THERAPY

Anecdote, Experience, or Evidence?

Gary R. Goldstein DDS, and Jack D. Preston, DDS

How does a practitioner determine what therapy to use? Often, the decision depends on the age of the practitioner and the experiences gained in practice. The younger practitioner depends mainly on what was taught in dental school. All dental schools have a core technique, usually derived by faculty consensus, that allows a student to develop competency in one approach to a therapeutic problem. Trying to teach a novice multiple techniques usually results in the student's mastering none. Educators have agreed that teaching one technique well allows the student to enter practice and satisfy the needs of the public. Unfortunately, dental schools have been unfairly criticized as teaching outdated and often unrealistic techniques. This criticism is not true. Dental school faculty almost universally teach time-tested and scientifically sound procedures. Ethics dictate that patients in dental schools be protected and not subject to whimsical trends in treatment. Internal review boards mandate that research be structured to ensure the patient's rights are preserved. The clinician, unencumbered by such constraints, often makes forays into other treatment modalities, some successful, others disappointing. Once in practice, the clinician is influenced by observations based on experience. Such observations, however, are often flawed, and associations thought to be causal are instead, only casual. Anecdotal evidence from colleagues may mold decision-making. With the broad communication now possible using with the internet, such anecdotes may come from a continent away and from a completely unknown

From the Department of Prosthodontics, New York University College of Dentistry; Department of Dental Material Science, New York University Graduate School of Arts and Sciences, New York, New York (GRG); and The University of Southern California School of Dentistry, Los Angeles, California (JDP)

person. Conversely, upon graduation some clinicians become comfortable with a particular procedure and may be wary of change.

As clinicians expand their knowledge through lectures and by reading journals, they constantly modify their clinical methods. There is always a new restorative material on the market, a new surgical technique, a new piece of equipment, a new toothbrush, and a new toothpaste. Detailers joke that dentists are gadget enthusiasts who buy a product, use it once or twice, and store it in some cabinet, finding it years later and not remembering when, where, or why it was purchased.

How do practitioners decide which treatments to use? Often, they are influenced by the prestige of the professor giving the lecture or writing the article. All too often, however, they are seduced by the *show* rather than by the science. Multiple projectors, enhanced digital presentations, or the glitz of the advertising become the main reasons for change. Companies market directly to the public who, with inadequate ability to evaluate the hype, pressure the practitioner to change therapy, often with inadequate research to justify the change.

An example is a patient with an edentulous area who presents with the request for implants. What the patient is really saying is, "I want teeth." It is the practitioner's responsibility to understand that the patient is requesting the ability to chew better, speak better, or look better. It is the practitioner's responsibility to determine the best therapy for that individual patient and to advise the patient of that therapy and any other suitable options. Another example is a patient who, having heard all the hype on tooth bleaching, requests the procedure when the problem is really recurrent decay around old, severely stained composite restorations that need replacement.

Any procedure involves some risk, and increasing risk usually accompanies more complex therapy. The practitioner should decrease that risk as much as possible without unduly burdening the patient. Patients have a moral, ethical, and legal right to know the risks and benefits of any therapy that is recommended.

Today, information may be obtained from a variety of sources. There are often newer procedures to supplant the approaches documented in textbooks. Reports in peer-reviewed journals are more current, depending upon the source and the publication delays. Today, many practitioners obtain information over the internet, through conversations with other practitioners, and through newsletters and non–peer-reviewed periodicals and journals. These *less scientific* sources can be useful. For example, the problems of root fracture when cementing dowels and the fracture of porcelain complete-coverage restorations when using the first-generation resin/ionomer cements were first made public in these forums. Regardless of how information is obtained, anyone seeking newer approaches to improve the delivery of dental service must apply the rules of evidence in evaluating a suggested technique. Failure to consider all aspects of a therapy have sometimes proven disastrous (e.g., the teratogenic effects of thalidomide) or merely ineffective after encouraging initial results (e.g., early treatments for

AIDS). Therefore, the alert practitioner walks a tightrope between endangering patients with a therapy that has an undetected accompanying risk and failing to provide optimal therapy that would be of substantial merit.

Anyone considering a new course of care must invoke the rules of evidence and evaluate the strengths of a report against its inadequacies. A technique or regimen may have statistical significance but lack clinical merit. How, then, does the practitioner maintain balance on that tightrope and best serve the patient? The rules of evidence have been well established; their benefit lies in their knowledgeable application. A report may not furnish all the information desired, or the data may be reported in such a manner that they are difficult to evaluate. Bias from well-meaning researchers is common, and dentistry is filled with volumes of pseudoscientific reports in which results have been derived from a false premise or a flawed research design.

Unfortunately, some established dental procedures have gained acceptance because a charismatic champion of the technique was a convincing advocate. Often, procedures that had merit were based on a falsely attributed cause and have been successful for reasons other than those to which their success has been ascribed.

When considering the merits of a report or lecture, the practitioner must clearly understand the purpose of the study and how the investigators sought to establish their premise. The results of the study must relate directly to this purpose statement. Anything not established as a purpose of the study should not be given primary consideration.

Subjects enrolled in the study must all have an equal chance to obtain the study parameter (e.g., drug, treatment regimen, material) rather than the alternative approach (e.g., placebo, previously accepted technique or regimen, no therapy). Those in the treatment and alternative groups must be equivalent in all pertinent respects.

Before being enrolled in a study, a person should go through a complex screening process that establishes them in the appropriate cohort. Patients have a dental or medical problem and choose a treatment facility. They may enter that facility at different stages of the disease process and hence may have a different prognosis. They may be motivated by cost, location, or the reputation of the facility or the treating doctor. After screening, the patient is referred to the researcher, whose study population is further filtered by the informed-consent or volunteer process. Investigators also tend to include persons perceived to be compliant to ensure their continuance in the study and to rule out apparently less-compliant or difficult candidates. An additional series of eligibility decisions are then made to reduce the population further. Inclusion and exclusion criteria must be clearly established. They are necessary but must be pragmatic and relevant. As investigators cull the potential population using appropriate demographic criteria, they also rule out persons with potentially confounding comorbidities. Ultimately, clinicians must ask if the results are applicable to their patient population. If the sample group is excessively homogenized, the study population may

not be representative of the clinicians' patients, and the study may have decreased validity.

For example, in a well-done study on IPS Empress inlays and onlays, the population (N = 130) consisted of 27 one-surface inlays, 38 two-surface inlays, 40 three-surface inlays,[8] and 25 onlays.[1] A significant percentage of the population consisted of class I restorations; therefore, the data may not be pertinent to a clinician who does not normally perform class 1 restorations.

Exactly what a study is to measure must be determined in advance, and the methods of measurement of the effects must be clearly and specifically stated. The precision of the measurements (or the converse, the error of the study) must be established before the study is initiated. It is not enough to know that the microscope used had a precision of 5 μm. The ability of investigators to repeat their measurement is crucial. How many persons were involved in making the measurements? Was their equipment calibrated to ensure that the measurements were equivalent? The method by which the study is to be analyzed must also be established a priori. Too often, investigators gather data only to find that statistical analysis is compromised by the procedures used.

The outcome assessment must be relevant. Investigators sometimes are encumbered by the dogma that the only legitimate way to do an experiment is to vary one factor at a time.[1] This univariate approach is at odds with the multivariate climate in which the clinician functions. For example, in reviewing the current literature for dental luting cements, Rosenstiel et al[6] listed 10 different clinically important parameters. A study that concentrates on only one factor may not supply enough information to warrant a change in material.

Readers must be acutely aware of the structure of the study before trying to ascertain its applicability to their patients. The design is determined by the direction of inquiry, who determines the therapy, and the presence of a control group. Prospective studies are those in which the therapy is initiated at the start of the study. The advantage of a prospective trial is that, theoretically, the investigator can control all aspects of the treatment and minimize the effect of confounding variables. Retrospective studies are those in which the therapy was initiated before the beginning of the study. The disadvantage of this study design is the inability of the investigator to control inadvertent confounding variables.

Studies can be further divided into comparative studies, (also called analytical studies) that have a control group, and descriptive studies with no control group.

The hierarchy of evidence can be listed as

1. Comparative studies
   - Prospective studies

     Randomized, controlled trials (RCTs) – assignment to therapy is under the control of the investigator

     Cohort study – two matched study groups (cohorts) are assem-

bled and followed. Because the patient self-selects the treatment, the assignment to therapy is not under the control of the investigator.

- Retrospective studies

  Case-control study – similar in design to the cohort study except that the outcome was present at the time the study began

2. Descriptive studies
   - Case series
   - Case report

The hierarchy of evidence gives the reader a primer to use when comparing conflicting evidence.

The RCT study is the standard for questions regarding therapy. Because the study is prospective, and the therapy is under the command of the investigator, an RTC minimizes the bias inherent in other designs. The control, usually the standard of care, allows the reader to make direct comparisons; hence, this design provides the best evidence. Understanding the terminology is more important than simply recognizing the terminology. Randomization, however, cannot make up for a poorly planned and implemented study. Just because the design is an RCT does not relieve the reader from the responsibility of examining the methodology.

Randomization is effective only when the study population is of compelling size. The appropriate study population size in a clinical trial varies for different questions. Determining the size of the study population requires a knowledgeable best guess by the researcher and consultation with a statistician to determine the power of the study.

Randomized, controlled trials are not always possible because of cost, time, or ethics. For example, it would not be ethical, in performing a study on the hazards of smoking, to randomize a cohort to a regimen that forced a participant to smoke two packs of cigarettes a day to see if there was a harmful result. Rather, matched cohort studies, although not ideal, are accepted as the norm to answer a question of harm.[4] Sackett et al have concluded that "Evidence-based medicine is not restricted to randomized trials and meta-analyses. It involves tracking down the best external evidence (from systematic reviews when they exist; otherwise from primary studies) with which to answer our clinical questions."[7]

The determination of an acceptable control refers back to the question the study is trying to answer. In a study to determine the efficacy of a new drug, a placebo could be an acceptable control. The pharmaceutic industry, however, is concerned about the extent of the placebo effect, which can be as high as 30%. A current trend is to have a three-group (instead of the typical two-group) RCT. One group receives the new drug, one group receives the placebo, and the third group (the control)

receives no treatment. The difference between the no-treatment group and the placebo group is the placebo effect, whereas the difference between the placebo group and the treatment group is called the therapeutic effect.[2] In therapy trials, the most useful control for the clinician is the current standard of care. The usefulness of a study on a new headache medication is enhanced if the control is aspirin, ibuprofen, or acetaminophen rather than a placebo, because most clinicians would not prescribe or take a placebo for a headache.

Unfortunately, most dental therapy articles are descriptive rather than comparative. In the typical case study or case series, practitioners evaluate their own work. Despite the integrity of the clinician, the study cannot have the same validity as one in which an independent, blinded observer assesses the outcome. Another problem of descriptive studies is that authors sometimes want to project their data beyond the scope of their project. For example, it is easy when doing a case series on implant product X to compare it with another case series done on implant product Y. This comparison is dangerous, because the two studies had different populations, in different settings, receiving different therapy from different investigators. The groups are almost always dissimilar, and treatment regimens are almost always different. Although such a comparison is acceptable in the discussion section, it should never appear in the conclusions. Conclusions can report only the results of the present study in answer to the initial question or hypothesis.

Journals and authors usually express results in positive numbers, a practice which can be misleading to the clinician. For example, an 85% success rate might sound impressive, but viewing the same results as a negative (a 15% failure rate) may have more impact on the decision making process.[3] An example is a recent article in the *Journal of the American Dental Association* evaluating Class V restorations with and without mechanical retention.[5] The authors claimed that "restoration of Class V lesions without using mechanical retention could be expected to succeed in seven of 10 restorations over a three year period," but clinicians must determine if a 30% failure rate after 3 years is an acceptable result in their practices.

Results are also presented in terms of statistical significance, and unfortunately statistical significance does not always relate to clinical significance. For example, an investigator may use an extremely accurate measurement device which can report attachment loss around teeth in tenths of a millimeter. After 2 years of treatment with drug X, the study shows a statistically significant attachment loss of 0.01 mm when compared with scaling and root planning. But are the results clinically significant, especially if the drug therapy causes an after-effect? The clinical relevance of a statistically significant finding is best determined by the clinicians reading the report and determining if the results are applicable to their patients.

The clinician must also understand the difference between a biologic response and a clinical response. A new mouthwash may demonstrate the ability to kill more bacteria or viruses (a biologic response), which

may have no clinical relevance. If, however, a clinical response such as a decrease in periodontal disease, caries, or malodor can be demonstrated with the use of the mouthwash, the data have relevance for the clinician. Another example is the family of glass ionomer restorations and cements. Fluoride release (a biologic response) is meaningless to the practitioner unless a documented decrease in caries (a clinical response) can be demonstrated.

When confronted with evidence that conflicts with the current standard of care, an experienced clinician can be biased in evaluating data. "I've been doing this for years and it works in my hands," may not be an acceptable excuse to disregard compelling studies.

The definition of success is controversial, and consensus is often difficult to achieve. How long should a restoration last? If it is still in place, but staining compromises the esthetics, is the procedure a success or failure? How long should an implant last? If there is a loss of osseointegration along one wall of the implant, but the fixture is rigid and there is no pain, is the implant a success or a failure? What is an acceptable success rate for molar endodontics? If there is a small periapical radiolucency, but the patient has no pain, is it a success or failure? If the patient has intermittent recurrent pain, but the radiograph demonstrates a perfect fill, is it a success or a failure?

Success is also tempered by the cause of the failure. Recurrent decay or periodontal problems that compromise a full-coverage restoration and are caused by poor home care are different from the same problems caused by defective margins. A patient fracturing a restoration by biting into an olive pit is different from a failure caused by an overlooked occlusal prematurity. Also, changes in the clinician's advice to patients can cause embarrassing moments if the dentist is not willing to admit that current good research has caused a change in thinking. Should a clinician restore a patient's lost molars? Patients need teeth to masticate, to phonate, and for esthetics. If the lost molars are not in the esthetic zone, the patient has no problem eating or speaking, and extrusion of the opposing dentition has not occurred or is not a concern (the opposing molars are also missing) why restore? This argument has intensified with an article by Witter et al[9] in the *Journal of Dental Research* that showed 9-year prospective evidence questioning the rationale for restoring the missing molars, and the controversy will persist as researchers supply more data.

SUMMARY

In dentistry, most changes in therapy come from new techniques and products that are introduced to the market. Clinicians (and patients) can be overwhelmed by advertisements and marketing, some obvious and some (e.g., paid clinical reports in non–peer-reviewed journals) not so obvious. Because most advances are made with small case studies, which are at a lower level of evidence, it is imperative that data clinicians

read or see have the greatest validity possible. This validity is imperative to achieve evidence-based dentistry that uses relevant, high-quality, clinically oriented research that provides better information for the clinician and better treatment for the patient.

References

1. Brunette DM: Critical Thinking. Understanding and Evaluating Dental Research. Chicago, Quintessence Publishing Co, 1996
2. Chambers D: The big placebo. J Am Coll Dent 66:2–5, 1999
3. Goldstein GR, Preston JD: How to evaluate an article about therapy. J Prosthet Dent 83:599–603, 2000
4. Jacob RF, Lloyd PM: How to evaluate a dental article about harm. J Prosthet Dent 84:8–16, 2000
5. McCoy RB, Anderson MH, Lepe X, et al: Clinical sucess of class V composite resin restorations without mechanical retention. J Am Dent Assoc 129:593–599, 1998
6. Rosenstiel SF, Land MF, Crispin BJ: Dental luting agents: A review of the current literature. J Prosthet Dent 80:280–301, 1998
7. Sackett D, Richardson WS, Rosenberg W, et al: Evidence-based Medicine: How to Practice and teach EBM. New York, Churchill Livingstone; 1997
8. Struder S, Lehner C, Brodbeck U, et al: Short-term results of IPS-Empress onlays and inlays. Journal of Prosthodontics 5:277–287, 1996
9. Witter DJ, Creugers NHJ, Kreulen CM, et al: Occlusal stability in shortened dental arches. J Dent Res 80:432–436, 2001

Address reprint requests to

Gary R. Goldstein, DDS
Division of Restorative and Prosthodontic Sciences
New York University College of Dentistry
345 East 24th Street
Clinic 5W
New York, NY 10010–4086

e-mail: gary.goldstein@nyu.edu