Effects of facemasks versus intraoral appliances in treating maxillary deficiency in growing patients: A systematic review and meta-analysis

Porównanie zastosowania masek twarzowych i aparatów wewnątrzustnych w leczeniu niedorozwoju szczęki u rosnących pacjentów – systematyczny przegląd piśmiennictwa oraz metaanaliza

Nasrin Farhadian1,B,E,F, Sepideh Soheilifar1,A–F, Masoud Abolvardi1,A,B,D–F, Amirfarhang Miresmailei1,E,F, Younes Mohammadi2,B,C

1 Department of Orthodontics, Dental School, Hamadan University of Medical Sciences, Iran
2 Modeling of Non-Communicable Diseases Research Center, Department of Epidemiology, School of Public Health, Hamadan University of Medical Sciences, Iran

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Abstract

Background. Class III malocclusion is one of the most challenging types of orthodontic problems.

Objectives. The aim of this study was to compare the dentoskeletal effects of facemasks and intraoral appliances in treating class III maxillary deficiency in growing patients through a systematic review of the available literature.

Material and methods. Electronic and manual searches were performed in the Cochrane Central Register of Controlled Trials (CENTRAL), the Cochrane Database of Systematic Reviews (CDSR), MEDLINE (PubMed), Embase (OVID), and Scopus to find all the relevant studies published by January 2018. All randomized controlled trials (RCTs) recruiting 5–12-year-old patients who received maxillary protraction treatment with any type of facemask and comparing the facemasks with any type of intraoral appliance were included. The primary outcome measure was changes in the A point-nasion-B point angle (ANB), and the secondary outcomes included changes in the overjet, upper-1 (U1) inclination, the mandibular plane angle, and treatment time. The meta-analysis was carried out using the inverse variance-weighted random effects model.

Results. Out of 1,629 articles found in the initial search, 5 studies met the inclusion criteria. The meta-analysis showed no differences in the duration of treatment or in any of the cephalometric variables, with the exception of the overjet.

Conclusions. It seems that intraoral appliances and facemasks are similar in terms of dentoskeletal effects in the treatment of class III malocclusion as well as treatment duration. However, due to a lack of a sufficient number of high-quality studies, these results should be viewed with caution. Further high-quality, long-term studies are recommended.

Key words: systematic review, meta-analysis, class III malocclusion treatment, facemask, intraoral appliance

Słowa kluczowe: systematyczny przegląd piśmiennictwa, metaanaliza, leczenie wady zgryzu klasy III, maska twarzowa, aparat wewnątrzustny
Introduction

The prevalence of class III malocclusion varies, with rates as high as 26.6% and a higher prevalence reported among East Asian populations, which means that a remarkable proportion of orthodontic patients are affected. The treatment of class III malocclusion is one of the most challenging kinds of orthodontic treatment. This type of malocclusion can occur for a variety of reasons, including maxillary deficiency, mandibular prognathism, or a combination of the two. Maxillary deficiency, with an incidence of approx. half of all cases, is the main cause. Therefore, maxillary protraction is considered to be the primary treatment option in class III cases. In growing patients, this goal can be achieved using either extraoral protraction facemasks or intraoral appliances.

Generally, a facemask is the most widely used orthopedic device to treat growing patients with maxillary deficiency. The therapeutic results are acceptable in terms of both skeletal and dentoalveolar aspects, and are achieved after a relatively short time. This device has also some disadvantages, including soft tissue irritation and abrasion, and bulkiness. These factors can compromise the patient compliance, definitely affecting the therapeutic results.

On the other hand, a large number of intraoral devices have been introduced to treat class III malocclusion, including Fränkel’s functional regulator, the reverse twin block, the modified Jasper Jumper device, the modified tandem traction bow appliance, and fixed tongue appliances and plates. They do not have bulky external appearance or the potential for soft tissue irritation, which are associated with a facemask. Fixed intraoral appliances also eliminate the patient compliance issues. Therefore, they seem to improve the results of maxillary protraction. The efficacy of these devices has been investigated in various studies. However, there is some controversy in the literature over the issue, with some studies observing differences between these modalities and preferring one over the other, and other studies showing no significant differences between these methods of treatment. Two recently published systematic reviews on the early treatment of class III malocclusion focused on the comparison of groups treated with various kinds of appliances against untreated control groups. Neither of these studies compared different therapies. Therefore, this systematic review and meta-analysis was designed and carried out to evaluate the efficacy of facemasks and selected intraoral devices in treating maxillary deficiency in growing class III malocclusion patients based on the studies previously published on the subject.

Material and methods

Protocol and registration

The study was carried out based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. It was registered in the PROSPERO database under the code CRD42017063810 (www.crd.york.ac.uk/PROSPERO).

Search strategy, the information sources and the type of included studies

All the relevant studies were searched for with no restrictions on the language or the date of publication until April 2017 (updated in January 2018) in the following databases: Cochrane Central Register of Controlled Trials (CENTRAL); Cochrane Database of Systematic Reviews (CDSR); MEDLINE (PubMed); Embase (OVID); and Scopus.

The search strategy was based on the one developed for PubMed (Table 1), but revised for each database according to their search strategy instructions.

Additionally, a manual search was carried out through the reference lists of the finally included articles, and the relevant systematic reviews and orthodontic journals not indexed in PubMed.

All searches were carried out by 2 independent investigators (S.S. and M.A.). Duplicate studies were removed in the 1st step of screening.

Only randomized clinical trials (RCTs) were considered for this systematic review and meta-analysis.

Criteria of eligibility

The inclusion criteria for the studies were chosen according to the PICO guidelines for research and evidence-based practice. Participants. Studies on 5–12-year-old patients with skeletal class III malocclusion, with no craniofacial syndromes or any type of facial clefts, who underwent early orthopedic/orthodontic maxillary protraction treatment with any removable or fixed appliances were included; studies with any bone-anchored device or any surgical procedure to facilitate protraction were excluded.

Intervention. Maxillary protraction treatment using any intraoral appliances, including functional appliances and tongue appliances, or the modified versions of them were sought, with no restrictions on the treatment protocol or duration.

Comparison. Maxillary protraction treatment with any type of facemask, i.e., Delaire, Petit, Nanda, a rail-type facemask, or any other modifications, were included, with no restrictions on the treatment protocol or duration.
Outcomes:

– the primary outcome criteria were the A point-nasion-B point angle (ANB) changes obtained with a face-mask vs intraoral devices, measured on a lateral cephalogram;

– the secondary outcomes sought were other skeletal and dental differences between the treatment modalities, differences in the duration of treatment and any complications that occurred with regard to either treatment.

Study selection and data extraction

After eliminating duplicates, all the studies selected by the search were examined in order to remove the irrelevant ones. Firstly, the names of authors and journals, and the result sections of the articles were deleted so as to prevent any possible reviewer or publishing bias. Then, the titles were screened. The articles with titles that made them potentially eligible were reviewed based on their abstracts. Then, the full texts of the selected articles were retrieved and screened for final inclusion based on the eligibility criteria. This selection of the studies was carried out by 2 investigators (S.S. and M.A.) independently and in duplicate. All doubts and disagreements were resolved after discussion.

A customized data extraction form was created by 2 researchers (S.S. and M.A.) and used by each of these investigators independently and in duplicate. The information collected in this form included the study and the author’s name, the year of publication, the type of study, the appliances used for intervention and comparison, the number and age of participants in each group, and the mean differences and standard deviations (SDs) of the outcome measures.

For any missing information, an attempt was made to contact the corresponding author.

Risk of bias and quality assessment

The risk of bias and methodological quality assessment was undertaken by 2 researchers (N.F. and S.S.) independently and in duplicate. It was carried out using the Cochrane Collaboration tool for assessing the risk of bias, comprising 6 categories: 1 – selection bias with a further division into ‘randomization sequence generation’ and ‘allocation concealment’; 2 – performance bias, which evaluated the blindness of the person who performed the intervention; 3 – detection bias, which evaluated the blindness of the person who detected or interpreted the results; 4 – attrition bias, which refers to the withdrawal of the patients; 5 – reporting bias, which refers to selective outcome reporting; and 6 – other sources of bias. The Review Manager (RevMan) software, v. 5.3 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014) was used to summarize the risk of bias. Each category was judged as high, low or unclear.

Data synthesis and analysis

The data analysis was carried out using comprehensive meta-analysis software (Comprehensive Meta-Analysis, v. 2; Borenstein M, Hedges L, Higgins J, Rothstein H. Biostat, Engelwood, NJ, 2005, 104). We calculated the summary measure with 95% uncertainty from different studies for each outcome. Due to the existence of heterogeneity, the meta-analysis was carried out with the inverse variance-weighted random effects approach. The \( \chi^2 \) test was used to assess heterogeneity between the studies. Statistical significance was set at \( p < 0.05 \). Funnel plots were used to assess any publication bias of the articles.
Results

Search findings and study selection

The initial search after eliminating duplicates identified 1,629 records. Screening by title and abstract resulted in 167 and 60 studies, respectively. After reviewing the full texts by means of detailed inclusion and exclusion criteria, 55 articles were excluded, as they did not fulfill the inclusion criteria of the study (case reports, case series, retrospective studies or non-randomized studies, lacking a comparison group, including surgical intervention, or including any craniofacial syndromes or any kind of facial clefts). Finally, 5 studies were included in the systematic review. Excluding Kurt et al., the 4 remaining studies were used for the meta-analysis after quality assessment (Fig. 1). The particulars of the selected studies are presented in Table 2.

Risk of bias and quality assessment of the included studies

The quality assessment and the risk of bias assessment criteria for RCTs were adapted from Cochrane Handbook for Systematic Reviews of Interventions (Fig. 2,3).29

Selection bias (sequence generation and allocation concealment)

Sequence generation was adequate for all studies, with the exception of that by Kurt et al. Husson et al.24 and Showkatbakhsh et al. (2012)24 used a random number generator. Showkatbakhsh et al. (2015)9 and Showkatbakhsh et al. (2013)8 used an unstratified subject allocation sequence. However, Kurt et al. assigned patients to treatment groups in accordance with their order of arrival at the clinic.22
### Table 2. Characteristics of the included studies

<table>
<thead>
<tr>
<th>Study design</th>
<th>Comparison(s)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>– single-blind</td>
<td>– single-blind – parallel-group</td>
<td>– SNB ≤ 80° – ANB ≤ 0° – Wits appraisal s−1 mm</td>
</tr>
<tr>
<td>– parallel-group</td>
<td>– fixed tongue appliance group (26 patients)</td>
<td>– normal mandibular growth pattern – ANB ≤ 0° – skeletal class III relationship with a concave profile</td>
</tr>
</tbody>
</table>
Allocation concealment was adequate for Showkatbakhsh et al. (2015)\(^9\) and Showkatbakhsh et al. (2013),\(^8\) whereas the studies by Husson et al.,\(^14\) Showkatbakhsh et al. (2012),\(^24\) and Kurt et al.,\(^22\) were judged as unclear on allocation concealment.

**Performance and detection bias**

Due to the nature of the studies, comparing 2 different treatment modalities, the blinding of participants and clinicians could not be performed, and therefore it was not assessed. Conversely, the blinding of outcome detectors and statisticians was possible in all studies. Regarding detection bias, there was a low risk of bias in Husson et al.,\(^14\) Kurt et al.,\(^22\) and in 2 studies by Showkatbakhsh et al. (2013, 2012),\(^8,24\) whereas it was unclear in the study by Showkatbakhsh et al. (2015), where no blinding was mentioned.\(^9\)

**Attrition bias**

The rates of withdrawal were clearly reported in the 2013 study by Showkatbakhsh et al.\(^8\) Interestingly, Husson et al.,\(^14\) and Showkatbakhsh et al. (2012)\(^24\) reported no dropouts in their studies. These 3 studies were judged as having a low risk of bias. The attrition bias of the study by Kurt et al. was judged to be unclear.\(^22\) On the other hand, Showkatbakhsh et al. (2015) reported that 4 patients had dropped out of the study due to personal reasons.\(^9\) In response to our e-mail, the authors declared that these 4 patients rejected the facemask due to its bulky size. Therefore, the risk of attrition bias in this study was assessed as high.

**Reporting bias**

The risk of reporting bias for all studies was judged to be low. All 5 RCTs reported the outcomes that they had set out to report and there was no obviously missing data.\(^8,9,14,22,24\)

### Overall risk of bias of randomized controlled trials

The study by Showkatbakhsh et al. (2013)\(^8\) showed an overall low risk of bias. Showkatbakhsh et al.'s (2015)\(^9\) and Kurt et al.'s studies\(^22\) were assessed as having a high risk of bias due to attrition bias and selection bias, respectively. The 2 remaining studies (by Husson et al.\(^14\) and by Showkatbakhsh et al. (2012)\(^24\)) were assessed to have an unclear risk of bias (Fig. 2,3).

### Summary of the studies and meta-analysis

The summary of findings of the final articles is reported in Table 3. Four cephalometric variables and treatment duration were selected, and the mean differences and SDs for each group were compared:

- overjet (Husson et al.\(^14\));
- ANB (all 4 studies);
- upper-1 (U1) inclination (all studies apart from Husson et al.\(^14\));
- mandibular plane angle (all 4 studies).

Furthermore, in order to investigate publication bias, funnel plots were constructed for each variable. These plots show symmetries in the publication of the studies (Fig. 4).

### Overjet

Since only 1 study (by Husson et al.\(^14\)) had measured this variable, the meta-analysis could not be performed for the RCT subgroup. In the study by Husson et al., the overjet was 3.53 ±1.24 mm in the facemask group and 2.52 ±1.45 mm in the intraoral device group.\(^14\) This difference was statistically significant (\(p = 0.04\)).

### A point-nasion-B point angle

The mean difference in ANB was 1.63° in the facemask group and 1.59° in the intraoral device group. The meta-analysis was performed for all 4 RCTs. The pooled estimate

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Group</th>
<th>Age [years]</th>
<th>ANB changes [°]</th>
<th>Mandibular plane angle changes [°]</th>
<th>Overjet changes [mm]</th>
<th>U1 inclination changes [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Husson et al. (2016)(^14)</td>
<td>RCT</td>
<td>MTA</td>
<td>7.98 ±0.68</td>
<td>1.88 ±0.72</td>
<td>1.00 ±0.97</td>
<td>2.25 ±1.45</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>facemask</td>
<td>8.11 ±0.76</td>
<td>2.13 ±1.09</td>
<td>3.50 ±1.41</td>
<td>3.53 ±1.24</td>
<td>–</td>
</tr>
<tr>
<td>Showkatbakhsh et al. (2015)(^9)</td>
<td>RCT</td>
<td>fixed tongue appliance</td>
<td>8.9 ±1.7</td>
<td>1.1 ±1.7</td>
<td>–0.9 ±2.4</td>
<td>–</td>
<td>43 ±7.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>facemask</td>
<td>8.5 ±1.4</td>
<td>1.4 ±1.4</td>
<td>−0.2 ±1.5</td>
<td>–</td>
<td>5.7 ±5.0</td>
</tr>
<tr>
<td>Showkatbakhsh et al. (2013)(^8)</td>
<td>RCT</td>
<td>tongue plate</td>
<td>9.1 ±0.9</td>
<td>1.8 ±1.2</td>
<td>0.4 ±1.8</td>
<td>–</td>
<td>49 ±4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>facemask</td>
<td>9.0 ±1.2</td>
<td>1.2 ±1.6</td>
<td>−0.4 ±1.6</td>
<td>–</td>
<td>6.5 ±6.0</td>
</tr>
<tr>
<td>Showkatbakhsh et al. (2012)(^24)</td>
<td>RCT</td>
<td>tongue appliance</td>
<td>10.1 ±0.7</td>
<td>1.6 ±1.6</td>
<td>−1.0 ±4.1</td>
<td>–</td>
<td>2.5 ±1.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>facemask</td>
<td>9.3 ±1.2</td>
<td>1.8 ±1.5</td>
<td>0.4 ±1.9</td>
<td>–</td>
<td>11.1 ±6.9</td>
</tr>
</tbody>
</table>

Data presented as mean ± standard deviation (SD).

U1 = upper-1.
was 0.050 (95% CI: −0.351–0.464; \( p = 0.787 \)). The difference between the groups was not statistically significant and the \( \chi^2 \) test showed heterogeneity (\( p < 0.001 \)) (Fig. 5).

**Upper-1 inclination**

All studies, except for that by Husson et al., had measured the U1 inclination changes in terms of each modality. The mean difference in the U1 inclination was 7.76° in the facemask group and 3.9° in the intraoral device group. The meta-analysis showed that the pooled estimate was 3.745 (95% CI: −0.583–8.074; \( p = 0.09 \)). The difference between the groups was not statistically significant and the \( \chi^2 \) test indicated heterogeneity (\( p < 0.001 \)) (Fig. 6).

**Mandibular plane angle**

All studies had measured the mandibular plane angle changes in terms of each treatment modality. The mean difference in the mandibular plane angle was 0.82° in the facemask group and −0.12° in the intraoral device group. The meta-analysis showed that the pooled estimate was 0.936 (95% CI: −0.669–2.542; \( p = 0.253 \)). The difference between the groups was not statistically significant and the \( \chi^2 \) test indicated heterogeneity (\( p < 0.001 \)) (Fig. 7).

**Duration of treatment**

All studies mentioned the exact duration of treatment in both the facemask and intraoral device groups. The mean of treatment duration was 15.1 months in the facemask group and 13.77 months in the intraoral device group. The meta-analysis showed that the pooled estimate was 1.266 (95% CI: −1.470–4.002; \( p = 0.364 \)). The difference between the groups was not statistically significant and the \( \chi^2 \) test showed heterogeneity (\( p < 0.001 \)) (Fig. 8).
Fig. 5. Meta-analysis of the differences in the A point-nasion-B point angle (ANB)
CI – confidence interval.

Fig. 6. Meta-analysis of the differences in the upper-1 (U1) inclination

Fig. 7. Meta-analysis of the differences in the mandibular plane angle

Fig. 8. Meta-analysis of the differences in treatment duration
Discussion

Summary of the main results

This systematic review was designed to appraise the available evidence comparing facemasks and intraoral devices in treating growing class III patients with maxillary deficiency. As shown in Table 2, 5 articles were included in this systematic review. Four articles had measured pre- and post-treatment cephalometric variables for the facemask and intraoral device groups, and had used them for meta-analysis. These variables were measured with regard to different references in different studies (Table 2), but by measuring the mean differences between the pre- and post-treatment values, this limitation was overcome. However, Kurt et al. compared class III growing patients treated with a facemask or an intraoral appliance (the modified Jasper Jumper device) in the case of temporomandibular disorders (TMDs). They concluded that both of these appliances were capable of reducing TMD-related symptoms, and thus could enhance the patient’s quality of life, without affecting the incidence of TMDs.

As shown in Table 3, in the remaining studies, 4 cephalometric variables were compared and analyzed, in addition to the treatment duration differences between the facemask and intraoral device groups. The comparison of the duration of treatment between these groups revealed no difference in the meta-analysis. On the other hand, no differences between the various cephalometric variables of the facemask and intraoral device groups were found, with the exception of the overjet.

Interestingly, although the analysis of the U1 inclination changes showed no differences, the results favored the intraoral device group. This might be attributed to the type of appliance used. All 3 of the relevant RCTs used a tongue plate/appliance. These appliances transferred the low functional forces of the tongue to the palate and the whole maxillary arch. These kinds of forces may produce fewer changes in the U1 inclination.

Although a recent systematic review identified the reverse overjet as the main reason for a class III malocclusion patient to seek treatment, in our research, the overjet was assessed only in 1 study. It seems that skeletal changes were more important to those authors, as they did not report the overjet changes throughout the treatment period. It appears that most of the treatment modalities can lead to the positive overjet and that most authors are focused on finding the cause of an increase in the overjet. To make systematic reviews more informative for patients as well as for researchers, reporting changes in the overjet seems to be quite helpful.

Due to the nature of a facemask appliance, patient compliance in wearing elastics clearly influences the final outcome. Therefore, it seemed that the type of the intraoral part of the facemask, i.e., removable or fixed, did not affect the treatment outcomes. Conversely, in the intraoral appliance group, some fixed appliances did not require patient compliance. Therefore, fixed or removable intraoral appliances might affect the final outcome. In 2 of the studies by Showkatbakhsh et al. (2013, 2012), removable intraoral appliances were used, but Showkatbakhsh et al. (2015) and Husson et al. combined a hyrax expansion screw with the protraction protocol. Therefore, they had to choose a fixed design for their intraoral appliances. The efficacy of maxillary protraction with or without expansion was studied by Foersch et al. in 2015 in a systematic review, with no significant improvement being observed in maxillary protraction with expansion. In the present review, due to a lack of new, high-level evidence regarding this issue since 2015, this variable was not assessed. Finally, it seems that there is no consensus regarding the best type of appliance in terms of fixed/removable and expansion/non-expansion.

Quality of the evidence

Only the study by Showkatbakhsh et al. (2013) was judged to have a low risk of bias. The studies by Husson et al. and Showkatbakhsh et al. (2012) were judged to have an unclear risk of bias. Showkatbakhsh et al’s study (2015) was assessed as having a high risk of bias.

Heterogeneity

Overall, facemask and intraoral appliances were effective in the treatment of class III malocclusion. Most of their effects seemed to be similar and they did not vary with respect to the duration of treatment. However, due to the high degree of heterogeneity in the pooled studies, the validity of this claim seems less conclusive.

Potential for bias in the reviewing process

By using a comprehensive search strategy specified for multiple databases, updating the search results as the process of systematic review continued, and having no restrictions on the language or the date of publication, the potential for bias was kept low in this systematic review.

Suggestion for future research

Due to a limited number of RCTs in the scope of this review, more well-designed RCTs seem to be necessary. The following suggestions should be considered:

– well-designed RCTs, which strictly follow the Consolidated Standards Of Reporting Trials (CONSORT) statement, especially regarding the use of fixed/removable appliances and the expansion/non-expansion protocols;
– ruling out patients with pseudo-class III;
– long-term follow-up in order to determine the stability of the treatment results;
– a uniform set of variables, including the overjet and the maxillary skeletal position, in addition to all of the variables mentioned in this review.
Conclusions

It seems that intraoral devices and facemasks have similar effects on class III malocclusion. However, due to a lack of a sufficient number of well-designed, high-quality RCTs, these conclusions must be viewed with caution.

ORCID iDs

Nasrin Farhadian https://orcid.org/0000-0001-7611-0685
Seepideh Soheilifar https://orcid.org/0000-0003-3515-1637
Masoud Abolvari https://orcid.org/0000-0002-8156-1525
Amirfarhang Miresmailei https://orcid.org/0000-0002-1796-7512
Younes Mohammadi https://orcid.org/0000-0001-7521-4909

References