Journal of Orthodontics, Vol. 36, 2009, 42–51

SCIENTIFIC SECTION

A systematic review of clinical trials of aligning archwires

Michael Riley
University of Manchester, UK

David R. Bearn
University of Dundee, UK

Objectives: This review aimed to identify the evidence for the efficacy of archwires used in the alignment stage of orthodontic treatment by undertaking a systematic review of the literature.

Data Sources: MEDLINE, the Cochrane Central Register of Controlled trials (CENTRAL), EMBASE, and the *meta* Register of Controlled Trials were searched up to July 2008. Reference lists of identified articles and relevant review articles were checked for further possible studies.

Review Methods: Controlled clinical trials and randomised clinical trials that compared aligning archwires and reported objective measures of alignment were selected for inclusion. Validity and quality assessment were undertaken to identify studies with a low risk of bias. Details of the study methodology and the reported results were then abstracted.

Results: 100 studies were identified by the searches and 7 of these were identified as meeting the selection criteria. Four studies were deemed, after quality assessment, to have a low risk of bias and data was extracted from these. No two studies shared a common methodology or common reporting of outcome. Meta-analysis was therefore not possible.

Conclusions: There is insufficient data in these studies to make clear recommendations regarding the most effective archwire for alignment. Recommendations on future study design have been made.

Key words: Orthodontics, archwires, systematic review

Received 25th August 2008; accepted 23rd Novermber 2008

Background

The aligning efficacy and effectiveness of many archwire materials and dimensions used on a daily basis by orthodontists have not been scrutinized by clinical trial. Proffit[1] described the ideal aligning archwire as one which has excellent strength, excellent springiness and a long range of action with force values of about 50 g (the optimum force for tipping the teeth into alignment). In addition to this, Kusy[2] suggested that such wires should be aesthetic and biocompatible. The archwire materials used to achieve levelling and alignment at the present time meet most of these ideals except for aesthetics and occasional biocompatibility problems (Nickel allergy). The commonly used materials currently used fall into two broad categories; stainless steel and nickel titanium. Newer materials which add Kusy's demand for aesthetics have also been developed.

The range of archwires available have been subjected to *in vitro* laboratory testing to ascertain their physical properties by too many investigators to list here. However, *in vitro* tests, no matter how complex will never be able to fully reproduce or predict the performance of an aligning archwire in clinical practice. To understand the clinical trials are required. However in 1996 Evans commented on this paucity of clinical evidence and added that manufacturers were in such a headlong rush to produce the ultimate aligning archwire that very little attention had been paid to the *in vivo* behaviour of these materials.[3] This review aimed to see if more than a decade later the situation had improved.

Objectives

To systematically evaluate all clinical trials which investigate the effectiveness of archwires for alignment and levelling, and if possible identify the most effective archwire through the use of meta-analysis. This review is reported according to the QUORUM statement recommendations.

Methods

Searching

A simple search of, for example, Medline is generally not considered adequate[4] and so a number of sources were searched to retrieve the relevant literature:

Address for correspondence: Professor David R Bearn, University of Dundee, UK.
Email: d.bearn@dundee.ac.uk
© 2009 British Orthodontic Society
DOI 10.1179/14653120722914

Electronic databases. All searches were initially conducted in January 2005 and no language restrictions were applied.

Two basic sets of terms were applied in the search. Firstly those terms used to identify records related to the health condition of interest (orthodontic alignment), and secondly those terms used to identify records related to the intervention being evaluated (archwire therapy). It is generally standard procedure to include a third basic set of terms which identifies those records related to the type of study designs to be included. This was not included in our final search because in pilot runs of the strategy it reduced the number of papers found by all databases to zero. Details of the searches of electronic databases are shown in Table 1.

Search strategies were saved on each database. Each of these strategies was automatically run every month up to and including July 2008. This updating revealed possible further studies for consideration, of which one met the selection criteria.

Other sources (reference lists). The bibliographies of papers and review articles identified were checked for studies published outside the electronically searched journals which may have been otherwise missed.

Selection

Study design. It was decided to include both randomized clinical trials and controlled clinical trials to have the best chance of finding all evidence of an acceptable level.

Participants. Children and adults who had aligning and/or levelling archwires used as part of orthodontic treatment.

Table 1 Search strategies

No.		Search history	Results
CENTRAL*	1	exp ORTHODONTICS/	615
	2	orthodont$.mp.	1049
	3	1 or 2	1149
	4	align$.mp.	345
	5	ORTHODONTIC WIRES/	42
	6	arch.mp.	471
	7	wire$.mp.	530
	8	6 and 7	36
	9	5 or 8	62
	10	3 and 4 and 9	10
MEDLINE	1	exp ORTHODONTICS/	31,737
	2	orthodont$.mp.	32,306
	3	1 or 2	35,798
	4	align$.mp.	86,042
	5	ORTHODONTIC WIRES/	1660
	6	arch.mp.	24,936
	7	wire$.mp.	21,658
	8	6 and 7	700
	9	5 or 8	2002
	10	3 and 4 and 9	89
EMBASE	1	exp ORTHODONTICS/	1813
	2	orthodont$.mp.	2783
	3	1 or 2	2783
	4	align$.mp.	29,454
	5	ORTHODONTIC WIRES/	341
	6	arch.mp.	16,079
	7	wire$.mp.	18,005
	8	6 and 7	201
	9	5 or 8	533
	10	3 and 4 and 9	11
mRCT†	1	Orthodontic% AND align% AND archwire%	6

Cochrane central register of controlled trials.
†Meta-register of controlled trials.

Intervention. Fixed orthodontic appliances consisting of brackets and archwires to achieve aligning and levelling. The type of archwires investigated would be used to put studies into homogenous groups, where applicable, for meta-analysis.

Outcome measures. Studies needed to report an objective measurement of alignment / irregularity to be included.

Validity assessment

Decisions on validity and quality assessment were made independently by two assessors. The factors which were considered when assessing the quality of the studies were:

- Was the sample size reported?
- Was the sample size based on a power calculation?
- Were the eligibility criteria described?
- Was assignment to groups random?
- Was treatment allocation concealed?
- Were groups similar at baseline in terms of prognostic factors?
- Was the care provider blinded?
- Were the patients blinded?
- Were outcome assessors blinded to treatment allocation?
- Were the point estimates and measure of variability presented for the primary outcome measure?
- Were the statistical methods used to compare the groups appropriate?
- Did the analysis include an intention to treat analysis?

On the basis of these questions, a score out of 12 was calculated for each study. A consensus meeting was then arranged to discuss which studies showed low risk of bias and should therefore be included for data extraction.

Data abstraction

Two assessors independently extracted data from the eligible trials and this was entered into customized data abstraction forms. The following details were recorded:

General details

- Name of paper (Author and year)
- Name of assessor

Care setting

- Location of care setting

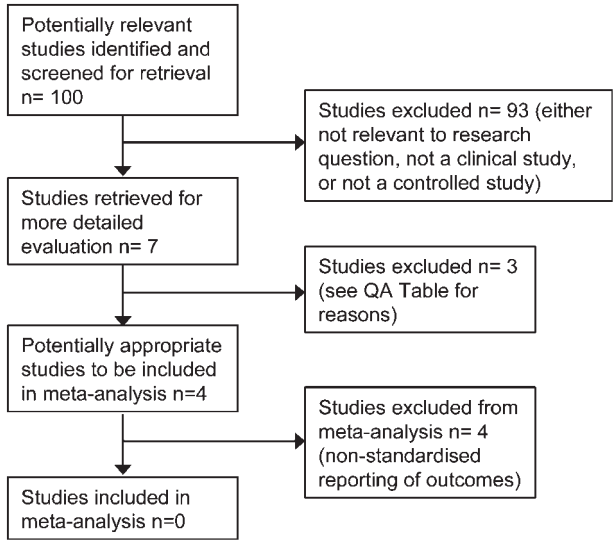

Figure 1 Study flow, as described in QUORUM statement

Participant details

- Age
- Gender
- Number of subjects in each group

Methodological quality

- Study design (RCT or CCT)

Interventions

- Types of archwire investigated

Outcome measures

- Which teeth were measured?
- How the teeth were were measured at baseline?
- What was measured after the intervention?
- What was used to take the measurements?

Results

- How much alignment was achieved with each archwire?
- Data was extracted according to the authors' method of assessment of alignment. This meant there were three broadly different possible types of data to extract for each trial:

  (i) What was the change in total irregularity for the given period of time?
  (ii) What was the change in contact point movement for the given period of time?

(iii) What was the duration taken to achieve 2 mm irregularity index?

- How many drop outs were there and what was the duration of the trial?

Meta-analysis. Once data extraction was completed, it was planned to enter this into Revman to undertake a meta-analysis and produce forest plots showing the overall effect of the archwire interventions.

Results

The study flow is shown in Figure 1 as described in the QUORUM statement.

Studies identified

The initial search strategy identified 89 studies, to which 11 further were added by the updated searches, giving a total of 100 studies. Inspection of the abstracts of these studies revealed seven[3,5–10] to be appropriate for further appraisal. A summary of the key methodological points of these studies can be seen in Table 2.

Quality assessment

Individual quality assessment results for each study are shown in Table 3. After discussion of the potential bias present in the studies a cut off score of 6 was decided on for inclusion as representing a level of methodological rigour that would mean low risk of bias in the results. The risk of bias in each study is discussed in more detail below. Detailed examination of the study reported by Pandis *et al.*[10] revealed that although different archwires were used in the two groups, different bracket systems were also used (this was the main question being addressed by the study) and this confounding factor was determined to introduce significant bias to the effect of archwire on alignment. This left four studies remaining for data abstraction.

Data extraction. A summary of the data extracted from the four remaining studies can be seen in Table 4, summarising the care setting, participant details, methodological quality, interventions and outcome measures and the results of each trial.

Data synthesis. Due to a lack of homogeneity between the studies, a meta-analysis was not possible.

Discussion

There have been only six controlled clinical trials of aligning archwires up to the time of the final search.[5–9,11] These trials aimed to test whether there was a difference in time taken to achieve alignment between archwires and compared either a multistranded stainless steel wire and nickel titanium archwire, or different types of NiTi archwire. Prior to these trials (pre-1990), various archwires were being used for levelling and aligning without having been scrutinized by clinical trial.

A summary of all of the controlled clinical trials of aligning archwires can be seen in Table 1.

Care setting

Samples seem to have been drawn exclusively from university teaching hospitals but it was only Cobb *et al.*[8] that actually made this explicit. When trials are undertaken in such a setting, it has been suggested that they might not be relevant to the real world as they measure the efficacy of the intervention rather than its effectiveness.[12] If patients are seen in less 'ideal' conditions in orthodontic practices than those in teaching hospitals with respect to archwire placement then this suggestion may be true. Conversely as many university departments are training establishments, it is likely that the operators in this setting, if trainees, were less experienced. The number of weeks between each adjustment of the appliance could well be longer in government funded practice than university teaching hospitals and this may have yielded different results between the archwires had the trials been carried out there instead.

Sample size

Sample sizes have ranged from 15–123 patients or 15–158 arches. Only two of the six trials reported power calculations, West *et al.*[7] and Evans *et al.*[11] Despite a power calculation, Evans *et al.* had a sample size below the threshold which had been set after drop outs. It is not known whether West *et al.* met their power calculation threshold as drop outs were not published. It is possible that samples in the remaining studies lacked the power to detect differences between the archwires.

Age and gender of subjects

Where reported, the ages of subjects in the trials were broadly similar.[7,8] The gender distribution was not widely reported – in fact, only West *et al.*[7] reported it, where females outnumbered males by 2 : 1. There is, however, no reason to suspect that there would be a

Table 2 Summary of the key methodological points of clinical trials of aligning archwires identified for detailed evaluation

		Sample		Method				Blinding		Statistics
Paper	Study type	Power calculation	Sample size*	Teeth measured	Measurement technique	Measurement error†	Follow up	Operator blind?	Examiner blind?	Analysis of distribution
O'Brien *et al.* 1990	RCT	No	40 pts 40 aws	↑ 3-3	Digitized (3D metrograph)	0.17 mm	35 days	No	Yes	No
Jones *et al.* 1990	RCT	No	43 pts 43 aws	↑/↓ 6-6	Digitized (3D metrograph)	Not assessed	35 days	No	Yes	No
West *et al.* 1995	RCT	Yes (74 aws)	62 pts 74 aws	↑/↓ 6-6	Digitized (Reflex Microscope)	0.17 mm	42 days	No	Yes	No
Cobb *et al.* 1998	RCT	No	123 pts 155 aws	↑/↓ 3-3	Digital calipers	<0.10 mm	51 days (to 2 mm LII)	No	No	Yes
Evans *et al.* 1998	RCT	Yes (100 aws)	51 pts 98 aws	↑/↓ 6-6	Digitized (Reflex microscope)	0.08mm	56 days	No	No	No
Dalstra *et al.* 2004	Split mouth	No	15 pts 15 aws	↑ 6-6	Computer aided photographic analysis	Not assessed	21–35 days	No	No	Yes
Pandis *et al.* 2007	RCT	No	54 pts 54 arches	↓ 3-3	Digital caliper	Not assessed	Mean of 103 days	No	No	Yes

*Pts = number of patients, aws = number of archwires.
†Root mean square error.

Table 3 Results of quality assessment of papers selected for detailed appraisal

	O'Brien *et al.* (1990)	Jones *et al.* (1990)	West *et al.* (1995)	Cobb *et al.* (1998)	Evans *et al.* (1998)	Dalstra *et al.* (2004)	Pandis *et al.* (2007)
Was sample size reported?	✓	✓	✓	✓	✓	✓	✓
Was sample size based on a power calculation?			✓		✓		
Were eligibility criteria described?	✓	✓	✓	✓	✓		✓
Was assignment to groups random?	✓		✓	✓	✓		✓
Was treatment allocation concealed							
Were groups similar at baseline in terms of prognostic factors?	✓			✓			
Was the care provider blinded?							
Were the patients blinded?							
Were outcome assessors blinded to treatment allocation?							
Were the point estimates and measure of variability presented for the primary outcome measure?	✓	✓	✓	✓	✓	✓	✓
Were the statistical methods used to compare the groups appropriate?	✓	✓	✓	✓	✓	✓	✓
Did analysis include an intention to treat analysis?	✓	✓				✓	
Score	7	5	6	6	6	4	5

difference in archwire performance between the genders so the absence of this data is not critical.

Quality assessment

The purpose of the quality assessment was to identify possible sources of bias, and make a judgment based on this as to which studies to include. The cut off score of 6 which was used is to some degree arbitrary, but in order to make the rationale behind this clear the possible biases identified are listed and discussed here.

Design of studies. This systematic review sought to find only controlled clinical trials and randomized controlled clinical trials. Given the number of different types of archwires currently available, it was somewhat surprising that only six of these types of trials were found (five randomized clinical trial and one controlled clinical trial).

Sample size and eligibility criteria. All six trials reported their sample size but only two of these samples had been based on power calculations. Eligibility criteria were described in all trials except Dalstra *et al.*[9]

Was assignment to groups random? Random allocation to treatment group is important because it gives groups that are likely to be balanced for known as well as unknown confounding variables. Jones *et al.*'s study[6] was not a randomized controlled trial and was therefore susceptible to selection bias. Dalstra *et al*. used a spit mouth design with random allocation to quadrants within each subject.

Was treatment allocation concealed? Without concealment of the randomly generated allocation sequence, the operator may choose to manipulate it so that subjects are allocated to the group that the operator chooses. This is called selection bias and is one of the most important factors which may distort treatment comparisons.[13] None of the six trials had their treatment allocation reported as concealed.

Were the groups similar at baseline in terms of prognostic factors? It would seem important that the groups were comparable at baseline with respect to alignment as otherwise differences between groups after intervention may be attributed (wrongly) to the archwire rather than the fact that the groups were not comparable at baseline. Only O'Brien *et al.*[5] and Cobb *et al.*[8] confirmed that both groups were similar in terms of alignment at baseline.

Was the care provider blinded? Blinding of the care provider is necessary as otherwise performance bias (a systematic difference in care provided apart from the intervention being evaluated) is possible. The operators were not blind in any of the trials leading to the possibility of performance bias. However, it would have been very difficult in most trials to blind the operator as the wires either had a different number of strands[6–8,11] or were a different colour.[9] In O'Brien *et al.*'s study[5] blinding would have been possible, although the operator may have realized which wire they were using from its handling characteristics.

Were the patients blinded? Blinding of patients to their group allocation helps to stop reporting bias. None of the trials had patient blinding but reporting bias is not a risk for these studies as the patients are not reporting the primary outcome.

Were the outcome assessors blinded to treatment allocation? Detection bias is a systematic difference between comparison groups in how outcomes are ascertained and can be minimized by blinding the outcome assessors.

The assessment of irregularity was not blind in three of the studies (Cobb *et al*.,[8] West *et al*.[7] and Dalstra *et al*.[9]). Cobb *et al*. measured Little's irregularity index directly on the patients allowing the operator to identify the archwire in use. West *et al.* digitized contact point displacement on study models which included the archwire – the multistranded could have been differentiated from the single stranded archwires. Dalstra *et al.* took an occlusal photograph on which the type of wire could be identified. In all of these trials there was the possibility of detection bias.

Were the point estimates and measure of variability presented for the primary outcome measure? Without the mean and a measure of variability for the primary outcome measure, data cannot be extracted and a meta-analysis performed. All studies included such data and from this point of view were admissible to the data abstraction stage.

Were the statistical methods used to compare the groups appropriate? Broadly speaking, the statistical methods used to compare the groups were appropriate. However, only Cobb *et al.*[8] and Dalstra *et al.*[9] analysed the distribution of the data. On both occasions this showed that alignment time/extent was not normally distributed. Consequently non parametric statistics were used. In the remaining four trials parametric statistics were used

Table 4 Details of study design, outcome measures reported and results for the four studies included in data abstraction

Author (year)	O'Brien *et al.* (1990)	West *et al.* (1995)	Cobb *et al.* (1998)	Evans *et al.* (1998)
Location	Not stated	Not stated	UNC	'2 Centres'
Participants details				
Sample size (number of arches)				
Rx group 1	20	36	–	–
Rx group 2	20	38	–	–
Rx group 3	–	–	–	–
All groups	40	74	158	112
Mean age of subjects (sd)				
Rx group 1	–	15.4 (5.8)	15.2 (3.8)	–
Rx group 2	–	14.4 (3.3)	17.3 (6.7)	–
Rx group 3	–	–	16.3 (5.1)	–
Gender of subjects (% male, % female)				
Rx group 1	–	–	–	–
Rx group 2	–	–	–	–
Rx group 3	–	–	–	–
All groups	–	33,66	–	–
Methodological quality				
Design of study	RCT	RCT	RCT	RCT
QA score	7	6	6	6
Intervention				
Rx group 1	0.016 Titanol	0.014 NiTi	0.016 ion implanted NiTi	0.016 × 0.022 active martensitic NiTi
Rx group 2	0.016 NiTi	0.0155 multistranded stainless steel	0.016 NiTi	0.0155 multistranded stainless steel
Rx group 3			0.0175 Multistranded stainless steel	0.016 × 0.022 active martensitic NiTi (graded force)
Outcome measures				
Teeth measured	Upper 3-3	Upper/lower 6-6	Upper/lower 3-3	Upper/lower 6-6
Measurements at baseline	3D contact point position with respect to palatal rugae	Sum of contact point displacements	Little's irregularity index	Sum of incisal edge irregularity 2-2 and sum of inter-bracket span 3–6
Measurements after intervention	3D contact point position with respect to palatal rugae	Sum of contact point displacements	Little's irregularity index	Sum of incisal edge irregularity 2-2 and sum of inter-bracket span 3–6
Device used to take measurement.	Reflex metrograph	Reflex microscope	Digital callipers	Reflex microscope
Results				
Mean duration of trial in days (sd)				
Rx group 1	34 (2)	42.7 (3.2)	–	57.1 (−)
Rx group 2	37 (2)	42.3 (1.2)	–	57.3 (−)
Rx group 3	–	–	–	57.5 (–)
All groups	–	–	–	–
n of arches at baseline				
Rx group 1	20	36	–	–
Rx group 2	20	38	–	–
Rx group 3	–	–	–	–
Total	40	74	158	112
n of arches at end of trial				
Rx group 1	20	36	–	32
Rx group 2	20	38	–	31
Rx group 3	–	–	–	35
Total	40	74	155	98

without confirming their suitability with distribution analysis.

Did the analysis include an intention to treat analysis? An intention to treat analysis is one where participants are analysed according to the group to which they were initially allocated. It is important to analyse data on an intention to treat basis because otherwise attrition bias (systematic differences between comparison groups in terms of withdrawals or exclusions of participants) is possible. O'Brien *et al.*[5] Jones *et al.*[6] and Dalstra *et al.*[9] analysed data on an intention to treat basis but this was only by virtue of the fact that there were no drop outs. In the remaining trials, West *et al.*[7] did not report on whether there were any drop outs and Evans *et al.*[11] had drop outs but omitted them from analysis.

Intervention. A summary of the different interventions in each of the four trials finally included can be seen in Table 5. Investigation of the same intervention by more than one author would potentially make a meta-analysis of this intervention possible. 0.016 NiTi was investigated by both O'Brien *et al.* and Cobb *et al.* whilst 0.0155 stainless steel was investigated by both West *et al.* and Evans *et al.* However, a meta-analysis is only possible if the outcomes measured and reported are the same.

Outcome measures

Teeth measured. Jones *et al.*,[6] West *et al.*,[7] Evans *et al.*[11] and Dalstra *et al.*[9] measured the movement of '6-6' so that both anterior and posterior teeth could be analysed. All of these trials were on the upper and lower arches except for Dalstra *et al.*'s[9] which was restricted to the upper arch only. O'Brien *et al.*[5] and Cobb *et al.*[8] measured '3-3' only. O'Brien *et al.* chose to measure just '3-3' because most crowding was present here and measurement of posterior contact points is subject to more error.[7,14]

Assessment of archwire performance. The extent of alignment achieved by each archwire was assessed in three broadly different ways:

1. 3D contact point position with respect to palatal rugae: O'Brien *et al.*[5] assessed archwire performance by measuring the 3D movement of each anatomical contact point. This was achieved by digitizing each anatomical contact point with respect to the medial palatal rugae at the beginning of the trial and then repeating these measurements at the end. The mean contact point movement was then calculated for each archwire.

Whilst the internal validity of this form of assessment appears to be good, the external validity is perhaps not as good. As a purchaser of archwires, an operator might want to know whether alignment with one archwire is faster than another. The form of assessment used by O'Brien *et al.* will tell us whether contact points move faster with one archwire than another but it won't tell us how many weeks faster this translates to.

2. Sum of contact point displacements: Jones *et al.*,[6] West *et al.*[7] and Cobb *et al.*[8] all summated anatomical contact point displacement to give a sum of contact point displacement both before and after archwire therapy.

Evans *et al.*[11] measured the displacement of incisal edges anteriorly whilst effectively measuring inter-bracket span posteriorly. This irregularity was then summated to give a sum of 'inter-tooth distances' before and after archwire therapy. The reason given for measuring irregularity in this way was that it reduced the measurement error posteriorly.

The internal validity of the assessment of alignment used by Jones *et al.*, West *et al.* and Cobb *et al.* is good; it quantifies the amount of irregularity before and after archwire therapy. However, the internal validity of Evans *et al.*'s assessment of alignment is questionable. Whilst measuring the inter-bracket span of posterior teeth may reduce measurement error, this is at the cost of not

Table 4 Continued

Author (year)	O'Brien *et al.* (1990)	West *et al.* (1995)	Cobb *et al.* (1998)	Evans *et al.* (1998)
Mean movement of contact points with respect to palatal rugae (sd)				
Rx group 1	1.7 (1.15)	–	–	–
Rx group 2	1.42 (0.79)	–	–	–
Rx group 3	–	–	–	–
Mean change in total irregularity in mm (sd)				
Rx group 1	–	–	–	2.54 (−)
Rx group 2	–	–	–	2.58 (−)
Rx group 3	–	–	–	2.65 (−)
Mean time to decrease irregularity to 2 mm				
Rx group 1	–	–	–	–
Rx group 2	–	–	–	–
Rx group 3	–	–	–	–

actually measuring the irregularity of the teeth themselves. For example, when considering correction of an individual tooth rotation during alignment, it is conceivable that the sum of 'inter-tooth distances' would remain the same despite the fact that considerable alignment has been achieved. A common example where this may occur is when two premolars are rotated towards each other prior to alignment, producing a very short inter-bracket span.

The external validity of assessing archwire performance by the degree of change in the sum of irregularity over time is reasonable but it is open to the same criticism as O'Brien *et al.*'s methodology – it does not tell us how many weeks will be saved by using one archwire over another. Cobb *et al.*[8] used the time taken to achieve 2 mm total irregularity as their endpoint. This approach has more external validity in that one can then tell how much time might be saved between archwires to reduce total initial irregularity to 2 mm.

3. Computer aided photographic analysis: Dalstra *et al.* assessed archwire performance in their split mouth study by measuring irregularity on each side of the mouth before and after archwire therapy. This was achieved by taking intra-oral photographs of the occlusal aspect of the upper arch at the beginning and end of alignment and analysing them. The photographs were taken with a custom made plexiglass plate fitted over the stable medial palatal rugae which had a grid of perpendicular lines drawn on it. The x and y co-ordinates of the bracket wings relative to the grid were measured pre and post alignment and this movement was compared between sides. There was no investigation of the validity or reproducibility of this technique.

Data abstraction

Duration of trial. The period over which alignment was assessed varied considerably from 21 to 56 days. Only Cobb *et al.* used a clinical end point for initial alignment (2 mm Little's irregularity index) rather than an arbitrary period of time. They found the median time to initial alignment was 51 days. Only Evans *et al.*[11] had a follow up of more than this and so the remaining trials[5–7] may not have fully reflected the alignment phase.

Meta-analysis. A meta-analysis was not possible because even though the same archwires were assessed in multiple studies, there was no homogeneity in the way that their performance was reported; no one methodology was used more than once. In addition to this, data could not be extracted from either West *et al.* or Evans *et al.*'s papers. In the case of West *et al.*, data were presented as geometric mean ratios of total irregularity of NiTi/Multistranded stainless steel with no data on the actual amount of irregularity for each archwire at each interval. In the case of Cobb *et al.*, the data was only presented in graphical form with insufficient detail for data to be extracted.

Only one of the four trials showed a significant difference between the wires studied and on both occasions the clinical significance was questionable; West *et al.*[7] showed that NiTi was significantly quicker at aligning teeth in the lower labial segment than multistranded stainless steel although the clinical significance of this difference appears to be small.

Evans *et al.* suggested that in the case of multistranded stainless steel and NiTi, differences may not have been found because the NiTi wires were not clinically deformed enough to take advantage of their superelastic properties.

In summary, it would appear that there have been little clinically significant differences found between the archwires studied.

Conclusions

Seven clinical trials have been undertaken investigating aligning archwires. After quality assessment, four were selected for data extraction but due to a lack of homogeneity, a meta-analysis was not possible. There is insufficient data in these studies to make clear

Table 5 Details of the archwires compared in the included studies

	Archwires compared		
Author	Rx group 1	Rx group 2	Rx group 3
O'Brien *et al.* 1990	0.016 Titanol	0.016 NiTi	–
West *et al.* 1995	0.014 NiTi	0.0155 multistranded stainless steel	–
Cobb *et al.* 1998	0.016 NiTi (ion implanted)	0.016 NiTi	0.0175 multistanded stainless steel
Evans *et al.* 1998	0.016 × 0.022 NiTi (active martensitic)	0.0155 multistranded stainless steel	0.016 × 022 NiTi (graded force active marstenitic)

recommendations regarding the most effective archwire for alignment.

Of the seven clinical trials evaluating aligning archwires, none have been ideal. It would seem that the best trial design to evaluate the clinical performance of aligning archwires would:

- be a randomized controlled trial;
- standardized appliance variables other than the archwires;
- have a power calculation to establish the sample size;
- have adequate randomization, allocation concealment and blinding where possible to protect against bias;
- use a valid, reproducible, quick and well recognized assessment of alignment such as Little's irregularity index;
- have a long enough follow up to accurately reflect the alignment phase. This could be a clinically defined endpoint e.g. 2 mm irregularity rather than an empirical time point;
- use an intention to treat analysis to protect against attrition bias;
- include an analysis of the distribution of data so that the most suitable statistical tests are used;
- report data in its raw form to make it more amenable to meta-analysis.

References

1. Proffit WR. *Contemporary Orthodontic*s, 3rd Edn. London: Mosby, 2000.
2. Kusy RP. The future of orthodontic materials: the long-term view. *Am J Orthod Dentofacial Orthop* 1998; **113**(1): 91–95.
3. Evans TJ, Durning P. Aligning archwires, the shape of things to come? – a fourth and fifth phase of force delivery. *Br J Orthod* 1996; **23**(3): 269–75.
4. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions 4.2.6.* Chichester: John Wiley and Sons Ltd, 2006.
5. O'Brien K, Lewis D, Shaw W, Combe E. A clinical trial of aligning archwires. *Eur J Orthod* 1990; **12**(4): 380–84.
6. Jones ML, Staniford H, Chan C. Comparison of superelastic NiTi and multistranded stainless steel wires in initial alignment. *J Clin Orthod* 1990; **24**(10): 611–13.
7. West AE, Jones ML, Newcombe RG. Multiflex versus superelastic: a randomized clinical trial of the tooth alignment ability of initial arch wires. *Am J Orthod Dentofacial Orthop* 1995; **108**(5): 464–71.
8. Cobb NW, 3rd, Kula KS, Phillips C, Proffit WR. Efficiency of multi-strand steel, superelastic Ni-Ti and ion-implanted Ni-Ti archwires for initial alignment. *Clin Orthod Res* 1998; **1**(1): 12–9.
9. Dalstra M, Melsen B. Does the transition temperature of Cu-NiTi archwires affect the amount of tooth movement during alignment? *Orthod Craniofac Res* 2004; **7**(1): 21–25.
10. Pandis N, Polychronopoulou A, Eliades T. Self-ligating vs conventional brackets in the treatment of mandibular crowding: a prospective clinical trial of treatment duration and dental effects. *Am J Orthod Dentofacial Orthop* 2007; **132**(2): 208–15.
11. Evans TJ, Jones ML, Newcombe RG. Clinical comparison and performance perspective of three aligning arch wires. *Am J Orthod Dentofacial Orthop* 1998; **114**(1): 32–39.
12. O'Brien K, Wright J, Conboy F, *et al*. Effectiveness of early orthodontic treatment with the Twin-block appliance: a multicenter, randomized, controlled trial. Part 1: Dental and skeletal effects. *Am J Orthod Dentofacial Orthop* 2003; **124**(3): 234–43; quiz 339.
13. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998; **317**(7167): 1185–90.
14. Richmond S. Feasibility of categorising treatment difficulty – the use of three dimensional plotting. Cardiff: University of Wales, 1984.