EJN fens EUROPEAN JOURNAL OF NEUROSCIENCE

European Journal of Neuroscience, Vol. 28, pp. 2363–2364, 2008 doi:10.1111/j.1460-9568.2008.06581.x

EDITORIAL
Reporting statistical methods and statistical results in EJN

Martin Sarter[1] and Jean-Marc Fritschy[2]
[1]Department of Psychology, University of Michigan, Ann Arbor, MI, USA
[2]Department of Pharmacology, University of Zurich, Zurich, Switzerland

EJN publishes research that represents the entire spectrum of the neurosciences. The content and style of the description of the statistical methods and the statistical results in the manuscripts submitted to EJN reflect the diverse customs of the many fields of neuroscientific research. Although EJN embraces these heterogeneous reporting and writing styles, we need to ensure a consistent description of statistical methods and results in our published papers. By definition, a scientific article should permit researchers to reproduce the study described in this article; to this end, an informative and complete description of statistical methods and results is an essential component of an effective scientific report.

The aim of this editorial is to provide guidance for the reporting of statistical methods and results. We consulted numerous specialized sources (referenced below) as well as comprehensive guidelines such as Lang & Secic (2006). Finally, the Associate Editors of EJN were consulted. As a result, the present recommendations constitute a component of the EJN Author Guidelines.

Statistical methods continue to be debated and to evolve as indicated, for example, by the ongoing discussion about mixed-effect regression models or planned comparisons as alternative approaches to traditional repeated measures ANOVAs and omnibus tests (Lavori, 1990; Gueorguieva & Krystal, 2004; Gonzalez, 2008). The present recommendations should not be viewed as a partisan stance with respect to this or other, sometimes more fundamental, discussions about the use of statistical methods and traditional null hypothesis testing (e.g. Shrout, 1997). The selection of the statistical approaches and the explanation and justification of methods remain the sole responsibilities of the authors. We expect that future 'Technical Spotlight' articles will bring some of the new developments in statistics to the readers of EJN.

Description of statistical methods

The objective of an effective description of statistical methods can be readily stated: 'Describe the statistical methods with enough detail to enable a knowledgeable reader with access to the original data to verify the reported results' (Bailar & Mosteller, 1988, p. 266).

Guidelines

– Provide a clear description of the design of your study or experiment; it is often informative to state the null hypothesis.
– Describe and justify the statistical approaches and the selection of statistical tests.

Correspondence: M. Sarter and J.-M. Fritschy, as above.
E-mail: editorial.office@ejn.uk.com; msarter@umich.edu; fritschy-ejn@pharma.uzh.ch

– If you report null (or negative) statistical results and attribute significance to such results, describe the methods used to determine adequate statistical power (Cohen, 1988; Murphy & Myors, 1998).
– Describe and justify data transformation procedures (e.g. arcsin transformation of percentage data).
– Define all within-subject and between-subject factors.
– Define planned comparisons.
– For multiple comparisons and multiple correlations (e.g. Curtin & Schulz, 1998) define measures taken to reduce Type I errors (e.g. Bonferroni- adjusted alpha levels).
– For repeated measures ANOVAs, define measures taken to control for violation of the sphericity assumption (Vasey & Thayer, 1987; Keselman, 1998); describe how you report results of corrected degrees of freedom statistics (see recommendation below).
– Determine the alpha level used as a significance criterion for your tests.
– State the name and version of the statistical software that was used (also company, city, state, country).

Description of statistical results

Generally, the Results section should provide a complete description of the results of the study. As such, the main statistical results should be described in this section. The reporting of statistical results in the figure legends should be limited to the results of *post hoc* multiple comparisons (that are often also indicated by symbols in the figures).

It is useful to reiterate the factors that were analysed by ANOVA in the Results section. To a reasonable degree, an effective Results section reiterates major aspects of the design of the study and avoids a 'dry' listing of statistical findings. It is often effective to state the main finding and its scientific meaning prior to describing the underlying statistical evidence. Readers should be able to understand the main statistical approaches and the meaning of the statistical results without returning to the section on statistical methods. Clearly, there is a fine line between reiterating statistical methods and providing sufficient information about the statistical approach that forms the basis for a particular statistical result; we encourage authors to err on the side of reiteration in order to generate an informative result section.

Guidelines

– Important descriptive statistics, such as mean and standard deviation (SD), or standard error of the mean (SEM), need to be either represented graphically or numerically in the text.
– All statements concerning significance must be qualified numerically.

– Always report the test statistic, the degrees of freedom, the test value, and the *P*-value that the result occurred at chance under the null hypothesis.
– For tests involving 1 degree of freedom (e.g. Student's *t*-test), state whether a directional or non-directional test was conducted.
– Round test statistics and individual *P*-value equalities to two siginificant figures, if applicable.
– To avoid ambiguities, all statistical variables should be italicized (*F, t, P*).
– In accordance with Greenwald *et al.* (1996) we recommend reporting individual *P*-values as equalities rather than as inequalities in relation to an alpha criterion (e.g. $P = 0.003$ as opposed to $P < 0.01$). However, inequalities may be useful for groups of data (e.g. in Tables or graphics).
– Ensure that all *P*-values defined in Figure legends and Table footnotes are linked (e.g. by symbols such as asterisks) to the corresponding data.
– The format of the description of the statistical results should follow these examples:

$$F_{1,32} = 22.32,\ P = 0.08$$

$$t_{27} = 7.85,\ P = 0.17$$

$$\chi^2_2 = 20.32\ (n = 62),\ P = 0.35$$

$$r_{28} = 0.73,\ P = 0.04$$

– We recommend the use of parentheses to maintain the readability of statements [e.g. 'Blockade of AMPA receptors attenuated the firing rate of ventral pallidal neurons (main effect of concentration of DNQX: $F_{1,32} = 28.32$, $P = 0.03$). However, the attenuation of firing rate was greater in animals lacking M1 muscarinic receptors when compared to wild type mice (interaction between effects of DNQX and genotype: $F = \ldots$; main effect of genotype: $F = \ldots$').].
– Corrected degrees of freedom statistics (for omnibus repeated measures ANOVAs and if corrections are required because of violation of the sphericity assumption; see above): in order to preserve the transparency of the statistical design, we recommend reporting the uncorrected degrees of freedom together with the corrected test value; authors may also report the correction factor ε to indicate the degree of sphericity.

Conclusions

These suggestions and guidelines are designed to ensure consistency in the description of statistical methods and results in EJN. Obviously, special circumstances, such as data that are not normally distributed or violations of variance homogeneity, may require special methods, or evidence indicating that the validity of the selected statistical analyses was not violated. Generally, we recommend that authors obtain expert assistance when analysing their data or have their manuscript reviewed by a statistician. Valid and conclusive statistical analyses and a transparent and complete description of the statistical results are essential components of a high quality scientific report.

Acknowledgements

We are grateful to the numerous suggestions provided by the Associate Editors of EJN. Additional comments were obtained from Drs Richard Gonzalez and Cindy Lustig (University of Michigan) and Dr Florence Crestani, University of Zurich.

References

Bailar, J.C. & Mosteller, F. (1988) Guidelines for statistical reporting in articles for medical journals. *Ann. Intern. Med.*, **108**, 266–273.
Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Lawrence Erlbaum Assoc., Hillsdale.
Curtin, F. & Schulz, P. (1998) Multiple correlations and Bonferroni's correction. *Biol. Psychiatry*, **44**, 775–777.
Gonzalez, R. (2008) *Data Analysis for Experimental Design*. Guilford Publications, New York, NY.
Greenwald, A.G., Gonzalez, R.T., Harris, R.J. & Guthrie, D. (1996) Effect sizes and *p* values: what should be reported and what should be replicated? *Psychophysiology*, **33**, 175–183.
Gueorguieva, R. & Krystal, J.H. (2004) Move over ANOVA. *Arch. Gen. Psychiatry*, **61**, 310–317.
Keselman, H.J. (1998) Testing treatment effects in repeated measures designs: an update for psychophysiological research. *Psychophysiology*, **35**, 470–478.
Lang, T.A. & Secic, M. (2006) *How to Report Statistics in Medicine: Annotated Guidelines for Authors, Editors and Reviewers*, 2nd Edn. ACP Press, Philadelphia, PA.
Lavori, P. (1990) ANOVA, MANOVA, my black hen. *Arch. Gen. Psychiatry*, **47**, 775–778.
Murphy, K.R. & Myors, B. (1998) *Statistical Power Analysis*. Lawrence Erlbaum Assoc., Mahaw, NJ.
Shrout, P.E. (1997) Should significance tests be banned? *Psychol. Sci.*, **8**, 1–2.
Vasey, M.W. & Thayer, J.F. (1987) The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. *Psychophysiology*, **24**, 479–486.