

Use of Statistical Procedures in Brazilian and International Dental Journals

Gláucia Maria Bovi AMBROSANO¹

André Figueiredo REIS²

Marcelo GIANNINI²

Antônio Carlos PEREIRA¹

¹Department of Community Dentistry

²Department of Restorative Dentistry – Operative Dentistry,
Piracicaba School of Dentistry, UNICAMP, Piracicaba, SP, Brazil

A descriptive survey was performed in order to assess the statistical content and quality of Brazilian and international dental journals, and compare their evolution throughout the last decades. The authors identified the reporting and accuracy of statistical techniques in 1000 papers published from 1970 to 2000 in seven dental journals: three Brazilian (*Brazilian Dental Journal*, *Revista de Odontologia da Universidade de São Paulo* and *Revista de Odontologia da UNESP*) and four international journals (*Journal of the American Dental Association*, *Journal of Dental Research*, *Caries Research* and *Journal of Periodontology*). Papers were divided into two time periods: from 1970 to 1989, and from 1990 to 2000. A slight increase in the number of articles that presented some form of statistical technique was noticed for Brazilian journals (from 61.0 to 66.7%), whereas for international journals, a significant increase was observed (65.8 to 92.6%). In addition, a decrease in the number of statistical errors was verified. The most commonly used statistical tests as well as the most frequent errors found in dental journals were assessed. Hopefully, this investigation will encourage dental educators to better plan the teaching of biostatistics, and to improve the statistical quality of submitted manuscripts.

Key Words: biostatistics, statistical testing, dental education, dental research.

INTRODUCTION

All dentists are challenged to keep abreast of advances in dental technology and basic scientific knowledge, which have increased dramatically in recent years. Experimental design and statistical analysis have become the major format in biomedical journals for providing answers to problems dentists face in daily practice (1). Statistical methods have become a core component of data analysis, manipulation, and expression in biomedical research (2). Successful practice requires that clinicians have skills that enable them to assess dental literature relevant to their work. Continued improvement in the quality of dental practice will depend heavily on the foundation of scientific literature on which clinical practice is based. Clinicians, dental students and researchers must have the ability to critically appraise and judge the validity of published re-

search, whereas researchers must know how to choose the statistical analysis designed for the data to be collected (3). We rely on statistics to distinguish significant associations from chance occurrences, whether the goal is to differentiate useful from useless techniques, assess risk factors for a disease, or critically evaluate new materials and drug performance (4).

The dental journal is the primary channel for disseminating dental information to the profession. The amount of information will continue to grow at an increasing rate. The practitioner who is determined to seek continuing education by reading dental journals is forced to choose those items on which he will focus and those he will ignore. One problem is that the statistical quality of some dental journals is questionable. The implications of inadequate statistical analysis are dramatic, with some treatments erroneously being “proven” to be effective, and with other promising treatments

Correspondence: Dra. Gláucia Maria Bovi Ambrosano, Departamento de Odontologia Social, Faculdade de Odontologia de Piracicaba, UNICAMP, Av Limeira 901, Areião, 13414-903 Piracicaba, SP, Brasil. Tel: +55-19-3412-5361; Fax: +55-1-3412-5218. e-mail: glaucia@fop.unicamp.br

being written off early in their development (5,6). Several surveys have been conducted in order to assess the statistical content of medical journals (2,4,7-8); however, little attention has been directed to dentistry and to the trends that have occurred in study design and statistical analysis.

The purpose of this study was to assess the most common statistical methods used in Brazilian and international dental research publications during the last decades, determining their accessibility to the reader, and recognizing potential sources of error.

MATERIAL AND METHODS

This study reviewed 1000 papers published from 1970 to 2000 in four international dental journals and three Brazilian journals: *Journal of the American Dental Association*, *Journal of Dental Research*, *Caries Research*, *Journal of Periodontology*, *Brazilian Dental Journal*, *Revista de Odontologia da Universidade de São Paulo* and *Revista de Odontologia da UNESP*, respectively.

A long period of time was chosen because the authors intended to show a long-term trend toward or away from the publication of statistically oriented reports in dental journals. The authors reviewed 310 papers published between 1970 and 1989 (118 Brazilian and 192 international), and 690 papers published between 1990 and 2000 (338 Brazilian and 352 international). Each of the reviewed articles was classified as to type of study. Articles were analyzed as to the presence or absence of statistical methods, if the statistical analysis was described, and, when described, what methodologies were applied. Statistical misuse or omission in the design or the analysis of each article were also recorded. Special attention was given to the types of statistical problems commonly seen in some medical literature reviews (9) namely inappropriate use of measures of variability, failure to use the correct statistical procedure, incomplete reporting of statistical methods, violation of the required assumptions of the statistical procedures chosen, failure to account for the use of multiple hypothesis testing, and mistakes in the use of graphics.

The effect that increasing statistical expertise would have on access to the dental literature was ascertained by determining the order of statistical methods that maximally increased the percentage of articles for

which readers would be acquainted with all of the statistical methods reported, if they learned one more method. The frequency with which each major test category was encountered was calculated. In order to analyze the papers accessibility to readers, the authors ordered the statistical techniques in a logical progression of learning for a reader without previous knowledge in biostatistics. The accumulated frequency of each method application in dentistry was calculated, reflecting the papers accessibility to readers with specific statistical methods knowledge.

RESULTS

Figure 1 depicts the distribution of the papers according to the type of study performed. Most of the articles were research papers. Of all articles published before 1990, 61% of the Brazilian (BP) and 66.7% of the international papers (IP) used some form of statistical technique. Among those which were statistically analyzed, 77.1% of BP and 87.9% of IP performed it correctly. The most frequent mistake was the use of *t* test for multiple comparisons when more than two groups were evaluated. Among BP published after 1990, 65.8% used statistics to quantitate their findings, compared with 92.6% found in IP. An increase to 82.3% of BP and to 89.9% of IP was noted for the

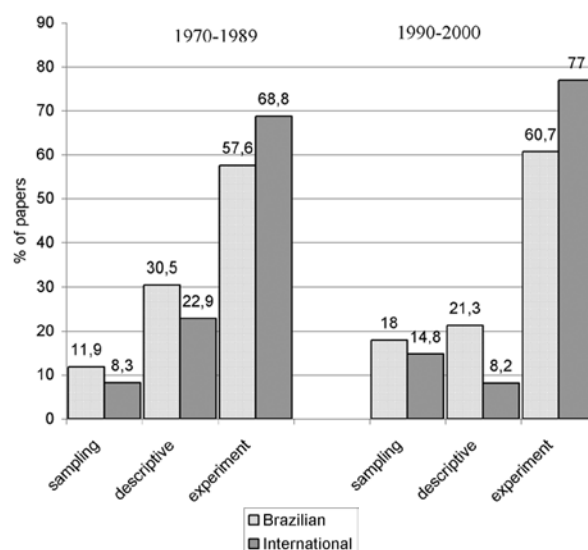


Figure 1. Distribution of Brazilian and international papers according to the type of study performed between 1970-1989 and 1990-2000.

correct use of statistical procedures for the type of variable and experiment design. The most common errors found not only for BP, but for IP as well, were: multiple comparison tests for quantitative variables (time, concentrations); split-plot experiments analyzed with only one error; experiments with more than two factors and significant interaction discussed as if no interaction was present; presentation of mean standard error as dispersion of a unique sample; presentation of standard deviation for ordinal data; and use of *t* test in experiments with more than two groups. It was also observed that, in many papers, no central tendency measure was presented. In addition, in numerous articles, only the statistical significance was discussed, while the clinical relevance of results was not informed.

From 1970 to 1989, only 22.0% of BP and 49.3% of IP used graphics to present results. After 1990, this number increased to 61.3% in IP, while in BP, this frequency remained low (24.6%). Mistakes were noted in 11.1% of the graphics presented in BP, and in 1.4% of IP before 1990. After this year, this percentage increased (33.3% BP and 14.3% IP), however, no significant correlation was found between publication year and mistake rates ($p > 0.05$). The most common mistakes were: use of too many abbreviations with no legends, absence of title on axis, no specifica-

tion of the presented variables, graphic bars for quantitative variables and line graphics for qualitative variables, and confusion between graphic bars and histogram.

Before 1990, most of the papers used only one kind of statistical procedure (67.4% BP and 64.3% IP). Among them, 84.1% of BP and 78.4% of IP used parametric tests, 15.9% of BP and 16.5% of IP used non-parametric tests, and 5.2% of IP used both. An increase in the use of both types of tests concomitantly was noted during the last decade.

The frequency and type of dispersion measurement used during the two time periods are presented in Table 1. Standard deviation (SD) was the most frequently used measure of variability. The coefficient of variation was only used in a few papers (6.10% in BP and 1.35% in IP until 1989, and 1.51% in BP and 1.24% in IP after 1990).

The frequency with which each major test was encountered and paper accessibility to readers of BP and IP are presented in Tables 2 and 3, respectively. The most frequently used statistical procedures in BP were: analysis of variance (ANOVA), Tukey test, frequency distribution and/or descriptive statistics, Kruskal-Wallis, *t* test, Pearson's correlation and χ^2 . Whereas in IP, the most commonly used procedures were: analysis

of variance (ANOVA), *t* test, paired *t* test, Pearson's correlation, frequency distribution and/or descriptive statistics, Wilcoxon and χ^2 , respectively.

Table 1. Distribution of papers (%) as function of dispersion measure presented.

Dispersion measure	Brazilian	International	Brazilian	International
	1970-1989		1990-2000	
No dispersion measure	79.27	56.76	77.8	35.5
Amplitude	0.00	0.00	0.38	0.00
Coefficient of variation	3.66	1.35	1.13	0.93
Coefficient of variation in the graphic	0.00	0.00	0.00	0.31
Coefficient of variation and Standard deviation	2.44	0.00	0.38	0.00
Standard deviation	8.54	37.16	17.29	39.2
Standard deviation and Standard Error	0.00	0.00	0.00	0.30
Confidence interval	1.22	1.35	0.00	0.93
Confidence interval and Standard deviation	0.00	0.00	0.38	0.00
Standard deviation in the graphic	2.44	0.00	0.38	20.98
Mean standard error	2.43	3.38	2.26	1.85

DISCUSSION

The authors identified the frequency of use of statistical techniques during the last decades. This survey consisted of journals thought to be commonly available to most dental practitioners, researchers and students, and, therefore, results reflect general statistical usage in dental research. This review provides a clear guideline of what should be taught to enhance understanding of the dental literature, a feature essential in continuing dental education.

Results showed that if all those who write and read dental literature

are up-to-date with some specific statistical procedures, a great extent of dental literature can be fully understood. It was verified that readers with basic knowledge of biostatistics (statistical techniques written in bold, Tables 2 and 3) would be able to understand 38.46% of BP and 43.47% of IP published after 1990, respectively, while readers with knowledge of ANOVA, Tukey and Duncan multiple comparison tests would be able to understand 82.05% of BP and 70.15% of IP, respectively. The most commonly used statistical tests were analysis of variance (ANOVA), Tukey test, frequency distribution and/or descriptive statistics, Kruskal-Wallis, Wilcoxon, *t* test, Pearson's correlation and χ^2 (Tables 2 and 3). This is valuable information for dental educators, clinicians and researchers, because with

knowledge of these techniques a large spectrum of dental literature may be interpreted and critically evaluated. A larger extent of BP would be accessible to the reader with knowledge of a few specific statistical procedures, reflecting that IP have been using more diversified methodologies. A trend toward the use of different methodologies was noted when both periods of time were compared. With knowledge of the same statistical procedures (ANOVA, Tukey and Duncan tests), a reader would have access to 94.81% of BP and 86.82% of IP before 1990. This evolution may be attributed to the widespread use of statistical software packages during the last decade, which facilitate the use of more complex methodologies (10). A consequence of this trend is that a greater level of statistical expertise

Table 2. Distribution of statistical procedures used in Brazilian journals and accessibility to readers.

Statistical method	Frequency					
	1970-1989			1990-2000		
	N	%	Accumulated %	N	%	Accumulated %
Not informed	15	19.48	19.48	9	2.88	2.88
Frequency distribution or descriptive statistics	4	5.19	24.68	47	15.06	17.95
<i>t</i> test	9	11.69	36.36	21	6.73	24.68
Paired <i>t</i> test	0	0	36.36	3	0.96	25.64
χ^2	4	5.19	41.56	14	4.49	30.13
Fisher	0	0	41.56	4	1.28	31.41
Pearson's correlation	2	2.6	44.16	16	5.13	36.54
Linear regression	0	0	44.16	6	1.92	38.46
ANOVA	26	33.77	77.92	83	26.60	65.06
Tukey	13	16.88	94.81	49	15.71	80.77
Duncan	0	0	94.81	4	1.28	82.05
Scheffé	0	0	94.81	1	0.32	82.37
Student-Newman-Keuls	0	0	94.81	1	0.32	82.69
Tukey-Kramer	0	0	94.81	8	2.56	85.26
Wilcoxon	0	0	94.81	6	1.92	87.18
Mann Whitney	0	0	94.81	8	2.56	89.74
Kruskal-Wallis	1	1.3	96.1	21	6.73	96.47
Friedman	1	1.3	97.4	1	0.32	96.79
Dunn	1	1.3	98.7	3	0.96	97.76
Miller	0	0	98.7	1	0.32	98.08
Mac Nemar	0	0	98.7	1	0.32	98.40
Kappa	1	1.3	100	1	0.32	98.72
Relative risk	0	0	100	1	0.32	99.04
Odds ratio	0	0	100	2	0.64	99.68
Logistic regression	0	0	100	1	0.32	100

Basic statistical methods are written in bold.

is expected of writers and readers of dental literature than some years ago.

Worthy of note is the gradual way in which increased knowledge of statistical techniques adds to the percentage of statistical uses understood and to the percentage of articles in which the reader has access to all the statistical methods used. It can be predicted that

knowledge of a few common statistical procedures would make a great part of the articles statistically accessible to the reader.

Publication of the results of a clinical investigation or laboratory finding may rapidly affect both the practice and research directions of dentistry. Some clinicians probably do not realize that publishing incor-

Table 3. Distribution of statistical procedures used in international journals and accessibility to readers.

Statistical method	Frequency					
	1970-1989			1990-2000		
	N	%	Accumulated %	N	%	Accumulated %
Not informed	2	1.55	1.55	11	2.17	2.17
Frequency distribution or descriptive statistics	15	11.63	13.18	32	6.32	8.49
t test	26	20.16	33.33	53	10.47	18.97
Paired t test	6	4.65	37.98	36	7.11	26.08
χ^2	10	7.75	45.74	29	5.73	31.81
Fisher	4	3.10	48.84	7	1.38	33.20
Pearson's correlation	9	6.98	55.81	32	6.32	39.52
Linear regression	13	10.08	65.89	20	3.95	43.47
ANOVA	24	18.60	84.50	107	21.15	64.62
Tukey	0	0.00	84.50	18	3.56	68.18
Duncan	3	2.33	86.82	10	1.98	70.15
Scheffé	3	2.33	89.15	11	2.17	72.33
Student-Newman-Keuls	4	3.10	92.25	7	1.38	73.71
Tukey-Kramer	0	0.00	92.25	2	0.40	74.11
Bonferroni	0	0.00	92.25	3	0.59	74.70
Dunnett	0	0.00	92.25	1	0.20	74.90
Wilcoxon	4	3.10	95.35	32	6.32	81.22
Mann Whitney	4	3.10	98.45	21	4.15	85.37
Kruskal-Wallis	1	0.78	99.22	11	2.17	87.55
Friedman	0	0.00	99.22	3	0.59	88.14
Dunn	0	0.00	99.22	1	0.20	88.34
Miller	0	0.00	99.22	0	0.00	88.34
Spearman's Correlation	0	0.00	99.22	9	1.78	90.11
Mac Nemar	0	0.00	99.22	2	0.40	90.51
Kappa	1	0.78	100	7	1.38	91.89
Relative Risk	0	0.00	100	1	0.20	92.09
Odds ratio	0	0.00	100	6	1.19	93.28
Specificity and Sensitivity	0	0.00	100	4	0.79	94.07
Logistic Regression	0	0.00	100	6	1.19	95.25
Analysis of covariance	0	0.00	100	5	0.99	96.24
ROC (Receiver Operating Curve)	0	0.00	100	1	0.20	96.44
Multivariate analysis	0	0.00	100	8	1.58	98.02
Multiple regression	0	0.00	100	10	1.98	100

Basic statistical methods are written in bold.

rect or misleading results is unethical by exposing patients to unjustified inconvenience, by wasting resources and time, and by carrying out unnecessary further works (11). When a published study fails to clearly report its research findings, a reader attempting to evaluate an article's scientific validity must largely rely on the authors' reputations or writing styles (12). A decrease in the number of problems present in statistical reporting of BP and IP was noticed. However, the number of BP that used some form of statistical procedures remains quite similar throughout the last decades (61% before 1990 and 65.8% after 1990). IP presented a notable increase in the number of articles that were statistically analyzed (from 66.7% to 92.6%), probably due to an increase in statistical training programs for dental researchers and students.

Austin and Attanasio (13) raised the concern that a common problem in scientific articles is the misuse of statistical analysis methods resulting in unjustified conclusions. This problem occurs in all professions, not only dentistry. Many times, the experiment is begun and collection of data is performed before enough thought has been given to the study design and to the method in which data will be statistically analyzed. Many researchers do not realize which test is appropriate and will go to computer statistical software for analysis. However, the computer will produce impressively prepared tables, charts, and results without regard to appropriateness (13). More than 10% of the papers presented inappropriate statistical analysis.

An interesting point is that a trend towards the use of parametric tests to evaluate ordinal data was noticed in this study. However, this "problem" must be dealt with caution because, according to Cohen (14), the use of parametric methods is valid, more powerful and more versatile than non-parametric tests for the analysis of dental indices, even those that are not normally distributed. Among the non-parametric tests, Kruskal-Wallis was the most used.

Substantial reduction could be made in the error rate if some suggestions were followed (9,15). In view of our results, we recommend that editorial boards standardize their format for reporting statistical data in their instructions for authors. A manuscript should describe statistical analyses in sufficient detail so a reader can reproduce calculations if data are available. For less commonly used statistical methods, a journal or textbook reference should be added. In addition,

every test of statistical significance should be accompanied by identification of technique and rationale. The clarification of the use of statistical methods has the potential of greatly enhancing the strength of a submitted manuscript. Finally, the statistician should participate from the beginning to the end of the study. It would be unfair to be overly critical of the papers cited as examples of errors since they often were the most clearly written and, furthermore, the error sometimes did not change the overall conclusion of the study. Despite some weakness, hypothesis testing remains today a useful approach for analyzing the results of a dental study; however, the purpose can only be reached by the proper use of tests, which means, on the one hand, the choice of a valid method for the given problem, and on the other, a correct interpretation of the test result (5).

The authors encourage all educators in dentistry to review their programs to ensure that an appropriate level of statistical prowess is provided to their students, residents and fellows. The authors also encourage current practitioners who wish to be more effective consumers of dental literature to review their own statistical skills, improving their familiarity with statistical techniques. Hopefully, it will encourage dental educators to better plan biostatistics teaching, and improve the quality of submitted manuscripts.

RESUMO

O objetivo deste estudo foi avaliar a qualidade e o conteúdo estatístico de periódicos odontológicos nacionais e internacionais, e comparar a sua evolução nas últimas décadas. Os autores verificaram a descrição e coerência das análises estatísticas em 1000 artigos publicados de 1970 a 2000 em sete periódicos odontológicos: três nacionais (*Brazilian Dental Journal*, *Pesquisa Odontológica Brasileira* e *Revista de Odontologia da UNESP*) e quatro internacionais (*Journal of the American Dental Association*, *Journal of Dental Research*, *Caries Research* e *Journal of Periodontology*). Os trabalhos foram divididos em dois períodos de publicação: de 1970 a 1989, e de 1990 a 2000. Foi verificado um ligeiro aumento no número de artigos que apresentaram alguma análise estatística em periódicos nacionais (de 61,0 a 66,7%), enquanto para os periódicos internacionais, um aumento notável foi observado (de 65,8 a 92,6%). Além disso, uma diminuição no número de erros nas análises estatísticas foi verificado. Os métodos estatísticos mais utilizados, assim como os erros mais frequentes foram analisados. Espera-se que este trabalho encoraje os educadores na área de Odontologia a planejar melhor o ensino de bioestatística, e a melhorar a qualidade estatística dos artigos submetidos para publicação.

ACKNOWLEDGEMENTS

This study was supported by CNPq (grants #463.230/00-0 and #300.785/94-1).

REFERENCES

1. Juzych MS, Shin DH, Seyedsadr M, Siegner SW, Juzych L. Statistical techniques in ophthalmic journals. *Arch Ophthalmol* 1992;110:1225-1229.
2. Schwartz SJ, Sturr M, Goldberg G. Statistical methods in rehabilitation literature: a survey of recent publications. *Arch Phys Med Rehabil* 1996;77:497-500.
3. Lorton L, Rethman MP. Statistics: curse of the writing class. *J Endodon* 1990;16:13-18.
4. Rosenfeld RM, Rockette HE. Biostatistics in otolaryngology journals. *Arch Otolaryngol Head Neck Surg* 1991;117:1172-1176.
5. Jamart J. Statistical tests in medical research. *Acta Oncol* 1992;37:723-727.
6. Murray GD. The task of a statistical referee. *Br J Surg* 1988;75:664-667.
7. Morris RW. A statistical study of papers in the *Journal of Bone and Joint Surgery [BR]* 1984. *J Bone Joint Surg [Br]* 1988;70-B:242-246.
8. Menegazzi JJ, Yearly DM, Harris JS. Methods of data analysis in the emergency medicine literature. *Am J Emerg Med* 1991;9:225-227.
9. Cruess DF (1989). Review of use of statistics in the *American Journal of Tropical Medicine and Hygiene* for January-December 1988. *Am J Trop Med Hyg* 1989;41:619-626.
10. Becker PJ, Vilijoen E, Wolmarans L, Ijsselmuiden CB. An assessment of the statistical procedures used in original papers published in the *SAMJ* during 1992. *S Afr Med J* 1995;85:881-884.
11. Altman DG, Gore SM, Gardner MJ, Pocock SJ. Statistical and ethics in medical research. Misuse of statistics is unethical. *Br Med J* 1980;281:1182-1184.
12. Hokanson JA, Bryant SG, Gardner Jr R, Luttman DJ, Guernsey BG, Bienkowski AC. Spectrum and frequency of use of statistical techniques in psychiatric journals. *Am J Psychiatry* 1986;143:1118-1125.
13. Austin D, Attanasio R. The implications and application of biostatistical analysis in craniomandibular and orofacial pain disorders. *J Cran Pract* 1991;9:280-285.
14. Cohen ME. Analysis of ordinal dental data: evaluation of conflicting recommendations. *J Dent Res* 2001;80:309-313.
15. Hokanson JA, Stiernberg CM, McCracken MS, Quinn Jr FB. The reporting of statistical techniques in otolaryngology journals. *Arch Otolaryngol Head Neck Surg* 1987;113:45-50.

Accepted February 18, 2003