JOURNAL OF THE SCIENCE OF FOOD AND AGRICULTURE

INSTRUCTIONS TO AUTHORS

GENERAL

Journal of the Science of Food and Agriculture publishes original research and critical reviews in agriculture and food science, with particular emphasis on interdisciplinary studies at the agriculture/food interface. The Journal covers fundamental and applied research including:

- production and processing of human food and animal feed
- properties of materials
- nutritional quality, sensory analysis, flavour, texture, safety and toxicity of foodstuffs, beverages and alcoholic drinks
- plant and animal husbandry, physiology, yield and quality
- agricultural production of industrial materials
- interaction of agriculture with forestry and the environment
- biotechnology and genetic engineering of plants, animals and microorganisms affecting agriculture, food production and processing, and fermentation science.

Concise contributions on experimental or theoretical investigations and critical reviews of the science of food and agriculture are invited for publication.

SUBMISSION OF PAPERS

All papers must be in English. Non-English speaking authors who do not have a good command of English are advised to seek assistance from someone who has. Typescripts in triplicate, including three copies of all tables etc, together with an original and two copies of all illustrations, but three original prints of photographs, should be sent to:

The Managing Editor,

Journal of the Science of Food and Agriculture,

SCI, 14/15 Belgrave Square,

London SW1X 8PS, UK

Tel: +44 (0)20 7235 3681

Fax: +44 (0)20 7235 0887

e-mail: jsfa@soci.org

Authors in North America should submit to:

Professor D S Reid,

Department of Food Science & Technology,

University of California, Davis,

1 Shields Avenue,

Davis, CA 95616-8571, USA

Tel: +1 530 752 8448

Fax: +1 530 752 4759

e-mail: dsreid@ucdavis.edu

Authors in Japan: If you require assistance, Wiley-Japan can provide a list of recommended services to check and improve the English in papers before submission. Please contact A Bocquet in the Wiley-Japan office by fax (+81 (0)3 3556 9763) or e-mail (wileyjpn@mb.kcom.ne.jp) for more information, stating which SCI/Wiley journal you wish to submit to.

In addition, please submit an electronic version saved on a PC disk, in a format compatible with MS Word 7 or lower.

Inadequately or incorrectly prepared typescripts may be rejected. Authors will receive an immediate acknowledgement of receipt of their paper followed, normally within four months, by a decision.

When preparing a manuscript, authors should refer to a recent (2001 onwards) issue of the Journal and follow the detailed instructions given below. **The corresponding author must obtain the consent of all the co-authors to the submission of the paper.** Papers will only be accepted on the understanding that their contents have neither been published nor are being offered for publication elsewhere.

MANUSCRIPT PREPARATION

Length

Papers should not normally exceed 6000 words, including relevant data. Any manuscript that is submitted which, in the opinion of the editor, is too long will be returned to the corresponding author for redrafting within a suggested maximum wordage.

Typescript

Type papers in double spacing on one side of A4 or 8½" × 11" paper with 30 mm wide left and right margins. Underlining to indicate italicised type (or use of italic type in the manuscript) should be restricted to genera and species names, chemical descriptors (eg cis, trans etc) and journal and book titles. Do not underline any headings. Footnotes should be kept to a minimum and indicated by * or †. Do not use full stops after abbreviations unless essential for clarity. Abbreviations of chemical and other names should be defined when first mentioned in the body of the paper, unless commonly used and internationally known and accepted. Each page should be numbered individually.

Units and nomenclature

Units Use **SI units** in accord with the recommendations of the International Organisation for Standardisation (ISO). **Use the form g kg^{-1} etc (not** %) **to specify content/composition/concentration.** Use % only to express proportional change. Note that the form g $100g^{-1}$ etc is not correct. Avoid the use of g per 100g, for example in food/feed composition, by using g kg^{-1}. Fertiliser rates should be presented in terms of the element applied. Further information on the ISO recommendations can be obtained from the following publication issued by the British Standards Institution, London: *Specification for SI units and recommendations for the use of their multiples and of certain other units*, BS 5555 : 1993 ISO 1000 : 1992.

Symbols Write all symbols, formulae and equations with great care. Unusual symbols (including Greek lettering) should be defined in words in the left margin at the first mention.

Scientific names Give the scientific names (with authority) for plants, animals, microorganisms, with generic names in full at the first mention, eg *Myzus persicae* (Sulzer). Thereafter abbreviate them in the text, eg *M persicae*; give them in full (without authority) in the headings of sections, tables, figures and key words. Where appropriate, cultivars should be specified.

Enzyme nomenclature Identify each enzyme together with its EC number, if available, at the first mention, following the recommendations of the latest edition of *Enzyme Nomenclature*.

Chemical nomenclature Use the current systematic IUPAC nomenclature throughout.

Statistical analyses

Particular care should be taken to ensure that the **appropriate** statistical analyses have been carried out. The methods used should be described concisely, yet with enough information to explain how the chosen methods have been applied to the data. The form of all experimental errors and their statistical significance must be given clearly. The statistical analyses should be used in the discussion to justify inferences made against the background of normal biological variation. Further information on recommended statistical procedures can be found in Appendix II, or printed in *J Sci Food Agric* **81** No 1 (2001); additional copies are available from the Managing Editor.

Layout

The main body of the paper should be divided into **unnumbered** sections and each given an appropriate heading. Main headings should be capitalised and centred over the text. Choice of headings will depend on the content, but the following is recommended for research papers:

Title This should be concise, specific and explain the nature of the work. State in a footnote if the paper was given, in whole or in part, at a scientific meeting.

Running title A running title of up to 80 characters should also be provided.

Authors' names Each must have the customary forename in full and initials (eg Arthur B Smith). Give the full address(es) **where the work was done.** If an author was on secondment or visiting from another address, or has since moved to a new address, this should be given in a footnote.

Abstract This must be informative yet concise, give essential information such as the purpose of the work, the data derived from it and their statistical significance, and be intelligible without reference to the paper itself. It should not normally exceed 150-200 words. Authors should remember that the abstract is often the only portion of a paper read, as in abstracting journals, and the use of unusual acronyms or abbreviations should be avoided.

Key words List all the main topics incorporated in the paper, including any already given in the title.

Introduction Include a clear description of the aims of the investigation (without summarising the work itself) and a brief statement of previous relevant work with references.

Experimental State clearly, in sufficient detail to permit the work to be repeated, the methods and materials used. Only new techniques and modifications to known methods need to be described in detail but known methods must have adequate references. Include the name, postal town, code and country of the supplier or manufacturer of any chemical or apparatus not in common use. Give the **statistical design (including replication)** of each experiment where appropriate (see also Statistical analyses, above).

Results Present these concisely, using tables or illustrations for clarity; do not list the results again in the text. State clearly the form of the experimental error and the statistical significance of the results (see also Statistical analyses, above). Do not overstate the precision of the measurements. Histograms or bar charts, unless prepared carefully, are inferior to tables. Only in exceptional circumstances will both tables and illustrations based on them be accepted. The Experimental and Results sections may be combined when appropriate.

Discussion The Results should be followed by a concise section to discuss and interpret them. Do not just repeat the results. A combined Results and Discussion section sometimes simplifies the presentation.

Conclusions Do not merely repeat content of preceding sections. The Discussion and Conclusions sections may be merged.

Acknowledgements Keep these to the absolute minimum. Avoid thanks for permission to publish.

References Check carefully for accuracy and follow the correct style. Refer to unpublished work only in the text (Smith A B unpublished), (Brown C D pers comm). Indicate literature references at the appropriate place in the text using superscript numbers in the order in which they appear and a full numerical list must appear at the end of the paper, giving all authors with initials after the respective surname. Ensure that all references in the list are cited in the text and vice versa. Give the date and full title of the paper in the language in which it appeared or an accurate English translation. Abbreviate all journal titles as in *Chemical Abstracts* or *Biological Abstracts* and the annual *BIOSIS List of Serials*, without using full stops after abbreviation. If the journal is not included, give its title in full. Volume numbers should be bold. Note the following style and order for journals:

1. Syers JK, Mackay AD, Brown MW and Currie LD, Chemical and physical characteristics of phosphate rock materials of varying reactivity. *J Sci Food Agric* **37:**1057-1064 (1986).

Quote books as follows:

2. Doyle J, *Altered Harvest. Agriculture, Genetics and the Fate of the World's Food Supply*. Viking Penguin Inc, New York, pp 136-158 (1985).

3. Thomas T, Barnes A and Hole CC, Modification of plant part relationships in vegetable crops, in *Chemical Manipulation of Growth and Development*, Ed by McLaren JS. Butterworths, London, pp 297-311 (1982).

When quoting patents give the name of the applicant, the year of publication, the title, the country and patent or application number thus:

4. Hagner MB and Wendt KL, Method of sorting seeds. UK Patent 1470133 (1977).

Tables Supply each table on a separate sheet. The table number (given as an Arabic numeral) should be given at the top, followed by a concise title. Give essential details as footnotes. Keep the number of columns to a minimum. Column headings should be brief, with the units of measurement clearly stated in parentheses. Where one unit applies to all the data in the body of the table include it in the title. The data should be easy to follow without horizontal lines between entries. A zero is often incorrect; use 'not detected' (ND) where appropriate, amplifying this, and trace (tr), where possible in a footnote, eg '<10 µg kg^{-1}'. For 'not significant' (NS) state the limiting level.

Illustrations Include only if essential, and number line drawings, figures and photographs in a single sequence in order of appearance using Arabic numerals. Keep lettering and numbering (characters) on illustrations to a minimum and include essential details in the legend. Write the authors' names and the title of the paper on the back of each illustration using soft pencil. Where any confusion is possible identify the top of the illustration clearly. Photomicrographs must have a scale bar. All legends should be combined in one list and typed on a separate sheet.

Where possible, illustrations should also be submitted in electronic format. Save each figure as a separate file, in TIFF or EPS format preferably, and include the source file.

Write on the disk the software used to create the files. Use dedicated illustration packages in preference to tools such as Excel or Powerpoint.

Line drawings and figures should be in a form suitable for direct reproduction, no larger than A4 or 8½" × 11", in black ink, with stencilled lettering (avoid using dry transfer, typewritten or handwritten lettering) all in proportion to the amount of detail. Computer-drawn diagrams must be prepared on a high quality laser or ink jet printer or plotter, **not** on a dot matrix printer or equivalent.

Use only essential characters and insert these and any other symbols clearly; explain all symbols used, and where a key to symbols is required, please include this in the artwork itself, not in the figure legend. On graphs, include labels and units on axes. Present logarithmic scales with arithmetic numbering 0.1, 1, 10, 100 rather than -1, 0, 1, 2. Avoid unnecessarily long axes that lead to large blank spaces on graphs.

Line drawings and figures should all require the same degree of reduction and all characters must be chosen so that after reduction they are at least 1.5 mm in height. The type area of the journal is 172 mm wide × 249 mm deep, in two columns each 81 mm wide, and the characters should therefore be large enough to be legible after reduction of the illustrations to fit the page or column width.

Photographs (halftones) should be supplied as glossy prints (four original prints of each) of good contrast, photocopies are **not** acceptable. Do not allow them to be damaged by paper clips, folding etc. Some loss of clarity may occur during reproduction.

Electrophoresis patterns These are complex. Photographs, which often lack clarity, should not be included except to make a particular point. Where the reporting of gel electrophoresis, SDS gels, immuno-electrophoresis, isoelectric focusing etc is essential, adhere to the following principles:

- a single zone requires only description in the text
- preferably claim homogeneity using a scan diagram
- preferably use a single gel to compare several tracks
- when scan diagrams are used accurate alignment is essential

Where photographs or scan diagrams must be used:

- number all zones and identify those common to more than one track
- give a molecular weight scale for SDS gels
- give experimental details and track identification in the legend

Chemical structures Prepare these on a separate sheet as described for illustrations and number the individual formulae with Roman numerals (I, II). All bonds, charges and free radicals should be accurately positioned. Indicate aromatic and unsaturated heterocyclic systems using double bonds. Preferably use general structures, distinguishing related compounds by substituents R^1, R^2 etc.

PROOFS

Proofs will be despatched to the corresponding author. These must be corrected and returned to the publishers **within 48 hours of receipt**. Author's corrections must be restricted to printer's errors. Give the full address for correspondence and proofs, including telephone and fax numbers and, where available, an e-mail address.

OFFPRINTS

There are no page charges. Twenty-five offprints of each article are supplied free of charge. Additional offprints can be ordered at current printing prices.

COPYRIGHT

Upon acceptance of a paper by the Journal, the author(s) will be asked to sign a transfer of copyright of the paper to SCI. The transfer will ensure the widest possible dissemination of information.

APPENDIX I: QUANTITIES, UNITS AND SYMBOLS

Recommended SI units and symbols

SI base units	*ame*	*ymbol*
Length	etre	
Mass	ilogram	g
Time	econd	
Electric current	mpere	
Thermodynamic temperature	elvin	
Amount of substance	ole	ol

Luminous intensity	candela	cd	

Multiples

To form decimal multiples of SI units the following prefixes may be used but for mass the prefix is added to the gram (g) and not the kilogram (kg).

Multiple	*Prefix*	*Symbol*	*Multiple*	*Prefix*	*Symbol*
10^{3}	kilo	k	10^{-6}	micro	μ
10	deca	da	10^{-9}	nano	n
10^{-1}	deci	d	10^{-12}	pico	p
10^{-2}	centi	c	10^{-15}	femto	f
10^{-3}	milli	m			

Derived units

Physical quantity	*Name*	*Symbol*	*Definition*
Energy	joule	kg m^2 s^{-2}	Nm
Force	newton	kg m s^{-2}	J m^{-1}

Pressure	pascal	a	$m^{-1} s^{-2}$	$N m^{-2}$
Power	watt		$m^2 s^{-3}$	$J s^{-1}$
Electrical charge	coulomb			A s
Electrical potential difference	volt		$m^2 s^{-3} A^{-1}$	$J A^{-1} s^{-1}$
Electrical resistance	ohm		$m^2 s^{-3} A^{-2}$	$V A^{-1}$
Electrical conductance	siemens		$m^{-2} A^2$	Ω^{-1}
Electrical capacitance	farad		$s^4 kg^{-1} m^{-2}$	A^2 $A s V^{-1}$
Magnetic flux	weber	b	$m^2 s^{-2} A^{-1}$	V s
Inductance	henry		$m^2 s^{-2} A^{-2}$	$V s A^{-1}$
Magnetic flux density	tesla		$s^{-2} A^{-1}$	$V s m^{-2}$
Luminous flux	lumen	m	sr	cd
Illuminance	lux	x	$sr m^{-2}$	cd
Frequency	hertz	z	$le s^{-1}$	$cyc s^{-1}$
Customary temperature	degree celsius	C		K
Mass	tonne		kg	10^3 Mg
Area	hectare	a	m^2	10^3

Volume	litre	10^{-3}	
	itre	m^3	dm^3

Other derived units

Physical quantity	*SI unit*	*Symbol*
Area	square metre	m^2
Volume	cubic metre	m^3
Density	kilogram per cubic metre	$g\ m^{-3}$
Velocity	metre per second	s^{-1}
Angular velocity	radian per second	$rad\ s^{-1}$
Acceleration	metre per second squared	s^{-2}
Kinematic viscosity	square metre per second	$m^2\ s^{-1}$
Dynamic viscosity	newton second per square metre	$s\ m^{-2}$
Electric field strength	volt per metre	m^{-1}
Magnetic field strength	ampere per metre	m^{-1}
Luminance	candela per square metre	$cd\ m^{-2}$
Entropy	joule per kelvin	

			K^{-1}
	Specific heat	joule per kilogram kelvin	J kg^{-1} K^{-1}
	Radiant intensity	watt per steradian	W sr^{-1}
	Molality	mole per kilogram	mol kg^{-1}
	Amount-of-substance concentration	mole per cubic metre	mol m^{-3}
	Stress	newton per square metre	N m^{-2}
	Young's modulus	newton per square metre	N m^{-2}
	Bulk modulus	newton per square metre	N m^{-2}
	Compressibility	square metre per newton	m^2 N^{-1}
	Moment of inertia	kilogram square metre	kg m^2

APPENDIX II: RECOMMENDATIONS OF THE EDITORIAL BOARD ON USE OF STATISTICS IN PAPERS SUBMITTED TO JSFA

GUIDELINES BASED ON A REPORT BY A W A MURRAY (*J SCI FOOD AGRIC* VOL 42 (1988)) AND UPDATED BY J W MCNICOL* (OCTOBER 1998)

*BioSS, SCRI, Invergowrie, Dundee, DD2 5DA, UK.

INTRODUCTION

This document is intended to provide some guidance on standards of statistical methodology and presentation for papers to be acceptable for publication in *Journal of the Science of Food and Agriculture*. It is not a definitive statistical text nor are any

rules intended to be inflexible. Good statistical practice, well presented in a manner appropriate to the journal style, should always be acceptable, but certain practices, which may either mislead or not do full justice to the authors' data, will be actively discouraged. These notes comprise some explanatory text and a list of useful references.

There are two somewhat interlinked aspects of the use of statistics in the context of scientific papers. These are methodology and presentation.

METHODOLOGY

The statistical methods used by authors should take account of the following points.

Experimental design

Replication of treatments

The statistical methods used in biology should assist in separating effects of interest from the 'noise' due to the inherent variability of biological systems, and also enable the estimation of quantities of interest together with appropriate measures of their precision of estimation. Authors must take account of the variability of their material when designing experiments. Adequate replication of treatments or samples must be made so that worthwhile inferences may be drawn in the presence of random variation. It is worthless to report 'no significant differences', or effects of similar size to error, when the replication is not adequate to detect differences of a size which would be of biological interest. Note that replicate experimental units for each treatment are required, rather than sub-samples within a bulk of similarly treated material. A treatment must be repeated on several experimental units to establish its repeatability. It cannot be shown that an effect of a treatment is repeatable by sub-sampling from within a single application of that treatment. Papers which seek to demonstrate differences where there is no valid estimate of error may be rejected on the grounds that such claims must be regarded as unsubstantiated. In general, the more variable the material and the smaller the threshold at which effects are of interest, the more replication must be made.

Arrangements of experimental units

Experimental units should be grouped (blocked) according to known or suspected sources of variation. Examples are the grouping of animals to take account of variation in weight at the beginning of a growth experiment, or the selection of blocks of land of similar fertility when laying out a field experiment. After grouping, experimental units must be randomly assigned to treatments. This is discussed in most standard texts on design. Where data are collected by sampling from some natural population then a suitable strategy should be used so that comparisons of interest have valid estimates of error.

Data analysis

Types of data

At least four distinct types of data could be collected in the course of scientific experiment or observation according to what is being investigated and the scale on which it is measured:

(i) Quantities, such as weight or length, measured on an interval scale. Such data are said to have a continuous distribution.

(ii) Variables measured on a semi-quantitative scale where observations can be ordered but the magnitude of the difference between successive points on the scale need not be constant. Examples are a score of an animal's body condition on (say) a ten point scale or a sensory judgement on (say) a four point scale from worst through to best. Under some circumstances (discussed later) the distributions of such variables might be approximately continuous but usually would not be so. Variables of this type are called ordinal categorical. An extreme example of this type is when data are ranks of observations.

(iii) Qualitative observations classified into two or more categories. Examples are classification of animals by sex or a food tasted as (say) bland, bitter, savoury or sweet. These data would follow some discrete distribution such as a binomial (for two categories) or a multinomial (for more than two categories). Such variables are called nominal categorical.

(iv) Counts of events occurring randomly in time such as radioactive disintegrations or cell divisions. Often such counts will follow a Poisson distribution.

It is important to make these distinctions and to recognise in any given instance what type of data are under consideration so that an appropriate method of statistical analysis can be chosen.

Analysis of variance

Data measured on an interval scale (type (i) above) arising from designed experiments should be subject to an analysis of variance according to the design used and the questions of interest to the experimenter. Such analyses provide estimates of error from which standard errors for tables of means can be calculated. The separate calculation of individual standard deviations for each cell of a table is rarely justified except where there is evidence of heterogeneity of variance between groups or when means reported in the table are based on widely differing numbers of samples. Each variable in the table should, of course, have its own standard error.

Structured treatments

Where experimental treatments have a factorial structure, the breakdown of sums of squares in the analysis of variance should follow this structure. Where factors have quantitative levels, it may be helpful to use polynomial sub-models to explore any trends which might be present in the data (see references). Where treatments at first appear unstructured, it is nevertheless often possible to test comparisons, such as 'untreated control against average of all novel treatments' or 'treatments of a certain type against those of some other type', by means of orthogonal contrasts for which sums of squares can be calculated in the analysis of variance table. Use of analysis of variance in this way provides a much better basis for inference than does an arbitrary selection of *t*-tests between cells of a table of means.

Comparisons between means

Multiple comparison procedures (MCP), such as Duncan's multiple range test (DMRT), are valid (and then not necessarily appropriate) only where it can be assumed that the treatment means are uncorrelated and have equal variance. These procedures are not appropriate when the treatments have quantitative levels or factorial structure. Because MCP ignore structure, they cannot provide satisfactory interpretation of structured treatments. Even where MCP might be valid they are apt to give rise to paradoxes such as 'A is not significantly different from B', 'B is not significantly different from C', but 'A is significantly different from C'! Furthermore, the often overlapping groups of means rarely provide a natural framework for the interpretation of data and it is clear from the pages of the journal that results of tests (such as DMRT) although reported are seldom discussed. Authors are referred to the extensive literature on this topic for a fuller discussion of the problems with MCP and suitable alternative methods (see references cited). It is hoped that authors will find these alternative methods more useful. They may well be called upon to justify their use of MCP if alternative methods could be used to better effect.

Least Significant Difference (LSD) is often used to make comparisons among a set of means. If used with care this can be a helpful informal technique in the exploratory phase of data analysis. Among many dangers is the temptation to make all possible pairwise comparisons. Other statistical considerations (see references) make this a statistic of little usefulness in supporting inferences from data. Authors are encouraged to use the more powerful and statistically valid methods mentioned in this paper.

Observations repeated on the same material

Where observations are repeated on the same experimental units at successive times, special methods of analysis may be required. This is due to the likely high correlation between successive observations made on the same experimental unit. For example, in an animal growth study the previous growth of an animal will very likely affect its future growth leading to correlated body weights at successive times of observation. See, for example, Rowell and Walters (1976) and Kenward (1987).

Regression analysis

If authors wish to establish the nature of a relationship between two or more quantitative variables then some sort of regression analysis is required. Simple analysis of variance with linear and quadratic orthogonal polynomials would suffice just to demonstrate trends in data. Regression analysis should never be done blind; graphical exploration of data is essential. When designing a study, care should be taken that the ranges of explanatory variables extend over useful intervals so that results will be as widely applicable as possible. The potential undue influence of isolated points is a danger. Examination of residuals from the fitted model will help in assessing how good the fit is. R^2 (adjusted or otherwise) is of limited use in measuring goodness of fit because it measures the goodness of fit only where the relationship is linear and also it is sensitive to the design of the study. Standard deviations about the regression line will give a clearer picture as to whether the fitted model is adequate. See Atkinson (1987) and Cook and Weisberg (1985) for a discussion of regression diagnostics.

Any non-linear models, such as exponential curves or logistic growth curves, should be fitted by least squares or the method of maximum likelihood using proven computer software (see below). Unlike simple linear regressions, these models require solution by iterative numerical methods.

Analysis of multivariate data

Where many distinct measurements on variables of a similar type (for example, absorbance at a number of wavelengths) have been made on each of the experimental units, authors are encouraged to consider use of a suitable multivariate analysis technique. These methods can sometimes give a much more concise summary of results than a set of univariate analyses in such circumstances. (See list of useful references.)

Analysis of categorical data and counts of discrete events

Ordinal categorical data

Data which are scores on some (possibly arbitrary) ordinal scale may require special methods of analysis (see references). However, if the data can be considered as from approximately a grouped continuous distribution then analysis of variance may be used. A rough rule as to when this assumption could be justified is where the scale has about ten or so points and where the increments between points on the scale are similar in size.

Nominal categorical data

Observations which are classified into two or more qualitative categories give rise to data in the form of counts. These data could be analysed by non-parametric methods or by analysis of χ^2 on contingency tables. More sophisticated methods such as Generalised Linear Models (GLM) are described below.

Statistical methods for data expressed as counts

Data which are counts or counts expressed as percentages will usually require transformation to stabilise variance. Transformation of percentages is helpful where the range of data covers the extremes near 0% or 100%; it might not be necessary if the range is small and between 30% and 70%. A transformation which may be suitable is the angular transformation, $\sin^{-1}\sqrt{x}/100$ where x is a percentage and $\sin^{-1}$ means 'the angle between 0° and 90° whose sine is'. Also a suitable transformation may enable data to be modelled in terms of additive treatment effects and thus allow a simple explanation. For example, where the effect of a treatment was to cause a proportional increase in, say, counts of bacteria, a log transform would enable such effects to be modelled on an additive scale. The logit transformation (or 'log odds'), $\log_e[p/(1-p)]$ where p is a proportion, may allow additive treatment effects to be fitted for data which are number responding out of a fixed total. For some data, where the twin objectives of stable variance and an additive scale for effects cannot be met by a single transformation, a Generalised Linear Model (GLM) could be useful. These GLM are like the familiar linear models of regression and analysis of variance but are more general in allowing error distributions other than normal, for example, Poisson or binomial, and data transformations to linearity (additivity) such as log, logit,

complementary log-log, power, and several others. Generalised Linear Models are usually straightforward to interpret and are incorporated in most general statistical packages (see below).

Statistical software

There are many excellent statistical packages which perform most of the standard analyses. The following list is not comprehensive: BMDP, Genstat, Minitab, SAS, Splus, SPSS, Stata and Statistica.

PRESENTATION

Description of statistical methods

Remember, unless your paper deals with new methodology or applications of statistics, that the statistics are supposed to be an aid to the interpretation of data, not an end in themselves. Include a clear, concise description of what you have done and only those details which you wish to refer to in the discussion of the results. It is rarely necessary to describe standard methods in full where reference can be made to a standard text. Equations may be helpful to make clear an adaptation of a standard method but *all* symbols used must be defined in the text.

Begin with a description of the design of the experiment or investigation. State the objectives of the analysis and justify any underlying assumptions in the context of your application. Describe how the chosen method has been applied to these data. If a statistical package has been used to process data on a computer, give a reference to the manual.

Making inferences from data

Statistical analyses should be used to estimate effects and to justify inferences against a background of biological variation. They should be discussed and not ignored. Lack of discussion strongly suggests that the statistics has been included solely as a concession to try to ensure publication. The achievement of 'significant' differences from a statistical test should not be regarded as an end in itself! Indeed, sameness can be just as important as difference. Statistical significance is judged according to arbitrary probability levels and its attainment may well depend on the level of replication in an experiment. A statistically significant effect is not necessarily a biologically interesting one, nor vice versa. The estimated size of effects of treatments and their standard errors are often of more interest than is statistical significance.

Tables

Deciding on precision

Do not be tempted to exaggerate the precision of results by giving too many decimal places. A useful rule of thumb is to round to approximately one tenth of the standard error and give standard errors with one more decimal place than this.

Choosing how to give errors of estimation

The use of '±' to indicate some sort of error is unsatisfactory, because usage varies in different countries. The type of error should be clearly stated; for example, standard error of mean (SEM), standard error of difference (SED), standard deviation (SD), and so on. The use of '±' will only be allowed where it is clearly stated, on first use, what is meant. As far as possible standard errors in tables should be based on pooled estimates of error. Separate standard errors for cells in a table are worthwhile only where there is evidence for heterogeneity of variance or where replication differs widely between classes.

Presenting data transformed for analysis

Where data have been transformed for analysis, results may be easier to comprehend if back-transformed. In this case confidence intervals, calculated on the transformed scale and also backtransformed, should be given for each entry in a table (although in general these are only approximate). There is little sense in back-transforming a standard error when the purpose of the transformation has been to stabilise variance. Use of Generalised Linear Models can provide summaries in terms of natural quantities, such as proportions, and so avoids problems of back-transformation.

Avoiding superfluous statistical tests

Tables should not be annotated with results of statistical tests unless these are referred to in the text of the paper. Routine annotation with superscripts according to Duncan's Multiple Range Test will not be acceptable (see above). Authors should themselves discuss all sensible inferences which can be drawn from their results; there is no need for any statistics other than those required to support these.

Graphs and regression analysis

Graphs are often helpful where regression results are presented. Error bars for standard errors of means or standard errors of prediction for fitted relationships are preferred. Do not use Least Significant Differences (LSD) for data where there is a quantitative structure to treatment levels as these should never be analysed by successive significance tests (see above for alternatives).

The use of R^2 (=sum of squares due to regression on *x*-variates ÷ total corrected sum of squares for *y*-variate) to measure goodness of fit is discouraged and is unacceptable in the case of multiple regression as R^2 takes no account of the number of terms in the model. Adjusted R^2 ($\bar{R}^2$=1-[residual mean square after fitting all *x*-variates ÷ total mean square with (*n*-1) degrees of freedom]) or percentage variance accounted for (=100×$\bar{R}^2$) should be used instead as this does allow for the number of parameters estimated. Prior to the calculation of any summary statistic for goodness of fit, however, the residuals should be plotted and examined to detect any patterns which invalidate the model assumptions.

ACKNOWLEDGEMENTS

I am grateful to John Gower (AFRC), David Brown (AFRC), Clive Moncrieff (AFRC), Peter Rothery (NERC) and other colleagues for many helpful comments on earlier versions of these notes.

USEFUL REFERENCES

Good Statistical Textbooks

Cochran WG, Cox GM, *Experimental Designs*. Wiley, New York (1992).

Cox DR, *Planning of Experiments*. Wiley, New York (1992).

Cox DR, Snell EJ, *Applied Statistics: Principles and Examples*. Chapman & Hall, London (1981).

Draper NR, Smith H, *Applied Regression Analysis* (3rd edn). Wiley, New York (1988).

Mead R, Curnow RN, Hasted A, *Statistical Methods in Agriculture and Experimental Biology* (2nd edn). Chapman & Hall, London (1993).

Snedecor GW, Cochran WG, *Statistical Methods* (7th edn). Iowa State University Press, Ames (1980).

Sokal RR, Rohlf FJ, *Biometry* (3rd edn). WH Freeman, San Francisco (1995).

Steel RGD, Torrie JH, Dickey D, *Principles and Procedures of Statistics*. McGraw (1996).

Weisberg S, *Applied Linear Regression* (2nd edn). Wiley, New York (1985).

Some General Papers

Chatfield C, The initial examination of data. *J Royal Statist Soc A* **148**:214-253 (1985).

Jeffers JNR, *Statistical Checklist 1: Design of Experiments*. Institute of Terrestial Ecology, Huntingdon, UK (1982).

Jeffers JNR, *Statistical Checklist 2: Sampling*. Institute of Terrestial Ecology, Huntingdon, UK (1982).

Preece DA, Good statistical practice. *The Statistician* **36**:397-408 (1987).

Polynomials and Contrasts

Mead R, Curnow RN, Hasted A, *Statistical Methods in Agriculture and Experimental Biology* (2nd edn). Chapman & Hall, London, pp 94-97 (1993).

Snedecor GW, Cochran WG, *Statistical Methods* (7th edn). Iowa State University Press, Ames, pp 224-228, 257-258, 307-315, 404-407 (1980).

Sokal RR, Rohlf FJ, *Biometry* (3rd edn). WH Freeman, San Francisco, pp 229-259, 521-530, 665-677 (1995).

Repeated Measures

Kenward MG, A method for comparing profiles of repeated measurements. *Applied Statistics* **36**:296-308 (1987).

Rowell JG, Walters DE, Analysing data with repeated observations on each experimental unit. *J Agric Sci* **87**:423-432 (1976).

Regression Diagnostics

Atkinson AC, *Plots, Transformations and Regression*. Clarendon Press, Oxford (1987).

Cook RD, Weisberg S, *Residuals and Influence in Regression* (2nd edn). Chapman & Hall, London and New York (1985).

Draper NR, Smith H, *Applied Regression Analysis* (3rd edn). Wiley, New York, Ch 3 (1998).

Weisberg S, *Applied Linear Regression* (2nd edn). Wiley, New York, Chs 4, 5, 6 (1985).

Multiple Comparisons

Letter to Editors and short articles that concentrate on illustrating inappropriate published uses of multiple comparison tests and on explaining the alternatives

Baker RJ, Multiple comparison tests. *Canadian J Plant Sci* **60**:325-327 (1980).

Bryan-Jones J, Finney DJ, On an error in instructions to authors. *Hort Science* **18**:279-282 (1983).

Little TM, Interpretation and presentation of results. *Hort Science* **16**:19-22 (1981).

Morse PM, Thompson BK, Presentation of experimental results. *Canadian J Plant Sci* **61**:799-802 (1981).

Pearce SC, Data analysis in agricultural experimentation. III Multiple comparisons. *Expl Agric* **29**:1-8 (1993).

Perry JN, Multiple comparison procedures - a dissenting view. *J Econ Ent* **79**:1149-1155 (1986).

Papers that either go into more detail about the principles and theory behind multiple comparison tests or are concerned with discussing the wider issues of experimental design and analysis, but include comments on multiple comparison tests

Chew V, Comparing treatment means: a compendium. *Hort Science* **11**:348-357 (1976).

Chew V, Testing differences among means: correct interpretation and some alternatives. *Hort Science* **15**:467-470 (1980).

Cox DF, Design and analysis in nutritional and physiological experimentation. *J Dairy Science* **63**:313-321 (1980).

Multivariate analysis

Barnett V (Ed), *Interpreting Multivariate Data*. Wiley, New York (1981).

Everitt BS, *Graphical Techniques for Multivariate Data*. North-Holland, Amsterdam (1978).

Krzanowski WJ, *Principles of Multivariate Analysis: a User's Perspective*. Clarendon (1990).

Ordinal Categorical Data

Agresti A, A survey of strategies for modelling cross-classifications having ordinal variables. *J Amer Statist Assoc* **78**:184-198 (1983).

Agresti A, *Introduction to Categorical Data Analysis*. Wiley, Chichester (1996).

Fingleton B, *Models of Category Counts*. Cambridge University Press, Cambridge (1984).

McCullagh P, Regression models for ordinal data. *J Royal Statist Soc B* **42**:109-142 (1980).

Generalised Linear Models

Cox DR, Snell EJ, *The Analysis of Binary Data*. Chapman & Hall (1988).

McCullagh P, Nelder JA, *Generalised Linear Models* (2nd edn). Chapman & Hall, London (1989).

Adjusted R^2 ($\bar{R}^2$)

Draper NR, Smith H, *Applied Regression Analysis* (2nd edn). Wiley, New York, pp 91-92 (1981).

Healy MJR, The use of R^2 as a measure of goodness of fit. *J Royal Statist Soc A* **147**:608-609 (1984).

Snedecor GW, Cochran WG, *Statistical Methods* (7th edn). Iowa State University Press, Ames, pp 358-360 (1980).

Weisberg S, *Applied Linear Regression* (2nd edn). Wiley, New York, p 188 (1985).