Acta Odontologica Scandinavica, 2009; 67: 206–211

informa
healthcare

ORIGINAL ARTICLE

Use of statistical methods in dental research: comparison of four dental journals during a 10-year period

HANNU VÄHÄNIKKILÄ[1], PENTTI NIEMINEN[2], JOUKO MIETTUNEN[3,4] & MARKKU LARMAS[1]

[1]*Institute of Dentistry and University Hospital, University of Oulu, Oulu, Finland,* [2]*Medical Informatics Group, University of Oulu, Oulu, Finland,* [3]*Department of Psychiatry, University of Oulu, Oulu, Finland and* [4]*Academy of Finland, Helsinki, Finland*

Abstract

Objective. To investigate development of the use of statistical methods in dental journals in 1996, 2001, and 2006 hypothesizing that methods are becoming more sophisticated. ***Material and methods.*** All original research articles in four dental journals in 1996, 2001, and 2006 were reviewed: *Journal of Dental Research (JDR), Caries Research (CR), Community Dentistry and Oral Epidemiology (CDOE) and Acta Odontologica Scandinavica (AOS)*. Evaluation covered 928 articles reporting original research findings based on systematic collection and statistical analysis of research data. Aspects measured were: research design, statistical methodology, description of procedures, and presentation of results. Percentage frequencies of reported statistical procedures were compared between journals and between years. ***Results.*** The main findings of the study are: 1) that use of multivariate or specific methods did not increase from 1996 to 2006, 2) that traditional statistical significance testing is still widely used in dental journals and that use of confidence intervals increased in the period 1996 to 2006, 3) that statistical procedures are usually extensively described, and 4) that dental journals have different profiles in their statistical content. ***Conclusions.*** The authors of articles in dental journals might apply these results when designing their dental research, i.e. to use statistical methods and to present results in line with the policy and presentation of the leading dental journals.

Key Words: *Articles, journals, research methodology, statistics*

Introduction

Over recent decades, there has been a great increase in the use of statistical methods in medical research in articles published in professional periodicals [1,2]. A trend toward the use of more sophisticated techniques can be observed, partly resulting from the availability of statistical software packages including new multivariate procedures. The use of statistics in medical journals has been the subject of considerable debate in recent years, and there is wide consensus that standards are generally low [3]. However, simple methods, such as chi-square tests or *t*-tests, are still important in dental research, although there is no systematic evaluation of this topic.

Surveys of original articles published in the *New England Journal of Medicine* (*NEJM*) 1978–1979, in 1989, and in 2004–2005, have revealed increasing sophistication of statistical methods over time [4–6]. This journal is one of the most visible scientific medical journals [impact factor (IF) = 51.296 (Journal Citation Report 2006, Institute of Scientific Information, Thomson Corp.)] and also one of the few scientific journals monitoring its own statistical level. The use of more sophisticated statistical methods may challenge clinicians' comprehension of new research and could slow dissemination of study results.

Strasak et al. [3] reported current practices in the use of statistics in medicine by contrasting the top journal in basic science (*Nature Medicine* (IF = 28.588)) with the top journal in clinical science (*NEJM*). They reported that a vast increase in the use and complexity of statistical methods could be observed for *NEJM* papers. This did not necessarily hold true for *Nature Medicine* papers, however, the

Correspondence: Hannu Vähänikkilä, Institute of Dentistry, P.O. Box 5281, FIN-90014 University of Oulu, Oulu, Finland. Tel: +358 8 537 5583. Fax: +358 8 537 5560. E-mail: hannu.vahanikkila@oulu.fi

(Received 11 November 2008; accepted 20 February 2009)

ISSN 0001-6357 print/ISSN 1502-3850 online
DOI: 10.1080/00016350902837922

results of their study indicating that basic science sticks with basic statistical methods. Horton et al. [6] reported that in more than half the articles in the *NEJM* in 2004 and 2005, relatively sophisticated statistical methods were used, such as survival analysis or multiple regressions. Their findings showed that there was a continued trend toward increased use of newer and more complex methods not typically included in introductory or second-level statistics courses in professional education at universities.

Statistical methods are important in dental research. However, there are only a few comprehensive studies concerning dental journals and their use of statistics. In evaluating the methodological qualities of reporting, Lesaffre et al. [7] concluded that split-mouth study design and analysis would benefit from development in statistical methods. The method of studying written communication by systematically measuring and analyzing research publications is bibliometrics [8]. Several medical journals have been reviewed for their statistical content [1,9]. Most authors have examined one single medical journal, a collection of sub-field oriented journals [3,10] or other samples of research articles. Thus far, there have not been many bibliometric studies in dental journals. One, however, was by Yang et al. [11] estimating the availability of statistical information within pediatric dentistry.

The aim of this present study was to analyze four dental journals and to evaluate what statistical methods were being used. Our purpose is not to judge whether the methods are used correctly, but to estimate the range of techniques reported and whether there is a trend toward the use of more sophisticated techniques. Comparisons are made between the journals in the years 1996, 2001, and 2006.

Articles were analyzed and compared, with attention focused on the following research hypotheses: 1. We assumed that there were differences in study planning (e.g. study design and sample size) of the articles between journals and also between publication years. 2. We assumed that only descriptive statistics were being used to summarize and determine the significance of the findings in basic dental research. 3. We assumed that the frequency of use of multivariate methods has increased from 1996 to 2006. 4. Because the journals approached dentistry from different aspects, we believed that there were differences between the four journals and over the 10-year span in the statistical content of published papers.

Material and methods

Set of articles

We limited our evaluation to articles published in four dental journals: *The Journal of Dental Research (JDR)*, *Caries Research (CR)*, *Community Dentistry and Oral Epidemiology (CDOE)* and *Acta Odontologica Scandinavica (AOS)*. *JDR* is the most visible and cited journal (IF = 3.475); CR has the highest impact factor of cariological journals (IF = 2.304); CDOE takes the epidemiological approach in dentistry (IF = 1.870); AOS is a European journal, mainly directed at and subscribed to in the Nordic countries (IF = 1.017), which lead in the use of preventive methods in dentistry.

The journals were scrutinized for original research articles published in 1996, 2001, and 2006. Editorials, letters, case reports, and review articles were excluded. The total number of reviewed articles reporting original studies based on statistical data analysis was 928, which was about 91% of all the articles in the four journals ($n = 1020$).

Definition of study design

Articles were classified within the following study design groups: non-experimental studies (including cross-sectional surveys, cohort studies, and case-control studies) and experimental and other studies (including reliability, methodology, and basic science studies). An experimental study consists of clinical trials and of studies on animals. An experimental study was defined when one or more variables were controlled by the investigator in order to monitor its effect on the process or outcome. A non-experimental study was defined when the investigator stood apart from the study arrangements. With slight modification we used the same classification as Nieminen et al. [12] in analyzing the psychiatric literature.

Categories of statistical procedures

All 928 papers were manually reviewed for their statistical content (Table I). Types and frequencies of statistical methods were systematically recorded and classified within 13 categories. Papers containing *t*-tests, simple contingency analysis, non-parametric methods, one-way ANOVA, and simple correlation techniques were considered as basic methods. Papers were categorized as using *p*-values if the results of statistical significance testing were reported. The use of confidence intervals and multiple comparisons was also recorded if these were reported. The use of statistical tables and figures was also assessed. Categories of statistical methods used, and a brief description of their content, are listed in Table I. We divided sample size into four groups. In the first group, sample size was under 30, in the second 30–99, in the third 100–300, and in the fourth over 300.

In addition, the following information was recorded: the extended description of procedures,

Table I. Categories of statistical procedures used to assess the statistical content of articles under study.

Category	Brief description
Basic methods	Descriptive statistics, comparison of means, cross-tabulation, non-parametric tests, and correlation
Multivariate methods	
Regression models	Linear, non-linear, logistic, Poisson and Cox's regression models
Other multivariate methods	Factor analysis, cluster analysis, log-linear models, discriminant analysis, etc.
Specific methods	
Reliability analysis	Percentage of agreement, sensitivity, specificity, kappa coefficient, intraclass correlation coefficients, ROC curve, Cronbach alpha
Epidemiological statistics	Relative risk, odds ratio, prevalence, incidence
Other specific methods	Survival analysis, effect size, etc.

references to statistical literature, and naming of the statistical software used.

Evaluation of articles and statistical analysis

An experienced biostatistician (P.N.) who has earlier used bibliometric methods in several studies [8,12] trained H.V. to classify the material in this study using the categorization described above. If there were problems interpreting a specific paper, its classification was appraised by the other authors. Reliability of the evaluation was checked by comparing the ratings of two reviewers, who independently reviewed 52 articles with 10 criteria; for example 'whether data analysis procedures were completely described in the methods part of the research report' or 'whether the statistical software used in the study was named in the report'. The kappa coefficient of reliability between raters ranged from 0.65 to 1.00. We used the chi-square test to evaluate differences in the use of statistical procedures between the four journals and the Cochran-Armitage test for trend between three years. SPSS for Windows version 15.0 (SPSS Inc., Chicago, Ill., USA) was used for the analysis.

Results

General design of analysis and sample size

General characteristics of the analyzed articles in different journals showed that only 27% were experimental studies (clinical trials or studies on animals); a clear majority were observational studies (Table II). The journals differed from each other in this respect: *JDR* had equal numbers of experimental and observational studies, *CR* had 25% experimental studies and 52% observational studies, while *AOS* and *CDOE* had only about 20% experimental and about 70% observational studies. There were significant differences in the sample sizes between the journals: in *JDR*, almost half of the articles had a sample size of under 30. *CR* and *AOS* published studies with small and large sample sizes equally. In *CDOE*, sample size in 69% of the articles was over 300.

There were more observational studies in 2006; however, differences between the years were not statistically significant in the trend test. The distribution of sample size was similar between 1996 and 2001, while in 2006 articles had smaller sample sizes ($p < 0.001$) (Table III).

Statistical methods and reporting

The statistical methods used in the journals are given in Table IV. In total, 37% of the articles reported only basic statistical methods, with the exception of *CDOE*, in which basic statistical methods were used in only 18%. Multivariate or specific methods were used in 39% of the articles, while in *CDOE* the figure was 72% of articles. *JDR* and *CR* used multiple regression methods more than did *AOS* and *CDOE*. Statistical significance testing was still a generally used method in all four journals: *p*-values were reported in 81% of all articles. Confidence intervals

Table II. Study design and sample size in the journals.

	Journal of Dental Research ($n=404$)	*Caries Research* ($n=193$)	*Acta Odontologica Scandinavica* ($n=159$)	*Community Dentistry and Oral Epidemiology* ($n=172$)	Total ($n=928$)	*p*-value
Study design						$p<0.001$
Experimental	132 (32.7%)	49 (25.4%)	31 (19.5%)	37 (21.5%)	249 (26.9%)	
Non-experimental	131 (32.4%)	101 (52.3%)	109 (68.5%)	124 (72.1%)	465 (50.1%)	
Other	141 (34.9%)	43 (22.3%)	19 (12.0%)	11 (6.4%)	214 (23.0%)	
Sample size						$p<0.001$
<30	134 (47.5%)	55 (32.0%)	27 (18.0%)	5 (2.9%)	221 (28.5%)	
30–99	63 (22.3%)	56 (32.6%)	41 (27.3%)	17 (9.9%)	177 (22.8%)	
100–300	37 (13.1%)	24 (13.9%)	31 (20.7%)	31 (18.1%)	123 (15.9%)	
>300	48 (17.0%)	37 (21.5%)	51 (34.0%)	118 (69.0%)	254 (32.8%)	

Table III. Study design and sample size used in different years.

	1996 ($n=310$)	2001 ($n=301$)	2006 ($n=317$)	p-value
Study design				$p=0.229$
Experimental	88 (28.4%)	97 (32.2%)	64 (20.2%)	
Non-experimental	147 (47.4%)	137 (45.5%)	181 (57.1%)	
Other	75 (24.2%)	67 (22.3%)	72 (22.7%)	
Sample size				$p<0.001$
<30	61 (24.3%)	61 (24.1%)	99 (36.5%)	
30–99	59 (23.5%)	51 (20.2%)	67 (24.7%)	
100–300	38 (15.1%)	46 (18.2%)	39 (14.4%)	
>300	93 (37.1%)	95 (37.5%)	66 (24.4%)	

were less frequently reported in *JDR*, and more frequently in *CDOE*. Research results were also often seen in graphic form, although there were large differences between the journals. The documentation of statistical methods and software used differed between the journals. An extended description of data analysis procedures was presented more frequently in *CDOE*, *CR* and *AOS* than in *JDR*. Software used was reported in 36% of the articles evaluated. A reference to statistical literature was at 50% of the evaluated articles in *CDOE*. In *CR* and *AOS*, statistical literature was referred to in 30% of the articles and in *JDR* in only 15%.

Statistical methods used in the period 1996 to 2006 are illustrated in Figure 1. In 1996, more articles used only basic statistical methods compared with 2001 and 2006. There were no statistically significant differences in the period between 1996 and 2006 in relation to the use of multivariate or specific methods. The use of multiple regression techniques significantly increased from 1996 to 2006, as did the use of p-values or significance tests and confidence intervals. Reporting results in tables was widespread, although this decreased from 1996 to 2006. From 1996 to 2006 the number of figures did not change in a statistically significantly manner. Both an extended description of procedures and of reporting software increased in the period. There were no differences between years (1996, 2001, and 2006) in references to statistical literature.

Discussion

Multivariate or specific methods

Because of the increasing availability of computer packages, there has been a trend towards use of more sophisticated statistical methods in many fields of medical science [8,13]. The results of this study indicate that these methods have not increased much in dental research. The reasons are not evident, but we believe that multivariate or specific methods may not have increased for the following reasons: first, that statistical knowledge is inadequate. Very few dental journals undergo statistical review and only a few publish their statistical guidelines for authors.

The second reason is that when calculating caries prevalence in epidemiological studies of dental health, the DMF index (decayed (D), missing (M), or filled (F)) has been used since 1938 [14]. When age cohorts have been compared with each other, the mean values of the DMF index have been used to describe dental caries, although this describes carious, restored, or extracted teeth, not patients with caries. More powerful and advanced statistical procedures, such as survival analysis methods, are seldom used in dental epidemiology.

Table IV. Statistical methods used in the journals.

	Journal of Dental Research ($n=404$)	*Caries Research* ($n=193$)	*Acta Odontologica Scandinavica* ($n=159$)	*Community Dentistry and Oral Epidemiology* ($n=172$)	Total ($n=928$)	p-value
Basic methods only	183 (45.3%)	67 (34.7%)	62 (39.0%)	31 (18.0%)	343 (37.0%)	<0.001
Multivariate or specific methods	105 (26.0%)	74 (38.3%)	62 (39.0%)	124 (72.1%)	365 (39.3%)	<0.001
Multiple comparisons	119 (29.5%)	55 (28.5%)	25 (15.7%)	15 (8.7%)	214 (23.1%)	<0.001
P-value	319 (79.0%)	162 (83.9%)	125 (78.6%)	141 (82.0%)	751 (80.9%)	=0.447
Confidence interval	54 (13.4%)	46 (23.8%)	39 (24.5%)	56 (32.6%)	195 (21.0%)	<0.001
Statistical figure	299 (74.0%)	120 (62.2%)	82 (51.6%)	67 (39.0%)	568 (61.2%)	<0.001
Statistical tables	267 (66.1%)	157 (81.3%)	146 (91.8%)	164 (95.3%)	734 (79.1%)	<0.001
Reference to statistical literature	60 (14.9%)	59 (30.6%)	50 (31.4%)	86 (50.0%)	255 (27.5%)	<0.001
Software reported	110 (27.2%)	89 (46.1%)	59 (37.1%)	76 (44.2%)	334 (36.0%)	<0.001
Extended description of procedures	229 (56.7%)	142 (73.6%)	110 (69.2%)	129 (75.0%)	610 (65.7%)	<0.001

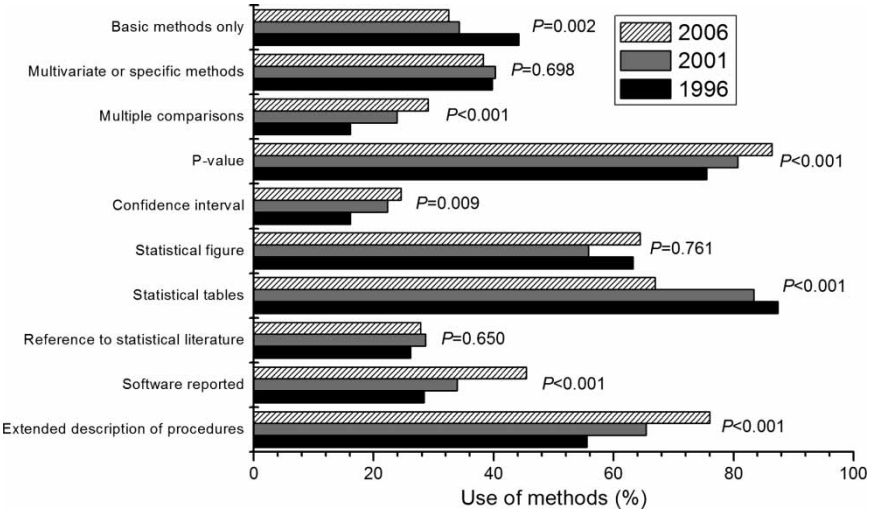

Figure 1. Statistical methods used in different years.

Caries data are usually collected employing the tooth or tooth surface as a unit, and the statistical methods applied aggregate data at the subject level to obtain a measure of the amount of disease at a certain age. The aggregated subject-based outcome measure is typically derived by summarizing the original tooth-based or surface-based observations, e.g. by means of calculating a DMF score per subject. As a result, important information may be lost because different tooth surfaces or teeth in the same mouth have different exposure times due to the great variation in their eruption times and different risks.

The third reason is that sample sizes have decreased from 1996 to 2006. Sophisticated statistical methods require larger sample sizes than basic methods. The reason for the decreasing sample size could be that studies on animals have been common and the costs much higher than, for example, cohort studies. It has also been noted that there were only a few articles with sample sizes calculated or justified. The lack of available large epidemiological data may also decrease the number of study subjects. The decreasing sample size is an alarming phenomenon in dental journals.

Statistical content of dental journals

We noted that dental journals had different profiles in their statistical analyses. For example, in *JDR*, multivariate or specific methods were used in 26% of articles, in *CR* and *AOS* in 39%, and in *CDOE* as much as in 72%. One explanation could be that statistical techniques differ between basic and clinical research, which can be seen when comparing statistical methods used between *NEJM* and *Nature Medicine* [3]. In studies on animals, experimental designs are used which include less intra-individual variation due to the use of genetically identical species. As a result, there is no need for the application of multivariate analysis to adjust for possible confounding factors, which is typical in clinical settings. For the same reason, sophisticated methods such as survival analysis, frequently observed in clinical research studies, are probably less likely to figure in basic research studies. Studies on animals have smaller sample sizes, which is one of the reasons for basic methods being used rather than more sophisticated techniques.

Conclusions

Our study reveals that the four journals selected here differ in their use of sophisticated multivariate methods and have diverging scientific fields of interest. This may explain our finding that these journals have different policies in statistical methodology. Our other important finding is that sample sizes have decreased over the years from 1996 to 2006.

This study has identified the commonly encountered statistical methods used in dental publications. With widespread use of different statistical techniques comes the need for better education on research methods and it is important that researchers, readers, and editors improve their statistical knowledge in order to understand the accuracy and validity of statistical results. The authors of articles in dental journals might apply these results and present their findings in line with policy and presentation of the leading journals. Authors who have knowledge of a variety of statistical methods will be better able to make important contributions to dental journals.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

[1] Altman DG. Statistical reviewing for medical journals. Stat Med 1998;17:2661–74.

[2] Altman DG. Statistics in medical journals: some recent trends. Stat Med 2000;19:3275–89.

[3] Strasak AM, Zaman Q, Marinell G, Pfeiffer KP, Ulmer H. The use of statistics in medical research: a comparison of the *New England Journal of Medicine and Nature Medicine*. Am Stat 2007;61:47–55.

[4] Emerson JD, Colditz GA. Use of statistical analysis in the *New England Journal of Medicine*. N Engl J Med 1983;309: 709–13.

[5] Emerson JD, Colditz GA. Use of statistical analysis in the *New England Journal of Medicine*. In: Bailar JC III, Mosteller F, editors. Medical uses of statistics, 2nd edn. Waltham, Mass.: NEJM Books; 1992. p. 45–57.

[6] Horton NJ, Switzer SS. Statistical methods in the journal. N Engl J Med 2005;353:1977–9.

[7] Lesaffre E, Garcia Zattera M-J, Redmond C, Huber H, Needleman I. Reported methodological quality of split-mouth studies. J Clin Periodontol 2007;34:756–61.

[8] Miettunen J, Nieminen P, Isohanni M. Statistical methodology in general psychiatric journals. Nord J Psychiatry 2002; 56:223–8.

[9] Walter SD. Methods of reporting statistical results from medical research studies. Am J Epidemiol 1995;141: 896–906.

[10] Rigby AS, Armstrong GK, Campbell MJ, Summerton N. A survey of statistics in three UK general practice journals. BMC Med Res Methodol 2004;4:28.

[11] Yang S, Needleman H, Niederman R. A bibliometric analysis of the pediatric dental literature in MEDLINE. Pediatr Dent 2001;23:415–8.

[12] Nieminen P, Rucker G, Miettunen J, Carpenter J, Schumacher M. Statistically significant papers in psychiatry were cited more often than others. J Clin Epidemiol 2007;60: 939–46.

[13] Wang Q, Zhang B. Research design and statistical methods in Chinese medical journals. J Am Med Assoc 1998;280: 283–5.

[14] Klein H, Palmer CE, Knutson JW. Dental status and dental needs of elementary school children. Public Health Reports 1938;53:751–5.