available at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/burns

Descriptive and inferential statistical methods used in burns research

Sammy Al-Benna [a,*], *Yazan Al-Ajam* [b], *Benjamin Way* [b], *Lars Steinstraesser* [a]

[a] *Department of Plastic Surgery and Burn Centre, BG University Hospital Bergmannsheil, Ruhr University Bochum, Buerkle-de-la Camp Platz 1, 44789 Bochum, Germany*
[b] *Department of Plastic and Reconstructive Surgery, Royal Free Hospital, London, United Kingdom*

ARTICLE INFO

Article history:
Accepted 21 April 2009

Keywords:
Statistics
Research report
Burns
Journals

ABSTRACT

Background: Burns research articles utilise a variety of descriptive and inferential methods to present and analyse data. The aim of this study was to determine the descriptive methods (e.g. mean, median, SD, range, etc.) and survey the use of inferential methods (statistical tests) used in articles in the journal *Burns*.
Methods: This study defined its population as all original articles published in the journal *Burns* in 2007. Letters to the editor, brief reports, reviews, and case reports were excluded. Study characteristics, use of descriptive statistics and the number and types of statistical methods employed were evaluated.
Results: Of the 51 articles analysed, 11(22%) were randomised controlled trials, 18(35%) were cohort studies, 11(22%) were case control studies and 11(22%) were case series. The study design and objectives were defined in all articles. All articles made use of continuous and descriptive data. Inferential statistics were used in 49(96%) articles. Data dispersion was calculated by standard deviation in 30(59%). Standard error of the mean was quoted in 19(37%). The statistical software product was named in 33(65%). Of the 49 articles that used inferential statistics, the tests were named in 47(96%). The 6 most common tests used (Student's *t*-test (53%), analysis of variance/co-variance (33%), χ^2 test (27%), Wilcoxon & Mann–Whitney tests (22%), Fisher's exact test (12%)) accounted for the majority (72%) of statistical methods employed. A specified significance level was named in 43(88%) and the exact significance levels were reported in 28(57%).
Conclusion: Descriptive analysis and basic statistical techniques account for most of the statistical tests reported. This information should prove useful in deciding which tests should be emphasised in educating burn care professionals. These results highlight the need for burn care professionals to have a sound understanding of basic statistics, which is crucial in interpreting and reporting data. Advice should be sought from professionals in the fields of biostatistics and epidemiology when using more advanced statistical techniques.

1. Introduction

Burns research has become increasingly sophisticated over recent years. It is becoming difficult to critically evaluate burns literature without a basic knowledge of statistics and experimental design. The advent of easy to use and powerful statistical packages and the increasing demands from burns journals for comprehensive statistical analysis of the literature has seen an

* *Corresponding author*. Tel.: +49 234 302 6135.
E-mail address: sammyalbenna@doctors.org.uk (S. Al-Benna).
0305-4179/$36.00
doi:10.1016/j.burns.2009.04.030

explosion in the use and reporting of biostatistics. A sound knowledge in study design and biostatistics has therefore become crucial in the ability of a health care professional to critically appraise medical literature. Undergraduate and graduate training allows little time for the study of these topics and a firm grasp of the subject still eludes many health care professionals, and as a consequence many are unable to detect errors in the literature they peruse and lack confidence in experimental design [1,2]. Although a part of most undergraduate teaching, biostatistics is seldom revisited in post-graduate curricula and with the advent of shortened training and increasing time constraints, biostatistics has taken a back seat in post-graduate training. In addition, healthcare professionals in training may perceive that statistics are conceptually difficult and view the task of developing sound knowledge in the subject as a great challenge. Burn care professionals are also faced with the additional problem of working variable shifts, adding to the difficulty in taking courses in biostatistics and research design. As a result, burn care professionals often perceive that the task of developing a statistical knowledge base as overwhelming.

The statistical guidelines in the Uniform Requirements for Manuscripts Submitted to Biomedical Journals (URMBJ) begin with: ''*Describe statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results* [3]''.

Description and reporting of statistical methods has been described in various medical specialties, but not in burns [4–9]. Burns research articles utilise a variety of descriptive and inferential methods to present and analyse data. The aim of this study was to determine the descriptive methods and survey the inferential statistics used in articles in the journal *Burns* in 2007. From these data, implications for authors, readers and, crucially, medical educators, of the results of the findings are discussed and recommendations for the design of statistical curriculum are suggested based on the identification of a manageable core of techniques to account for the majority of data analysis methods used.

2. Methods

The study defined its population as all original articles published in the journal *Burns* in 2007. Original research articles were defined as those reporting on studies that included primary data collection. Therefore, case reports, burn care in practice, letters to the editor, literature review papers, editorials, personal reports, analyses of secondary data, and theoretical articles without data were not considered as eligible. Study characteristics, objectives and design, type of statistical software employed, use of descriptive statistics and the number and types of statistical methods employed were evaluated. Frequency distributions for each statistical test, together with cumulative distributions, were tabulated. The appropriateness of statistical methods based on study design was not assessed.

3. Results

51 articles met the inclusion criteria. Of the 51 articles analysed, 11(22%) were randomised controlled trials, 18(35%) were cohort studies, 11(22%) were case control studies and 11(22%) were case series.

Table 1 – Relative and cumulative distributions of inferential statistics from the 49 articles in the journal *Burns* in 2007. Note that the 'Percent Articles' add up to more than 100 and the 'Number of Articles' exceeds 49 as some articles employed more than one statistical method. Values in italics are subtotals for specific statistical methods.

Method	Number of articles	Percent articles	Cumulative percent
Student's *t*-test	26	53	26
Contingency tables		39	
χ^2	13	(27)	39
Fisher's exact test	6	(12)	45
ANOVA/ANCOVA	16	33	61
Nonparametric tests		27	
Mann–Whitney	8	(16)	59
Wilcoxon	3	(6)	72
Kruskal–Wallis	2	(4)	74
Confidence intervals	9	18	83
Multiple comparison		20	
Bonferroni	4	(8)	88
Turkey–Kramer	2	(4)	90
Tamhane	2	(4)	92
Dunnett	2	(4)	94
Regression techniques		8	
Multiple regression	4	(8)	98
Correlation techniques		4	
Pearson product-moment	2	(4)	100

The study objectives and study design was described in all articles. All papers used descriptive statistics and all made use of continuous data. Inferential statistics were used in 49(96%) articles. Data dispersion was calculated by standard deviation in 30(59%) and standard error of the mean was quoted in 19(37%). The statistical software product was named in 33(65%). Of the 49 articles that made use of inferential statistics, a specified significance level was named in 43(88%) with the exact significance levels reported in 28(57%). The tests were named in 47(96%) articles. The 6 most common tests used were Student's *t*-test (53%), analysis of variance/co-variance (ANOVA/ANCOVA) (33%), χ^2 test (27%), Wilcoxon & Mann–Whitney tests (22%) and Fisher's exact test (12%). The relative and cumulative distributions of statistical techniques employed are summarized in Table 1.

4. Discussion

Burns research has become increasingly sophisticated over the years, and a sound basic understanding of statistics and experimental design has become essential in critically evaluating medical literature. A reader who is conversant with just descriptive statistics (percentages, means and standard deviations of means, medians, ranges, etc.) will have very limited statistical access to any of the articles; knowledge of only 6 major statistical methods will allow readers to understand over 70% of all statistical tests used in the journal *Burns*. Burn care professionals should have sound knowledge of basic descriptive statistics, including description of averages, and should be familiar with the common pitfall of using the standard error of the mean in place of standard deviation to erroneously describe data dispersion around a sample mean. Whereas standard deviation describes the variability between individuals in a sample, standard error of the mean describes the uncertainty of how the sample mean represents the population mean [10]. 37% of articles used standard error of the mean in place of standard deviation.

Of the 49 articles that used inferential statistics, exact significance levels were reported in only 28(57%), the rest quoting the *p* value as either greater or less than 0.05, or simply as 'non-significant'. A specified significance level (in all cases set at $p < 0.05$) was named in 43(88%) of cases; in the remaining articles, it was assumed that it was set at the 5% level. Journals have different styles of presenting statistical material. There remains no consensus on whether to use P, p, *P* or *p* values. While not of major importance, a standard notation would be desirable. Most journals continue to use the '$\pm$' notation to connect means and either standard deviations or standard errors despite frequent ambiguity and the misleading implications of such notation.

Student's *t*-test was the most commonly used statistical method, used in 53% of articles. Our findings are in keeping with previous reports in the medical literature [7,11]. Surprisingly, confidence intervals (CIs) were only reported in 18% of articles. CIs aid in interpretation of data by placing upper and lower limits on the likely size of any true effect [12–16]. This is echoed in the recommendations of the International Committee of Medical Journal Editors (ICMJE), whose guidelines for statistical reporting advises "*When possible, quantify findings and present them with appropriate indicators of measurement error or uncertainty (such as confidence intervals)*", and "*Avoid sole reliance on statistical hypothesis testing, such as the use of p values, which fails to convey important quantitative information* [17]". It appears that there is a tendency to focus on the means generated within a single study rather than paying heed to the breadth of values the mean could reasonably have assumed if the study were repeated [12–16]. In addition, the ICMJE recommends specifying the computer software used [17]. This was reported in only 65% of articles.

As ordinal data is often gathered during clinical burns research (for example, visual analog pain scales), nonparametric tests were used in 27% of original articles. Regression techniques, used in 8% of the articles are useful for predicting the outcome of one variable based on data obtained through another variable. This is also a technique that is often of interest to the burn care professionals in trying to decide which factors might influence outcome.

This study did have limitations. First, it represents only 1 year of the journal *Burns* in 2007 and may not be representative. However, as it included all original research in that recent year, it is likely to adequately represent the current quality of research reports in the burns literature. Second, this evaluation did not include burns articles published outside the journal *Burns*, which may or may not have higher quality research reports. An additional consideration is that other evaluations of research reports have typically been restricted to clinical research, whereas our study included basic science, clinical and educational research. It is also worth noting that of the 6 major statistical tests used, most of them, including one-way ANOVA, are quite simple. However, ANOVA is a generic term that covers a huge variety of statistical models that can be far more complex. This has implications when designing statistical curricula.

The Uniform Requirements for Manuscripts Submitted to Biomedical Journals is a widely available document and is recommended for authors submitting to journals [17]. It is not clear if authors do not look at these requirements or do not understand the requirements. Adhering to other items in the statistical guidelines would also help authors increase the quality of data reporting. These include "Put a general description of methods in the Methods section. When data are summarized in the Results section, specify the statistical methods used to analyze them. Restrict tables and figures to those needed to explain the argument of the paper and to assess its support ... do not duplicate data in graphs and tables. Avoid non-technical uses of technical terms in statistics, such as "random", ..., "normal," "significant," "correlations," and "sample". Define statistical terms, abbreviations, and most symbols."

In view of the above findings, we recommend that authors and readers should have a sound knowledge of statistical methods in order to improve their ability to design studies and critically evaluate medical literature. The findings are equally useful for educators in designing a biostatistical curriculum for burn care professionals in training. Medical education has come to play a central role in resident training since the radical reform of the post-graduate medical system in the United Kingdom in 2005 [18]. With the introduction of the Intercollegiate Surgical Curriculum Project, knowledge of basic concepts in statistics and study design has become part of the curriculum. Surgical trainees are expected to have a broad and specific understanding of statistical significance and confidence intervals, and to know applications of parametric, nonparametric, multivariate and χ^2 analysis [19].

A recent study looking at residents' understanding of the biostatistics and results in medical literature in 11 different residency programs in the United States concluded that most residents in that study lacked the knowledge in biostatistics needed to interpret many of the results in published clinical research [1]. Our empirical experience with residents in Germany and the United Kingdom suggest similar findings.

Improving the statistical understanding of health care researchers is high on statisticians' agenda as the shortage of statisticians in medical schools worldwide is a major worry [20,21]. There are strong arguments for increasing the number of statisticians in medical research [22]. They would not eliminate statistical errors, but their direct and indirect influence should certainly be of major benefit to the quality of medical research and thus the quality of published papers.

While familiarity with these few tests will be sufficient for most health care professionals in training, junior faculty, research fellows and faculty will likely need more extensive training in experimental design. Health care professionals also need to familiarise themselves with additional techniques more commonly used in epidemiological studies and reports generated by quality assurance investigations.

Information from the findings of this report has been used to draw up an outline of a potential burns care statistical

Table 2 – Possible statistical curriculum.

Lecture	Topic
1	Basic overview of statistics and study design
2	Descriptive statistics
3	Inferential statistics: parametric and nonparametric
4	multivariate, correlation and regression, Student's
5	t-test and χ^2 analysis
6	Suitability of test for specific data
7	Common pitfalls and errors
8	Misuse of statistics
9	Exercises in applied biostatistics
10	Examination

curriculum (Table 2). The curriculum can be based on a series of short lectures, with examples drawn from articles in the medical literature to illustrate the use of various statistical methods. Crucial to the series of lectures is the inclusion of common pitfalls, errors and misuse of statistical techniques. An examination at the end of the course can be used to assess the knowledge gained and to refine the structure of future courses.

By developing a basic knowledge of these 6 statistical tests, burn care professionals can greatly improve their ability to critically evaluate burns literature. Information provided by this study should prove useful when deciding which tests should be emphasized in educating burn care professionals. In an environment where evidence based medicine is playing an increasingly important role in clinical practice, a sound understanding of biostatistics in order to critically appraise burns literature is imperative. Fewer than 10 key statistical techniques accounted for almost 90% of the statistical techniques reported in the journal *Burns* in 2007. The authors call for the introduction of a burns statistical curriculum and suggest an outline of subjects to be included in light of the findings of this study. Finally advice should be sought from professionals in the fields of biostatistics and study design when using more advanced statistical methods.

Conflict of interest statement

None of the authors has any financial and personal relationships with other people or organisations that could inappropriately influence (bias) this work.

Acknowledgements

Contributors: Sammy Al-Benna: conception and design of the study, acquisition of data, analysis and interpretation of data, drafted the paper, final approval of the version to be submitted. Yazan Al-Ajam: acquisition of data, interpretation of data, revising it critically for important intellectual content, final approval of the version to be submitted. Sammy Al-Benna and Yazan Al-Ajam are joint first authors. Benjamin Way: acquisition of data, interpretation of data, revising it critically for important intellectual content, final approval of the version to be submitted. Lars Steinstraesser: acquisition of data, analysis and interpretation of data, revising it critically for important intellectual content, final approval of the version to be submitted.

REFERENCES

[1] Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. JAMA 2007;298:1010–22.
[2] West CP, Ficalora RD. Clinician attitudes toward biostatistics. Mayo Clin Proc 2007;82:939–43.
[3] International Committee of Medical Journal Editors. Uniform requirements for manuscripts submitted to biomedical journals. Ann Intern Med 1997; 126:36–47.
[4] Petrie A. Statistics in orthopaedic papers. J Bone Joint Surg 2006;88:1121–36.
[5] Rigby A, Armstrong G, Campbell M, Summerton N. A survey of statistics in three UK general practice journal. BMC Med Res Methodol 2004;4:28.
[6] Huang W, LaBerge J, Lu Y, Glidden D. Research publications in vascular and interventional radiology: research topics, study designs, and statistical methods. J Vasc Interv Radiol 2002;13:247–55.
[7] Menegazzi J, Yealy D, Harris J. Methods of data analysis in the emergency medicine literature. Am J Emerg Med 1991;9:225–7.
[8] Cruess DF. Review of use of statistics in The American Journal of Tropical Medicine and Hygiene for January–December 1988. Am J Trop Med Hyg 1989;41:619–26.
[9] 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.
[10] Nagele P. Misuse of standard error of the mean (SEM) when reporting variability of a sample. A critical evaluation of four anaesthesia journals. Br J Anaesth 2001;90:514–6.
[11] Emerson JD, Colditz GA. Use of statistical analysis in The New England Journal of Medicine. N Engl J Med 1983;309:709–13.
[12] Young KD, Lewis RJ. What is confidence? Part 1. The use and interpretation of confidence intervals. Ann Emerg Med 1997;30(3):307–10.
[13] Rothman KJ. A show of confidence. N Engl J Med 1978;299:1362–3.
[14] Bulpitt CJ. Confidence intervals. Lancet 1987;14:494–7.
[15] Gardner MJ, Altman DG. Confidence intervals rather than P values: estimation rather than hypothesis testing. Br Med J Clin Res Ed 1986;292:746–50.
[16] Simon R. Confidence intervals for reporting results of clinical trials. Ann Intern Med 1986;105:429–35.
[17] International Committee of Medical Journal Editors Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication, section. http://www.icmje.org [accessed 14th December, 2008].
[18] Modernising Medical Careers. http://www.mmc.nhs.uk/default.aspx?page=503 [accessed 14th December, 2008].
[19] Intercollegiate Surgical Curriculum Project. http://www.iscp.ac.uk/Syllabus/Structure.aspx [accessed 14th December, 2008].
[20] Bland JM, Altman DG, Royston JP. Statisticians in medical schools. J Roy Coll Phys Lond 1990;24:85–6.
[21] Altman DG, Bland JM. Improving doctors' understanding of statistics (with discussion). J Roy Stat Soc A 1991;154:223–67.
[22] Vere-Jones D. The coming of age of statistical education. Int Stat Rev 1995;63:3–23.