

# 1

## Do we need to improve oral health research?

Dominique Declerck and Emmanuel Lesaffre

### 1.1 Introduction

The goal of research is to expand our knowledge. Based on the insights obtained, decisions and choices can be made in order to organize our society. The ultimate aim of oral health (OH) research, therefore, must be to accomplish this at the level of oral health and related aspects. Ideally, prevention of disease development, treatment of existing disease and the organization of care delivery should be based on high-level evidence obtained from high-quality research. Statistical and methodological issues determining research quality in the field of OH research are covered in this book.

Research involves different stages, it starts with the planning of the study and it ends with the dissemination of the conclusions. In each of the stages, there is the risk of taking wrong decisions thereby blurring the outcome of the research. In this chapter we elaborate on what might go wrong in the different stages of OH research. We hope that this chapter is an appetizer for the subsequent chapters where most methodological aspects of oral health research will be explained and critically examined. Explicit references to chapters will be made which hopefully will guide the reader throughout the book.

## 1.2 Is there a problem?

Several publications can be found reporting on quality issues in different fields of OH research, even in the recent literature. Statistical aspects of design and analysis of randomized controlled clinical trials (RCTs) for the prevention of caries were considered in the paper by Burnside *et al.* (2006). These authors concluded that in recent RCTs of topical fluoride interventions the design of the study was not taken into account. Indeed, while a cluster-randomized design was chosen, the statistical analysis afterwards (see Chapters 6 and 13) most often ignored the clustering of subjects. In addition, essential information was often lacking in the reports not allowing judgement of presence of e.g. possible consent bias. Robinson *et al.* (2006) discussed the quality of reports of RCTs comparing manual and powered toothbrushes. Several important shortcomings were identified: inadequate generation of the randomization sequence, inadequate concealment of the allocation of treatments, inappropriate use of the split-mouth design, etc. A similar finding was reported by Lesaffre *et al.* (2007) on split-mouth trials. Split-mouth studies are popular e.g. in the area of dental materials research. In Lesaffre *et al.* (2007) a review was made of the appropriateness of the statistical approach and on the quality of reporting. Surprisingly, many of the studies did not motivate the use of the split-mouth design and in many papers no full advantage was taken of its design in the statistical analysis. Further, the majority of the studies had a major flaw in their statistical analysis and/or in the reporting of the results. In periodontal research, Tu and co-workers (2006a) evaluated the quality of trials that compared guided tissue regeneration with use of enamel matrix derivatives. Most trials did not meet the majority of their design criteria, so that they placed doubt on the value of and the conclusions from these trials. Eckert and co-workers (2005) presented a literature review on the quality of evidence comparing the clinical performance of dental implant systems. They concluded that the evidence supporting implant therapy was generally derived from case series rather than cohort studies or RCTs. Reports on a direct comparison of different implant systems were not available. In orthodontic research, statistical problems are also common as indicated by Tu *et al.* (2006b).

There are not only problems with the design and the statistics, the problems also exist at the level of taking the measurements. Robinson *et al.* (2006) indicated that plaque data were reported using ten different indices and gingivitis with nine indices. In a recent review of methodological aspects of caries experience screening in epidemiological surveys, Agbaje *et al.* (2009) reported that detailed information was lacking in a considerable number of reports. Moreover, when the use of standardized survey methodology was mentioned (e.g. WHO guidelines), deviations from the original recommendations were often present hampering comparisons between surveys.

### 1.3 Is oral health research unique?

The problems in quality are not limited to OH research but are commonly encountered in many other fields of medical research. For instance, in a review article by Sjögren and Halling (2002), the quality of RCTs was evaluated in different areas of dental and medical research. Again, the authors concluded that there was a clear need to improve the quality of trial reporting.

Since the methodological issues are not limited to OH research, why should we especially be bothered about the methodology in oral health studies? There are several reasons indeed why we should have a special focus on OH research. The first reason is an obvious but important one, i.e. oral health data are often complex in nature. For instance, in caries research dentists are interested in lesions at tooth surfaces. There are more than 100 tooth surfaces on permanent teeth but, unlike data that are obtained on subject level (stroke, cardiac mortality, weight, etc.), the information that one tooth surface brings to us is not independent of the information from other tooth surfaces in the same mouth. We say that the information from tooth surfaces is 'not independent' or also 'is correlated'. The statistical analysis of simple research questions but involving correlated data is challenging and needs special care and statistical expertise, much more so than most medical research.

Despite the great need, the training in basic methodological and statistical skills at European (and probably also elsewhere) medical faculties is in general rather poor. The situation in dental training is certainly no exception on this trend, on the contrary. The need of methodological expertise is great in OH research.

Another problem is that OH research is somewhat isolated from medical research. Consequently certain trends that pop up in medical research appear much later in OH research. For instance, the CONSORT guidelines (see Chapter 3) were adopted in three medical journals (JAMA, BMJ, Lancet) in 1996 but it was only in 1999 that a dental journal (*British Dental Journal*) adopted the CONSORT policy of reporting. Other examples of a different attitude in medical and oral health research are given in Chapter 2 (cf. surrogate markers).

As medical research, OH research has become increasingly complex with need for multidisciplinary. The need for technical assistance is obvious and recognized, e.g. when dedicated computer hard- and/or software needs to be developed. Also, when OH research involves new technical innovations, engineers are a necessity. But the need for collaborative efforts with methodologists/statisticians is not sufficiently appreciated. Partly the blame for this attitude is the poor refereeing process on methodology of many dental journals, not forcing the oral researchers to collaborate.

The message is not 'medical research is great and oral health research is poor', but there is no doubt a discrepancy. In the next sections, we review the risks of taking the wrong decisions in performing OH research. We distinguish three stages: (1) planning of the study, (2) conduct of the study and (3) the analysis stage of the

study: statistical analysis of the data, reporting of the results, interpretation of the results and conclusions.

## **1.4 What to do to improve the quality of oral health research?**

### **1.4.1 Planning stage**

An example of poor research is a retrospective study (1) where data is collected from medical records by students in their free available time; (2) without a clear view of what will be investigated; (3) without any idea of how much time and resources are needed to perform the research and (4) what expertise is needed to yield high quality output. All of the necessary steps for performing high quality (OH) research are described in Chapter 4. By means of 20 questions the authors guide the researchers to take the correct steps in planning their research. The authors' emphasis is on setting up OH RCTs, but much of their advice also applies to epidemiological studies.

When considering a RCT or an epidemiological study, the reader is referred to Chapter 6 where the characteristics of a RCT are laid down and to Chapter 7 where the focus is on epidemiological research. The reader is also advised to consult Chapter 2 where reflections are given on what information the different types of research are bringing us.

As explained in Chapter 4, researchers need to reflect on (1) the research question and the associated primary and secondary outcomes, (2) which hypotheses to test, (3) the study design, (4) how to recruit study subjects and how many individuals are needed, (5) ethical aspects, (6) financial aspects, (7) which and how much personnel is needed, etc.

In Chapter 8, the principles of qualitative research are explained. It appears that this kind of research has a lot to offer to the researcher in helping to ask the right kind of questions to the patients, in other words to collect the relevant information for the research questions at hand.

The involvement of statistics/statisticians starts at the planning phase when deciding for the primary endpoint, the choice of statistical analyses, the necessary sample size, etc. Statistical principles are also involved during the conduct of the study, e.g. when an interim analysis of the data is envisaged, but in general statistics are more manifest when planning the study, analyzing the data and interpreting the results. In Chapter 6, a review is given of the different statistical aspects involved in RCTs.

### **1.4.2 Conduct of the study**

In Chapter 5 the authors reflect on building up a multidisciplinary team and managing it to yield a successful outcome. In this respect, the authors elaborate on the basic team management principles. For instance, they reflect on the necessary

qualities of the principal investigator, the co-investigators, the study coordinator, the clinical research monitor, etc. But also aspects of recruitment of patients, how to practically organize randomized allocation and blinding are treated. As in Chapter 4 the authors focus primarily on RCTs, but the chapter is also useful to epidemiological studies.

While data management is becoming more and more computerized, there are still a lot of quality issues that need to be considered in this context. The level of sophistication differs from study to study depending on factors such as the type of study, the financial resources, etc. For example, electronic data entry is becoming a standard in the conduct of RCTs, while in hospital-based studies or small-scale epidemiological surveys data are often recorded on paper and later transferred to the electronic database. Both systems require special attention to guard the quality of the data. Issues that relate to data management problems are also treated in Chapter 5.

A careful planning of the study involves the reflection on which measurements to take. Reasons for not choosing a particular measurement on board are: unclear relationship to the primary questions, (perceived) lack of quality and validity of the measurement process, high likelihood to suffer from missing data, etc. All of these aspects are treated in detail in Chapter 9.

### **1.4.3 Analysis stage of the study**

#### **1.4.3.1 Statistical analysis of the data**

Upon collection of the data the exciting part of the research can start, i.e. the statistical analysis of the data leading eventually to the answers to the research questions. However, without a basic understanding of statistics it is impossible to perform any data analysis. Chapter 10 introduces the basic statistical concepts and represents the absolute minimum knowledge that an OH researcher should have for (1) either to do some basic analyses or (2) more importantly, to be prepared for the confrontation with the statistician. This chapter describes a variety of univariate statistical techniques, i.e. methods that analyse each measurement separately. In epidemiological studies, though, it is imperative to correct for imbalances of the risk groups at baseline, leading automatically to regression models. The mother of all regression models is the (multiple) linear regression model where the response is a continuous measurement. Statistical analyses that pertain to this approach are described in Chapter 11 as well as the related correlation approaches. The most popular epidemiological regression models, the logistic regression model and the Cox proportional hazards (survival) model, are also introduced in this chapter.

The statistical implications of the hierarchical structure of oral health data are explained in Chapter 13. Dedicated statistical approaches that deal with such correlated data in an appropriate manner are discussed. These methods were not especially developed for the analysis of oral health data, but they are especially useful for these data. Chapter 15 describes approaches in survival analyses that take into account the correlated nature of the data. The authors also elaborate on

many issues in survival analysis that are especially important in the analysis of survival data.

Coarsened data appear everywhere in medical research, as well as – and perhaps even more so – in OH research. By coarsened data we mean missing data and data that are measured roughly such as the timing of caries development (only known to occur in a time interval; called interval-censored data). The impact and treatment of missing data are treated in Chapter 14. The second type of data are treated in Chapter 15. Less dramatic than missing data, but still problematic, are data that are measured with error. When a (binary, categorical) disease state, e.g. caries experience (Y/N), or a (binary, categorical) risk factor like taking sweets in between meals (Y/N) is recorded with error one speaks of misclassification. Chapter 16 explores what the harm is of measurement error/misclassification and looks for ways to deal with it. Chapter 12 deals with methods that measure the agreement between scoring methods and with methods that measure the predictive ability of imperfect measures vis-à-vis a benchmark examiner or gold standard.

Let it be clear that the application of sophisticated statistical techniques can not correct for a badly planned study or for a poor data collection. This is too often misunderstood by OH researchers but also by medical researchers. For instance, no sophisticated statistical technique can repair (completely) the damage that is done by lacking data (problem of missing data). Oral health data have inherently a complex nature and therefore one cannot always hope for simple analysis methods even to solve simple research questions. Thus the OH researcher is bound to collaborate with statisticians.

We end this section by pointing out that two chapters – Chapter 17 on the analysis of genetic studies and Chapter 18 on the Bayesian methodology – are included to introduce the reader in these popular and rapidly evolving areas.

### **1.4.3.2 Reporting of the results**

Results need to be reported and summarized such that an accurate picture emerges of what data have been collected, what problems were encountered and what statistical methods have been applied. The CONSORT (or similar) guidelines introduced in Chapter 3 help the researcher to write down in detail his/her results.

### **1.4.3.3 Interpretation of the results and conclusions**

The final steps of the research are the interpretation of the statistical analyses and the final conclusions. With a detailed protocol and statistical analysis plan, the freedom in interpreting the results is limited but only for the primary (and secondary) endpoint(s). But such limitations do not hold for all other information that has been collected, or for epidemiological studies with their relatively exploratory character. Moreover, in practice the researchers too often have a biased view upon the results. This is not surprising, since they initiated the research with certain beliefs and hopes. Also, the output of statistical analyses is often only barely understood by clinicians so that opportunities are created to interpret the results in a convenient manner. Examples of such a biased attitude are the phrase ‘there is a trend in

the data' when the P-value is close to 0.05 but still higher, or statements like 'the result would have been definitely significant if the sample size had been greater'. Chapter 2 is in this respect quite useful to obtain an idea of the dangers of making biased conclusions.

Conclusions of the research are typically found in the abstract of the paper and in the Conclusions Section. All of us lack time to read properly our scientific literature. We usually browse through journal issues thereby reading at most the abstract and the conclusions. Also in a further reading of the paper many skip the methodological part of the paper because they lack the technical knowledge to have a good understanding of strengths and weaknesses of the methods applied. It is therefore vital that the conclusions are formulated in a correct manner thereby using the right wording with enough nuances. It is often seen that results are discussed in an appropriate manner in the Results section, but conclusions are too far fetched not supported by the data. Again the CONSORT guidelines can help us in this respect.

## 1.5 How to assess quality in research?

The assessment of research quality is challenging involving the readers of the paper, but also the referees and the editor of the journal. Especially the referees and the editor have a major responsibility in improving the quality of research. But how can a reader assess the quality of a paper? And what other means do referees and the editor have to judge the quality of reported research?

**Indirect assessment of quality** A popular way of evaluating the quality of research, especially by the readers, is considering the journal in which the report was published. The Science Citation Index (SCI) was originally introduced in 1961 by the Institute of Scientific Information (ISI). The impact factor of a journal measures the frequency with which the 'average article' in a journal has been cited in a particular year. In this way the overall quality and significance of a given journal's contents are reflected. Journal ranking is increasingly used as a quantitative tool for evaluating journal quality. However, the journal impact factor should only be used as a rough indicator of scientific quality of an individual article (Andersen, 2006). Also, impact factors are greatly influenced by the type of articles (review versus original research), subject speciality and novelty of the research field.

**Quality in reporting research** Probably the best (we do not claim the perfect, though) tool is the set of guidelines formulated in the CONSORT statement. Several studies reported on the improvement of the quality of reporting in journals where the CONSORT statement was adopted (Moher *et al.*, 2001; Kane *et al.*, 2007). Mills *et al.* (2005) showed that some recommendations were frequently reported, but reporting of others remained suboptimal. However, it should be kept in mind that possible discrepancies between published reports and actual conduct of randomized clinical trials may still exist (Hill *et al.*, 2002). In addition, the quality of

reporting does not guarantee research integrity and cannot protect against scientific misconduct. An example of this is the publication by Sudbo *et al.* (2005) in *The Lancet*, where he stated that the long-term use of non-steroidal anti-inflammatory drugs was associated with a reduced incidence of oral cancer but also with an increased risk of death due to cardiovascular disease. A few weeks after publication of the report some concerns were expressed about this study (Horton, 2006a). It soon became evident that the data had been manipulated: fictitious patients were included in the study. The publication was retracted shortly afterwards and an investigation was started to find out whether other reports by the same author also showed evidence of fraud (Horton, 2006b).

In other research areas, similar initiatives have followed, e.g. for the reporting of meta-analyses of observational studies in epidemiology (MOOSE recommendations – Stroup *et al.*, 2000), meta-analyses of randomized trials (the QUORUM statement – Moher *et al.*, 1999), diagnostic studies (the STARD initiative – Bossuyt *et al.*, 2003) or observational studies in epidemiology (the STROBE statement – von Elm *et al.*, 2007).

## 1.6 Which actions to take

There are a lot of research groups that are doing excellent research based on solid methods. But, there is also a lot of variability in the quality of medical and OH research. We do not expect this to change overnight. If things improve they can only do so gradually whereby first of all the problem should be recognized. Recognition as well as the intention to improve is a must, then comes educational efforts in methodology but also the collaboration with methodologists, not only statisticians. Current research is multidisciplinary, the earlier one realizes this in life the better the research output will be. This implies that efforts should be undertaken to organize joint meetings with oral health researchers and statisticians/methodologists to eventually create networks which will promote collaborative activities and increase the quality of oral health research.

## References

- Andersen, J., Belmont, J., & Cho, C.T. (2006) Journal impact factor in the era of expanding literature. *J Microbiol Immunol Infect* **39**: 436–43.
- Agbaje, O., Lesaffre, E., & Declerck, D. (2009) Assessment of caries experience in epidemiological surveys. *Community Dental Health* (submitted).
- Burnside, G, Pine, C.M., & Williamson, P.R. (2006) Statistical aspects of design and analysis of clinical trials for the prevention of caries. *Caries Res* **40**: 360–5.
- Bossuyt, P.M., Reitsma, J.B., & Bruns, D.E., *et al.* (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* **138**: 40–4.
- Eckert, S.E., Choi, Y.G., Sanchez, A.R., & Koka, S. (2005) Comparison of dental implant systems: quality of clinical evidence and prediction of 5-year survival. *Int J Oral Maxillofac Implants* **20**(3): 406–15.

- Hill, C.L., La Valley, M.P., & Pelson, D.T. (2002) Discrepancy between published report and actual conduct of randomized clinical trials. *J Clin Epidemiol* **55**: 783–6.
- Horton, R. (2006a) Expression of concern: non-steroidal anti-inflammatory drugs and the risk of oral cancer. *Lancet* **367**: 196.
- Horton, R. (2006b) Retraction – non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study. *Lancet* **367**: 382.
- Kane, R.L., Wang, J., & Garrard, J. (2007) Reporting of randomized clinical trials improved after adoption of the CONSORT statement. *J Clin Epidemiology* **60**: 241–9.
- Lesaffre, E., Garcia-Zattera, M., Redmond, C., Huber, H., & Needleman, I. (2007) A review of the reported methodological quality of split-mouth studies. *Journal of Clinical Periodontology* **34**: 756–61.
- Mills, E.J., Wu, P., Gagnier, J., Devereaux, P.J. (2005) The quality of randomized trial reporting in leading medical journals since the revised CONSORT statement. *Contemp Clin Trials* **26**(4): 480–7.
- Moher, D., Cook, D.J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D.F. (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. Quality of Reporting of Meta-analyses. *Lancet* **354**: 1896–1900.
- Moher, D., Jones, A., & Lepage, L. (2001) Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* **285**: 1992–5.
- Robinson, P.G., Damien Walmsley, A., Heanue, M., *et al.* (2006) Quality of trials in a systematic review of powered toothbrushes: suggestions for future clinical trials. *J Periodontol* **77**(12): 1944–53.
- Sjögren, P. & Halling, A. (2002) Quality of reporting randomized clinical trials in dental and medical research. *Br Dent J* **192**(2): 100–3.
- Stroup, D.F., Berlin, J.A., Morton, S.C., *et al.* (2000) Meta-analysis of observational studies in epidemiology. A proposal for reporting. *JAMA* **283**(15): 2008–12.
- Sudbø, J., Lee, J.J., Lippman, S.M., *et al.* (2005) Non-steroidal anti-inflammatory drugs and the risk of oral cancer: a nested case-control study. *Lancet* **366**: 1359–66.
- Tu, Y.K., Maddick, I., Kellett, M., Clerehugh V, & Gilthorpe MS. (2006a) Evaluating the quality of active-control trials in periodontal research. *J Clin Periodontol* **33**(2): 151–6.
- Tu, Y.K., Neslon-Moon, Z.L., & Gilthorpe, M.S. (2006b) Misuses of correlation and regression analyses in orthodontic research: the problem of mathematical coupling. *Am J Orthod Dentofacial Orthop* **130**(1): 62–8.
- von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C., & Vandembroucke, J.P. for the STROBE initiative (2007) The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* **370**: 1453–7.

