Injury, Int. J. Care Injured (2008) 39, 659—665

INJURY
INTERNATIONAL JOURNAL OF THE CARE OF THE INJURED

www.elsevier.com/locate/injury

Graphical data presentation

Dirk Stengel [a,*], Georgio M. Calori [b], Peter V. Giannoudis [c]

[a] Centre for Clinical Research, Department of Orthopaedic and Trauma Surgery, Unfallkrankenhaus Berlin, Warener Str. 7, 12683 Berlin, and University Hospital of Greifswald, Sauerbruchstr., 17475 Greifswald, Germany
[b] Academic Department of Trauma and Orthopaedic Surgery, School of Medicine, University of Milano, Italy
[c] Academic Department of Trauma and Orthopaedic Surgery, School of Medicine, University of Leeds, UK

Accepted 21 January 2008

KEYWORDS
Manuscript writing;
Graphical data;
Figures;
Charts

Summary Figures and charts are the most influential vehicles for distributing scientific information, for affecting decisions as to the acceptance or rejection of a manuscript, and for attracting the attention of the scientific community to study results. Graphical excellence is mainly defined, first, by the highest possible data density (that is, the amount of information provided per graph area); second, by a low ink-to-data ratio (the avoidance of unnecessary shading, three-dimensionality, gridlines and what is often called 'chartjunk'); and third, by clear and unequivocal labelling of axes. The researcher's essential graphical toolbox should contain histograms, bar charts (always with measures of error), box-and-whiskers plots, scatter plots and forest plots.

Introduction

'Try graphics first' is a basic scientific principle when analysing the data from a biomedical investigation. Before proceeding with formal statistical analyses, a graph gives a first impression of the effect size and the centre and distribution of values and outliers. As physicians, specifically surgeons, we are visual people, and often we grab the essentials from a clinical study more effectively by graphical than by numerical presentation.

On the other hand, and in addition to the manipulation of photographic images, graphs are the most common vehicles for rigging information and intentionally leading readers and the scientific community down the wrong track. For example, in a review of 74 pharmaceutical advertisements, numerical distortion (e.g. improper scaling leading to visual over- or underestimation of effect sizes) and redundancy were observed in 36% and 46% of all investigated graphs.[3] Thus, reading and interpreting figures properly has

* Corresponding author. Tel.: +49 30 5681 3030; fax: +49 30 5681 3034.
E-mail address: dirk.stengel@ukb.de (D. Stengel).

0020-1383/$ — see front matter
doi:10.1016/j.injury.2008.01.050

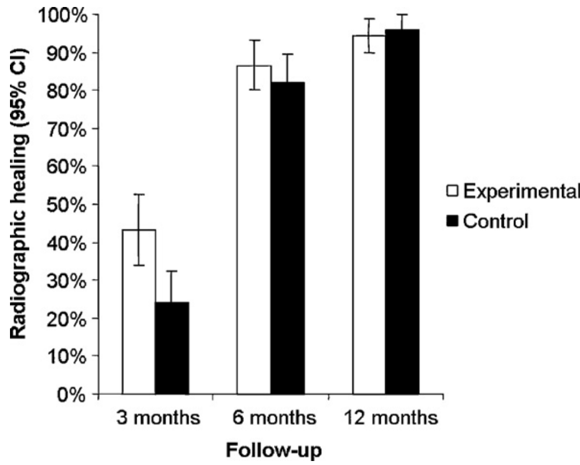


Figure 1 A simple bar chart option for presenting the data from a comparative study of two interventions for fracture treatment. The figure allows for: intragroup, longitudinal assessments of healing rates; intergroup comparisons of treatment effectiveness; and rough statistical inferences (with no overlap of the 95% confidence intervals at the 3-month follow-up, it is unlikely that the advantage of the experimental over the control treatment was produced by chance).

become an important skill for evidence-based practice.

A figure that presents the key findings from a study in a comprehensive and clear fashion is more than a necessary add-on to the manuscript—it may be a decisive factor in the acceptance or rejection of the paper during peer review, and immediately indicates scientific professionalism and honesty. Outstanding diligence in the choice of graph format and its design is warranted.

Many guidelines to authors in high-circulation journals still lack precise recommendations on how to prepare figures. For example, *Injury*'s author instructions print eleven lines on figures, mainly on technical issues and not on the preferred types of graph, labelling of axes or suitable graphical elements. The most important advice comes in a single sentence: 'Figures should be limited to those considered essential.' Authors should strategically plan partitioning of study results into figures, tables and text. For example, data related to the primary hypothesis may elegantly be presented in a first-order graph such as a bar chart, box plot or scatter plot, together with appropriate measures of error and distribution. Whenever possible, 95% confidence intervals should be added to allow the reader to assess both the relevance and the significance of the findings. Data from subgroups and stratified analyses or those related to secondary hypotheses can be presented graphically as well, e.g. box plots for different strata, or tabulated. Further results considered noteworthy, conflicting with current evidence or otherwise hypothesis-generating, may be explained in the text.

Principles of graphical excellence

Figures must replace but not repeat written text. As a rule of thumb, a figure is needed if a written passage is far more complex to comprehend than an illustration, e.g. 'After 3, 6 and 12 months of follow-up, radiographic healing was noted in 45/104 (43%), 90/104 (87%) and 98/104 (94%) fractures in the experimental group. In the control group, these numbers were 24/100 (24%), 82/100 (82%) and 96/100 (96%).' An example how to present this information graphically is sketched in Fig. 1.

Extensive work on the use and misuse of graphics in the biomedical literature has been conducted by

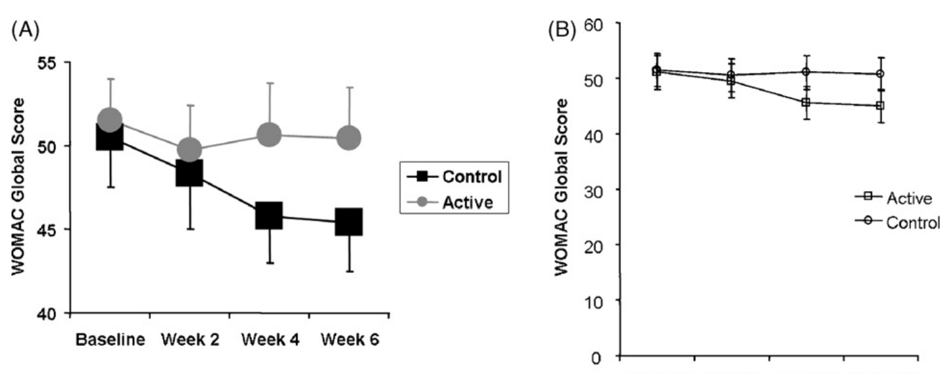


Figure 2 Misguidance of readers on the effectiveness of magnetic pulse treatment compared with placebo for knee osteoarthritis (A).[8] The WOMAC (Western Ontario and McMaster University OA Index) may range from 0 (perfect health) to 96 (maximum pain, stiffness and functional impairment). The original *y*-axis depicts only the interval from 40 to 55, thereby visually inflating the effect size. The artificial difference is pronounced by one-sided error bars (standard deviations, as indicated in the statistics section of the manuscript). After rescaling (B), the ineffectiveness of magnetic pulse treatment becomes obvious.

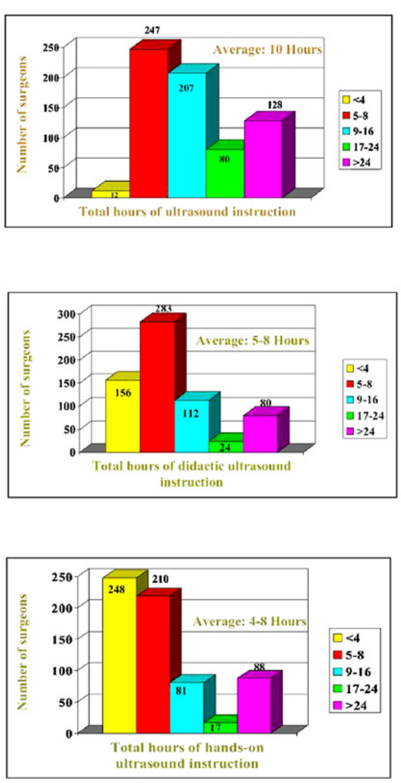
Tukey,[12] Tufte,[11] Wainer and Velleman,[13] Cooper et al.,[1,2] Schriger and Cooper[9] and others. They all contributed significantly to a pragmatic framework for achieving graphical excellence. The following minimum standards and rules of graphical data presentation can thus be defined.

1. The data-to-ink ratio and the data density index, i.e. the number of entries in the data matrix per area of data graphic, must be as high as possible. Any 'chartjunk', i.e. unnecessary shading, gridlines, three-dimensionality or overlap, must be avoided. A figure tells a story, and byplay seriously distracts readers from the key message. Although technical editing of accepted manuscripts may already delete much non-informative ballast, it remains the authors' responsibility to check whether figures convey a maximum of information with a minimum number of graphical elements.
2. Although the selective use of colour may greatly enhance information flow and highlight the key message in a slide presentation, it has little, if any, meaning in a scientific manuscript (except for photographic images, such as histological sections). Since most journals do not publish colour figures because of the high printing costs, it makes almost no sense to submit them for peer review. The graphical features of commercial software such as Microsoft Excel® or advanced statistical packages are seductive, and researchers may feel that colour jazzes up their figures or makes them more impressive. However, if a figure needs colour to attract attention, it is likely to be useless. Scientific professionalism is not expressed by turning graphs into artwork, but by as simple and comprehensible a design as possible. If information is new and important, a black-and-white line drawing will be sufficient and self-explanatory.
3. All axes and elements of a figure must be unequivocally labelled.
4. Natural scales with the entire range of values are strongly recommended. Error bars must extend in both directions. Fig. 2 illustrates how the combination of a fragmented *y*-axis and one-sided error bars falsely suggest a large difference between two treatment methods.[8]
5. Font sizes should be adapted to the size of the graph and the graph area.

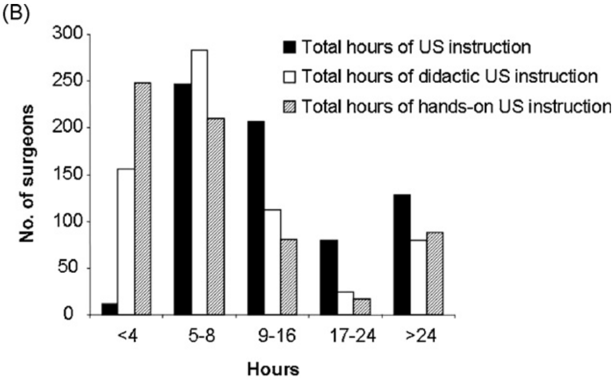
Examples from the literature

Even in recent manuscripts, remarkably inefficient graphical presentation can be found. Care must be

(A)



(B)



(C)

Total Hours	US instruction n=674	Didactic US instruction n=655	Hands-on US instruction n=644
<4	12 (1.8%)	156 (23.8%)	248 (38.5%)
5-8	247 (36.6%)	283 (43.2%)	210 (32.6%)
9-16	207 (30.7%)	112 (17.1%)	81 (12.6%)
17-24	80 (11.9%)	24 (3.7%)	17 (2.6%)
>24	128 (19.0%)	80 (12.2%)	88 (13.7%)

Figure 3 Inefficient data presentation by multicolour, three-dimensional bar charts (A).[10] Data density can be increased by aggregating the graphs (B). However, a table provides far more clarity, and the full range of information can be achieved by including both absolute numbers and percentages (C).

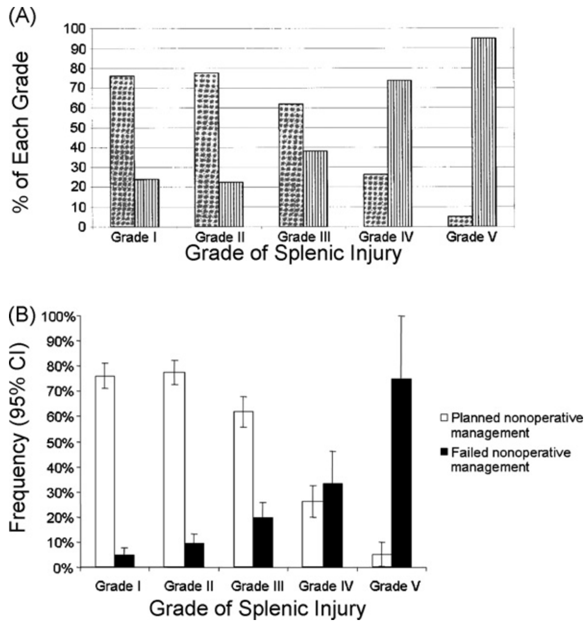


Figure 4 Information redundancy (A).[7] Note that, in addition to the text, the same information is provided three times in the figure; percentages of operative management are expressed in the legend and by striped bars and are repeated by shaded bars (non-operative management = 1 − operative management). (B) Aims at increasing data density by combining two figures—the rate of non-operative management plus the failure rate with non-operative management, which appeared in another illustration on the same page. Confidence intervals were added to allow for some rough predictions. Reformatting revealed another problem; there were 276 grade I, 299 grade II, 247 grade III, 194 grade IV and 78 grade V splenic injuries. Of these, 210, 232, 153, 51 and 4 (!), respectively, were scheduled for non-operative management. Thus, three of four non-operative managements for grade V splenic ruptures failed, but the calculation of a percentage is not justified with this small sample. A solution would be to sum grade IV and V injuries.

taken not to dilute interesting data with poor illustration. Fig. 3A shows examples from a published report of a mail survey among surgeons who had participated in the American College of Surgeons ultrasound educational programme.[10] Four three-dimensional colour bar charts were placed at either edge of one page, taking up 179 cm^2 (41%) of 441 cm^2 of effective page space. Note the low data-to-ink ratio and the non-informative use of colour, even with axis labels. To increase data density, three figures could have been condensed into one (Fig. 3B), or simply replaced by a table (Fig. 3C).

An example of serious information redundancy is shown in Fig. 4.[7] Percentages of intended operative or non-operative treatment for blunt splenic injury were provided in the text, the figure legend and the figure twice, i.e. percentage/grade and (100% − percentage)/grade. The figure occupies 37% of effective page space and the paper contains multiple figures of similar size with redundant information. Fig. 4B was designed to bring two key messages of the EAST multi-institutional spleen study into context, i.e. non-operative management was less likely to be chosen with higher grades of injury and more likely to fail with severe injuries.

Pie charts belong to the most useless graph types, and must be avoided in a scientific manuscript. Fig. 5 sketches an example, published in the report of a randomised trial of hook pins versus AO screws for internal fixation of cervical hip fractures.[6] The figure does not allow for reading proportions, and should have been replaced by a histogram (Fig. 5B). A stacked bar chart was presented at the bottom of the same page (Fig. 6). Stacked bars have both advantages and disadvantages. They may easily compile different outcomes but, since they produce cumulative percentages, the relevant information is difficult to extract. In the present example, the risk

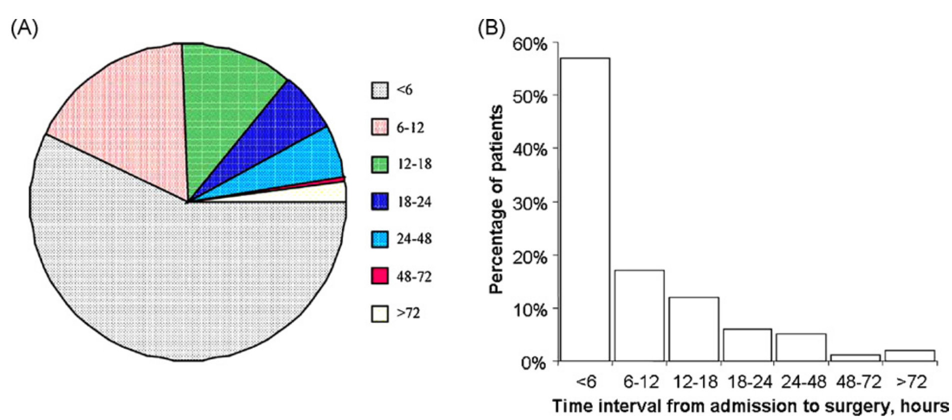


Figure 5 Inefficient data presentation by a multicolour pie chart (A).[6] Although intended to show the distribution of time intervals between admission and fracture fixation, percentages cannot be traced from the diagram. A histogram would have easily depicted the distribution of lag time (B). Unfortunately, no stratification was made for the interventions under investigation.

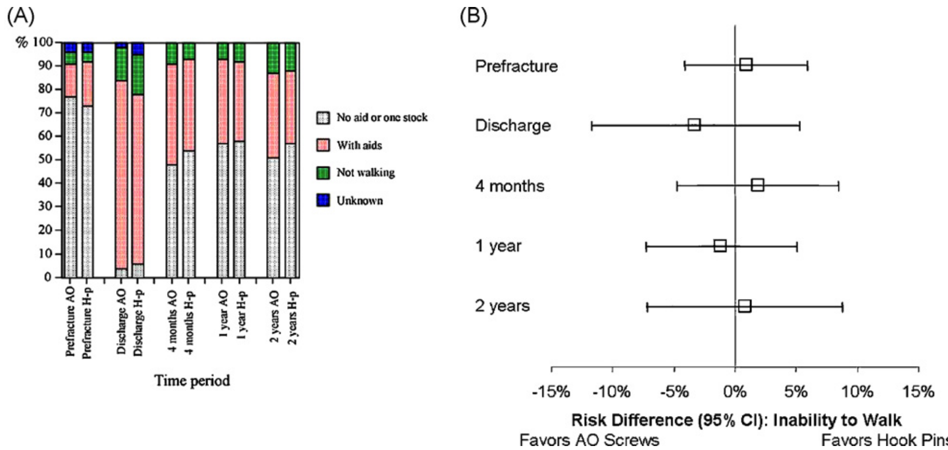


Figure 6 Information hidden in a stacked bar chart (A).[6] A forest plot (B) shows no evidence of a difference in the risk of not walking between the interventions under investigation.

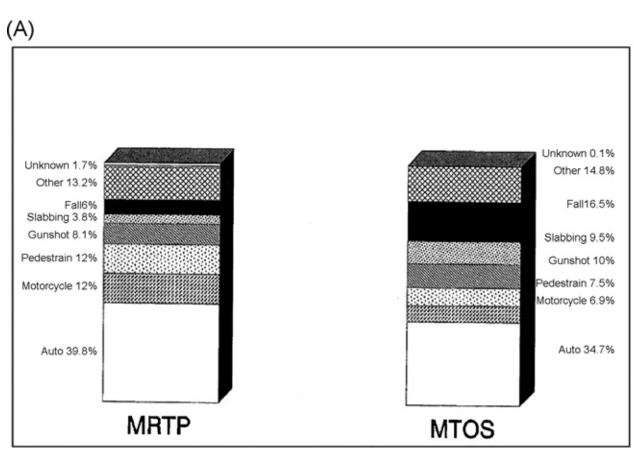
of walking inability may be of particular interest. Forest plots, the classic graphs to summarise the results from meta-analyses, are increasingly used to display differences between treatment methods. They may illustrate outcomes after distinct periods of follow-up and within subgroups (Fig. 6B). They enable the observer to eyeball both the effect size and its statistical meaning, i.e. if the 95% confidence interval includes the null (or one, in case of a ratio) there is no evidence of a difference at the two-sided *p*-value of 0.05. Also, stacked bar charts may be confusing if they contain more than two or three categories, as shown in Fig. 7A.[5] Probably the best alternative to this space-consuming, uninformative graph may be a table (Fig. 7B).

Schriger and Cooper stressed the need for distinguishing between unpaired and paired observations.[9] Fig. 8A shows the pre- and postoperative ranges of motion (ROM) in elbow joints of 14 people undergoing surgical resection of heterotopic ossifications.[4] Again, the figure contains much chartjunk. Two box plots would have pictured the gain in ROM more simply and clearly than the original figure (Fig. 8B). However, summary measures, e.g. mean, median, may obscure the worsening of function in individual cases (Fig. 8C). In the case of small sample sizes, i.e. 20—25 subjects, one-way plots may reveal both overall trends and individual patients' courses.

The essential graphical toolbox

The following list covers the essential graphical tools needed to illustrate almost any study result.

1. Histograms. Although requiring more space than other figures, these are useful for displaying the



(B)

Mechanism of injury	MRTP	MTOS
	n=266	n=80544
Unknown	5 (1.7%)	81 (0.1%)
Other	35 (13.2%)	11921 (14.8%)
Fall	16 (6.0%)	13290 (16.5%)
Stabbing	10 (3.8%)	7652 (9.5%)
Gunshot	22 (8.1%)	8054 (10.0%)
Pedestrian	32 (12.0%)	6041 (7.5%)
Motorcycle	32 (12.0%)	5558 (6.9%)
Auto	106 (39.8%)	27949 (34.7%)

Figure 7 Low data density of a three-dimensional stacked bar chart (A) that compares mechanisms of injury among patients enrolled in the Modal Rural Trauma Project (MRTP) and the Major Trauma Outcome Study (MTOS).[5] It does also not account for the huge difference in sample sizes between both studies. Again, a table allows for a more valid comparison (B).

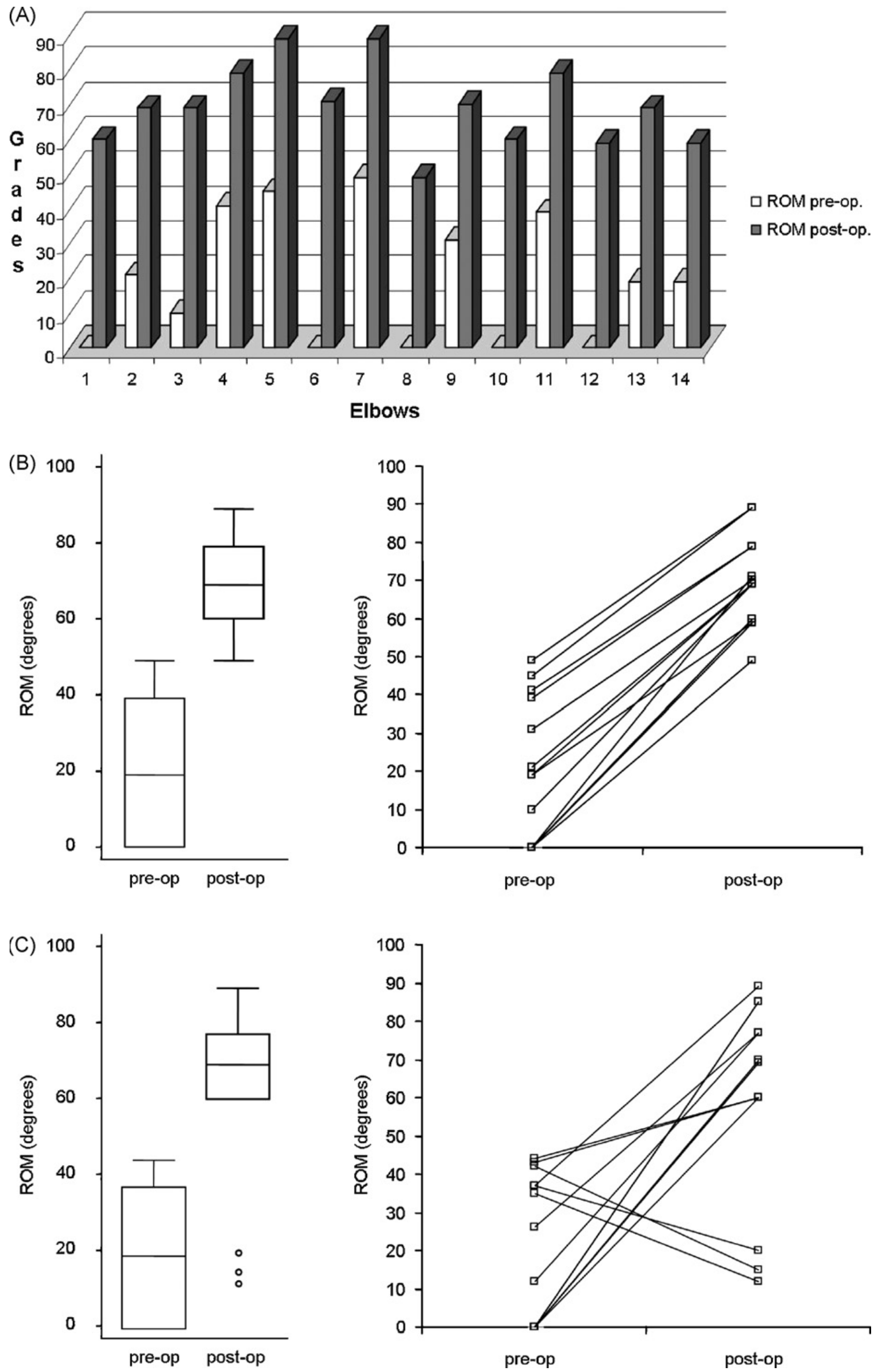


Figure 8 Handling of paired data. (A) Original figure, as published by de Palma et al.[4] (B) Box-and-whiskers plots of pre- and postoperative ranges of motion (ROM) provide equivalent information with a much higher data density index. The full scope of information, i.e. the individual effect of surgical resection in all patients, is provided by a one-way plot. (C) Fictitious dataset. Although the box-and-whiskers plots resemble those shown in (B), they obscure worsening of function in three cases, and moderate effects in another two.

entire range of data and detecting bimodal distributions.

2. Bar charts (always add measures of error, describe whether and why they consider 95% confidence intervals).
3. Box-and-whiskers plots. By convention, boxes always include the median and the interquartile range. Since there is diversity in the meaning of the whiskers, and different software packages may produce different outputs, this has to be

specified in the figure legend, e.g. 1.5 times the interquartile range, 10th and 90th percentile, 5th and 95th percentile, minimum and maximum values.

4. Scatter plots. If regression lines are added, they should be surrounded by a 95% confidence interval.
5. Study flow diagrams, according to the CONSORT statement, and other recommendations for reporting trials.
6. One-way plots for paired data.
7. Survival curves.
8. Forest plots for risk ratios, odds ratios, risk differences and differences in means.
9. Receiver operating characteristics for diagnostic test data. These depict the trade-off between sensitivity and 1 − specificity.

Conflict of interest

None.

References

1. Cooper RJ, Schriger DL, Close RJ. Graphical literacy: the quality of graphs in a large-circulation journal. Ann Emerg Med 2002;40:317—22.
2. Cooper RJ, Schriger DL, Tashman DA. An evaluation of the graphical literacy of Annals of Emergency Medicine. Ann Emerg Med 2001;37:13—9.
3. Cooper RJ, Schriger DL, Wallace RC, et al. The quantity and quality of scientific graphs in pharmaceutical advertisements. J Gen Intern Med 2003;18:294—7.
4. de Palma L, Rapali S, Paladini P, Ventura A. Elbow heterotopic ossification in head-trauma patients: diagnosis and treatment. Orthopedics 2002;25:665—8.
5. Karsteadt LL, Larsen CL, Farmer PD. Analysis of a rural trauma program using the TRISS methodology: a three-year retrospective study. J Trauma 1994;36:395—400.
6. Mjorud J, Skaro O, Solhaug JH, Thorngren KG. A randomised study in all cervical hip fractures osteosynthesis with Hansson hook-pins versus AO-screws in 199 consecutive patients followed for two years. Injury 2006;37:768—77.
7. Peitzman AB, Heil B, Rivera L, et al. Blunt splenic injury in adults: multi-institutional study of the Eastern Association for the Surgery of Trauma. J Trauma 2000;49:177—87.
8. Pipitone N, Scott DL. Magnetic pulse treatment for knee osteoarthritis: a randomised, double-blind, placebo-controlled study. Curr Med Res Opin 2001;17:190—6.
9. Schriger DL, Cooper RJ. Achieving graphical excellence: suggestions and methods for creating high-quality visual displays of experimental data. Ann Emerg Med 2001;37:75—87.
10. Staren ED, Knudson MM, Rozycki GS, et al. An evaluation of the American College of Surgeons' ultrasound education program. Am J Surg 2006;191:489—96.
11. Tufte ER. The visual display of quantitative information, 2nd ed., UK: Graphics Press; 2001.
12. Tukey JW. Data-based graphics: visual display in the decades to come. Stat Sci 1990;5:327—39.
13. Wainer H, Velleman PF. Statistical graphics: mapping the pathways of science. Annu Rev Psychol 2001;52:305—35.