Evidence-based Dentistry: Research Design and Levels of Evidence

The principles and methods of evidence-based dentistry give dentists the opportunity to apply

- 3 relevant research findings to the care of their patients. The key to finding evidence is to start with a focused, well-built clinical question. A clear question will help you to identify key words for use in your strategic search. Once evidence has been found, you need to decide if the results are believable
- 6 and whether the findings can be applied to your patient. Assessing the *validity* (closeness to the truth) and the *relevance* (importance and usefulness) of the evidence is called critical appraisal. The purpose of this paper is to discuss the concept and rationale for "levels of evidence" and the types of
- 9 research designs that are appropriate for answering the clinical questions most commonly encountered in dental practice.

The Evidence Hierarchy

- 12 Evidence-based practice involves tracking down the available evidence, assessing its validity and then using the "best" evidence to inform decisions regarding care. Rules of evidence have been established to grade evidence according to its strength. Systematic reviews and randomized
- 15 controlled trials represent the highest levels of evidence, whereas case reports and expert opinion are the lowest. This "ladder of evidence" was developed to a large extent for questions related to interventions or therapy. For questions related to diagnosis, prognosis or causation, other study
- 18 designs such as cohort studies or case-control studies will often be more appropriate. For these types of studies, it is useful to think of the various study designs not as a hierarchy, but as categories of evidence, where the strongest design which is possible, practical and ethical should be used.

21 The evidence ladder

- 1. High-quality systematic reviews.
- 2. Large randomized trials with clear-cut results.
- 24 3. Small randomized trials with uncertain results (i.e., positive trends without statistical significance).
 - 4. Nonrandomized trials with contemporary controls .
 - 5. Nonrandomized trials with historical controls.
 - 6. Cohort studies.

27

- 7. Case-control studies.
- B. Dramatic results from uncontrolled studies (e.g., the treatment of infections with penicillin in the 1940s).
 - 9. Case series and other descriptive studies.
- 33 10. Reports of expert committees and opinions of respected authorities, based on clinical experience.

It should be noted that using the "rules" or categories of evidence only helps classify studies based on the type of research design. The quality of each individual study still needs to be assessed for

- on the type of research design. The quality of each individual study still needs to be assessed for strengths and weaknesses using the techniques of critical appraisal.
 Basic Concepts of Research Design
- 39 Clinical research can be experimental or observational. In experimental studies, the intervention is under the control of the researcher, whereas in observational studies, the researcher observes patients at a point in time (cross-sectional studies) or over time (longitudinal studies). If the
- 42 observations are made by looking forward and gathering new data, the study is prospective; if the data already exist (for instance, in dental records or as census data), the studies are retrospective. **Experimental Studies**
- 45 Experimental studies can be either controlled (there is a comparison group) or uncontrolled. Uncontrolled studies provide very weak evidence and should not be used to guide practice. These studies may be carried out early in an area of research to explore the safety of a new intervention, to
- 48 identify unanticipated effects and to gather baseline data for the planning of more definitive trials. For similar purposes, a study may use a historical control group, where data would be gleaned from a chart review or a previous study. These designs are generally weak because many factors may

have changed since the data were gathered and there are no assurances that bias was not introduced in the collection, recording or retrospective interpretation of the data.

3 Randomized Controlled Trials

Randomized controlled trials (RCTs) are the gold standard by which all clinical research is judged. The fact that randomization keeps study groups as similar as possible from the outset, together with

- 6 other features of the design, such as blinding, sample size justification, appropriate outcome measures and statistical analysis, means that RCTs have the greatest potential to minimize bias. Bias is any factor or process that acts to deviate the results or conclusions of the study away from
- 9 the truth, causing either an exaggeration or an underestimation of the effects of an intervention. In fact, methodology research has shown that, most often, bias and weak designs cause trials to conclude that a treatment is effective when it really may not be, and to overestimate the effect, even

12 when it is true. Randomization of treatment allocation is what makes the RCT one of the simplest and most powerful tools of scientific research. In any study involving people there are potentially many

- 15 unknown factors genetic or lifestyle factors, for example which can have a bearing on the outcome. Randomization, if done properly, reduces the risk that these unknown factors will be *seriously* unbalanced in the various study groups. The allocation sequence *must* be randomly
- 18 allocated. This can be by the flip of a coin, or more usually, by using random number tables or computer-generated sequences. Dates of birth (even or odd), chart numbers or any other alternating type of sequence is inappropriate, because there is the potential for people associated with the study,
- 21 either directly or indirectly, to guess the sequence. Blinding is another key feature of RCTs. The "double-blind" trial is one in which both the researcher and the patient do not know whether the patient is in the experimental group or the
- 24 control group. This design is most useful when the control group is receiving an identical placebo drug or "sham" intervention, but falls down in many types of important studies. Few patients would agree to participate in a study where the control group received "sham" orthognathic or TMJ
- 27 (Temporomandibular joint) surgery. Surgical trials, by necessity, are "open" trials, since both the investigator performing the surgery and the patient know the intervention. However, there are 3 other groups or individuals who can be blinded. The investigator evaluating the outcome must not
- 30 be the surgeon who performed the operation and should be kept unaware of the intervention (the patient must be thoroughly informed about the importance of not "dropping hints"). Although a surgical scar is usually a giveaway, the outcome measure should be planned with this in mind. The
- 33 other 2 groups who can be kept blinded are the statistician(s) doing the data analysis and the investigators who write the results of the trial. To do this, the allocation code is not broken until these components are completed. Although blinding of the statistician is being done with increasing 36 frequency, blinding of the investigator writing the report is rarely done.

Observational Studies

39

51

RCTs cannot answer all clinical questions. There are situations where they may not be necessary, appropriate, ethical or feasible, or they simply may not have been done yet. In general, questions of therapy are best answered by RCTs, or even better, meta-analyses if available, whereas questions of

diagnosis, prognosis and causation may be best addressed by observational (sometimes called
"epidemiological") studies. Observational studies, which are frequently undertaken in dentistry, can be even more challenging to design and execute, in terms of controlling bias. Therefore, it is very

important to use critical appraisal methods to assess the validity of these studies.

45 To a large extent, the type of observational study done depends on the rarity of the disease or condition and on issues related to human resources and economics. Usually several methods of answering the question are possible and the strongest design should be used. The following are some of the most common types of observational studies.

The Cohort Study

In a cohort study, it is known at the outset whether people have been exposed or not to a treatment or possible causal agent (i.e., a vaccine, a drug or an environmental toxin) and are divided into groups or cohorts (treated or exposed versus nontreated or nonexposed) on this basis. They are then followed forward in time (prospectively) for years or even decades to see how many in each group

- 3 develop a particular disease or other outcome. These studies are usually less expensive and easier to administer than RCTs. They may also be ethically more acceptable, because a potentially beneficial treatment is not withheld, and conversely, a possibly harmful treatment is not given. The major
- 6 disadvantage is that we can never be sure that the cohorts are well matched and that there are not other factors, such as social class or occupational exposure that may influence the results. In addition, for rare disorders, the sample size or length of follow-up needed to show an effect may be
- 9 prohibitively large.
 One of the most famous cohort studies followed 40,000 British doctors in 4 cohorts (non-smokers, light, moderate and heavy smokers) for 40 years, from 1951 to 1991. This study, which achieved
- 12 94% follow-up, was instrumental in establishing the causal link between smoking and lung cancer and other diseases, as well as the dose–response relationship between smoking and lung cancer. This study showed the tremendous strength of a well-designed cohort study.
- 15 A variation of a cohort study is a longitudinal study in which there is only one group. Included in the group (called the *inception cohort*) are people who have a positive screening test (for example, for a new genetic marker) or who have all been diagnosed with an early stage of a disease (for
- 18 example, multiple sclerosis). They are then followed and evaluated on a repeated basis to assess the development of the disease (i.e., in the example of the genetic marker), or the time frame for particular outcome measures, in the case of a chronic disease.

21 The Case-control Study

In this type of study, people with a particular condition (the "cases") are matched with a group of people who do not have the disorder (the "controls") and the researchers look back in time to

- 24 determine the proportion of people in each group who were exposed to the suspected causal factor. This is a relatively quick and inexpensive study and is often the best design for rare disorders or when there is a long time lag between the exposure and the outcome. An example of case-control
- 27 studies is the ones that examined the role of water fluoridation in the prevention of dental caries. A major disadvantage of these studies was that some of them were performed without randomization. Cross-sectional Studies
- 30 This design attempts to establish an association between a possible causal factor and a condition, by determining an exposure to the factor and "caseness" at the same time. For instance, a large cross-section of women might be interviewed to determine if they had given birth to a baby with a cleft
- 33 palate *and* if they had taken a particular drug during pregnancy. Although this type of study is relatively easy and inexpensive to carry out and ethically acceptable, it can only establish an association, not a cause and effect relationship. In addition, both "exposure" and "caseness" may
- 36 depend on accurate recall of past events.Case Reports and Case Series

39

Case reports and case series are often used to describe a condition (like a new or rare disorder), a new treatment or innovation, or adverse effects of an intervention. They often provide information which cannot be conveyed in a trial. The description of cases may alert the world to important new

- which cannot be conveyed in a trial. The description of cases may alert the world to important new problems and then allow hypotheses to be developed, leading to focused studies of stronger design.
 42 Case studies and case series are relegated to the lowest rungs of the evidence ladder, however,
- 42 Case studies and case series are relegated to the lowest rungs of the evidence ladder, however, because isolated observations are collected in an uncontrolled, unsystematic manner and the information gained cannot be generalized to a larger population of patients.
- 45 **Warning:** Basing important clinical decisions on single trials, especially when the result is a change in treatment policy, is risky. Because of the numbers of patients needed to detect small to moderate differences for clinically important outcome measures, definitive answers may not be
- 48 found in single studies, unless they are systematic reviews or well-designed "large simple trials." These "mega" trials, which usually involve many thousands of patients, have rarely been carried out in dentistry.
- 51 Author: Susan E. Sutherland, DDS (*J Can Dent Assoc* 2001; 67:375-8)