ORIGINAL ARTICLE

Long-term (≥ 15 years) outcome quality after Class II:1 bionator or Herbst multibracket appliance treatment

A comparison

Niko Christian Bock[1] · Rebecca Jungbauer[2] · Ingrid Rudzki[3] · Peter Proff[2] · Sabine Ruf[1]

Abstract

Purpose To compare the long-term outcome quality (≥ 15 years) of Class II:1 treatment using either a bionator (BIO) or a Herbst–multibracket appliance (HMB).

Methods Patients who underwent functional treatment during the ideal treatment period for the respective approach (prepuberty vs. peak/postpeak) were assessed. Inclusion criteria were overjet ≥ 4 mm, skeletal Class II and availability of study casts from before, after and ≥ 15 years after treatment. The study casts were assessed using the Peer Assessment Rating (PAR) index and standard orthodontic cast measurements.

Results During treatment, PAR score, overjet and sagittal occlusal relationship improved significantly in all groups. Long-term, there was a significant increase of incisor irregularity in the upper (HMB) and lower (BIO) arch and a significant decrease of lower arch width 3–3 (BIO). PAR score, overjet, and sagittal occlusal relationship remained stable long-term. Intergroup comparisons revealed significant differences between the BIO and HMB groups in terms of lower arch width (6–6), upper and lower arch width (3+3/3–3) as well as sagittal molar relationship.

Conclusions The achieved improvement in PAR score, overjet, and sagittal occlusion remained comparably stable long-term in all groups. The long-term changes are probably a consequence of natural aging.

Keywords Long-term stability · Functional treatment · Peer Assessment Rating index · Overjet · Occlusal relationship

Qualität der Langzeitergebnisse (≥ 15 Jahre) nach Klasse-II:1-Behandlung mit Bionator bzw. Herbst-Multibracket-Apparatur

Ein Vergleich

Zusammenfassung

Ziel Der Vergleich der langfristigen Ergebnisqualität (≥ 15 Jahre) einer Klasse-II:1-Behandlung mit einem Bionator (BIO) oder einer Herbst-Multibracket-Apparatur (HMB).

✉ Dr. Rebecca Jungbauer
   rebecca.jungbauer@ukr.de

1   Department of Orthodontics, Justus Liebig University, Giessen, Germany

2   Department of Orthodontics, University Medical Centre Regensburg, Franz-Josef-Strauß-Allee 11, 93053 Regensburg, Germany

3   Faculty of Medicine, Ludwig Maximilian University, Munich, Germany

Methoden Patienten, die während des idealen Behandlungszeitraums für den jeweiligen Therapieansatz (vor der Pubertät bzw. Peak/Post-Peak) funktionskieferorthopädisch behandelt worden waren. Die Einschlusskriterien waren: Overjet ≥ 4 mm, skelettale Klasse II, Studienmodelle von vor, nach und ≥ 15 Jahre nach der Behandlung vorhanden. Die Studienmodelle wurden anhand des Peer Assessment Rating (PAR) index und kieferorthopädischer Standardmessungen ausgewertet.

Ergebnisse Während der Behandlung verbesserten sich PAR-Score, Overjet und sagittale Okklusionsbeziehung in allen Gruppen signifikant. Langfristig kam es zu einer signifikanten Zunahme des Engstandes der Schneidezähne im oberen (HMB) bzw. unteren (BIO) Zahnbogen und zu einer signifikanten Abnahme der Breite des unteren Zahnbogens 3 – 3 (BIO). PAR-Score, Overjet und sagittale Molarenrelation blieben langfristig stabil. Die Intergruppenvergleiche ergaben signifikante Unterschiede zwischen der BIO- und der HMB-Gruppe in Bezug auf die untere Zahnbogenbreite (6 – 6), die obere und untere Zahnbogenbreite (3 + 3/3 – 3) sowie die sagittale Molarenbeziehung.

Schlussfolgerungen Die erzielte Verbesserung des PAR-Scores, des Overjets und der sagittalen Molarenrelation blieb in allen Gruppen langfristig vergleichbar stabil. Die langfristigen Veränderungen sind wahrscheinlich eine Folge physiologischer Alterungsprozesse.

Schlüsselwörter Langzeitstabilität · Funktionelle kieferorthopädische Apparaturen · Peer Assessment Rating index · Overjet · Okklusion

Introduction

Class II:1 malocclusion is known to have a prevalence of 8.1–16.2% [32, 66, 67]. There are several therapeutic approaches that are sometimes discussed controversially; one of them is functional orthopaedics. Basically, functional appliances can be divided into two main groups: removable and fixed appliances.

Removable functional appliances have a very long history dating back into the 19th century [41]. Over the years, variable appliance designs have been described such as the activator [1] and bionator appliance [7] and their further developments and modifications, e.g. the Balter's bionator in the modification by Ascher [36, 38]. Most of these appliances are indicated predominantly in growing patients [10, 22] presenting a Class II:1 with a favourable facial morphology and favourable inherited growth potential of the mandible [34, 36].

The original basic idea behind the classical functional appliance is to eliminate all habits and parafunctions as early as possible so that the setting of the final inherited neutral (Class I) occlusion is not endangered during puberty. So, with respect to the literature, the ideal treatment (Tx) period for this specific functional therapeutic approach is prepuberty, in order to enable maximal mandibular growth [5, 39, 44, 68].

Fixed functional appliances were introduced slightly later [31] but their routine use started only a few decades ago [40, 48, 50, 53]. These appliances are mainly used in the permanent dentition and after the peak of pubertal growth. The effects of the Herbst appliance have been shown to be a combination of skeletal and dental changes in both the maxilla and the mandible [49], with the immediate Tx results being independent of the growth pattern [6, 11, 52, 53, 58, 59, 62], but without changing the inherited vertical growth pattern long-term [20, 52, 58].

For both removal and fixed functional appliances, controversial opinions exist in terms of their effectiveness and the stability of Tx results [17, 19, 42, 70]. While a certain amount of data exist on the immediate Tx effects and short-term stability [19, 42, 70], data regarding the long-term effects during adulthood are scarce for both Tx options. This applies also for comparisons between different appliances regarding the respective post-Tx changes and long-term outcome quality.

Aim

Therefore, the aim of the present investigation was to compare the long-term outcome quality in patients who underwent functional (removable vs. fixed) Class II Tx during the ideal Tx period recommended for the corresponding Tx approach (prepuberty vs. at peak/postpeak of the pubertal growth).

Subjects and methods

Orthodontic study casts and lateral cephalograms of patients who were treated with

- A Balter's bionator modified by Ascher [5] only, prepuberty (group BIO) and
- A Herbst appliance followed by a multibracket appliance peak/postpeak (group HMB) were retrospectively collected from the Departments of Orthodontics, University Medical Centre Munich and University of Giessen, both

Fig. 1 Flowchart exhibiting the effect of the exclusion criteria on the two original patient samples. *MBA* multibracket appliance, *n.a.* not applicable
Abb. 1 Flussdiagramm, das die Auswirkung der Ausschlusskriterien auf die beiden ursprünglichen Patientengruppen zeigt. *MBA* Multibracket-Apparatur, *n.a.* nicht zutreffend
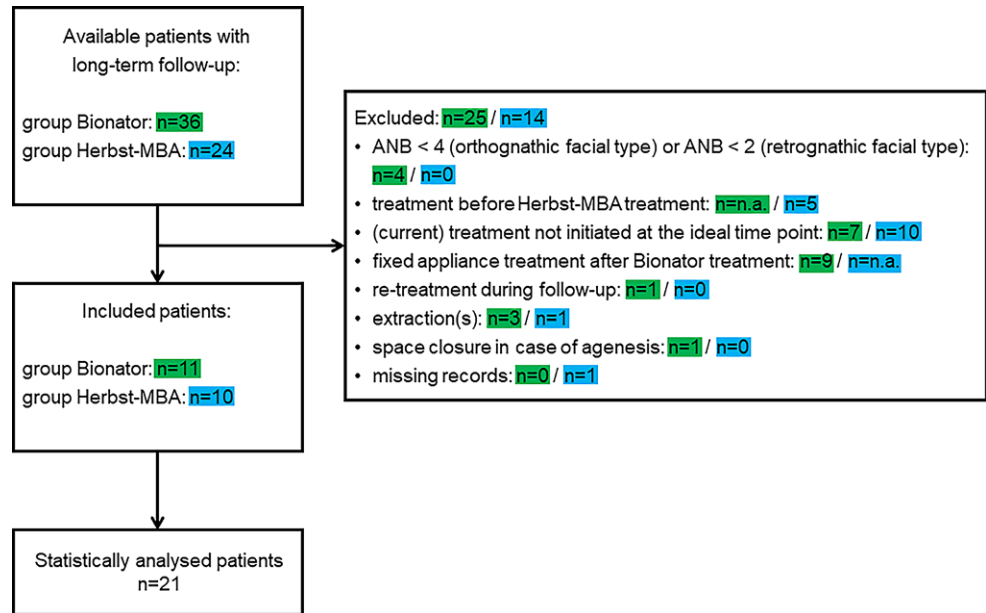
Germany. To be included in the investigation patients had to meet the following inclusion criteria (Fig. 1):

– Class II:1 with an overjet ≥ 4 mm before Tx,
– Skeletal Class II (ANB > 2° in retrognathic/4° in orthognathic/6° in prognathic facial types [63]),
– Class II Tx according to one of the protocols given above performed during the ideal skeletal maturity period recommended for the respective Tx approach,
– Hand–wrist radiographs from before Tx (T0) available,
– Lateral cephalograms from before Tx (T0) available,
– Study casts from before Tx (T0), after treatment (T1) and ≥ 15 years after treatment (T2) available and
– No agenesis of permanent teeth, extraction therapy or space opening/closure, no craniofacial syndromes.

All subjects participated in a long-term follow-up examination at one of the two departments of orthodontics: the group BIO at the University of Munich and the HMB group at the University of Giessen. Ethical approval was obtained separately at the two locations (Munich 77/97 and Giessen 146/13). All patients gave written informed consent. A brief overview of the Tx protocols is given in Fig. 2, while detailed descriptions can be found elsewhere for the BIO group [38, 57] and the HMB group [12].

Methods

Existing study models from all time points (T0, T1, T2) were assessed. The peer assessment rating (PAR) index [55, 56] was used in order to generate objective data on outcome quality; all ratings were performed by two certified/calibrated assessors according to the respective guidelines. In addition, the following variables were assessed:
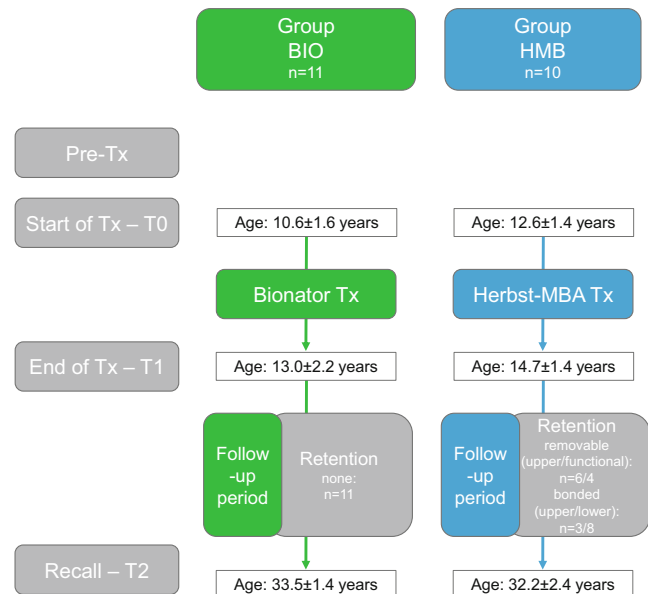

Fig. 2 Flowchart describing the observation time points and periods of the two groups. *BIO* bionator, *HMB* Herbst–multibracket appliance (MBA), *Tx* treatment
Abb. 2 Flussdiagramm zur Beschreibung der Beobachtungszeitpunkte und -zeiträume der beiden Gruppen. *BIO* Bionator, *HMB* Herbst–Multibracket-Apparatur (MBA), *Tx* Therapie

- Sagittal molar relationship (right, left) in cusp widths (cw),
- Sagittal canine relationship (right, left) in cusp widths (cw),
- Overjet in mm,
- Overbite in mm,

- Mandibular and maxillary incisor irregularity measured according to Little's irregularity index [45],
- Upper arch perimeter in mm,
- Lower arch perimeter in mm,
- Upper arch width (molar, canine) in mm and
- Lower arch width (molar, canine) in mm.

Visual ratings of the sagittal molar and canine relationships were performed to the nearest 0.25 cusp widths (cw) and classified as Class I, II or III. Linear measurements on dental casts were made using a digital calliper (HSL 246-15, Hammacher, Solingen, Germany) with a precision of 0.01 mm. Prior to the statistical evaluation all measurements were rounded to the first decimal.

Measurement error

To minimize the error of the method, all assessments were performed twice—once each by N. B. and R. J.; the mean value of both measurements was used as the final measurement value.

The method error (ME) was calculated (a) using the formula of Dahlberg [18]:

$ME = \sqrt{\frac{\sum d^2}{2n}}$, where d is the difference between two registrations and n is the sample size and (b) using the intraclass correlation coefficient (ICC, two-way mixed, absolute agreement) to determine interrater reliability. The data are shown in the supplementary Table 1.

Statistical assessment

Due to the explorative character of the study, no sample size calculation was performed. The arithmetic mean (M), standard deviation (SD), and median (MD) as well as interquartile range (IQR) were calculated for each variable. The data were tested regarding normal distribution using the Shapiro–Wilk test and visual assessment of the histograms showing a nonnormal distribution in more than 5%. Therefore, nonparametric Mann–Whitney U tests were applied for intergroup comparison long-term and Friedman tests followed by post hoc pairwise comparison for longitudinal changes within groups. For the post hoc test the Bonferroni p-value correction was applied. The following levels of significance were utilized: $p < 0.001$, $p < 0.01$ and $p < 0.05$; $p \geq 0.05$ was considered as not significant (n. s.). In case the global test indicated significant changes, but the post hoc test yielded p-values $p \geq 0.05$ due to the Bonferroni correction, the results of the global test were rated nonsignificant. The effect size (r) was calculated and interpreted in accordance with Cohen: 0.1 = small, 0.3 = medium 0.5 = large [16].

Results

Based on the inclusion criteria a total of 21 patients could be included (Fig. 1): BIO group: 11 (6 female/5 male), HMB group: 10 (7 female/3 male). Their baseline characteristics are presented in Fig. 3.

Intragroup comparison

Active Tx (T1–T0)

In both groups a significant reduction of the mean PAR score (14.8–23.2), overjet (3.1–5.0 mm) and overbite (1.2–2.4 mm) occurred. According to the PAR score categorisation, the ratio greatly improved/improved/worse or no different at T1 was 18%/73%/9% in the BIO group and 70%/30%/0% in the HMB group. Sagittal molar and canine relationships were significantly improved (0.2–0.7 cw), although sagittal canine relationship improvement in the BIO group was not significant. In the HMB group a significant reduction of incisor irregularity in both dental arches (maxilla: 3.8 mm; mandible: 1.5 mm) could be detected, whereas the corresponding reduction in the group BIO (maxilla: 2.9; mandible: 0.9 mm) was not significant (Fig. 3). The group BIO, however, was the only one to show a significant increase of upper arch width 3 + 3 distance.

Long-term follow-up (T2–T1)

None of the mean post-Tx changes of the variables PAR score (0.0–1.3), sagittal molar and canine relationships (–0.1 to 0.1 cw) or overjet (0.5–0.8 mm) were significant in any of the groups (Fig. 3). At T2, the PAR score categorisation (greatly improved/improved/worse or no different) revealed a ratio of 18%/82%/0% in the BIO and 50%/50%/0% in the HMB group.

The HMB group exhibited a significant average decrease in upper arch perimeter (1.4 mm), even though the decrease was similar in the BIO group (1.7 mm). All groups showed a significant decrease of the mean lower arch perimeter (1.4–3.2 mm). In the BIO group the mean lower arch width (3 – 3 distance) also decreased significantly (2.1 mm). A significant increase in mean mandibular incisor irregularity was seen in the BIO (2.2 mm) group (Fig. 3).

Intergroup comparison

Long-term follow-up (T2–T1)

For the majority of variables, no significant group difference was obvious. This, however, was not true for upper and lower arch width 3 + 3/3 – 3 distance, where a mean decrease of 1.1/2.1 mm was seen in the BIO group, while an

Variable	Group	n	T0 M ±SD	T0 MD [IQR]	T1 M ±SD	T1 MD [IQR]	T2 M ±SD	T2 MD [IQR]	T1-T0 M ±SD	T1-T0 MD [IQR]	T2-T1 M ±SD	T2-T1 MD [IQR]	Global test p-value	Post hoc tests adjusted p-value T1-T0 (r)	Post hoc tests adjusted p-value T2-T1 (r)
Total PAR score	BIO	11	24.7 ±10.2	23.5 [10.0]	9.9 ±6.9	10.0 [7.0]	11.1 ±7.0	10.0 [8.0]	-14.8 ±7.5	-13.0 [9.0]	1.3 ±1.6	1.5 [3.0]	<0.001	<0.001(0.89)	0.602
	HMB	10	26.3 ±7.8	27.0 [14.3]	3.1 ±3.4	2.0 [1.3]	3.1 ±2.2	2.8 [2.3]	-23.2 ±8.4	-23.5 [16.0]	0.0 ±2.1	0.5 [2.0]	<0.001	0.002 (0.78)	1.000
													0.072	p-value intergroup comparison	
Sagittal relationship M1, right/left pooled [cw]; positive value (+)=Class II, negative value (-)=Class III, 0=Class I	BIO	11	0.5 ±0.3	0.4 [0.5]	0.0 ±0.3	0.0 [0.2]	-0.1 ±0.4	0.0 [0.1]	-0.5 ±0.1	-0.5 [0.2]	-0.1 ±0.3	-0.1 [0.1]	<0.001	0.009 (0.68)	0.859
	HMB	10	0.7 ±0.2	0.8 [0.4]	0.0 ±0.1	0.0 [0.1]	0.1 ±0.1	0.0 [0.1]	-0.7 ±0.2	-0.8 [0.4]	0.1 ±0.1	0.1 [0.1]	<0.001	<0.001 (0.90)	0.221
													0.043 (0.45)	p-value intergroup comparison (r)	
Sagittal relationship C, right/left pooled [cw]; positive value (+)=Class II, negative value (-)=Class III, 0=Class I	BIO	10	0.4 ±0.3	0.5 [0.2]	0.2 ±1.4	0.2 [0.1]	0.3 ±0.2	0.2 [0.2]	-0.2 ±0.2	-0.2 [0.3]	0.0 ±0.2	0.0 [0.1]	0.023	0.539	0.539
	HMB	10	0.8 ±0.1	0.8 [0.2]	0.2 ±0.1	0.2 [0.1]	0.2 ±0.1	0.2 [0.2]	-0.6 ±0.1	-0.6 [0.2]	0.0 ±0.2	0.0 [0.2]	0.001	0.002 (0.75)	1.000
													0.557	p-value intergroup comparison (r)	
Overjet [mm]	BIO	11	6.1 ±1.9	5.7 [2.6]	3.2 ±1.7	2.7 [1.4]	3.7 ±1.6	2.8 [1.9]	-3.1 ±1.4	-2.7 [2.0]	0.5 ±1.3	0.3 [1.3]	<0.001	0.001 (0.77)	1.000
	HMB	10	7.2 ±2.4	6.7 [3.6]	2.3 ±0.4	2.3 [0.5]	3.1 ±0.4	3.1 [0.7]	-5.0 ±2.1	-4.4 [3.0]	0.8 ±0.6	1.1 [1.0]	<0.001	<0.001 0.95)	0.221
													0.132	p-value intergroup comparison (r)	
Overbite [mm]	BIO	11	3.2 ±1.0	3.2 [1.9]	2.1 ±1.3	1.6 [1.7]	3.2 ±1.0	3.0 [1.2]	-1.2 ±1.0	-1.3 [1.3]	1.2 ±1.0	1.1 [1.5]	0.012	0.017 (0.59)	0.057
	HMB	10	3.9 ±1.5	3.8 [2.5]	1.4 ±0.4	1.3 [0.6]	2.5 ±1.1	2.4 [1.5]	-2.4 ±1.5	-2.1 [2.8]	1.1 ±1.1	0.7 [1.0]	0.001	<0.001 (0.85)	0.076
													0.654	p-value intergroup comparison (r)	
Upper arch width 6+6 distance [mm]	BIO	11	45.5 ±2.4	45.1 [3.7]	47.8 ±2.6	48.6 [4.5]	47.1 ±2.4	47.1 [3.8]	2.3 ±2.5	2.6 [5.1]	-0.7 ±2.3	-1.3 [2.1]	0.060	NA	NA
	HMB	10	43.2 ±2.3	43.7 [4.4]	44.7 ±1.9	44.8 [3.2]	44.5 ±2.3	44.9 [3.6]	1.4 ±1.9	1.9 [3.5]	-0.2 ±0.8	-0.4 [1.4]	0.150	NA	NA
													0.173	p-value intergroup comparison (r)	
Lower arch width 6-6 distance [mm]	BIO	11	47.1 ±2.3	47.3 [3.2]	49.3 ±2.1	49.3 [3.0]	48.1 ±1.9	47.8 [3.0]	2.2 ±2.9	2.1 [4.4]	-1.2 ±2.3	-1.4 [2.1]	0.066	NA	NA
	HMB	10	45.3 ±2.5	46.4 [4.2]	46.0 ±1.6	46.1 [2.3]	46.3 ±2.4	46.2 [3.3]	0.7 ±2.1	0.4 [3.9]	0.3 ±1.5	0.1 [2.0]	0.905	NA	NA
													0.043 (0.45)	p-value intergroup comparison (r)	
Upper arch width 3+3 distance [mm]	BIO	7	31.0 ±3.2	29.8 [5.9]	34.3 ±2.7	34.7 [5.8]	33.2 ±2.1	33.1 [1.8]	3.4 ±1.4	2.7 [4.0]	-1.1 ±1.4	-1.3 [1.3]	0.028	0.023 (0.71)	0.544
	HMB	9	31.4 ±2.4	32.1 [3.1]	32.0 ±1.7	32.1 [1.3]	32.5 ±2.2	33.4 [3.0]	0.6 ±1.5	0.8 [2.5]	0.5 ±1.2	0.9 [2.0]	0.045	1.000	0.178
													0.017 (0.52)	p-value intergroup comparison (r)	
Lower arch width 3-3 distance [mm]	BIO	10	26.9 ±2.1	26.2 [3.7]	26.9 ±2.1	26.1 [4.1]	24.8 ±2.2	24.7 [4.5]	-0.1 ±1.5	0.3 [2.2]	-2.1 ±1.2	-1.9 [2.5]	0.001	1.000	0.001 (0.80)
	HMB	9	25.8 ±1.9	25.3 [2.5]	25.7 ±1.5	26.6 [2.4]	24.9 ±1.6	25.4 [3.2]	0.0 ±1.9	-0.6 [2.0]	-0.8 ±0.9	-1.1 [1.5]	0.097	NA	NA
													0.035 (0.46)	p-value intergroup comparison (r)	
Upper arch perimeter [mm]	BIO	11	75.7 ±5.1	73.7 [6.4]	77.2 ±4.5	76.4 [8.1]	75.4 ±2.9	75.2 [6.3]	1.5 ±2.0	0.6 [3.9]	-1.7 ±2.0	-1.2 [3.2]	0.148	NA	NA
	HMB	10	74.5 ±4.4	74.4 [4.7]	73.8 ±3.6	73.6 [3.5]	72.4 ±3.5	72.5 [4.3]	-0.8 ±2.7	-0.2 [5.1]	-1.4 ±1.1	-1.8 [1.2]	0.007	1.000	0.042 (0.55)
													0.918	p-value intergroup comparison (r)	
Lower arch perimeter [mm]	BIO	11	67.2 ±5.0	65.9 [9.2]	66.9 ±3.2	67.3 [6.5]	63.7 ±2.9	62.8 [5.5]	-0.3 ±5.4	-1.1 [1.3]	-3.2 ±2.4	-2.7 [3.0]	0.001	1.000	0.009 (0.64)
	HMB	10	63.6 ±3.8	63.8 [6.5]	63.9 ±3.5	64.5 [3.4]	62.5 ±3.4	63.0 [4.3]	0.3 ±.2.3	-0.4 [3.5]	-1.4 ±0.9	-1.6 [1.1]	0.008	1.000	0.02 (0.60)
													0.085	p-value intergroup comparison (r)	
Maxillary incisor irregularity [mm]	BIO	11	5.6 ±2.6	5.7 [5.3]	2.8 ±1.8	2.8 [2.8]	3.0 ±1.8	2.9 [2.6]	-2.9 ±2.4	-2.1 [2.8]	0.3 ±1.1	0.2 [1.4]	0.035	0.057	1.000
	HMB	10	5.1 ±2.6	4.8 [4.6]	1.3 ±0.9	1.0 [1.6]	1.9 ±0.7	1.9 [1.1]	-3.8 ±3.1	-3.7 [4.6]	0.6 ±1.0	0.5 [1.8]	0.006	0.005 (0.70)	1.000
													0.468	p-value intergroup comparison (r)	
Mandibular incisor irregularity [mm]	BIO	11	2.9 ±1.7	2.7 [1.8]	2.0 ±1.3	1.7 [1.5]	4.2 ±1.9	4.1 [2.6]	-0.9 ±2.1	-0.5 [1.7]	2.2 ±1.5	2.2 [2.9]	0.010	0.210	0.009 (0.64)
	HMB	10	2.2 ±1.9	1.7 [1.8]	0.8 ±0.6	0.8 [0.6]	2.2 ±1.7	2.0 [2.6]	-1.5 ±2.3	-0.9 [2.0]	1.4 ±1.6	1.0 [1.9]	0.025	0.042 (0.55)	0.076
													0.223	p-value intergroup comparison (r)	

Fig. 3 Descriptive and analytical outcome of the intra- and intergroup comparison during the different time periods. Global changes were tested with nonparametric Friedman's two-way analysis of variance by ranks for intragroup changes (rows) followed by post hoc pairwise comparisons and with Mann–Whitney U tests for intergroup comparison long-term (columns), r is the effect size (< 0.3 small effect, 0.3–0.5 medium effect, > 0.5 large effect). *T0* before treatment, *T1* after treatment, *T2* ≥ 15 years after treatment, *n* numbers analysed, *M* mean, *SD* standard deviation, *MD* median, *IQR* interquartile range, *M1* first molars, *C* canines, *cw* cusp widths, *NA* not applicable

Abb. 3 Deskriptive und analytische Ergebnisse des Intra- und Intergruppenvergleichs über die verschiedenen Zeiträume. Globale Veränderungen wurden mit nichtparametrischen Friedman-Tests hinsichtlich gruppeninterner Veränderungen (Zeilen), gefolgt von paarweisen Post-Hoc-Vergleichen und mit Mann-Whitney-U-Tests für den Langzeitvergleich zwischen den Gruppen (Spalten) getestet, r ist die Effektstärke (< 0,3 schwacher Effekt, 0,3–0,5 mittlerer Effekt, > 0,5 starker Effekt). *T0* vor der Behandlung, *T1* nach der Behandlung, *T2* ≥ 15 Jahre nach der Behandlung, *n* Anzahl der analysierten Patienten, *M* Mittelwert, *SD* Standardabweichung, *MD* Median, *IQR* Interquartilsbereich, *M1* erste Molaren, *C* Eckzähne, *cw* Prämolarenbreiten, *NA* nicht zutreffend

increase of 0.5/decrease of 0.8 mm occurred in the HMB group; and the lower arch width 6–6 exhibited a decrease in the BIO (1.2 mm) and an increase in the HMB group (0.3 mm). The sagittal molar relationship showed also a significant difference with a further decrease towards a slight Class III occlusion (0.1 cw) in the BIO and a slight increase towards Class II occlusion (0.1 cw; Fig. 3).

Measurement error

For the PAR index the interrater reliability was good or excellent (0.8–1). The method error calculated by using Dahlberg's formula ranged from 0.1 to 1.74. The detailed results are presented in supplementary Table 1.

Discussion

The present retrospective investigation is the first to compare long-term (≥ 15 years) outcome and stability after functional Class II Tx using different approaches. For natural reasons, the two groups differed in terms of age and skeletal maturity, as Tx was mandatory to have taken place during the ideal skeletal maturity period recommended for the respective Tx approach (prepuberty and peak/postpeak, respectively). Consequently, intergroup comparison was only performed for the long-term interval. Due to this inclusion criterion and the necessity of study casts from ≥ 15 years after Tx, the sample size of the individual groups was rather small. Nevertheless, the Tx approach was identical within each of the groups. The investigation is based on study casts which were evaluated by two examiners. For most variables the interrater reliability can be considered as very good.

Looking at the results in terms of outcome quality (total PAR score), both groups showed significant improvement during Tx (T0–T1). Nevertheless, the post-Tx PAR score remains distinctly higher in the BIO group than in the HMB group. This difference can be explained by the fact that the patients in the BIO group did not undergo additional multibracket appliance Tx for finishing after functional appliance therapy. Franscisconi et al. reported a mean reduction of the PAR score by 24.7 ± 6.6 as a result of Tx with a bionator and a multibracket appliance [25], which is a similar reduction as in the HMB group. Nevertheless, both groups showed only small mean changes of total the PAR score (range 0.0–1.3) during the long-term post-Tx observation period (T2–T1), which is in accordance with previous studies [12, 25, 38].

In terms of the sagittal and vertical occlusal variables both groups showed similar post-Tx values (T1). On average, the molar relationship was Class I and the canine relationship was up to 0.25 cw Class II which can still be considered normal [2].

Post-Tx mean overjet and overbite values were within a "physiological" range as well (2.3–3.2 mm and 1.4–2.1 mm, respectively). All parameters (molar relationship, overjet and overbite) showed rather good stability during the long-term post-Tx observation period (T1–T2) with the final values being similar as in untreated subjects without orthodontic Tx need during adolescence [12, 29]. Although there was a significant group difference long-term considering molar relationship with minimal tendency towards Class II occlusion in the HMB group (+0.1 cw) and towards Class III occlusion (−0.1 cw) in the BIO group, the amount of the change is not relevant from a clinical perspective. In the BIO group, only patients with favourable inherited growth potential of the mandible were included, which might explain the tendency towards a Class III occlusion which naturally takes place during the residual growth phase. Thus, in terms of stability of the achieved occlusal outcome it does not seem to make a difference long-term whether to treat early during the mixed dentition using a removable functional appliance or to do so at a later stage in the permanent dentition with a fixed functional appliance.

Looking at the transverse dimension, for upper and lower arch width 6±6 distance none of the groups exhibited a significant change, neither during Tx nor during the long-term observation period. The bionator group was the only group that was not treated with any kind of fixed appliances and the gain in the posterior transverse distance is very likely only due to development and natural growth [46]. The minor, nonsignificant, changes during the long-term interval are in accordance with the findings by Bondevik et al. [13]. Nevertheless, there was a significant group difference long-term for the lower intermolar width (6–6) with a decrease only in the BIO group. Henrikson et al. reported a signif-

icantly greater decrease of lower intermolar width in male compared to female between the age of 13.6 and 31.1 years [30]. In the present study the proportional relation female/male in the BIO group was 6/5, whereas in the HMB group 7/3, which might be the reason for the greater decrease in the BIO compared to the HMB group. Furthermore, the BIO group was without any appliance or retention for a longer time interval compared to the HMB group (20.5 and 17.5 years).

For upper arch width 3 + 3 distance, however, a significant increase occurred during T1–T0 in the BIO group, probably due to natural growth and development of the dental arches as they were treated prepuberty [46, 47]. And this variable showed a significant difference between the groups BIO and HMB (T2–T1) with a decrease only in the BIO group. In the HMB group, however, 3 patients had bonded upper retainers and in addition all patients wore either an upper removable or a functional appliance, whereas none of the BIO patients wore any appliance for retention after the end of Tx. Regarding lower arch width 3 – 3, a significant change (decrease) was only determined for the BIO group during the long-term observation period. It is well known from literature that lower arch width 3 – 3 remains stable after the eruption of the permanent canines during the growth period and decreases over the years which is very likely related to aging processes [13, 15, 21, 46, 64]. In contrast to the BIO group most of the subjects in the Herbst group had been given a bonded cuspid retainer at T1, which is probably the reason for the significant long-term group difference. Nevertheless, a reduction of the 3 – 3 distance also occurred in the Herbst group. It remains unclear whether the reduction of lower arch width 3 – 3 in the HMB group was caused by failures of the retainers, removal of the bonded retainers or unwanted tooth movement in spite of retainers.

Looking at the lower arch perimeter, a significant mean decrease (range 1.4–3.5 mm) occurred in both groups during the long-term observation period. In the HMB group, however, the amount was much smaller, which might be explained by the fact that—in comparison to the other group—many subjects (60%) still wore a bonded retainer at the time of the recall. Regarding the upper arch perimeter, an average decrease of 1.4–1.7 mm was determined in both groups, but was significant in the HMB group only. The physiological age dependent reduction of both the upper and lower arch perimeters long-term is in accordance with the literature [9, 15, 28, 69].

Maxillary and mandibular incisor irregularity showed a significant mean decrease during T0–T1 in the HMB group. It can be assumed that due to no fixed appliance Tx, changes in the bionator group the reduction was less pronounced. Nevertheless, especially the incisor irregularity in the upper arch improved, which might be owing to the gain of space due to the significant increase in the upper arch with 3 + 3. During the long-term observation, a significant mean increase occurred in the BIO group. The group HMB showed slightly less increase, which might also be attributed to the fact that the majority of subjects in this group still had a bonded retainer in place at the time of the recall. Again, the increase of incisor irregularity in the mandible related to aging is well known [13, 15, 51, 64] and is very likely associated with the decrease in arch perimeter and intercanine width [69]. But of course, it might also be kind of a side effect related to incisor proclination during active Tx—particularly in the HMB group; on the other hand, no such long-term effect was seen when investigating the effects of Herbst Tx on the mandibular incisor segment [27].

In the literature, it is often recommended to treat patients with functional appliances during pubertal growth [3, 8, 23, 33, 54]. The original therapeutic approach of removable Class II functional appliances, however, was less to directly stimulate mandibular growth but primarily to enable the full development of the inherited pubertal growth potential by eliminating all habits and parafunctions as early as possible, namely in the prepuberty phase [4, 5, 7, 35, 37, 39]. Therefore, for the present investigation, the ideal period for bionator Tx was considered as prepuberty. Compliance with removable functional appliances can be limited resulting in prolonged Tx or even impeding unsuccessful Tx [14, 65], but seems to be substantially better during prepuberty [24, 43]. In contrast, fixed functional appliances are mainly used during later stages of dental and skeletal development, respectively. According to the literature the ideal period for Herbst Tx is in the permanent dentition at or just after the peak of pubertal growth [61]. Even if mandibular growth stimulation has been shown to be possible even in young adults [60], the largest amount of sagittal condylar growth was demonstrated in subjects treated around the peak of the pubertal growth period [26].

Retrospective long-term follow-up studies have several limitations to consider. As it is very often impossible to locate and include all patients and not every potential patient agrees to take part in the follow-up, this kind of study is inherently biased but also the number of subjects very limited. Furthermore, patients that were satisfied with the former Tx might be more likely to be willing to participate as those that were not satisfied with the Tx result. The different groups were treated by different orthodontists, which might be considered as further possible bias. Due to the retrospective design and limited sample size there is of course a limited power and generalizability; in addition, no untreated control group with similar long-term data was available. Another limitation is that long-term cephalometric images were not available for ethical reasons. Therefore, only dental parameters could be evaluated, which surely limits the validity of the Tx effectiveness within the Tx groups.

The aim of this study was to evaluate and compare the long-term stability of the different Tx approaches and not the comparison their Tx effectiveness, nor to investigate their equivalence. BIO or HMB treatments are indicated at different stages of growth and therefore are not comparable or interchangeable. For each patient, an individual decision must be made as to which treatment approach is best, depending on the skeletal, not the dental, morphology, and growth stage, as well as compliance.

Although the results of the present study need to be interpreted with caution, the data are unique.

Conclusion

Both functional treatment approaches (bionator treatment, Herbst–multibracket appliance treatment) resulted in substantial reduction in Peer Assessment Rating (PAR) score, overjet and sagittal occlusion. Long-term, these treatment results exhibited good stability in all groups. Long-term changes that have occurred in the dentition are most likely a consequence of the aging process.

Declarations

References

1. Andresen V, Häupl K, Petrik L (1953) Funktionskieferorthopädie. JA Barth Verlag, München
2. Andrews LF (1989) Straight wire. The concept and appliance. Wells, San Diego
3. Antunes CF, Bigliazzi R, Bertoz FA et al (2013) Morphometric analysis of treatment effects of the Balters bionator in growing Class II patients. Angle Orthod 83(3):455–459. https://doi.org/10.2319/072512-605.1
4. Ascher F (1971) Kontrollierte Ergebnisse der Rückbissbehandlung mit funktionskieferorthopädischen Geräten. Fortschr Kieferorthop 32(2):149–159. https://doi.org/10.1007/BF02170640
5. Ascher F (1968) Praktische Kieferorthopädie. Urban & Schwarzenberg, München, Berlin, Wien
6. Atresh A, Cevidanes LHS, Yatabe M et al (2018) Three-dimensional treatment outcomes in Class II patients with different vertical facial patterns treated with the Herbst appliance. Am J Orthod Dentofacial Orthop 154(2):238–248.e1. https://doi.org/10.1016/j.ajodo.2017.11.037
7. Balters W (1960) Ergebnis der gesteuerten Selbstheilung von kieferorthopädischen Anomalien. Dtsch Zahnärztl Z 15:241
8. Bigliazzi R, Franchi L, Bertoz AP et al (2015) Morphometric analysis of long-term dentoskeletal effects induced by treatment with Balters bionator. Angle Orthod 85(5):790–798. https://doi.org/10.2319/080714-554.1
9. Bishara SE, Jakobsen JR, Treder J et al (1998) Arch length changes from 6 weeks to 45 years. Angle Orthod 68(1):69–74. https://doi.org/10.1043/0003-3219(1998)068⟨0069:ALCFWT⟩2.3.CO;2
10. Bishara SE, Ziaja RR (1989) Functional appliances: a review. Am J Orthod Dentofacial Orthop 95(3):250–258. https://doi.org/10.1016/0889-5406(89)90055-3
11. Bock N, Pancherz H (2006) Herbst treatment of Class II division 1 malocclusions in retrognathic and prognathic facial types. Angle Orthod 76(6):930–941. https://doi.org/10.2319/100605-352
12. Bock NC, Saffar M, Hudel H et al (2018) Long-term (≥15 years) post-treatment changes and outcome quality after Class II:1 treatment in comparison to untreated Class I controls. Eur J Orthod 40(2):206–213. https://doi.org/10.1093/ejo/cjx051
13. Bondevik O (2015) A longitudinal study of dental arches and occlusal changes in adults from 22 to 33, and 33 to 43 years of age. J Orofac Orthop 76(1):79–89. https://doi.org/10.1007/s00056-014-0230-z
14. von Bremen J, Lorenz N, Ludwig B et al (2018) Increased BMI in children-an indicator for less compliance during orthodontic treatment with removable appliances. Eur J Orthod 40(4):350–355. https://doi.org/10.1093/ejo/cjy007
15. Carter GA, McNamara JA (1998) Longitudinal dental arch changes in adults. Am J Orthod Dentofacial Orthop 114(1):88–99
16. Cohen J (2013) Statistical power analysis for the behavioral sciences, 2nd edn. Taylor and Francis, Hoboken
17. Cozza P, Baccetti T, Franchi L et al (2006) Mandibular changes produced by functional appliances in Class II malocclusion. A systematic review. Am J Orthod Dentofacial Orthop 129(5):599.e1–599.12. https://doi.org/10.1016/j.ajodo.2005.11.010
18. Dahlberg G (1940) Statistical methods for medical and biological students. Br Med J 2(4158):358–359
19. D'Antò V, Bucci R, Franchi L et al (2015) Class II functional orthopaedic treatment. A systematic review of systematic reviews. J Oral Rehabil 42(8):624–642. https://doi.org/10.1111/joor.12295

20. Deen E, Woods MG (2015) Effects of the Herbst appliance in growing orthodontic patients with different underlying vertical patterns. Aust Orthod J 31(1):59–68

21. DeKock WH (1972) Dental arch depth and width studied longitudinally from 12 years of age to adulthood. Am J Orthod 62(1):56–66

22. DiBiase AT, Cobourne MT, Lee RT (2015) The use of functional appliances in contemporary orthodontic practice. Br Dent J 218(3):123–128. https://doi.org/10.1038/sj.bdj.2015.44

23. Faltin KJ, Faltin RM, Baccetti T et al (2003) Long-term effectiveness and treatment timing for Bionator therapy. Angle Orthod 73(3):221–230. https://doi.org/10.1043/0003-3219(2003)073〈0221:LEATTF〉2.0.CO;2

24. Fleischer-Peters A, Scholz U (1985) Psychologie und Psychosomatik in der Kieferorthopädie. Carl Hanser, München

25. Franchisconi MF, Henriques JFC, Janson G et al (2013) Stability of Class II treatment with the Bionator followed by fixed appliances. J Appl Oral Sci 21(6):547–553. https://doi.org/10.1590/1679-775720130002

26. Hägg U, Pancherz H (1988) Dentofacial orthopaedics in relation to chronological age, growth period and skeletal development. An analysis of 72 male patients with Class II division 1 malocclusion treated with the Herbst appliance. Eur J Orthod 10(3):169–176. https://doi.org/10.1093/ejo/10.3.169

27. Hansen K, Koutsonas TG, Pancherz H (1997) Long-term effects of Herbst treatment on the mandibular incisor segment: a cephalometric and biometric investigation. Am J Orthod Dentofacial Orthop 112(1):92–103. https://doi.org/10.1016/S0889-5406(97)70279-8

28. Harris EF (1997) A longitudinal study of arch size and form in untreated adults. Am J Orthod Dentofacial Orthop 111(4):419–427

29. Heikinheimo K, Nyström M, Heikinheimo T et al (2012) Dental arch width, overbite, and overjet in a Finnish population with normal occlusion between the ages of 7 and 32 years. Eur J Orthod 34(4):418–426. https://doi.org/10.1093/ejo/cjr025

30. Henrikson J, Persson M, Thilander B (2001) Long-term stability of dental arch form in normal occlusion from 13 to 31 years of age. Eur J Orthod 23(1):51–61. https://doi.org/10.1093/ejo/23.1.51

31. Herbst E (1910) Atlas und Grundriß der Zahnärztlichen Orthopädie. J.F. Lehmanns, München

32. Ingervall B, Mohlin B, Thilander B (1978) Prevalence and awareness of malocclusion in Swedish men. Community Dent Oral Epidemiol 6(6):308–314. https://doi.org/10.1111/j.1600-0528.1978.tb01172.x

33. Jacobs T, Sawaengkit P (2002) National institute of dental and craniofacial research efficacy trials of bionator Class II treatment: a review. Angle Orthod 72(6):571–575. https://doi.org/10.1043/0003-3219(2002)072〈0571:NIODAC〉2.0.CO;2

34. Janson I (1977) A cephalometric study of the efficiency of the Bionator. Transactions of the European orthodontic society, pp 283–297

35. Janson I (1977) Zeitpunkt der Bionatorbehandlung in Abhängigkeit vom Wachstum. Fortschr Kieferorthop 38(4):435–451. https://doi.org/10.1007/BF02163222

36. Janson I (1987) Morphologische Kriterien für die Indikation einer Behandlung mit der Bionatormodifikation nach Ascher. Fortschr Kieferorthop 48(2):71–86. https://doi.org/10.1007/BF02175610

37. Janson I (1983) Skeletal and dento-alveolar changes in patients treated with a Bionator during puberty and prepuberty growth. In: McNamara JA Jr., Ribbons KA, Howe R (eds) Moyers symposium proceedings: "the clinical alteration of the growing face". USA, Ann Arbor, Michigan

38. Jungbauer R, Koretsi V, Proff P et al (2020) Twenty-year follow-up of functional treatment with a bionator appliance: a retrospective dental cast analysis. Angle Orthod 90(2):209–215. https://doi.org/10.2319/042419-292.1

39. Kantorowicz A (1959) Von der Frühbehandlung zur Vorbeugung der erworbenen kieferorthopädischen Anomalien. Dtsch Zahnärztl Z 14:227–238

40. Keim RG, Gottlieb EL, Nelson AH et al (2002) 2002 JCO study of orthodontic diagnosis and treatment procedures. Part 1. Results and trends. J Clin Orthod 36(10):553–568

41. Kingsley NW (1880) A treatise on oral deformities as a branch of mechanical surgery. D. Appleton and Company, New York

42. Koretsi V, Zymperdikas VF, Papageorgiou SN et al (2015) Treatment effects of removable functional appliances in patients with Class II malocclusion. A systematic review and meta-analysis. Eur J Orthod 37(4):418–434. https://doi.org/10.1093/ejo/cju071

43. Kressel D (1999) Fragebogenstudie zum Compliance in der Kieferorthopädie

44. Levrini A (2003) The masters of functional orthodontics. Quintessenz, Berlin

45. Little RM (1975) The irregularity index. A quantitative score of mandibular anterior alignment. Am J Orthod 68(5):554–563

46. Moorrees CF, Grøn A-M, Lebret LM et al (1969) Growth studies of the dentition: a review. Am J Orthod 55(6):600–616. https://doi.org/10.1016/0002-9416(69)90037-2

47. Moorrees CFA (1959) The dentition of the growing child. Harvard University Press, Cambridge

48. Pancherz H (1979) Treatment of Class II malocclusions by jumping the bite with the Herbst appliance. A cephalometric investigation. Am J Orthod 76(4):423–442. https://doi.org/10.1016/0002-9416(79)90227-6

49. Pancherz H (1982) The mechanism of Class II correction in Herbst appliance treatment. Am J Orthod 82(2):104–113. https://doi.org/10.1016/0002-9416(82)90489-4

50. Pancherz H (1985) The Herbst appliance—its biologic effects and clinical use. Am J Orthod 87(1):1–20. https://doi.org/10.1016/0002-9416(85)90169-1

51. Pancherz H, Bjerklin K, Lindskog-Stokland B et al (2014) Thirty-two-year follow-up study of Herbst therapy: a biometric dental cast analysis. Am J Orthod Dentofacial Orthop 145(1):15–27. https://doi.org/10.1016/j.ajodo.2013.09.012

52. Pancherz H, Fackel U (1990) The skeletofacial growth pattern pre- and post-dentofacial orthopaedics. A long-term study of Class II malocclusions treated with the Herbst appliance. Eur J Orthod 12(2):209–218. https://doi.org/10.1093/ejo/12.2.209

53. Pancherz H, Ruf S (2008) The Herbst appliance. Research based clinical management. Quintessence, London

54. Pavoni C, Lombardo EC, Lione R et al (2018) Treatment timing for functional jaw orthopaedics followed by fixed appliances: a controlled long-term study. Eur J Orthod 40(4):430–436. https://doi.org/10.1093/ejo/cjx078

55. Richmond S, Shaw WC, O'Brien KD et al (1992) The development of the PAR index (peer assessment rating). Reliability and validity. Eur J Orthod 14(2):125–139

56. Richmond S, Shaw WC, Roberts CT et al (1992) The PAR index (peer assessment rating). Methods to determine outcome of orthodontic treatment in terms of improvement and standards. Eur J Orthod 14(3):180–187

57. Rudzki-Janson I, Noachtar R (1998) Functional appliance therapy with the bionator. Semin Orthod 4(1):33–45. https://doi.org/10.1016/S1073-8746(98)80040-6

58. Ruf S, Pancherz H (1996) The effect of Herbst appliance treatment on the mandibular plane angle: a cephalometric roentgenographic study. Am J Orthod Dentofacial Orthop 110(2):225–229. https://doi.org/10.1016/S0889-5406(96)70113-0

59. Ruf S, Pancherz H (1997) The mechanism of Class II correction during Herbst therapy in relation to the vertical jaw base relationship: a cephalometric roentgenographic study. Angle Orthod 67(4):271–276. https://doi.org/10.1043/0003-3219(1997)067〈0271:TMOCIC〉2.3.CO;2

60. Ruf S, Pancherz H (1998) Temporomandibular joint growth adaptation in Herbst treatment: a prospective magnetic resonance

imaging and cephalometric roentgenographic study. Eur J Orthod 20(4):375–388

61. Ruf S, Pancherz H (2003) When is the ideal period for Herbst therapy—early or late? Semin Orthod 9(1):47–56. https://doi.org/10.1053/sodo.2003.34024

62. Schiavoni R, Grenga V, Macri V (1992) Treatment of Class II high angle malocclusions with the Herbst appliance: a cephalometric investigation. Am J Orthod Dentofacial Orthop 102(5):393–409. https://doi.org/10.1016/S0889-5406(05)81186-2

63. Segner D, Hasund A (2003) Individualisierte Kephalometrie, 4th edn. Segner, Hamburg

64. Sinclair PM, Little RM (1985) Dentofacial maturation of untreated normals. Am J Orthod 88(2):146–156

65. Stefanovic NL, Uhac M, Brumini M et al (2021) Predictors of patient compliance during Class II division 1 malocclusion functional orthodontic treatment. Angle Orthod 91(4):502–508. https://doi.org/10.2319/090820-780.1

66. Tausche E, Luck O, Harzer W (2004) Prevalence of malocclusions in the early mixed dentition and orthodontic treatment need. Eur J Orthod 26(3):237–244. https://doi.org/10.1093/ejo/26.3.237

67. Thilander B, Myrberg N (1973) The prevalence of malocclusion in Swedish schoolchildren. Scand J Dent Res 81(1):12–21. https://doi.org/10.1111/j.1600-0722.1973.tb01489.x

68. Tränkmann J (1988) Kieferorthopädische Mißerfolge nach übersehenen orofazialen Dyskinesien. Fortschr Kieferorthop 49:388–399

69. Tsiopas N, Nilner M, Bondemark L et al (2013) A 40 years follow-up of dental arch dimensions and incisor irregularity in adults. Eur J Orthod 35(2):230–235. https://doi.org/10.1093/ejo/cjr121

70. Zymperdikas VF, Koretsi V, Papageorgiou SN et al (2016) Treatment effects of fixed functional appliances in patients with Class II malocclusion: a systematic review and meta-analysis. Eur J Orthod 38(2):113–126. https://doi.org/10.1093/ejo/cjv034