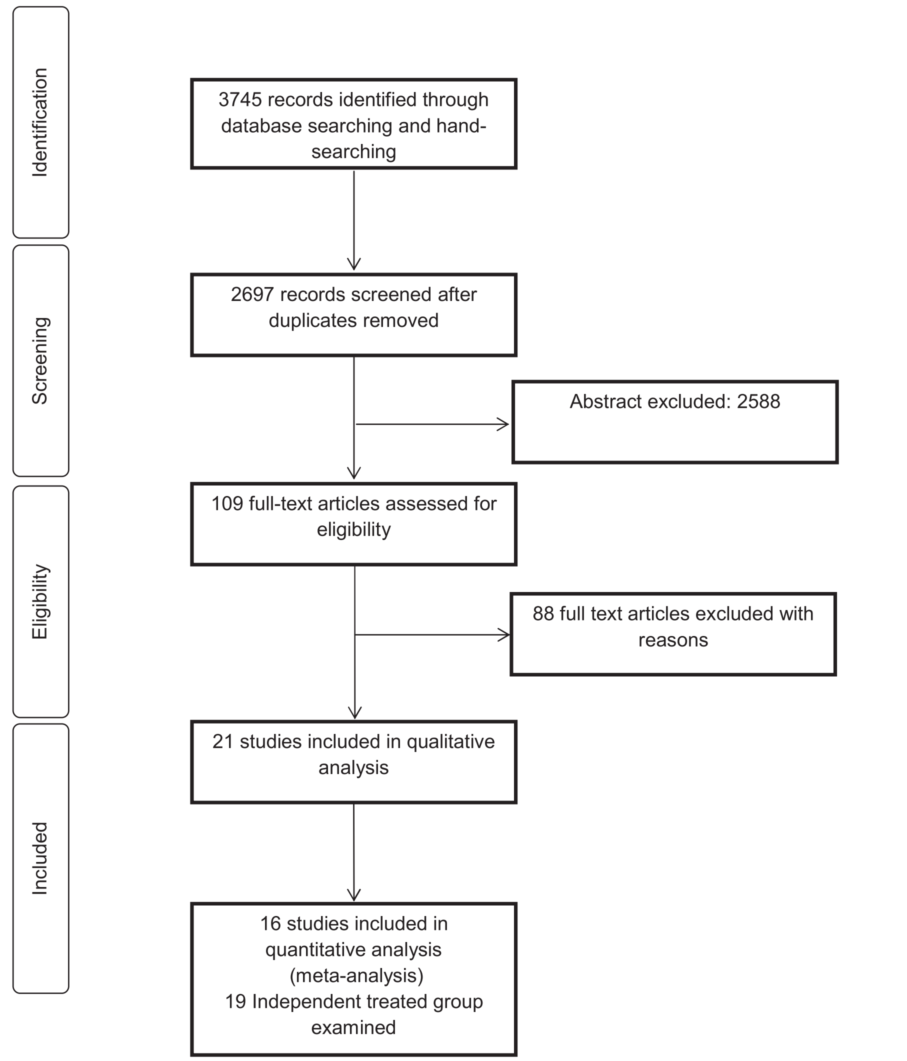

Table 1. Flow diagram of the included study

biggest increase of maxillary length (+4·4 mm) was in Falck and Zimmermann-Menzel's (32) study using FR-III, while with RME+FM the greatest increase was reached in Vaughn's study (43) (+4·29 mm). The highest control of mandibular length was in the Yagci and Uysal study (44) with RME+FM (−6·4 mm). Two studies reported significant effects only on the upper jaw (31, 42) (Tables S5 and S6).

In the FM meta-analysis, strong effects were found on ANB, SNA, SNB, with a significant Egger's test: ANB (P < 0·001); SNA (P < 0·001); SNB (P = 0·016) (Fig. 1a–c). Some moderators explained the heterogeneity of ANB and SNA, in fact for both variables, the Ret studies presented better results (ANB, P = 0·001; SNA, P = 0·004) (ANB, SMD = 6·63 CI = 3·82–9·44, SNA, SMD = 3·68 CI = 2·17–5·20) than CCT (ANB, SMD = 3·09 CI = 2·18–4, SNA, SMD = 1·74 CI = 0·95–2·54) and RCT (ANB, SMD = 1·90 CI = 1·30–2·5, SNA, SMD = 1·05 CI = 0·51–1·59) (Figure S1a,b). Moreover, changes in SNA were lower (P = 0·038) with expansion than without expansion (EXP, SMD = 1·41 CI = 0·68–2·13, NO EXP, SMD = 2·64 CI = 1·73–3·55) (Figure S1c).

For maxillary length a significant increase was found with a significant Egger's test (P < 0·001) (Fig. 2a). The high heterogeneity might be explained by an effect of the study design (P = 0·008) as Ret studies showed higher values (SMD = 3·54 CI = 1·26–5·81), than the CCT (SMD = 0·71 CI = 0·26–1·15) and RCT (SMD = 1·51 CI = 0·99–2·03) (P = 0·008)

Table 2. Risk of bias analysis of the included studies according to the Downs and Black scale (25)

	Reporting (11)	External validity (3)	Internal validity (Bias) (6)	Internal validity (Selection bias) (6)	Power (1)	Total	
Abdelnaby and Nassar (28)	8	3	4	4	0	19	Medium quality
Arun and Erverdi (12)	7	3	5	4	0	19	Medium quality
Atalay and Tortop (29)	8	3	5	4	0	20	Medium quality
Chen *et al.* (30)	9	3	4	3	0	19	Medium quality
Cozza *et al.* (31)	8	3	4	3	0	18	Medium quality
Falck and Zimmermann-Menzel (32)	6	2	4	3	0	15	Low quality
Göyenc and Ersoy (13)	10	3	5	3	0	21	Medium-high quality
Kajiyama *et al.* (33)	8	3	3	2	0	16	Low quality
Kajiyama *et al.* (34)	8	3	2	3	0	16	Low quality
Kiliçoğlu and Kirliç (35)	8	3	4	3	0	18	Medium quality
Mandall *et al.* (14)	10	3	6	6	1	26	High quality
Saleh *et al.* (37)	9	3	6	6	1	25	High quality
Sar *et al.* (38)	9	3	4	3	0	19	Medium quality
Sar *et al.* (39)	6	2	5	3	1	17	Medium quality
Tortop *et al.* (40)	9	3	5	3	0	20	Medium quality
Tuncer *et al.* (41)	9	3	3	2	0	17	Medium quality
Ucem *et al.* (42)	8	3	5	3	0	19	Medium quality
Vaughn *et al.* (43)	8	3	4	4	0	19	Medium quality
Yagci and Uysal (44)	10	3	5	4	1	23	Medium-high quality
Yuksel *et al.* (45)	9	3	4	3	0	19	Medium quality
Zhao *et al.* (46)	10	2	4	2	0	18	Medium quality

In the brackets the maximum score achievable for each domain.
≤ 16: low quality; $16 < x \leq 20$: medium quality; $20 < x \leq 24$: medium-high quality; $24 < x \leq 27$: high quality.

Table 3. Cochrane risk of bias tool (26)

	Abdelnaby and Nassar (28)	Arun and Erverdi (12)	Atalay and Tortop (29)	Kiliçoğlu and Kirliç (35)	Mandall *et al.* (14)	Saleh *et al.* (37)	Vaughn *et al.* (43)
Sequence generation	Unclear	Low risk	Low risk	Unclear	Low risk	Low risk	Low risk
Allocation concealment	Unclear	High risk	High risk	Unclear	Low risk	Low risk	Unclear
Blinding of participants, personnel and outcome assessors	Unclear	High risk	High risk	Unclear	Low risk	Low risk	Low risk
Incomplete outcome data	Unclear	Low risk	Low risk	Unclear	Low risk	Low risk	Unclear
Selective outcome reporting	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Other sources of bias	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Overall risk of bias	High risk	High risk	High risk	Unclear	Low risk	Low risk	Unclear

(Figure S2a). The FM also produced an effect on mandibular length (Fig. 2b), and this parameter was unaffected by publication bias ($P = 0{\cdot}51$). None of the analysed moderators was able to explain the medium heterogeneity.

In CC meta-analysis, significant changes were found for ANB, SNA and SNB (Fig. 1a–c), while there was no effect on mandibular length (Fig. 2b). No data were found on maxillary length.

The GRADE scale showed a low level of evidence that FM produced a decrease of SNB and a very low level of evidence that it corrected ANB and SNA, due to the high chance of a publication bias. Moreover, there was a very low level of evidence that the FM

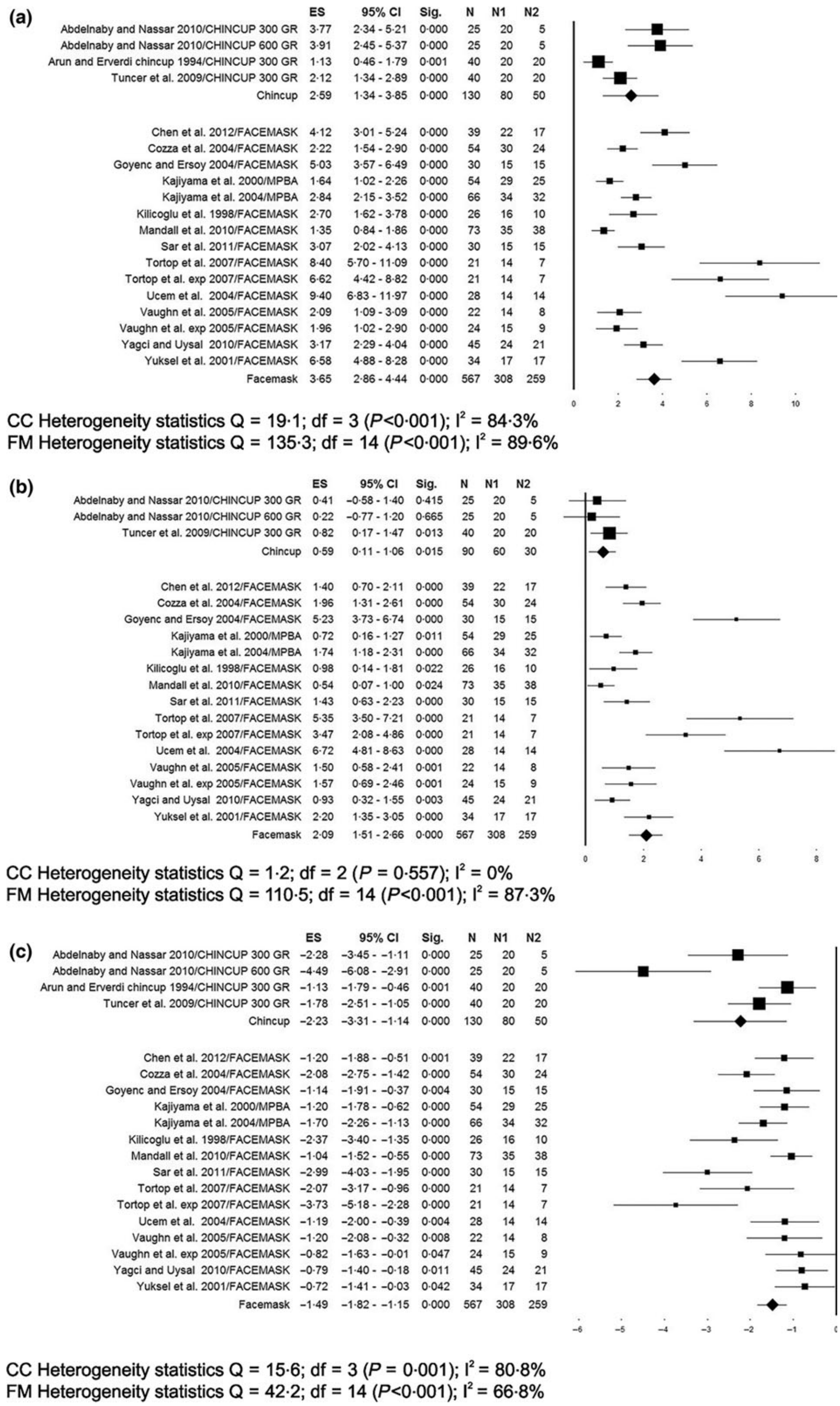

Fig. 1. Meta-analyses of ANB (a), SNA (b) and SNB (c) changes when comparing the orthopaedic treatment with the chin cup (CC) and with the facemask (FM) with no treatment. Forest plot for the standard mean including the source studies, effect sizes with 95% confidence intervals, statistical significance, number of total (*N*), treated (*N*1) untreated (*N*2) participants and assessments of heterogeneity (I^2).

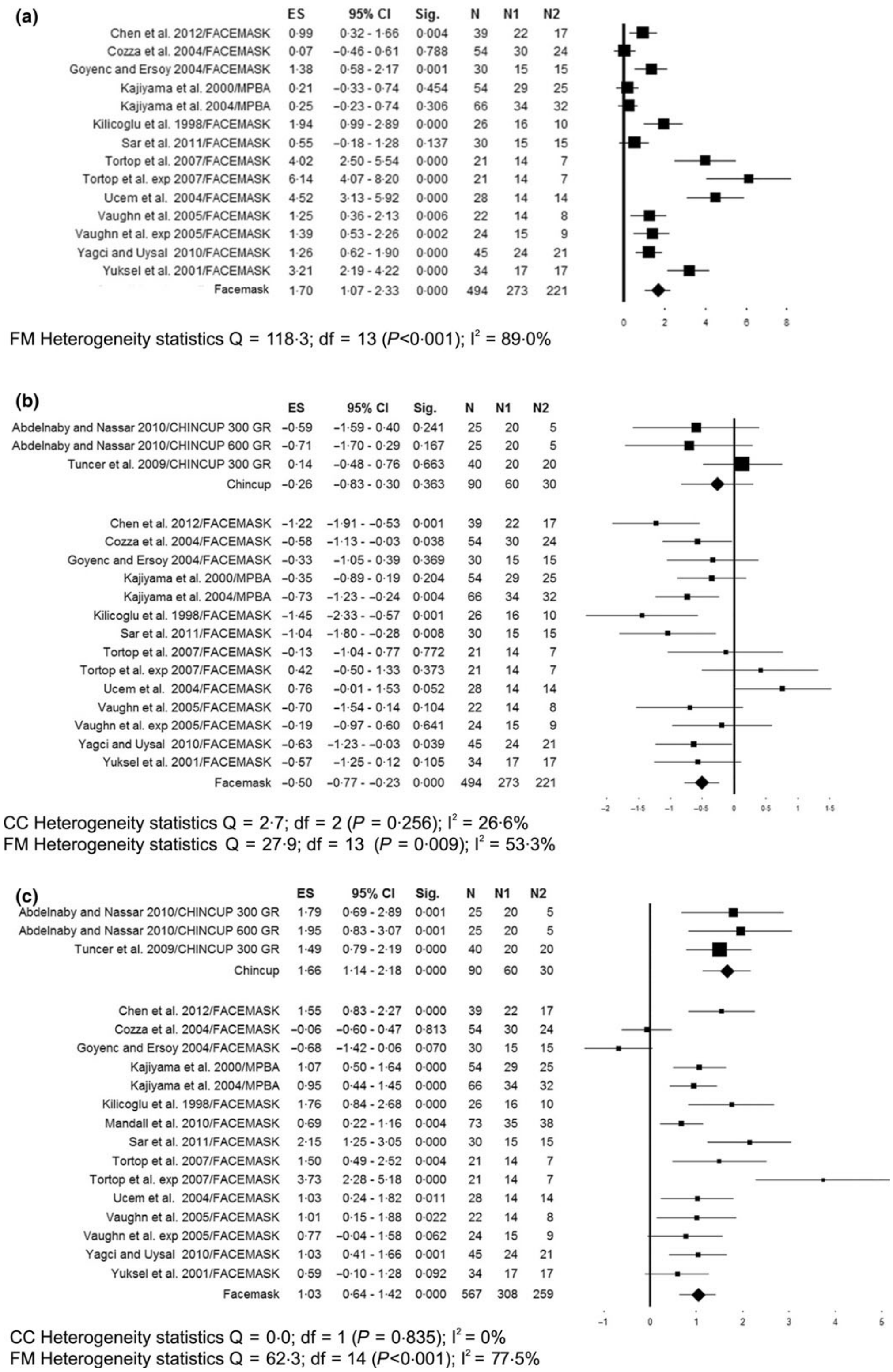

Fig. 2. Meta-analyses of maxillary length (a), mandibular length (b) and mandibular divergence (c) changes when comparing the orthopaedic treatment with chin cup (CC) and with the facemask (FM) with no treatment. Forest plot for the standard mean including the source studies, effect sizes with 95% confidence intervals, statistical significance, number of total (*N*), treated (*N*1) untreated (*N*2) participants and assessments of heterogeneity (I^2).

controlled mandibular growth and promoted maxillary growth due to the high risk of bias and the small sample size of the considered studies (Table 4). For CC the level of evidence was very low for the following outcomes: ANB, SNA, SNB and mandibular length for the high risk of bias (Table 5).

Regarding vertical changes 6 of 21 studies showed a counterclockwise rotation of the mandible (32, 37) or stability of patient divergence (13, 31, 42, 45). The greatest counterclockwise rotation was in Falck's study (32) with a reduction of the gonial angle of −7·17° and of −5·15° for the mandibular plane angle. On the other hand, 15 studies reported a significant increase in mandibular divergence. Yagci and Uysal (44) reported the highest increase in SN-Mp +4·2° (Tables S5 and S6).

In the meta-analysis of FM there was an increase in mandibular divergence with a significant (Fig. 2c) Egger's linear regression ($P = 0{\cdot}031$). The mandibular divergence was influenced by the age of treatment: the older the patient, the higher the degree of clockwise rotation (Slope = 0·24; $P = 0{\cdot}027$) (Figure S2c). In addition, CC caused a clockwise rotation of the mandible (Fig. 2c).

Finally, according to the GRADE evaluation, there was a low level of evidence that FM increased the mandibular divergence and a very low level of evidence for CC (Tables 4 and 5).

Dental effects

One study did not evaluate dental effects of the orthopaedic treatment (41). The overjet was always corrected during facemask treatment, but only eight studies reported data (14, 29, 38–40, 42, 45, 46) (Table 5; Table S5). In FM meta-analysis, there was a significant correction of the overjet (Fig. 3a). Two of the moderators assessed explained in part the high heterogeneity; in fact, retrospective studies ($P < 0{\cdot}001$) (Ret, SMD = 11·79 CI = 9·07–14·51, $P < 0{\cdot}001$; CCT, SMD = 6·27 CI = 4·53–8·01, $P < 0{\cdot}001$; RCT, SMD = 1·87 CI = 1·32–2·42, $P < 0{\cdot}001$) and no expansion ($P = 0{\cdot}011$) (NO EXP, SMD = 12·67 CI = 9·40–15·94, $P < 0{\cdot}001$; EXP, SMD = 5·67 CI = 1·35–9·99, $P < 0{\cdot}01$) reported greater values of overjet correction (Figure S3a,b). Overbite was evaluated in seven studies (29, 38–40, 42, 45, 46), and in six of seven studies the overbite significantly decreased. Four studies included in FM meta-analysis showed a significant decrease (Fig. 3b). The subgroup analysis revealed that retrospective studies ($P < 0{\cdot}001$) (Ret, SMD = −3·95 CI = −4·93 to 2·98, $P < 0{\cdot}001$; CCT, SMD = −0·95 CI = −1·71 to 0·20, $P = 0{\cdot}013$) and no expansion ($P = 0{\cdot}026$) (NO EXP, SMD = −4·32 CI = −5·28 to 3·36, $P < 0{\cdot}001$; EXP, SMD = −1·87 CI = −3·8 to 0·06, $P = 0{\cdot}026$) had a greater decrease in overbite (Figure S3c,d).

Two studies did not evaluate incisor inclinations (37, 45) and another evaluated only upper incisor inclination (43). Four papers did not find any change in incisors inclination after treatment with CC, MHG or FM (12, 31, 40, 44). At the end of the functional treatment, six studies found a retroclination or a stable position of upper incisors (14, 35, 38, 39, 43, 44) while eleven showed a proclination (13, 29, 30, 32–34, 38, 39, 42, 45, 46); three found a proclination or a stable position of lower incisors (13, 30, 39) while nine showed a retroclination (13, 28, 29, 33, 34, 38, 39, 42, 46) (Tables S5 and S7).

The meta-analysis for FM found a significant proclination on the upper incisors (Fig. 3c) without a publication bias ($P = 0{\cdot}943$). The study design ($P = 0{\cdot}028$) affected the result with RCTs (SMD = 0·10 CI = −0·32 to 0·52), showing no effects together with Ret (SMD = 0·72 CI = −0·48 to 1·93) while CCT showed a higher effect (SMD = 0·82 CI = 0·50–1·15) (Figure S4a). Only 10 of 13 studies included in the meta-analysis of FM evaluated the inclination of the lower incisors and did not find a significant effect of the appliance (Fig. 3d), and this parameter was not affected by a publication bias ($P = 0{\cdot}82$). None of the analysed moderators explained the heterogeneity.

For CC no data were available on overjet and overbite, and only one study (28) was included in the meta-analysis, which showed a significant proclination of upper incisors and retroclination of lower incisors (Fig. 3c,d).

According to the GRADE, there was a moderate level of evidence that FM corrected the overjet, a low level of evidence that FM did not have any effects on the upper incisors and a very low level that it produced a retroclination of the lower incisors (Table 4). Also for CC, the level of evidence was very low for the dental inclinations (Table 5).

Soft tissue effects

Only 10 studies, one with a removable mandibular retractor (RMR) (37), one with a magnetic orthopaedic

Table 4. Summary of findings table of the facemask studies. Should orthopaedic treatment with facemask vs. no treatment be used for correction of Class III malocclusion?

Quality assessment							No of patients		Effect		
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthopaedic treatment with facemask	No treatment	Absolute (95% CI)	Quality	Importance
ANB (assessed with: ANB angle)											
3	Randomised trials	Very serious*	Not serious	Not serious	Not serious	Publication bias strongly suspected†	80	65	SMD 1·9 higher (1·3 higher to 2·5 higher)	⊕◯◯◯ Very low	Critical
SNA (assessed with: SNA angle)											
3	Randomised trials	Very serious*	Not serious	Not serious	Not serious	Publication bias strongly suspected†	80	65	SMD 1·05 higher (0·51 higher to 1·59 higher)	⊕◯◯◯ Very low	Critical
SNB (assessed with: SNB angle)											
3	Randomised trials	Very serious*	Not serious	Not serious	Not serious	None	80	65	SMD 1·27 lower (1·83 lower to 0·71 lower)	⊕⊕◯◯ Low	Important
Maxillary length (assessed with: Maxillary length in mm)											
2	Randomised trials	Very serious‡	Serious§	Not serious	Serious¶	None	72	45	SMD 1·51 higher (0·99 higher to 2·03 higher)	⊕◯◯◯ Very low	Important
Mandibular length (assessed with: Mandibular length in mm)											
2	Randomised trials	Very serious‡	Not serious	Not serious	Serious¶	None	72	45	SMD 0·75 lower (1·47 lower to 0·04 lower)	⊕◯◯◯ Very low	Important
Mandibular divergence (assessed with: Different mandibular plane angle)											
3	Randomised trials	Very serious*	Not serious	Not serious	Not serious	None	80	65	SMD 0·96 higher (0·52 higher to 1·39 higher)	⊕⊕◯◯ Low	Critical
Overjet correction (assessed with: Overjet)											
1	Randomised trials	Not serious	Not serious	Not serious	Not serious	Publication bias strongly suspected‡‡	35	38	SMD 1·87 higher (1·32 higher to 2·42 higher)	⊕⊕⊕◯ Moderate	Critical
Upper incisor inclination (assessed with: U1 angle)											
3	Randomised trials	Very serious*	Not serious	Not serious	Not serious	None	80	65	SMD 0·1 higher (0·32 lower to 0·52 higher)	⊕⊕◯◯ Low	Important

(continued)

Table 4. (continued)

Quality assessment							No of patients		Effect		
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthopaedic treatment with facemask	No treatment	Absolute (95% CI)	Quality	Importance
Lower incisor inclination (assessed with: L1 angle)											
2	Randomised trials	Very serious**	Not serious	Not serious	Serious††	None	51	48	SMD 0·73 lower (1·19 lower to 0·26 lower)	⊕◯◯◯ Very low	Important
Upper lip position (assessed with: UL position in mm)											
2	Randomised trials	Serious**	Serious§	Not serious	Serious¶	None	45	27	SMD 2·32 higher (0·86 higher to 3·77 higher)	⊕◯◯◯ Very low	Critical
Low lip position (assessed with: LL position in mm)											
1	Randomised trials	Very serious‡	Not serious	Not serious	Serious¶	Publication bias strongly suspected‡‡	16	10	SMD 0·27 lower (1·07 lower to 0·52 higher)	⊕◯◯◯ Very low	Critical
Profile (assessed with: Facial angles)											
1	Randomised trials	Very serious‡	Not serious	Not serious	Serious¶	Publication bias strongly suspected‡‡	16	10	SMD 2·45 lower (3·49 lower to 1·42 lower)	⊕◯◯◯ Very low	Critical

CI, Confidence interval; SMD, Standardised mean difference.
*Two of the three studies had no blinding and no allocation concealment.
†A publication bias, detected by Egger's linear test, was found in the meta-analysis.
‡The RCTs did not have the allocation concealment and blinding.
§High heterogeneity $I^2 > 75$.
¶Some studies had no sufficient sample size.
**One of two studies had no blinding and no allocation concealment.
††One of two studies had no sufficient sample size.
‡‡There is only one RCT that investigated this aspect.

Table 5. Summary of findings table of the chin cup studies. Should orthopaedic treatment with chin cup vs. no treatment be used for correction of Class III malocclusion?

Quality assessment							No of patients		Effect		
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Orthopaedic treatment with chin cup	No treatment	Absolute (95% CI)	Quality	Importance
ANB (assessed with: ANB)											
2	Randomised trials	Very serious*	Serious§	Not serious	Not serious	Publication bias strongly suspected¶	60	30	SMD 2·86 higher (0·78 higher to 4·93 higher)	⊕◯◯◯ Very low	Critical
SNA (assessed with: SNA)											
1	Randomised trials	Very serious*	Not serious	Not serious	Serious†	Publication bias strongly suspected‡	40	10	SMD 0·31 higher (0·38 lower to 1·01 higher)	⊕◯◯◯ Very low	Critical
SNB (assessed with: SNB)											
2	Randomised trials	Very serious*	Serious§	Not serious	Not serious	None	60	30	SMD 2·51 lower (4·29 lower to 0·74 lower)	⊕◯◯◯ Very low	Important
Mandibular length (assessed with: Ar-Me)											
1	Randomised trials	Very serious*	Not serious	Not serious	Serious†	Publication bias strongly suspected‡	40	10	SMD 0·65 lower (1·35 lower to 0·06 higher)	⊕◯◯◯ Very low	Important
Mandibular divergence (assessed with: SNGoMe)											
1	Randomised trials	Very serious*	Not serious	Not serious	Serious†	Publication bias strongly suspected‡	40	10	SMD 1·87 higher (1·09 higher to 2·65 higher)	⊕◯◯◯ Very low	Critical
Upper incisor inclination (assessed with: U1-SN)											
1	Randomised trials	Very serious*	Not serious	Not serious	Serious†	Publication bias strongly suspected‡	40	10	0·87 higher (0·16 higher to 1·59 higher)	⊕◯◯◯ Very low	Important
Lower incisor inclination (assessed with: L1-MP)											
1	Randomised trials	Very serious*	Not serious	Not serious	Serious†	Publication bias strongly suspected‡	40	10	SMD 2·14 lower (2·95 lower to 1·32 lower)	⊕◯◯◯ Very low	Important

CI, Confidence interval; SMD, standardised mean difference.
*Blinding and allocation concealment were not described.
†Sample size was not sufficient.
‡A publication bias was suspected because there was just one RCT on this outcome.
§High heterogeneity $I^2 > 75$.
¶A publication bias, detected by Egger's linear test, was found in the meta-analysis.

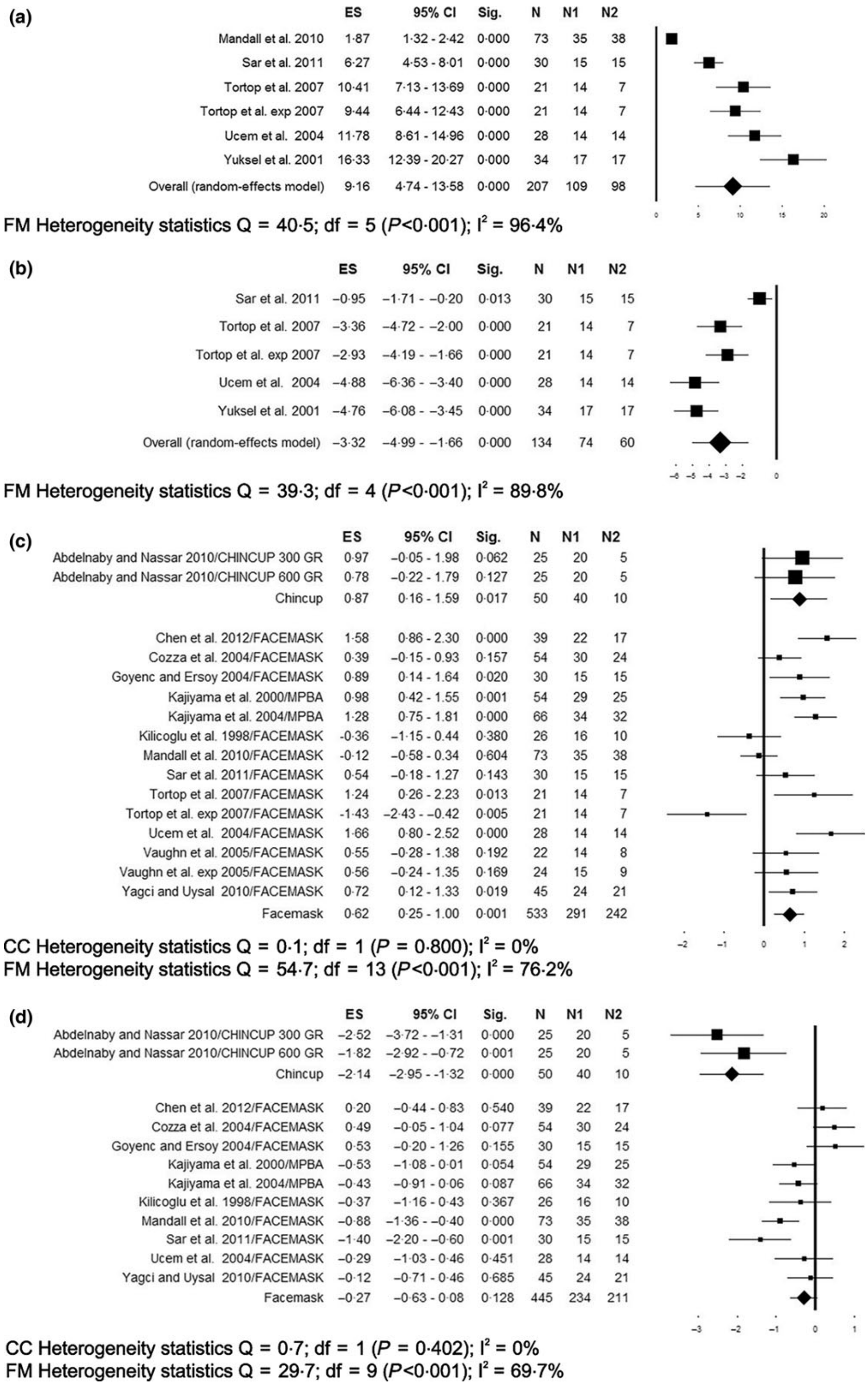

Fig. 3. Meta-analyses of overjet (a), overbite (b), upper incisor inclination (c) and lower incisor inclination (d) changes when comparing the orthopaedic treatment with chin cup (CC) and with the facemask (FM) with no treatment. Forest plot for the standard mean including the source studies, effect sizes with 95% confidence intervals, statistical significance, number of total (*N*), treated (*N*1) untreated (*N*2) participants and assessments of heterogeneity (I^2).

appliance (MOA-III) (46) and eight with an FM (13, 31, 35, 38, 39, 42–44), reported effects on the profile. All studies described positive profile changes and just one showed a stable position of the lower lip (46) (Tables S5 and S8). The meta-analysis on FM showed a significant advancement of the upper lip, control of the lower lip and improvement of the profile (Fig. 4a–c). Furthermore, in the moderator analysis, the upper lip position was influenced by age at treatment start; in fact, the FM might cause a higher improvement in younger patients (Slope = −0·73; $P = 0{\cdot}003$) (Figure S4b).

The level of evidence assessed by the GRADE was very low for the three parameters (Table 4).

The meta-analysis for CC did not present any study that assessed soft tissue changes.

Follow-up data

Only two studies have follow-up information and both used FM. Chen *et al.* (30) reported a 2-year follow-up of ten patients. Among these, six had stable mandibular growth and decreased ANB, while four had maxillary retrusion, mandible protrusion and horizontal mandibular growth direction. Mandall *et al.* (36) reported a 3-year follow-up with increased ANB in 86% of the initial sample, downward and backward rotation of the maxilla, upward and forward rotation of the occlusal plane and correction of the overjet, without vertical changes (Table S5).

Discussion

The aim of this systematic review was to summarise the effects of the orthopaedic treatment with different appliances on growing skeletal Class III patients compared with an untreated control group, assessing skeletal, dental or soft tissues variables, and to estimate the effect of study design. Class III orthopaedic therapy includes several appliances, not always assessed by RCTs (13, 32, 46). Other systematic reviews comprised randomised and non-randomised studies without evaluating if the study design could affect the results (18–22). Indeed, an interesting finding of this review was that some results may be overestimated due to the study design. All the retrospective studies that chose the treatment group with a successful criterion, or without a concurrent control group, were excluded due to the chance of a misinterpretation and invalidation of the study results (47, 48). These data should be carefully assessed in orthodontics because most of our knowledge is based on CCT and/or retrospective studies that could not provide an adequate evidence-based support (49).

Characteristics of the studies

Still many controversies are present in the early treatment of Class III malocclusion. The age for starting treatment varied from 5 to 13 years old, with a large range in treatment time. In addition, the forces applied and time-wear of the appliances differed among studies. Only in 13 studies, the inclusion criteria for the diagnosis of the skeletal Class III malocclusion were based on cephalometric data or soft tissue evaluation, while others used dental parameters. Therefore, one crucial limit is the scarce emphasis given to skeletal and profile evaluation, both for the diagnosis and the treatment outcome.

Quality assessment

The Downs and Black checklist (25) was chosen to evaluate external validity and report domains, which are not present in other tools. Only two RCTs (14, 37) were considered of high quality with a low level of bias. On the other hand, two CCTs (13, 44) were of medium-high quality, higher than other RCTs, meaning that RCTs may not always have the best quality. The quality of the studies analysed was generally medium mainly for the lack of adequate statistics and follow-up data.

Skeletal effects

All selected studies reported sagittal skeletal changes, suggesting that orthopaedic therapy is effective to correct Class III malocclusions with a low or very low level of evidence. Consistent with previous reviews (16, 17, 19–21), ANB showed a significant change. Not all the appliances had effects on SNA and maxillary length. Only the FM in the meta-analysis, in accordance with other studies (16, 17, 20, 21, 50, 51), had a strong effect in the short term on both SNA and maxillary length, with a very low level of evidence, Also the CC increased SNA, but this finding was not consistent with another meta-analysis (19). Similarly, SNB in the meta-analysis, for FM and CC,

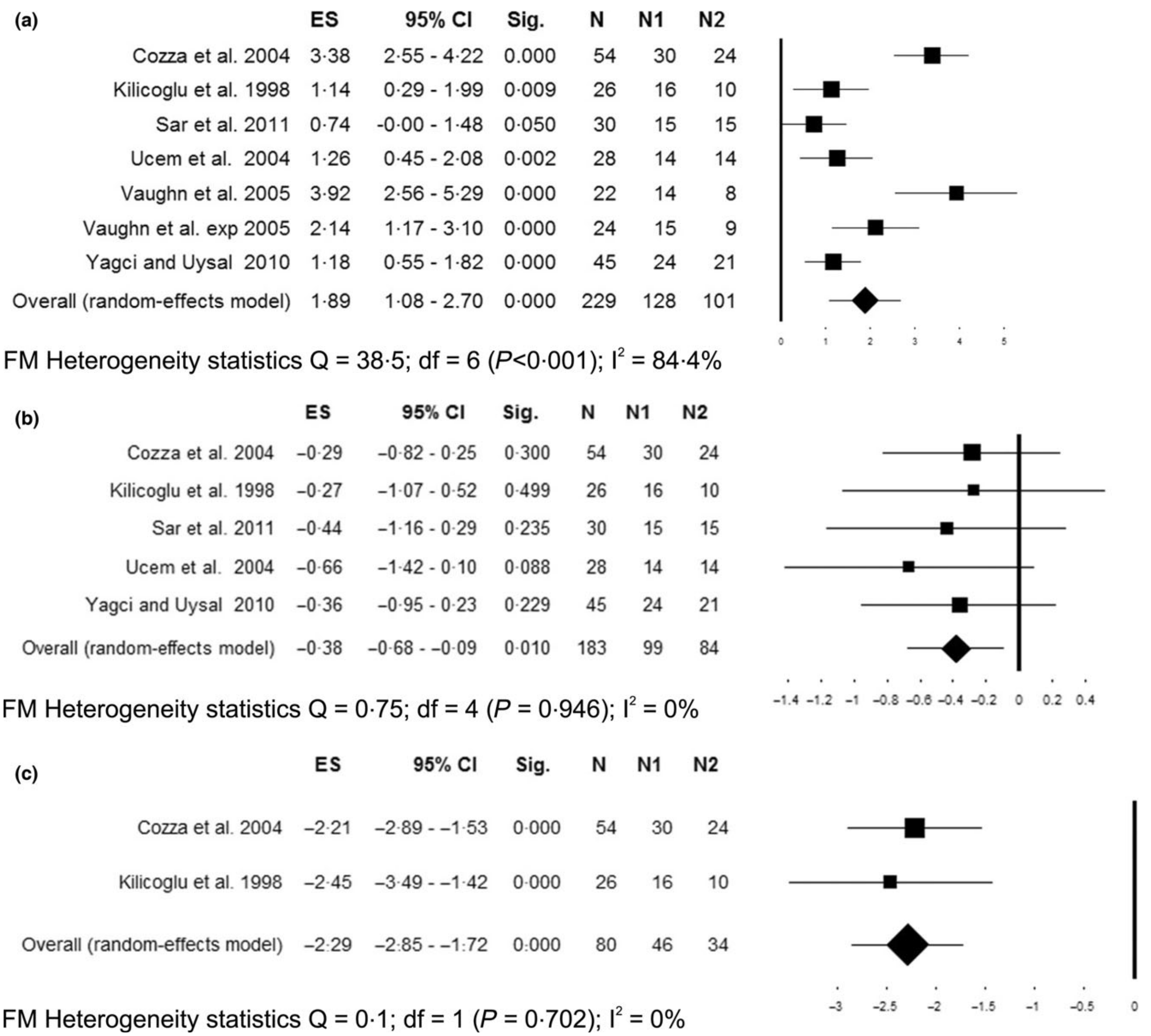

Fig. 4. Meta-analyses of upper lip position (a), lower lip position (b) and profile (c) changes when comparing the orthopaedic treatment with chin cup (CC) and with the facemask (FM) with no treatment. Forest plot for the standard mean including the source studies, effect sizes with 95% confidence intervals, statistical significance, number of total (*N*), treated (*N*1) untreated (*N*2) participants and assessments of heterogeneity (I^2).

and Mandibular length only for FM, showed a significant control of the mandible. Analogous results were found for FM (16, 17, 20, 21) and for CC (19). However, it must be stressed that the sagittal control of the mandible assessed by angular measurements (SNB, ANB) suffers from the influence of a clockwise rotation of the mandible, enhancing the apparent amount of sagittal effect (52). Indeed, except for FR-III (32), supported by one study of low quality, and for RMR (37), supported by one RCT of high quality, all the other appliances determined a clockwise rotation of the mandible. This effect was also confirmed in other reviews (16, 17, 19–21). Furthermore, it should be take into account that most of the studies started treatment in older patients and, as showed in the moderator analysis, older patients have higher increases in the divergence.

Hence, even if the early orthopaedic treatment of Class III malocclusion could be effective in the short term, it should not be recommended in hyperdivergent and older patients in case the increase of the mandibular divergence is unwanted.

Dental effects

There was a moderate level of evidence that FM corrects the overjet while there was no evidence, according to the GRADE, that FM or CC affect the overbite. Nonetheless, in six of seven studies there was a decrease in the overbite often associated with an increase in mandibular divergence. This confirms the importance of being warned about the vertical morphology when treating skeletal Class III malocclusion.

Controversial data were reported for incisor inclinations. The meta-analysis for the FM showed a significant proclination of upper incisors, which was confirmed in another study (20), while, according to one RCT (28) CC proclined the upper incisors and retroclined the lower incisors. Nevertheless, the GRADE revealed a level of evidence from very low to low.

Hence, the orthopaedic treatment corrects the overjet, but due to the controversial data, there is still a lack of evidence on the molar relationship and other dental effects.

Soft tissue effects

To our knowledge, this is the first meta-analysis of the effects on soft tissues. Few studies reported improvement on soft tissue (13, 31, 35, 37–39, 42–44, 46). However, the GRADE showed a very low level of evidence for FM. It is well known that, from the patient's point of view, the success of the therapy is strongly related to aesthetic improvement (53). Also a patient's quality of life is influenced by aesthetics (54, 55). There is the need for future studies to focus on the objective and subjective evaluations of soft tissue changes by taking advantage of new 3D technology (56, 57).

Hence, although the level of evidence is very low, the orthopaedic treatment seems to improve the facial profile mainly in younger patients.

Follow-up

There is still insufficient follow-up evidence. Two (30) and 3-year (36) follow-up data showed a relapse in about 15% of patients. One 6-year follow-up study suggested that class III protraction facemask treatment reduces the need for orthognathic surgery in adulthood, indeed the group without treatment showed an odd of needing surgery 3·5 times higher than the group treated with FM. Moreover, 68% of patient treated with FM maintained a positive overjet. Nevertheless, no improvement in quality of life in treated patients respect to untreated subjects was reported (9).

Hence, due to the scarce available information, it is not possible to establish if the early functional treatment prevents the relapse and the need for surgery in adulthood, and if the early treatment has an impact in the quality of life of skeletal Class III subjects.

One limit of this systematic review was the use of SMD for the meta-analysis. Even though the SMD decreases the possible discrepancies in terms of magnification, variables assessed and study method error, interpretation by clinicians is difficult. Moreover, the inclusion of non-randomised studies could be considered a limit (58, 59), but it was supported to provide a wider overview on this topic.

Conclusions

1. The quality of the primary studies was medium-low. Patient selection, blinding assessment and statistical analysis were often inadequate.
2. The study design might lead to an overestimation of the results; hence, there is a need for well designed RCTs.
3. An improvement of the sagittal skeletal relationship was reported with all the orthopaedic appliances, but only a few appliances were analysed by more than one study.
4. The FM seems to correct Class III discrepancies, but it might determine a clockwise rotation of the lower jaw and a decreased overbite.
5. There was controversial evidence on the dental effects of the orthopaedic appliances; however, there was a moderate level of evidence that the FM corrects overjet.
6. Soft tissue improvements were reported in all the studies assessing this outcome.
7. There was insufficient evidence to assess the long-term stability of Class III orthopaedic treatment.
8. The level of evidence supporting the efficacy of FM or CC varied from very low to moderate.

Further studies are needed to achieve enough information in early treatment of Class III malocclusion;

the scarce presence of follow-up data in high quality studies does not allow for an evaluation on stability and utility of orthopaedic treatment in the long term.

Acknowledgments

The authors thank Dott. Paolo Chiodini for his support for the statistical analysis. This study did not receive any funding.

Disclosure

The authors have stated explicitly that there are no conflict of interests in connection with this article.

References

1. Sugawara J, Mitani H. Facial growth of skeletal class III malocclusion and the effects, limitations, and long-term dentofacial adaptations to chincap therapy. Semin Orthod. 1997;3:244–254.
2. Reyes BC, Baccetti T, McNamara JA Jr. An estimate of craniofacial growth in class III malocclusion. Angle Orthod. 2006;76:577–584.
3. Baccetti T, Franchi L, McNamara JA Jr. Growth in the untreated class III subject. Semin Orthod. 2007;13:130–142.
4. Baccetti T, Reyes BC, McNamara JA Jr. Craniofacial changes in class III malocclusion as related to skeletal and dental maturation. Am J Orthod Dentofacial Orthop. 2007;132:171.e1–12.
5. Alexander AE, McNamara JA Jr, Franchi L, Baccetti T. Semilongitudinal cephalometric study of craniofacial growth in untreated class III malocclusion. Am J Orthod Dentofacial Orthop. 2009;135:1–14.
6. Georgalis K, Woods MG. A study of class III treatment: orthodontic camouflage vs orthognathic surgery. Aust Orthod J. 2015;31:138–148.
7. Lee CH, Park HH, Seo BM, Lee SJ. Modern trends in class III orthognathic treatment: a time series analysis. Angle Orthod. 2017;87:269–278.
8. Harrington C, Gallagher JR, Borzabadi-Farahani A. A retrospective analysis of dentofacial deformities and orthognathic surgeries using the index of orthognathic functional treatment need (IOFTN). Int J Pediatr Otorhinolaryngol. 2015;79:1063–1066.
9. Mandall N, Cousley R, DiBiase A, Dyer F, Littlewood S, Mattick R *et al.* Early class III protraction facemask treatment reduces the need for orthognathic surgery: a multicentre, two-arm parallel randomized, controlled trial. J Orthod. 2016;43:164–175.
10. Fränkel R. The treatment of class II, division 1 malocclusion with functional correctors. Am J Orthod. 1969;55:265–275.
11. Sugawara J, Asano T, Endo N, Mitani H. Long-term effects of chincap therapy on skeletal profile in mandibular prognathism. Am J Orthod Dentofacial Orthop. 1990;98:127–133.
12. Arun T, Erverdi N. A cephalometric comparison of mandibular headgear and chin-cap appliances in orthodontic and orthopaedic view points. J Marmara Univ Dent Fac. 1994;2:392–398.
13. Göyenç Y, Ersoy S. The effect of a modified reverse headgear force applied with a facebow on the dentofacial structures. Eur J Orthod. 2004;26:51–57.
14. Mandall N, DiBiase A, Littlewood S, Nute S, Stivaros N, McDowall R *et al.* Is early class III protraction facemask treatment effective? A multicentre, randomized, controlled trial: 15-month follow-up. J Orthod. 2010;37:149–161.
15. De Clerck H, Cevidanes L, Baccetti T. Dentofacial effects of bone-anchored maxillary protraction: a controlled study of consecutively treated class III patients. Am J Orthod Dentofacial Orthop. 2010;138:577–581.
16. Cordasco G, Matarese G, Rustico L, Fastuca S, Caprioglio A, Lindauer SJ *et al.* Efficacy of orthopedic treatment with protraction facemask on skeletal class III malocclusion: a systematic review and meta-analysis. Orthod Craniofac Res. 2014;17:133–143.
17. Watkinson S, Harrison JE, Furness S, Worthington HV. Orthodontic treatment for prominent lower front teeth (class III malocclusion) in children. Cochrane Database Syst Rev. 2013;9:CD003451.
18. Yang X, Li C, Bai D, Su N, Chen T, Xu Y *et al.* Treatment effectiveness of Fränkel function regulator on the class III malocclusion: a systematic review and meta-analysis. Am J Orthod Dentofacial Orthop. 2014;146:143–154.
19. Chatzoudi MI, Ioannidou-Marathiotou I, Papadopoulos MA. Clinical effectiveness of chin cup treatment for the management of class III malocclusion in pre-pubertal patients: a systematic review and meta-analysis. Prog Orthod. 2014;2:62.
20. Foersch M, Jacobs C, Wriedt S, Hechtner M, Wehrbein H. Effectiveness of maxillary protraction using facemask with or without maxillary expansion: a systematic review and meta-analysis. Clin Oral Investig. 2015;19:1181–1192.
21. Zhang W, Qu HC, Yu M, Zhang Y. The effects of maxillary protraction with or without rapid maxillary expansion and age factors in treating class III malocclusion: a meta-analysis. PLoS ONE. 2015;10:e0130096.
22. Yepes E, Quintero P, Rueda ZV, Pedroza A. Optimal force for maxillary protraction facemask therapy in the early treatment of class III malocclusion. Eur J Orthod. 2014;36:586–594.
23. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine: how to practice and teach EBM. New York (NY): Churchill Livingston; 1997.
24. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol. 2009;62:e1–e34.
25. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of

randomised and non-randomised studies of health care interventions. J Epidemiol Community Health. 1998;52:377–384.
26. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions. Version 5.1.0. The Cochrane Collaboration [Internet]. Available from: www.cochrane-handbook.org, accessed March 2014.
27. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J *et al.* GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64:383–394.
28. Abdelnaby YL, Nassar EA. Chin cup effects using two different force magnitudes in the management of class III malocclusions. Angle Orthod. 2010;80:957–962.
29. Atalay Z, Tortop T. Dentofacial effects of a modified tandem traction bow appliance. Eur J Orthod. 2010;32:655–661.
30. Chen L, Chen R, Yang Y, Ji G, Shen G. The effects of maxillary protraction and its long-term stability–a clinical trial in Chinese adolescents. Eur J Orthod. 2012;34:88–95.
31. Cozza P, Marino A, Mucedero M. An orthopaedic approach to the treatment of class III malocclusions in the early mixed dentition. Eur J Orthod. 2004;26:191–199.
32. Falck F, Zimmermann-Menzel K. Cephalometric changes in the treatment of class III using the Fränkel appliance. J Orofac Orthop. 2008;69:99–109.
33. Kajiyama K, Murakami T, Suzuki A. Evaluation of the modified maxillary protractor applied to class III malocclusion with retruded maxilla in early mixed dentition. Am J Orthod Dentofacial Orthop. 2000;118:549–559.
34. Kajiyama K, Murakami T, Suzuki A. Comparison of orthodontic and orthopedic effects of a modified maxillary protractor between deciduous and early mixed dentitions. Am J Orthod Dentofacial Orthop. 2004;126:23–32.
35. Kiliçoglu H, Kirliç Y. Profile changes in patients with class III malocclusions after Delaire mask therapy. Am J Orthod Dentofacial Orthop. 1998;113:453–462.
36. Mandall N, Cousley R, DiBiase A, Dyer F, Littlewood S, Mattick R *et al.* Is early class III protraction facemask treatment effective? A multicentre, randomized, controlled trial: 3-year follow-up. J Orthod. 2012;39:176–185.
37. Saleh M, Hajeer MY, Al-Jundi A. Short-term soft- and hard-tissue changes following class III treatment using a removable mandibular retractor: a randomized controlled trial. Orthod Craniofac Res. 2013;16:75–86.
38. Sar C, Arman-Özçırpıcı A, Uçkan S, Yazıcı AC. Comparative evaluation of maxillary protraction with or without skeletal anchorage. Am J Orthod Dentofacial Orthop. 2011;139:636–649.
39. Sar C, Sahinoğlu Z, Özçirpici AA, Uçkan S. Dentofacial effects of skeletal anchored treatment modalities for the correction of maxillary retrognathia. Am J Orthod Dentofacial Orthop. 2014;145:41–54.
40. Tortop T, Keykubat A, Yuksel S. Facemask therapy with and without expansion. Am J Orthod Dentofacial Orthop. 2007;132:467–474.
41. Tuncer BB, Kaygisiz E, Tuncer C, Yüksel S. Pharyngeal airway dimensions after chin cup treatment in class III malocclusion subjects. J Oral Rehabil. 2009;36:110–117.
42. Ucem TT, Ucuncü N, Yüksel S. Comparison of double-plate appliance and facemask therapy in treating class III malocclusions. Am J Orthod Dentofacial Orthop. 2004;126:672–679.
43. Vaughn GA, Mason B, Moon HB, Turley PK. The effects of maxillary protraction therapy with or without rapid palatal expansion: a prospective, randomized clinical trial. Am J Orthod Dentofacial Orthop. 2005;128:299–309.
44. Yagci A, Uysal T. Effect of modified and conventional facemask therapy on condylar position in class III patients. Orthod Craniofac Res. 2010;13:246–254.
45. Yüksel S, Uçem TT, Keykubat A. Early and late facemask therapy. Eur J Orthod. 2001;23:559–568.
46. Zhao N, Feng J, Hu Z, Chen R, Shen G. Effects of a novel magnetic orthopaedic appliance (MOA-III) on the dentofacial complex in mild to moderate skeletal class III children. Head Face Med. 2015;14:34.
47. Antoun JS, Cameron C, Sew Hoy W, Herbison P, Farella M. Evidence of secular trends in a collection of historical craniofacial growth studies. Eur J Orthod. 2015;37:60–66.
48. Papageorgiou SN, Koretsi V, Jäger A. Bias from historical control groups used in orthodontic research: a meta-epidemiological study. Eur J Orthod. 2017;3:98–105.
49. Papageorgiou SN, Xavier GM, Cobourne MT. Basic study design influences the results of orthodontic clinical investigations. J Clin Epidemiol. 2015;68:1512–1522.
50. Kim JH, Viana MA, Graber TM, Omerza FF, BeGole EA. The effectiveness of protraction face mask therapy: a meta-analysis. Am J Orthod Dentofacial Orthop. 1999;115:675–685.
51. Jäger A, Braumann B, Kim C, Wahner S. Skeletal and dental effects of maxillary protraction in patients with angle class III malocclusion. A meta-analysis. J Orofac Orthop. 2001;62:275–284.
52. Hussels W, Nanda RS. Analysis of factors affecting angle ANB. Am J Orthod. 1984;85:411–423.
53. Tuncer C, Canigur Bavbek N, Balos Tuncer B, Ayhan Bani A, Çelik B. How do patients and parents decide for orthodontic treatment-effects of malocclusion, personal expectations, education and media. J Clin Pediatr Dent. 2015;39:392–399.
54. Bucci R, Rongo R, Zito E, Galeotti A, Valletta R, D'Antò V. Cross-cultural adaptation and validation of the Italian psychosocial impact of dental aesthetics questionnaire (PIDAQ). Qual Life Res. 2015;24:747–752.
55. Benson PE, Da'as T, Johal A, Mandall NA, Williams AC, Baker SR *et al.* Relationships between dental appearance, self-esteem, socio-economic status, and oral health-related quality of life in UK school children: a 3-year cohort study. Eur J Orthod. 2015;37:481–490.
56. Antoun JS, Lawrence C, Leow A, Rongo R, Dias G, Farella M. A three-dimensional evaluation of Māori and New Zealand European faces. Aust Orthod J. 2014;30:169–175.
57. Rongo R, Antoun JS, Lim YX, Dias G, Valletta R, Farella M. Three-dimensional evaluation of the relationship between jaw divergence and facial soft tissue dimensions. Angle Orthod. 2014;84:788–794.

58. Bucci R, D'Antò V, Rongo R, Valletta R, Martina R, Michelotti A. Dental and skeletal effects of palatal expansion techniques: a systematic review of the current evidence from systematic reviews and meta-analyses. J Oral Rehabil. 2016;43:543–564.
59. D'Antò V, Bucci R, Franchi L, Rongo R, Michelotti A, Martina R. Class II functional orthopaedic treatment: a systematic review of systematic reviews. J Oral Rehabil. 2015;42:624–642.

Correspondence: Roberto Rongo, Department of Neurosciences, Reproductive Sciences and Oral Sciences, Division of Orthodontics, University of Naples 'Federico II', Via S. Pansini 5, 80131 Naples, Italy.
E-mails: roberto.rongo@unina.it; roberto.rongo@gmail.com

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Subgroup analysis of ANB for study design (a), SNA for study design (b) and of SNA for expansion phase (c).

Figure S2 Subgroup analysis of maxillary length for study design (a), and metaregression of mandibular divergence for age at the start of treatment.

Figure S3 Subgroup analysis of overjet for study design (a), overjet for expansion phase (b) overbite for study design (c) overbite for expansion phase (d).

Figure S4 Subgroup analysis of upper incisor inclination for study design (a), and metaregression of upper lip position for age at the start of treatment.

Table S1 Research strategy, database used and search results up to February 2016.

Table S2 Inclusion and exclusion criteria used in this meta-analysis.

Table S3 List of the articles excluded on basis of full-text with exclusion reasons.

Table S4 Characteristics of the 21 studies included in the systematic review.

Table S5 Brief description of the results found for the 21 included studies.

Table S6 Cephalometric parameters of the skeletal changes, between before treatment and after treatment, reported in the primary studies.

Table S7 Cephalometric parameters of the dental changes, between before treatment and after treatment, reported in the primary studies.

Table S8 Cephalometric parameters of the soft tissue changes, between before treatment and after treatment, reported in the primary studies.