

ORIGINAL ARTICLE

The effects of headgear treatment timing on costs and orthodontic outcome: a follow-up study

Matti Hannula^{a,b}, Mimmi Tolvanen^c, Pertti Pirttiniemi^{a,b}, Kirsi Pirilä-Parkkinen^{b,d}, Anna-Sofia Silvola^{a,b,d,e,f} and Johanna Julku^{a,b,d}

^aResearch Unit of Population Health, Faculty of Medicine, University of Oulu, Oulu, Finland; ^bMedical Research Center, Oulu University Hospital and University of Oulu, Oulu, Finland; ^cNational Rescue Services College, Kuopio, Finland; ^dDepartment of Oral and Maxillofacial Diseases, Oulu University Hospital, Oulu, Finland; ^eFaculty of Medicine and Health Technology, Tampere University, Tampere, Finland; ^fPirkanmaa Wellbeing Services County, Tampere, Finland

ABSTRACT

Objective: To analyze the effects of cervical headgear (CHG) timing on costs and orthodontic outcome.

Material and methods: Sixty-seven participants with Class II malocclusion were randomized into two groups. In the early group (EG, $n = 33$), CHG treatment began after the eruption of first upper molars. In the later-timed group (LG, $n = 34$), CHG treatment began nearly 2 years later. Orthodontic outcome was evaluated when CHG ended and at the end of follow-up using the Peer Assessment Rating (PAR) index and Little's Irregularity Index (LII). Costs were calculated using medical records, estimated chairside times and average personnel and material costs.

Results: The mean total costs including personnel and material costs during the follow-up were higher in the EG (€416) compared with the LG (€364) ($p = 0.043$). No statistically significant differences between the groups were found in the PAR and in the LII scores at the end of follow-up. The number of visits during CHG treatment was higher in the EG (13.3) compared with the LG (10.4) ($p = 0.01$).

Conclusions: Earlier initiation of CHG treatment increased the personnel and material costs, number of appointments, and chairside time. Based on the PAR and in the LII scores, the orthodontic outcome of CHG treatment was equal regardless of the timing of the treatment. If orthodontic treatment during mixed-dentition stage is considered, it would be justified from an economic perspective to favor treatment during late mixed-dentition.

ARTICLE HISTORY

Received 13 January 2026
Accepted 23 June 2026

KEYWORDS

Cervical headgear; costs; outcome

Introduction

The optimal treatment timing of Class II malocclusion has been studied extensively. Opinions on the correct timing of treatment vary, and there are results supporting various approaches. Early treatment during the mixed-dentition stage has been stated to reduce skeletal jaw discrepancy [1–3] and prevent incisor trauma in Class II malocclusion [4, 5]. On the other hand, early therapy for Class II malocclusion has been stated to extend the total treatment time, increase the number of appointments, and incur additional costs [6]. Early treatment often leads to two-phased treatment and a longer total treatment time [7]. However, there are also findings suggesting that the timing of Class II malocclusion treatment has no effect on costs or the treatment outcome [8].

Orthodontic treatment causes costs for society, patients, and the patients' guardians. Cost-effectiveness research in orthodontics has focused on comparing the outcomes and costs of different orthodontic appliances, aiming to find the most cost-effective method to achieve the desired outcome [9, 10]. Cost-minimization analyses (CMA) identify cost differences when the orthodontic outcome of treatment is the same [11].

In addition, cost-benefits can also be evaluated within orthodontic appliances based on their timing of use [6, 8]. While it is not reasonable to choose an orthodontic treatment method solely from an economic point of view, the costs of treatments must be considered as one aspect in treatment planning. Health economic research may help in designing orthodontic treatment plans and identifying effective treatment pathways.

In economic evaluations, relevant costs can be identified from different perspectives. From the perspective of healthcare, attention is focused on the costs incurred by the service provider and the cost-benefits of the treatment. From the patient's perspective, the costs incurred by the patient or their guardians, such as service fees and absences from work, are emphasized. According to a recently published article, the societal perspective, which considers all costs to society, is the correct approach to a full economic evaluation [10]. The most common point of view in high-quality cost analyses of orthodontic treatment is the societal perspective (50%), the second most common is healthcare (38%), and the rest are various combinations of healthcare/patient and healthcare/societal perspectives [10]. Administrators, healthcare providers, patients,

and the patients' parents all view the costs of orthodontic treatment from their own perspectives. Although it may not be possible to determine or measure all costs, it is important to recognize them all [11].

In addition to cost comparison, it is important to reliably measure the outcome of the treatment. The assessment of orthodontic outcomes has generally been carried out using the subjective opinion and experience of clinicians [12]. Occlusal indices are used to minimize subjectivity when evaluating treatment needs and the outcomes of orthodontic treatment [13–15]. Even though validated instruments are used to assess occlusion, some variation is always present among professionals [13, 14]. The Peer Assessment Rating (PAR) is a validated method to evaluate orthodontic outcomes and the extent to which malocclusion deviates from normal occlusion and alignment [16–18]. The Little's Irregularity Index (LII) has been demonstrated to be an appropriate measure for quantifying lower anterior crowding [19–21].

The primary aim of this study was to investigate the effect of cervical headgear (CHG) timing on treatment costs by evaluating the number and duration of appointments as well as personnel and material costs. The secondary aim was to assess the benefits of headgear timing on treatment outcomes by using occlusal indices.

Material and methods

Trial design and ethical approval

This CMA is a secondary analysis of the data from a previous randomized controlled trial [22–26]. The economic perspective of

the study was added later, and therefore it has not been mentioned in the original registration.

The regional medical research ethics committee of the Wellbeing services county of North Ostrobothnia, Finland (EETMK: 46/2003) and health center supervisors approved the trial set-up. The trial was conducted according to the Declaration of Helsinki.

Sample size calculation

A pre-study power analysis for an RCT was calculated with GPower 3.1 Software using 80% power and a 0.05 significance level. The sample size calculation was based on a CHG trial with a similar study design using the means and standard deviations (SDs) of the ANB angle values [27]. Eleven participants per group were needed to reach sufficient power. To tolerate patient drop-outs during the long-term follow-up, the number of participants was increased.

Participants, inclusion criteria and exclusion criteria

Participants were recruited to this trial (from February 2004 to June 2008) from three public health centers in Northern Finland. The pool of subjects comprised 270 7-year-old children who were assessed for eligibility by orthodontists. The inclusion criteria for this trial were Angle Class II molar relationships, deep bite, and overjet over 6 mm. Patients with earlier orthodontic treatment, obstructive sleep apnea, continuous airway infections, inborn facial asymmetry, facial syndrome, or the maxillary-mandibular plane angle (MMPA) over 35 degrees were excluded from this trial. The need for orthodontic treatment was

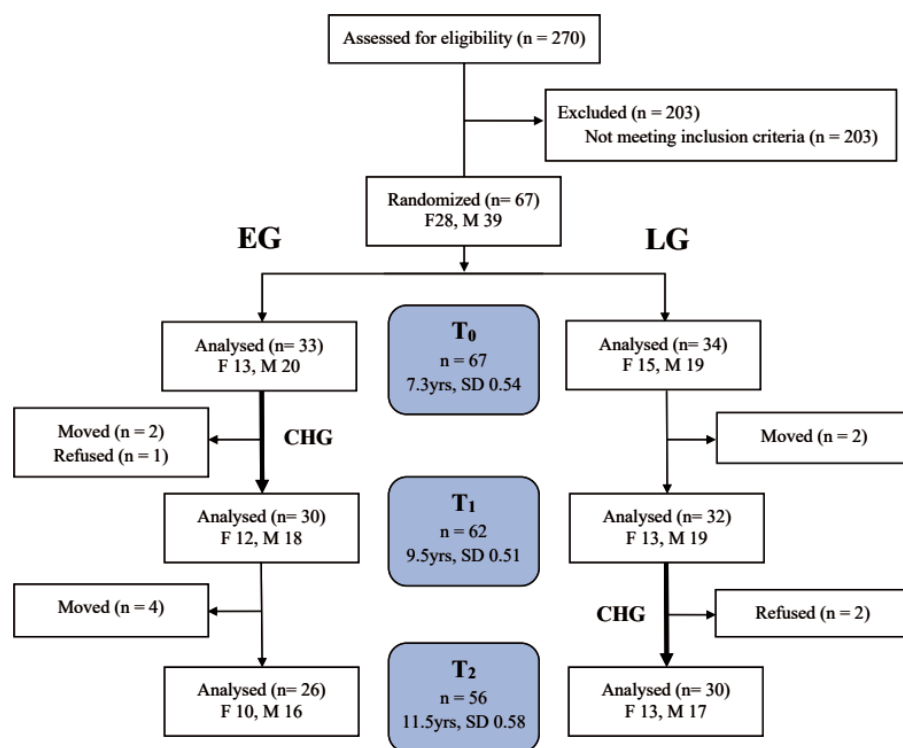


Figure 1. Flow chart of the trial. F: female; M: male; EG: early group; LG: later group; CHG: cervical headgear.

determined by a 10-grade scale based on the Treatment Priority Index (TPI) [28].

The flow of participants is presented in Figure 1. The total number of subjects fulfilling the inclusion criteria was 67 (early group [EG], $n = 33$; later group [LG], $n = 34$). The mean value of ANB was 5.2 (SD 1.92) in the EG and 5.4 (1.83) in the LG at T_0 [22]. Based on cephalometric analysis, there were no skeletal differences between the groups at baseline [22]. No differences were observed between the groups in dental arch dimensions at baseline [25].

At the end of follow-up (T_2), the total number of participants was 56 (EG, $n = 26$; LG, $n = 30$). The dropouts included eight children who had moved and three subjects who refused the treatment.

Randomization and blinding

Oral and written information about the study protocol was given to all the children fulfilling the eligibility criteria and their guardians, and written consent was obtained before the

study. The participants were randomized (1:1 ratio) into two equal-sized groups, the early and later groups. The randomization was made using sealed and opaque envelopes including numbered notes. The participants picked and opened one of the envelopes and gave it to the clinician to inform the treatment group. More specific details of the randomization process have been published in earlier papers concerning this trial [22–26].

Data analysts were blinded. All the analyzed material was coded so that the participants or groups could not be identified. Blinding of the personnel, patients, and the patients' guardians could not be achieved during CHG treatment.

Interventions

Appliance settings for the CHG treatment were identical among all participants. Activated force magnitude was 500 g. Children were instructed to wear CHG for 8–10 hours a day while sleeping. The inner bow of the CHG was expanded 5 mm, and the outer bow was bent 10 degrees upwards.

Table 1. Description of the participants: mean and median values and IQRs for age at the beginning of CHG (years), total duration of CHG (months), costs (€), PAR score, LII, and number of visits according to group and gender.

Description of the participants	EG				LG			
	Female		Male		Female		Male	
	Mean	Md (IQR)	Mean	Md (IQR)	Mean	Md (IQR)	Mean	Md (IQR)
Age at the beginning of CHG (years)	7.9	8.0 (7.4–8.3)	7.7	7.7 (7.2–8.1)	9.4	9.4 (9.2–9.6)	9.9	9.6 (9.2–10.0)
Total duration of CHG (months)*	29.2	28 (16–38)	26.7	21 (18–35)	19.9	19 (12–28)	21.4	21 (13–29)
COSTS (€)								
Personnel cost before CHG	79 ^b	78 (74–90)	64 ^b	64 (46–73)	98	95 (75–108)	99	95 (74–120)
Personnel cost during CHG (scheduled)	262	234 (188–359)	257	244 (185–310)	187	148 (139–243)	199	185 (148–259)
Personnel cost, emergency during CHG	33	24 (15–46)	28	21 (9–44)	19	12 (0–24)	26	30 (6–37)
Personnel cost, No-show during CHG	13	0 (0–19)	15	0 (0–19)	11	0 (0–19)	13	0 (0–19)
Total personnel cost during CHG	308	314 (213–403)	300	302 (225–350)	216	173 (139–280)	238	217 (172–302)
Total personnel cost (before + during)	387	41 (281–489)	364	367 (289–418)	314	287 (235–369)	337	327 (272–388)
Material costs	36	30 (30–39)	38	39 (30–44)	35	32 (30–37)	33	30 (30–35)
Material cost, emergency	4	4 (0–6)	6	4 (0–9)	2	0 (0–2)	4	2 (0–4)
Total material cost	40	35 (32–45)	44	43 (34–52)	38	34 (30–42)	36	34 (30–37)
TOTAL COSTS during CHG	348	362 (244–454)	344	347 (255–402)	254	202 (174–313)	274	247 (202–338)
TOTAL COSTS (during + before)	427	457 (313–536)	408	414 (319–471)	352	320 (267–408)	373	366 (305–426)
ORTHODONTIC OUTCOME								
PAR when CHG ended**	18.2	17 (11–21)	18.5	17 (12–22)	16.0	15 (13–19)	18.6	16 (14–21)
PAR at T_2	13.1	12 (9–18)	16.5	15 (12–21)	16.0	15 (13–19)	18.6	16 (14–21)
LII when CHG ended**	5.1 ^a	5.2 (4.3–6.3)	3.4 ^a	2.8 (2.0–3.7)	3.2	2.0 (1.4–4.7)	4.1	3.5 (1.8–6.9)
LII at T_2	5.3	4.1 (2.9–6.8)	3.7	3.2 (1.7–5.8)	3.2	2.0 (1.4–4.7)	4.1	3.5 (1.8–6.9)
VISITS								
Scheduled visits before CHG	4.5	4 (4–5)	4.0	4 (3–4)	5.8	6 (4–7)	5.8	6 (5–7)
Scheduled visits during CHG	13.8	13 (9–19)	10.1	12 (10–16)	10.1	10 (7–12)	10.6	9 (9–14)
Emergency visits related to CHG	2.2	2 (1–4)	1.8	2 (1–2)	1.2	1 (0–2)	0.7	2 (1–2)
No-show visits during CHG	0.7	0 (0–1)	0.8	0 (0–1)	0.5	0 (0–1)	1.5	0 (0–1)
Total visits	21.1	22 (15–26)	19.6	19 (15–23)	17.7	18 (14–21)	18.6	17 (16–22)

Gender differences evaluated using the Mann-Whitney U test and the χ^2 test. CHG: cervical headgear; IQR: interquartile range; LII: Little's Irregularity Index; PAR: Peer Assessment Rating; EG: early group; LG: later-timed group.

*Total duration of CHG treatment included active phase and reduced use phase.

**In the EG, CHG ended at T_1 . In the LG, CHG ended at T_2 .

^aStatistically significant gender differences in EG: LII when CHG ended ($p = 0.008$).

^bStatistically significant gender differences in EG: Personnel cost before ($p = 0.008$).

In the EG, the CHG treatment started after eruption of the first upper molars (7.8 yrs, interquartile range [IQR] 7.3–8.2). The EG was treated with CHG between time points T_0 and T_1 . In the LG, the CHG therapy started nearly 2 years later compared with the EG after T_1 (9.7 yrs, IQR 9.2–9.7). The LG was treated with CHG between time points T_1 and T_2 . The treatment protocol was similar in both groups. The CHG treatment was carried out until Angle Class I molar relationships were achieved. Reduced use of CHG was applied when needed.

Orthodontic treatment was planned and supervised by an orthodontist who further delegated the procedures to general dentists and dental hygienists. During the orthodontic visits, general dentists and dental hygienists performed only orthodontic procedures.

Outcomes

Main outcomes of the present study were total costs and the PAR score. Additional outcomes were LII, number of visits and chairside time during CHG treatment.

Data handling

All scanning and measuring of dental casts were performed by one of the authors (M.H.). Appointment details for the cost

analysis were collected from the electric patient records by the same author (M.H.).

Cost measures

The approach chosen for the economic evaluation was to use the perspective of healthcare system. In this study, the focus was on personnel and material costs, number of appointments, and chair time. The material cost calculations were based on average commercial prices in 2022. The personnel costs in 2021 were collected from the Statistics Finland open database [29]. Treatment time was calculated by using the number of appointments and estimated average duration of an orthodontic appointment in each personnel category. The expenses of facilities, equipment, and maintenance were excluded from this analysis because they depend strongly on local factors and determining them accurately proved to be impossible. Indirect costs resulting from the guardian's absence from work and loss of income when accompanying their child to the orthodontic appointments were excluded from this study due to the perspective chosen.

Outcome measures

Dental casts were manufactured using alginate impressions and bite indexes, which were taken at T_1 (9.5 years, SD 0.50) and at T_2

Table 2. Mean and median values and IQRs for age at the beginning of CHG (years), total duration of CHG (months), costs (€), PAR score, LII, and number of visits according to group, and significances of the differences between the groups.

Costs, outcome and visits	EG		LG		<i>p</i>
	Mean	Md (IQR)	Mean	Md (IQR)	
Age at the beginning of CHG (years)	7.8	7.8 (7.3–8.2)	9.7	9.4 (9.2–9.7)	<0.001
Total duration of CHG (months)*	27.7	21 (18–35)	20.8	21 (13–28)	0.081
COSTS (€)					
Personnel cost before CHG	70	70 (58–84)	99	95 (76–120)	<0.001
Personnel cost during CHG (scheduled)	259	241 (185–333)	194	174 (147–259)	0.004
Personnel cost, emergency during CHG	30	24 (12–43)	23	19 (0–37)	0.254
Personnel cost, no-show during CHG	14	0 (0–19)	12	0 (0–19)	0.987
Total personnel cost during CHG	304	314 (222–374)	229	212 (148–285)	0.004
Total personnel cost (before + during)	374	373 (293–470)	327	315 (250–371)	0.051
Material costs	37	37 (30–41)	34	32 (30–37)	0.100
Material cost, emergency	5	4 (0–8)	3	1 (0–4)	0.035
Total material cost	42	41 (34–51)	37	34 (30–38)	0.024
TOTAL COSTS during CHG	346	352 (255–422)	265	247 (183–319)	0.002
TOTAL COSTS (before + during)	416	520 (414–584)	364	350 (287–407)	0.043
ORTHODONTIC OUTCOME					
PAR when CHG ended**	18.4	17 (12–22)	17.5	16 (14–19)	0.655
PAR at T_2	15.0	13 (10–17)	17.5	16 (14–19)	0.146
LII when CHG ended**	4.1	3.7 (2.7–5.6)	3.7	3.3 (1.6–5.8)	0.325
LII at T_2	4.4	3.8 (2.3–6.6)	3.7	3.3 (1.6–5.8)	0.369
VISITS					
Scheduled visits before CHG	4.2	4 (3–5)	5.8	6 (5–7)	<0.001
Scheduled visits during CHG	13.3	12 (10–18)	10.4	9 (8–12)	0.010
Emergency visits related to CHG	2.0	2 (1–3)	1.4	1 (0–2)	0.232
No-show visits during CHG	0.8	0 (0–1)	0.6	0 (0–1)	0.967
Total visits	20.2	19 (15–24)	18.2	18 (15–21)	0.193

Group differences evaluated using the Mann-Whitney *U* test. CHG: cervical headgear; IQR: interquartile range; LII: Little's Irregularity Index; PAR: Peer Assessment Rating; EG: early group; LG: later-timed group.

*Total duration of CHG treatment included active phase and reduced use phase.

**In the EG, CHG ended at T_1 . In the LG, CHG ended at T_2 .

(11.5 years, SD 0.62). Dental casts were scanned into digital 3D models (3Shape, R700TM Orthodontic Scanner, Denmark). The occlusion indices were measured from 3D models with an analyzer program (3Shape, OrthoAnalyzer™ 2012, Denmark).

Evaluation of orthodontic outcome was based on the PAR index and LII. The PAR index consists of five components: anterior segments, buccal segments, overjet, overbite, and midline. These components are weighted and summed to the PAR total [16–18]. The occlusion is classified as an almost ideal ($PAR < 5$), acceptable ($5 \leq PAR \leq 10$), or unacceptable occlusion ($PAR > 10$) [16]. The LII describes mandibular anterior alignment [19, 20]. The lower anterior segment is determined into perfect alignment ($LII = 0$), minimal irregularity ($1 \leq LII \leq 3$), moderate irregularity ($4 \leq LII \leq 6$), severe irregularity ($7 \leq LII \leq 9$), and very severe irregularity ($LII > 10$).

Intra-examiner reliability of the PAR and LII measurements was determined by duplicating the measurements of 20 dental casts.

Statistics

Differences in age at the beginning of CHG, total duration of CHG, PAR scores, LII scores, number of visits, costs, and chairside time between groups and genders were evaluated using the Mann-Whitney U test for continuous variables and the χ^2 test for categorical variables. Some of the variables were non-normally distributed, and therefore Mann-Whitney U test was chosen, and both mean and median values as well as IQR were reported. The intraclass correlation coefficient (ICC) was evaluated using two-way analysis of variance (ANOVA). Analyses were carried out using IBM SPSS version 27 (SPSS, Inc., Chicago, IL, USA).

Results

The treatment was carried out between time points T_0 – T_1 in the EG and between T_1 and T_2 in the LG. The active phase lasted 1.6 years (SD 0.76) in the EG and 1.4 years (SD 0.80) in the LG.

Intra-examiner reliability

The ICC value was 0.941 for PAR measurements and 0.989 for LII measurements. The ICC values indicated an excellent level of intra-examiner reliability.

Costs

The personnel costs before CHG treatment were lower in the EG (€70) compared with the LG (€99) ($p < 0.001$; Table 2). In the EG, the personnel costs before CHG treatment were higher for females (€79) compared with males (€63) ($p = 0.008$; Table 1). The personnel costs during CHG treatment were higher in the EG (€259) compared with the LG (€193) ($p = 0.004$; Table 2). The total personnel costs during CHG treatment, including emergency and no-show visits, were higher in the EG (€303) compared with the LG (€228) ($p = 0.004$; Table 2).

The mean total material costs of CHG treatment and emergency visits were slightly higher in the EG compared with the LG ($p = 0.043$, $p = 0.035$, respectively) although the absolute cost differences were small and not clinically meaningful (€42 vs. €36 and €5 vs. €3, respectively).

The mean total costs including personnel and material costs during the follow-up were €416 for the EG and €363 for the LG ($p = 0.043$; Table 2). During the CHG treatment, the mean total costs including personnel and material costs were higher in the EG (€346) compared with the LG (€265) ($p = 0.002$; Tables 1 and 2; Figure 2).

Orthodontic outcome

No statistically significant differences between groups or genders were found in the PAR scores at T_2 (Tables 1 and 2; Figure 2). There were no significant differences in the LII scores between the groups at T_2 . In the EG, the LII scores were higher for females (5.1) compared with males (3.4) when CHG ended ($p = 0.008$; Table 1).

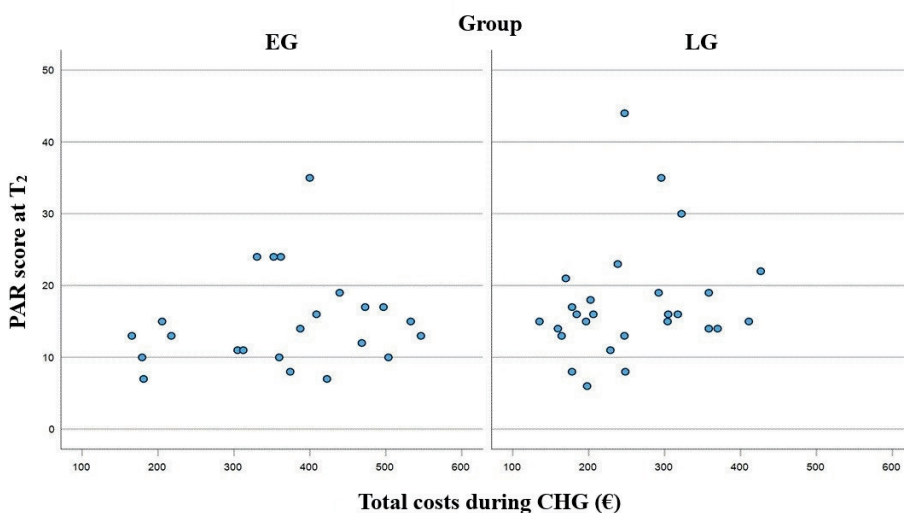


Figure 2. Dot-dash plot visualizing PAR scores at T_2 and total costs during CHG (€) according to group and gender of participants. EG: early group, LG: later group; PAR: Peer Assessment Rating; CHG: cervical headgear.

Number of appointments

The mean number of visits before CHG treatment was lower in the EG (5.2) compared with the LG (5.8) ($p < 0.001$; Table 2). The number of visits during CHG treatment was higher in the EG (13.3) compared with the LG (10.4) ($p = 0.01$; Table 2). No statistically significant differences were found between groups in emergency and no-show visits.

Chairside time

The mean total chairside time was 270 minutes for the EG and 207 minutes for the LG ($p = 0.006$; Table 3). Among all participants, the total chairside time for visits to general dentists was longer in the EG (215 minutes) compared with the LG (152 minutes) ($p = 0.004$; Table 3). The mean total chairside time among females was 278 minutes for the EG and 200 minutes for the LG ($p = 0.019$; Table 3). Among females, the total chairside time for appointments with general dentists was longer in the EG (216 minutes) compared with the LG (144 minutes) ($p = 0.039$; Table 3). No significant differences were found between the groups in the total chairside time for visits to orthodontists or dental hygienists.

Harms

No harms were encountered.

Discussion

Main findings

The results of the present study show that the personnel costs, number of appointments, and chairside time during CHG treatment were higher in the EG compared with the LG. This study also shows that the timing of CHG treatment has no significant effect on the overall orthodontic outcome based on evaluation of indices.

Interpretation

Earlier start of orthodontic treatment has been found to increase the number of visits when considering the total orthodontic treatment [30]. Present results showed that the earlier timing of CHG treatment increased the number of appointments. The same effect can be found with the twin-block appliance treatment [6]. By contrast, the timing of headgear activator treatment was found to have no effect on the number of visits [8]. One possible explanation for the increased number of visits in the EGs is the longer need for reduced use of orthodontic appliances during follow-up. Therefore, more appointments do not necessarily mean improved outcomes.

In addition to the increased number of visits, personnel costs were higher in the EG compared with the LG. Similar effects have been observed with the costs of earlier timing of twin-block appliance and overall earlier timing of orthodontic treatment [6, 30]. This is likely due to the typical challenge of early orthodontic treatment, where there is more growth remaining and treatment is started well before the growth spurt. Overall, the results revealed that total costs during CHG treatment were 30% higher in the EG. Cost differences may be small for individual patients, but when viewed at the population level, the differences can become significant. Therefore, orthodontists should have a general understanding of the costs associated with orthodontic treatments.

The results of the present study showed that the timing of CHG had no significant effect on the overall orthodontic outcome, which is consistent with earlier studies focused on twin-block appliance and headgear activator [6, 8]. Previous studies have demonstrated that the timing of headgear treatment has effects on individual dental measurements [22–26]. When examining individual measurements, the timing of orthodontic treatment may show differences between groups. However, when assessing the overall occlusion, occlusal indices are more comprehensive methods to describe orthodontic outcomes than individual measurements.

Table 3. Mean and median values and IQRs for chairside time during CHG (min) including scheduled appointments and emergency appointments according to group and gender of participants, and dentistry personnel.

Chairside time during CHG (min)		EG		LG		<i>p</i>
		Mean	Md (IQR)	Mean	Md (IQR)	
All	Dental hygienist	52	45 (0–75)	47	45 (15–75)	0.924
	General dentist – Dental nurse	215	210 (165–270)	152	143 (105–214)	0.004
	Orthodontist – Dental nurse	3	0 (0–0)	8	0 (0–6)	0.320
	Total	270	240 (195–315)	207	185 (150–255)	0.006
Female	Dental hygienist	58	60 (8–90)	50	45 (15–75)	0.687
	General dentist – Dental nurse	216	225 (150–293)	144	120 (98–195)	0.039
	Orthodontist – Dental nurse	5	0 (0–5)	6	0 (0–3)	0.960
	Total	278	270 (198–368)	200	165 (128–250)	0.019
Male	Dental hygienist	48	38 (0–75)	45	30 (15–60)	0.832
	General dentist – Dental nurse	214	210 (161–259)	158	150 (105–225)	0.053
	Orthodontist – Dental nurse	2	0 (0–0)	9	0 (0–10)	0.303
	Total	263	233 (195–304)	212	195 (155–265)	0.110

Group differences evaluated using the Mann-Whitney *U* test. CHG: cervical headgear; IQR: interquartile range; EG: early group; LG: later-timed group.

The PAR index is a widely-used method to assess orthodontic outcomes and the extent to which malocclusion deviates from normal alignment and occlusion [16–18]. The LII index has been found suitable for assessing lower anterior crowding, but it has also been criticized due to challenges with repeatability [19–21]. However, the PAR index has been used in previous CHG studies to describe the outcome of orthodontic treatment [7, 31], and the LII has been used to measure lower anterior alignment changes during CHG treatment [31].

Current results showed that the mean total chairside time during CHG treatment was longer in the EG compared with the LG. A previous study on the timing of headgear activator treatment found no differences in the chairside time between the groups [8]. The number of visits and the chairside time have a direct relationship with the costs: as the number of visits and chairside time increases, costs generally increase, and vice versa. Another finding in this study was that statistically significant differences in chairside time were observed between groups among females. Females in the LG had shorter chairside time compared to females in the EG, which could point to earlier dental development in females compared to males. The orthodontist's chairside time was minimal, which can be attributed to the fact that most of the patients were not seen by the orthodontist during the headgear treatment. The orthodontic treatment was performed by the general dentists and dental hygienists, and the role of the orthodontists was to plan and supervise the treatment.

The average coverage of the orthodontic treatment among 0–18-year-olds is reported to be 11% in Finnish public healthcare, but the range between municipalities is wide, varying between 2 and 43% [32]. Class II malocclusion is the third most common reason for orthodontic treatment after anterior and posterior crossbite among children aged 7–9 years, and the most common reason among children aged 10–13 years within public healthcare services [33]. In Finland, the CHG is one of the most used appliances in the orthodontic treatment of children [32, 34].

Orthodontic treatment methods should not be chosen solely based on economic factors, but the costs of treatments must be considered due to limited resources. If earlier treatment timing is considered, it is reasonable to expect to have advantages over later treatment timing. When the outcomes are the same, it is reasonable to favor a shorter, more efficient, and lower-cost treatment approach. The burden of treatment on patients and parents cannot be ignored. This study provides important information on the impact of CHG treatment timing on the orthodontic outcome and costs, so that CHG could be used as efficiently as possible.

Limitations, strengths, and generalization

The main limitation of this study is that it does not constitute a full economic evaluation. Thus, resource allocation cannot be based on this study. Another limitation is that the PAR and LII scores could not be measured at baseline, due to the early mixed-dentition phase. However, previous studies using the

same dataset have not found differences in individual dental arch measurements between groups at baseline, supporting the validity of the outcome comparisons [22–26]. Limitation is also that the use of CHG was not monitored, so participants' adherence to the treatment instructions cannot be confirmed. In this study, limitations are potentially caused by the fact that material was collected between 2004 and 2012, and in addition, the publication of the results has been delayed. The age of the material must be considered when interpreting the results.

The strength of this article is the use of an RCT design, and the outcome measures used have been shown to be valid and reliable [16–20].

Single-center design leads to limited generalizability and may cause potential bias. It is worth noting that the costs depend on local factors and cannot be extrapolated to other locations.

Interest in conducting an economic evaluation emerged after the initiation of the study. Economic evaluation was not mentioned in the study registration of the preceding RCT, which may have led to selective reporting bias.

This study represents a secondary analysis, and the sample size was not originally determined based on the present study outcomes. The sample size is too small to achieve adequate statistical power. Consequently, the analysis is at increased risk of Type II error, meaning that true effects may remain undetected, and this may have influenced the findings of the present study.

No correction for multiple comparisons was applied, but if all the null hypotheses are true, 5% of comparisons are expected to yield uncorrected p -values < 0.05 [35, 36]. With 78 statistical tests, this corresponds to about four chance significant results. In the present analyses, however, 16 p -values < 0.05 were observed. This number is greater than the ~4 expected by random variation alone, providing evidence that the observed effects are unlikely to be due to chance and instead support our study hypotheses.

Conclusion

The current results showed that earlier timing of CHG treatment increased personnel and material costs, number of appointments, and chairside time. Regardless of the timing of CHG treatment, the overall orthodontic outcomes based on PAR and LII scores were equal. If orthodontic treatment during mixed-dentition stage is considered, it would be justified from an economic perspective to favor treatment during late mixed-dentition.

Acknowledgements

We would like to thank all children, their parents, and the health-care staff for participating in and making this study possible.

Funding

This work was supported by the Finnish Dental Society Apollonia under Grant [20210043].

Disclosure of interest

The authors report there are no competing interests to declare.

Data availability

The data underlying this paper will be shared by the corresponding author upon reasonable request.

References

- [1] Wheeler TT, McGorray SP, Dolce C, Taylor M, King G. Effectiveness of early treatment of Class II malocclusion. *Am J Orthod Dentofacial Orthop.* 2002;121(1):9–17. <https://doi.org/10.1067/mod.2002.120159>
- [2] Tulloch JF, Phillips C, Koch G, Proffit W. The effect of early intervention on skeletal pattern in Class II malocclusion: a randomized clinical trial. *Am J Orthod Dentofacial Orthop.* 1997;111(4):391–400. [https://doi.org/10.1016/S0889-5406\(97\)80021-2](https://doi.org/10.1016/S0889-5406(97)80021-2)
- [3] Tulloch JF, Proffit WR, Phillips C. Influences on the outcome of early treatment for Class II malocclusion. *Am J Orthod Dentofacial Orthop.* 1997;111(5):533–42. [https://doi.org/10.1016/S0889-5406\(97\)70290-7](https://doi.org/10.1016/S0889-5406(97)70290-7)
- [4] Thiruvengkatachari B, Harrison J, Worthington H, O'Brien K. Early orthodontic treatment for Class II malocclusion reduces the chance of incisal trauma: results of a Cochrane systematic review. *Am J Orthod Dentofacial Orthop.* 2015;148(1):47–59. <https://doi.org/10.1016/j.jado.2015.01.030>
- [5] Batista K, Thiruvengkatachari B, Harrison J, O'Brien K. Orthodontic treatment for prominent upper front teeth (Class II malocclusion) in children and adolescents. *Cochrane Database Syst Rev.* 2018;3: CD003452. <https://doi.org/10.1002/14651858.CD003452.pub4>
- [6] O'Brien K, Wright J, Conboy F, Appelbe P, Davies L, Connolly I, et al. Early treatment for Class II Division 1 malocclusion with the twin-block appliance: a multi-center, randomized, controlled trial. *Am J Orthod Dentofacial Orthop.* 2009;135(5):573–9. <https://doi.org/10.1016/j.jado.2007.10.042>
- [7] Pirttiniemi P, Kantomaa T, Mäntysaari R, Pykäläinen A, Krušinskiene V, Laitala T, et al. The effects of early headgear treatment on dental arches and craniofacial morphology: an 8 year report of a randomized study. *Eur J Orthod.* 2005;27(5):429–36. <https://doi.org/10.1093/ejo/cji025>
- [8] Kallunki J, Bondemark L, Paulsson L. Comparisons of costs and treatment effects – an RCT on headgear activator treatment of excessive overjet in the mixed and late mixed dentition. *Eur J Orthod.* 2022;44(1):86–94. <https://doi.org/10.1093/ejo/cjab026>
- [9] Sollenius O, Petré S, Björnsson L, Norlund A, Bondemark L. Health economic evaluations in orthodontics: a systematic review. *Eur J Orthod.* 2016;38(3):259–65. <https://doi.org/10.1093/ejo/cjv040>
- [10] Jermyn O, Bister D, Jin H. Cost-effectiveness of orthodontics: a systematic review. *Eur J Orthod.* 2022;44(5):566–77. <https://doi.org/10.1093/ejo/cjac019>
- [11] Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. *Methods for the economic evaluation of health care programmes.* 4th ed. Oxford: Oxford University Press; 2015.
- [12] Dyken RA, Sadowsky PL, Hurst D. Orthodontic outcomes assessment using the peer assessment rating index. *Angle Orthod.* 2001;71(3):164–9. [https://doi.org/10.1043/0003-3219\(2001\)071<0164:OOAUP>2.0.CO;2](https://doi.org/10.1043/0003-3219(2001)071<0164:OOAUP>2.0.CO;2)
- [13] Richmond S, Daniels CP. International comparisons of professional assessments in orthodontics: part 2 – treatment outcome. *Am J Orthod Dentofacial Orthop.* 1998;113(3):324–8. [https://doi.org/10.1016/S0889-5406\(98\)70304-X](https://doi.org/10.1016/S0889-5406(98)70304-X)
- [14] Richmond S, Daniels CP. International comparisons of professional assessments in orthodontics: part 1 – treatment need. *Am J Orthod Dentofacial Orthop.* 1998;113(2):180–5. [https://doi.org/10.1016/S0889-5406\(98\)70290-2](https://doi.org/10.1016/S0889-5406(98)70290-2)
- [15] Khandakji MN, Ghafari JG. Evaluation of commonly used occlusal indices in determining orthodontic treatment need. *Eur J Orthod.* 2019;1:107–14. <https://doi.org/10.1093/ejo/cjz042>
- [16] Richmond S, Shaw WC, O'Brien K, Buchanan I, Jones R, Stephens C, et al. The development of the PAR Index (Peer Assessment Rating): reliability and validity. *Eur J Orthod.* 1992;14(2):125–39. <https://doi.org/10.1093/ejo/14.2.125>
- [17] Shaw W, Richmond S, O'Brien K. The use of occlusal indices: a European perspective. *Am J Orthod Dentofacial Orthop.* 1995;107(1):1–10. [https://doi.org/10.1016/S0889-5406\(95\)70151-6](https://doi.org/10.1016/S0889-5406(95)70151-6)
- [18] Richmond S, Shaw W, Roberts C, Andrews M. The PAR Index (Peer Assessment Rating): methods to determine outcome of orthodontic treatment in terms of improvement and standards. *Eur J Orthod.* 1992;14(3):180–7. <https://doi.org/10.1093/ejo/14.3.180>
- [19] Little RM. The irregularity index: a quantitative score of mandibular anterior alignment. *Am J Orthod Dentofacial Orthop.* 1975;68(5):554–63. [https://doi.org/10.1016/0002-9416\(75\)90086-X](https://doi.org/10.1016/0002-9416(75)90086-X)
- [20] Bernabe E. Estimating arch length discrepancy through Little's Irregularity Index for epidemiological use. *Eur J Orthod.* 2005;28(3):269–73. <https://doi.org/10.1093/ejo/cji112>
- [21] Macauley D, Garvey T, Dowling A, Fleming G. Using Little's Irregularity Index in orthodontics: outdated and inaccurate? *J Dent.* 2012;40(12):1127–33. <https://doi.org/10.1016/j.jdent.2012.09.010>
- [22] Julku J, Pirilä-Parkkinen K, Pirttiniemi P. Airway and hard tissue dimensions in children treated with early and later timed cervical headgear – a randomized controlled trial. *Eur J Orthod.* 2018;40(3):285–95. <https://doi.org/10.1093/ejo/cjx088>
- [23] Julku J, Pirilä-Parkkinen K, Tolvanen M, Pirttiniemi P. Comparison of effects of cervical headgear treatment on skeletal facial changes when the treatment time is altered: a randomized controlled trial. *Eur J Orthod.* 2019;41(6):631–40. <https://doi.org/10.1093/ejo/cjz053>
- [24] Hannula M, Tolvanen M, Pirttiniemi P, Pirilä-Parkkinen K, Julku J. Effects of headgear timing on dental arch changes from 7 to 18 years of age: a follow-up study. *Eur J Orthod.* 2023;45(5):496–504. <https://doi.org/10.1093/ejo/cjad045>
- [25] Julku J, Hannula M, Pirilä-Parkkinen K, Tolvanen M, Pirttiniemi P. Dental arch effects after early and later timed cervical headgear treatment – a randomized controlled trial. *Eur J Orthod.* 2019;41(6):622–30. <https://doi.org/10.1093/ejo/cjy083>
- [26] Kasmä K, Silvola AS, Vuollo V, Julku J. Effects of early and later timed cervical headgear treatment on the eruption timing and pattern of permanent upper canines and molars from 7 to 18 years of age: follow-up of a randomized controlled trial. *Eur J Orthod.* 2025;47(4):cjaf056. <https://doi.org/10.1093/ejo/cjaf056>
- [27] Mäntysaari R, Kantomaa T, Pirttiniemi P, Pykäläinen A, Krušinskiene V, Laitala T, et al. The effects of early headgear treatment on dental arches and craniofacial morphology: a report of a 2 year randomized study. *Eur J Orthod.* 2004;26(1):59–64. <https://doi.org/10.1093/ejo/26.1.59>
- [28] National Board of Health. Instruction letter no. 6/88: organization of orthodontic treatment. Helsinki: National Board of Health; 1988.
- [29] Statistics Finland. Local government sector wages and salaries [Internet]. 2021 [cited 2025 Mar 18]. Available from: https://pxweb2.stat.fi/PxWeb/pxweb/en/StatFin/StatFin__ksp/statfin_ksp_pxt_11ap_f.px/
- [30] Pietilä I, Pietilä T, Svedstrom-Oristo AL, Varrelä J, Alanen P. Orthodontic treatment practices in Finnish municipal health centres with differing timing of treatment. *Eur J Orthod.* 2009;31(3):287–93. <https://doi.org/10.1093/ejo/cjn109>
- [31] Krušinskiene V, Kiuttu P, Julku J, Silvola A, Kantomaa T, Pirttiniemi P. A randomized controlled study of early headgear treatment on occlusal stability – a 13 year follow-up. *Eur J Orthod.* 2008;30(4):418–24. <https://doi.org/10.1093/ejo/cjn021>

- [32] Pietilä I, Pietilä T, Varrela J, Pirttiniemi P, Alanen P. Trends in Finnish Public Orthodontic Care from the professionals' perspective. *Int J Dent*. 2009;2009:945074. <https://doi.org/10.1155/2009/945074>
- [33] Pietilä I, Pietilä T, Pirttiniemi P, Varrela J, Alanen P. Orthodontists' views on indications for and timing of orthodontic treatment in Finnish public oral health care. *Eur J Orthod*. 2007;30(1):46–51. <https://doi.org/10.1093/ejo/cjm085>
- [34] Arpalahti A, Saarnio-Syrjäläinen A, Laaksonen S, Arponen H. Early orthodontic treatment in a Finnish public health centre: a retrospective cross-sectional study. *Acta Odontol Scand*. 2023;81(5):396–401. <https://doi.org/10.1080/00016357.2022.2161623>
- [35] Rothman KJ. No adjustments are needed for multiple comparisons. *Epidemiology*. 1990;1:43–6. <https://doi.org/10.1097/00001648-199001000-00010>
- [36] Saville J. No multiple comparison procedures: the practical solution. *Am Stat*. 1990;44:174–80. <https://doi.org/10.1080/00031305.1990.10475712>