SCIENTIFIC ARTICLE

Assessment of the Quality of Reporting Observational Studies in the Pediatric Dental Literature

Yogita Butani, BDS, MS[1] Arthur Hartz, MD, PhD[2] Steven Levy, DDS, MPH[3] Catherine Watkins, DDS, PhD[4] Michael Kanellis, DDS, MS[5] Arthur Nowak, DDS, MA[6]

Abstract

Purpose: The purpose of this assessment was to evaluate reporting of observational studies in the pediatric dental literature.

Methods: This assessment included the following steps: (1) developing a model for reporting information in clinical dentistry studies; (2) identifying treatment comparisons in pediatric dentistry that were evaluated by at least 5 observational studies; (3) abstracting from these studies any data indicated by applying the reporting model; and (4) comparing available data elements to the desired data elements in the reporting model.

Results: The reporting model included data elements related to: (1) patients; (2) providers; (3) treatment details; and (4) study design. Two treatment comparisons in pediatric dentistry were identified with 5 or more observational studies: (1) stainless steel crowns vs amalgams (10 studies); and (2) composite restorations vs amalgam (5 studies). Results from studies comparing the same treatments varied substantially. Data elements from the reporting model that could have explained some of the variation were often reported inadequately or not at all.

Conclusions: Reporting of observational studies in the pediatric dental literature may be inadequate for an informed interpretation of the results. Models similar to that used in this study could be used for developing standards for the conduct and reporting of observational studies in pediatric dentistry. (Pediatr Dent 2006;28:66-71)

KEYWORDS: COMPARATIVE STUDIES, EPIDEMIOLOGIC RESEARCH DESIGN, META-ANALYSIS, REPRODUCIBILITY OF RESULTS, REPORTING, CRITERIA FOR OBSERVATIONAL STUDIES

Received July 11, 2005 Revision Accepted November 11, 2005

Several studies have reported the relative lack of outcomes research necessary for evidence-based dental practice[1-3] and pediatric dental practice.[4] Development of evidence-based standards could help reduce the inconsistencies in clinical practices.[5] Improved standards for conduct and reporting, however, are necessary for the research to have the greatest positive impact. Although there has been recent attention to adopting standards for reporting of randomized controlled trials in dentistry,[6,7] there has been little discussion about standards for conducting and reporting observational studies.

Observational studies (OSs) are those in which investigators collect data, but treatment decisions are determined by the clinicians and patients. They have several advantages compared to controlled trials in which treatments are assigned by the investigators based on a list of random numbers. In general, OSs:

1. cost less;
2. can be completed more quickly; and
3. do not require patients or providers who are willing to be randomized to treatments.

For these reasons, they constitute the majority of dental research.[2,4,5,8] OSs, however, may give invalid results. At a minimum, these studies must take into account multiple factors that influence outcomes.[9]

In conducting an OS, it is helpful to consider a model such at that presented in Figure 1. This model summarizes

[1]Dr. Butani is postdoctoral fellow, Department of Preventive and Restorative Dental Sciences, School of Dentistry, University of California, San Francisco; [2]Dr. Hartz is professor, Department of Family Medicine, Carver College of Medicine, [3]Dr. Levy is Wright-Bush-Shreves professor and graduate program director, Department of Preventive and Community Dentistry, and professor, Department of Epidemiology, College of Public Health, [4]Dr. Watkins is adjunct assistant professor, Department of Preventive and Community Dentistry, College of Dentistry, [5]Dr. Kanellis is professor, graduate program director, and assistant dean for patient care, and [6]Dr. Nowak is professor emeritus, Department of Pediatric Dentistry, College of Dentistry, all at the University of Iowa, Iowa City, Iowa; [6]Dr. Nowak is also executive director, American Board of Pediatric Dentistry, Iowa City.
Correspond with Dr. Hartz at arthur-hartz@uiowa.edu

4 domains that influence dental outcomes . How the study manages these domains can influence whether the results are valid (ie, whether there is selection, measurement, or confounding bias[10]) and whether they are relevant for a patient who has specific risk factors, treatment options, and outcome goals. For adequate interpretation of an observational study, all the information in these domains should be reported.

Examples of factors that could influence validity are size of cavity or oral hygiene level. If these factors are not equivalent in the treatment groups, then they might cause confounding by influencing the differences in the outcomes between the 2 treatment groups. Other factors might be the same for the 2 treatment groups, but could influence the relative effectiveness of the 2 treatments. For example, in studies with primarily inexperienced practitioners, treatment A could appear better than treatment B. In a study with primarily experienced practitioners, however, treatment B might appear better than treatment A.

The present study assessed the quality of reporting of observational studies in the pediatric dental literature. It compared what was reported to what ideally should have been reported, based on a model of factors that can influence a study's validity and/or a treatment's relative effectiveness. This model is a necessary first step toward developing standards for conducting and reporting observational studies in dentistry.

Methods

Pediatric dentistry studies were reviewed to assess how well observational studies of treatment comparisons reported factors required for interpretation. Because the study was based on a literature review, IRB approval was not required. The types of studies examined were those that compared treatments (eg, materials, procedures, or techniques). In addition to the treatment, these comparisons should take into account other factors that influence outcomes. One of these factors is patient status, which includes the teeth's condition, the ability to comply with treatment in the office and afterwards, and other health and lifestyle characteristics. The outcome can also be influenced by the provider's skill and available resources. Even if these domains are constant, the study's results can be influenced by design issues, such as how treatment failure is defined and measured, or the length of patient follow-up.

To obtain articles for this study, the literature was searched for observational studies reported from 1985 through 1998. Although MEDLINE is now indexed for highly sensitive searches for randomized, controlled trials, observational studies is not an indexable concept in MEDLINE, and there is no search term for observational studies (Wright N, National Library of Medicine: personal communication). Therefore, a text-word strategy was used to search for the terms "observational," "cohort," "retrospective," "cross-sectional," and "nonrandomized."

This search identified 12 comparisons of dental treatments that were evaluated with observational studies. Five comparisons were excluded because they were not full-length studies in dental journals, written in English and utilizing patients less than 13 years of age. This left the following 7 topics:

1. amalgam vs composite restorations in posterior teeth;
2. use of space maintainers vs no treatment for prematurely lost primary teeth;
3. formocresol vs ferric sulfate pulpotomy;
4. stainless steel crowns (SSCs) vs multisurface amalgam;
5. amalgam vs crowns for large carious lesions;
6. amalgam vs glass ionomer restorations; and
7. chlorhexidine vs sodium fluoride mouthrinse as anticaries rinse.

Medline was then searched for observational studies evaluating each of these treatment comparisons.

Articles were abstracted that compared treatments evaluated by at least 5 observational studies. Requiring several observational studies of a particular comparison made it easier to identify key factors that should be reported and assess whether results and reporting varied across studies. Only 2 treatment comparisons met all of the inclusion criteria: (1) comparisons between SSCs and multisurface amalgams; and (2) comparisons between amalgam and composite restorations in children.

To identify all additional studies that evaluated SSCs and amalgams, the following search words were used: (1) "stainless steel crown(s)"; (2) "preformed metal crown(s)"; (3) "preformed crown(s)"; and (4) "ion-chrome crown(s)." Articles that compared composite and amalgam restorations were searched using the following terms: (1) "composite"; (2) "tooth-colored restoration"; and (3) "composite resin."

From each of the identified articles one of the authors abstracted the types of information shown in Figure 1. Specifics of this information are shown in Table 1. They were determined based on clinical experience of the authors and from the information recorded by at least one of the reviewed articles. Abstracted information was tabulated to assess the consistency of reporting the type of information shown in the comprehensive model.

Chi-square tests of 2 by k contingency tables were used to determine the statistical significance of variation across studies of failure rates for a given treatment. The difference between 2 odds ratios was tested using the equation $Z=(Ln_1-Ln_2)/\sqrt{(SE_1 + SE_2)}$, where:

1. Z has a normal distribution with a mean of 0 and a variance of 1;
2. Ln_1 and Ln_2 are the natural logarithms of the 2 odds ratios; and
3. SE_1 and SE_2 are the standard errors of these logarithms.

Variation among odds ratios was tested with the Breslow-Day test for homogeneity at the $P<.05$ level.

Results

For comparisons of SSCs and amalgams in children, a total of 13 studies reported in 16 articles[11-26] was found; 10 of the 13 studies[11-23] described the outcome definition. In general, when studies were reported in more than one article, the data differed primarily in the length of follow-up. For amalgam and composite restorations (CRs) in primary teeth, 5 studies reported in 8 articles[20,22,23,27-31] were found that compared restorations in children. Other studies were found evaluating restorations in adults, which were not included in this review.

Failure rates for the reviewed studies are shown in Tables 2 and 3. In both tables, there is substantial and statistically significant variation ($P<.001$) in failure rates for each treatment. There is also substantial and significant variation in the odds ratios comparing the 2 treatments.

The next step in the analysis was to determine whether there was adequate reporting of the study features that influenced variation. To do this, the following characteristics listed in Table 1 were abstracted from every article: (1) provider characteristics; (2) patient characteristics; (3) materials and procedure characteristics; and (4) study characteristics. The reporting of each of these characteristics is described below.

Two provider characteristics were generally presented:

1. operator type was described in all 5 of the CR studies and 9 of the 10 SSC studies; and
2. practice type was reported in all 15 of the studies.

None of the studies reported information about the provider demographics or number of procedures the provider performed in the past year.

The patient characteristic most often reported was the type of tooth restored, which was reported in all 15 studies. Some information about age was reported in 14 of the 15 studies, but none of the studies gave an actual age distribution that would make it possible to compute the number of patients younger than a particular age. Information on caries activity was reported in 3 of the SSC studies and none of the CR studies, and information on fluoride history was reported in 1 of the SSC studies and none of the CR studies. No studies reported on structural damage to the teeth, the quality of the occlusion, the patient's cooperation, or oral hygiene.

More information was reported about the characteristics of materials and procedures. The type of material used was reported for 3 SSC and all CR studies, use of pharmacological agents was reported for 5 SSC and all CR studies, and the use of rubber dams was reported for 3 SSC and 3 CR studies.

The best reported study characteristics were those that described the study design and sample size. Some information was always reported about length of follow-up, but the percentage of patients who had a given length of

Table 1. Specific Factors That Could Influence Outcomes

Provider characteristics	Patient characteristics	Materials and procedure characteristics	Study characteristics
Practice type (private practice or hospital setting)	Age at the time of placement of restoration	Composition of material(s)	Study design: (eg cohort, case-control study)
Practice site (country)	Patient cooperation	Isolation (use of rubber dam)	Sample size at the start of study
Operator type (provider training)	Oral health status	Preoperative diagnostic radiographs	Patients lost to follow-up
No. of operators	Caries activity	Use of pharmacological management techniques	Length of follow-up
Provider expertise: ability to manage patients (eg, patient satisfaction); ability to manage materials and techniques (eg, direct observation or success rate of a common procedure)	Dietary history		Patient inclusion/exclusion criteria
	Fluoride history		Outcome measure: clinical, radiological, or histological aspects
	Medical history		
	Type of tooth restored: anterior/posterior; primary/permanent; maxillary/mandibular		Unit of analysis
	Occlusion		
Provider demographics: years in practice; graduation year	Condition of the tooth prior to restoration (size of lesion, endodontic treatment)		Statistica analysis
	Location of restorations being compared		
	Expected length of service of the tooth and restoration		
	Expected length of service of the tooth and restoration		

follow-up was never reported. Exclusion criteria were not reported, perhaps because the studies included all patients who had a given procedure for a given purpose. Failure was not defined for 3 SSC studies.[24-26] The other studies clearly defined failure, although the definitions varied substantially in ways that could influence the results. For example, one SSC study considered recementation as the only criterion for crown failure.[12] Another study added the following criteria: amalgam repair of crown, new crown, or extraction,[11] and a third study defined failure as a restoration that needs to be replaced due to recurrent decay.[22-23]

Statistical analysis in the reviewed studies was primarily limited to comparison of the overall results. The only exceptions were 3 studies that analyzed the data stratified by age.[15-18] No studies used regression analysis to adjust for the many factors shown in Table 1 that could have influenced differences between treatment groups. There was also no analysis of statistical interaction that evaluates whether the relative effectiveness of the treatments depended on patient type.

Discussion

For both the SSC and CR treatment comparisons, there was substantial and highly significant variation in the failure rates of any given treatment and in the odds ratios comparing the failure rates for 2 treatments. Because of the amount of this variation, it is clear that there must be substantial variation in the study features that influence outcomes. The major purpose of this study was to assess the quality of reporting of these study features.

The list of features that should have been reported is shown in Table 1. This list was derived by applying the comprehensive reporting model shown in Figure 1 to the specific treatments evaluated. A few elements were reported by most studies, and other elements were reported by some studies, but several study features that could influence the interpretation of the study were never reported. These include extent of structural damage to the tooth, patient oral hygiene, quality of the occlusion, and patient behavior. Information on patient age and length of follow-up was almost always collected but not reported with as much detail as would be helpful.

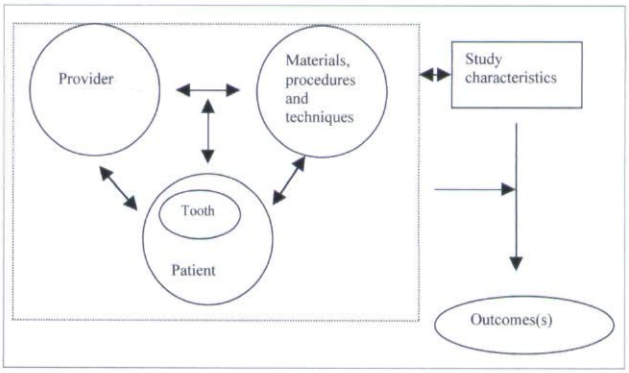

Figure 1. Conceptual model of how study characteristics influence outcomes.

The conclusions on the reporting of observational studies in the dental literature were based on a small percentage of studies that may not be representative of all dental literature. There is no apparent reason, however, why the quality of reporting would be lower for these studies than for others. Even if the studies examined in this report are not representative of all studies in the pediatric dental literature, they may be representative of a substantial percentage of studies.

There are several ways that the factors described in Table 1 can influence results. One is that some of these factors can cause confounding. Confounding occurs when 2 conditions are met:

1. subjects who have one treatment have different characteristics than subjects who have another treatment; and
2. these characteristics influence the risk of failure.

For this reason, it is important to compare the subject's risk factors prior to treatment and, if necessary, use statistical methods to take these risk factors into account when comparing treatment outcomes. It was rare that subject risk factors were compared or that statistical methods were used to take these risk factors into account.

Table 2. Variation in Failure Rates for Studies of Amalgams vs. Stainless Steel Crowns

Author	Failure (amalgam) (%)	Failure (crowns) (%)	Odds ratio
1. Braff[11]	133/150 (89%)	23/76 (30%)	18.02
2. Dawson et al[12]	53/102 (52%)	8/64 (13%)	7.57
3. Paunio et al[13]	66/104 (63%)	22/104 (21%)	6.72
4. Messer et al[16]	242/1117 (22%)	40/331 (12%)	2.34
5. Wong et al[17]	124/233 (53%)	0/18 (0%)	
6. Roberts et al[18]	82/706 (12%)	13/673 (2%)	5.94
7. O'Sullivan et al[19]	17/106 (16%)	7/210 (3%)	5.55
8. Papathanasiou et al[20]	58/198 (29%)	37/183 (20%)	1.65
9. Einwag et al[21]	38/66 (58%)	4/66 (6%)	21.04
10. Tate et al[22]	140/669 (21%)	69/862 (8%)	3.04

Table 3. Variation in Failure Rates for Studies of Amalgams Vs Composites			
Author	Failure (amalgam) (%)	Failure (composites) (%)	Odds ratio
1. Gibson et al[27]	33/61(54%)	35/61 (57%)	.83
2. Derkson et al[30]	18/90 (20%)	51/94 (54%)	.21
3. Tonn and Ryge[31]	4/76 (5%)	11/76 (15%)	.33
4. Papathanasiou et al[20]	58/198 (29%)	79/173 (46%)	.50
5. Tate et al[22]	141/669 (21%)	110/367 (30%)	.63

The second way a factor can influence the results is by changing the effect of one or both treatments. For example, it is possible that a certain treatment could be relatively more effective than standard care if it is performed by a provider with special skill or on a tooth with a particular structural defect. Therefore, to determine whether the results of a given study are likely to apply to a particular practice or to determine why one study gave very different results than another, it is important to know the characteristics of the providers and subjects in the study.

Results also depend on the specifics of the treatment (eg, changes in the materials used or how a procedure was performed could influence the outcome) and characteristics of the study design (eg, how long subjects were followed, whether the data were obtained from dental records or specifically collected for the study, and how a good outcome was defined).

Factors that could influence the study's interpretation even if they don't influence the validity can be referred to as modulating factors. The only modulating factor that was examined in any of these studies was subject age (ie, a few studies examined whether the relative effectiveness of the 2 treatments differed for subjects according to age.

Characteristics that should be reported because they are potentially confounding or modulating factors are those that: (1) influence outcomes; (2) vary across treatments or practices; and (3) can be assessed with reasonable effort. These characteristics are modeled in Figure 1. The model was proposed as a first step in the development of standards for observational studies in the dental literature. There has been a great deal of attention paid to developing standards for conducting and reporting randomized controlled trials in the dental literature by way of adopting the CONSORT statement.[32] The majority of pediatric dental treatment studies, however, have used an observational design, and it is unlikely that this will change in the near future.[4] Better reporting of these studies will make it easier to explain variation of study results and to correctly apply the results in practice.

No method for conducting observational studies is likely to make these studies as good as randomized controlled trials. Observational studies can provide accurate information,[33] however, and most clinical dental research relies on observational studies. Therefore, careful attention should be paid to standards for conducting and improving observational studies.

Conclusions

Based on this study's results, the following conclusions can be made:

1. Observational studies comparing the same 2 treatments may give substantially different results.
2. Studies do not generally report enough information to explain the variation in the results.
3. A framework should be developed for observational studies that will guide:
   a. what data elements should be collected;
   b. how the data should be analyzed; and
   c. how the results should be reported.
4. This framework should include the factors presented in Figure 1 and Table 1.

Acknowledgements

This study was supported in part by grant 1 RO1 HS10739-01 from the DHS-Agency for Healthcare Research and Quality. The article was derived from a thesis by Yogita Butani, DDS, MS, in partial fulfillment of her requirements for an MS in Dental Public Health from the University of Iowa Graduate College, Iowa City, Iowa. The authors appreciate the assistance of Laurie Wallace, who helped with manuscript preparation, and the consultative advice of Steven Hedlund, DDS, and David Foreman, DDS. The authors continue this work in memory of their late colleague, Keith E. Heller, DDS, PhD.

References

1. Bader JD, Ismail AI. A primer on outcomes in dentistry. J Public Health Dent 1999;59:131-135.
2. Bader JD, Ismail AI, Clarkson J. Evidence-based dentistry and the dental research community. J Dent Res 1999;78:1480-1483.
3. Burt BA. Evidence-based dentistry. Its development and use in clinical practice. N Y State Dent J 1999;65:34-40.
4. Butani Y, Levy SM, Nowak AJ, Kanellis MJ, Heller K, Hartz AJ, Dawson DV, Watkins CA. Overview of the evidence for clinical interventions in pediatric dentistry. Pediatr Dent 2005;7:6-11.
5. Bader JD, Shugars DA. Variation in treatment outcomes, and practice guidelines in dental practice. J Dent Ed 1995;59:61-95.
6. Needleman I. CONSORT: Consolidated Standards of Reporting Trials. Br Dent J 1999;13:207.

7. List of organizations supporting CONSORT. Available at: http://www.consort-statement.org/endorsements/journals/journals.html. Accessed February 6, 2006.
8. Nainar SMH. Profile of journal of dentistry for children and pediatric dentistry journal articles by evidence typology: thirty-year time trends (1969-1998). Pediatr Dent 2000;22:475-478.
9. Crall JJ. Pediatric dental treatment outcomes: The importance of multiple perspectives. Pediatr Dent 1998;20:219-220.
10. West S, King V, Carey TS, et al. Systems to rate the strength of scientific evidence. Evidence report/technology assessment no. 47. Prepared by the Research Triangle Institute—University of North Carolina Evidence-based Practice Center under contract no. 290-97-0011. In: AHRQ publication no. 02-E016. Rockville, Md: Agency for Healthcare Research and Quality; 2002.
11. Braff MH. A comparison between stainless steel crowns and multisurface amalgams in primary molars. J Dent Child 1975;42:474-478.
12. Dawson LR, Simon JF, Taylor PP. Use of amalgam and stainless steel restorations for primary molars. J Dent Child 1981;48:420-422.
13. Paunio P, Erikson AL, Isotupa K. Suitability of ion-crowns for the restoration of deciduous molars. Proc Finn Dent Soc 1980;76:25-29.
14. Eriksson AL, Paunio P, Isotupa K. Restoration of deciduous molars with ion-crowns: Retention and subsequent treatment. Proc Finn Dent Soc 1988;84:95-99.
15. Levering NJ, Messer LB. The durability of primary molar restorations: II. Observations and predictions of success of stainless steel crowns. Pediatr Dent 1988;10:74-80.
16. Messer LB, Levering NJ. The durability of primary molar restorations: II. Observations and predictions of success of stainless steel crowns. Pediatr Dent 1988;10:81-84.
17. Wong FSL, Day SJ. An investigation of factors influencing the longevity of restorations in primary molars. J Int Assoc Dent Child 1990;20:11-16.
18. Roberts JF, Sherriff M. The fate and survival of amalgam and preformed crown molar restorations placed in a specialist pediatric dental practice. Br Dent J 1990;169:237-244.
19. O'Sullivan EA, Curzon MEJ. The efficacy of comprehensive dental care for children under general anesthesia. Br Dent J 1991;171:56-58.
20. Papathanasiou AG, Curzon MEJ, Fairpo CG. The influence of restorative material on the survival rate of restorations in primary molars. Pediatr Dent 1994;16:282-288.
21. Einwag J, Dunninger P. Stainless steel crown vs multisurface amalgam restorations: An 8-year longitudinal clinical study. Quintessence Int 1996;27:321-323.
22. Tate AR, Wai Ng M, Needleman HL, Acs G. Failure rates of restorative procedures following dental rehabilitation under general anesthesia. Pediatr Dent 2002;24:69-71.
23. Wai Ng M, Tate A, Needleman HL, Acs G. The influence of medical history on restorative procedure failure rates following dental rehabilitation. Pediatr Dent 2001;23:487-490.
24. Gruythuysen RJ, Weerheijm KL. Calcium hydroxide pulpotomy with a light-cured cavity-sealing material after two years. J Dent Child 1997;64:251-253.
25. Farooq NS, Coll J, Kuwabara A, Shelton P. Success rates of formocresol pulpotomy and indirect pulp therapy in the treatment of deep dentinal caries in primary teeth. Pediatr Dent 2000;22:278-286.
26. Holan G, Fuks AB, Keltz N. Success rate of formocresol pulpotomy in primary molars restored with stainless steel crowns vs amalgam. Pediatr Dent 2002;24:212-216.
27. Gibson GB, Richardson A, Patton RE, Waldman R. A clinical evaluation of occlusal composite and amalgam restorations: one and two year results. J Am Dent Assoc 1982;104:335-337.
28. Derkson GD, Richardson. A, Waldman R. Clinical evaluation of composite and amalgam posterior restoration: 1-year results. J Can Dent Assoc 1982;48:45-47.
29. Derkson GD, Richardson. A, Waldman R. Clinical evaluation of composite resin and amalgam posterior restorations: 2-year results. J Can Dent Assoc 1983;49:277-279.
30. Derkson GD, Richardson. A, Waldman R. Clinical evaluation of composite resin and amalgam posterior restorations: 3-year results. J Can Dent Assoc 1984;50:478-480.
31. Tonn EM, Ryge G. Two-year clinical evaluation of light cured composite resin restorations in primary molars. J Am Dent Assoc 1985;111:44-48.
32. Moher D, Schultz KF, Altman DG. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. Ann Intern Med 2001;134:657-662.
33. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. N Engl J Med 2000;342:1878-1886.