European Journal of Orthodontics 27 (2005) 105–114
doi: 10.1093/ejo/cji038

Guest Editorial

What do prospective randomized clinical trials tell us about the treatment of Class II malocclusions? A personal viewpoint

Murray C. Meikle
Discipline of Orthodontics, Faculty of Dentistry, University of Otago, Dunedin, New Zealand

SUMMARY The prospective randomized clinical trial (RCT) is seen by many to be the 'gold standard' for analysing treatment outcome and the only valid source of clinical data. In orthodontics, most RCTs have been designed to resolve the controversy surrounding the ability of functional appliances to significantly modify dentofacial growth. Given the variability in the timing, magnitude and duration of pubertal dentofacial growth, differing levels of motivation and patient compliance, the inherent inaccuracy of cephalometry and the questionable validity of the measurements themselves used to quantitate change, it is not surprising that the conclusions have not been as clear-cut as anticipated. Unlike a laboratory experiment, in which it is possible to limit the differences between experimental and control groups to the single factor being investigated, in a clinical trial an orthodontic appliance is just one of several variables affecting the outcome. Furthermore, RCTs are expensive and time-consuming; by the time the end-point of the study has been reached, the appliance may no longer be in widespread use, the fate of more than one recent well-publicized RCT.

What RCTs have shown is that functional appliances such as the Herbst and twin block, based on the principle of 'jumping the bite', are more effective at modifying dentofacial growth and reducing overjets than headgear and more passive appliances such as the Andresen activator and its variants. However, if one asks whether RCTs have achieved their objective, or provided knowledge not previously available from retrospective studies or animal experimentation, then the answer would have to be no; it is also hard to justify the cost. What is particularly interesting is that knowledge based on years of clinical experience has been disregarded and then announced as if it was something completely new.

Introduction

The prospective randomized clinical trial (RCT) is seen by many to be the 'gold standard' for analysing treatment outcome and the only valid source of clinical data. The scientific worth of the retrospective study, the traditional method used to evaluate the effects of orthodontic treatment, has been criticized for a number of reasons (including selection bias, inadequate sample size, lack of contemporaneous controls, poor research design) and therefore has fallen out of favour. Before administering retrospective clinical studies the last rites, however, the article by Johnston (2002) is recommended.

Over the past decade, having listened to several talks on the findings of RCTs designed to test the hypothesis that certain orthodontic appliances can modify dentofacial growth, one is struck by the often quite marked difference between the conclusions of such investigations and one's own clinical experience. Most have been designed to resolve once and for all the question as to whether headgear (HG) and particularly functional appliances (bionator, Herbst, twin block) have the ability to exert significant orthopaedic as well as orthodontic effects on the dentofacial skeleton. It would appear that RCTs have only partially achieved that ambition and it has to be asked if the rejection of retrospective data is fully justified. The aim of this commentary, therefore, is to discuss the reasons for reaching these conclusions and why, given the complexity of the problem and the evidence of the relevant literature, such an outcome was not entirely unexpected.

Remodelling the dentofacial skeleton

Whether or not growth of the jaws can be altered by orthodontic treatment is a controversial subject of long standing. Prior to the introduction of cephalometric radiography, most clinicians believed the teaching of the Angle school. Malocclusion of the teeth and jaws was the consequence of inadequate bone growth that could be corrected by orthodontic treatment. In other words, orthodontic appliances could stimulate the growth of bone. This is why the cephalometric investigations of Broadbent (1937) and Brodie *et al.* (1938) had such a profound impact on orthodontic thinking at the time, and gave rise to the linked concepts of: (1) the immutability of the facial or morphogenetic pattern of the individual, and (2) the inability of the clinician to alter it in any way.

The old dogma was replaced by a new one: orthodontic treatment was restricted to tooth movement alone. Some clinicians still believe this.

The first evidence that changes in facial growth could be achieved clinically came from retrospective cephalometric studies of patients who had worn extraoral traction (Moore, 1959; Ricketts, 1960). These demonstrated that HG could restrain forward growth of the maxilla and in some cases move it distally in relation to the anterior cranial base. The skeletal and dental effects of a variety of appliances originating with the Norwegian method of functional jaw orthopaedics (Andresen and Häupl, 1942) have also been reported in numerous, largely retrospective, investigations (Korkhaus, 1960; Fränkel, 1966; Marschner and Harris, 1966; Demisch, 1972; Pancherz, 1979; Weislander, 1984; McNamara *et al.*, 1985; Hägg and Pancherz, 1988; Baccetti *et al.*, 2000). The volume of literature is such that during the period 1980–1990, more than 130 articles reporting data on groups of 10 or more patients with Class II malocclusions, treated with 14 different appliance systems, were published in the four most widely circulated English language orthodontic journals (Tulloch *et al.*, 1997a). Nearly all reported successful treatment of the malocclusion, but whether the correction involved the modification of facial growth, and in particular mandibular growth to any significant extent, has remained a matter of dispute.

Numerous well-documented experimental studies carried out on non-human primates have shown that craniofacial sutures and the temporomandibular joints (TMJs) are amenable to mechanically induced remodelling activity (Breitner, 1940; Meikle, 1970; Stöckli and Willert, 1971; Moffett, 1971; Adams *et al.*, 1972; Elder and Tuenge, 1974; Brandt *et al.*, 1979). As might be expected from its structure and function, the TMJ is morphologically adapted to resist the impact of mechanical loading and, therefore, more difficult to remodel. In contrast, the fibrous articulations that attach the maxillary complex to the rest of the skull can be remodelled with comparative ease. Although out of fashion these days for moral and financial reasons, such animal experimentation does enable informed judgements to be made regarding the likely effects of clinical treatment on the dentofacial skeleton. In practical terms this means that if a force of appropriate magnitude and duration is applied to the jaws, it should be possible to produce an orthopaedic effect and remodel both the sutures of the maxillary complex and the TMJs. Whether these changes can be utilized clinically in the correction of skeletal Class II malocclusions is the question that most RCTs have been designed to answer.

The evidence of RCTs

In orthodontics, RCTs of Class II treatment have been limited in number. The first was undertaken by Jakobsson (1967), in which 60 subjects aged 8–9 years with a Class II division 1 malocclusion were randomly assigned from triads of children at the same stage of dental development to either an Andresen activator, HG, or the control group. Both HG and activator treatment had a distalizing effect on the maxilla, but the study did not support the hypothesis that activator treatment affected condylar growth. A similar methodology was used by Nelson *et al.* (1993) in a study of 42 children aged 10–13 years, one subject in each triad being randomly assigned to either a Fränkel functional regulator (FR-2), a Harvold activator or the control group. They could find no evidence to support the view that either appliance was capable of altering the size of the mandible, the main effect being to increase the height of the face.

In 1988, the National Institute of Dental and Craniofacial Research (NIDCR) funded two RCTs at the Universities of North Carolina (UNC) and Florida designed to evaluate the effectiveness of HG and bionator treatment in the early management of Class II division 1 malocclusions. In the UNC trial, which has been widely reported in the literature and at clinical meetings (Tulloch *et al.*, 1997a,b, 1998), 166 patients in the mixed dentition with an overjet greater than 7 mm were randomly assigned to early treatment with either a HG or bionator, or to observation. There was considerable variation in the pattern of change in all three groups (Figure 1), with about 80 per cent of the treated children responding favourably. The HG group showed restricted forward movement of the maxilla of about 1 mm (–0.92 versus 0.26 mm). In the bionator group, a small increase in mandibular length was observed (3.69 versus 2.36 mm),

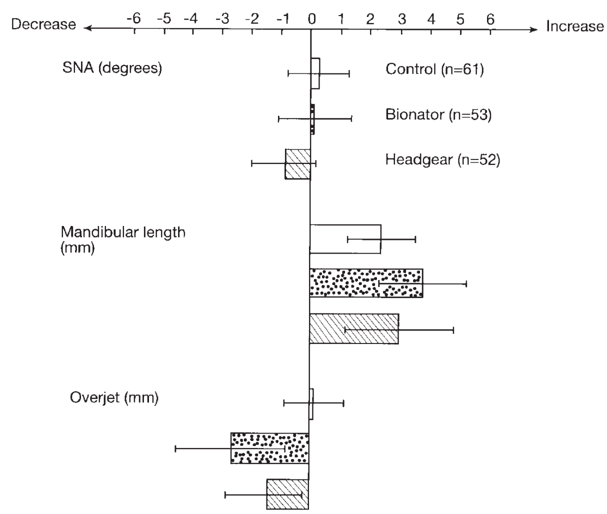

Figure 1 Mean ± standard deviation (annualized change in mm) for three cephalometric measurements during phase 1 of the University of North Carolina clinical trial. One hundred and sixty-six patients in the mixed dentition with a Class II division 1 malocclusion were randomly assigned to either a headgear, bionator or control group. Data from Tulloch *et al.* (1997a).

although it is not clear how mandibular length was measured. The Florida study involved 249 participants and showed that compared with controls, facial growth was altered during the course of both bionator and HG/bite plate therapies; both appliance systems enhanced mandibular growth (2.4 and 2.2 mm, respectively, versus 1.6 mm) without detectable relapse 12 months after the end of active treatment (Keeling *et al.*, 1998). In both studies these values were expressed as the average annualized change.

The Andresen monobloc and its variants, such as the bionator and Harvold activator, are essentially passive in action. With the popularity of a new generation of active functional appliances such as the Herbst and Clarke twin block, based on the principle of 'jumping the bite', attention has focussed on how they might achieve dento-facial change. The Herbst appliance has been shown to be very effective in the treatment of Class II malocclusions. In a retrospective study it was concluded that besides a possible stimulating effect on mandibular growth, Herbst therapy resulted in a redirection of maxillary growth and significant tooth movement in both jaws (Pancherz, 1979). In a subsequent prospective study of 42 Class II division 1 patients, 22 were treated with a Herbst appliance for 6 months (Pancherz, 1982). A Class I occlusal relationship was achieved in all cases and equally divided between skeletal and dental changes. Overjet correction averaging 5.2 mm was mainly the result of a mean increase in mandibular length of 2.2 mm (3.0 versus 0.8 mm; Figure 2) and a 1.8 mm proclination of the mandibular incisors.

The twin block has been the subject of several investigations in recent years, some of which have been conducted as RCTs. Lund and Sandler (1998), in a study involving 63 participants aged 10–14 years treated with a twin block for an average period of 9 months, could find no significant HG effect. However, treatment resulted in an increase in the mandibular dimension articulare–pogonion (Ar–Pog) of 5.1 mm, compared with a control group increase of 2.7 mm; they were unable to decide if this was due to an increase in mandibular length or a repositioning of the mandible.

In a multicentre trial carried out recently in the UK (O'Brien *et al.*, 2003a,b,c), 174 children aged 8–10 years with a Class II division 1 malocclusion were randomly allocated to a twin block treatment group or to an untreated control group. Data were collected at the start of treatment and after 15 months. The results showed that early treatment with a twin block resulted in a reduction in overjet and molar relationship and a reduction in the severity of the malocclusion. These authors concluded that most of the correction was dentoalveolar, and that the small change in skeletal relationship might not be considered clinically significant. It is difficult to understand how they reached this conclusion, given that the combined maxillary and mandibular skeletal changes amounted to 2.43 mm (Figure 3) and the dental changes 5.06 mm (maxillary incisor retroclination –3.03 mm; mandibular incisor proclination +2.03 mm). Interestingly, these changes are similar to the findings of previous retrospective studies of twin block treatment (Mills and McCulloch, 1998; Baccetti *et al.*, 2000). These attributed 50 per cent of overjet correction to skeletal change involving the mandible, and 50 per cent to alterations in the position of the upper and lower incisors. That these were retrospective studies and did not use contemporaneous controls does not necessarily invalidate their findings.

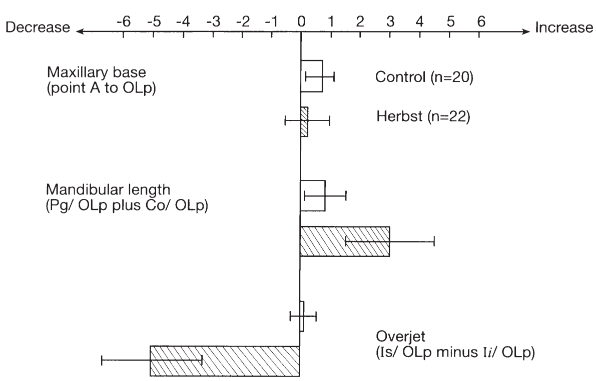

Figure 2 Mean ± standard deviation for three cephalometric measurements (see Figure 6) after 6 months of treatment with a Herbst appliance. Forty-two patients with a Class II division 1 malocclusion were assigned to a Herbst treatment group or to a control group. Two lateral skull radiographs were taken, one in occlusion and the other with the mouth wide open to aid identification of the condylar head. Temporomandibular joint radiographs indicated that post-treatment, the condyle–fossa relationship was unchanged. Data from Pancherz (1982).

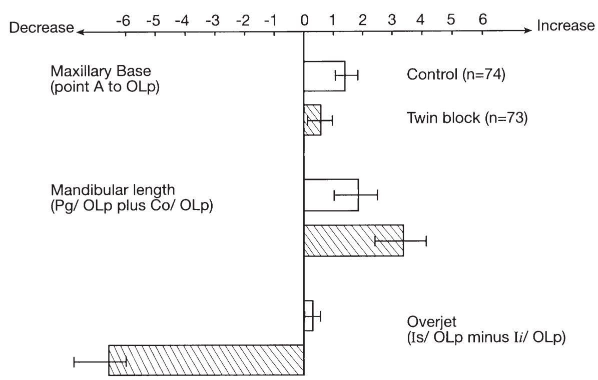

Figure 3 Means plus 95 per cent confidence intervals for three cephalometric measurements using Pancherz's (1982) analysis in a randomized clinical trial of 174 children aged 8–10 years with a Class II division 1 malocclusion. The participants were randomly assigned to a twin block treatment group or to a control group; 147 completed the trial. Data from O'Brien *et al.* (2003b).

What is being measured?

On the face of it, an RCT designed to test the clinical effectiveness of an orthodontic appliance looks fairly straightforward. After deciding on selection criteria, patients are recruited onto the programme, randomly assigned to experimental and control groups, and cephalometric radiographs are taken at the beginning and end of treatment. Various outcome measures are chosen and analysed numerically to establish whether significant differences can be detected between the groups. From these data, conclusions regarding the ability of the appliance to produce dentofacial change can be derived. One needs to be absolutely clear, however, about what is being measured. Is one measuring, for example (1) patient compliance; (2) the magnitude and/or timing of pubertal facial growth; (3) the skill and personality of the clinician; or (4) the ability of the appliance to alter the pattern of maxillary and/or mandibular growth? The answers to the parameters 1–3 will clearly influence the outcome and raise some important questions. Should participants in the trial who fail to wear the appliance be excluded? Are multicentre trials designed to eliminate the operator effect more valid than single centre trials? Those not familiar with the terminology, may find the distinction between efficacy (does it work?) and effectiveness (how well does it work in the field?) somewhat Jesuitical. Furthermore, how does one take into account the fact that the timing and magnitude of pubertal spurts in facial growth are highly variable and difficult to predict? And finally, are the measurements chosen to quantify differences in the various skeletal or dental parameters before and after treatment accurate and/or valid?

The validity of mandibular growth measurements

Changes in the growth of various facial dimensions are customarily based on measurements of the linear distance between two cephalometric landmarks. For a measurement of the anterior cranial base such as sella–nasion (S–N), this does not present too much of a problem. However, for an angular bone, such as the mandible, the use of linear dimensions such as condylion–gnathion (Co–Gn) or its surrogate, articulare–gnathion (Ar–Gn), to quantify changes in mandibular growth has considerable potential for error. An additional and potentially more severe problem is that condylion, obscured as it is by the temporal bone and basiocciput, is notoriously difficult to identify accurately on conventional cephalometric radiographs. A headfilm taken with the mouth wide open will aid location of the condylar head, a technique used by Pancherz (1982), but an additional radiographic exposure for this purpose is unlikely to be approved by ethical committees.

When Björk superimposed mandibular profiles on metallic implants he found that the longitudinal growth of the mandible was confined to the condyle, and that the direction of condylar growth was highly variable (Björk, 1963; Björk and Skieller, 1972). The linear dimensions Co–Gn or Ar–Gn will thus be strongly influenced by the condylar growth direction of the subject being measured (Figure 4). Unfortunately, none of the recently reported RCTs seems to have taken mandibular growth rotation into account when calculating mandibular change; the fact that it may be difficult to measure does not mean it can be ignored. It is disappointing to find that the work of Björk, one of the key figures of cephalometric growth studies, is apparently being ignored.

The significance of the problem was demonstrated by Hägg and Attström (1992), who quantified differences between conventional cephalometric methods for estimating mandibular length, and what they termed the 'scientific method'. The scientific method was based on changes in the position of condylion on cephalograms from the Björk collection, orientated by metallic implants. They found that: (1) the amount of growth estimated by the distance condylion–pogonion (Co–Pog) was on average 3.3 mm less than the scientific method, and (2) the distance Ar–Pog was on average 3.9 mm less than the scientific method (Figure 5). By any criteria, these are significant differences. Measurements of mandibular growth based on standard cephalometric methods that consistently underestimate condylar growth are therefore

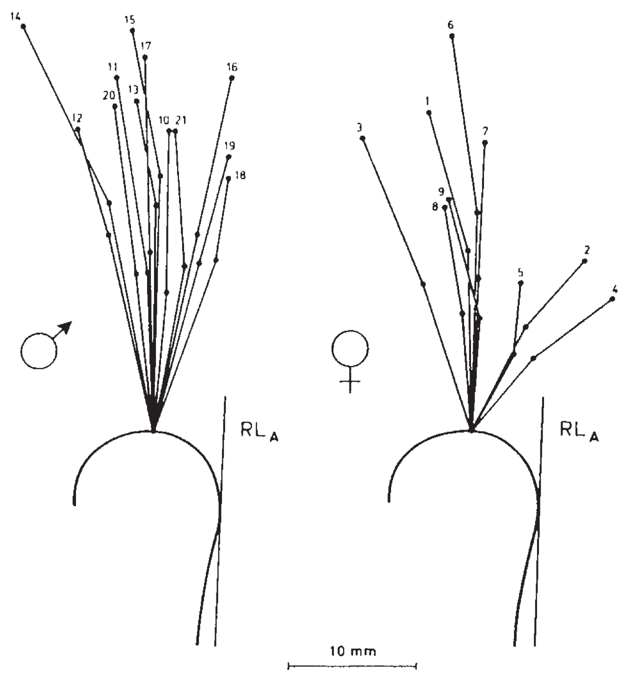

Figure 4 Illustration of the wide variability in condylar growth direction (revealed by superimposition of cephalometric tracings on mandibular implants at 3 year intervals) in 12 boys and nine girls over a 6 year period in relation to the ramus line (RL). The growth in most cases followed a path that curved forwards and was strongly correlated with rotation of the mandible; in case 4 there was a backward curvature of 13 degrees and in case 2 of 15 degrees. Reprinted from Björk and Skieller (1972) with permission from the American Association of Orthodontists.

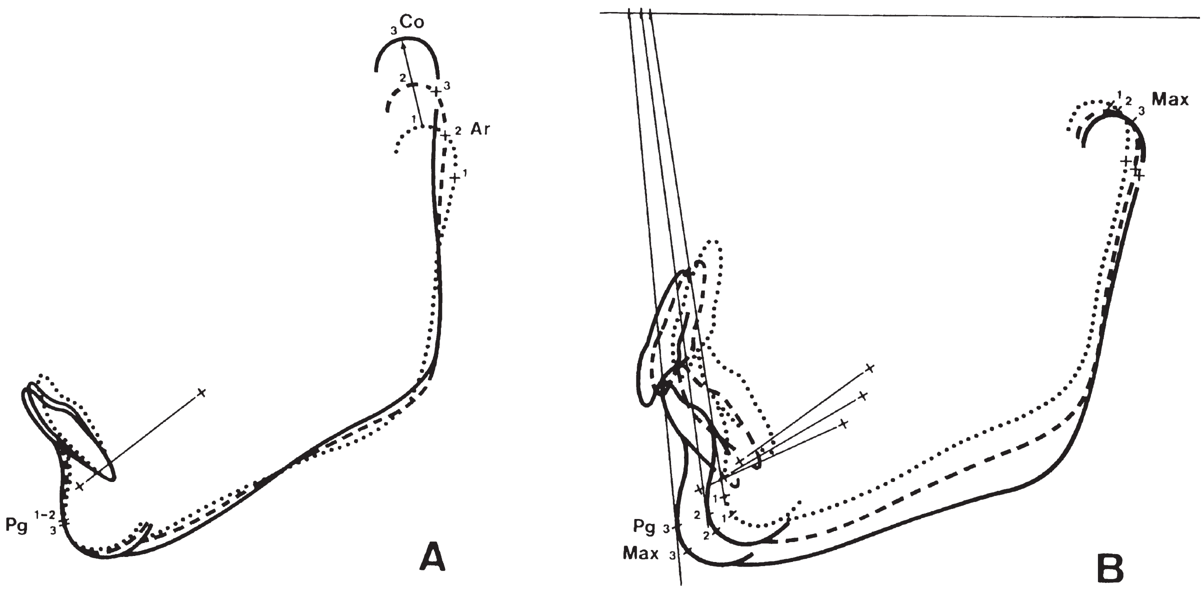

Figure 5 Pronounced vertical condylar growth pattern in a girl from the Copenhagen Growth Study. (A) Localization of the landmarks used to estimate mandibular growth by the scientific method (condylion), and for two of the standard cephalometric methods [condylion–pogonion (Co–Pog) and articulare–pogonion (Ar–Pog)]. (B) Points used for the estimation of maximum mandibular length. It also shows the three locations of pogonion (Pog). Reprinted from Hägg and Attström (1992) with permission from the American Association of Orthodontists.

not valid. The analysis of Pancherz (1982), also used to evaluate treatment effects in the RCT reported by O'Brien *et al.* (2003a,b), will similarly underestimate mandibular growth as it is a linear measurement that does not take into account individual variation in condylar growth rotation (Figure 6).

Individual variation in the pattern of dentofacial growth

It was Lande (1952), in one of the classic papers of the cephalometric literature, who reported that between 3 and 18 years of age there was a decrease in the

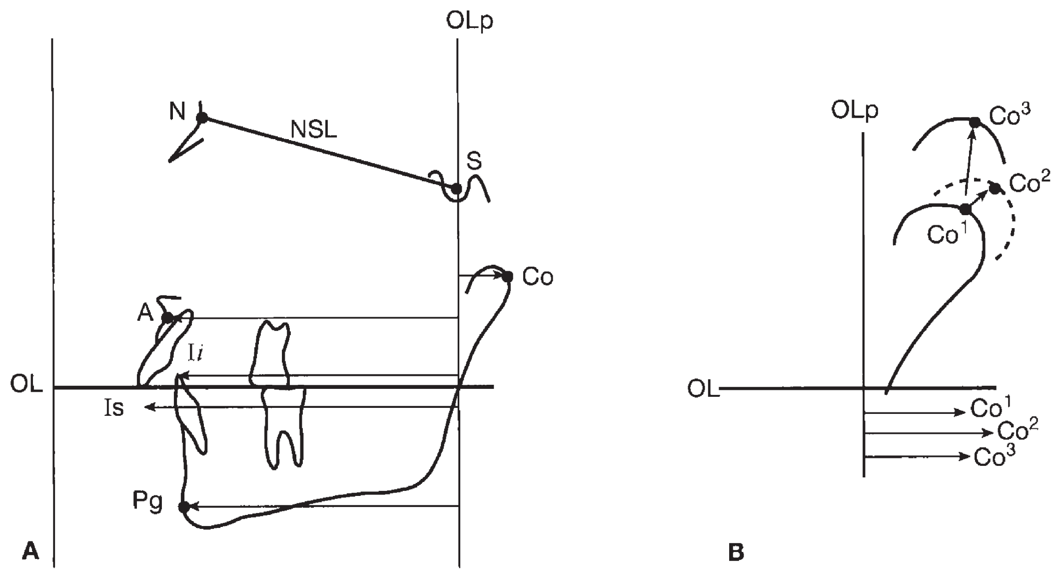

Figure 6 Pancherz's analysis. (A) Linear measurements of sagittal skeletal and dental changes are made from the reference line OLp (occlusal line perpendiculare), a line perpendicular to the occlusal line (OL) through sella. Mandibular length is represented by Pg/OLp plus Co/OLp. Redrawn from Pancherz (1982). (B) Pancherz's method for measuring mandibular length will underestimate the contribution of condylar growth (Co/OLp) compared with a direct measurement from Co^1 to Co^2 or Co^3. The difference will be less significant the more horizontal the growth pattern of the condyles. However, Figure 4 suggests that most will grow in the direction suggested by Co^3, in which case the difference between the direct and indirect measurements will be considerable.

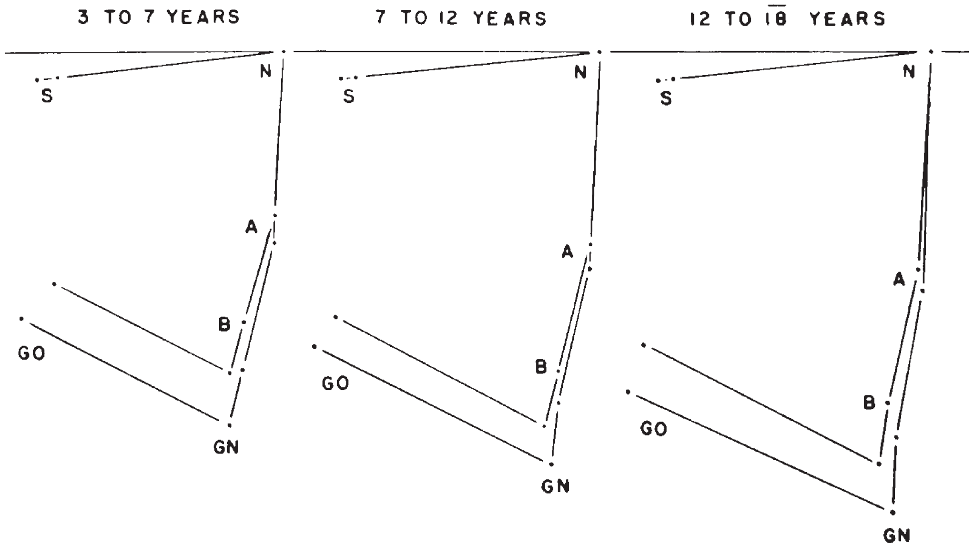

Figure 7 Mean changes in the profile of 34 male subjects from 3 to 18 years of age. From 3 to 7 years no significant change in the anteroposterior position of points A, B or gnathion (Gn) was recorded. From 7 to 12 years, Gn moved forward 1.3 mm, while points A and B remained unchanged. From 12 to 18 years, point A moved forward 1 mm, point B 2.2 mm, and Gn 3.7 mm (Lande, 1952).

convexity of the facial profile represented by the angle nasion–point A–gnathion (N–A–Gn). Most of the patients also showed a decrease in the angular relationship between the lower border of the mandible and the S–N plane (Figure 7). The increase in mandibular prognathism was similar to that previously reported by Björk (1947), but what makes this study important is that special attention was focussed on individual variation within the sample. Expressing data numerically as mean ± standard deviation may be statistically *de rigeur*, but does not have the same visual impact as Figure 8, in which the subjects are ranked according to the degree of forward movement of gnathion. The amount of movement in the position of gnathion in an anteroposterior direction ranged from –3.7 (B 2049) to +12.75 mm (B 1168). Anteroposterior change in the position of point A showed a comparable wide variation and ranged from –4.25 (B 2005) to +3.25 mm (B 2252). Ten of the 34 subjects had a Class II dental base relationship, but in none of these did the face tend to become more retrognathic. However, data from Tulloch *et al.* (1997a) for a mixed dentition Class II population suggest that without treatment there is approximately a 30 per cent chance of a favourable change in jaw relationship, a 50 per cent chance of no change, and a 15 per cent chance that the Class II will worsen. It is not surprising, therefore, that given the wide variation and unpredictability in the pattern of dentofacial growth, the response of individual patients to treatment can be equally variable and unpredictable.

Growth velocity curves for growth of the mandibular condyle in a French-Canadian population (Buschang *et al.*, 1999) have shown, for males, that mean yearly velocities of condylar growth can range from 2.1 to 3.1 mm/year with a pubertal peak at 14.3 years; females showed a more constant rate of growth that ranged from 2.0 to 2.7 mm/year with a pubertal peak at 12.2 years. There was a substantial variation in condylar growth. For a male patient on the 90th percentile, for example, condylar growth will average 5 mm/year, while for another on the 25th percentile the annual increment will be as little as 1–2 mm. This will have a significant impact on treatment outcome in patients with skeletal discrepancies, as shown by the study of DeVincenzo (1991).

The role of statistics

It is a common assumption that most of the participants in a clinical investigation will exhibit the mean (or average) change. In practice this is not the case, although it is true that 60 per cent of a normally distributed population sample will lie within ± 1 SD of the mean. With the relatively modest numbers of patients in the experimental and control groups in RCTs, very few of the participants are therefore likely to show the average change. To their credit, some investigators recognize this. Tulloch *et al.* (1997a) pointed out that the variability of individual responses is often obscured by a focus on mean changes. Unfortunately, this is what audiences, and orthodontic trainees in particular, seem to remember, and the reason why statistical data have limited impact on clinical opinion when it comes to discussing treatment outcome. It may be unfashionable to say so, but clinical opinion is still strongly influenced by anecdotal evidence and the training and experience of the clinician, not always by the statistical artefact represented by the mean. Or put

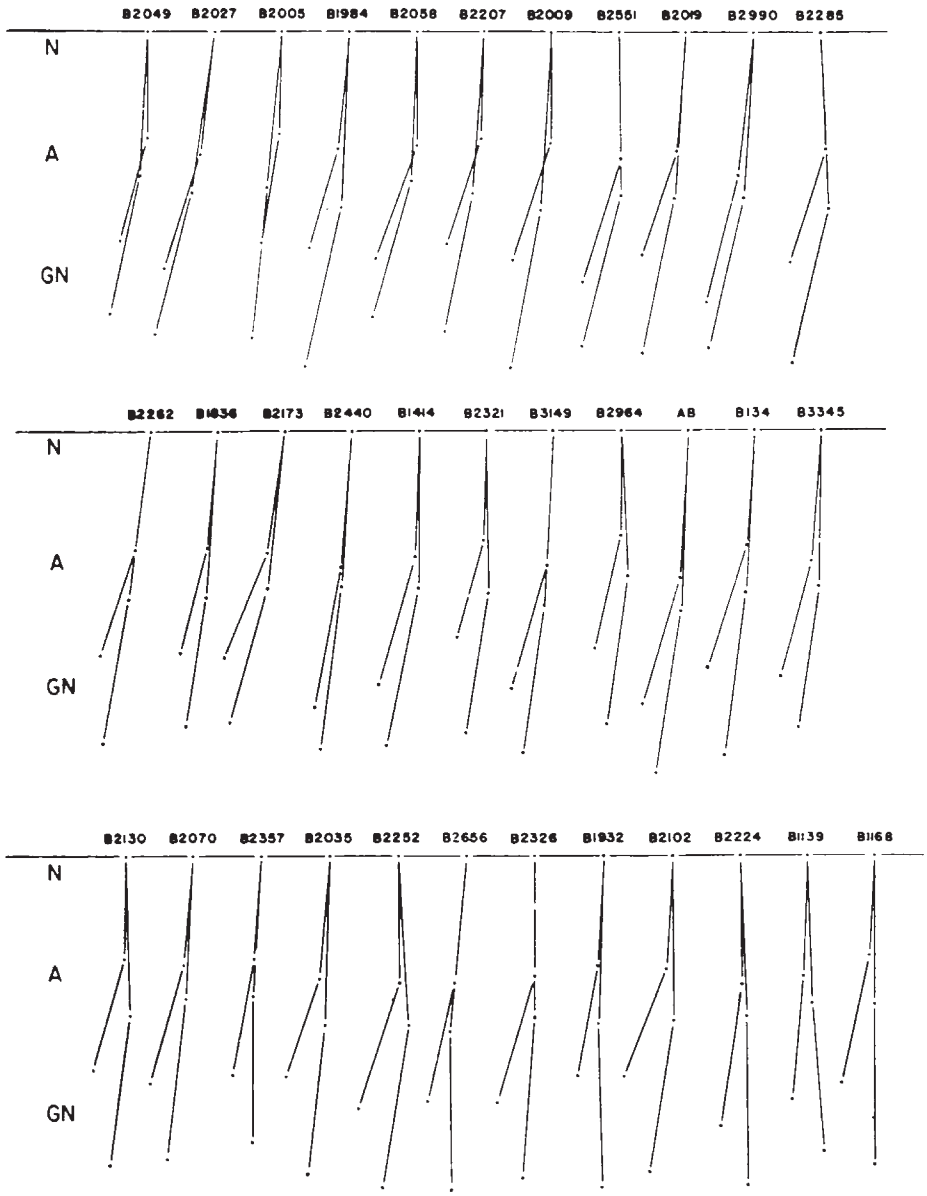

Figure 8 Profile changes [nasion–point A–gnathion (N–A–Gn)] in each of the 34 male subjects ranked in order of increasing forward movement of Gn. This figure illustrates the wide range of individual variation within the group compared with the mean profile changes shown in Figure 7 (Lande, 1952).

another way, many orthodontists are sceptical about the findings of RCTs, for the simple reason that they do not seem to reflect their own clinical experience.

Can the TMJ be remodelled clinically?

On the basis of the histological evidence, it is clear that functional anterior displacement of the mandible in non-human primates produces remodelling of the surface contours of both the condylar head and the glenoid fossa (Breitner, 1940; Adams *et al.*, 1972; Brandt *et al.*, 1979; McNamara and Carlson, 1979; Woodside *et al.*, 1987). There is also some evidence to suggest that condylar growth can be directed in a more posterior direction. Remodelling the TMJ in monkeys is one thing, remodelling it clinically is another. Nevertheless, there is some evidence from patients treated with the Herbst appliance that suggests that it might be achievable.

In a recent systematic review of the literature regarding the effects of Herbst treatment on TMJ morphology, Popowich *et al.* (2003) identified 80 studies related to the topic. Publications that used transpharyngeal radiographs to document morphological change were excluded, which eliminated the major report of Paulsen (1997) on 100 consecutive Class II patients. That left five publications meeting the selection criteria and of those four were from Giessen, Germany. In one of these, Ruf and Pancherz (1998) used magnetic resonance imaging (MRI) to analyse TMJ growth adaptation in 15 consecutive Class II patients treated for a period of 7 months. After 6–12 weeks of treatment, MRI signs of condylar remodelling were seen at the posterosuperior border in 29 of the 30 condyles, while glenoid fossa remodelling was noted in 22 joints. In another study, Croft *et al.* (1999) used cephalometry and TMJ tomograms to evaluate treatment changes. They did not examine the effects of treatment on condylar morphology and found little evidence of glenoid fossa remodelling. They did, however, conclude that Herbst treatment in the mixed dentition produces significant long-term improvements in dental and skeletal relationships as a result of dentoalveolar changes and orthopaedic effects in both jaws.

These findings suggest that at least with the Herbst appliance (and probably the twin block), TMJ remodelling and growth modification can be regarded as a possibility. It is important, however, to regard these changes as an adaptation to the altered functional activity of the joint (as in any other joint) and not growth stimulation. Despite claims to the contrary, there is no convincing experimental proof for condylar growth stimulation in either rodent or primate models of mandibular displacement. Whether these changes can be regarded as clinically significant will depend upon the individual practitioner. Equally, the question of long-term stability has not been considered.

Conclusions

Most RCTs have been designed to resolve, once and for all, the controversy surrounding the ability of orthodontic appliances to significantly modify dentofacial growth. Given the variability in the timing, magnitude and duration of pubertal dentofacial growth, differing levels of motivation and patient compliance, the inherent inaccuracy of the cephalometric method (the technical errors in cephalometry may be greater than the growth changes one is hoping to identify), the questionable validity of the measurements themselves used to quantitate change, not to mention the personality and competence of the clinicians involved, it is not surprising that the conclusions have not been as clear-cut as might have been anticipated. Unlike a laboratory experiment, in which it is possible to limit the difference between experimental and control groups to the single factor being investigated, in a clinical trial an orthodontic appliance is just one of several variables affecting the outcome. The greater the variation between individuals, the harder it is to demonstrate that a difference in treatment effect is significant or does in fact exist, and limitless variation is the normal human condition. RCTs are also expensive to conduct and given the time required to reach a conclusion, may be investigating an appliance no longer in widespread use; this seems to have been the fate of the UNC and Florida studies.

What RCTs have shown is that functional appliances such as the Herbst and twin block, based on the principle of 'jumping the bite', are more effective in modifying dentofacial growth and reducing overjets than HG and more passive appliances such as the Andresen activator and its variants. However, if one asks whether RCTs have achieved their intended objective, or provided knowledge not previously available from retrospective studies or animal experimentation, then the answer would have to be no. It is also hard to justify the cost. What is particularly interesting is that knowledge based on years of clinical experience has been disregarded and then announced as if it was something completely new. It is also doubtful whether the outcome of larger or better trials would be any more informative. Some form of selection from these population samples to establish why some patients have a successful outcome while others do not would be of greater use to the clinician. The present situation appears to be not unlike the relationship between cross-sectional data (in which individual characteristics are smoothed out by grouping) and longitudinal data in studies of physical growth at adolescence.

Address for correspondence

M. C. Meikle
Faculty of Dentistry
University of Otago
PO Box 647
Dunedin
New Zealand
Email: oral-sciences@otago.ac.nz

Acknowledgements

The author wishes to thank Jonathan Sandy, Dirk Bister, Jeremy Breckon and Jules Keiser for their useful comments and suggestions regarding the manuscript.

References

Adams C D, Meikle M C, Norwick K W, Turpin D L 1972 Dentofacial remodelling produced by intermaxillary forces in *Macaca mulatta*. Archives of Oral Biology 17: 1519–1535

Andresen J, Häupl K 1942 Funktionkieferorthopädie. Die Grundlage des Norwegischen Systeme. Verlag von J A Barthe, Leipzig

Baccetti T, Franchi L, Toth L R, McNamara Jr J A 2000 Treatment timing for twin-block therapy. American Journal of Orthodontics and Dentofacial Orthopedics 118: 159–170

Björk A 1947 The face in profile. An anthropological X-ray investigation on Swedish children and conscripts. Svensk Tandlakare-Tidskrift 40: 1–180

Björk A 1963 Variations in the growth pattern of the human mandible: longitudinal radiographic study by the implant method. Journal of Dental Research 42: 400–411

Björk A, Skieller V 1972 Facial development and tooth eruption. An implant study at the age of puberty. American Journal of Orthodontics 62: 339–383

Brandt H C, Shapiro P A, Kokich V G 1979 Experimental and post-experimental effects of posteriorly directed extraoral traction in adult *Macaca fascicularis*. American Journal of Orthodontics 75: 301–317

Breitner C 1940 Bone changes resulting from experimental orthodontic treatment. American Journal of Orthodontics and Oral Surgery 26: 521–546

Broadbent B H 1937 The face of the normal child. Angle Orthodontist 7: 183–208

Brodie A G, Downs W B, Goldstein A, Myer E 1938 Cephalometric appraisal of orthodontic results. Angle Orthodontist 8: 261–351

Buschang P H, Santos-Pinto A, Demirijan A 1999 Incremental growth charts for condylar growth between 6 and 16 years of age. European Journal of Orthodontics 21: 167–173

Croft R S, Buschang P H, English J D, Meyer R 1999 A cephalometric and tomographic evaluation of Herbst treatment in the mixed dentition. American Journal of Orthodontics and Dentofacial Orthopedics 116: 435–443

Demisch A 1972 Effects of activator therapy on the craniofacial skeleton in Class II division 1 malocclusion. Transactions of the European Orthodontic Society, pp. 295–310

DeVincenzo J P 1991 Changes in mandibular length before, during, and after successful orthopedic correction of Class II malocclusions, using a functional appliance. American Journal of Orthodontics and Dentofacial Orthopedics 99: 241–257

Elder J R, Tuenge R H 1974 Cephalometric and histologic changes produced by extra-oral high-pull traction to the maxilla in *Macaca mulatta*. American Journal of Orthodontics 66: 599–617

Fränkel R 1966 The theoretical concept underlying treatment with functional correctors. Transactions of the European Orthodontic Society, pp. 233–250

Hägg U, Attström K 1992 Mandibular growth estimated by four cephalometric measurements. American Journal of Orthodontics and Dentofacial Orthopedics 102: 146–152

Hägg U, Pancherz H 1988 Dentofacial orthopaedics in relation to chronological age, growth period and skeletal development. An analysis of 72 male patients with Class II division 1 malocclusion treated with the Herbst appliance. European Journal of Orthodontics 10: 169–176

Jakobsson S O 1967 Cephalometric evaluation of treatment effect on Class II division 1 malocclusions. American Journal of Orthodontics 53: 446–456

Johnston L E 2002 Moving forward by looking back: 'retrospective' clinical studies. Journal of Orthodontics 29: 221–226

Keeling S D *et al.* 1998 Anteroposterior skeletal and dental changes after early Class II treatment with bionators and headgear. American Journal of Orthodontics and Dentofacial Orthopedics 113: 40–50

Korkhaus G 1960 Present orthodontic thought in Germany. Experiences with the Norwegian method of functional orthopedics in the treatment of distocclusion. American Journal of Orthodontics 46: 207–287

Lande M J 1952 Growth behavior of the human bony facial profile as revealed by serial cephalometric roentgenology. Angle Orthodontist 22: 78–90

Lund D I, Sandler P J 1998 The effects of twin blocks: a prospective controlled study. American Journal of Orthodontics and Dentofacial Orthopedics 113: 104–110

Marschner J F, Harris J E 1966 Mandibular growth and Class II treatment. Angle Orthodontist 36: 89–93

McNamara Jr J A, Carlson D S 1979 Quantitative analysis of temporomandibular joint adaptations to protrusive function. American Journal of Orthodontics 76: 593–611

McNamara Jr J A, Bookstein F L, Shaughnessy T G 1985 Skeletal and dental changes following functional regulator therapy on Class II patients. American Journal of Orthodontics 88: 91–110

Meikle M C 1970 The effect of a Class II intermaxillary force on the dentofacial complex in the adult *Macaca mulatta* monkey. American Journal of Orthodontics 58: 323–340

Mills C M, McCulloch K J 1998 Treatment effects of the twin block appliance: a cephalometric study. American Journal of Orthodontics and Dentofacial Orthopedics 114: 15–24

Moffett B C 1971 Remodelling of the craniofacial articulations by various orthodontic appliances in rhesus monkeys. Transactions of the European Orthodontic Society, pp. 207–216

Moore A W 1959 Observations on facial growth and its clinical significance. American Journal of Orthodontics 45: 397–423

Nelson C, Harkness M, Herbison P 1993 Mandibular changes during functional appliance treatment. American Journal of Orthodontics and Dentofacial Orthopedics 104: 153–161

O'Brien K *et al.* 2003a Effectiveness of treatment for Class II malocclusion with the Herbst or twin-block appliances: a randomized, controlled trial. American Journal of Orthodontics and Dentofacial Orthopedics 124: 128–137

O'Brien K *et al.* 2003b Effectiveness of early orthodontic treatment with the twin-block appliance: a multicenter, randomized, controlled trial. Part 1: Dental and skeletal effects. American Journal of Orthodontics and Dentofacial Orthopedics 124: 234–243

O'Brien K *et al.* 2003c Effectiveness of early orthodontic treatment with the twin-block appliance: a multicenter, randomized, controlled trial. Part 2: Psychosocial effects. American Journal of Orthodontics and Dentofacial Orthopedics 124: 488–495

Pancherz H 1979 Treatment of Class II malocclusions by jumping the bite with the Herbst appliance: a cephalometric investigation. American Journal of Orthodontics 76: 423–442

Pancherz H 1982 The mechanism of Class II correction in Herbst appliance treatment. A cephalometric investigation. American Journal of Orthodontics 82: 104–113

Paulsen H U 1997 Morphological changes in the TMJ condyles of 100 patients treated with the Herbst appliance in the period of puberty to adulthood: a longitudinal radiographic study. European Journal of Orthodontics 19: 657–668

Popowich K, Nebbe B, Major P W 2003 Effect of Herbst treatment on temporomandibular joint morphology: a systematic literature review. American Journal of Orthodontics and Dentofacial Orthopedics 123: 388–394

Ricketts R M 1960 The influence of orthodontic treatment on facial growth and development. Angle Orthodontist 30: 103–133

Ruf S, Pancherz H 1998 Temporomandibular joint adaptation in Herbst treatment: a prospective magnetic resonance imaging and cephalometric roentgenographic study. European Journal of Orthodontics 20: 375–388

Stöckli P W, Willert H G 1971 Tissue reactions in the temporomandibular joint resulting from anterior displacement of the mandible in the monkey. American Journal of Orthodontics 60: 142–155

Tulloch J F C, Phillips C, Koch G, Proffit W R 1997a The effect of early intervention on skeletal pattern in Class II malocclusion: a randomized clinical trial. American Journal of Orthodontics and Dentofacial Orthopedics 111: 391–400

Tulloch J F C, Proffit W R, Phillips C 1997b Influences on the outcome of early treatment for Class II malocclusion. American Journal of Orthodontics and Dentofacial Orthopedics 111: 533–542

Tulloch J F C, Phillips C, Proffit W R 1998 Benefit of early Class II treatment: progress report of a two-phase randomized clinical trial. American Journal of Orthodontics and Dentofacial Orthopedics 113: 62–72

Weislander L 1984 Intensive treatment of severe Class II malocclusions with a headgear–Herbst appliance in the early mixed dentition. American Journal of Orthodontics 86: 1–13

Woodside D G, Metaxas A, Altuna G 1987 The influence of functional appliance therapy on glenoid fossa remodeling. American Journal of Orthodontics and Dentofacial Orthopedics 92: 181–198