Current Products and Practice Section
Evidence-based Orthodontics—How do I assess the evidence?

JAYNE E. HARRISON, B.D.S., F.D.S.(ORTH.), M.ORTH., M.DENT.SCI.
Orthodontic Department, Liverpool University Dental Hospital, Pembroke Place, Liverpool L3 5PS, U.K.

Introduction

In the first article on Evidence-based Orthodontics I described the stages involved in applying the evidence-based approach to clinical practice, where and how to find evidence and introduced you to the Oral health group of the Cochrane Collaboration (Harrison, 2000). Having found the papers that you hope will provide the information to help answer a clinical question, how do you assess whether these papers are worth reading, provide valid information and should influence your current clinical practice? In this article I will provide you with guidelines that will allow you to assess systematically the methods, analysis and interpretation of research papers.

Strength of Evidence

The research method used in a study will depend on what question the study is addressing and for any given clinical question some research designs will provide information that is more valid than others (Oxman *et al*., 1993; Greenhalgh, 1997; Table 1).

Hierarchy of Evidence

Well designed randomized controlled trials (RCTs), confirming the same hypothesis, have, for many years, been recognized as providing the strongest level of evidence of the treatment effect of therapeutic interventions (Green and Byar, 1984; O'Brien and Craven, 1995). However, with the development of systematic review and meta-analytic techniques we now see systematic reviews as the foundation stone of our pyramidal hierarchy of evidence (Deeks and Sheldon, 1995; Guyatt *et al*., 1995; Harrison *et al*., 1996; Antczak-Bouckoms, 1998; Table 2.) Although considerable weight is placed on the evidence from RCTs and systematic reviews of RCTs, these research methods are not appropriate to answer every question. It must be remembered that valuable information can be obtained from other levels of evidence and each has its role to play in providing evidence about the treatment we provide for our patients (Table 1). *Case reports* have been used constructively to alert clinicians of serious side effects of interventions (e.g. face bow injuries, Booth-Mason and Birnie, 1988).

At some time we have all taken part in a *survey* using a questionnaire and these can be used usefully to assess current clinical practice. The results from such surveys can then help clinicians to identify whether they are in our out of step with current practice (e.g. management of the orthodontic patient at risk from infective endocarditis, Hobson and Clark, 1995). Clinical surveys are the preferred method to assess disease prevalence or a new diagnostic criteria (e.g. to assess the need for orthodontic treatment in a given population, Holmes, 1992). The information from such surveys can then be used by health care providers to assess manpower requirements or eligibility for treatment within existing resources.

Case Series have been used effectively to describe new treatments that have the potential to improve the management and prognosis for certain patient groups where previously there had been no treatment (e.g. alveolar bone grafting for children with cleft palate, Boyne and Sands, 1972).

Cohort studies with literature or historical controls and analyses of *computer data bases* are frequently used research methods in orthodontics and can provide valuable

TABLE 1 Guidelines for selecting articles that are most likely to provide valid results to a given clinical question [modified from Oxman et al. (1993) and Greenhalgh (1997)]

Question	Most appropriate research method	Key Questions
Therapy	Clinical trial	1. Was the allocation of treatments to patients randomized? 2. Were the patients, clinicians, and/or assessors blind to treatment allocation? 3. Were all the patients who entered the trial accounted for and attributed at its conclusion?
Diagnosis or screening	Cross-sectional survey	1. Was there an independent, blind comparison with a reference standard? 2. Did the patient sample include an appropriate range of the sort of patients to whom the diagnostic/screening test will be applied in clinical practice?
Prognosis	Cohort study or longitudinal survey	1. Was there a representative patient sample, at a well defined point in their disease? 2. Was the follow-up sufficiently long and complete?
Causation	Case control study or cohort study	1. Were there clearly identified comparison groups that were similar with respect to important determinants of the outcome of interest? 2. Were outcomes and exposures measured in the same way in the groups being compared?
Summary of evidence	Systematic review	1. Did the review address a focused clinical question? 2. Were the criteria used to select articles for inclusion appropriate?

0301-228X/00/020000+00$02.00

TABLE 2 Hierarchy of evidence

Anecodotal case report
Cross-sectional survey
Case series without a control
Case-control observational study
Cohort study with a literature control
Analyses using computer databases
Cohort study with a historical control group
Cohort study with a contemporary control group
Unconfirmed randomized controlled clinical trial
Confirmed definitive randomized controlled clinical trials
Systematic review of randomized controlled clinical trails

information that can then be used to plan prospective clinical trials. However, these alternative research methods should not be considered easier routes to quicker (and cheaper) answers, but as second choice methods when there are compelling reasons why an RCT cannot be carried out (Ellenberg, 1981). The key features of the main research methods used in clinical orthodontic research, together with examples of papers where each method has been used, are given in Table 3.

Critical Appraisal

Reading journal articles can be time consuming, and in our busy lives we need to be able to identify those papers which

TABLE 3 The key features and examples of different research methods

Research method		Key features	Examples
Survey		Describes how things are now. Sample may include all or a random sample of the population of interest. Do not usually have separate control groups, but internal comparisons can be made.	
Cross-sectional		Data collected from sample members on one occasion.	Prevalence of orthodontic treatment need, (Holmes, 1992).
Longitudinal		Data collected from sample members on two or more occasions.	Changes in the orthodontic treatment need in the same sample over 4 years (Tarvit and Freer, 1998).
Cohort study		Describes what happens to patients without actively intervening with the treatment they receive. Can be prospective or retrospective. May have a separate control group or be uncontrolled.	
Uncontrolled	*Case Series*	An uncontrolled cohort study describing the outcome of treatment for a group of patients.	Alveolar bone grafting for children with cleft palate (Boyne and Sands, 1972).
	Case Report	A small case series describing the outcome of treatment of a few (<5–10) cases or reporting potential problems with treatment.	Outcome of treatment with a specific appliance (Harrison, 1998), injuries caused by face-bows (Booth-Mason and Birnie, 1988).
Controlled	*Literature*	Comparison made to information on patients in a published paper or growth study. Prone to chronological and/or geographical bias.	Comparison made to values calculated from several published studies (Stucki and Ingervall, 1998).
	Historical	Comparison made with patients treated previously in the same unit/place. Prone to chronological bias.	New appliance compared with one used in the past in the same department (Buchannan *et al*., 1996).
	Matched	Comparison made with patients who are similar in respect to one or two specific criteria. Prone to allocation bias.	Pairs of boys matched for maturity and skeletal discrepancy (Ömblus *et al*., 1997).
	Concurrent	Control group treated at the same time as the study group. Prone to allocation bias.	Patients treated at the same time but in different locations (Fox *et al*., 1997).
Clinical trial		Assess whether one health care intervention is better than another, a placebo or no treatment. Are prospective and controlled. Allocation to test/control groups is predetermined.	Competing interventions (Ash and Hay, 1996), a placebo (Anderson *et al*., 1997), no treatment (Harradine *et al*., 1998).
	Random	Allocation to patient/quadrant/tooth according to a sequence generated from a table of random numbers or its electronic equivalent. Minimizes risk of all forms of bias.	A list of randomly generated numbers was used to allocate patients to extraction or no treatment group (Harradine *et al*., 1998).
	Quasi-random	Allocation to alternate patients or according to date of birth, case note number, day of week, side of mouth, quadrant. Prone to allocation bias.	Allocation of different interventions to a specific quadrant (Erverdi *et al*., 1997).
	Haphazard	A group of patients is divided into groups. Prone to allocation bias.	Sample divided into two groups, (Sidhu *et al*., 1995).
Case control study		Asks what makes a group of individuals different with respect to treatment received or environmental factors. Retrospective and look back in time.	Factors influencing root resorption following fixed appliance treatment (Kaley and Phillips, 1991).
	Multi-variant methods	Identify factors which have a significant influence on the outcome of interest.	Factors associated with the standard and length of treatment (Taylor *et al*., 1996).
Review article		Summarizes information from several previously published papers on a specific topic	
	Narrative review	Based on haphazard selection of papers related to the subject of the review.	Impacted maxillary canines (Bishara, 1992).
	Systematic review	Papers are identified, critically appraised and the results synthesized according to a defined protocol.	Orthodontic treatment for posterior crossbites (Harrison and Ashby, 1998).
	Meta-analysis	Combines the results from several different clinical trials to obtain an overall estimate of the effectiveness of a particular intervention.	Orthodontic treatment for posterior crossbites (Harrison and Ashby, 1998).

TABLE 4 Structure of a paper

Introduction	Why the research was done
Method	How the study was done
	Who and what was studied
	Where the study was carried out
	How the results were analysed
Results	What was found
Discussion	What the results mean
	How they relate to the findings of others

are worth reading and disregard those that will add little to our knowledge or impact on our clinical practice. Critical appraisal is the process of assessing and interpreting evidence by systematically considering its validity, results and relevance to your own work [Critical Appraisal Skills programme (CASP), 1997]. It is possible for critical appraisal skills to be learnt by all members of the orthodontic team including those with no clinical training or prior knowledge of research methods (Milne *et al*., 1995). These skills can then be used by the individual to improve the effectiveness of their personal reading or in the group situation, e.g. departmental journal club, to direct reading and help prevent the all too common 'trashing' of papers.

Most scientific papers are written in a standard IMRaD format (Table 4). The decision on whether a paper is worth reading should be based on the design of the methods section, rather than the hypothesis, *P* value or speculation found in the discussion (Watson, 1980; Hall, 1994; Greenhalgh, 1997). The assessment of methodological quality of papers has been covered in many books on critical appraisal and evidence based medicine (Crombie, 1996; Greenhalgh, 1997; Sacket *et al*., 1997) and in a series of papers collectively titled 'Users' guide to the medical literature' (Oxman, Guyatt and colleagues 1993–1995). Although written for medics the principles contained within these books and articles can equally be applied to dentistry and, in turn, to orthodontics. When assessing a paper there are a series of questions that can be asked of its contents which can direct the reader to make an informed assessment of the methodological quality, results and relevance of the information reported. There are preliminary and standard appraisal questions that can be asked of all papers and then secondary questions related to the specific research method used (Crombie, 1996; Greenhalgh, 1997). These questions can be asked informally or incorporated into a structured checklist. Checklists direct the appraisal of an article to ensure that all areas are covered and can be used by individuals or in the group (journal club) environment.

Preliminary Questions

Question 1: Why was the study done and what question were the researchers asking? This information should be easily identified from the introduction which should state briefly what the background to the study was and why it was carried out. The objectives of the study are often contained towards the end of the *Introduction* or at the start of the *Methods* section.

Question 2: What type of study was done? The information to answer this question should be provided in the *Methods* section. Using this information the study can then be classified according to one of the designs described in Table 3.

Question 3: Was this research method appropriate to the question being asked? This answer to this question can be obtained by integrating the information gained about the 'question asked' and 'research method used' as demonstrated in Table 1.

If the answers to these preliminary questions are clear and positive it may then be worth reading the paper in more detail. Having decided which research method has been used checklists that include the standard appraisal questions together with those specific to each research method can be used to assess the methodological quality, results, and relevance of the information reported.

Standard Appraisal Questions

Abstract

Is the abstract structured? An abstract should summarize the paper to allow the reader to quickly assess whether the paper is of interest to them. This is best achieved in a structured abstract where authors give details about the objectives, research design, setting, participants, interventions, main outcome measures, results and conclusion of the study in a systematic way (Haynes *et al*., 1990; Harrison *et al*., 1996). As I described in the first article on Evidence-based Orthodontics, structured abstracts also facilitate recognition of relevant articles when using computerised searches, contain more information for the reader than non-structured ones and are now being used in the *Journal of Orthodontics* (Jones, 1998; Harrison, 2000).

Introduction Section

Are the aims clearly stated? The reasons for the study being carried out and the question(s) being addressed by the study should be clearly stated and precise. This allows the reader to assess whether the research is investigating an important topic. Precise aims suggest that the study has been designed to answer specific questions which have been asked before the study began. If the aims are rather ambiguous it may suggest that 'date dredging' has been carried out with questions being posed after the event to match the interesting findings.

Methods Section

The methods section of a paper should state clearly how the study was carried out including who was studied, how they were selected and assessed and how the data was analysed.

Who was studied, how were they recruited, and where was the study carried out? It is important that the reader is told who was included in the study, how were they recruited and where was the study carried out to allow an assessment of whether the of the findings of the study can be generalized. Characteristics of the patients studied and details of how the participants were recruited gives an indication of whether the subjects were likely to be typical of the population of interest or were likely to be different in any way. The location of the study will give the reader an indication as to whether the findings are applicable to their own clinical situation.

How many participants were studied and was the sample size justified? Any research should include sufficient participants to have a high chance of detecting a difference between groups if there is one and be reasonably sure that one doesn't exist if none is found by the study. This is known as the *power* of the study. Small studies tend to be under powered and are unable to detect an important difference in effect even if there is one present. This is known as a *false negative* result or *Type II or* β error. The sample size required to detect a difference in effect should be calculated at the protocol writing stage of a project. It can be determined using statistical formulae, tables, software or nomograms (Pocock, 1983; Altman, 1991).

Was the study adequately controlled? For clinical trials, comparing two or more interventions, the control group should ideally be determined by randomization which has a specific meaning and is different from allocating patietns in a haphazard way. Random allocation should only be claimed when it is performed using a predetermined, concealed method, and the clinician responsible for recruiting patients does not know, and cannot predict or influence which treatment the next patient would get if recruited to the trial. Random allocation ensures that the treatment and control groups are balanced, within the limits of chance variation, with respect to all known and unknown confounding factors (Newcombe, 1994). Other methods of constructing a control group, e.g. haphazardly, allocation to alternate patients or related to date of birth or case note number, have the potential to bias the allocation to treatment and control interventions. These biases may then be greater than any difference in treatment effect and discredit the findings of the trial (Altman and Bland, 1999). For other research designs, (surveys, case-control studies, cohort studies) alternative or no controls may be appropriate. Surveys (cross-sectional or longitudinal) are often uncontrolled but it is important that the patients are *selected* randomly, be it on an individual or group (e.g. school, clinic, practice) basis, so as not to bias the 'type' of patients included in the survey. The control group selected for case-control studies needs to be as similar as possible to the affected group with respect to their exposure to the environmental conditions or intervention (e.g. extraction pattern, appliance system, operation), which is thought to have caused the disease or side-effect being studied. Cohort studies should have concurrent controls, rather than historical or literature controls which tend to be biased with respect to differences in time, location, and/or population compared to the intervention group. Case series and case reports are, by their very nature, uncontrolled.

Were the assessors blind to the interventions received and are any measurements taken likely to be valid and reliable? Prior knowledge of the intervention received or the stage of assessment can consciously or subconsciously bias the assessors which may result in an over- or under-estimation of the true measurement. In many drug trials it is possible for the patient, clinician and/or assessor to be blinded to the drug received or the stage of assessment (double or triple blind trials). However, in orthodontics this is harder as the interventions we are assessing (e.g. extractions, appliances) are difficult to camouflage. Whilst accepting that total blinding cannot always be achieved in orthodontics, we need to ensure the greatest degree of blinding possible and make every effort to blind assessors when assessing radiographs or study models. The validity of measurements (extent to which it measures what it is supposed to measure) can be a problem where assessments are subjective (e.g. aesthetics, pain) or influenced by other factors (e.g. overjet measurement and mandibular posturing) and efforts should be made to use scales that have been validated when measuring such outcomes. Reliability is important especially if assessments are made on different occasions and/or by different examiners. Studies should state the method for assessing the reliability of any measurements taken and what the inter- and intra-examiner reliability was (Roberts and Richmond, 1997).

Are the statistical methods described and are they appropriate? What comparisons of data, sub-group analyses and the statistical approaches needed to analyse them should be determined at the protocol stage of a study. The use of inappropriate statistical methods can produce misleading results and multiple significance testing increases the likelihood that spurious significance will be found. The use of over complicated or obscure methods should also be viewed with suspicion.

The type of methods required are determined by the type of data collected. Key questions about data that need to be asked to determine the most appropriate statistical methods include:

1. Is the data normally distributed?
2. Is the data continuous or categorical?
3. Is the data taken from independent samples?

The inter-relationship between type of data, most appropriate statistical method and it's purpose together with an example is described well in Greenhalgh (1997).

Results Section

The results section should present the data on what the study found. The results should be presented in a logical order with the basic data and simple analyses being presented first before proceeding to more complex comparisons and analyses (if appropriate).

Were the basic data about the sample described and baseline comparisons made? All studies should report the number of participants at the start of a study, together with details of how many of them completed the study and reasons for incomplete follow-up of participants. All participants should be accounted for so the number included in analyses are either consistent or any variations are explained. Informal comparison of the baseline characteristics of participants should be made so that if any differences are found they can be compensated for and differences between the groups at the end of the study can be attributed to the intervention being assessed, rather than pre-existing differences between the groups.

Was the statistical significance of the results assessed? The statistical significance of a result gives an indication of the probability of that result having occurred by chance alone. In scientific papers the level of significance that is taken to be significant is usually $P < 0{\cdot}05$. This is equivalent to a

chance of 5 in 100 or 1 in 20 that such a result could have occurred by chance alone. However, a statistically significant result does not rule out the possibility that the result has arisen by chance. If the level of statistical significance is set at $P < 0{\cdot}05$, for every 20 statistical tests of significance that are done, one will be significant by chance alone. This is of particular concern in some orthodontic studies when numerous cephalometric measures are tested for significance. One of the easiest mistakes to make (or ways of 'cheating' with statistics) when analysing the results of a study is to put all your data into a computer statistical package and report any results that emerge as significant whilst ignoring those where $P > 0{\cdot}05$. For this reason it is best to limit the number of variables that are assessed, specify which these are going to be at the protocol stage of the study and quote the calculated P value, rather than just whether it is greater or less than 0·05.

Probability (P) values are the most usual way of reporting statistical significance, but increasingly confidence intervals (CIs) are the preferred way to present this information (Gardner and Altman, 1986; Crombie, 1996; Greenhalgh, 1997). CIs are an alternative to quoting P values, whilst also providing information on the limits within which we can expect the true result to lie, how large we can expect the effect to be and how precise the results are. It is usual to use the 95 per cent CI which is informally interpreted as the range within which we are 95 per cent certain that the true value lies. If the 95 per cent confidence interval lies one or other side of zero (e.g. 15 to 45 or −25 to −10) this corresponds to a statistically significant difference between the groups at the 5 per cent level. However, if a confidence interval spans zero (e.g. −4 to 16) this indicates that there is little evidence for a difference between the groups and is equivalent to $P > 0{\cdot}05$. The point estimate—uually the midpoint of the interval—is an indication of the size of the difference between the groups. The width (or range) of the CI is evidence of how precise the results are and is related to the standard error and hence the standard deviation and sample size. A narrow CI indicates good precision whilst a wide CI should be viewed with caution because the meaning of the estimated size of effect is questionable. Increasing the sample size increases precision and reduces the width of the CI. Generally, to halve the width of the CI requires increasing the sample size by a factor of four.

Discussion Section

The discussion section of a paper should include a summary of the main findings of the study and then relate them to any deficiencies in the study design or problems in the conduct of the study. They can then be related to previous work in the area, whether they can be generalised and their clinical implications. The interpretation of data is not as clear cut as it may seem and several factors have to be taken into account when trying to determine what the results actually mean.

What are the main findings and does the data support them? A summary of the main findings should be given and then the size of each effect examined to assess whether it is clinically significant and, if so, to whom are the results applicable. Statistical significance does not always imply clinical significance but confidence intervals are helpful in providing information on the range within which the true value lies. The results then need to be interpreted in light of any factors which may have biased the results.

Did any untoward events occur during the study? Unplanned events can happen at any stage of the study and may result in subjects who should have been included being missed, measurements not taken or subjects lost to follow-up. All these factors can lead to data being missing from the final analysis. Whilst some missing data is expected in most studies, those where there is a substantial amount of missing data should be read with caution. In a survey it cannot be assumed that non-responders are similar to responders and every effort should be made to minimise the amount of missing data.

How are null findings interpreted? Apparently non-significant results need to be interpreted with care. Lack of evidence of a difference in effect does not necessarily mean that there was no difference in effect (Altman and Bland, 1995). The same can be said for studies into causation (case control studies). Lack of evidence that A causes B does not necessarily mean that this is the case. Again confidence intervals are useful in assessing the precision of the results. Narrow CIs that span the point of zero difference, suggest that the study results can be viewed with a degree of certainty that there was no difference in effect or causative link found. However, if the CIs are wide and span the point of zero difference it can be indicative that the study is inconclusive and may have been too small (inadequately powered) to detect a difference in effect or causation even if one existed.

How do the findings of this study relate to previous work in the area? It is unusual to find that there have not been any other studies carried out in any area of research so the results of a single study should not be seen in isolation, but interpreted in the light of other studies. It is important to give a balanced view of previous work and see the results of the new study in context of previous work. Where there is a considerable body of knowledge it is tempting for the author to overemphasise studies that support his findings and play down those that don't.

Who are the results applicable to and will they affect my clinical practice? This is often the bottom line of critical appraisal and involves integrating information gained at all the other stages of the assessment. Key factors include the

(1) population to whom the results will apply;
(2) setting in which the study was carried out;
(3) quality of the study design, conduct, and analysis;
(4) clinical significance of the results;
(5) likelihood that the result are valid.

All these points need integrating and an assessment made as to whether you can expect the results of the study to apply to the patients you treat in your particular clinical circumstances.

Specific Questions for Each Research Method

Having answered the preliminary and key questions and identified the research method used, specific questions re-

TABLE 5 Specific questions for each research method [modified from Crombie (1996)]

Design	Questions	Justification
Survey	*Who was studied?*	Allows assessment of whether the results can be generalized and are relevant to your patients.
	How was the sample obtained?	Determines whether the sample is potentially representative of the population of interest and that each member has the same chance of being selected to participate in the survey.
	What was the response rate?	A response rate of less than 80 per cent could threaten the validity of the survey as non-responders or those lost to follow-up may differ in some way that may bias the results.
Cohort study	*Who exactly has been studied?*	Allows assessment of whether the results are representative of the treatment effects that can be expected for all patients with a particular disease (malocclusion) or whether they could have been biased (positively or negatively) by the patient group studied.
	What type of control group was used?	A control or comparison group that receives either an alternative treatment, placebo or no treatment is necessary to allow a meaningful assessment of the treatment effect of a particular intervention. Most types of control have limitations and can be prone to bias e.g. chronological, environmental, racial, geographic.
	How adequate was the follow-up?	This should assess three factors (1) the proportion of patients followed-up; (2) whether assessment of outcome was made 'blind' to the stage of treatment or treatment received; (3) whether the length of follow-up was appropriate.
Clinical trial	*Were the interventions allocated randomly?*	Random allocation to the test and control groups minimizes the risk of there being systematic differences between the baseline characteristics of the comparison groups and thus reduces bias.
	Were all participants accounted for?	All participants should be accounted for and note made of the number lost to follow-up so that an assessment can be made on the impact of these losses on the results and their interpretation.
	Were the outcomes assessed blind?	It is important that the assessor is blind to the intervention received so that the risk of systematic differences in outcome assessment occurring is minimized.
Case control study	*How were the cases obtained?*	Selection of the 'diseased' and 'control' cases is the area that is most prone to bias. The source of cases, disease stage, and severity should be defined to allow assessment of whether selection bias could have occurred.
	Was the control group appropriate?	The control group should be selected from the same source as the 'disease' group and be as similar as possible with respect to all factors except the disease/side effect of interest.
	Were data collected in the same way for cases and controls?	Data should be collected in the same way (e.g. from case notes, interview) from both groups and be obtained by persons 'blind' to which group the participants belong.
Review article	*How were the papers identified?*	Details of the method for identifying papers will allow an assessment of the degree to which bias could have occurred in selecting papers that were included in the review.
	How was the quality of the papers assessed?	It is important that evidence from good quality studies is given weight over that from studies that are methodologically weak.
	How were the results summarized?	It is important to check the quality of the studies, comparability of subjects, settings, interventions, and outcome measures in order to assess whether it is appropriate to combine the results of several studies in a meta-analysis.

lated to each method can be incorporated into the appraisal process or checklist. These questions and the reason for asking them are presented in Table 5.

Checklists

In this article I have described how to identify the research design used in a study, how to assess the strength of the evidence it provides and the process of critical appraisal. The process of critical appraisal includes a series of questions that start with preliminary questions that allow you to identify whether the paper is worth reading. The next questions are those that are specific to the research method used and these are followed by the standard questions that can be asked of any paper. Questions can also be asked that are related to the quality of the abstract, the interpretation of the results and their implication on the clinical practice of the reader(s). All these questions can be brought together to form a checklist that can be used by an individual or in a group (journal club) environment to allow papers to be assessed in a systematic way. An example of a complete checklist for use with surveys is shown in Appendix 1. This has been derived from those published by Crombie (1996), Greenhalgh (1997) and Oxman, Guyatt and colleagues in the *Users' Guide to Medical Literature* (1993–1995). Checklists that have been devised for other research methods can be obtained directly from the author.

Acknowledgements

I would like to thank Dr R. Newcombe for his constructive comments and help in modifying previous versions of this article.

Recommended Reading

Crombie, I. K. (1996)
The Pocket Guide to Critical Appraisal,
British Medical Journal Publishing Group, London.

Greenhalgh, T. (1997)
How to Read a Paper: the basis of evidence based medicine,
British Medical Journal Publishing Group, London.

Sackett, D. L., Richardson, W. S., Rosenberg, W. and Haynes, R. B. (1997)
Evidence-based Medicine,
Churchill Livingstone, Edinburgh.

References

Altman, D. G. (1991)
Practical Statistics for Medical Research,
Chapman & Hall, London

Altman, D. G. and Bland, J. M. (1995)
Absence of evidence is not evidence of absence,
British Medical Journal, **311**, 485.

Altman, D. G. and Bland, J. M. (1999)
Testament allocation in controlled trials: why randomise?
British Medical Journal, **318**, 1209.

Anderson, G. B., Bowden, J., Morrison, E. C. and Caffesse, R. G. (1997)
Clinical effects of chlorhexidine mouthwashes on patients undergoing orthodontic treatment,
American Journal of Orthodontics and Dentofacial Orthopedics, **111**, 606–612.

Antczak-Bouckoms, A. (1998)
The anatomy of clinical research,
Clinical Orthodontics and Research, **1**, 75–79.

Ash, S. and Hay, N. (1996)
Adhesive pre-coated brackets, a comparative clinical study,
British Journal of Orthodontics, **24**, 325–329.

Bishara, S. E. (1992)
Impacted maxillary canines: a review,
American Journal of Orthodontics and Dentofacial Orthopedics, **101**, 159–171.

Booth-Mason, S. and Birnie, D. J. (1988)
Penetrating eye injury from orthodontic headgear—a case report,
European Journal of Orthodontics, **10**, 111–114.

Boyne, P. J. and Sands, N. R. (1972)
Secondary bone grafting of residual alveolar and palatal defects,
Journal of Oral and Maxillofacial Surgery, **30**, 87–92.

Buchannan, I. B., Russell, J. I. and Clark, J. D. (1996)
Practical application of the PAR Index: an illustrative comparison of the outcome of treatment using two fixed appliance techniques,
British Journal of Orthodontics, **23**, 351–357.

Critical Appraisal Skills Programme (CASP) (1997)
Making Sense of Evidence About Clinical Effectiveness. CASP Workshop Pack—glossary,
Critical Appraisal Skills Programme (CASP), Oxford.

Crombie, I. K. (1996)
The Pocket Guide to Critical Appraisal,
British Medical Journal Publishing Group, *London.*

Deeks, J. D. and Sheldon, T. A. (1995)
Guidlines for Undertaking Systematic Reviews of Effectiveness, Version 4,
York Centre for Reviews and Dissemination, York.

Ellenberg, S. S. (1981)
Studies to compare treatment regimens: the randomised clinical trial and alternative strategies,
Journal of the American Medical Association, **246**, 2481–2482.

Erverdi, N., Koyuturk, O. and Kucukkeles, N. (1997)
Nickel-titanium coil springs and repelling magnets: a comparison of two different intra-oral molar distalisation techniques,
British Journal of Orthodontics, **24**, 47–53.

Fox, N. A., Richmond, S., Wright, J. L. and Daniels, C. P. (1997)
Factors affecting the outcome of orthodontic treatment within the General Dental Service,
British Journal of Orthodontics, **24**, 217–221.

Gardner, M. J. and Altman, D. G. (1986)
Confidence interval rather than *p* values: estimation rather than hypothesis testing,
British Medical Journal, **292**, 746–750.

Green, S. B. and Byar, D. P. (1984)
Using observational data from registries to compare treatments: the fallacy of omnimetrics,
Statistics in Medicine, **3**, 361–370.

Greenhalgh, T. (1997)
How to Read a Paper: the basis of evidence based medicine,
British Medical Journal Publishing Group, London.

Guyatt, G. H., Sackett, D. L. and Cook, D. C. (1993)
Users' guide to the medical literature: II. How to use an article about therapy or prevention. A. Are the results of the study valid?,
Journal of the American Medical Association, **270**, 2598–2601.

Guyatt, G. H., Sackett, D. L. and Cook, D. C. (1994)
Users' guide to the medical literature: II. How to use an article about therapy or prevention. B. What are the results and will they help me in caring for my patients?,
Journal of the American Medical Association, **271**, 59–63.

Guyatt, G. H., Sackett, D. L., Sinclair, J. C., Hayward, R. and Cook, D. C. (1995)
Users' guide to the medical literature: IX. A method for grading healthcare recommendations,
Journal of the American Medical Association, **274**, 1800–1804.

Hall, G. M. (1994)
How to Write a Paper,
British Medical Journal Publishing Group, London.

Harradine, N. W. T., Pearson, M. H. and Toth, B. (1998)
The effect of extraction of third molars on late lower incisor crowding: a randomized controlled trial,
British Journal of Orthodontics, **25**, 117–122.

Harrison, J. E. (1998)
Early experiences with the Tip-Edge appliance,
British Journal of Orthodontics, **25**, 1–9.

Harrison, J. E. and Ashby D. (1999)
Orthodontic treatments for posterior crossbites (Cochrane Protocol),
In: The Cochrane Library, Issue 4,
Update Software, Oxford.

Harrison, J. E., Ashby, D. and Lennon, M. A. (1996)
An analysis of papers published in the British and European journals of orthodontics,
British Journal of Orthodontics, **23**, 203–209.

Haynes, R. B., Mulrow, C. D., Huith, E. J., Altman, D. G. and Gardner, M. J. (1990)
More informative abstracts revisited
Annals of Internal Medicine, **113**, 69–76.

Hayward, R. S. A., Wilson, M. C., Tunis, S. R., Bass, E. B. and Guyatt, G. H. (1995)
Users' guide to the medical literature: VIII. How to use clinical practice guidelines. A. Are the recommendations valid?,
Journal of the American Medical Association, **274**, 570–574.

Hobson, R. S. and Clark, J. D. (1995)
Management of the orthodontic patient 'at risk' from infective endocarditis,
British Dental Journal, **178**, 289–295.

Holmes, A. (1992)
The prevalence of orthodontic treatment need,
British Journal of Orthodontics, **19**, 177–182.

Jaeschke, R., Guyatt, G. H. and Sackett, D. L. (1994)
Users' guide to the medical literature: III. How to use an article about a diagnostic test: A. Are the results of the study valid?
Journal of the American Medical Association, **271**, 389–391.

Jones, M. (1998)
Editorial—a new look for 1998,
British Journal of Orthodontics, **25**, 55–56.

Kaley, J. and Phillips, C. (1991)
Factors related to root resorption in edgewise practice,
Angle Orthodontist, **61**, 125–132.

Laupacia, A., Wells, G., Richardson, W. S. and Tugwell, P. (1994)
Users' guide to the medical literature: V. How to use and article about prognosis,
Journal of the American Medical Association, **272**, 234–237.

Levine, M., Walter, S., Lee, H., Haines, T., Holbrook, A. and Moyer, V. (1994)
Users' guide to the medical literature: IV. How to use an article about harm,
Journal of the American Medical Association, **271**, 1615–1619.

Milne, R., Donald, A. and Chambers, L. (1995)
Piloting short workshops on critical appraisal of reviews,
Health Trends, **27**, 120–123.

Newcombe, R. G. (1994)
Research in orthodontics—A statistical perspective,
British Journal of Orthodontics, **21**, 299–302.

O'Brien, K. and Craven, R. (1995)
Pitfalls in orthodontic health service research,
British Journal of Orthodontics, **22**, 353–356.

Ömblus, J., Malmgren, O., Pancherez, H., Hagg, U. and Hansen, K. (1997)
Long-term effect of Class II correction in Herbst and Bass therapy,
European Journal of Orthodontics, **19**, 185–193.

Oxman, A. D., Cook, D. C. and Guyatt, G. H. (1994)
Users' guide to the medical literature: VI. How to use an overview,
Journal of American Medical Association, **272**, 1367–1371.

Oxman, A. D., Sackett, D. L. and Guyatt, G. H. (1993)
Users' guide to medical literature: I. How to get started,
Journal of the American Medical Association, **270**, 2093–2095.

Pocock, S. J. (1983)
Clinical Trials: a practical approach,
John Wiley & Sons Ltd. Chichester.

Richardson, W. S. and Detsky, A. S. (1995a)
Users' guide to the medical literature: VII. How to use a clinical decision analysis. A. Are the results of the study valid?,
Journal of the American Medical Association, **273**, 1292–1295.

Richardson, W. S. and Detsky, A. S. (1995b)
Users' guide to the medical literature: VII. How to use a clinical decision analysis. B. What are the results and will they help me in caring for my patients?
Journal of the American Medical Association, **273**, 1610–1613.

Roberts, C. T. and Richmond, S. (1997)
The design and analysis of reliability studies for the use of epidemiological and audit indices in orthodontics,
British Journal of Orthodontics, **24**, 139–147.

Sackett, D. L., Richardson, W. S., Rosenberg, W. and Haynes, R. B. (1997)
Evidence-based Medicine,
Churchill Livingstone, Edinburgh.

Sidhu, M. S., Kharbanda, O. P. and Sidhu, S. S. (1995)
Cephalometric analysis of changes produced by a modified Herbst appliance in the treatment of Class II division 1 malocclusion,
British Journal of Orthodontics, **22**, 1–12.

Stucki, N. and Ingervall, B. (1998)
The use of the Jasper Jumper for the correction of Class II malocclusion in the young permanent dentition,
European Journal of Orthodontics, **20**, 271–281.

Tarvit, D. J. and Freer, T. J. (1998)
Assessing malocclusion—the time factor,
British Journal of Orthodontics, **25**, 31–34.

Taylor, P. J. S., Kerr, W. J. S. and McColl, J. H. (1996)
Factors associated with the standard and duration of orthodontic treatment,
British Journal of Orthodontics, **23**, 335–341.

Watson, W. G. (1980)
Anatomy of a manuscript,
American Journal of Orthodontics and Dentofacial Orthopedics, **77**, 106–108.

Wilson, M. C., Hayward, R. S. A., Tunis, S. R., Bass, E. B. and Guyatt, G. H. (1995)
Users' guide to the medical literature: VIII. How to use clinical practice guidelines. B. What are the recommendations and will they help you in caring for your patients?
Journal of the American Medical Association, **274**, 1630–1633.

Appendix 1: checklist for use with surveys

ABSTRACT	Yes	No	Un-clear	Comments
Are the objectives of the survey stated?				
Is the design of the survey stated?				
Is the setting of the survey stated?				
Is the source of the subjects surveyed stated?				
Is the sample size stated?				
Is the response rate stated?				
Are the outcomes of interest stated?				
Are any results given?				
Are any conclusions stated?				
Is the abstract informative?				
PRELIMINARY QUESTIONS				
Why was the survey done?				
What question were the researchers asking?				
What type of survey was done?				
Was this research method appropriate?				
SPECIFIC QUESTIONS				
Who was studied?				
How was the sample obtained?				
What was the response rate?				

STANDARD QUESTIONS	Yes	No	Un-clear	Comments
Are the aims clearly stated?				
Where was the survey carried out?				
How many participants were surveyed?				
Was the sample size justified?				
Was the survey adequately controlled?				
Were the sub-groups described?				
Was the follow-up of non-responders adequate?				
Are the outcome measures clinically relevant?				
Were measurements likely to be valid and reliable?				
Are the statistical methods described?				
Were the basic data adequately described?				
Do the numbers add up?				
Was the statistical significance assessed?				
Were confidence intervals given?				
Could the findings have occurred by chance?				
Did any untoward events occur during the survey?				
Could selection bias have occurred?				
INTERPRETATION				
What are the main findings?				
Does the data support the results?				
How are null findings interpreted?				
How do the findings relate to previous work?				
Who are the results of this survey applicable to?				
How will the findings affect my clinical practice?				

Resorbable Implants (Plates and Screws) in Orthognathic Surgery

I. K. MOHAMED-HASHEM, B.D.S., F.D.S.R.C.S. (ENG.) D. A. MITCHELL, M.B.B.S., B.D.S., F.D.S.R.C.P.S., F.R.C.S (ED.), F.R.C.S. (MAXFAC)

Oral & Maxillofacial Surgery Department, Pindergields General Hospital, U.K.

Introduction

The aim of internal fixation of traumatic and iatrogenic skeletal fracture is to achieve undisturbed fracture healing. The need for plates and screws for fixation is only temporary, until the fracture has united. Accordingly surgeons including those of AO-ASIF school recommend that all metallic implants used for fixation of fractures be removed in due course (Müller *et al*., 1979). Reasons for removal of the implants include the possibility of bone atrophy due to stress shielding by rigid bone plates and screws (Pavolainen *et al*., 1978). Other disadvantages re hypothesized carcinogenic potential, the possibility of corrosion, disturbance in normal growth pattern, and implant migration in children (Simon *et al*., 1978). Internationally, the removal of the metallic hardware varies from routine removal from all patients in some countries to selective removal only from patients who have symptoms as is usual in the UK (Chapman and Woo, 1988). The use of biologically inert resorbable implants would eliminate the need for a second operation for their removal, and offers major clinical advantages for the fixation of facial bone fractures in trauma and orthognathic surgery. They would be enormously advantageous in paediatric craniofacial surgery. Clinical studies have shown that absorbable implants have been used successfully as a rigid fixation device in mandibular osteotomy and craniofacial surgery and that normal growth pattern is probably not disturbed by use of these implants (Simon *et al*., 1978; Suurohen *et al*., 1992). The resorbable plates and screws available from one company have been widely used in our combined trauma service (Leeds/Wakefield).

History of Absorbable Implants

The use of absorbable implants in the repair of bone fractures began in the late 1960s. Fabrication of implants was accomplished by melt moulding and extrusion of polymer into pins and rods. Subsequently more complex designs such as screws and small plates became possible in the late 1970s and early 1980s (Böstman, 1991).

Chemical composition of Absorbable Implants

Alpha compounds such as polyglycolic acid, polylactic acid, and polyesters polyparadioxanon are organic macromolecular compounds that are degradable and absorbable by the body. They also possess the chemical and physical properties necessary for internal fixation devices (Böstman, 1991).

Experimental Studies

Several investigations have shown that these polymers are completely absorbable within bony tissue and that new bone is deposited on and within the implants as degradation proceeds. The degradation procedure appears to be mainly by hydrolytic activity and to a lesser extent through non-specific enzymatic action. The rate of degradation is dependant on the molecular weight, crystallinity, thermal history, and geometry of the implant, as a porous thin sheet depolymerizes much more rapidly than a dense block.

The degradation process in itself does not imply immediate absorption of an implant, as experiments show that 70 per cent of the material from the implant remains *in situ* for 3 months. Studies also show that the principal route of ultimate elimination is respiration with excretion in the urine and faeces playing only a minor role (Böstman, 1991).

The main clinical complication reported associated with the use of polylactic acid and polyglycolic acid implants is the development of inflammatory foreign body reaction. Clinical reports show that these problems may be due to delayed resorption rate of the polymer (Brady *et al*., 1973; Böstman, 1991).

Products Available

The commercially available resorbable polymers include pure polyglycolic (PGA) acid in the form of pins and screws. pure poly-L-lactic acid (PLLA) and a co-polymer of PLLA and PGA. The last gathered the best physical and chemical properties of both PLLA and PGA, and experimental studies have shown that the fixation devices made from this copolymer maintain most of their strength for 8 weeks and will completely resorb in the body in 12–15 months, with no complications reported in their usage to date (Investigational Products in the United States, 1995).

In the Oral and Maxillofacial Department at Pinderfields General Hospital we currently use Lactosorb resorbable craniomaxillofacial fixation system in facial trauma and are in process of using it in orthognathic surgery. Lactosorb is a patented copolymer of PLLA (82 per cent) and PGA (18 per cent), and offers a good balance between initial strength and resorption rate.

Features of this system include:

(1) tensile and flex strength are comparable to titanium plating system (Bergsma *et al*., 1993);
(2) plates are easy to adapt with aid of heat pack;
(3) a wide selection of implant sizes and shapes are available (Figures 1 and 2);
(4) a convenient hex-drive breakaway delivery system simplifies screw placement;
(5) eliminates growth restriction and implant migration for paediatric craniofacial reconstruction;
(6) resorbs completely and may eliminate the need for second operation;
(7) does not induce late stage inflammatory reaction.

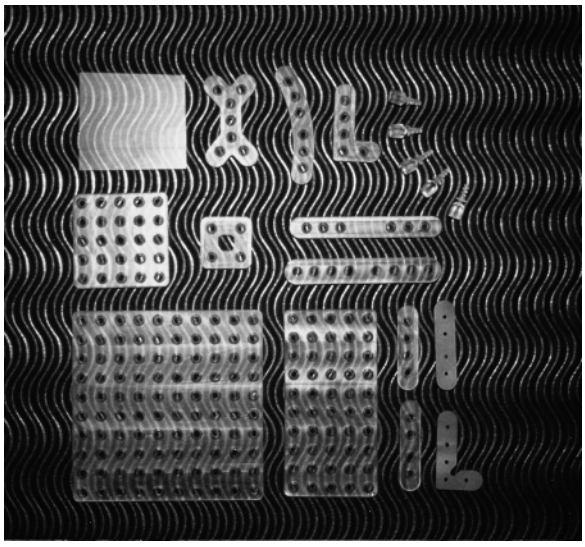

FIG. 1 Shows different sizes and shapes of plates and screws.

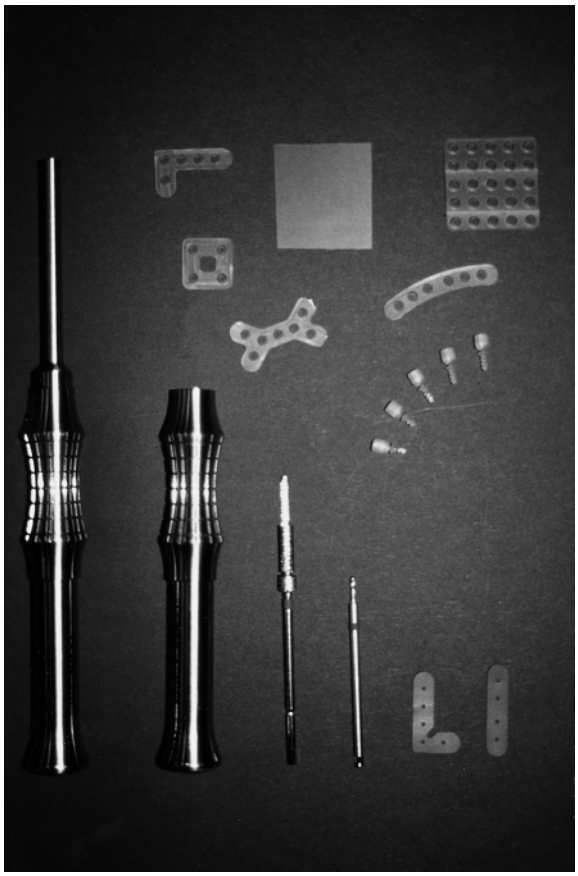

FIG. 2 Shows the hex-drive breakaway delivery system that simplifies screw placement.

Current Costs

Unit price for 1·5-mm plate in lactosorb system ranges from £52·50 to £94·50 depending on the shape and number of holes compared with the conventional metal A-O system, which ranges from £34·85 to £61·50. A 1·5-mm screw ranges in price from £27·30 to £30·75 for lactosorb system depending on length compared with A-O 1·5-mm screws at £9·40.

Unit price for a 2·00 mm. plate in lactosorb system ranges from £52·50 to £94·05 compared with the A-O system at £35·50 to £49·05 and for the matching screw prices range from £23·65 to £34·15 for lactosorb compared with the A-O system at £10·75.

Lactosorb resorbable craniomaxillofacial fixation system is provided by Poly-Medics and supplied by Walter Lorenz Surgical, Ins. The U.K. representative is Athrodax health-care international Ltd, Great Western Court, Ross-on-Wye, Herefordshire HR9 7XP, U.K.

References

Bergsma, E. J., Rozema, F. R., Bos, R. R. and deBruign, E. C. (1993)
Foreign body reactions to resorbable poly (L-lactide) bone plates and screws used for fixation of unstable zygomatic fractures,
Journal of Oral and Maxillofacial Surgery, **51**, 666–670.

Böstman, O. M. (1991)
Current concepts review, absorbable implants for the fixation of fractures,
Journal of Bone and Joint Surgery, **73-A**, 148–155.

Brady, J. M., Cutright, D. E., Miller, R. A. and Battistone, G. C. (1973)
Resorption rate, route of elimination and ultra structure of implant site of polylactic acid in the abdominal wall of the rat,
Journal of Biomedical Materials Research, **7**, 155–166.

Chapman, M. W. and Woo, S. L. Y. (1988)
In: principles of Fracture Healing in Operative Orthopaedics,
M. W. Chapman and Michael Madison (Eds)
J. B. Lippincott, Lippincott, pp. 115–123.

Investigational Products in the United States: (1995)
Compay data Polyamides Inc,
Warsaw, Indiana 46580
Biomet, Inc form No. Y-INT-022/051525/HP.

Müller, M. E., Allgöwer, M., Schneider, R. and Willenegger, H. (1979)
Manual of internal fixation. Techniques recommended by the AO Group, 2nd edn,
Springer, New York.

Pavolainen, P., Karajarju, E., Slatis, P., et al. (1978)
Effect of rigid plate fixation on structure and mineral content of cortical bone.
Clinical Orthopedics and Related Research, **136**, 287.

Simon, B. R., Woo, S. L-Y., McChartly, M., et al. (1978)
Parametric study of bone remodeling beneath internal fixation plates of varying stiffness,
Journal of Bioengineering,. **2**, 543.

Suurohen, R., Laine, P., Sarkiala, G., et al. (1992)
Sagittal split osteotomy fixed with biodegradable, self reinforced poly-L-lactide screws,
International Journal of Oral and Maxillofacial Surgery, **21**, 303–308.