Meta-analysis of skeletal mandibular changes during Fränkel appliance treatment

Letizia Perillo*, Rosangela Cannavale*, Fabrizia Ferro*, Lorenzo Franchi**, Caterina Masucci**, Paolo Chiodini*** and Tiziano Baccetti**

Departments of Orthodontics, *Second University of Naples, **University of Florence and ***Department of Medicine and Public Health, Second University of Naples, Italy

Correspondence to: Dr Letizia Perillo, Department of Orthodontics, Second University of Naples, Via De Crecchio, 80100 Naples, Italy. E-mail: letizia.perillo@unina2.it

SUMMARY The purpose of this study was to perform a meta-analysis of articles to verify the mandibular changes produced by the Fränkel-2 (FR-2) appliance during the treatment of growing patients with Class II malocclusions when compared with untreated growing Class II subjects.

The literature published from January 1966 to January 2009 was reviewed with search engines. A quality analysis was performed. The effects on primary end points were calculated with random-effect models. Heterogeneity was assessed using Q statistic and investigated using study-level meta-regression.

A total of nine articles were identified. The quality of the studies ranged from low to medium. Meta-analysis showed that the FR-2 was associated with enhancement of mandibular body length [0.4 mm/year 95 per cent confidence interval (CI) 0.182–0.618], total mandibular length (1.069 mm/year, 95 per cent CI 0.683–1.455), and mandibular ramus height (0.654 mm/year, 95 per cent CI 0.244–1.064). A consistent heterogeneity among studies was found for all the considered linear measurements.

The FR-2 appliance had a statistically significant effect on mandibular growth. Nevertheless, the heterogeneity of the FR-2 effects, the quality of studies, the differences in age, skeletal age, treatment duration, and the inconsistent initial diagnosis seem to overstate the benefits of the FR-2 appliance.

An evidence-based approach to the orthodontic outcomes of FR-2 appliance is needed, by selecting and comparing groups of children with the same cephalometric characteristics with and without treatment.

Introduction

Class II malocclusions occur in a variety of skeletal and dental configurations (Cozza et al., 2006) among which the most common appears to be mandibular skeletal retrusion (McNamara, 1981; Pancherz et al., 1997). A therapy aimed at enhancing mandibular growth is indicated in these patients.

Since the 1930s, a wide range of functional appliances designed to increase mandibular growth gained popularity in Europe and then throughout the rest of the world (McNamara et al., 1996; McNamara and Brudon, 2001; Chen et al., 2002; Cozza et al., 2006). One of the most popular and well-characterized functional appliances is the functional regulator (FR-2; Fränkel, 1966, 1969a,b, 1973, 1983; Falck and Fränkel, 1989; Perillo et al., 1996; Tulloch et al., 1997, 1998; Johnston, 1998; Chen et al., 2002). Unlike other functional appliances, the FR-2 has a mode of action based on orthopaedic principles that consider exercise and muscle training to be important factors in the normal development of osseous tissues (Fränkel, 1966, 1969a,b, 1973, 1983; Perillo et al., 1996). A specific indication for FR-2 is represented by a Class II division 1 malocclusion associated with mandibular deficiency (Fränkel, 1983; Perillo et al., 1996).

The FR-2 treatment approach has led to disparate outcomes in studies on humans (McNamara et al., 1996; McNamara and Brudon, 2001; Chen et al., 2002; Cozza et al., 2006). Some authors (Reey and Eastwood, 1978; Luder, 1982; Pancherz, 1982, 2005; Birkebak et al., 1984; McNamara et al., 1985, 1990; Haynes, 1986; Jakobsson and Paulin, 1990; Mamandras and Allen, 1990; Windmiller, 1993; Perillo et al., 1996; Pancherz et al., 1997; Tulloch et al., 1997; Franchi et al., 1999; Toth and McNamara, 1999; Tümer and Gülten, 1999; Baccetti et al., 2000; Mills and McCulloch, 2000; De Almeida et al., 2002; Bascifci et al., 2003; Faltin et al., 2003; Pangrazio-Kulbersh et al., 2003; Cozza et al., 2004) have suggested that mandibular growth can be increased, whereas others stated that mandibular length cannot be altered (Jakobsson, 1967; Vargervik and Harvold, 1985; Nelson et al., 1993; Illing et al., 1998; Chadwick et al., 2001; Janson et al., 2003; O’Brien et al., 2003) with significant treatment effects restricted to dentoalveolar changes (Tulley, 1972; Robertson, 1983; McNamara et al., 1985, 1990; Fränkel and Fränkel, 1989; Perillo et al., 1996; Toth and McNamara, 1999; Chadwick et al., 2001; McNamara and Brudon, 2001; De Almeida et al., 2002; Cozza et al., 2006).
Two systematic reviews (Chen et al., 2002; Cozza et al., 2006) have been performed on the efficacy of functional appliances on mandibular growth. Chen et al. (2002) found no significant differences between an untreated control group and subjects treated with functional appliances (FR-2 included) with the exception of linear measurements related to articulare (Ar) point. Moreover, the control and treated group did not appear to differ in the angulation of the lower incisors. The more recent systematic review, on the other hand, reported significant supplementary elongation in total mandibular length, induced by functional appliances. The FR-2 appliance, however, showed one of the lowest coefficients of efficiency when compared with other functional appliances (Herbst, twin-block).

Both systematic reviews analysed a wide variety of appliances with different modes of action (mechanical, orthopaedic, etc.) and did not focus specifically on the effects of a true functional appliance (the FR-2). Neither of the two systematic reviews performed a meta-analysis of treatment outcomes; meta-analysis investigates heterogeneity and provides a summary measure of study results.

The aim of the present investigation was, therefore, to perform a systematic review and a meta-analysis of FR-2 studies in order to assess the dental or skeletal changes induced by this appliance in growing patients with Class II malocclusions compared with changes in untreated Class II growing subjects. Moreover, because in the previous systematic reviews, only very few randomized clinical trials (RCTs) were found, the present search also included controlled clinical trials and retrospective investigations to determine weak methodologies used in those studies.

Materials and methods

Search strategy

A systematic review of the literature was developed to identify articles that address the effects of FR-2 appliance on mandibular growth. The literature search was carried out using PubMed, the Cochrane Central Register of Controlled Trials (CENTRAL), Scirus, Lilacs, Embase, and Scopus. The following search terms were used: ‘Class II malocclusion’ and ‘Fränkel appliance’ or ‘FR-2 appliance’, alternatively, and ‘Fränkel II’ and ‘Frankel-2’ alone.

To improve the search, the ‘related articles’ tool was used in the PubMed search and references of retrieved studies were checked by a research librarian.

Inclusion criteria

Studies were selected if they satisfied all the following inclusion criteria: publication date from January 1966 to January 2009; original studies based on humans; prospective and retrospective longitudinal studies, RCTs, systematic reviews, meta-analyses; studies conducted on growing patients with Class II malocclusions; concurrent untreated growing subjects, historical controls with Class II malocclusions; studies with cephalometric measurements; no restrictions were set for language.

Data collection and quality analysis

A full text version was obtained for the studies considered adequate on the basis of the abstract and for those where the abstract was inconclusive. Data were collected on the following items for the retrieved studies: year of publication, origin, study design, materials (study sample, control sample), age at the start of treatment, methods of measurement, appliance wear, treatment/observation duration, age and gender matching, and reported outcomes.

A quality evaluation of the methodological soundness of each article was performed according to a modified version of the method described by Jadad et al. (1996). The following parameters were evaluated: prospective design, randomization, prior estimate of sample size, method error analysis, blinding in measurements, and adequate statistics. Quality score was calculated by compounding previous items. Scores ranged from 0 to 6 with higher scores indicating a better methodologic quality.

The methods and results sections of each article were read and scored by two independent blind readers (LP, RC). The evaluators discussed their findings, and when disagreement occurred, it was resolved through further discussion and re-reading.

Analysis of reported outcomes

To provide a quantitative appraisal of skeletal and dental modifications in Class II patients treated with the FR-2 appliance when compared with untreated Class II controls, the following data were evaluated for each retrieved study: total mandibular length (measured as Co–Gn or Co–Pg or Ar–Gn or Ar–Pg), mandibular body length (measured as Go–Gn or Go–Me or Go–Pg), and ramus height (Ar–Go or Co–Go). Mandibular changes were annualized to accommodate variations in treatment duration thereby allowing comparison with data of other investigations. Statistically significant differences between the cephalometric evaluations of treated and untreated subjects were obtained.

Statistical analysis

From each study, the mean difference and standard deviation (SD) of the difference of each variable of interest were extracted. Only in one study (Chadwick et al., 2001) were SDs not reported and were estimated on the basis of reported confidence intervals (CIs). In the study of McNamara et al. (1985), data reported for very young and young subjects were considered separately.

Meta-analysis was performed according to a modified version of the method proposed by Curtin et al. (2002). For
each outcome, the weighted mean difference, assessed by means of inverse variance method (fixed-effects model), was calculated separately for each measurement performed. Heterogeneity was assessed using $Q$ statistic (Deeks and Bradburn, 2001) in studies that used the same cephalometric measurements, among different cephalometric measurements, and for all studies (Curtin et al., 2002); a $P$ value of $Q$ statistic lower than 0.10 was considered significant (Deeks and Bradburn, 2001). In the absence of substantial heterogeneity, the computation of the overall combined effect was based on the fixed-effect model, while if there was evidence of heterogeneity, outcomes were pooled using the random-effect model of DerSimonian and Laird (Deeks and Bradburn, 2001).

Explanatory analysis of associations between quality score and the effect of the FR-2 on cephalometric measurements was investigated using study-level meta-regression (Van Houwelingen et al., 2002). Meta-regression is a model that relates the treatment effect to study-level covariates while assuming additivity of within- and between-study components of variance. Restricted maximum likelihood estimators were used (Thompson and Sharp, 1999). Permutation testing (using 1000 Monte Carlo simulations, Stata Corp., College Station, Texas, USA) was used to calculate $P$ values and to reduce the chance of spurious false-positive findings (Higgins and Thompson, 2004).

Analysis was performed using Stata, version 9.0 (Stata Corp.), and R 2.4.1 (R Foundation for Statistical Computing, Vienna, Austria). The quality of outcomes of meta-analysis (Meta-analysis of Observational Studies in Epidemiology; Stroup et al., 2000) was used for evaluation of the results.

**Results**

The PubMed search strategy identified 97 articles: 36 were selected on the basis of the abstract information. Three other articles were selected on the basis of the references of these 36 papers, so that 39 full-text articles were finally selected for further evaluation. Thirty-two of the 39 papers were excluded as they did not meet the criteria of inclusion, while seven articles were considered eligible for inclusion in the review.

From the Scirus search, five full-text articles were selected on the basis of the abstract information. Three were excluded as they did not meet the inclusion criteria, while two were included in the review. The search in the CENTRAL gave no results. Lilacs identified 38 studies: 20 full-text articles selected on the basis of the abstract information were excluded as they did not meet the inclusion criteria. Embase located 54 studies: none of these could be used because they were double publications. Scopus found 118 studies: the chosen 15 full-text articles did not meet the inclusion criteria.

The main reasons for exclusion of the studies were Class III malocclusions, no cephalometric analysis, expert opinion, case reports and double publications, other functional appliances, no data on linear mandibular measurements, adult patients, no control groups, and no quantitative data.

The review process identified a total of nine studies as reported in Figure 1.

Characteristics of the included studies (McNamara et al., 1985; Haynes, 1986; Falck and Fränkel, 1989; Nelson et al., 1993; Perillo et al., 1996; Toth and McNamara, 1999;

![Figure 1](image-url)  
**Figure 1** Flow chart of the selection process of studies on mandibular linear changes during functional regulator appliance treatment.
Chadwick et al., 2001; De Almeida et al., 2002; Janson et al., 2003) are reported in Table 1.

The total number of patients treated with the FR-2 appliance in the nine studies was 366 and 320 for the controls.

**Quality analysis of the studies**

The quality of the studies ranged from low to medium (Table 1).

Seven studies (McNamara et al., 1985; Falck and Fränkel, 1989; Perillo et al., 1996; Toth and McNamara, 1999; Chadwick et al., 2001; De Almeida et al., 2002; Janson et al., 2003) were retrospective clinical trials, while only two were prospective clinical trials and reported previous estimates of sample size (Haynes, 1986; Nelson et al., 1993). Randomization was used in only one study (Nelson et al., 1993). Two studies (McNamara et al., 1985; Haynes, 1986) did not include a method error analysis, and only one (Chadwick et al., 2001) used blinding in measurements. Two studies (Chadwick et al., 2001; Janson et al., 2003) used appropriate statistical methods, while seven (McNamara et al., 1985; Haynes, 1986; Falck and Fränkel, 1989; Nelson et al., 1993; Perillo et al., 1996; Toth and McNamara, 1999; De Almeida et al., 2002) applied parametric tests in samples that were not tested for normality.

**Quantitative analysis of mandibular changes**

**Changes in mandibular body length.** Eight studies (McNamara et al., 1985; Haynes, 1986; Falck and Fränkel, 1989; Nelson et al., 1993; Perillo et al., 1996; Toth and McNamara, 1999; De Almeida et al., 2002; Janson et al., 2003) evaluated changes in mandibular dimension (Table 2 and Figure 2) using the following cephalometric measurements: Go–Pg, Go–Me, and Go–Gn. Significant heterogeneity of changes in mandibular body length was found among different cephalometric measurements ($P = 0.03$) and for all studies but not within studies that used the same cephalometric measurements (Table 3). Therefore, a random-effect model was used to estimate the overall effect. The FR-2 was associated with significant enhancement of mandibular body length [0.400 mm/year, 95 per cent confidence interval (CI) 0.182–0.618; Figure 2] compared with untreated subjects.

**Changes in mandibular total length.** Nine studies (McNamara et al., 1985; Haynes, 1986; Falck and Fränkel, 1989; Nelson et al., 1993; Perillo et al., 1996; Toth and McNamara, 1999; Chadwick et al., 2001; De Almeida et al., 2002; Janson et al., 2003) considered this outcome (Table 2 and Figure 2) as indicated by the following cephalometric measurements: Co–Gn, Ar–Pg, Ar–B, Ar–M, Co–Pg, and Ar–Gn.

Significant heterogeneity of changes in mandibular total length was found within and among different cephalometric measurements ($P = 0.021$) and for all studies (Table 3). Thus, to estimate the overall effect, a random-effect model was used. FR-2 was associated with significant enhancement of mandibular total length (1.069 mm/year, 95 per cent CI 0.683–1.455; Figure 2) compared with untreated subjects.

**Changes in mandibular ramus height.** Nine studies (McNamara et al., 1985; Haynes, 1986; Falck and Fränkel, 1989; Nelson et al., 1993; Perillo et al., 1996; Toth and McNamara, 1999; Chadwick et al., 2001; De Almeida et al., 2002; Janson et al., 2003) analysed this outcome (Table 2 and Figure 2) as indicated by the following cephalometric measurements: Co–Go and Ar–Go. Significant heterogeneity of mandibular ramus height was found between different cephalometric measurements ($P = 0.012$) and for all studies (Table 3). Therefore, a random-effect model was used to estimate the overall effect. FR-2 was associated with a significant enhancement of mandibular ramus height (0.654 mm/year, 95 per cent CI 0.244–1.064; Figure 2) compared with untreated subjects.

**Association of quality score with the FR-2 effect.** Despite the relevant heterogeneity of the FR-2 effect among studies, an exploratory analysis of associations between quality score and FR-2 effect was made. A significant negative association (Figure 3) was found between the effect of the FR-2 and quality score for changes in mandibular total length ($P = 0.047$) and mandibular ramus height ($P = 0.005$) but not for mandibular body length ($P = 0.087$).

**Discussion**

The aim of the present investigation was to perform a systematic review and meta-analysis of studies on the mandibular skeletal effects of the FR-2 appliance in growing patients with a Class II malocclusion versus changes occurring in untreated Class II growing subjects. More specifically, the present study aimed to determine whether the FR-2 appliance had an impact on the dimensions of the mandible in treated patients versus untreated controls. While this impact was statistically significant, it can be considered to have a modest clinical effect. The average Class II malocclusion requires molar correction of approximately 4 to 6 mm (Johnston, 1986). Thus, treatment with an appliance such as the FR-2 that produces approximately 1 mm of supplementary mandibular growth per year constitutes a partial contribution to the expected Class II correction. The mandibular change is also less than the average growth deficiency in mandibular length during the circumpubertal period in Class II subjects when compared with those with a normal occlusion, which is about 3 mm (Stahl et al., 2008).
Table 1  Summarized data of the nine retrieved studies (M, male; F, female).

<table>
<thead>
<tr>
<th>Author</th>
<th>Origin</th>
<th>Number of patients treated with Fränkel-2</th>
<th>Type of selection of Fränkel patients</th>
<th>Initial mean age (years/months)</th>
<th>Appliance wear</th>
<th>Mean period of treatment (months)</th>
<th>Number of untreated controls</th>
<th>Type of selection of untreated controls</th>
<th>Initial mean age (years/months)</th>
<th>Mean observation period (months)</th>
<th>Quality score</th>
</tr>
</thead>
<tbody>
<tr>
<td>McNamara et al.</td>
<td>USA</td>
<td>51 (20M, 31 F) 49 (23M, 26 F)</td>
<td>Private practices</td>
<td>8.8 11.6</td>
<td>Not specified</td>
<td>23</td>
<td>36 (17M, 19 F)</td>
<td>University of Michigan, Elementary and Secondary Growth Study (UMESSGS)*</td>
<td>8.4 11</td>
<td>26 22</td>
<td>0</td>
</tr>
<tr>
<td>Haynes (1986)</td>
<td>UK</td>
<td>29</td>
<td>National Health Service</td>
<td>9.6 ± 0.8</td>
<td>12 h/day</td>
<td>36.5</td>
<td>29</td>
<td>National Health Service who would not accept FR treatment**</td>
<td>9.6 ± 0.8</td>
<td>36.8 2</td>
<td>2</td>
</tr>
<tr>
<td>Falck and Fränkel</td>
<td>Germany</td>
<td>60 (28M, 32F)</td>
<td>Department of Orthodontics School children referred to the Department of Orthodontics</td>
<td>8/0 ± 12</td>
<td>Not specified</td>
<td>14.7</td>
<td>50 (33M, 17F)</td>
<td>Department of Orthodontics School children referred to the Department of Orthodontics</td>
<td>7/11 ± 15</td>
<td>14.3 1</td>
<td>1</td>
</tr>
<tr>
<td>Nelson et al. (1993)</td>
<td>New Zealand</td>
<td>13 (7M, 6F)</td>
<td>University of Naples and private practitioner</td>
<td>11.7 ± 0.68</td>
<td>Minimum 14 h/day</td>
<td>18</td>
<td>17 (11M, 6F)</td>
<td>University of Naples**</td>
<td>11.5 ± 0.93</td>
<td>18 4</td>
<td>4</td>
</tr>
<tr>
<td>Perillo et al. (1996)</td>
<td>Italy</td>
<td>14 (6M, 8F)</td>
<td>University of Naples and private practitioner</td>
<td>8.7</td>
<td>At least 18 h/day</td>
<td>29</td>
<td>14 (7M, 7F)</td>
<td>University of Naples**</td>
<td>8.7 29</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Toth and McNamara (1999)</td>
<td>USA</td>
<td>40 (21M, 19F)</td>
<td>Private practices and Graduate Orthodontic Clinic</td>
<td>10/2</td>
<td>24 h/day</td>
<td>24</td>
<td>40</td>
<td>UMESSGS*</td>
<td>9/11 23</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>de Almeida et al. (2002)</td>
<td>Brazil</td>
<td>22 (11M, 11F)</td>
<td>Orthodontic graduate programme</td>
<td>9</td>
<td>24 h/day</td>
<td>17</td>
<td>22 (11M, 11F)</td>
<td>Files of the longitudinal growth study of the University of São Paulo Patients declining FR-2 treatment**</td>
<td>8/7 13</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Chadwick et al. (2001)</td>
<td>UK</td>
<td>70 (33M, 37F)</td>
<td>Patients seen at a single centre by a single operator</td>
<td>11/6</td>
<td>Not specified</td>
<td>19.92</td>
<td>68 (31M, 37F)</td>
<td>Files of the longitudinal growth study at the University of São Paulo**</td>
<td>10/10 22.32</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Janson et al. (2003)</td>
<td>Brazil</td>
<td>18 (10M, 8F)</td>
<td>Orthodontic Department graduate clinic</td>
<td>9/3</td>
<td>Not specified</td>
<td>28</td>
<td>23 (13M, 10F)</td>
<td>Files of the longitudinal growth study at the University of São Paulo**</td>
<td>9/3 28</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

*Matched for age. **Matched for age and gender.
This investigation also showed the limitations of the published papers: they were heterogeneous, mostly non-randomized and retrospective, of low to medium quality, and almost all with a poorly defined skeletal diagnosis.

Heterogeneity of the effects

A consistent heterogeneity among studies was found for all the considered outcomes (Table 3). The heterogeneity was assessed also within studies that used analogous cephalometric measurements. Several considerations can be postulated to explain this heterogeneity.

Differential diagnosis. Almost all the studies identified in this systematic review lacked initial differential diagnosis in order to identify Class II malocclusions associated with mandibular deficiency. The selection criteria were generic Class II malocclusion in five papers (McNamara et al., 1985; Falck and Fränkel, 1989; Nelson et al., 1993; Toth and McNamara, 1999; De Almeida et al., 2002), Class II division 1 malocclusions in two papers (Haynes, 1986; Janson et al., 2003), an overjet greater than 6 mm in one study (Chadwick et al., 2001), and mandibular deficiency associated with an aberrant muscular pattern in one paper (Perillo et al., 1996).

Cephalometric measurements. Part of the heterogeneity in the results can be explained by the localization of different cephalometric landmarks in different studies, which is one of the major confounding problems in cephalometrics. In this systematic review, both condylion (Co) and Ar were accepted as the posterior end point in measuring mandibular total length and ramus height.

It is reported that measurements with Ar as an end point, such as Ar–Pg or Ar–Gn, might give significant values for supplementary mandibular growth, without a corresponding increase in Co–Pg and Co–Gn, as Co is the most accurate end point landmark (Nelson et al., 1993). This meta-analysis showed significant changes for mandibular ramus height and total length using both Co and Ar points.

Age, skeletal maturation, and treatment variation. In two studies (McNamara et al., 1985; Toth and McNamara, 1999), treated and control subjects were matched for age, in four papers for both age and gender (Haynes, 1986; Perillo et al., 1996; Chadwick et al., 2001; Janson et al., 2003), while in three investigations the two groups were not matched.

The mean age at the start of treatment ranged from 8.0 to 11.7 years (Table 1). Although there was some overlap in the ages, these differences resulted in some problems when comparing studies.

Furthermore, growth does not occur at a constant rate, especially in young children. Even children of the same chronological age might not have equivalent skeletal
Table 3  Meta-analyses of selected studies.

<table>
<thead>
<tr>
<th></th>
<th>Number of studies</th>
<th>P value for heterogeneity*</th>
<th>Pooled estimate (mm/year)</th>
<th>95% confidence interval (mm/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mandibular body length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Go–Gn</td>
<td>Fixed effect (FE)</td>
<td>2</td>
<td>0.731</td>
<td>0.699</td>
</tr>
<tr>
<td>Go–M</td>
<td>FE</td>
<td>1</td>
<td>0.479</td>
<td>-0.010, 0.968</td>
</tr>
<tr>
<td>Go–Pg</td>
<td>FE</td>
<td>6</td>
<td>0.119</td>
<td>0.251</td>
</tr>
<tr>
<td>All studies</td>
<td>Random effect (RE)</td>
<td>9</td>
<td>0.088</td>
<td>0.400</td>
</tr>
<tr>
<td><strong>Mandibular total length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ar–Gn</td>
<td>FE</td>
<td>1</td>
<td>—</td>
<td>1.500</td>
</tr>
<tr>
<td>Ar–Pg</td>
<td>FE</td>
<td>1</td>
<td>—</td>
<td>0.761</td>
</tr>
<tr>
<td>Co–Gn</td>
<td>FE</td>
<td>6</td>
<td>0.003</td>
<td>1.016</td>
</tr>
<tr>
<td>Co–Pg</td>
<td>FE</td>
<td>2</td>
<td>0.018</td>
<td>1.593</td>
</tr>
<tr>
<td>All studies</td>
<td>RE</td>
<td>10</td>
<td>&lt;0.001</td>
<td>1.069</td>
</tr>
<tr>
<td><strong>Ramus height</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ar–Go</td>
<td>FE</td>
<td>3</td>
<td>0.914</td>
<td>0.197</td>
</tr>
<tr>
<td>Co–Go</td>
<td>FE</td>
<td>6</td>
<td>&lt;0.001</td>
<td>0.881</td>
</tr>
<tr>
<td>All studies</td>
<td>RE</td>
<td>9</td>
<td>&lt;0.001</td>
<td>0.654</td>
</tr>
</tbody>
</table>

*From Q test; a value <0.10 was considered significant.

**Figure 2**  Meta-analysis results. Mean change in mandibular body length, mandibular total length, and mandibular ramus height with the functional regulator-2 appliance treatment in Class II malocclusion growing patients versus untreated Class II growing subjects. The overall effect represents the pooled estimate of mean change. The size of each square is inversely proportional to the variance of the study estimate. (A) Very young and (B) young subjects from McNamara et al. (1985).

**Figure 3**  Meta-regression results. Association of quality score with mean change in mandibular body length, mandibular total length, and mandibular ramus height. The size of each square is inversely proportional to the variance of the study estimate.
the maturational stage that was presumably pubertal (age at start of treatment 11.6 years; McNamara et al., 1985) showed clinically relevant outcomes. Treatment duration differed widely among the studies, ranging from 14.7 to 36.5 months. Individual changes were annualized (i.e. expressed as change per year) to accommodate variation in intervals between radiographs (Table 1) and thus allow comparison with the data of other investigations. However, if change does not occur uniformly during the entire treatment time, this process can skew the analysis of treatment outcomes (Chen et al., 2002).

**Meta-analysis limitations**

This meta-analysis may have some limitations. First, the inclusion of published data alone may overestimate the treatment effects. This problem can be overcome using a funnel plot, which is a graphical method to detect publication bias (Lau et al., 2006). However, this method was not employed because simulation studies of funnel plots have found that bias may be incorrectly inferred if studies are heterogeneous (Schwartzet al., 2002; Terrin et al., 2003).

Second, the internal validity of a meta-analysis can only be as good as the quality of the studies reviewed. Nine studies were identified by the review process with 366 patients treated with the FR-2 and 320 controls. These studies were judged to be of low/medium quality. The reason for a low/medium-quality score is that some studies had some methodological limitations (method error not reported, lack of blinding in measurement, etc.). Only one study was an RCT. In the meta-regression, it was found that low-quality studies reported higher FR-2 effects than medium-quality studies. Nevertheless, quality is only one component of heterogeneity and has an uncertain role in explaining any treatment-effect differences. Thus, quality-related differences in the treatment effect should be treated as hypothesis-generating observations.

**Conclusions**

The findings of the present study were as follows:

1. The FR-2 appliance had a statistically significant effect on mandibular growth. Specifically, it appeared to have an effect on total mandibular length with a low-to-moderate clinical impact.

2. The heterogeneity of the FR-2 effects, the quality of studies, the differences in age, skeletal age, treatment duration, and the inconsistent initial diagnosis seem to overstate the benefits of the FR-2 appliance.

3. This investigation also serves to highlight the limitations of the reviewed papers on FR-2 therapy. This information can be used proactively as a platform to achieve more methodologically sound investigations.

**References**


Fränkel R 1966 The theoretical concept underlying the treatment with function correctors. Transactions of the European Orthodontic Society pp. 233–254


Fränkel R, Fränkel C 1989 Orofacial orthopedics with the function regulator. S Karger, Basel

Haynes S 1986 A cephalometric study of mandibular changes in modified function regulator (Fränkel) treatment. American Journal Orthodontics 88: 308–320


