

Commentary

Why does clustering matter in orthodontic trials?

Jayne E. Harrison* and Girvan Burnside**

*Department of Orthodontics, Liverpool University Dental Hospital, Liverpool and **Department of Biostatistics and School of Dentistry, University of Liverpool, UK

Correspondence to: Jayne E. Harrison, Department of Orthodontics, Liverpool University Dental Hospital, Pembroke Place, Liverpool L3 5PS, UK. E-mail: Jayne.Harrison@rlbuht.nhs.uk

SUMMARY Clustering in RCTs occurs when participants or units are allocated to an intervention in a group rather than independently or when multiple measurements are taken from the same individual. Cluster RCTs occur frequently in clinical orthodontic research; however, only a quarter of published trials take account of the effects of clustering in the design and analysis of these trials. The effects of clustering needs to be considered when calculating the sample size required to detect a difference in treatment effect, obtaining consent for participation in the trial and finally the analysis of the data.

Introduction

The paper by [Koletsis *et al.* \(2011\)](#) highlights the frequent occurrence and problems associated with the design and analysis of cluster randomized controlled trials (RCTs) in orthodontics. A cluster RCT is one where research participants or units are not allocated to an intervention independently but in a group ([Bland, 2004](#)). The units may, for example, be children in a class, patients at a practice, or teeth in a mouth of an individual patient as is common in orthodontic trials. Teeth in a mouth may be allocated individually or in groups within the mouth, e.g. quadrants as occurs in split mouth trials. Clustering also occurs when multiple measurements are taken from the same individual, e.g. plaque scores over a course of treatment; multiple bond failures in the same patient, or growth measurements over time. In accordance with CONSolidated Standards Of Reporting Trials statement extension to cluster RCTs ([Campbell *et al.*, 2004](#)) reports of cluster RCTs should include: the following information:

- the rationale for adopting a cluster design,
- how the effects of clustering were incorporated into the sample size calculations,
- how the effects of clustering were incorporated into the analysis, and
- the flow of both clusters and individuals through the trial, from assignment to analysis.

There are also issues surrounding consent in cluster RCTs that need to be addressed and reported appropriately.

The effect of clustering has to be taken into account in the design, conduct, analysis, and interpretation of the trials to allow for the likelihood that, for example, children within the same class or school or teeth within a mouth will respond

similarly because they are exposed to a similar environment. This means that data from each child or tooth cannot be assumed to be independent of each other. As identified by [Koletsis *et al.* \(2011\)](#), many cluster trials in orthodontics are incorrectly designed or analysed and treat data as though the unit of allocation had been the individual tooth rather than the patient. This can lead to spurious statistical significance due to the overinflated sample sizes or problems of interpretation due to confusion of the results from the unit compared with those from the group. In statistical terms, the patient is the sampling unit (or unit of investigation) and should therefore be the unit of analysis ([Altman and Bland, 1997](#)). However, in some situations, e.g. development of caries during orthodontic treatment, useful information about which teeth and which sites on individual teeth undergo demineralization, is lost if we only look at the number of patients affected.

Rationale for cluster RCTs

When designing a cluster RCT, there must be a justifiable rationale for adopting the design. Reasons for designing a trial as a cluster RCT include:

- concerns about contamination of the randomized groups,
- where the intervention is designed to be delivered to a group,
- situations where multiple body parts are being assessed in an individual or
- multiple measurements from each individual are being made over time, and
- logistic or administrative problems in delivering the intervention to an individual.

Conversely, trials of interventions that contaminate other units within a cluster should not be designed as a cluster RCT and are better designed individual patient trials, e.g. fluoride leaching cements to prevent demineralization around orthodontic brackets. In addition, for ethical reasons and under the principles of the Declaration of Helsinki, the decision to undertake a cluster RCT should not be taken lightly because they do need more patients in order to gain adequate power and it is unethical to expose people unnecessarily to the risks of research (World Medical Association, 1997).

Sample size considerations for cluster RCTs

A key consideration in the design of a cluster RCT is the sample size calculation because it must take into account the fact that the individuals or units within a group or cluster are not independent. This means that by randomizing by group or cluster rather than the individual or unit, there will be some loss of power (Kerry and Bland, 1998) and a cluster RCT will always require a larger sample size than that of a comparative non-clustered trial (Christie *et al.*, 2009). The ratio between the numbers of individual participants needed when using a cluster design and that needed when using a conventional design is called the design effect and increases with the size of the cluster.

When undertaking a sample size calculation for a conventional RCT with a continuous outcome, the expected difference in the means of the outcome between the two groups and the variance in outcome is used to calculate how many participants will be required to give the trial adequate power to detect a difference in treatment effect if there is one. With a cluster RCT, however, you have to take into account the variance in the outcome within each cluster and also between clusters. This is called the intracluster correlation coefficient (ICC) which is defined as the proportion of the total variation which can be attributed to the variation between clusters. The value of the ICC can range from 0 to 1. An ICC of 0 would mean that all observations within a cluster were independent, i.e. there is no cluster effect. An ICC of 1 would arise, when all observations within a cluster are identical, i.e. there is no variation within clusters (Burnside *et al.*, 2006). The ICC therefore describes the extent to which two members of one cluster are more similar than two members from different clusters. It can be used to calculate the effective sample size, which is defined as the number of participants in an individually randomized trial which would give the same power as the cluster randomized trial.

The sample size is also related to the number of units per cluster because as the number of units per cluster increases so does the number of total number of units required. This means that, in two trials with the same sample size, the trial with a larger number of clusters and fewer individuals in each cluster will be able to distinguish between the treatment

effect of two interventions better than a trial with fewer clusters and but larger numbers of individuals in each cluster (Christie *et al.*, 2009). However, if the number in each cluster is kept small, then the number of clusters required is large which may have logistical problems in terms of recruitment of the clusters.

Data analysis considerations for cluster RCTs

Bland and Kerry (1997) say that 'there is a price to be paid for this design at the analysis stage'. This is because in a cluster RCT, the size of the standard errors increases thus widening the confidence intervals and increasing the *P* values compared to a conventional trial of the same size there by reducing the power as the effective sample size is reduced.

When analysing the data, account must be taken of the units of analysis and multiplicity (Altman and Bland, 1997). In orthodontic terms, the unit can be the tooth, the different surfaces of the tooth, the quadrant, or the mouth that, in turn, may contribute multiple data or a single summary measure for each patient. Whichever unit is chosen, account must be taken that the data are not independent because they are derived from the same patient. If each tooth was taken as an independent unit, then the sample size would be artificially inflated and a false-positive result would be obtained. However, if data are summarized, tooth- and surface-specific information is lost. These data could provide important clinical information as interventions may be more effective on particular teeth or surfaces within the mouth, depending on the intervention. Therefore, if the data analysis is to be performed at tooth or surface level, the clustering within participants should be accounted for to ensure accurate conclusions (Burnside *et al.*, 2006). This can be done, without losing the information at the individual level, using various methods, such as multilevel modelling.

When considering multiplicity in the orthodontic environment, we must also take into account that multiple measurements, taken from the same patient over time, are related and cannot be considered independently. The simplest method of accounting for clustering within a mouth or quadrant is to take a summary statistic for each cluster and then analyse these summary values (Kerry and Bland, 1998). Alternately, more complex approaches such as repeated measures analysis of variance or multilevel modelling may be used to compare the means. The impact of not taking into account clustering when analysing data were demonstrated in a Cochrane review of adhesives for fixed orthodontic braces (Mandall *et al.*, 2003) where 10 of the 22 excluded studies were excluded because they analysed the number of bond failures by tooth rather than on a patient basis or included multiple failures per tooth.

Ethical and consent considerations for cluster RCTs

As mentioned previously, cluster RCTs do give rise to ethical concerns due the involvement of more participants in a trial

than a non-clustered. The other ethical issue is that of obtaining consent in cluster trials where the level of randomization is, for example, a school or a practice rather than an individual. If consent is obtained at the cluster level, then the cluster guardian, e.g. a head teacher or practice principal, may be signing consent for his/her cluster to participate rather than consent being obtained individual participants within the cluster (Taljaard *et al.*, 2011). If consent is also obtained from participants within the cluster, there may be differential refusal to participate within the clusters that can give rise to consent bias. However, as most cluster trials in orthodontics are clusters of teeth within an individual participant, this is rarely an issue for orthodontic researchers.

Conclusions

In conclusion, we would like to advise that a statistician is involved in designing and analysing the data from a cluster RCT to ensure that all the design and statistical issues arising from them are addressed. It is disappointing that Koletsi *et al.* (2011) found that only a quarter of trials, where clustering was evident, had taken into account the clustering effects during statistical analyses meaning that the presented results may not have been valid. It is to be hoped that by raising awareness of this common problem that the design, analysis, and reporting of orthodontic trials where there is clustering will improve.

References

- Altman D G, Bland J M 1997 Statistics notes. Units of analysis. *British Medical Journal* 314: 1874
- Bland J M 2004 Cluster randomised trials in the medical literature: two bibliometric surveys. *BioMedCentral Medical Research Methodology* 4: 21
- Bland J M, Kerry S M 1997 Statistics notes. Trials randomised in clusters. *British Medical Journal* 315: 600
- Burnside G, Pine C M, Williamson P R 2006 Statistical aspects of design and analysis of clinical trials for the prevention of caries. *Caries Research* 40: 360–365
- Campbell M K, Elbourne D R, Altman D G; CONSORT group 2004 CONSORT statement: extension to cluster randomised trials. *British Medical Journal* 328: 702–708
- Christie J, O'Halloran P, Stevenson M 2009 Planning a cluster randomized controlled trial: methodological issues. *Nursing Research* 58: 128–134
- Kerry S M, Bland J M 1998 Sample size in cluster randomisation. *British Medical Journal* 316: 549
- Koletsi D, Pandis N, Polychronopoulou A, Eliades T 2011 Does published orthodontic research account for clustering effects during statistical data analysis? *European Journal of Orthodontics* [Epub ahead of print]
- Mandall N A, Millett D T, Mattick C R, Hickman J, Macfarlane T V, Worthington H V 2003 Adhesives for fixed orthodontic brackets. *Cochrane Database Systematic Reviews* (2): CD002282
- Taljaard M *et al.* 2011 Inadequate reporting of research ethics review and informed consent in cluster randomised trials: review of random sample of published trials. *British Medical Journal* 342: d2496. doi: 10.1136/bmj.d2496
- World Medical Association declaration of Helsinki. 1997 Recommendations guiding physicians in biomedical research involving human subjects. *Journal of the American Medical Association* 277: 925–926