Hawley retainers full- or part-time? A randomized clinical trial

M. Shawesh*, B. Bhatti**, T. Usmani** and N. Mandall*

*School of Dentistry, University of Manchester and **Heaton Mersey Specialist Orthodontic Practice, Manchester, UK

SUMMARY The aim of this trial was to compare two different orthodontic retention regimens: is nightonly wear of upper and lower Hawley retainers for 1 year as effective as 6 months full-time followed by 6 months night-only wear? Sixty-seven consecutive patients attending for orthodontic debond were randomly allocated to wear upper and lower Hawley retainers either for 1 year night-only (group 1) or for 6 months full-time followed by 6 months night-only (group 2). In group 1, 41.2 per cent were males and 58.8 per cent were females and their mean age was 15.6 years [standard deviation (SD) 1.6 years]. In group 2, 24.2 per cent were males and 75.8 per cent were females and their mean age was 15.8 years (SD 1.2 years). Study models were taken at the start (T0) and end (T1) of treatment and 1 year post-debond (T2). Digital callipers were used to measure upper and lower labial segment irregularity using Little's index and upper and lower labial segment crowding. To evaluate differences between groups 1 and 2 *t*-tests were used.

There were no statistically significant differences between the two retention regimens at T2 for labial segment irregularity or crowding (P > 0.05). Since both retention regimens were equally effective during the 1 year retention period, it would seem clinically acceptable to ask patients to wear their retainers at night only.

Introduction

The aim of this randomized clinical trial was to determine whether there is any difference between two orthodontic retention regimens: whether night-only wear of an upper and lower Hawley retainer for 1 year is as effective as 6 months full-time wear followed by 6 months night-only wear.

Most orthodontic clinicians will carry out supervised retention for at least 1 year after active treatment has ceased. The scientific rationale for this is that Reitan (1959, 1967) suggested that the gingival fibre network typically took 4–6 months to remodel and periodontal fibres took at least 262 days to re-organize, thus necessitating a means of maintaining teeth in their new post-treatment position. Use of retainers, theoretically, prevents the tendency of teeth to return to their pre-treatment positions not only from the influence of periodontal and gingival fibres but also from occlusal and soft tissue forces and continued dentofacial growth.

Southard *et al.* (1992) investigated the potential role of periodontal transseptal fibres, which were thought to be the prime force in exerting compression between mandibular contact points. They showed, using a digital tension transducer to record the interproximal force, that elastic supracrestal fibres continued to exert significant forces between mandibular contact points, possibly contributing to post-treatment changes in tooth position.

There is a wide variation in the retention regimen used by orthodontists, varying from immediate night-only wear of retainers to a period of 3–6 months full-time wear followed by night-only wear.

There are limited prospective studies that investigate this question and the problem of lack of scientific evidence has been highlighted in a systematic review (Littlewood et al., 2006). However, Destang and Kerr (2003) investigated maxillary retention in two parallel groups to determine whether a longer retention period would decrease the relapse potential and increase stability. Twenty patients were allocated to a 6 month retention regimen using an upper Hawley retainer for 3 months full-time and 3 months night-only. The second group of 18 patients followed a 1 year retention regimen, with the same retainer, wearing it for 6 months full time followed by 6 months night-only. They found that the second group who experienced an overall retention regimen for 1 year showed less post-retention irregularity relapse of the maxillary anterior teeth compared with the group who had only worn a retainer for 6 months. They concluded that retention for 1 year, rather than 6 months, was clinically more beneficial.

There are numerous studies of possible variables that may influence orthodontic retention and relapse; these include: continued post-treatment facial growth and development (Björk and Skieller, 1972; Lopez-Gavito *et al.*, 1985; Little *et al.*, 1990; Wieslander, 1993), proclination of the lower labial segment and expansion of the intercanine width during orthodontic treatment (Mills, 1968; Little *et al.*, 1981; Felton, 1987), arch length deficiency (Richardson, 1996), tooth fibre discrepancy and a triangular shape of the lower incisors (Peck and Peck, 1972), and the mesial drift theory (Richardson, 1979) and the third molar theory. However, several published studies suggest that the latter plays a very minor role in long-term changes to the dental arch (Richardson, 1989; Ades *et al.*, 1990; Harradine *et al.*, 1998). It is still unclear whether it may be clinically acceptable for patients to wear their retainers for 1 year at night only or whether it is necessary for an initial period of full-time wear followed by night-only wear. This question, therefore, formed the focus of this study.

The null hypothesis tested was that there is no difference in the effectiveness of upper and lower Hawley retainers whether worn night-time only for 1 year or 6 months fulltime followed by 6 months night-only in terms of the incisor irregularity index (Little, 1975) and incisor crowding.

Subjects and methods

Sample size calculation

The sample size for each group was calculated as n = 23, based on an alpha significance level of 0.05 and a beta of 0.1. This gave a power of 90 per cent to detect a clinically significant difference (if one existed) of 2 mm in labial segment alignment between group 1 (1 year night-only) and group 2 (6 months full-time followed by 6 months night-only), assuming that the common standard deviation is 2 mm using a two-group *t*-test with a 0.05, two-sided significance level. This gave a total sample size of 46 patients required for the study.

The sample was obtained from patients attending for orthodontic debond appointments at one specialist practice in Heaton Mersey, Manchester, United Kingdom. The patients were consecutively approached to take part in the study and the patient and parent signed a consent form. The study protocol had previously been approved by the Central Manchester Research Ethics Committee (reference: 03/07/2307).

The inclusion criteria were 10–16 years of age, labial segment crowding or tooth contact point displacement at the start of orthodontic treatment, clinically acceptable labial segment alignment at the end of active treatment, and good oral hygiene.

The exclusion criteria were lack of consent, severe rotations or midline diastema suggesting the need for a bonded retainer, and patients with a restorative need in the labial segment, e.g. implant, bridges, or missing teeth.

Randomization

The subjects were randomly allocated to one of the two retention regimen groups using a restricted randomization technique, in blocks of 12, to ensure that equal numbers were allocated to each group. The allocation was decided by throwing an unweighted die where throws of 1, 2, or 3 = group 1 and 4, 5, or 6 = group 2. From this random list, the retention regimen was recorded alongside a patient identification number; the random allocation was sealed in numbered opaque envelopes and held in a central place. Thus, neither the clinician nor the patient knew their group allocation prior to consenting to the study.

Outcome measures

The outcome measures assessed were the upper and lower labial segment irregularity index (Little, 1975) and upper and lower labial segment crowding.

Appliance design and management

Upper Hawley retainer: Adams' cribs were placed on both upper first permanent molars and a long labial bow was taped and soldered to, and extended from, the bridges of these cribs. The anterior part of the labial bow was covered in acrylic to engage the embrasures between the incisors. The base plate was manufactured with acrylic that contacted the palatal surface of all teeth around the entire arch.

Lower Hawley retainer: Adams' cribs were placed on both lower first permanent molars. The labial bow was constructed to extend from the lower permanent canine to the canine and the labial aspect of the bow was covered in acrylic that engaged the embrasures between the incisors. The base plate was manufactured with acrylic that contacted the lingual surface of the teeth all the way round the arch.

All retainers, for all patients, were fitted 3 days post-debond and standardized instructions were given on appliance care.

Patient records. Study models were obtained at the following time intervals: T0, commencement of orthodontic treatment; T1, end of active orthodontic treatment (debond); and T2, end of the 1 year retention period. The patients were registered in this retention study at debond; therefore, T0 study models had previously been taken and were retrieved from the model store for measurements to be made.

Measurement methods. One author (MS), calibrated in the use of the digital dial callipers [Absolute Digimatic, Mitutoyo (Wednesbury, UK), Ltd.; www.jlindustrial.co.uk], measured the study models to an accuracy of 0.1 mm. To calculate incisor irregularity, the sum of the distances between the anatomic contact points from the mesial of the left canine to the mesial of the right canine in each labial segment was measured as described by Little (1975). These distances were summed to obtain a total irregularity index for the upper and lower labial segments.

Incisor crowding was calculated by measuring the difference between the sum, in millimetres, of the canine-to-canine tooth widths and the space in the labial segment from canine to canine. The available space in the labial segment was measured by dividing the labial segment into two straight-line segments, extending from the distal contact point to the canine on each side to the midpoint between the central incisors.

Method error

Systematic error. Although neither the operator nor the patient could be blinded to group allocation, the study model assessor (MS) was blind to the retention regimen used. Each

model was given a number and models were measured in random order, ensuring that no patient's stage models were measured consecutively.

Random error. Each model was measured twice for each outcome, and a mean value was calculated to reduce random error.

Error associated with the impression technique. Any potential error associated with the impression technique and model preparation has been investigated by O'Brien *et al.* (1990). This technique has been shown to have a 97 per cent coefficient of reliability.

Intra-examiner reliability. This was assessed by remeasuring 20 models after an interval of at least 1 week.

Statistical analysis

Data were entered into the Statistical Package for Social Sciences, version 11.0 (SPSS Inc., Chicago, Illinois, USA), checked for normality, and summary descriptive statistics produced. The difference between the retention regimen groups for both outcomes, and also examiner calibration, was evaluated using an independent sample *t*-test with a 0.05, two-sided significance level. Intraclass correlation coefficients (ICC) were used to assess intra-examiner reliability.

Results

A total of 67 subjects were entered into the trial from October 2003 to April 2004. During recruitment, four patients declined to take part; no reasons were given. In group 1, 41.2 per cent were males and 58.8 per cent were females and their mean age was 15.6 years (SD 1.6 years). In group 2, 24.2 per cent were males and 75.8 per cent were females and their mean age was 15.8 years (SD 1.2 years). The trial profile is shown in Figure 1.

Further descriptive statistics are summarized in Table 1 to show incisor classification and extraction/non-extraction according to retention group. Intra-examiner reliability for the calibrated examiner was high, with ICC as follows: Little's index/lower arch, 0.96; Little's index/upper arch, 0.95; labial segment crowding/lower arch, 0.90; and labial segment crowding/upper arch, 0.81.

Little's irregularity index

Both retention groups showed a considerable decrease in upper and lower labial segment irregularity from T0 to T1. This was followed by a trend to increasing irregularity from T1 to T2. Both groups, statistically, had equivalence at T0; however, clinically, slightly more lower labial segment irregularity was exhibited by group 2. Despite this, at T1 and T2, there was no statistically significant difference



Figure 1 Trial profile showing the numbers of patients registered and the reason for dropouts. *Patients did not attend appointments.

between the two retention regimens for the labial irregularity index (P > 0.05). The data for lower and upper labial segments are presented in Table 2.

Labial segment crowding

All patients showed a decrease in upper and lower labial segment crowding from T0 to T1 with a tendency for crowding to increase again from T1 to T2. This was particularly so in the lower arch, although there was slight or no change for this variable in the upper arch. There was no

 Table 1
 Incisor classification of groups 1 and 2 and extraction/ non-extraction treatment at T0.

	Group 1 (retainers night-time only for 1 year), <i>n</i> (%)	Group 2 (retainers 6 months full time followed by 6 months night only), <i>n</i> (%)
Incisor classification		
Class I	12 (36.4)	11 (32.4)
Class II division 1	13 (39.4)	13 (38.2)
Class II division 2	4 (12.1)	7 (20.6)
Class III	4 (12.1)	3 (8.8)
Extraction/non-extract	tion	
Lower arch		
Extraction	25 (76)	15 (44)
Non-extraction	8 (24)	19 (56)
Upper arch		
Extraction	29 (88)	26 (77)
Non-extraction	4 (12)	8 (23)

	Group 1: retainers night-time only for 1 year (SD)	Group 2: retainers 6 months full-time followed by 6 months night only (SD)	Mean difference	P value	95% confidence interval of the difference
Mean lower labial segment score					
Pre-treatment (T0)	6.1 (3.3)	7.8 (4.2)	1.64	0.09	-0.27 to 3.50
Debond (T1)	1.4 (0.8)	1.2 (0.5)	-0.17	0.33	-0.50 to 0.17
1 year post-debond (T2)	2.0 (1.0)	1.8 (0.7)	-0.26	0.39	-0.88 to 0.36
Mean upper labial segment score					
Pre-treatment (T0)	10.8 (5.3)	10.6 (4.1)	-0.15	0.89	-2.52 to 2.21
Debond (T1)	1.8 (0.8)	1.6 (0.7)	-0.24	0.23	-0.64 to 0.16
1 year post-debond (T2)	2.0 (0.7)	2.0 (0.8)	-0.02	0.94	-0.57 to 0.53

 Table 2
 Little's irregularity index (mm): lower and upper labial segment.

statistically significant difference between the two retention groups for lower or upper labial segment crowding at T2 (Table 3a,b). When T0 equivalence was considered, the upper labial segment retention groups exhibited T0 equivalence for crowding. However, in group 2, there was approximately 1 mm more lower labial segment crowding at T0 (P = 0.02). Although, statistically, ideal pre-treatment equivalence was not obtained, it is unlikely that a 1 mm difference would be clinically significant at baseline, particularly if the 1 mm of crowding was spread over several contact points.

Discussion

The results of this trial revealed that both retention regimens (night-only wear of an upper and lower Hawley retainer for 1 year or 6 months full-time followed by 6 months night-only wear) were equally effective during a 1 year retention period.

It is difficult to make comparisons with previous investigations due to the limited number of prospective studies investigating post-treatment tooth movement as well as the effects of retention regimen on stability. Comparison with the study of Destang and Kerr (2003) is difficult because of differences in study design, number of arches studied (maxillary arch only), and differences in retention protocols.

Maxillary and mandibular labial segment irregularity

Much of the literature related to Little's irregularity index has investigated the lower labial segment retrospectively. There are few prospective randomized clinical trials. The results of this present investigation show relatively small increases in the irregularity index from T1 to T2. This finding was also reported by Sadowsky *et al.* (1994), who found a mean increase in the irregularity index from post-treatment to post-retention of 1.4 mm for mandibular incisors and 1.1 mm for maxillary incisors. Any relapse in tooth position during the retention period may be explained either by the retainers allowing small movements or by the patients not complying with retainer wear.

Maxillary and mandibular labial segment crowding

The results of the present study show a relatively small amount of labial segment crowding at T2, in the region of 0.5 mm for both groups. Unfortunately, most of the previous literature has not evaluated this variable and so objective comparison is difficult. The current study revealed no statistically significant differences between the retention regimens in terms of labial segment crowding.

Clinical significance

The results of this research indicate that it may be possible for clinicians to advocate night-time use only of upper and lower Hawley retainers for 1 year as a retention regimen. Thus, where there was previous uncertainty, with many orthodontists opting for full-time initial wear for either 3 or 6 months, it would now seem acceptable for Hawley retainers to be worn at night only from the time of debond.

 Table 3
 Labial segment crowding (mm): lower and upper labial segment.

		Group 2: retainers 6 months full-time followed by 6 months night only (SD)	Mean difference	P value	95% confidence interval of the difference
	Group 1: retainers night-time only for 1 year (SD)				
Mean lower labial segment score					
Pre-treatment (T0)	2.3 (1.5)	3.4 (2.0)	1.08	0.02	0.17 to 1.98
Debond (T1)	0.7 (0.4)	0.7 (0.4)	0.04	0.68	-0.14 to 0.22
1 year post-debond (T2)	1.0 (0.5)	1.2 (0.5)	0.16	0.33	-0.17 to 0.50
Mean upper labial segment score		()			
Pre-treatment (T0)	2.6 (2.0)	3.3 (2.7)	0.73	0.23	-0.49 to 1.96
Debond (T1)	0.6(0.4)	0.7(0.5)	0.12	0.28	-0.09 to 0.33
1 year post-debond (T2)	0.7 (0.5)	0.8 (0.3)	0.03	0.82	-0.26 to 0.33

It does not, however, seem reasonable to extrapolate these findings to other types of removable retainers, particularly as Rowland *et al.* (2007) showed greater changes in incisor irregularity with a Hawley retainer compared with a vacuum-formed retainer when patients were followed up for 6 months. It is also important to consider that, presently, it is more unusual to use a lower Hawley retainer, while upper Hawley retainers are commonly used. However, in this study, it was considered important to use the same retainers in the upper and lower arches to minimize the confounding effect of retainer type.

Clinically, a 1 year retention period is the minimum advocated. Since, at the end of 1 year, patients were given the choice of either continuing or not with their Hawley retainers, the effect of full- versus part-time retainer wear on long-term stability could not be evaluated. It would further add to our clinical knowledge if long-term follow-up could be arranged but with the understanding of the difficulty of asking patients, who may want to continue wearing their retainers, to stop after 1 year.

A randomized clinical trial was conducted as the aim was to minimize any bias when comparing alternative clinical techniques. All potentially known and unknown confounding clinical factors that may affect the results have the best chance of being split evenly between the two retention groups. Therefore, the closest measure of true clinical performance is provided by clinical trials of this nature. While it would have been ideal for both operators and subjects to be blind to the retention regimen used, the difference in retention regimen did not lend itself to a double-blind design.

In any clinical trial, it is important that the characteristics of the dropouts of the study are taken into account. This was investigated by statistical comparison of the study model data for the two outcomes at T0 and T1, which showed that the dropouts were no different for labial segment irregularity or labial segment crowding compared with those patients who remained in the study (P = 0.35). Therefore, although there was some attrition of the sample, the loss of information was unlikely to bias the findings from the data of those patients remaining in the trial.

As the findings suggested that one retention regimen was not more effective than the other, it is necessary to consider the power of the study. When power was calculated, a meaningful difference between the retention regimens of 2 mm of labial segment irregularity or labial segment crowding was set. This may be considered high, but the intention was to reveal clinically important, not statistical, differences. Thus, it is considered that the study had sufficient power to reveal a meaningful difference between the two retention regimens.

One of the advantages of a randomized trial design is that the randomization process generally ensures that confounding variables such as start irregularity or crowding are equally divided between groups, so ensuring pretreatment equivalence. The data from this trial supported this concept except for lower labial segment crowding and irregularity. Group 2 had around 1 mm more crowding and nearly 2 mm more irregularity than group 1 at T0. It is difficult to explain why this occurred despite strict randomization and no statistical evidence of bias from dropouts. As a result, the data from this trial should be interpreted with caution.

Conclusions

No statistically significant differences were found between the two retention regimens, group 1 (night-only wear of upper and lower Hawley retainers for 1 year) versus group 2 (6 months full-time followed by 6 months night-only wear), in terms of upper and lower labial segment alignment and crowding. Therefore, clinicians could advise their patients to wear their retainers at night only from the time of debond, and a period of full-time retainer wear is not necessary.

Address for correspondence

Dr N. A. Mandall Orthodontic Department Tameside General Hospital Fountain Street Ashton-under-Lyne Lancashire OL6 9RW UK E-mail: nicky.mandall@tgh.nhs.uk

References

- Ades A G, Joondeph D R, Little R M, Chapko M K 1990 A long-term study of the relationship of third molars to changes in the mandibular dental arch. American Journal of Orthodontics and Dentofacial Orthopedics 97: 323–335
- Björk A, Skieller V 1972 Facial development and tooth eruption: an implant study at the age of puberty. American Journal of Orthodontics 62: 88–97
- Destang D L, Kerr W J S 2003 Maxillary retention: is longer better? European Journal of Orthodontics 25: 65–69
- Felton J M 1987 A computer analysis of the shape and stability of mandibular arch form. American Journal of Orthodontics and Dentofacial Orthopedics 92: 478–483
- Harradine N W, Pearson M H, Toth B 1998 The effects of extraction of third molars on late lower incisor crowding: a randomized controlled trial. British Journal of Orthodontics 25: 117–122
- Little R M 1975 The irregularity index: a quantitative score of mandibular anterior alignment. American Journal of Orthodontics 68: 249–257
- Little R M, Wallen T R, Riedel R A 1981 Stability and relapse of mandibular anterior alignment—first premolar extraction cases treated by traditional edgewise orthodontics. American Journal of Orthodontics 80: 349–364
- Little R M, Riedel R A, Stein A 1990 Mandibular arch length increase during the mixed dentition: post retention evaluation of stability and relapse. American Journal of Orthodontics and Dentofacial Orthopedics 97: 393–404
- Littlewood S J, Millett D T, Doubleday B, Bearn D R, Worthington H V 2006 Orthodontic retention: a systematic review. Journal of Orthodontics 33: 205–212
- Lopez-Gavito G, Wallen T R, Little R M, Joondeph D R 1985 Anterior open bite malocclusion: a longitudinal 10-years post retention evaluation of orthodontically treated patients. American Journal of Orthodontics 87: 175–186

- Mills J R E 1968 The stability of the lower labial segment. Dental Practitioner 18: 293–305
- O'Brien K, Lewis D, Shaw W, Combe E 1990 A clinical trial of aligning archwires. European Journal of Orthodontics 12: 380–384
- Peck H, Peck S 1972 Crown dimensions and mandibular incisor alignment. American Journal of Orthodontics 62: 384–401
- Reitan K 1959 Tissue rearrangement during retention of orthodontically rotated teeth. Angle Orthodontist 29: 105–113
- Reitan K 1967 Clinical and histologic observation on tooth movement during and after orthodontic treatment. American Journal of Orthodontics 53: 721–790
- Richardson M E 1979 Late lower arch crowding: facial growth or forward drift. European Journal of Orthodontics 1: 219–225
- Richardson M E 1989 The role of third molar in the cause of late lower arch crowding: a review. American Journal of Orthodontics and Dentofacial Orthopedics 95: 79–83

- Richardson M E 1996 Late lower arch crowding in the relation to direction of eruption. European Journal of Orthodontics 18: 341–347
- Rowland H *et al.* 2007 The effectiveness of Hawley and vacuum-formed retainers: a single-center randomized controlled trial. American Journal of Orthodontics and Dentofacial Orthopedics 132: 730–737
- Sadowsky C, Schneider B J, BeGole E A, Tahir E 1994 Long term stability after orthodontic treatment: non-extraction with prolonged retention. American Journal of Orthodontics and Dentofacial Orthopedics 195: 243–249
- Southard T, Southard K A, Tolley E A 1992 Periodontal force: a potential cause of relapse. American Journal of Orthodontics and Dentofacial Orthopedics 101: 221–227
- Wieslander L 1993 Long-term effect of treatment with the headgear-Herbst appliance in the early mixed dentition: stability or relapse? American Journal of Orthodontics and Dentofacial Orthopedics 104: 319–323

Copyright of European Journal of Orthodontics is the property of Oxford University Press / UK and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.