**ABSTRACT**

**Objective:** To evaluate morphologic stability and patient satisfaction at least 5 years after orthodontic treatment.

**Materials and Methods:** Published literature was searched through the PubMed and Cochrane Library electronic databases from 1966 to January 2005. The search was performed by an information specialist at the Swedish Council on Technology Assessment in Health Care. The inclusion criteria consisted of a follow-up period of at least 5 years postretention; randomized clinical trials, prospective or retrospective clinical controlled studies, and cohort studies; and orthodontic treatment including fixed or removable appliances, selective grinding, or extractions. Two reviewers extracted the data independently and also assessed the quality of the studies.

**Results:** The search strategy resulted in 1004 abstracts or full-text articles, of which 38 met the inclusion criteria. Treatment of crowding resulted in successful dental alignment. However, the mandibular arch length and width gradually decreased, and crowding of the lower anterior teeth recurred postretention. This condition was unpredictable at the individual level (limited evidence). Treatment of Angle Class II division 1 malocclusion with Herbst appliance normalized the occlusion. Relapse occurred but could not be predicted at the individual level (limited evidence). The scientific evidence was insufficient for conclusions on treatment of cross-bite, Angle Class III, open bite, and various other malocclusions as well as on patient satisfaction in a long-term perspective.

**Conclusions:** This review has exposed the difficulties in drawing meaningful evidence-based conclusions often because of the inherent problems of retrospective and uncontrolled study design.

**KEY WORDS:** Long-term stability; Patient satisfaction; Orthodontic treatment results; Systematic review

---

**INTRODUCTION**

The maintenance of dental alignment after orthodontic treatment has been and continues to be a challenge to the orthodontic profession. Usually, the goal of orthodontic treatment is to produce a normal or so-called ideal occlusion that is morphologically stable and esthetically and functionally well adjusted. There is, however, a large variation in treatment outcome because of the severity and type of malocclusion, treatment approach, patient cooperation, and growth and adaptability of the hard and soft tissues. Follow-up studies of treated cases have shown that although “ideal” occlusion and dental alignment have been achieved, there is a tendency for relapse toward the original malocclusion posttreatment.1-4 Long-term stability of orthodontic treatment results has to be considered in relation to aging, periodontal disease, caries, and various types of dental restorations. With these factors in mind, and in relation to the duration, effort,
Table 1. Search Strategy PubMed 1966 to January 2005a

<table>
<thead>
<tr>
<th>MeSH-terms (Medical Subject Headings)</th>
<th>AND</th>
<th>RCT</th>
<th>CCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthodontics Malocclusion/therapy, surgery</td>
<td>Orthodontics Malocclusion/therapy, surgery</td>
<td>Comparative study AND treatment outcome</td>
<td>Follow-up studies</td>
</tr>
<tr>
<td>Follow-up studies</td>
<td>Longitudinal studies</td>
<td>Evaluation study</td>
<td></td>
</tr>
<tr>
<td>Treatment outcome</td>
<td>Recurrence</td>
<td>Cephalometry</td>
<td></td>
</tr>
<tr>
<td>Consumer satisfaction</td>
<td>Longitudinal/Ti</td>
<td>Follow-up/Ti</td>
<td></td>
</tr>
<tr>
<td>Long-term/Ti</td>
<td>Postretention/Ti</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stability/Ti</td>
<td>Evaluation/Ti</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Limits: All child 0–18 y | Human | English | Swedish | Norwegian | Danish | Finnish |

/ indicates subheading; Ti, in title.

and cost invested in orthodontic therapy, the choice of a follow-up period of at least 5 years after completed retention seems reasonable when stability of orthodontic treatment is evaluated.1,5

When evaluating postretention changes, it is of vital importance to take into account the natural growth changes seen in individuals who have had no orthodontic treatment. Moreover, because orthodontic treatment improves facial and dental appearance, assessment of the long-term outcome of orthodontic treatment should also include patient satisfaction with respect to dental and facial appearance in treated as well as in untreated groups. It is therefore appropriate and advisable to use matched untreated control groups.

To date, several studies have been published concerning long-term stability of orthodontic treatment, and a systematic review of the present knowledge is motivated. In 2002, the Swedish Council on Technology Assessment in Health Care (SBU) commissioned a project group to undertake a systematic review of malocclusions and orthodontic treatment in an oral health perspective. The aim of this part of the larger systematic review was to evaluate morphologic stability and patient satisfaction at least 5 years after orthodontic treatment.

MATERIALS AND METHODS

Search Strategy

To identify all the studies that examined morphologic stability at least 5 years after orthodontic treatment, a literature survey was performed by applying the PubMed and Cochrane Library databases from 1966 to January 2005. An information specialist at SBU performed the search for specific search fields by preview or index in PubMed. The search strategy is presented in Table 1.

A total of 1004 studies (abstracts or full-text articles if abstracts were missing) were identified and printed. Relevant studies were selected independently by two of the authors (Drs Bondemark and Holm). Case reports, review articles, letters, editorials, gray literature, papers describing surgical or cleft lip or palatal treatment, and obviously irrelevant literature were excluded. A study was ordered in full text if at least one of the two reviewers considered it to be potentially relevant. Reference lists of the studies were hand searched for additional relevant studies not found in the database search. Another 19 studies were identified by this method. Human studies in Swedish, Danish, Norwegian, Finnish, and English were considered. For studies in which more than one study had been published on the same material at different times, only...
the last publication based on the longest follow-up period was included. The two independent reviewers assessed all the studies with respect to the inclusion criteria, and any interexaminer conflicts were resolved by discussion of each study to reach consensus. A total of 109 studies were thus selected. Of these, 38 were included and 71 were excluded (Table 2) according to the following inclusion criteria:

- Randomized clinical trials, prospective or retrospective clinical controlled studies, and cohort studies
- A follow-up period of at least 5 years postretention
- Orthodontic treatment including fixed or removable appliances, selective grinding, or extractions

**Endpoints**

Studies using endpoints relevant to the patient (ie, a treatment result that was functionally or esthetically satisfying) were nonexistent. Therefore, measurements of tooth positions and relations between upper and lower jaws made before and after treatment, such as using measurements on dental casts or cephalometric measurements or using the Peer Assessment Rating (PAR) index or Little’s index to assess lower incisor positions, were accepted as endpoints.

**Evaluation of Studies and Level of Evidence**

The two reviewers used a data extraction form to independently read the 38 remaining studies. The external and internal validity as well as the quality of methodology, statistics, and performance of each study were assessed, and the studies were graded with a score of A to C according to predetermined criteria (Table 3). In the event of disagreement between the two reviewers, the study was discussed within the whole group to reach consensus. Based on the evaluated studies, the final level of evidence for each conclusion was judged according to the protocol of the SBU (Table 4), which is based on the criteria for assessing study quality from Centre for Reviews and Disseminations in York, UK.

**RESULTS**

Thirty-eight studies met the inclusion criteria. Of these, 20 were graded as moderate value of evidence (grade B) and served as a basis for the conclusions,3,4,71,78–94 18 were graded as low value of evidence (grade C) (Table 5), and none were graded as a high value of evidence (grade A).

**Treatment of Crowding**

Long-term morphologic stability after treatment of crowding was studied in 12 studies: eight studies3,78–84 were graded B (Table 6), whereas four12,96,108,109 were graded C. All eight studies graded B were retrospective, and seven used control or reference groups. One
Table 6. Treatment of Crowding

<table>
<thead>
<tr>
<th>Author, Year, and Country</th>
<th>Population, Number, % Female, and Age</th>
<th>Orthodontic Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ades et al,78 1990, United States</td>
<td>University clinic, T1: 32, with third molars, T2: 17, third molars in retention, T3: 17, no third molars, T4: 34, third molars extracted, female unknown, 15 y</td>
<td>T1 extraction/no extraction. T2 extraction/no extraction. T3 extraction/no extraction. T4 extraction/no extraction. Fixed appliance in all groups.</td>
</tr>
<tr>
<td>Little et al,80 1990, United States</td>
<td>University clinic T1: 30, T2: 30, 80%, 8–12 y</td>
<td>T1: mixed dentition, early treatment and expectancy, later fixed appliance. T2: permanent dentition, late extraction, and fixed appliance.</td>
</tr>
<tr>
<td>McReynolds and Little,7 1991, United States</td>
<td>University clinic, T1: 14 (early extraction), T2: 32 (late extraction), T1: 85%, T2: 55%, T1: 11.6 y, T2: 12.6 y</td>
<td>T1: mixed dentition, second premolar extraction and expectancy, later fixed appliance. T2: permanent dentition, second premolar extraction, and fixed appliance.</td>
</tr>
<tr>
<td>Surbeck et al,84 1998, United States</td>
<td>University clinic, T1: 30 (spaced teeth at follow-up), 27%, 13 y, T3: 49 (with crowding at follow-up), 70%, 12.8 y, T3: 28 (with well aligned teeth at follow-up), 61%, 12.6 y</td>
<td>Fixed appliance with and without extractions.</td>
</tr>
</tbody>
</table>

study was a cohort study.83 The follow-up period varied between 7 and 20 years, and fixed or removable appliances had been used in all but one study.82 Most of the studies focused on morphologic changes in the lower jaw, particularly the position of the mandibular incisors.3,78–80 Only two studies analyzed changes in the maxilla.81,84

It was evident from the studies that the arch length and intercanine width of the mandible were reduced and crowding of the mandibular incisors frequently re-occurred during the follow-up period. The studies showed large individual variations, and parameters such as gender, age at start of orthodontic treatment, initial diagnosis, extraction or nonextraction as part of the treatment, the length of the treatment and retention periods, and the presence or absence of third molars could not be used to predict stability changes in the mandible. Many authors considered lifelong bonded mandibular canine retainers to be the only way to obtain a result of the treatment that is morphologically stable.

Treatment of Angle Class II Malocclusion

In 15 studies, the long-term effect after treatment of Angle Class II malocclusion was evaluated, and nine were graded C97–102,110–112 (Table 5). Six studies were graded B85–90 (Table 7), and in these the follow-up period varied between 5 and 11 years. In one study90 the treatment was performed with an Andresen activator, and in all other studies the Herbst appliance was used. Two studies had control groups (historical controls).87,88 Two other studies used subgroups where stable and nonstable cases were compared,85,89 and one compared treatment outcome before, during, and after pubertal growth maximum.86 The studies concluded that the Herbst appliance normalized dentition and occlusion. Relapse occurred but could not be predicted at the individual level.
Table 6. Extended

<table>
<thead>
<tr>
<th>Follow-up After Retention, Y</th>
<th>Endpoints</th>
<th>Results After Treatment</th>
<th>Results After Retention Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>Dental cast, and cephalometric data.</td>
<td>Crowding and horizontal and vertical overbite normalized.</td>
<td>Renewed crowding of lower front teeth. No difference with or without third molars.</td>
</tr>
<tr>
<td>15</td>
<td>Dental cast analysis.</td>
<td>An acceptable alignment in both groups.</td>
<td>Moderate renewed crowding of lower front teeth. Less relapse after early treatment.</td>
</tr>
<tr>
<td>15</td>
<td>Dental cast and cephalometric data.</td>
<td>Clinically satisfactory results. Crowding adjusted.</td>
<td>Renewed crowing in lower jaw, intercanine distance and arch length decrease. No difference between early and late extraction.</td>
</tr>
<tr>
<td>7</td>
<td>Cephalometric data.</td>
<td>Molars tipped distally and showed downward-backward translation.</td>
<td>Total relapse, at follow-up no difference in position of molars between test and control groups.</td>
</tr>
<tr>
<td>20</td>
<td>Dental cast and cephalometric data.</td>
<td>Effect directly after treatment not analyzed.</td>
<td>Significant alignment of teeth and space closure after extractions. No effect on overjet or overbite. In spite of extractions, crowding in lower jaw as in the control group.</td>
</tr>
<tr>
<td>15.7</td>
<td>Dental cast analysis.</td>
<td>Clinically satisfactory results.</td>
<td>Renewed crowding of incisors in both jaws occurred and more in the mandible. Contributing factors: large teeth, pronounced crowding, short and thin arches, expansion of dental arches, persistent Class II and III molar relation after treatment.</td>
</tr>
<tr>
<td>12–14</td>
<td>Dental cast analysis.</td>
<td>Clinically satisfactory results with alignment of upper front teeth.</td>
<td>Contact point displacement, maxillary incisor rotations and spacing are significant risk factors for postretention relapse of alignment.</td>
</tr>
</tbody>
</table>

* All studies were retrospective, and all were graded B (moderate value of evidence).

Treatment of Cross-bite

Treatment of unilateral cross-bite was investigated in three studies. One study103 was graded C (Table 5) and two were graded B.91,92 In one randomized study91 it was shown that 79% of forced unilateral cross-bites in the primary dentition could be corrected with selective grinding, and after 5 years 50% were still stable. In the control group, spontaneous correction was seen in 17%. In a retrospective study,92 the majority of cases treated in the mixed dentition with quadhelix and expansion plate were stable after a long-term follow-up. The studies were too few for any evidence-based conclusions.

Treatment of Angle Class III Malocclusion

In two studies, long-term follow-up of treatment of Angle Class III was analyzed. One of the studies105 was graded C (Table 5) and the other93 was graded B. This study showed that treatment with rapid maxillary expansion and face mask could, in the long-term (5½ years follow-up), correct Angle Class III malocclusions in 8-year-old children.93 No evidence-based conclusions could be drawn.

Treatment of Open Bite

The long-term effect of treatment of open bite was assessed in two studies. One study106 was graded C (Table 5) and the other94 was graded B. In this retrospective study it was shown that a fixed appliance after extraction of premolars could normalize the open bite (at least 1 mm) in young teenagers and that 74% had a clinically stable open-bite correction after 8½ years. No evidence-based conclusions were possible.
### Table 7. Treatment of Angle Class II Malocclusion

<table>
<thead>
<tr>
<th>Author, Year, and Country</th>
<th>Population, Number, % Female, and Age</th>
<th>Orthodontic Treatment</th>
<th>Follow-up After Retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pancherz, 1991, Sweden</td>
<td>Consecutive cases grouped in, T1: 14 stable, T2: 15 instable, 24%, 12.7 y</td>
<td>Herbst</td>
<td>5–10</td>
</tr>
<tr>
<td>Hansen et al</td>
<td>Consecutive cases grouped in, T1 x prepeak, T2: y peak, T3: z postpeak, 0%, T1: 12.2 y, T2: 12.9 y, T3: 14.2 y</td>
<td>Herbst</td>
<td>6.6</td>
</tr>
<tr>
<td>Hansen and Pancherz, 1992, Sweden</td>
<td>Consecutive cases, T: 32, K: 32 (historical controls), 50%12.5 y</td>
<td>Herbst</td>
<td>5.7</td>
</tr>
<tr>
<td>Pancherz and Anehus-Pancherz, 1993, Sweden</td>
<td>Consecutive cases, T: 45, K: 32 (historical controls), 24%, 12.4</td>
<td>Herbst</td>
<td>6.4</td>
</tr>
<tr>
<td>Pancherz and Anehus-Pancherz, 1994, Sweden</td>
<td>T1: 49 stable cases, T2: 20 unstable cases, female unknown, 12.7 y</td>
<td>Herbst</td>
<td>5–10</td>
</tr>
<tr>
<td>Pancherz, 1976, Sweden</td>
<td>University clinic, T1: 64 with extractions, 63%, 11 y, T2: 45 without extractions, 49%, 11 y</td>
<td>Andresen activator</td>
<td>11</td>
</tr>
</tbody>
</table>

*All studies were retrospective, and all were graded B (moderate value of evidence).*

### Treatment of Various Other Malocclusions

Three studies were identified where the long-term effect of orthodontic treatment of various types of malocclusions were investigated. Two studies were graded C (Table 5). The third study was a cohort study graded B and used the PAR index for analysis of morphologic changes of occlusion. This study reported 45% of the achieved orthodontic treatment results to be stable 10 years postretention. No evidence-based conclusions could be drawn.

### Patient Satisfaction

Two studies were excluded because the follow-up period was less than 5 years (Table 2). Two other studies were based on the same patient material; one was thus excluded and the other was graded B. No evidence-based conclusions were possible.

### DISCUSSION

This systematic review of the scientific literature dealing with long-term stability after orthodontic treatment resulted in some consistent findings. After treatment of crowding, a continuous decrease in dental arch length and intercanine width of the mandible resulted in reoccurrence of crowded anterior teeth postretention. The degree of individual misalignment was not predictable. Treatment of Angle Class II division 1 malocclusion with the Herbst appliance resulted mainly in normal occlusion. Relapse occurred but could not be predicted at the individual level. The scientific evidence was too weak for conclusions regarding specific occlusal traits such as open bite, Angle Class III malocclusion, and unilateral posterior cross-bite.

In view of the present knowledge, it was impossible to identify if relapse posttreatment was the result of orthodontic treatment alone or of physiologic changes in the dentition and surrounding tissues during the follow-up period. It has been shown that craniofacial alterations occur in adults and are accompanied by compensatory changes in the dentition. To evaluate the relapse, where several factors may act at different time intervals together with natural craniofacial alterations and compensatory changes in the dentition, the researchers have to focus on and use prospective well-designed follow-up studies with untreated controls. Efforts should be made to avoid bias by using well-defined and sufficiently large samples.

Today, the systematic literature search, data extraction, and subsequent quality assessment of included studies are well-established measures in evidence-based medicine and dentistry. However, the precise methods for the process can differ among various systematic reviews. The methodology used in this review was adopted from the guidelines of the SBU. Many studies were excluded primarily because of a lack of control group, a large number of dropouts, or a follow-up period of less than 5 years postretention. Other excluded studies were those dealing with treatment of crowding where sample selection had not been based on type of malocclusion, which resulted in samples with a wide variation in skeletal relationship and growth pattern.

The restrictions concerning the number of databas-
Table 7. Extended

<table>
<thead>
<tr>
<th>Endpoints</th>
<th>Results After Treatment</th>
<th>Results After Retention Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cephalometric data.</td>
<td>Correction of sagittal relation, overjet, and overbite.</td>
<td>Some relapse of sagittal relation and overjet because of unstable occlusion, lip, and tongue thrust.</td>
</tr>
<tr>
<td>Cephalometric data.</td>
<td>In all groups, good dental and skeletal correction.</td>
<td>In the long-term, no difference if treatment performed prepeak, peak, or postpeak. It is advisable to treat postpeak because less growth is left and it is easier to establish good occlusion in the permanent dentition.</td>
</tr>
<tr>
<td>Cephalometric data.</td>
<td>At short sight, good dental and skeletal correction.</td>
<td>In long-term, normalization of the occlusion, whereas the skeletal relation improves but does not normalize.</td>
</tr>
<tr>
<td>Cephalometric data.</td>
<td>Maxillary molars were moved distally and intruded. The occlusal plane tilted backward as a consequence of the Herbst treatment.</td>
<td>The molar movement was temporary, like other movements. Retention is necessary.</td>
</tr>
<tr>
<td>Cephalometric data.</td>
<td>Reduction of the hard and soft tissue profile convexity.</td>
<td>Mostly stable results but unpredictable relapse at the individual level.</td>
</tr>
<tr>
<td>Dental cast analysis.</td>
<td>Overjet, overbite, and sagittal relation improved.</td>
<td>Difficult to separate relapse from normal growth changes. Better dentoalveolar stability after treatment without than with extractions. In both groups, minimal relapse in sagittal relation and more relapse in vertical than in horizontal dimension. Crowding in the mandible increases in both groups.</td>
</tr>
</tbody>
</table>

es and languages when searching the literature might imply that some studies were not identified. Studies that are difficult to find are, however, often of lower quality. The strength of the evidence in a systematic review is probably more dependent on assessing the quality of the included studies than on the degree of comprehensiveness.115

A notable finding was that none of the selected studies were graded A (high value of evidence) and only one randomized controlled trial was identified. Instead, a majority of the studies had a retrospective design; from an evidence-based point of view, the scientific value of a retrospective study is limited. Some authors have argued that well-designed prospective or retrospective studies should not be ignored when assessing scientific literature.116 Nevertheless, it should be emphasized that the randomized controlled trial is the most powerful tool to evaluate treatment, and the quality of the trial significantly affects the validity of the conclusions.

Overall, randomized controlled trials have been rarely used in orthodontics.117 One reason might be the practical difficulty to identify many patients with a certain malocclusion. Other important reasons could be ethical or logistic, for patients in randomized controlled trials do not have the right to influence the choice of treatment, or they may be designated to an untreated control group in which the treatment is postponed during the study period and, as a consequence, refuse to participate in the trial. In addition, in a long-term study there is always a risk of dropouts. Thus, performing randomized controlled trials demands an enthusiastic research team, well-motivated patients and parents, and in many cases sufficient financial resources.

On the other hand, the fact that evidence for a method’s efficacy is limited does not necessarily imply that it is ineffective or should not be used. For some treatments, the caregivers might have to accept that high levels of evidence could not be obtained.

The quality of treatment outcome has traditionally been assessed by applying professionally established metric or categorical scales with measurements obtained from dental casts, radiographs, and clinical examinations. As health services exist primarily to benefit the patient, an important variable for measuring outcome would be overall patient satisfaction with the care provided.118 Therefore, it was astonishing that only a few studies were found on patient satisfaction in the long-term, and furthermore most of them showed low scientific evidence and no conclusions could be drawn. This review of the literature has thus exposed a great need for future studies in this area.

CONCLUSIONS

• Treatment of crowding resulted in successful dental alignment. However, the mandibular dental arch length and intercanine width gradually decreased in the long-term, and crowding of the lower anterior teeth reoccurred postretention. This condition was unpredictable at the individual level (evidence-level 3).
• Treatment of Angle Class II division 1 malocclusion with a Herbst appliance normalized the occlusion.
Relapse occurred but could not be predicted at the individual level (evidence-level 3).

- The scientific evidence was insufficient for conclusions on long-term stability after treatment of cross-bite, Angle Class III, open bite, and various other malocclusions as well as on patient satisfaction in the long-term after orthodontic treatment.

- Despite a large number of studies on long-term stability after orthodontic treatment, this systematic review has shown that evidence-based conclusions were few. This was most often due to inherent problems with retrospective and inferior study design. There is a great need for well-designed prospective studies with untreated control groups; sufficient sample sizes; and sample selection according to type of malocclusion, age, and growth pattern.

REFERENCES


52. Janson G, Valarelli FP, Henriques JF, de Freitas MR, Can-


105. Yoshida I, Ishii H, Yamaguchi N, Mizoguchi I. Maxillary protraction and chincap appliance treatment effects and...


