

SCIENTIFIC
SECTION

Development and use of a taxonomy to carry out a systematic review of the literature on methods described to effect distal movement of maxillary molars

G. J. Atherton

Halifax, UK

A.-M. Glenny and K. O'Brien

University Dental Hospital of Manchester, UK

Abstract

Objectives To devise a taxonomy for the assessment of the orthodontic literature on methods described to effect distal movement of maxillary molars, to test the taxonomy for inter-assessor reliability, and to use it to classify studies in a systematic review of the literature.

Data sources Articles appearing in the *American Journal of Orthodontics and Dentofacial Orthopedics*, *Angle Orthodontist*, *(British) Journal of Orthodontics*, *European Journal of Orthodontics*, and the *Journal of Clinical Orthodontics* between 1988 and 1998.

Data selection Articles describing or evaluating the effect of appliances known or thought to have a distalizing effect on maxillary molars.

Data extraction A taxonomy was designed, tested by two reviewers independently to assess levels of agreement, and then used to record the features of the articles in a systematic review of the literature.

Data synthesis Kappa scores were used to assess the level of agreement between reviewers and found to be satisfactory. Studies were grouped according to study design and features of their methodology quantified.

Conclusions Having devised and tested the taxonomy, we found that the quality of evidence for any method of moving maxillary molars distally was not high.

Index words:

distal movement
maxillary molars
systematic review

Received 9 April 2001; accepted: 14 March 2002

Introduction

Currently much emphasis is being placed on dental and medical practice based upon sound evidence. An integral part of assessing evidence for a given intervention is an evaluation of the quality of research providing that evidence. It is in the light of this that our aim in this study was to conduct a systematic review of the orthodontic literature, published over an 11-year period, on methods described to move maxillary molars distally and assess the quality of evidence supporting the techniques described.

One of the primary goals of orthodontic treatment is

attainment of an 'ideal occlusion' which was first defined by Angle¹ and later refined and to an extent redefined by Andrews.² An integral part of achieving this involves placing the first molars in a Class I relationship. Class II malocclusions are among the most common presenting for orthodontic treatment and are often accompanied by a 'pre-normal' molar relationship. Whether to correct this relationship is an important decision in treatment planning for this group of patients. A Class I molar relationship can be achieved in a number of ways, some of which involve distal movement of the maxillary first molars.

Techniques advocated for distal movement of the

maxillary first molars include extra-oral traction, used either alone or in combination with fixed or removable appliances, fixed or removable appliances on their own, and fixed intra-oral devices. The last group has become popular due to an increased interest in 'non-compliance' treatment and most, such as the Jones Jig,³ rely on a Nance button attached to banded first premolars for anchorage. Functional appliances may also be used to correct a Class II molar relationship. However, with many such appliances and with 'Class II' inter-arch elastics, there is a distal force applied to the maxillary arch that could result in distal movement of the molars.

Appraisal of the literature

The process of critical appraisal involves assessing and interpreting evidence from research by considering a study's validity, relevance and results. Only by a full assessment of the methodology, to ensure that the findings of the research are valid, is it possible to identify interventions that are truly effective.

An important preliminary stage in assessment of the literature is to grade it according to a hierarchy of strength of evidence based on the designs of the studies. The highest level of evidence for a given procedure is derived from a well conducted systematic review of good quality randomized controlled trials. Systematic reviews can establish the 'generalizability' of an intervention, i.e. the degree to which the results of a study hold true for situations other than those pertaining in the study. They can also increase the power of research findings particularly if there are a number of studies that are homogeneous enough to allow pooling of data by meta-analysis.

Experimental studies provide the next best source of evidence. They aim to standardize conditions under which the effect of a particular intervention is observed and recorded, and are conducted prospectively. They are considered to provide better evidence than observational studies, because the intervention and treatment conditions are under the control of the investigator. The best type of experimental study, theoretically least open to bias, is a randomized controlled trial (RCT), in which subjects are allocated by a random process to different treatment groups. The Consolidated Standards of Reporting Trials (CONSORT) statement⁴ lists 21 items that should be included in the report of an RCT to ensure that the reader is provided with enough information to make a judgement about the reliability of the findings.

Further down the hierarchy are observational studies, such as cohort and case-control studies. Cohort studies identify groups of people according to whether or not they have been exposed to a particular intervention/factor or not. These groups are then followed forward in time to measure the development of different outcomes. A case-control study identifies groups of participants according to whether or not they have the outcome of interest. The researcher then looks back in time to establish exposure status.

When examining the results from the literature, greater credence should be given to studies of a high level of evidence. However, within each level of the hierarchy, there is likely to be variation in the quality of the studies and it is therefore necessary to assess certain features for each study design. Guidelines are available for how to develop validity checklists (or *taxonomies*) for assessing the quality of the research literature. In orthodontics, taxonomies have been used to assess papers appearing in the orthodontic literature,⁵ evaluating treatment of cross-bites,⁶ and to examine correlations between overjet and traumatic dental injuries⁷.

No critical evaluation has been carried out to date to determine the most effective method or methods of moving maxillary molars distally. We proposed, therefore, to assess the strength of evidence found in the literature for this type of tooth movement.

Aims of the study

The aims of the study were to:

- devise a taxonomy for a systematic assessment of papers appearing in the orthodontic literature describing or measuring distal movement of the maxillary molars;
- test the taxonomy for inter-assessor agreement;
- when a satisfactory level of agreement had been achieved, use it to evaluate studies included in a systematic review of the literature on methods described to effect this tooth movement;
- pool data from different studies if appropriate, to enhance the strength of evidence for any treatment modality.

Method

Our initial step was to develop criteria for a hand search of the relevant literature. We did this by carrying out an

electronic search with *Medline*, using ‘molar distalization’ as a search term. This generated a number of headings, which we combined in pairs. The studies identified were assessed using the following criteria:

- articles which described or measured distal movement of maxillary molars;
- articles describing or evaluating the effect of appliances known or thought to have a distalizing effect on maxillary molars.

We defined distal movement to be where, during the intervention, the maxillary molars had been moved distally relative to a vertical reference line.

It was necessary to devise a taxonomy to classify the papers. An initial taxonomy was developed from assessment of features of the papers themselves and those of a well conducted RCT. This was tested for inter-examiner agreement between two reviewers (GJA and A-MG) on a random sample of 30 papers from those identified in the hand search. This led to further discussion and development to reduce areas of ambiguity in interpretation and improve its objectivity. The final version of the taxonomy was then tested on the same sample of 30 hand searched articles and on a sample of 30 articles from the *Journal of Clinical Periodontology*, as this is used by the Cochrane Oral Health Group as a test journal for all new hand searches. Levels of agreement for the classification of papers were tested using the Kappa statistic.

Once the exact format of the taxonomy, detailed as Appendix 1 (which is included on the web version of this paper <http://ortho.oupjournals.org/>), had been finalized, we conducted a hand search of the following journals, published between 1988 and 1998:

- *American Journal of Orthodontics and Dentofacial Orthopedics*;
- *Angle Orthodontist*;
- *British Journal of Orthodontics*;
- *European Journal of Orthodontics*;
- *Journal of Clinical Orthodontics*.

The title, author, journal, volume and year of publication of all relevant papers were recorded and each article given an identification number. Details of the author and, where possible, the journal were removed from a photocopy of the article. The taxonomy was then applied and used to provide an indication of the quality of evidence available for the techniques of molar distalization.

Results

The taxonomy

Kappa scores and levels of agreement between the two reviewers for the items of the taxonomy are shown in Table 1. When assessing the level of agreement in use of the taxonomy between two reviewers for the sample of the hand searched papers, there was ‘good’ or ‘very good’ agreement in 11 categories (69 per cent) and ‘moderate’ agreement in the other five (31 per cent). When assessing the articles in the *Journal of Clinical Periodontology*, there was ‘good’ or ‘very good’ agreement in eight categories (50 per cent), ‘moderate’ agreement in four (25 per cent) and ‘fair’ agreement in four. We concluded that the final version of the taxonomy provided a sufficient level of agreement to justify its use in assessment of articles for a systematic review.

Assessment of hand searched articles

Classification of articles. Table 2 classifies the 105 articles identified that included some discussion or assessment of maxillary molar movement.

Study design. Table 3 classifies the studies according to design. Fifty-eight studies offered a higher level of evidence than a case report. The most common study design used to assess distal movement of maxillary molars was a cohort study, of which there were 23 (40 per cent). However none appeared to be prospective. In addition, there were 18 case series (31 per cent), 11 clinical trials (19 per cent) and six of what we have called ‘comparisons of case series’ (10 per cent), where records of groups of patients who were different at the outset were compared. Of the clinical trials, eight (14 per cent) were classed as controlled clinical trials (CCTs) and three (5 per cent) as RCTs—where it stated in the method that there was random allocation of subjects. In 22 studies, it was not clear whether they were retrospective or prospective.

Appendix 2 (<http://ortho.oupjournals.org/>) lists the study designs used in these articles, the quality of inclusion or exclusion criteria, the controls used, whether pre-treatment equivalence was assessed, whether the groups were equivalent at the outset of the studies, the amount of mean distal movement of the molars recorded, the appliances assessed and whether an untreated control was used.

None of the three RCTs, no cohort studies, only three of the eight CCTs and nine case series set out specifically

Table 1 Kappa scores measuring levels of agreement between two reviewers in use of the taxonomy to assess the listed features of a sample of 30 of the papers from the hand search and 30 papers from the *Journal of Clinical Periodontology*

Feature	Hand searched articles		<i>Journal of Clinical Periodontology</i>	
	Kappa value	Level of agreement	Kappa value	Level of agreement
Article type	0.420	Moderate	0.510	Moderate
Study type	0.407	Moderate	0.630	Good
Study design	0.460	Moderate	0.667	Good
Inclusion criteria	0.792	Good	0.518	Moderate
Exclusion criteria	1.000	Very good	0.615	Good
Sample (1)	0.815	Very good	0.649	Good
Sample (2)	0.745	Good	0.433	Moderate
Sample (3)	0.576	Moderate	0.391	Fair
Untreated control	0.845	Very good	0.389	Fair
Power considerations	0.967	Very good	0.672	Good
Random allocation	0.783	Good	0.911	Very good
Control	0.829	Very good	0.704	Good
Pre-treatment equivalence	0.605	Good	0.421	Moderate
Groups equivalent?	0.542	Moderate	0.373	Fair
Method error	0.683	Good	0.261	Fair
Duration of study	0.626	Good	0.815	Very good

Table 2 Classification of articles discussing or assessing maxillary molar movement

	Number (per cent)
Observational	73 (70)
Experimental	23 (22)
<i>In vitro</i>	1 (0.9)
Non-clinical	5 (4.7)
Review	1 (0.9)
Not clear	2 (1.9)
Total	105

to examine molar distalization techniques. The others assessed the effect of functional appliances and also recorded movement of the maxillary molars during the investigation.

Outcome of treatment. We recorded distal movement as 'significant' if it were greater than 1mm. Using this criterion, only the Bass appliance effected any significant distal movement of the maxillary molars in an RCT, with a mean of 1.6 mm, whereas the Fränkel appliance, Harvold activator and bionator provided no distal movement. According to the description of the method, there were deficiencies in the design of all the RCTs when assessed relative to the CONSORT guidelines.

In the CCTs, Ni-Ti coil springs produced the most distal movement, with 3.8 mm, and the acrylic splint

Herbst appliance, with 0.5 mm, the least. The Herbst appliance produced the most distal movement (2.7 mm) and the bioactivator resulted in most mesial movement in cohort studies, while in comparisons of case series, the Herbst appliance again provided the most distal movement and combination-pull headgear with an edgewise appliance the least.

The most distal movement of the maxillary first molars was reported in use of the *en masse* appliance with headgear, with a mean of 5.7 mm and this was in a case series. The least distal movement in this type of study was reported in use of cervical pull headgear. As no appliance was examined in an RCT more than once, no data were available to be pooled in a meta-analysis.

Discussion

The most important finding of this study is that the literature, published over a recent 11-year period, cannot be considered to offer strong evidence for the use of any appliance to bring about distal movement of maxillary molars. Most studies were, or appeared to be, retrospective cohort studies which used control samples that varied in quality. We found it necessary to introduce a classification of 'comparison of case series' to define a number of the studies satisfactorily. Most studies appear to represent searches through records of patients who

Table 3 Classification of articles according to study design and proportion of designs giving a greater strength of evidence than a case report

	Number	
Not relevant	7	
Technique description	11	
Technique and cases	18	
Case report	11	
Case series	18	18 (31 per cent)
Comparison of case series	6	6 (10 per cent)
Cohort study	23	23 (40 per cent)
Controlled clinical trial	8	8 (14 per cent)
Randomized controlled trial	3	3 (5 per cent)
Total	105	58

had previously undergone treatment to identify those matching chosen selection criteria.

We did not use the taxonomy devised and published by Harrison,⁵ which was used to assess and compare papers published in the *British* and *European Journals of Orthodontics*. We found, in the early stages of our assessment of the papers, that it did not appear to provide adequately for the types of paper we encountered nor did it assess some of the features of the methodology we wanted to quantify. However, the taxonomies are very similar in many respects, including how they assess the study designs, but we wanted to try and include some simple, admittedly subjective, assessment of the quality of inclusion and exclusion criteria and assessment of a number of other features of the study design.

We found no RCTs had been carried out over this recent 11-year period, which looked specifically at distal movement of maxillary molars and there were deficiencies in the methodology of those few that did measure distal movement of the molars. However, these papers were accepted for publication prior to dissemination of the CONSORT guidelines in 1998.⁴ It was an important finding that, in many of the articles analysed (22 of 58), it was not clear whether the study was prospective or retrospective. It is also relevant to consider that historical control groups were used in more than a quarter of the studies (12 of 41). This type of control group may not be equivalent to the study group as it is not derived from the same population and is unlikely to have had similar malocclusions at the outset. These observations echo those of Tulloch *et al.* in their review of growth modification with functional appliances that found that no RCTs had been undertaken and the studies that had been carried out were of poor quality.⁸

So can we move molars distally?

In the light of the above comments about the quality of the papers, we feel we cannot make any specific recommendations for any appliance with regard to its ability to move molars distally. Bearing in mind that any biases are going to be in the direction of successful treatment, in assessment of the papers our impression was that the most distal movement of the molars that could be achieved was no more than 2–2.5 mm. For most patients, this represents no more than a half unit improvement in the molar relationship. If any greater correction is required, then some mesial movement of the mandibular molars will be required in addition to achieve it.

It is clear from this investigation that the evidence underpinning one of the most commonly used orthodontic procedures is weak. In order to improve our knowledge of the most effective means of moving molars distally, there is a requirement for good quality research to provide strong evidence. If it is accepted that the best source of evidence for the effectiveness of a given intervention is from the findings of a properly conducted and reported RCT, it is incumbent upon the orthodontic speciality to promote the establishment of such trials.

Conclusions

- A taxonomy has been devised and used in a systematic review to assess papers published in the orthodontic literature describing distal movement of maxillary molars.
- The quality of research carried out over a recent eleven-year period has not provided a high level of evidence upon which to base clinical decisions.

Acknowledgements

We are grateful to Mrs Jean Wright for her invaluable assistance with the statistical analysis.

References

1. Angle EH. Treatment of malocclusion of the teeth and fractures of the maxillae. *Angle's System*, 6th edn, pp. 5–15. Philadelphia: S. S. White Dental Manufacturing Company, 1900.
2. Andrews LF. The six keys to normal occlusion. *Am J Orthod* 1972; **62**: 296–309.
3. Jones RD, White JM. Rapid Class II molar correction with an open-coil jig. *J Clin Orthod* 1992; **26**: 661–664.

4. Begg C, Cho M, Eastwood S, *et al.* Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *J Am Med Ass* 1996; **276**: 637–639.
5. Harrison JE, Ashby D, Lennon MA. An analysis of papers published in the British and European Journals of Orthodontics. *Br J Orthod* 1996; **23**: 203–209.
6. Harrison JE, Ashby D. Orthodontic treatments for posterior crossbites (Cochrane Review). In: *The Cochrane Library*, Issue 1, 1999. Oxford: Update Software.
7. Nguyen QV, Bezemer PD, Habets L, *et al.* A systematic review of the relationship between overjet size and traumatic dental injuries. *Eur J Orthod* 1999; **21**: 503–515.
8. Tulloch JFC, Medland W, Tuncay OC. Methods used to evaluate growth modification in Class II malocclusion. *Am J Orthod Dentofac Orthop* 1990; **98**: 340–347.